



# ГЕНЫ И ГЕНОМЫ

М. СИНГЕР, П. БЕРГ

ИЗДАТЕЛЬСТВО "МИР"

1

## **ГЕНЫ И ГЕНОМЫ**

# **GENES & GENOMES**

A changing Perspective

## **Maxine Singer**

President, Carnegie Institution of  
Washington Scientist Emeritus,  
National Institutes of Health

## **Paul Berg**

Willson Professor of Biochemistry  
Director, Beckman Center for Molecular  
and Genetic Medicine,  
Stanford University School of Medicine

**М.СИНГЕР, П.БЕРГ**

# **ГЕНЫ И ГЕНОМЫ**

**В двух томах**

**Том 1**

Перевод с английского  
д-ра биол. наук Т.С. Ильиной  
д-ра биол. наук Ю.М. Романовой

под редакцией  
д-ра биол. наук Н.К. Янковского



**МОСКВА «МИР» 1998**

УДК 575.113/118  
ББК 28.04  
С38

С38 Авторы: Сингер М., Берг П.  
**Гены и геномы:** В 2-х т. Т. 1. Пер. с англ. М.: Мир,  
1998.— 373 с., ил.

ISBN 5-03-002849-8

Университетское руководство по молекулярной биологии, написанное выдающимися американскими учеными, членами Национальной академии наук (П. Берг - лауреат Нобелевской премии). Книгу отличают общебиологический подход, глубина теоретических обобщений, изящная и наглядная форма подачи материала.

В 1-м томе рассматриваются следующие вопросы: строение и функционирование биологических молекул, участвующих в работе генетического аппарата; процессы репликации и экспрессии генов; широкий круг вопросов, относящихся к технологии рекомбинантных ДНК (ферментные системы, системы вектор—хозяин, манипуляции с рекомбинантными ДНК).

Для молекулярных биологов, преподавателей и студентов университетов, специалистов.

**ББК 28.04**

*Редакция литературы по биологии*



Издание осуществлено при поддержке  
Российского фонда фундаментальных исследований  
по проекту № 98-04-62107

ISBN 5-03-002848-X (русск.)  
ISBN 5-03-002849-8  
ISBN 0-935702-17-2 (англ.)

© 1991 by University Science Books  
© перевод на русский язык издательство  
«Мир», 1998

# ОТ РЕДАКТОРА ПЕРЕВОДА

Книга «Гены и геномы» отличается сочетанием традиционных разделов молекулярной генетики, таких, как молекулярные основы наследственности, технология рекомбинантных ДНК, анатомия, экспрессия и регуляция генов, с новым взглядом на молекулярно-генетические проблемы – рассмотрением их в контексте организации и функционирования целых геномов.

Авторы книги М. Сингер и лауреат Нобелевской премии по химии П. Берг – всемирно известные специалисты в области молекулярной биологии, одни из создателей геной инженерии, методической основы современной молекулярной биологии и генетики. Читатель найдет в книге подробные сведения о структуре и функциях ДНК, РНК, белков; репликации и функционировании генома, обратной транскрипции; модификациях, репарации и рекомбинации ДНК; о транскрипции и трансляции мРНК в клетках про- и эукариот; регуляции экспрессии генов; технологии рекомбинантных ДНК. Все эти вопросы обсуждаются в первых двух частях книги.

Особый интерес для нашего читателя представляет описание молекулярной анатомии эукариотических геномов, интегрирующее современные молекулярно-биологические знания в данной области (часть III), поскольку в отечественной литературе информация по этим проблемам носит фрагментарный характер. Авторы анализируют структуру и функции генома как целого и роль геномных перестроек в функционировании генома в норме и патологии, подчиняя этой задаче рассмотрение его элементов (тандемных и диспергированных повторов, центромерных и теломерных областей, транспозонов и др.). В заключительной, четвертой части

обсуждаются принципы функционирования и перспективы дальнейших исследований сложных биологических систем (мультигенные семейства, тканеспецифичная экспрессия, эмбриональное развитие и дифференцировка).

Каждой части предшествует краткий обзор общих принципов и концепций рассматриваемых в этом разделе вопросов молекулярной генетики; за ним следует углубленное описание экспериментальных работ, на основе которых эти концепции были сформулированы. Завершаются части списком первоисточников по ключевым экспериментальным работам данного направления.

Книга может служить руководством для подготовки студентов и аспирантов на биологических и медицинских факультетах вузов, а также поможет научным работникам ориентироваться в новых областях исследования, связанных с изучением геномов. Понимание текста облегчают многочисленные оригинальные иллюстрации и таблицы. Они будут также полезны преподавателям вузов в качестве иллюстративного материала при чтении лекций.

Издание русского перевода книги «Гены и геномы» особенно своевременно, поскольку развитие современной молекулярной биологии идет во многом через исследования структуры и функций геномов, проводимые в рамках всемирной программы «Геном человека».

Книгу переводили: Т.С. Ильина (введение к ч. II, гл. 4–7), Е.А. Кабанова (введение к ч. III, гл. 8), А.А. Лушникова (гл. 9, 10, ч. III), Ю.М. Романова (предисловие, благодарности, введение к ч. I, гл. 1–3).

*Н.К. Янковский*

# ПРЕДИСЛОВИЕ

Эта книга посвящена молекулярным структурам и механизмам, лежащим в основе передачи и использования генетической информации сложными организмами. Представляя ее, мы надеемся, что у читателя возникнет чувство предвкушения того удовольствия, которое он получит, знакомясь с открытиями в области рекомбинантных ДНК.

Чтобы подготовить читателя к восприятию этих новейших достижений, мы остановились во введении к гл. 1–3 на классических генетических и биохимических исследованиях, проведенных в период с начала века до примерно 1972 г. Эти исследования были выполнены преимущественно на микроорганизмах, и благодаря им удалось установить структуру и функции генов и геномов. Разработка в начале 70-х годов методов получения рекомбинантных ДНК и связанный с этим прогресс в технологии быстрого определения нуклеотидной последовательности ДНК позволили вскоре применить молекулярные методы и для исследования более сложных геномов. Ядро книги составляют работы, выполненные менее чем за два десятилетия, которые показали, какие широкие возможности дает использование этих методов в биологии.

Книга создавалась довольно долго. Идея ее написания возникла после того, как один из нас (*Л. Б.*) прочел серию публичных лекций в университете Питтсбурга в 1979 г. Попытка подготовить публикацию лекций показала несостоятельность этого замысла. Когда такие лекции читаются в аудитории слушателей, не являющихся биологами, приходится опускать многие интересные детали, и книга, лишенная таких тонкостей, конечно, не может принести удовлетворения. По мере расширения концептуальной основы книги становилась очевидной необходимость разделить тяжкий труд по ее написанию с кем-нибудь еще. Так нас стало двое. Мы вместе прорабатывали весь материал, пытаясь добиться унифицированного, согласованного подхода и стиля. Мы поняли всю безрассудность попытки дать всестороннее, всеобъемлющее изложение материала; эта задача была успешно решена до нас большими группами авторов в нескольких недавних блестящих изданиях. Нам приходилось постоянно пересматривать уже написанные главы, чтобы их окончательный вариант соответствовал последующим разделам, еще находящимся в работе. Чтобы нас не захлестнул поток постоянно поступающей новой информации, мы ограничили наши задачи,

сосредоточив внимание на тех проблемах, которые уже довольно глубоко изучены и наглядно иллюстрируют достигнутый прогресс.

Наша главная задача состояла в том, чтобы раскрыть сущность и глубину экспериментальных подходов науки, которая была названа молекулярной генетикой, применительно к эукариотическим организмам. Чтобы решить эту задачу, а также облегчить понимание материала читателями, обладающими ограниченным объемом знаний по биохимии, клеточной биологии и генетике, мы постарались изложить основы этих направлений биологии двумя способами. Во-первых, в гл. 1, 2 и 3 суммирована наиболее важная информация о структуре ДНК, РНК и белков; о различных клеточных процессах, протекающих с участием ДНК (репликация, репарация и рекомбинация); об основных механизмах транскрипции, трансляции и контроле экспрессии генов. Читатели, хорошо ориентирующиеся в данных вопросах, могут пропустить эти главы. Во-вторых, во введениях к частям I, II и III даны исторические экскурсы и общий взгляд на проблемы, изложенные в главах, составляющих эти части. В них не говорится детально о том, как были открыты и доказаны те или иные положения, а делается попытка объяснить, как на основе различных исследований в области биохимии, генетики, микробиологии, клеточной и эволюционной биологии был выстроен интеллектуальный каркас современной биологии. Так, во введении, предворяющем гл. 1, 2 и 3, прослеживается исторический путь, приведший нас к современному взгляду на наследственность. Мы знакомимся с концепцией гена, трансмиссией и сегрегацией генов, с логическим переходом от первичного картирования генетических детерминант к точной локализации генов на хромосоме, с идентификацией генов как дискретных участков молекулы дезоксирибонуклеиновой кислоты и информационными взаимоотношениями между ДНК, РНК и белками.

В части II (гл. 4–7) описаны подходы к конструированию рекомбинантных молекул ДНК, их клонированию, отбору и характеристике, а также другие экспериментальные методы, использующиеся в молекулярной генетике. Основное внимание уделяется принципам этих методов, а не лабораторным приемам. Читатели, которых больше интересуют методические детали, могут обратиться к замечательным сборникам, вышедшим за послед-

ние 10 лет; ссылки на них даны в книге. Во введении, предваряющем гл. 4–7, особое внимание уделяется истории развития современных подходов и методов, разработанных на основе огромного количества казалось бы не связанных между собой фундаментальных исследований в энзимологии, генетике бактерий и бактериофагов и биохимии нуклеиновых кислот.

Основная тема книги раскрывается в части III, где изложены концепции, лежащие в основе наших представлений о генетических информационных системах эукариот. Введение к этой части знакомит нас с отличительными особенностями геномов эукариот. Речь идет об интронах и сложных сигналах, регулирующих транскрипцию, а также о множестве повторяющихся последовательностей в геноме и связанной с этим ролью обратной транскрипции в происхождении сегментов эукариотической ДНК. Раздел заканчивается описанием некой объединяющей концепции, рассматривающей биологическую эволюцию как процесс, в основе которого лежит перестройка нуклеиновых кислот и, следовательно, изменение структуры белковых молекул.

В первой из трех глав части III (гл. 8) приведены данные о структуре генов эукариот и современные представления о механизме их экспрессии, в частности сведения о сложных сигналах регуляции транскрипции, а также о происхождении, локализации и структуре интронов и тех механизмах, с помощью которых интроны удаляются из первичных транскриптов при сплайсинге. Очень существенным здесь явилось применение обратной генетики—введение специфических мутаций в определенные сегменты ДНК и последующий анализ структурно-функциональных взаимоотношений в генах эукариот. В гл. 9 основное внимание сосредоточено на организации сложных эукариотических геномов. Рассмотрено расположение генов и других элементов в молекуле ДНК, в частности в центромерных и теломерных областях. Красной нитью через всю главу проходит концепция генома как летописи эволюционной истории. В заключение дано описание геномов внутриклеточных органелл—митохондрий и хлоропластов. В гл. 10 представлены механизмы случайных и неслучайных перестроек геномной ДНК. Речь идет об амплификациях, делециях и транспозициях—как незапрограммированных и приводящих к мутагенезу, так и запрограммированных в геноме и осуществляющих точную регуляцию генной экспрессии, например изменение типов спаривания у дрожжей и образование генов иммуноглобулинов.

В части IV дана краткая иллюстрация применения общих принципов, рассмотренных ранее, к специфическим сложным биоло-

гическим системам. Показано, что гены работают в составе сложных многокомпонентных взаимодействующих систем. В каждой из таких систем указаны общие принципы использования по-своему, что приводит к огромному разнообразию живого. Ранее это разнообразие описывалось чисто феноменологически. Теперь феноменологическое описание уступает место описанию регуляции генной экспрессии во времени и пространстве на молекулярном уровне. В этой части рассматривается также, как с развитием методологии рекомбинантных ДНК биология превратилась из науки описательной в практическую. Генотипы, а следовательно, и фенотипы на уровне индивидуальных белков, клеток и организмов можно изменять, что позволяет в будущем исследовать фундаментальные биологические процессы, а также решить острые вопросы, стоящие перед человечеством и планетой, которую мы населяем. Многие из предоставляющихся возможностей можно будет реализовать только после того, как мы лучше узнаем структуру геномов. Это касается определения точного расположения генов и получения более полной информации об их нуклеотидных последовательностях. В части IV излагаются также положения, существенные для картирования и секвенирования геномов некоторых видов, в том числе и человека.

В 1980 г., когда нам впервые пришла мысль о написании книги, мы, как и многие другие биологи, только что пережили период, когда научные издания стали привлекать к себе столь же пристальное внимание, как и общественно-политические. Ошелмляющие успехи, достигнутые благодаря применению методологии рекомбинантных ДНК, были только одним из аспектов того, что называли «революцией в биологии». Наряду с приходом новой эры в постижении сути живого эта революция вызвала беспрецедентное беспокойство общественности о последствиях биологических экспериментов.

Вначале общественный интерес был связан с тем, что группа ученых-биологов подняла вопрос о безопасности экспериментов с рекомбинантными ДНК. Людей, работающих с патогенными микроорганизмами, тоже беспокоила безопасность проведения лабораторных исследований. Подобная ситуация уже возникала в 1971 г., когда во всем мире стали широко использоваться вирусы, индуцирующие образование опухолей у экспериментальных животных. Жизненный цикл этих вирусов отличался от жизненного цикла более привычных вирусов (например, вирусов кори или полиомиелита), а некоторые из них, вероятно, могли индуцировать рак у человека. Это заставляло обратить самое серьезное внимание на потенциальную опасность таких вирусов для ученых и студентов. Аналогичные опа-



сения высказывались и в связи с появлением методов получения рекомбинантных ДНК. Общая атмосфера озабоченности усугублялась глубоким чувством социальной ответственности, возникшим и окрепшим в Соединенных Штатах за предыдущее десятилетие борьбы за гражданские права и против войны во Вьетнаме.

Первые эксперименты по созданию рекомбинантной ДНК состояли в соединении ДНК, выделенной из вирусов, которые вызывали опухоли у мелких лабораторных грызунов, с фрагментами ДНК из хорошо изученных вирусов. Намерения ввести эти новые молекулы ДНК в бактериальные клетки вызвали серьезные нарекания, но вскоре было принято решение приостановить подобные попытки, и все страсти улеглись. Следующим крупным шагом было конструирование рекомбинантной ДНК, содержащей ген, который обеспечивал устойчивость бактерий к антибиотикам. Возможность генетической трансформации живых клеток путем введения необычных молекул ДНК в бактерии поставила вопрос о том, что такие бактерии могут вызывать развитие опухоли или приобретут резистентность к важным в медицинском отношении антибиотикам. Эти вопросы обсуждались как на научных конференциях, так и в частных беседах. На одной из таких конференций в июне 1973 г. участвующие в ней ученые решили привлечь внимание Национальной академии наук к этой перспективной области исследований и попросить провести направленные исследования для оценки потенциальной опасности манипуляций с рекомбинантными ДНК. Чтобы придать гласность этому обращению, ученые опубликовали письмо в журнале "Science".

Академия, как обычно, сформировала комитет, куда вошли ученые, активно работающие с рекомбинантными ДНК и в смежных областях. Члены комитета встретились в апреле 1974 г. и предложили два основных решения, ставших, к их удивлению, новостью номер один. Во-первых, они призвали объявить всеобщий мораторий на те эксперименты с рекомбинантными ДНК, в которых используется ДНК опухолеродных вирусов или в бактерии вводятся гены, детерминирующие токсины или ответственные за резистентность бактерий к антибиотикам. Во-вторых, они призвали к широкому и свободному международному обсуждению этих проблем на конференции, которую предполагалось провести следующей зимой.

Несмотря на ропот недовольства и даже обвинения в адрес ученых США в том, что они таким образом пытаются затормозить работу ученых других стран и выиграть гонку в соревновании за главные открытия, которые могут быть сделаны благодаря развитию новых методов, мораторий, насколько это известно, был принят повсеместно.

На конференцию были приглашены молекулярные биологи, вирусологи, микробиологи и биохимики США и других стран, а также представители научной администрации, журналисты и юристы.

Встреча состоялась в феврале 1975 г. в Асиломаре, шт. Калифорния, в конференц-центре на берегу Тихого океана, несколькими милями южнее Гопкинсовской океанологической станции в Пасифик-Гроу (при Станфордском университете). Организационный комитет провел предварительные заседания рабочих групп, и эксперты в различных областях подготовили документы для представления на конференции. Состоялась оживленная, иногда жаркая, дискуссия о том, насколько реальна опасность подобных экспериментов. Большинство (но не все) участников признали, что риск появления особенно опасных организмов существует, но вероятность этого очень мала. Принятие окончательного решения о дальнейших действиях было отложено до тех пор, пока не скажут свое слово юристы. Юристы же возложили персональную юридическую ответственность за все последствия, даже экстраординарные, на ученых. Они также напомнили нам, что общество должно быть застраховано от риска, если опасность последствий, пусть крайне малая, существует. Позиции прояснились. Необходимо было следовать именно этому основанному на здравом смысле курсу. В последний день работы конференции был предложен предварительный отчет Организационного комитета. Он всесторонне обсуждался и был в конце концов принят. Рекомендации Асиломарской конференции получили широкую огласку в прессе и были опубликованы позже в ряде научных журналов. Несколькими неделями ранее к аналогичным заключениям пришел и возглавляемый лордом Эшби правительственный комитет Великобритании.

Рекомендации, принятые в Асиломаре, стали основой и импульсом для официальных мероприятий в США, начавшихся на следующий же день после закрытия конференции. Комитет, организованный Национальными институтами здоровья (НИЗ), приступил к разработке директив по контролю за всеми экспериментами с рекомбинантными ДНК, проводимыми в институтах своего подчинения. Первые инструкции были весьма строгими в расчете на то, что по мере накопления опыта и знаний они будут пересмотрены. По сей день не выявлено ни одного нежелательного инцидента ни с персоналом лабораторий, ни с кем-либо из других людей, причиной которого были бы эксперименты, проведенные с рекомбинантными ДНК; а таких экспериментов насчитывается десятки тысяч. Содержащиеся в документе требования к большинству рутинных экспериментов с рекомбинантной ДНК изъятые или сделаны менее строгими. Жесткие тре-

бования, согласно рекомендациям НИЗ, сохраняются только к тем экспериментам, в которых рекомбинантная ДНК содержит протяженные участки, полученные из ДНК особо патогенных микроорганизмов (этих рекомендаций придерживаются и во многих других странах). Интересно отметить, что использование технологии рекомбинантных ДНК сделало изучение некоторых важных, но опасных возбудителей инфекционных заболеваний человека и животных вполне доступным и безопасным.

Интерес общественности к экспериментам с рекомбинантными ДНК сохраняется по сей день и теперь распространяется также на генную инженерию целых организмов—бактерий, растений и животных. Сначала этот интерес и беспокойство были сосредоточены на том же, на чем основывались интерес и беспокойство большинства ученых — на возможности создания в эксперименте болезнетворных агентов. Местные власти и государственные органы издавали законы и указы, аналогичные рекомендациям НИЗ или даже более строгие. В конгрессе обсуждались предложения о придании рекомендациям НИЗ силы закона и установлении наказания за их несоблюдение, но ни одно из них не было принято. Огромное количество независимых дебатов—в разных местах и на разных уровнях—в конце концов узаконили причастность НИЗ к данной проблеме как в том, что касается оценки риска, так и в смысле административной организации.

Позже дискуссии приняли несколько иной характер. Некоторых ученых волнует вопрос о возможных эволюционных последствиях проникновения ДНК через видовые барьеры, например введение ДНК человека в клетки бактерий. Эти же вопросы задают и люди, не являющиеся учеными, однако данный аспект признан не заслуживающим большого внимания. В природе возникает неимоверное число возможностей для обменов ДНК подобного рода, и, по-видимому, они действительно происходят. Эксперименты с рекомбинантными ДНК лишь ненамного увеличивают их вероятность. Кроме того, большинство организмов, несущих рекомбинантные молекулы ДНК, мало жизнеспособны вне лаборатории. Этот последний довод приводился и при недавнем обсуждении вопросов, связанных с необходимостью правительственного контроля за преднамеренным распространением в окружающей среде важных в сельскохозяйственном отношении организмов, полученных методом генной инженерии. Конечно, подобные проблемы требуют внимания, но предъявление слишком строгих требований (например, использование защитной одежды, подобной костюму астронавта) с научной точки зрения необоснованно.

Постоянно обсуждается вопрос и о возможности применения техники рекомбинантной ДНК для создания биологического оружия. США присоединились к международной конвенции 1972 г., запрещающей проведение подобных работ, а Министерство обороны продолжает финансировать исследование по защите от биологического оружия, которое может быть создано в других странах. Это направление тоже интенсивно обсуждается, поскольку трудно определить грань, разделяющую работы, которые направлены на создание средств нападения и защиты. Такие проблемы одинаково волнуют как ученых, так и людей, не имеющих отношения к науке, поскольку они носят политический и социальный характер, а не чисто научный, как и другие проблемы, связанные с развитием новых направлений биологии. Что принесет человечеству генная терапия, если станет возможным ее применение,— пользу или вред? Как справиться обществу и отдельным гражданам с растущим потоком информации о человеческих генах, которые могут быть приобретены в результате применения новых технологий? Каковы этические проблемы, связанные с возможным введением генов человека в организмы животных? Надо ли патентовать животных и растения, полученные с помощью генной инженерии?

Озабоченность общественности, связанная с наступлением биологической революции, породила негативное отношение к данным исследованиям. Биологи опасались худшего: введения строгих правил или инструкций, которые могли серьезно помешать дальнейшей экспериментальной работе и, следовательно, не позволили бы получить новые многообещающие результаты и внедрить их в медицину, сельское хозяйство и промышленность. Многие ученые выражали сожаление о прекращении свободных дискуссий на эту тему. Подобные дискуссии противостояли демагогическим рассуждениям критиков и стремлениям устраивать шумиху в средствах массовой информации вокруг этих вопросов вместо внимательного их анализа. Но все встало на свои места. Свидетельством тому служат научные достижения, описанные в этой книге, и огромное количество очень полезных и важных продуктов, полученных в новой, бурно развивающейся биотехнологической промышленности. Может быть, мораторий и введенные вначале ограничения и задержали прогресс, но ненадолго. Возникшее в то время предубеждение против подобных исследований, по-видимому, можно оправдать неведением, и теперь с уверенностью можно сказать, что риск был гораздо меньше, чем мы могли предполагать.

*Максин Сингер,  
Пол Берг*

# БЛАГОДАРНОСТИ

Вопреки широко распространенному мнению, наука—это коллективный род деятельности, и многое в ней зависит от хорошо налаженного сотрудничества и общения между учеными. Эта книга — не исключение. Мы благодарим за сотрудничество и помощь наших коллег, прочитавших отдельные разделы и главы книги и высказавших критические замечания. Это David Finnegan, Claude Klee, Arthur Kornberg, I.R. Lehman, Howard Nash, Bruce Paterson, Carl Schmid, Robert Tjian. Мы благодарны также тем, кто предоставил нам неопубликованные данные и рукописные материалы.

Пятеро наших коллег прочитали почти всю книгу в первой редакции. Окончательный ее вариант был значительно улучшен благодаря их ценным критическим замечаниям. Мы весьма признательны за этот труд Barbara H. Bowman, David Dressier, Paul Schimmel, Jean O. Thomas, William B. Wood. Нам очень повезло, что мы имели возможность воспользоваться советами Carol Dempster—одного из тех редких людей, кто с энтузиазмом и в высшей степени профессионально вникает во многие научные проблемы. Bruce Armbruster, издатель, президент издательства University Science Books, познакомил нас с Carol Dempster, а также с другими замечательными членами группы, которые помогли нам выпустить эту книгу: Sylvia Stein Wright, редактором рукописи; Gary Head, техническим редактором; Mary Miller, менеджером. Но самым важным было постоянное внимание и неослабевающий энтузиазм самого Bruce Armbruster, которые он проявлял, когда другие дела отвлекали нас от работы над книгой и она замедлялась и практически останавливалась; мы очень ему за это благодарны.

Молекулярные генетики, говоря о генах и геномах, представляют их зримо. Мы «разглядываем» ДНК мысленным взором и пытаемся понять, как она функционирует. Поэтому схемы, приводимые в научных статьях и книгах, подобных этой, очень важны для понимания и запоминания материала. Прекрасные рисунки и диаграммы в этой книге были сделаны Charlene Kornberg и ее коллегами из Медицинской школы Станфордского университета. Эти иллюстрации являются неотъемлемой частью текста, поясняя и дополняя его. Наши наивные представления о процессе создания рисунков и наше настойчивое желание, чтобы их было как можно больше, сделали работу над рисунками обременительной и долгой. Мы благодарны всем, кто так терпеливо работал под руководством Charlene Kornberg: Meryle Colten, Butch Colyear, Mike May-

stead, Eunice Ockermam, Lois Schoen, Kelly Solis-Navarro, Karen Sullivan, Linda Toda. И только мы, а не они несем ответственность за некоторые оставшиеся неисправленными ошибки.

Среди тех неоспоримых преимуществ, которые получили современные биологи, необходимо отметить прекрасные библиотеки, предусмотрительно собранные предыдущими поколениями. При написании этой книги мы пользовались тремя такими библиотеками и работали, огражденные от телефонных звонков и повседневных дел. Все сотрудники библиотеки Лаборатории биологии моря в Вудс-Холе (шт. Массачусетс), Джексоновской лаборатории в Бар-Харбор (шт. Мэн) и Гопкинсовской океанологической станции в Пасифик-Гроу (при Станфордском университете, шт. Калифорния) были неизменно гостеприимны и заботливы.

Harry Woolf, тогдашний директор Института новейших исследований при Принстонском университете (шт. Нью-Джерси), тоже любезно предоставил нам возможность поработать в таком спокойном месте. Его доброжелательный интерес, гостеприимство и дружба, а также прекрасное угощение, которым нас потчевал «мастер на все руки» Franz Moehn, сделали те дни незабываемыми. Как это ни парадоксально, мы имели возможность работать в Принстоне, не отвлекаясь, также благодаря тому, что институтские физики и математики почти не выражали желания беседовать с биологами. Спокойные дни провел также один из нас (П. Б.), работая в Клэр-Холл-Колледже Кембриджского университета. Живое, интеллектуальное общество во главе с Michael Stoker создавало благоприятный, способствующий работе климат.

Пока мы писали эту книгу, наука, которая, мы надеемся, увлечет и вас, шла вперед семимильными шагами. Непрерывные исправления рукописи ложились тяжелым бременем на людей, ответственно относящихся к своему делу и обладающих к тому же чувством юмора, позволявшим им печатать и перепечатывать несчетное множество черновых вариантов. Без Eleanor Olson и Dot Potter (в Станфордском университете) и May Liu и Gail Gray (в Бетесде) работа никогда не была бы доведена до конца.

Милли Берг и Дэн Сингер иногда выражали недовольство и нетерпение (что вполне понятно), но всегда с готовностью поддерживали нас. С неизменными любознательностью, энтузиазмом, гордостью и любовью относилось к нашей работе и юное поколение семей Берг и Сингер, хотя ни один из них не был ни биохимиком, ни молекулярным биологом.

# МОЛЕКУЛЯРНЫЕ ОСНОВЫ НАСЛЕДСТВЕННОСТИ: ОБЗОР

## ВВЕДЕНИЕ

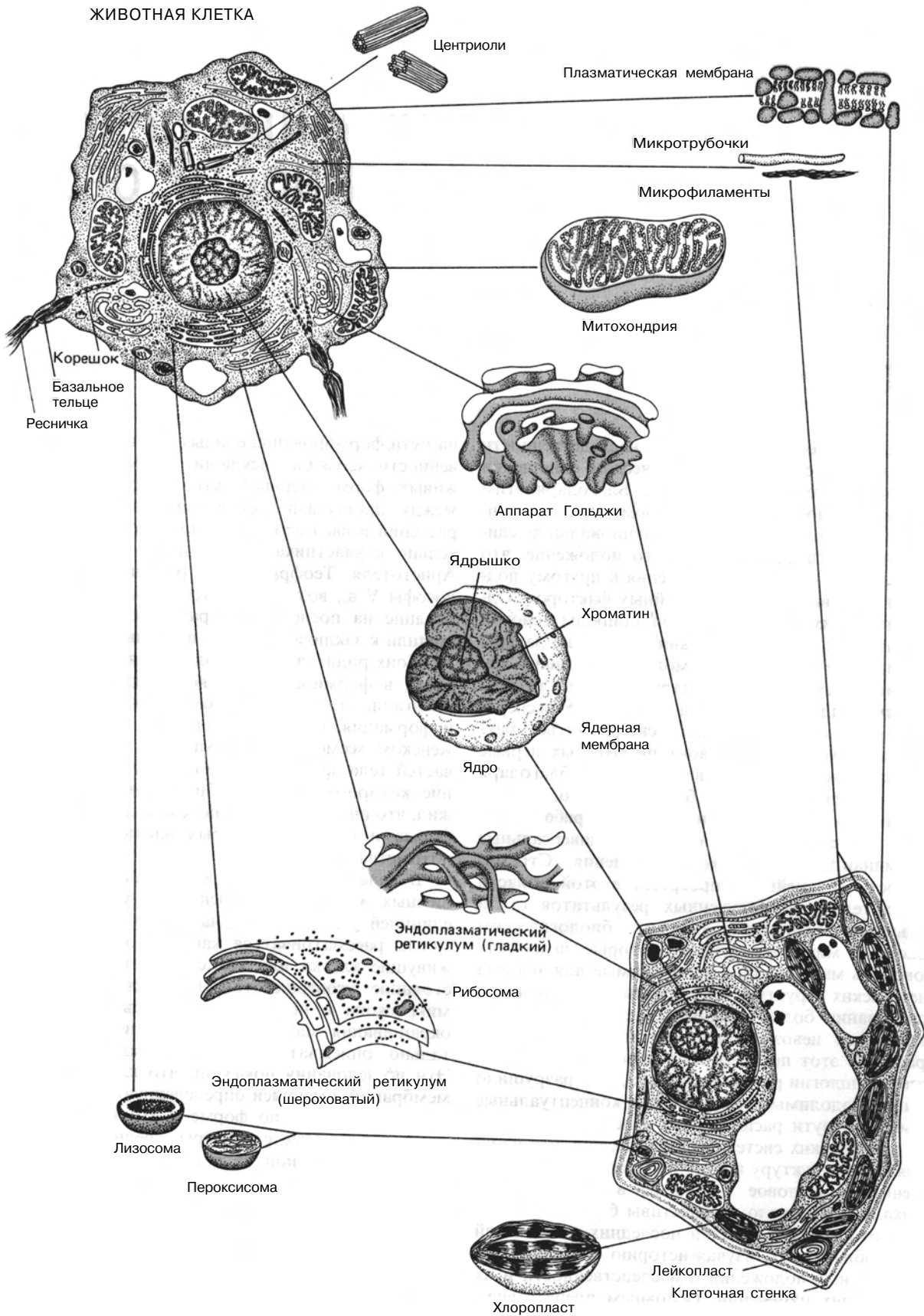
Систематическое изучение наследственности началось со сложных в генетическом отношении объектов—растений и животных. Благодаря этим ранним исследованиям была сформулирована концепция неделимого гена как функциональной единицы наследственности и принято положение, что перенос генов от одного поколения к другому подвержен действию разных случайных факторов. Однако до понимания химической природы генов и механизма их функционирования было еще далеко. Исследование генетических молекул и тонких механизмов регуляции наследственности стало возможным лишь тогда, когда в качестве экспериментальных моделей начали использоваться бактерии и вирусы, о существовании которых первые генетики даже не подозревали. Только благодаря этим организмам впервые было показано, что **дезоксирибонуклеиновая кислота (ДНК), рибонуклеиновая кислота (РНК) и белок**—универсальные детерминанты генетического поведения. Стремительность дальнейшего прогресса в этой области и убедительность полученных результатов стали реальными благодаря особым биологическим свойствам микроорганизмов, которые позволяли проводить манипуляции, необходимые для анализа генетических структур. Аналогичные аналитические исследования более сложных генетических систем тогда были невозможны, поэтому на животных и растениях этот прогресс не распространялся. Развитие технологии рекомбинантных ДНК разрушило труднопреодолимые технические и концептуальные барьеры на пути расшифровки и понимания сложных генетических систем. Неудивительно, что наши взгляды на структуру и функцию генов значительно изменились, а новое мышление в свою очередь радикально изменило перспективы биологии.

Некоторые предпосылки последних достижений можно обнаружить, изучая историю создания фундаментальных положений о наследственности и их последующих изменений. Основным препятствием

на пути формирования единых принципов наследственности служило исключительное разнообразие живых форм. Первым, кто проследил аналогии между процессами воспроизведения животных и растений и ввел слова «самец» и «самка» применительно к участникам этого процесса, был ученик Аристотеля—Теофраст. Еще раньше греческие философы V в., воззрения которых оказали заметное влияние на последующее развитие научных идей, пришли к заключению, что, поскольку дети похожи на обоих родителей, оба пола вносят определенный вклад в формирование нового индивидуума. Они полагали, что этим вкладом является своего рода информация, сконцентрированная в мужском или женском «семени» и поступившая туда из разных частей тела зрелых индивидуумов. Демокрит, мнение которого не было общепринятым, предположил, что информация заключена в частицах, размер, форма и строение которых влияют на свойства потомства.

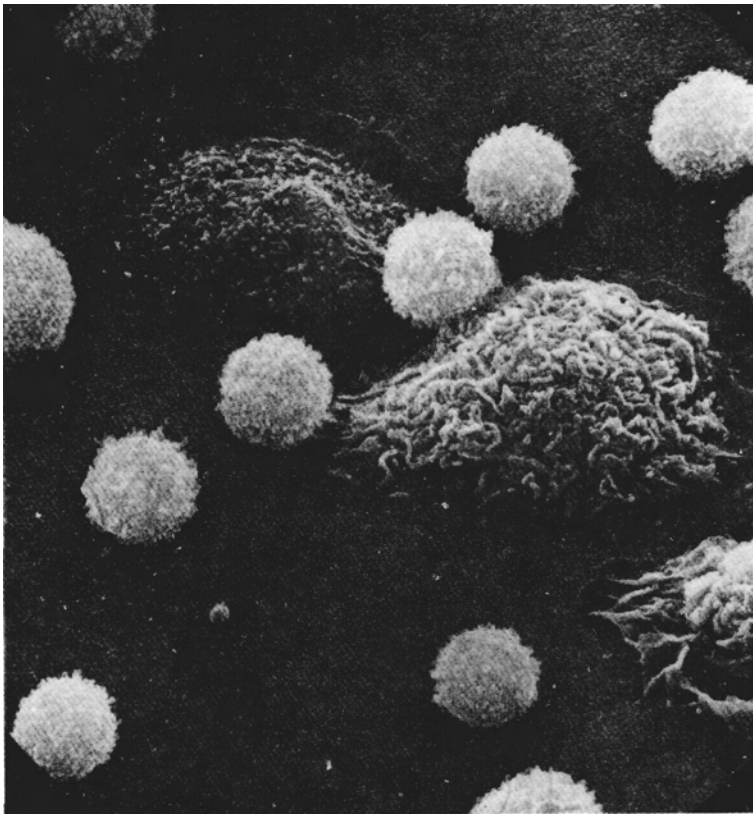
В начале XIX в., после создания более совершенных микроскопов, основной унифицирующей единицей в биологии стала клетка. Все организмы могли рассматриваться как одиночные, свободно живущие клетки или как сообщество клеток. Постоянное усовершенствование оптических систем микроскопа и новаторские методы подготовки и окрашивания материала позволяли все более детально описывать содержимое клеток (рис. I.1). Эти исследования показали, что клетки окружены мембраной, имеющей определенную структуру, что они неодинаковы по форме и что в них имеются четко различимые структуры, среди которых наиболее удивительной является **ядро**, содержащее в свою очередь небольшие, хорошо окрашиваемые структуры, названные **хромосомами**. Позднее обнаружили, что некоторые микроорганизмы (бактерии) не имеют ядра. Было установлено, что новые клетки появляются только в результате деления предсуществующих клеток.

ЖИВОТНАЯ КЛЕТКА





Б



В

**РИС. 1.1.**

А. Схематическое изображение типичных животных и растительных клеток. Б. Схематическое изображение бактериальной клетки. В. Макрофаги (клетки большего размера) и лимфоциты (клетки меньшего размера), сфотографированные в электронном микроскопе при увеличении 6000. (С любезного разрешения Emma Shelton, Jan Ornstein.)

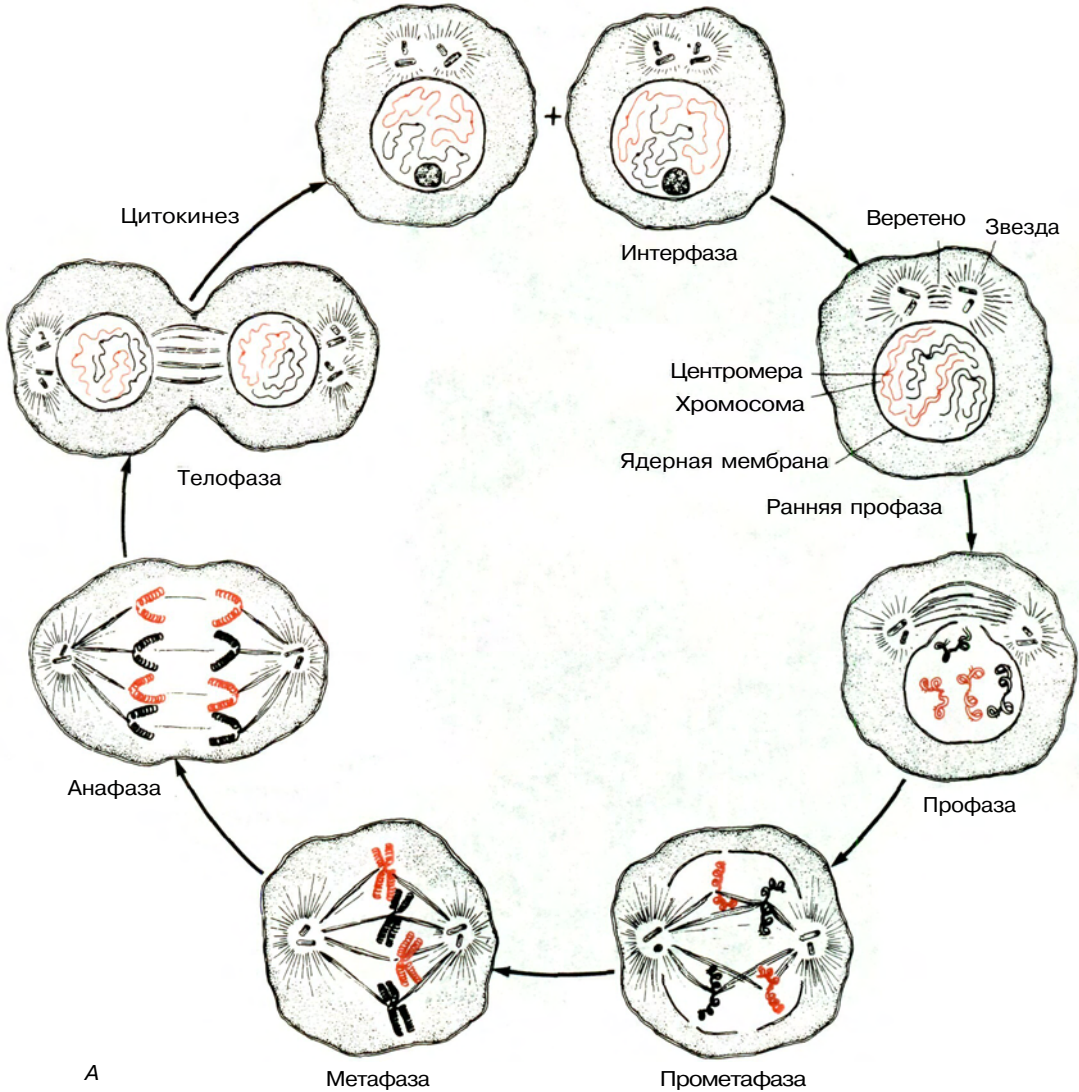
В настоящее время все живые организмы подразделяют на две группы. Первая — **эукариоты** — многоклеточные организмы, клетки которых содержат оформленное ядро; внутри ядра заключены хромосомы — хранители генетической информации. Вторая — **прокариоты** — представлена одноклеточными бактериями, лишенными ядра, с хромосомами, находящимися в цитоплазме. За немногими исключениями, все клетки многоклеточного орга-

низма содержат одинаковый полный набор хромосом. Эукариотические организмы имеют более сложное строение и, как правило, содержат больше генетической информации. Кроме того, эукариоты способны к истинному половому воспроизведению и для многих из них этот способ обязателен для образования потомства. Одним из важных моментов процесса полового размножения является наличие в дочерних ядрах двух копий каждой хро-

мосомы; такие эукариотические клетки называются **диплоидными**. Прокариоты, содержащие только одну хромосому, называются **гаплоидами**. При некоторых обстоятельствах у прокариот наблюдаются процессы, аналогичные по результату процессу оплодотворения у эукариот, вследствие которых они могут стать частично диплоидными; эти

процессы широко используются в генетических исследованиях.

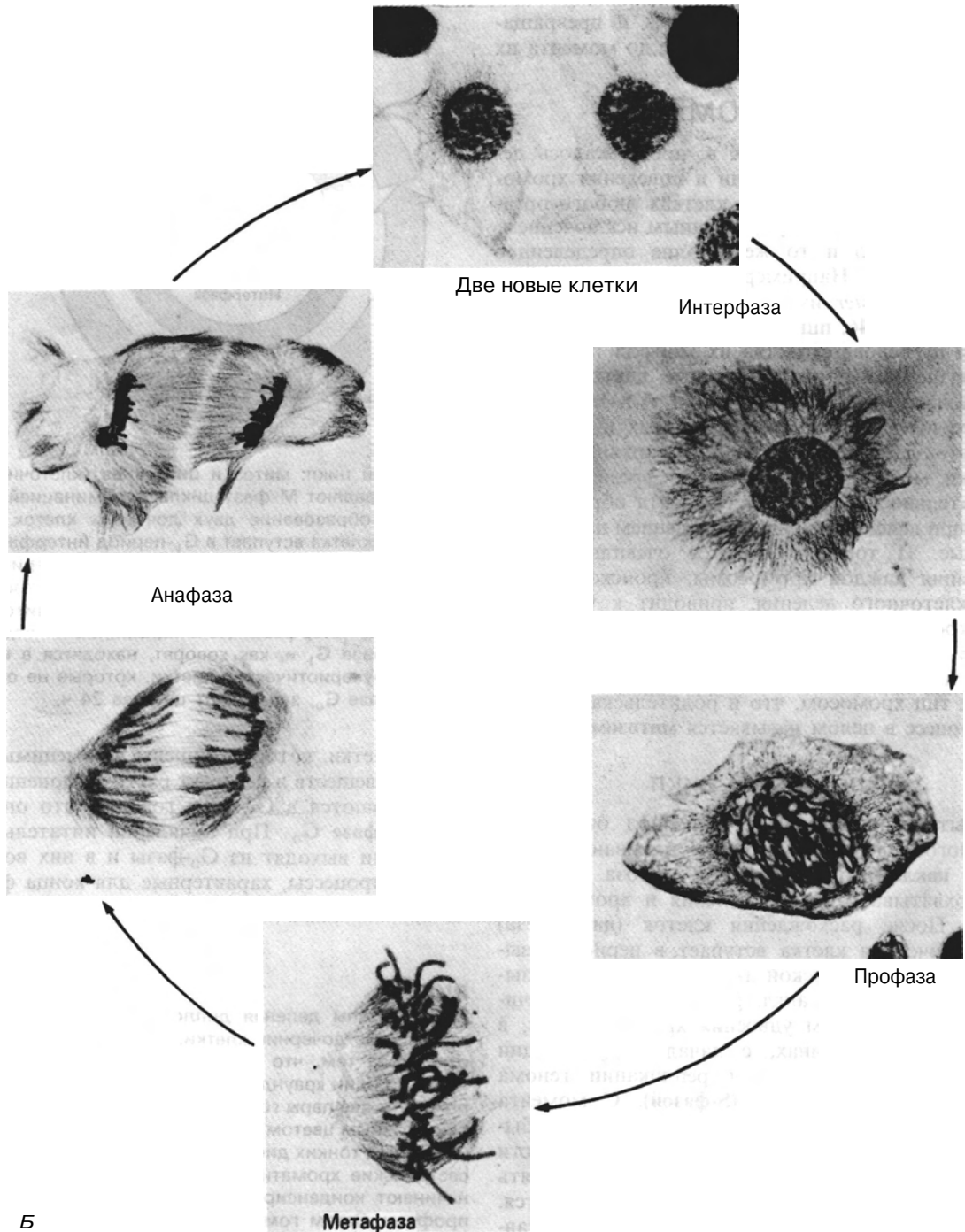
Сразу после принятия клеточной теории в изучении живых организмов выделились три направления: исследование хромосом, статистический анализ наследования одиночных признаков, выделение и характеристика компонентов хромосом. Эти на-



**РИС. 1.2.**

Митоз: этапы деления диплоидной клетки. А. На схеме показаны две пары гомологичных хромосом (они выделены разным цветом). Каждый член пары проходит через митоз как независимая единица. Во время интерфазы хромосомы имеют вид тонких, диффузных нитей, которые в норме трудно визуализировать. В это время происходит дупликация хромосом. Реплицированные хромосомы конденсируются в дискретные структуры, которые в стадии профаза легко различимы. В проме-

тафазе хромосомы еще более конденсируются, ядерная оболочка разрушается и хромосомы оказываются связанными с веретеном. Дуплицированные структуры остаются соединенными друг с другом в районе, называемом центромерой (на рисунке она отмечена точкой). Обе копии каждой хромосомы – сестринские хроматиды – разделяются по всей длине, за исключением центромеры. К метафазе пары сестринских хроматид выстраиваются у центра веретена, но их центромеры еще



Б

остаются соединенными. В анафазе пары сестринских хроматид разделяются и каждый член пары движется по направлению к полюсу веретена. В это же время и нити веретена, и клетка начинают растягиваться. Когда в телофазе хроматиды достигают противоположных полюсов, вокруг каждого набора хроматид формируется новая ядерная оболочка и начинается деконденсация хромосом. Наконец, плазматическая мембрана разде-

ляет два ядра и окружающую цитоплазму на две клетки (цитокнез). Хромосомы приобретают растянутую, диффузную форму, типичную для интерфазы, и процесс деления начинается снова. Б. Микрофотографии митоза в клетках лилии *Haemanthus katherinae*. Клетки окрашены иммунозолотом/серебром. Увеличение 600. (С любезного разрешения А.С. Вайер.)



правления параллельно развивались и превращались в важные научные дисциплины до момента их слияния в середине нашего века.

## ХРОМОСОМЫ

Во второй половине XIX в. продолжалось детальное изучение морфологии и поведения хромосом. Оказалось, что во всех клетках любого организма, за одним лишь существенным исключением, содержится одно и то же, вполне определенное число хромосом. Например, плодовая мушка *Drosophila melanogaster* имеет 8 хромосом, человек и летучая мышь—46, пшеница—20, носорог—84. Хромосомы на основе сходства их морфологии могут быть разделены на гомологичные пары: 4 пары у *D. melanogaster*, 23—у человека и т.д. Микроскопическое исследование фиксированных и окрашенных клеток дает лишь статическую картинку, но эти картинки можно расположить во временной последовательности, начиная с момента образования клетки при делении и кончая ее делением на две себе подобные. И тогда становится очевидным, что дубликация каждой хромосомы, происходящая в цикле клеточного деления, приводит к удвоению числа хромосом (рис. 1.2). При делении этот удвоенный набор распределяется таким образом, что каждая из двух дочерних клеток получает такое же число и тип хромосом, что и родительская клетка. Весь процесс в целом называется **митозом**.

## Клеточный цикл

События, происходящие в период от одного клеточного деления до другого, называются **клеточным циклом** (рис. 1.3). Фаза митоза (**М-фаза**) цикла охватывает период деления и хромосом, и клеток. После расхождения клеток (**цитокинеза**) каждая дочерняя клетка вступает в период повышенной биосинтетической активности—в так называемую **G<sub>1</sub>-фазу** (от англ. gap). G<sub>1</sub>-фаза заканчивается перед началом удвоения хромосом, или, в молекулярных терминах, с началом дубликации хромосомной ДНК; период репликации генома называется фазой синтеза (**S-фазой**). С момента завершения S-фазы в клетках инициируются события, характерные для митотической профазы,— части цикла, называемой **G<sub>2</sub>-фазой**. В конце концов опять начинаются митоз и цитокинез, и цикл повторяется. Как правило, G<sub>1</sub>-, S- и G<sub>2</sub>-периоды, вместе составляющие **интерфазу**, занимают около 90% времени клеточного цикла, а М-фаза—менее 10%. Полное время прохождения клеточного цикла в клетках разного типа сильно варьирует (от минут до суток) в зависимости от условий роста. Основным показателем продолжительности всего цикла является продолжительность G<sub>1</sub>-фазы. Например, покоя-

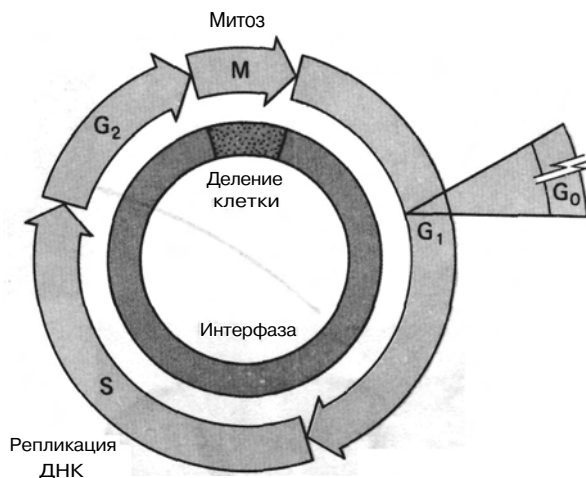


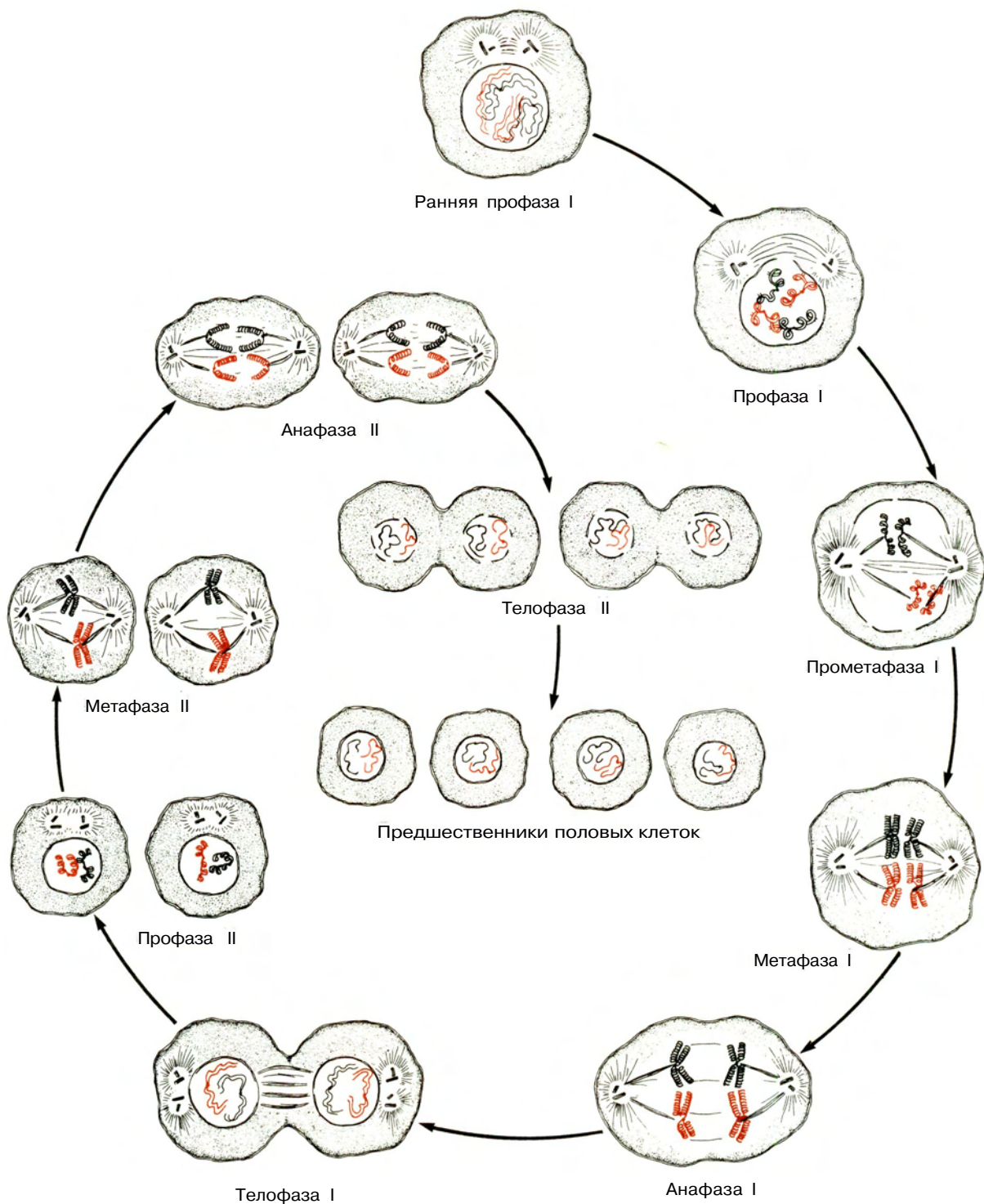
РИС. 1.3.

Клеточный цикл: митоз и цитокинез (клеточное деление) составляют М-фазу цикла, кульминацией которой является образование двух дочерних клеток. Каждая дочерняя клетка вступает в G<sub>1</sub>-период интерфазы и может начать новый клеточный цикл. За периодом G<sub>1</sub> следует S-фаза, во время которой ДНК и хромосомы дублируются, и далее—фаза G<sub>2</sub>. Начало митоза означает конец интерфазы. Покоящиеся клетки задерживаются в фазе G<sub>1</sub> и, как говорят, находятся в фазе G<sub>0</sub>. Обычно эукариотические клетки, которые не остановились в фазе G<sub>0</sub>, завершают цикл за 24 ч.

щиеся клетки, которые лишены незаменимых питательных веществ и не могут расти экспоненциально, задерживаются в G<sub>1</sub>-фазе; говорят, что они находятся в **фазе G<sub>0</sub>**. При появлении питательных веществ они выходят из G<sub>0</sub>-фазы и в них возобновляются процессы, характерные для конца фазы G<sub>1</sub> с переходом в S-фазу.

РИС. 1.4.

Мейоз: этапы деления диплоидной клетки на четыре гаплоидные дочерние клетки. Этот процесс отличается от митоза тем, что включает два клеточных деления и только один «раунд» репликации хромосом. На схеме показаны две пары гомологичных хромосом (они выделены разным цветом). Во время интерфазы хромосомы имеют вид тонких диффузных нитей. После репликации сестринские хроматиды остаются тесно связанными и начинают конденсироваться, что указывает на начало профазы. Затем гомологичные пары сестринских хроматид приходят в тесное соприкосновение, образуя тетрады; этот процесс называется синапсисом. Начало мейотической метафазы I характеризуется дальнейшей конденсацией хромосом и дезинтеграцией ядерной мембраны. В анафазе I члены гомологичной пары сестринских хроматид начинают перемещаться к разным полюсам удлиняющейся клетки. К концу телофазы I и клеточного деления I образуются две дочерние клетки,



в каждой из которых имеется по одной гомологичной паре сестринских хроматид. Второй раунд клеточного деления происходит без дополнительной дупликации хромосом и начинается с профазы II, с переходом в метафазу II. В стадии анафазы II две сестринские хроматиды, которые оставались до этого момента вмес-

те, начинают перемещаться к противоположным концам удлиняющейся клетки. После телофазы II и клеточного деления II образуются четыре гаплоидные клетки – предшественники половых клеток. В каждую дочернюю клетку попадает только по одной хромосоме из исходных гомологичных пар.

## Мейоз и образование гамет

Образование яйцеклеток и сперматозоидов подразумевает уменьшение нормального числа хромосом ровно в половину; этот процесс называется **мейозом** (рис. I.4). Гаметы, или половые клетки, гаплоидны, т.е. в них содержится по одному члену каждой пары гомологичных хромосом, и, таким образом, только половинное число хромосом каждого из родителей попадает во все другие, **соматические**, клетки организма потомка. Распределение хромосом в мейозе происходит случайно, поэтому любой из членов гомологичной пары может оказаться во вновь образовавшихся зародышевых клетках.

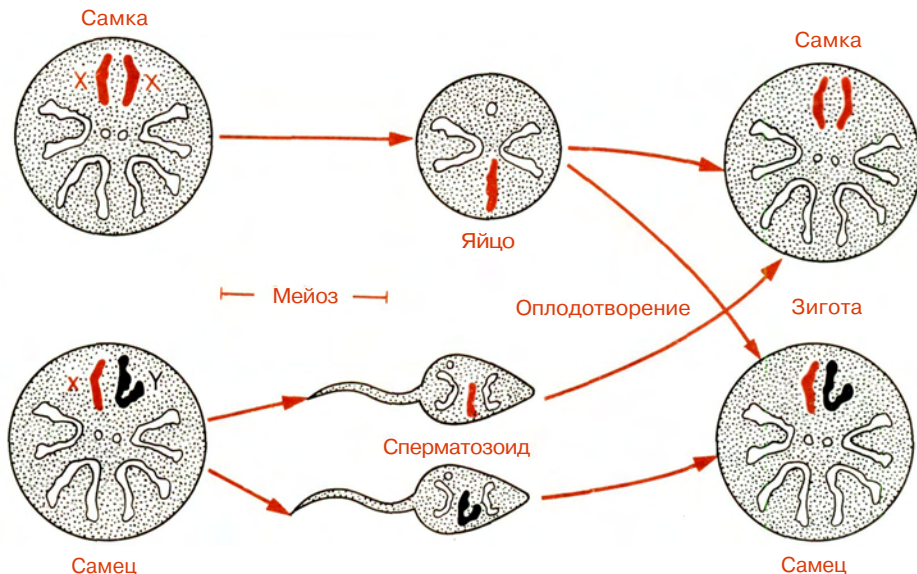
При оплодотворении гаплоидные наборы хромосом сперматозоидов и яйцеклеток объединяются (рис. I.5). Таким образом восстанавливается полный набор гомологичных хромосомных пар, каждый из членов которых произошел из яйцеклетки и из сперматозоида соответствующих родителей. Диплоидное состояние оплодотворенной яйцеклетки (**зиготы**) поддерживается далее во всех соматических клетках механизмом митотического деления. Иногда зрелые организмы могут развиваться из неоплодотворенных гаплоидных яйцеклеток или из оплодотворенных яйцеклеток с неполным набором родительских хромосом. Как уже отмеча-

лось, любой из членов гомологичной пары может попасть в функциональную гамету. В зрелую яйцеклетку или сперматозоид попадает по одному члену каждой пары в процессе редукции числа хромосом в мейозе.

## Строение хромосом

Легче всего наблюдать метафазные хромосомы. Под микроскопом их фотографируют или зарисовывают (рис. I.6). В этой стадии хромосомы наиболее сконденсированны и образуют дискретные структуры. У многих организмов индивидуальные хромосомы и их гомологи легко различимы по размеру и форме. Каждая метафазная хромосома действительно состоит из двух идентичных частей, называемых **сестринскими хроматидами**, поскольку дупликация хромосомной ДНК протекает как раз перед метафазой, в S-фазе клеточного цикла.

У хромосомы имеется перетяжка, называемая **центромерой**. Положение центромеры для каждой хромосомы строго определено. С центромерой связаны специфические хромосомные функции; это последняя точка, соединяющая плечи сестринских хроматид перед полным расхождением при митотическом или II мейотическом делении (рис. I.2 и I.4). Сами плечи имеют вид отдельных образований задолго до расхождения центромер в анафазе.

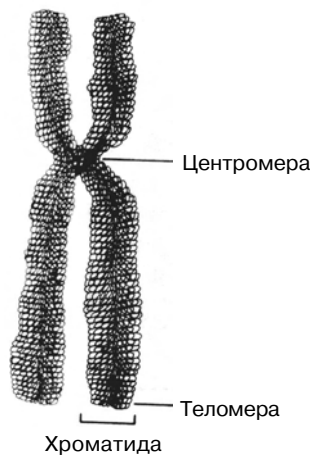


**РИС. I.5.**

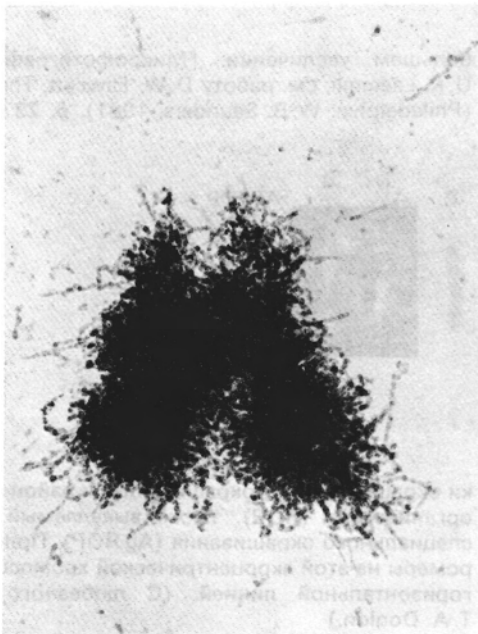
Образование гаплоидных гамет при мейозе и слияние двух гамет с образованием диплоидной клетки при оплодотворении. Обратите внимание на то, что у *D. melanogaster*, рассмотренной здесь в качестве примера, как и у других организмов, включая млекопитающих, две половые хромосомы у самца (X и Y) не гомологичны друг другу. При мейозе формируются два типа

сперматозоидов, из которых один несет X-, а другой – Y-хромосому. У самок, несущих пару X-хромосом, в результате мейоза образуются гаметы одного типа. Пол потомков зависит от того, какую из хромосом – X или Y – несут оплодотворяющие сперматозоиды. У некоторых организмов (например, у птиц) негомологичную, определяющую пол хромосому несет самка.

Различие между областью центromеры и плечами хромосом становится очевидным после обработки определенными красителями. После окрашивания центromеры выглядят более плотными и компактными по сравнению с плечами (рис. 1.7). Такие плотные, интенсивно окрашиваемые хромосомные области называются **гетерохроматиновыми**. Гетерохроматин центromеры можно наблюдать после окрашивания даже в плохо различимых интерфазных хромосомах. Другие, негетерохроматиновые области хромосом принято называть **эухроматиновыми**. Эухроматиновые области окрашиваются гораздо менее интенсивно, чем гетерохроматиновые.



А



Б

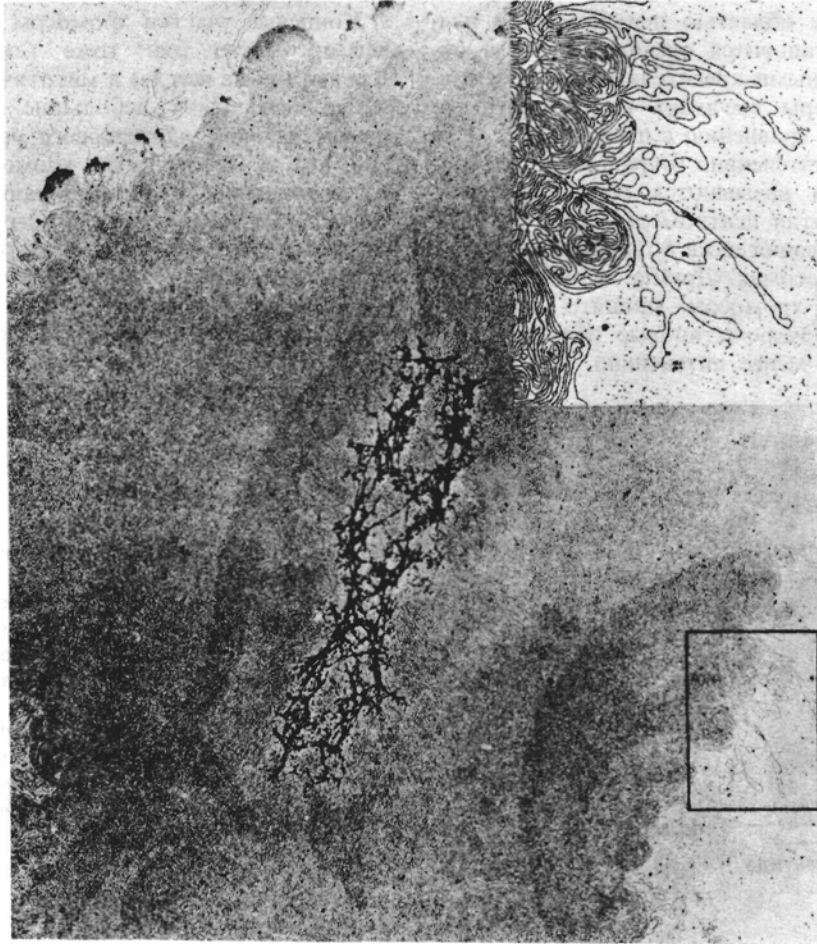
Концевые участки хромосом называются **теломерами**. Часто они тоже гетерохроматиновые. Нередко (но не всегда) в митотических хромосомах можно наблюдать небольшие перетяжки, называемые **районом ядрышкового организатора (ЯОР)** (рис. 1.8). В мейотических хромосомах они имеют вид утолщений. В пределах данного вида районы ядрышковых организаторов встречаются на одной или нескольких специфических хромосомах (и их гомологах), и если они есть, то всегда находятся в одном и том же месте. В  $G_1$ -фазе клеточного цикла некоторые ядрышковые организаторы начинают разрастаться; если их больше, чем один, то такие разросшиеся области объединяются в одну

### РИС. 1.6.

Свойства метафазных хромосом эукариот. А. Схематическое изображение двух копий, или хроматид, дублированной хромосомы. Две хроматиды удерживаются вместе центромерой, которая в данном случае находится примерно в центре хромосомы (метацентрическая хромосома). Б. Электронная микрофотография субметацентрической хромосомы; увеличение 30000. (С любезного разрешения G.F. Bahr.) В. Фотография некоторых хромосом человека, полученная с помощью сканирующего электронного микроскопа. Четко видно, что хромосомы скручены. [С любезного разрешения J.B. Rattner. См. J.B. Rattner, C.C. Lin, *Cell*, **42**, (1985), p. 291.]



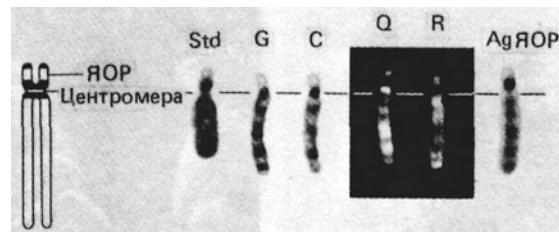
В



Продолжение рис. 1.6 Г

Г. Хромосома, представленная на этой фотографии, содержит меньшее количество гистонов и выглядит, как спутанный клубок. Вверху справа показана ДНК при

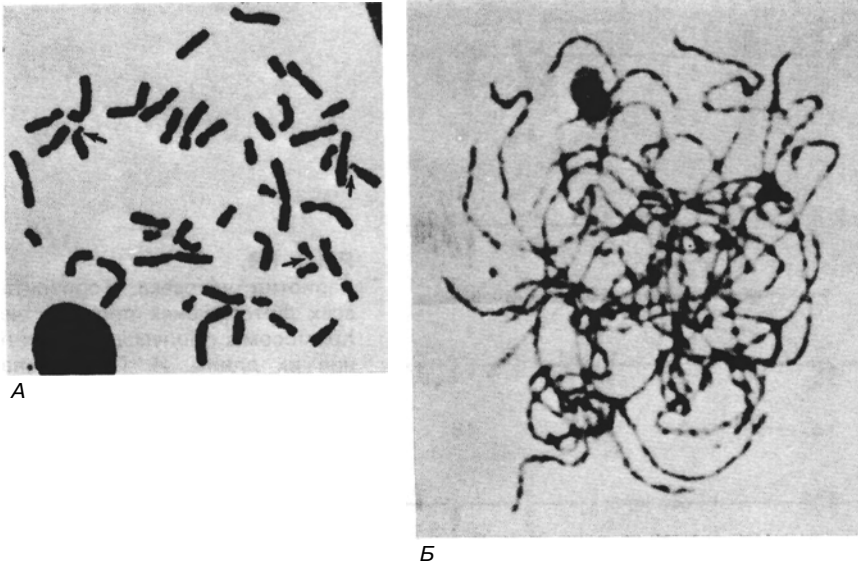
большем увеличении. [Микрофотография получена U.K. Laemmli; см. работу D.W. Fawcett, *The Cell*, 2nd ed. (Philadelphia: W.B. Saunders, 1981), p. 237.]



### РИС. 1.7.

Фотография метафазной хромосомы человека N 13. Видны общее строение хромосомы (Std), характер полос в эухроматиновых плечах, выявляемый после специального окрашивания (три разных метода окрашивания G, Q и R), гетерохроматиновая область в центромере, наблюдаемая благодаря применению особой техни-

ки окрашивания (С-окрашивание) и район ядрышкового организатора (ЯОР), также выявляемый с помощью специального окрашивания (AgЯОР). Положение центромеры на этой акроцентрической хромосоме отмечено горизонтальной линией. (С любезного разрешения Т.А. Donlon.)

**РИС. 1.8.**

Фотографии хромосом, на которых видны районы ядрышковых организаторов (ЯОР). А. Метафазные хромосомы человека, обработанные таким образом, что ЯОР не окрашиваются (показаны стрелками). Видны не все ЯОР. (С любезного разрешения Т.С. Hsu.) Б. Мейотические хромосомы человека, на которых ЯОР имеют вид утолщений после окрашивания серебром. [С. Mirre, M. Hartung, S. Stahl, Proc. Natl. Acad. Sci. USA, **77** (1980), p. 6019.]

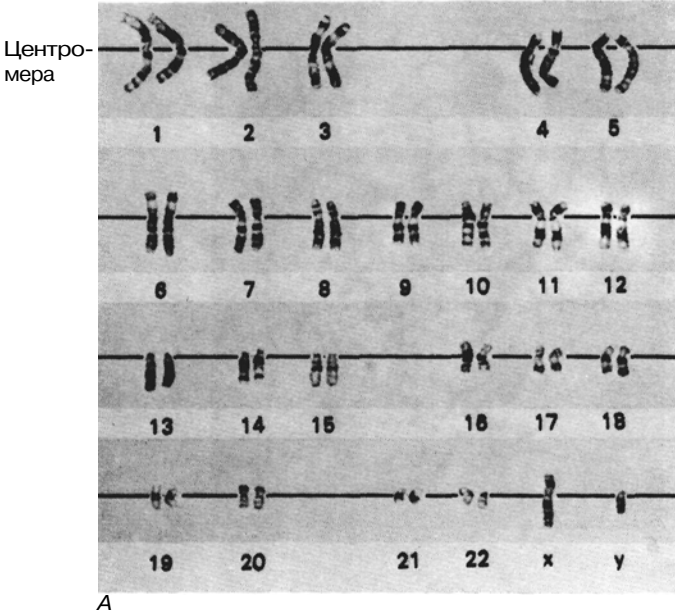
или несколько больших, почти сферических структур — **нуклеолей** (рис. 1.1). Часто в интерфазном ядре только нуклеолы и можно видеть, но с переходом в профазу они постепенно исчезают.

Применение специальных красителей и особых способов окрашивания, разработанных в последние несколько десятилетий, дало возможность выявить достаточно тонкие детали в структуре прометафазных и метафазных хромосом, даже если это довольно мелкие хромосомы млекопитающих. Итак, после окрашивания в каждой хромосоме можно наблюдать уникальное чередование светлых и темных полос; гомологичные хромосомы имеют идентичный рисунок (рис. 1.7). Этот рисунок достоверно воспроизводится, и каждую хромосому в наборе можно идентифицировать. На рис. 1.9 представлен полный набор прометафазных хромосом в клетке человека. На этом изображении, называемом **кариотипом** человека, отражены относительный размер и форма хромосом наряду с положением центromеры и характерным видом полос.

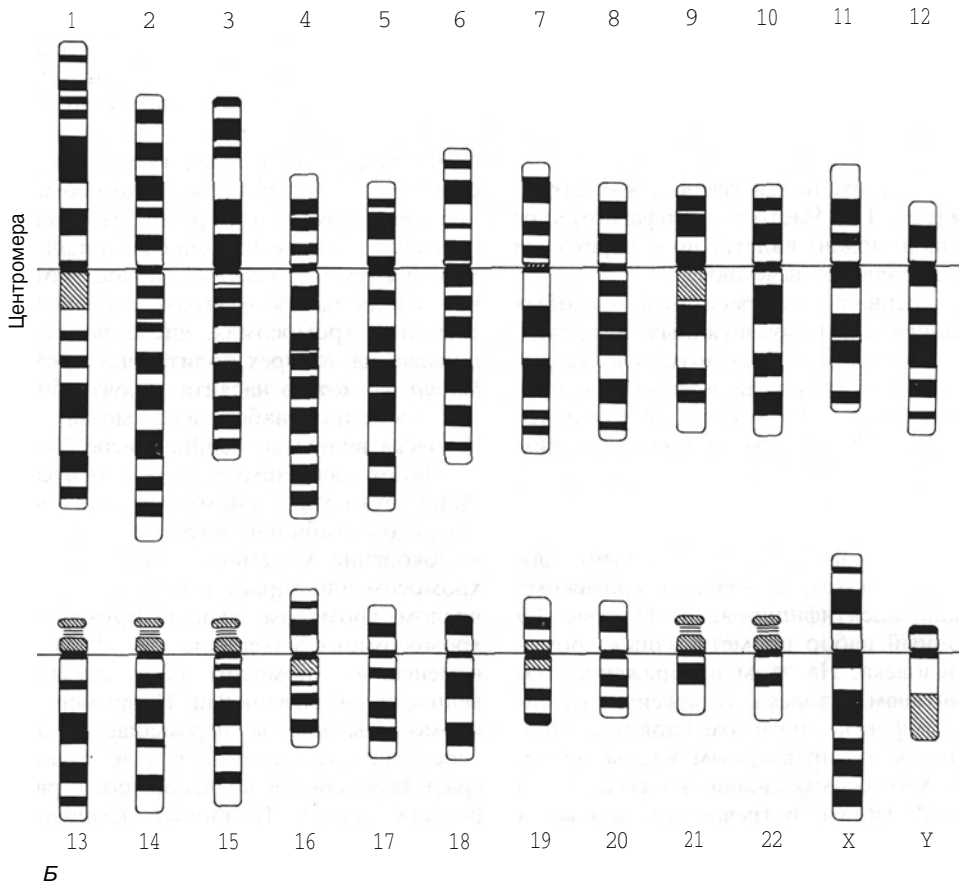
В интерфазе хромосомы сильно растягиваются и, как правило, не видны. Встречаются, однако, и существенные исключения, которые уже много лет интенсивно исследуются. Секреторные клетки ли-

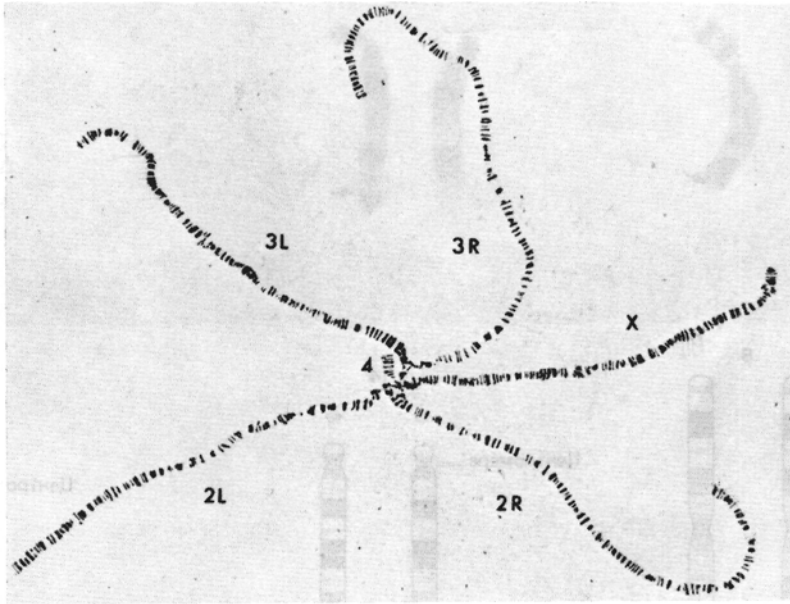
чинок некоторых насекомых (например, *D. melanogaster*) разрастаются до огромных размеров и проходят несколько S-фаз без митоза и клеточного деления. В результате формируется комплекс из множества, иногда вплоть до тысячи, хроматид, которые остаются сцепленными и лежат рядом друг с другом, образуя толстые нити, называемые **политенными** хромосомами (рис. 1.10). Так же как и все интерфазные хромосомы, политенные хромосомы растянуты значительно сильнее, чем конденсированные метафазные хромосомы. При окрашивании политенных хромосом специальными красителями выявляется определенный рисунок чередования темных и светлых полос. В отличие от того, что наблюдается в высококонденсированных метафазных хромосомах, число полос огромно. Например, на четырех политенных хромосомах *D. melanogaster* можно насчитать почти 5000 темных полос, а в полном наборе из 23 метафазных хромосом человека видны по крайней мере 2000 полос.

Четко различимые морфологические признаки индивидуальных прометафазных и политенных хромосом стабильно воспроизводятся из поколения в поколение у данного вида. Необычная форма хромосом или характер полос наряду с атипичным числом хромосом сигнализируют о повреждении хромосомного материала (рис. 1.11). Наличие таких измененных хромосом часто связано с наследственными заболеваниями. Например, сегмент одной хромосомы иногда перемещается на совершенно неродственную хромосому, и такие перестройки сразу выявляются по необычному размеру или характеру полос. Подобные транслокации иногда бывают реципрокными, т.е. две неродственные хромосомы могут обмениваться фрагментами. Дру-

**РИС. 1.9.**

Кариотип человека. Горизонтальные линии на всех фотографиях проходят через центромеры. Хромосомы пронумерованы в порядке уменьшения их длины. А. Полный набор метафазных хромосом мужчины, окрашенных способом, который выявляет особенности рисунка сегментации. Представлены оба члена каждой гомологичной пары. (С любезного разрешения Uta Francke.) Б. Идиограмма хромосом, построенная на основании кариотипа, представленного на рис. А. [D.G. Harnden, H.P. Klinger, eds., *An International System for Human Cytogenetic Nomenclature* (Basel: S. Karger, AG., 1985), p. 50.]



**РИС. 1.10.**

Комбинированная фотография (фотомонтаж), на которой показан полный набор гигантских политенных хромосом, обнаруженных в клетках слюнных желез *D. melanogaster*. Все хромосомы соединены своими центромерами (хромоцентр). Каждая из этих интерфазных хромосом состоит из плотно сплетенных множественных хроматид, которые являются репликами обоих членов гомологичных пар, лежащих параллельно

друг другу. Полосы выявляются при окрашивании; их характер и относительная ширина уникальны для каждой хромосомы. Обратите внимание на то, что каждая полоса проходит через весь пучок. [ G. Lefevre, Jr., as published in Chapter 2, Figure 3, in M. Ashburner, E. Novitski, eds., *The Genetics and Biology of Drosophila*, volume 1a (London: Academic Press, 1976).]

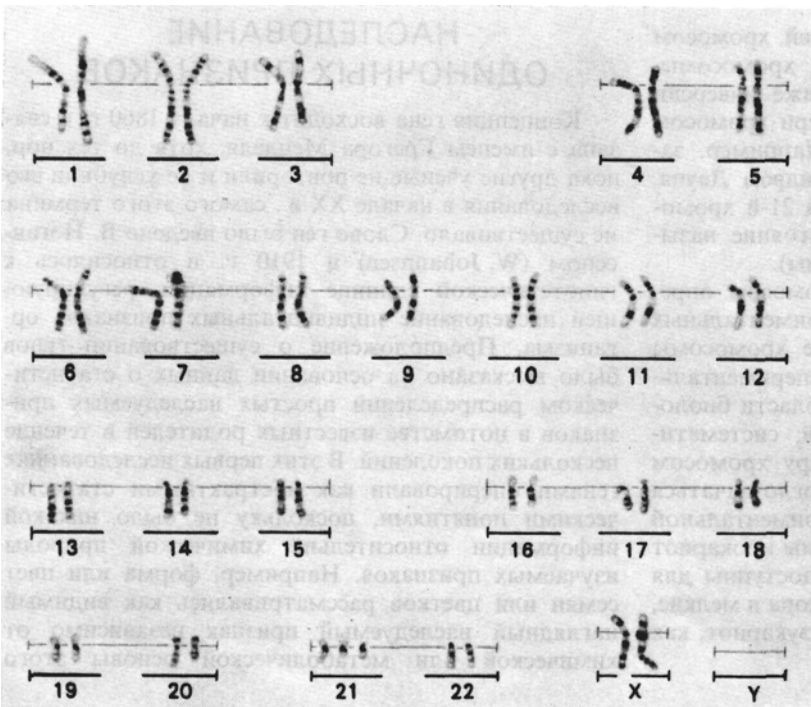
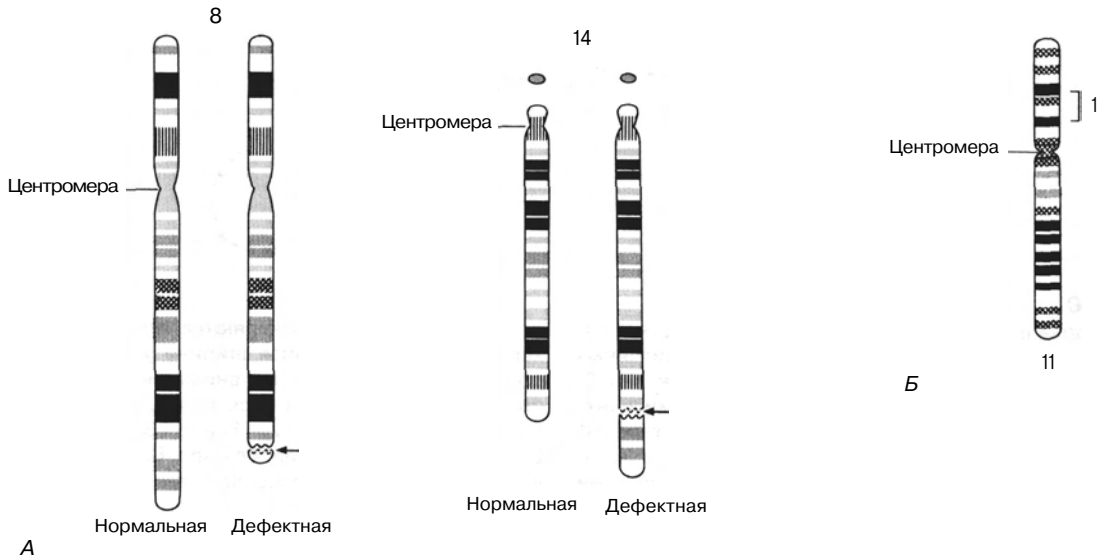
гим примером изменений, или **аббераций**, хромосом служат делеции части нормальной хромосомы, дубликации некоторых областей и даже инверсии сегментов. Иногда наблюдаются потери хромосом или, напротив, появление лишних. Например, заболевание человека, известное как **синдром Дауна**, обусловлено присутствием трех копий 21-й хромосомы вместо обычных двух (это состояние называют также трисомией 21-й хромосомы).

Успехи в изучении структуры хромосом определялись выбором подходящих экспериментальных объектов. Так, огромные политенные хромосомы *D. melanogaster* стали излюбленной экспериментальной системой еще на заре развития области биологии, именуемой теперь **цитогенетикой**; систематическое изучение небольших по размеру хромосом человека и других млекопитающих могло начаться лишь с усовершенствованием экспериментальной техники в начале 50-х годов. Хромосомы прокариот не видны в световом микроскопе; недоступны для анализа с помощью светового микроскопа и мелкие, диффузные хромосомы таких низших эукариот, как дрожжи и трипаносомы.

## НАСЛЕДОВАНИЕ ОДИНОЧНЫХ ПРИЗНАКОВ

Концепция гена восходит к началу 1860 г. и связана с именем Грегора Менделя, хотя до тех пор, пока другие ученые не повторили и не углубили его исследования в начале XX в., самого этого термина не существовало. Слово **ген** было введено В. Йогансенем (W. Johannsen) в 1910 г. и относилось к гипотетической единице информации, регулирующей наследование индивидуальных признаков организма. Предположение о существовании генов было высказано на основании данных о статистическом распределении простых наследуемых признаков в потомстве известных родителей в течение нескольких поколений. В этих первых исследованиях генами оперировали как абстрактными статистическими понятиями, поскольку не было никакой информации относительно химической природы изучаемых признаков. Например, форма или цвет семян или цветков рассматривались как видимый наглядный наследуемый признак независимо от химической или метаболической основы этого





**РИС. 1.11.**

Аберранные хромосомы, являющиеся причиной некоторых заболеваний человека. На рисунках А и Б вверху показаны пары нормальных (слева) и аберранных (справа) хромосом, а внизу — соответствующие идиограммы. На этих идиограммах рисунок сегментации представлен более детально, чем на рис. 1.9. А. Транслокация (стрелка) фрагмента хромосомы 8 на хромосому 14 при лимфоме Беркитта. Б. Делеция (квадратная скобка) части хромосомы 11 в клетках опухоли Вилмса у человека. (Рис. А и Б любезно предоставлены U. Franke.) В. Три копии 21-й хромосомы в кариотипе пациентки с синдромом Дауна. (С любезного разрешения Т.А. Donlon.)

свойства. Тем не менее логический интеллектуальный фундамент, заложенный Менделем и его последователями, вполне соответствует нашим теперешним представлениям о химической структуре генов и тому, как эта структурная информация воплощается в свойства организма.

## Независимая сегрегация и независимое комбинирование

Взгляд Менделя на наследственность у эукариот определялся двумя главными обнаруженными им явлениями. Первое—существование независимой сегрегации. Любой организм содержит пару генов для любого одиночного наследуемого признака, при этом каждый из членов пары имеет либо отцовское, либо материнское происхождение. В каждом поколении члены каждой пары генов расходятся с образованием новых яйцеклеток или сперматозоидов, и во время оплодотворения формируются новые пары генов. Теперь члены пары называются **аллелями** и особенность признака зависит от объединения одинаковых аллелей (про организм, несущий идентичные аллели, говорят, что он **гомозиготен**) или различных (тогда организм **гетерозиготен**). Так, аллельные пары, детерминирующие некий признак, могут быть  $a^1a^1$ ,  $a^1a^2$  или  $a^2a^2$  у разных индивидуумов. Тогда сперматозоиды или яйцеклетки должны содержать аллель  $a^1$  или  $a^2$ . И хотя в каждом отдельном организме имеется не более двух разных аллелей определенного гена, в популяции данного вида циркулирует много различных аллелей. Например, могут существовать множественные формы гена  $a$ :  $a^1$ ,  $a^2$ ,  $a^3$ ,  $a^4$  и т.д., поэтому отдельные индивидуумы могут содержать такие пары, как  $a^1a^2$ ,  $a^2a^2$ ,  $a^3a^2$ ,  $a^1a^4$ ,  $a^4a^5$  и  $a^4a^4$ .

Второе важное наблюдение Менделя касалось независимого комбинирования различных аллельных пар генов, каждая из которых определяет различные признаки. Например, яйцеклетки или сперматозоиды в организме, содержащем аллельные пары  $a^1a^2$  для признака  $a$  и  $b^1b^2$  для  $b$ , могут иметь сочетания аллелей  $a^1b^1$ ,  $a^1b^2$ ,  $a^2b^1$  или  $a^2b^2$ . При образовании гамет сегрегация аллельных пар « $a$ » не зависит от сегрегации аллельных пар « $b$ »; то же самое можно сказать и о других аллельных парах.

## Связь между генами и хромосомами

В начале XX в. была обнаружена корреляция между физическим поведением хромосом и положениями менделевской генетики. Каждый член аллельной пары генов мог быть ассоциирован с одной из хромосом пары, а независимое распределение аллелей можно было объяснить, если считать, что различные аллельные пары находятся

на разных хромосомах. Томас Гент Морган и его коллеги доказали, что у *D. melanogaster* гены ассоциированы с хромосомами. Они выбрали этот организм для генетических исследований, поскольку короткое время генерации и большое число особей в потомстве, получаемом от каждого скрещивания, делает генетический анализ удобным и точным; кроме того, хромосомы *D. melanogaster* легко различимы в световом микроскопе. В ходе экспериментов было установлено, что наследование аллеля, приводящего к появлению у потомства белых, а не обычных красных глаз, всегда сцеплено с наследованием X-хромосом и никогда—Y-хромосомы. Были обнаружены и другие аллели, коррелирующие с разными признаками, также связанные с наследованием X-хромосом, и аллели, наследуемые совместно, сцепленными группами, но независимо от X-хромосомы. Таким образом, стало очевидным, что число групп совместно наследуемых аллелей соответствует числу хромосомных пар. В ходе исследования было установлено, что аллели, ассоциируемые с различными хромосомами, распределяются в потомстве независимо, а группы аллелей, связанные с определенной хромосомой, остаются сцепленными и в потомстве.

## Рекомбинация

Почти одновременно с выявлением групп сцепления были обнаружены и неожиданные исключения. Например, такие аллели, как  $a^1$  и  $b^1$  или  $a^2$  и  $b^2$ , как правило, наследовались сцепленно, но иногда появлялись новые сочетания,  $a^1b^2$  и  $a^2b^1$ , которые наследовались в последующих поколениях. С помощью цитогенетического анализа было установлено, что при мейозе гомологичные хромосомы обвиваются друг вокруг друга, поэтому Морган предположил, что они могут обмениваться между собой частями, давая тем самым новые комбинации сцепленных аллелей (рис. 1.12). Этот процесс получил название **кроссинговера** или **рекомбинации**. Совершенно не зная химической природы этого явления, генетики использовали феномен рекомбинации в качестве основного инструмента генетических исследований. Определение частот рекомбинации между сцепленными парами аллелей у *D. melanogaster* позволило сделать три важных заключения: гены расположены в линейном порядке, и члены аллельных пар обычно занимают одинаковое относительное положение на гомологичных хромосомах; рекомбинация происходит только внутри одной группы сцепления (т.е. между гомологичными хромосомами); частота, с которой два разных сцепленных аллеля перекрещиваются (скажем, связываются узелком), зависит от расстояния между ними на хромосоме (чем дальше они

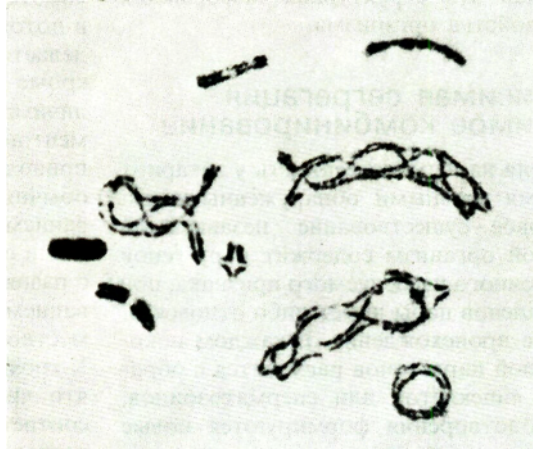
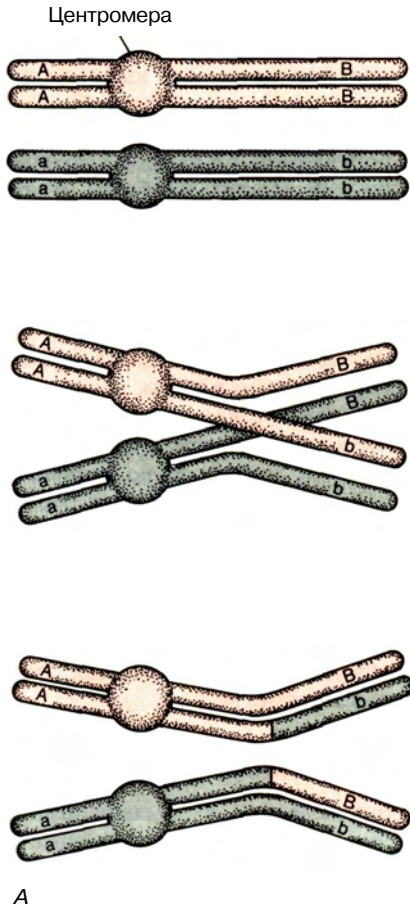


РИС. I.12.

Кроссингвер хромосом, происходящий при мейозе. А. Схематическое представление процесса. А и а, В и b — пары аллелей. Б. Фотография хромосом кобылки (*Chorthippus parallelus*) в диплотенной стадии мейоза (профаза I). Видны гомологичные пары уже дуплицированных хромосом (см. схему). На некоторых парах видно несколько точек перекреста (хиазмы). Отдельные пары уже уплотнились в преддверии метафазы I. (С любезного разрешения В. John.)

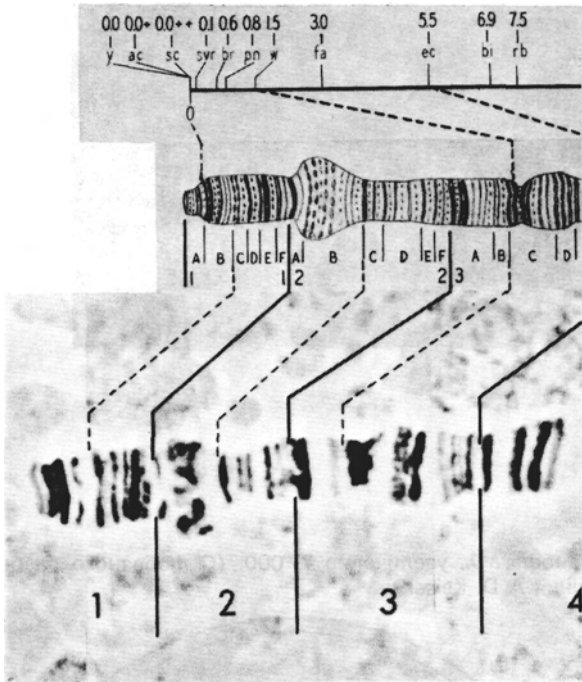
друг от друга, тем выше эта частота). Относительное положение различных генов на хромосоме *D. melanogaster*, а в дальнейшем и других организмов было установлено именно исходя из этих принципов. К 1922 г. Морган и его коллеги смогли картировать несколько сотен генов на четырех хромосомах *D. melanogaster* (рис. I.13).

### Связь между генами и белками

Одно из первых предположений о том, как информация, заключенная в генах, проявляется в специфических свойствах клетки и целого организма, было высказано еще до того, как изобрели слово «ген». В первом десятилетии XX в. английский врач Арчибальд Гаррод заметил, что наследование некоторых метаболических идиосинкразий и других расстройств у людей происходит в соответствии с правилами Менделя. Он предположил, что причиной подобных наследственных расстройств служит недостаток или отсутствие особых ферментов, необходимых для нормального метаболизма.

Тогда же Гаррод высказал гипотезу, что детерминанты наследственности контролируют образование ферментов. Таким образом, способность к синтезу особых ферментов или даже их свойства связывались с генами. Предположение казалось очень заманчивым даже при отсутствии экспериментальных доказательств, потому что оно связывало имеющиеся в то время генетические данные, полученные для мух и растений, с биологией человека.

Дальнейшее развитие идеи Гаррода могли получить лишь с появлением новых экспериментальных подходов. В конце 1930-х годов такие подходы появились благодаря использованию в качестве экспериментальных объектов микроорганизмов. Вначале в центре внимания исследователей оказались низшие грибы из родов *Aspergillus* и *Neurospora*. Эти организмы хорошо росли в определенных условиях культивирования и достаточно быстро размножались. К середине 40-х годов было накоплено и проанализировано достаточно генетических и биохимических данных для того, чтобы прийти к выводу, что наличие или отсутствие фермента



**РИС. I.13.**

Генетическая карта части X-хромосомы (политенной) *D. melanogaster*. Обратите внимание, что каждому гену соответствует определенный участок политенной хромосомы. Мутация в гене *w* (область 3С) обуславливает белый цвет глаз у мухи. [С.В. Bridges, *Journal of Heredity*, **26** (1935), p. 60.]

наследуемо и зависит от экспрессии одного гена. Джордж Бидл и Эдвард Татум (George Beadle, Edward Tatum) обобщили связь между ферментом и геном в виде постулата **один фермент—один ген**. Поскольку ферменты—это белки, а многие белки состоят из более чем одного типа полипептидных цепей, постулат в дальнейшем стал формулироваться как «один полипептид—один ген». Проведенные исследования показали, что некоторые гены кодируют белки, не являющиеся ферментами (в том числе гормоны и структурные белки), а другие гены контролируют образование молекул РНК, которые необходимы для синтеза белков.

Для обоих компонентов этой информационной цепочки часто используются такие термины, как **генотип** и **фенотип**, относящиеся соответственно к гену и признаку, который им кодируется. В общем виде генотип иногда трактуется как вся генетическая информация отдельной клетки или организма. Аналогично и термин фенотип применяется более широко для описания видимых свойств клетки или организма, будь то особые белки или функции либо

морфологические и даже поведенческие признаки. Фенотип, как правило, является результатом взаимодействия между генетической информацией и условиями окружающей среды, в которой она реализуется. Термин **геном** применяется к совокупности хромосом (на молекулярном уровне—к ДНК), свойственных отдельному организму (или любой клетке внутри организма), в отличие от термина генотип, который относится к информации, заключенной в этих хромосомах (или ДНК).

К 1950 г. была обнаружена еще более заманчивая и многообещающая экспериментальная система для исследования связей между генами и функциями клетки. Обычная кишечная бактерия *Escherichia coli* (*E. coli*) имеет примитивные питательные потребности и делится каждые 20–60 мин (в зависимости от условий культивирования), давая в потомстве огромное число клеток ( $10^9$  в 1 мл). У нее было обнаружено множество легко выявляемых генетически контролируемых физиологических признаков. Кроме того, использование мутантов, которые достаточно просто выделить и охарактеризовать, позволило идентифицировать гены, кодирующие специфические функции клетки. Таким образом был открыт путь для более формального генетического анализа и создания генетической карты единственной хромосомы *E. coli*. Еще одним преимуществом *E. coli* оказалось то, что эта бактерия является хозяином для нескольких вирусов (бактериофагов) (рис. I.14), для которых в свою очередь характерно значительное генетическое разнообразие инфекционных свойств.

Бактериофаги, или, для краткости, фаги, оказались еще более удобной системой для генетических исследований. Два или даже больше фагов могут обмениваться фрагментами своих гомологичных геномов, порождая фаговое потомство с новыми генетическими свойствами (рис. I.15). Фаговые геномы даже способны к обратимой интеграции с бактериальной хромосомой. При выщеплении из хромосомы фаг может включить в свой геном часть бактериального генома и, таким образом, стать носителем бактериальных генов. Анализ подобного обмена генетическим материалом показал, что даже такие примитивные организмы обладают упорядоченным геномом и индивидуальные гены могут составить генетическую карту.

## ГЕНЫ И ДНК

Современная биохимическая генетика ведет свое начало от открытия ДНК в 1869 г. Фридрихом Мишером (Friedrich Miescher). Он установил, что вещество, экстрагируемое из гнойной массы и клеточных ядер, химически отличается от белков как по содержанию органического фосфора, так и по



А

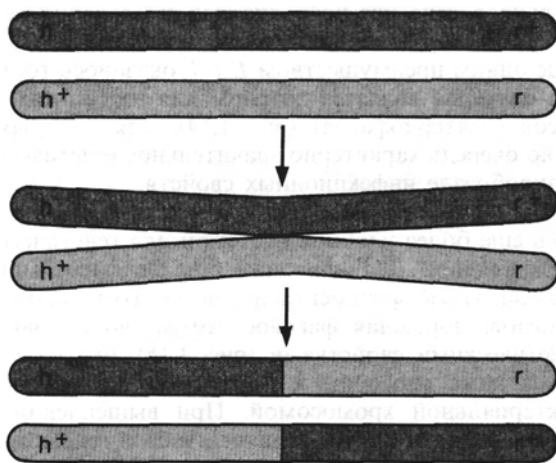


Б

**РИС. I.14.**

А. Электронная микрофотография бактериофага Т4, увеличение 200000. (С любезного разрешения R.C. Williams.) Б. Электронная микрофотография бак-

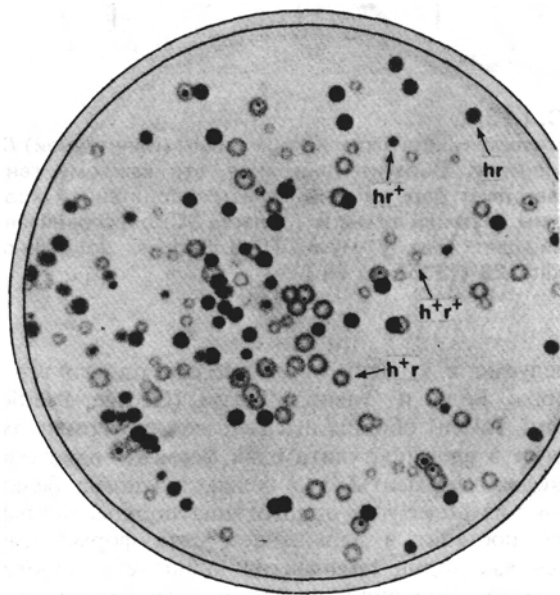
териофага  $\lambda$ , увеличение 77000. (С любезного разрешения A.D. Kaiser.)



А

**РИС. I.15.**

Альтернативные судьбы фаговых хромосом после инфекции. А. Две фаговые хромосомы, несущие различные пары аллелей генов  $h$  и  $r$  ( $hr^-$  или  $h^+r$ ), инфицируют одну бактериальную клетку. В фаговом потомстве после рекомбинации можно получить четыре разных типа хромосом: две родительские,  $hr^+$  и  $h^+r$  (без рекомбинации), и две рекомбинантные,  $hri$  и  $h^+r^+$ . Б. Фаги с разным типом хромосом образуют разные бляшки. Бляшки имеют вид относительно прозрачных пятен на бактериальном газоне. Каждая бляшка начинает образовываться в тот момент, когда из одиночной инфицированной клетки высвобождаются фаговые частицы, которые

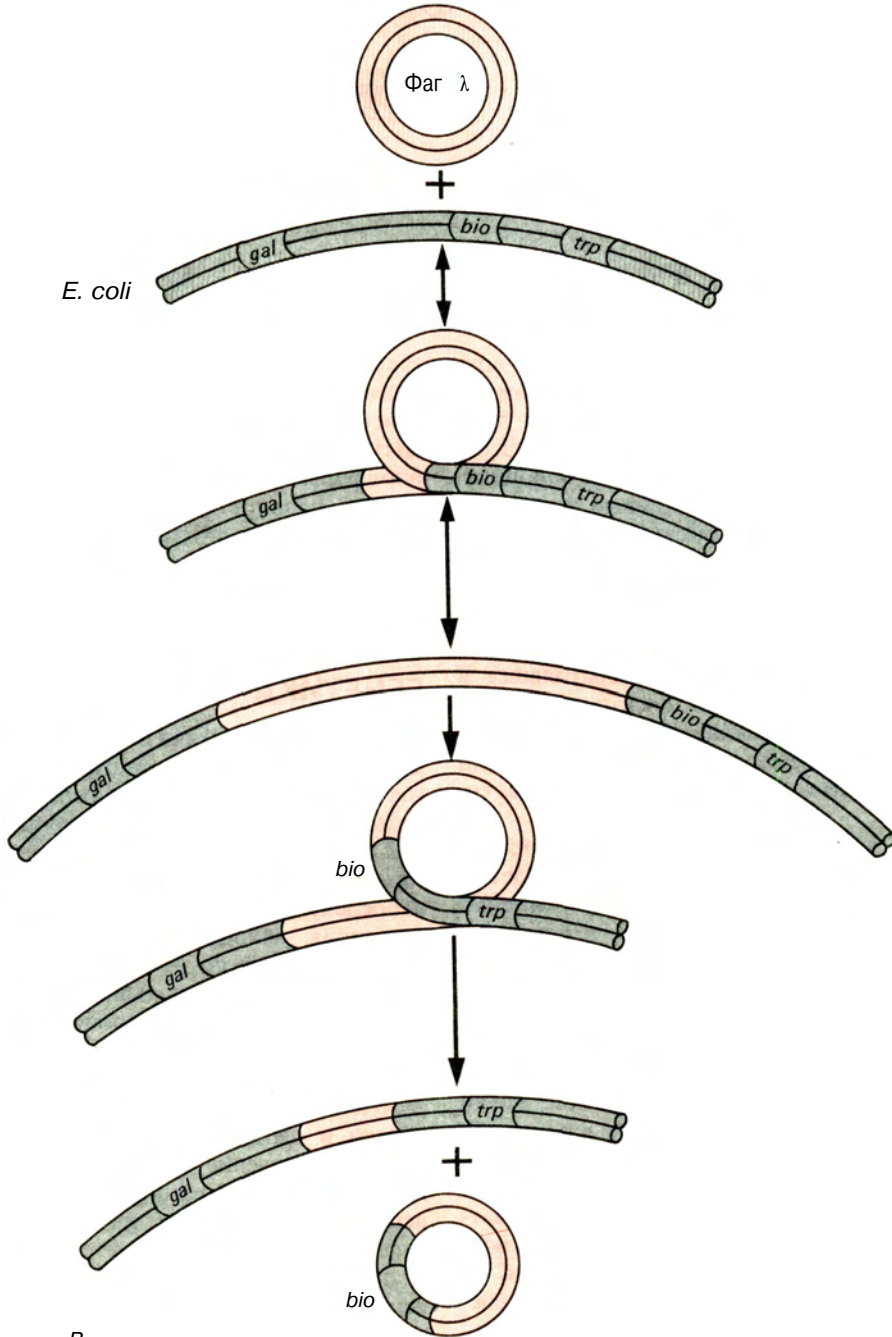


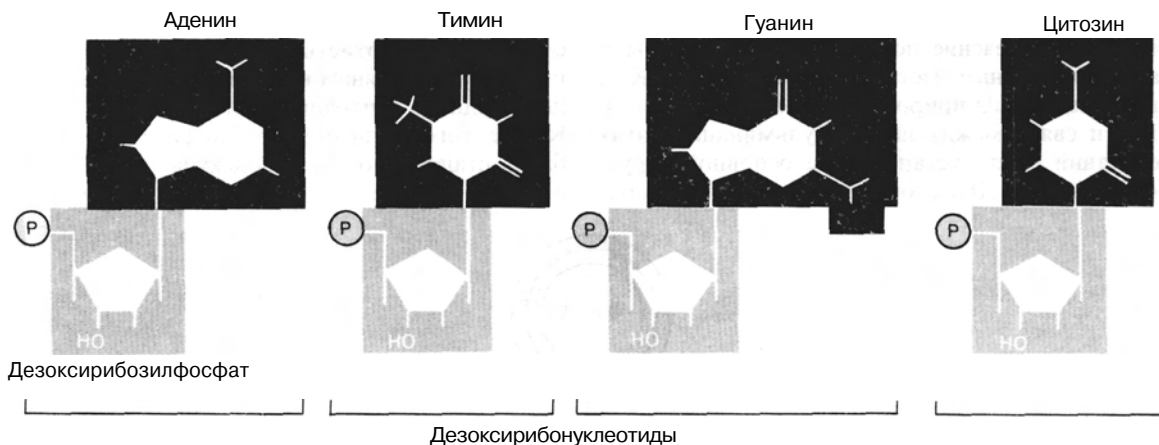
Б

затем заражают окружающие бактерии и лизируют их. [Из G.S. Stent, Molecular Biology of Bacterial Viruses (San Francisco: W.H. Freeman, 1963), p. 335.] В. Рекомбинация между фаговой и бактериальной ДНК. В каждой из ДНК происходит одиночный разрыв, а затем разорванные концы соединяются так, что фаговая ДНК встраивается в бактериальную. Процесс может быть обратимым: фаговая ДНК может выщепляться из бактериальной, причем если это выщепление происходит неточно, то фаговая ДНК приобретает один или несколько бактериальных генов, а какие-то фаговые гены остаются в бактериальной хромосоме.

устойчивости к расщеплению протеолитическими ферментами. В течение последующих 85 лет были разработаны разные методы выделения ДНК с целью исследования природы ее химических составляющих и связи между ними. Кульминацией этих исследований стало установление основной структурной единицы ДНК: она состоит из фосфори-

лированного сахара, дезоксирибозофосфата, соединенного с азотистым основанием—либо с одним из пуринов (аденином или гуанином), либо с одним из пиримидинов (цитозином или тимином) (рис. 1.16). Кроме того, с помощью биофизических методов было установлено, что молекула ДНК—это очень длинная цепочка, остов которой построен из де-



**РИС. I.16.**

Основные структурные элементы ДНК—дезоксирибонуклеотиды.

зоксирибозофосфатных единиц, соединенных друг с другом **фосфодиэфирными** мостиками; к каждой дезоксирибозной единице цепи присоединено либо пуриновое, либо пиримидиновое основание (рис. I.17). В 1953 г. Джеймс Уотсон и Фрэнсис Крик обобщили накопленные к тому времени данные о составе и структуре ДНК, построив ставшую теперь классической теорию двойной спирали ДНК. Импульсом к ее созданию послужило обогатившее науку открытие Освальда Эвери (Oswald Avery) и его коллег, а также Альфреда Херши (Alfred Hershey) и Маргарет Чейз (Margaret Chase), состоявшее в том, что только ДНК является носителем генетической информации. Центральная роль в наследственности, приписываемая хромосомам, могла быть теперь отнесена к ДНК, которую они содержат.

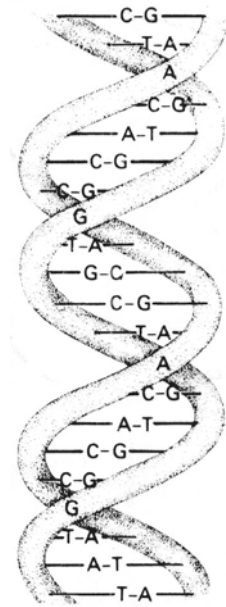
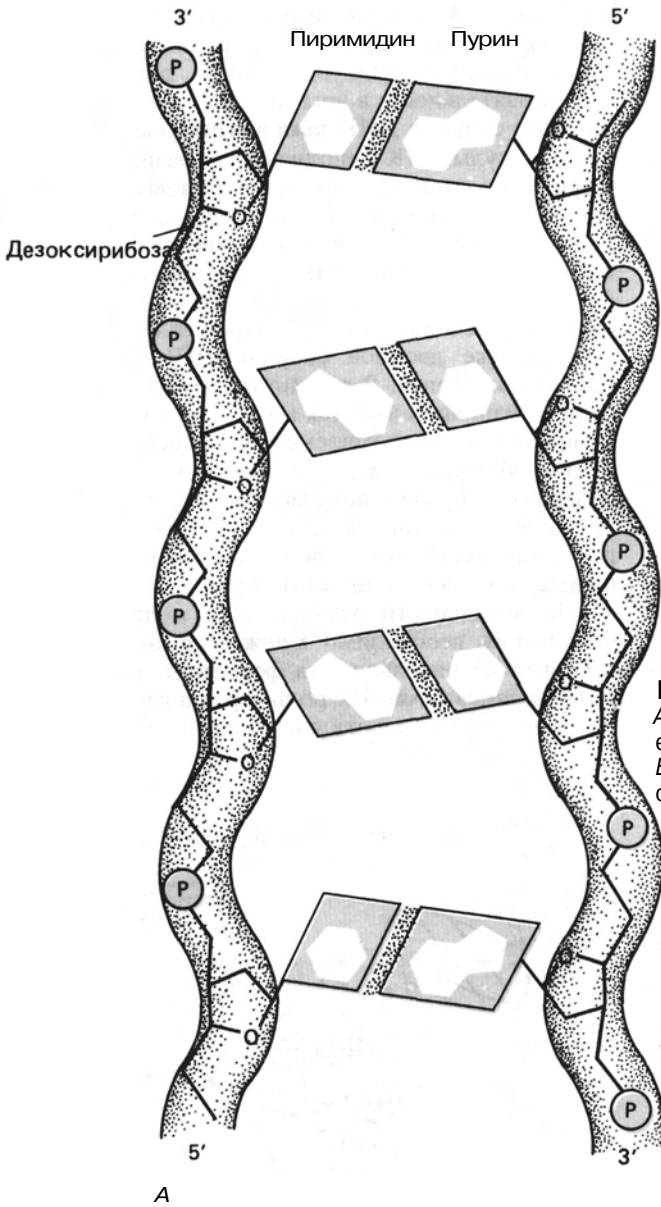
ДНК—не единственная нуклеиновая кислота (или полинуклеотид), обнаруживаемая в клетке. Близкородственные молекулы—рибонуклеиновые кислоты (РНК)—отличаются от ДНК в основном тем, что вместо дезоксирибозы содержат рибозу и чаще имеют одноцепочечную структуру.

Расшифровка структуры ДНК и установление ее центральной роли в наследственности увенчали накопленные наукой данные и позволили генетике из статистической и феноменологической науки превратиться в науку с преобладанием химических и молекулярных направлений развития. Незамедлительная бурная реакция ученых на открытие двойной спирали свидетельствовала об ее адекватности. Модель структуры ДНК не только соответствовала химическим и физическим данным, но и полностью отвечала функциям, присущим генетическому материалу. В линейной последовательности четырех пуринов и пиримидинов могло быть

закодировано огромное количество информации, и в принципе эта структура могла обеспечить свою собственную репликацию. Расшифровка структуры ДНК проливали свет на самые разные аспекты биологии и создавала основу для объяснения многих разноречивых данных, полученных ранее. Она обеспечила фундаментальную целостность при интерпретации огромного многообразия жизненных форм. Раз и навсегда наследственность связывалась с определенной молекулярной структурой.

## ГЕНЕТИКА—МОЛЕКУЛЯРНАЯ НАУКА

Проблемы механизмов переноса, перераспределения и экспрессии генетических признаков, долгое время не находившие решения, с начала 50-х годов перешли на молекулярный и химический уровни. Как реплицируются и рекомбинируют молекулы ДНК? Каким образом они сохраняются в последующих поколениях? Каким способом информация, закодированная в ДНК, обеспечивает образование фенотипических продуктов—белков? Как регулируется считывание информации, закодированной в ДНК, в процессе роста клеток или развития организма и при других физиологических состояниях? Как нарушаются эти процессы при заболеваниях? Эти и еще многие другие вопросы стояли в центре молекулярно-генетических исследований в течение последних 35 лет. Бурный прогресс в первые 20 из них был достигнут благодаря использованию систем прокариот и связан с идентификацией молекулярных структур, участвующих в процессах хранения, поддержания, передачи и использования генетической информации.



Б

**РИС. 1.17.**

А. Цепи ДНК состоят из дезоксирибонуклеотидов, соединенных друг с другом фосфодиэфирными связями. Б. Две такие цепи, закрученные вокруг общей оси, образуют двойную спираль.

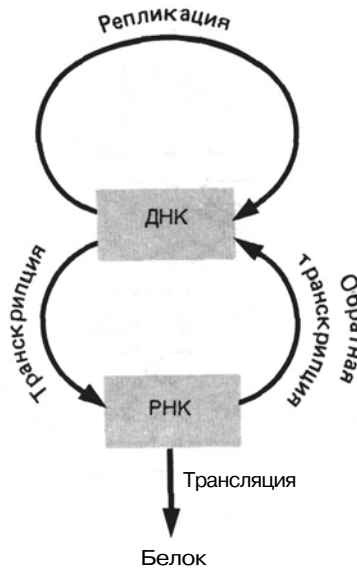
## Перенос генетической информации в клетке

Информационные взаимоотношения между ДНК, РНК и белками теперь точно установлены (рис. 1.18). **Репликация**, с помощью которой создаются идентичные копии родительской молекулы ДНК, обеспечивает генетическую непрерывность в ряду поколений. **Транскрипция** ДНК с образованием РНК опосредует **трансляцию** (перевод) этой информации на уровень белков. Итак, ДНК выполняет две основополагающие функции. Первая—это осу-

ществление своей собственной репликации. Вторая—это формирование фенотипа через образование молекул РНК, участвующих в трансляции информации, содержащейся в ДНК, на язык белков. И, насколько это известно, только у эукариот информация может передаваться в обратном направлении, от РНК к ДНК, посредством процесса, именуемого **обратной транскрипцией**.

В основе переноса информации от ДНК к РНК или от РНК к ДНК лежит универсальная способность нуклеиновых кислот служить **матрицей**. Нуклеиновые кислоты направляют сборку идентич-



**РИС. I.18.**

Информационная связь между ДНК, РНК и белком и процессы, с помощью которых осуществляется передача информации. Заметим, что обратная транскрипция наблюдается только у эукариот.

ных или родственных молекул и непосредственно участвуют в процессе синтеза белка. Насколько известно, информация не передается от белков к нуклеиновым кислотам (т.е. обратной транскрипции не обнаружено). Однако белки помимо самосборки осуществляют важнейшую функцию катализа и информационного переноса между нуклеиновыми кислотами.

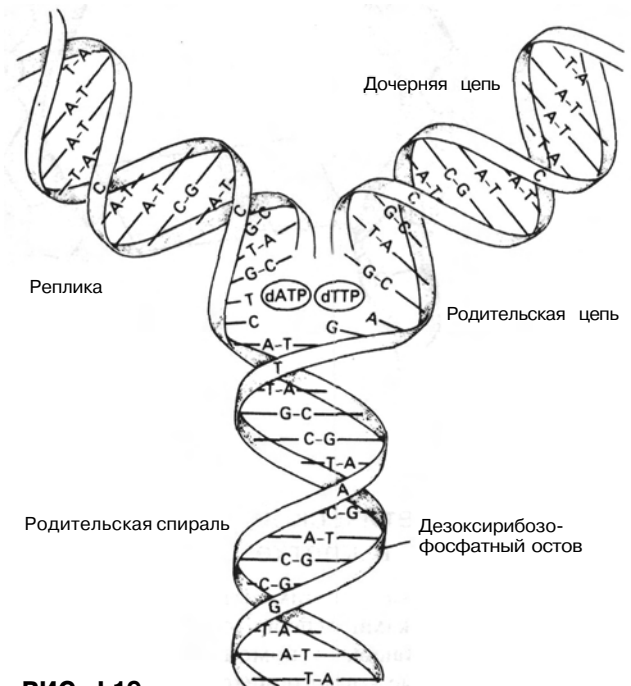
Далее мы рассмотрим вкратце ключевые характеристики генетического аппарата и его функционирования: структурные особенности важнейших компонентов молекул—ДНК, РНК и белков — и то, как они работают, обеспечивая сохранение целостности генома и трансляцию генотипа организма в его фенотип. Эти вопросы детально рассматриваются в гл. 1, 2 и 3, составляющих первую часть книги.

## Структура и сохранение геномной ДНК

Все клеточные ДНК состоят из двух полинуклеотидных цепей, закрученных вокруг общей оси с образованием двойной спирали (рис. I.19). Наружную поверхность спирали составляет остов каждой цепи, состоящий из повторяющихся остатков дезоксирибозы. Цепи удерживаются вместе благодаря водородным связям между пуриновыми основаниями одной цепи и пиримидиновыми — другой: аденин всегда спарен с тиминном, а гуанин —

цитозином. В результате образования таких практически инвариантных пар последовательность оснований одной цепи однозначно определяет их последовательность в другой — иными словами, цепи двойной спирали ДНК **комплементарны**.

Молекулы ДНК выполняют две разные функции. Первая — последовательность пуриновых и пиримидиновых оснований каждой цепи служит матрицей, с которой копируется новая цепь. Вторая — гены, составляющие ДНК, детерминируют синтез ферментов и других белков, необходимых для синтеза новых молекул ДНК. При репликации в особом участке двойной спирали ДНК происходит расплетание цепей. В результате каждая цепь начинает функционировать как матрица, на которой синтезируется новая, комплементарная цепь (рис. I.19). Таким образом, каждая из обеих образовавшихся дочерних спиралей получает одну цепь от родительской спирали, а другую — образованную в результате синтеза *de novo*. Несмотря на кажущуюся логическую простоту, процесс репликации в действительности очень сложен и для его осуществления необходимо множество белков. Важнейшими из них являются ферменты, называемые **ДНК-полимеразами**. Их роль в репликации состоит в сборке полинуклеотидных цепей из отдельных

**РИС. I.19.**

Схематическое изображение реплицирующейся молекулы ДНК. Нижняя часть молекулы еще имеет форму двойной спирали со спаренными основаниями. В верхней части двойная спираль уже расплетена, и на каждой родительской цепи синтезируется новая цепь.

моноклеотидов (дезоксинуклеотидтрифосфатов). Все ДНК-полимеразы удлиняют полинуклеотидную цепь последовательным добавлением отдельных дезоксинуклеотидов.

Выбор нуклеотида, который должен быть присоединен к цепи, определяется способностью входящего в его состав основания образовывать комплементарную пару со следующим свободным основанием цепи-матрицы. Высокая надежность процесса репликации гарантирует практически безошибочную передачу генетической информации в ряду поколений.

Одно из открытий, сделанных при изучении простейших геномов, состояло в том, что они кодируют аппарат для собственного увековечения и сохранения (т.е. репликации и репарации геномной ДНК). Более того, генетическая программа допускает возможность перестроек ДНК, и хотя при этом часто образуются невыгодные, неблагоприятные перестройки, создаваемые новые комбинации генов являются материалом для эволюционного экспериментирования. Все геномы содержат информацию, необходимую для синтеза РНК, ферментов и различных белков, участвующих в этих процессах. Один из таких процессов—генетическая рекомбинация, в результате которой происходит обмен между сегментами гомологичных хромосом. Ранее мы отмечали, что генетические обмены связаны, по-видимому, со спариванием хромосом в мейозе; более того, процесс кроссинговера можно визуализировать. Если рассматривать эти события на молекулярном уровне, то рекомбинация происходит в местах перекреста и состоит в разрыве и воссоединении цепей в пределах соответствующих (го-

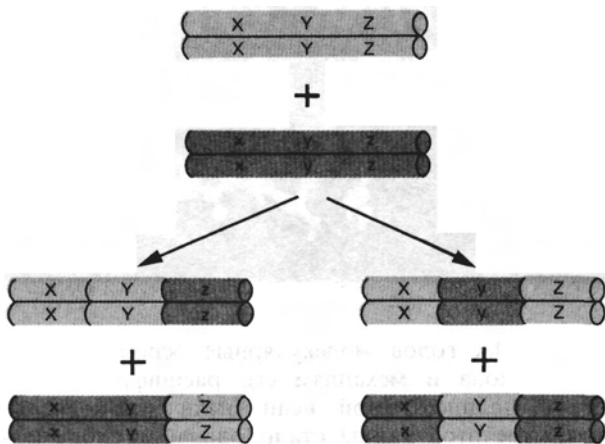
мологичных) областей ДНК рекомбинирующих хромосом (рис. 1.20). Рекомбинация, также генетически детерминированная, может происходить и между определенными участками ДНК негомолгичных хромосом; в результате создаются новые связи между генетическими структурами. Для осуществления различных процессов рекомбинации, обнаруженных у прокариот, требуется целая армия ферментов, обеспечивающих спаривание гомологов или особых последовательностей и катализирующих разрывы и воссоединение цепей.

Существуют также и специальные механизмы репарации повреждений ДНК. Облучение клеток ультрафиолетовым светом или рентгеновскими лучами либо обработка различными химическими агентами приводят к повреждениям, затрагивающим основания или остов молекулы ДНК. В ДНК закодирована информация о синтезе репарирующих ферментов и белков, поддерживающих целостность генома любого организма.

## Экспрессия и регуляция генов

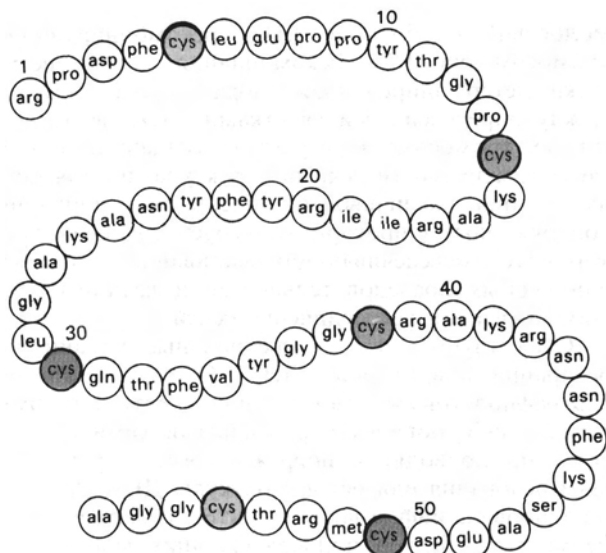
Белки—основные детерминанты фенотипа организма. Из них построены и ферментативный аппарат, обеспечивающий метаболическую, энергетическую и биосинтетическую активность всех клеток, и регуляторные элементы, координирующие эти виды активности в ответ на эндогенные и экзогенные сигналы. Белки являются также основными компонентами многих структурных элементов, характеризующих морфологию клетки и опосредующих ее движение. Говоря в двух словах, организмы—это в конечном счете те белки, которые они сами и производят.

Постулат «один ген—один полипептид» создал концептуальную базу для анализа связи генотипа организма с его фенотипом. Но до решения проблемы структурной организации белков и ДНК, т.е. до начала 50-х годов, эта теория не имела молекулярной основы. С разработкой новых методов анализа белковой структуры было установлено, что каждый белок обладает уникальной линейной аминокислотной последовательностью (рис. 1.21, А). Эта последовательность, называемая **первичной структурой**, определяет характер укладки полипептидной цепи с образованием биологически активной трехмерной формы (рис. 1.21, Б и В). Таким образом, структура белка определяется его аминокислотной последовательностью, которая в свою очередь кодируется генами. Доказательством этому служит тот факт, что мутации в гене приводят к изменению аминокислотной последовательности соответствующего белка. Более того, последовательности мутантных сайтов в генах и последовательности измененных аминокислот в соответ-

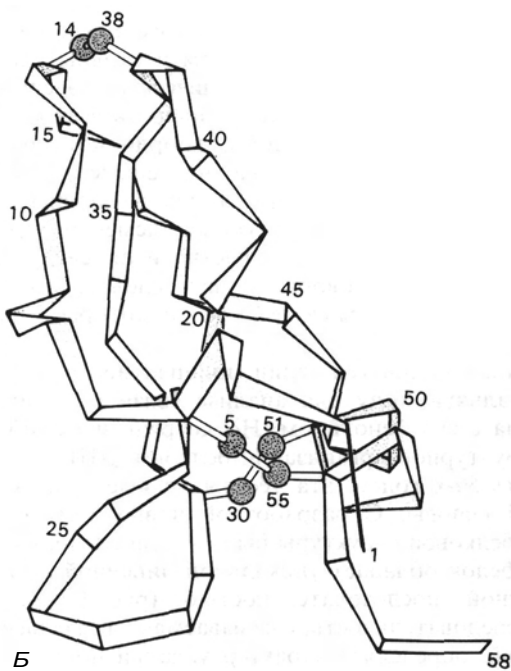


**РИС. 1.20.**

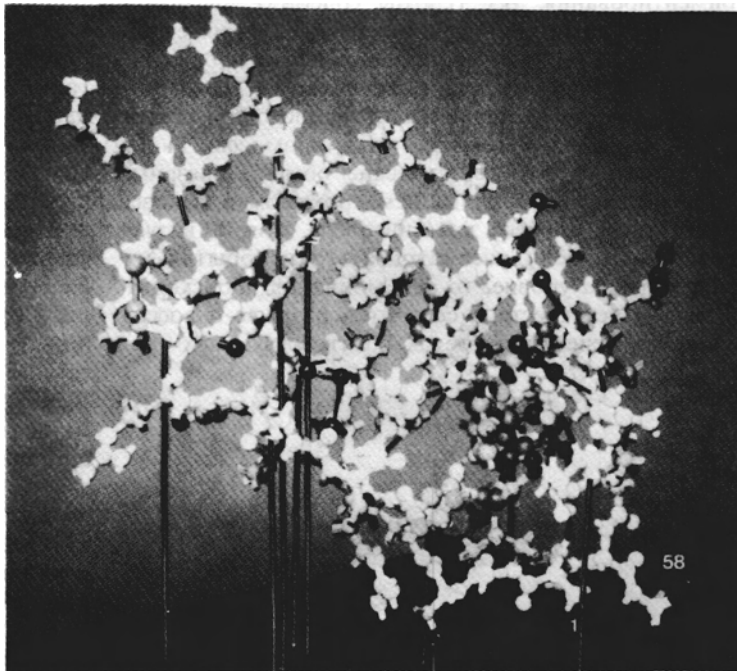
Рекомбинация между двумя гомологичными участками родительских хромосом, несущими разные аллели генов X, Y и Z. Показаны различные возможные продукты рекомбинации.



А



Б



В

РИС. 1.21.

А. Первичная структура белка-ингибитора панкреатического трипсина—из клеток быка. Б. Трехмерная структура белка (ленточная диаграмма). Видны локальные области вторичной структуры (например, спиральные участки с 47-го остатка по 56-й), а также изгибы цепи с образованием третичной структуры, стабилизируемой взаимодействиями между аминокислотами, которые не являются ближайшими соседями вдоль полипептидной цепи. Эти взаимодействия осуществляются между остатками цистеина (цветные шарики) с помощью дисульфидных связей. В. Молекулярная модель белка. Ленточная диаграмма на рис. Б представляет собой вид на эту модель снизу. [С любезного разрешения Т.Е. Creighton, J. Mol. Biol., **95** (1975), p. 167.]

вующих белках коллинеарны, т.е. порядок их следования одинаков. Таким образом, было показано, что линейное расположение нуклеотидов в ДНК и аминокислот в белках взаимосвязано, т.е. одна из характеристик генетического кода установлена.

Идея генетического кода подразумевает существование определенного механизма перевода нуклеотидной последовательности ДНК в аминокислотную последовательность белков. С середины 50-х до

начала 60-х годов молекулярные основы генетического кода и механизм его расшифровки при сборке полипептидной цепи были установлены. Раскрытие этой тайны стало одним из монументальных достижений молекулярной генетики. Неожиданно код оказался очень простым и абсолютно одинаковым для всех жизненных форм. Более того, выяснилось, что универсальны и общие правила трансляции генетически закодированных посланий.

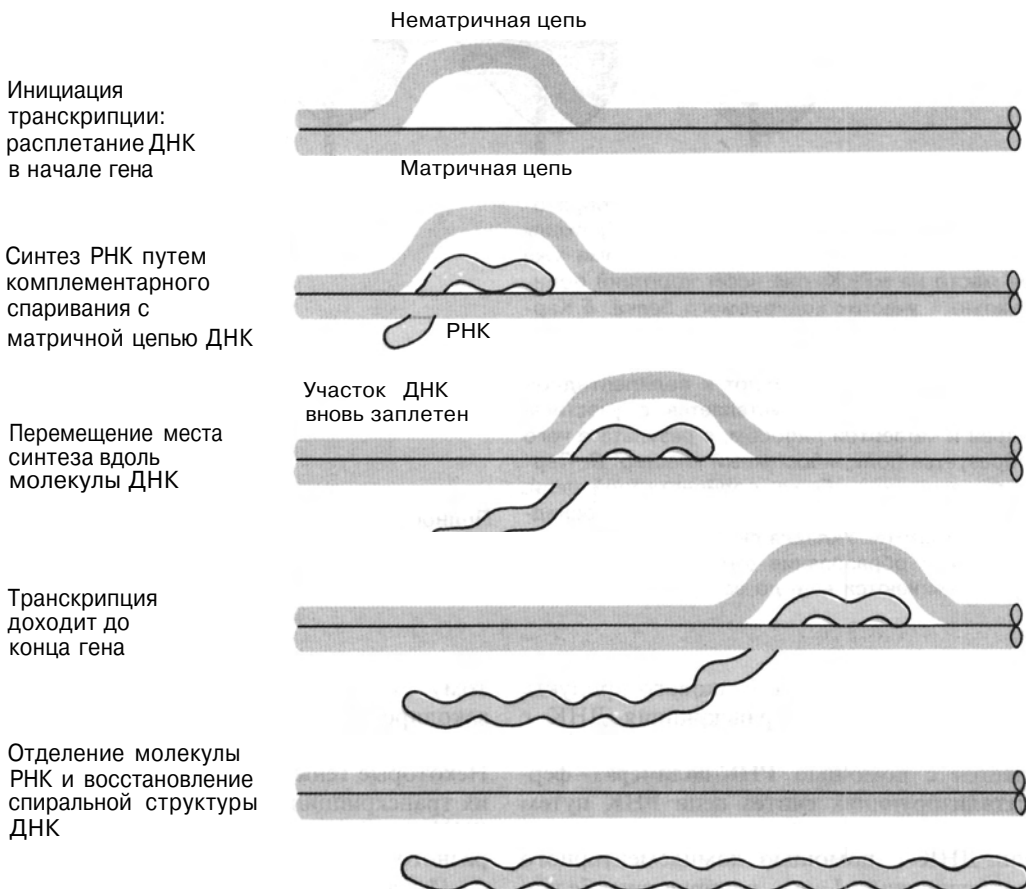
AGA										TTA			AGT								
AGG										TTG			AGC								
GCA CGA						GGA				CTA			CCA TCA ACA								GTA
GCG CGG						GGG		ATA	CTG				CCG TCG, ACG								GTG TAA
GCT CGT GAT AAT TGT GAA CAA						GGT CAT ATT CTT AAA				TTT			CCT TCT ACT								TAT GTT TAG
GCC CGC GAC AAC TGC GAG CAG						GGC CAC ATC CTC AAG				ATG	TTC		CCC TCC ACC	TGG	TAC						GTC TGA
Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ileu	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	stop	

**РИС. I.22.**

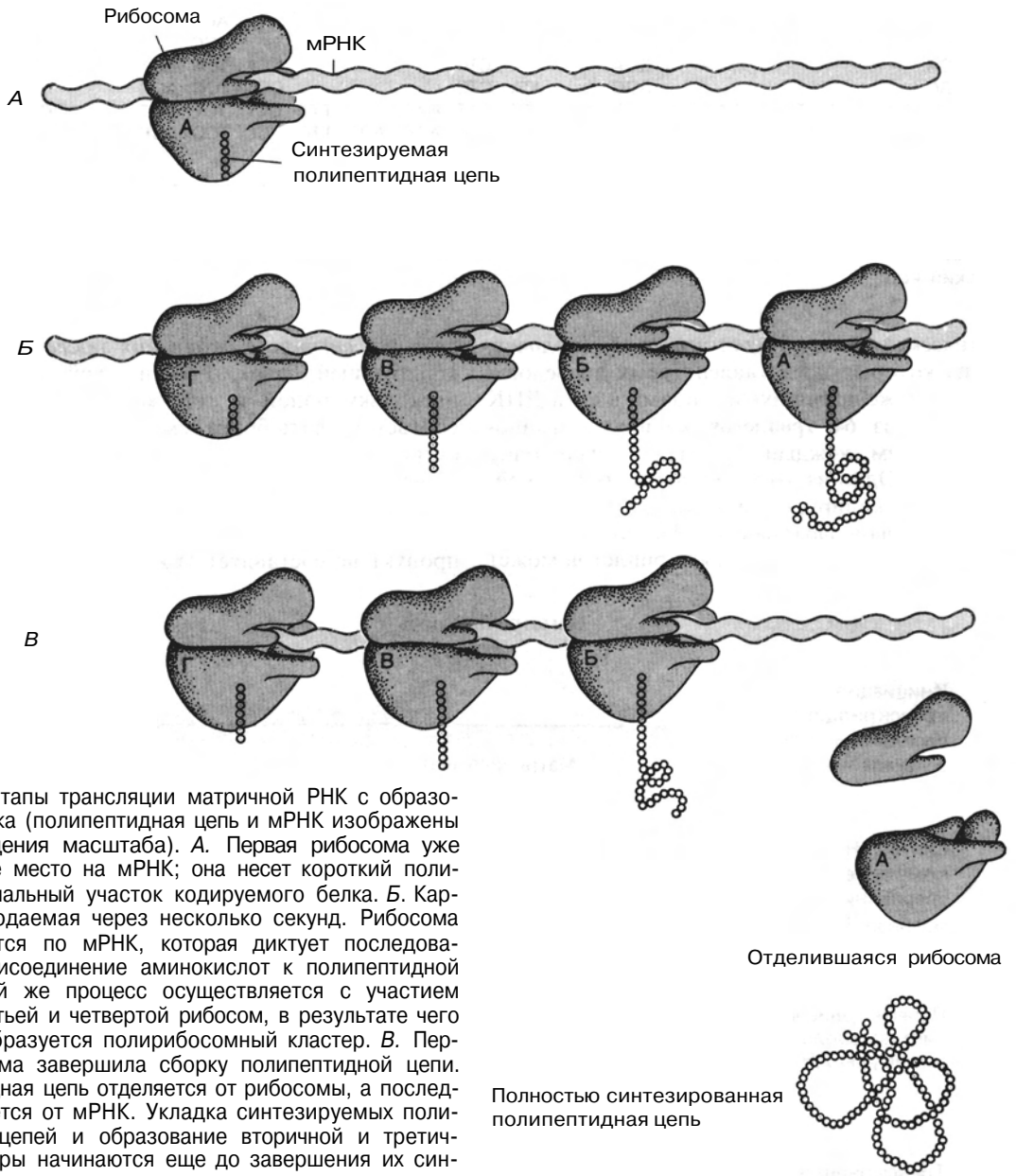
Генетический код.

Генетический словарь состоит из 64 **кодонов**, каждый из которых представлен тремя последовательно расположенными нуклеотидами в цепи ДНК (рис. I.22). 61 из 64 **триплетов** кодируют аминокислоты, причем каждый триплет—только одну аминокислоту. Один из этих триплетов имеет двойную функцию: кодирует аминокислоту метионин и обозначает начало фрагмента ДНК, кодирующего белок. Каждый из трех остальных триплетов может

служить сигналом окончания последовательности, кодирующей белок. Генетический код вырожден, поскольку одной и той же аминокислоте может соответствовать более чем один кодон; но, с другой стороны, код не двусмысленный, потому что любой кодон обозначает только одну аминокислоту. Если известен словарь кодонов, то перевести генную последовательность в соответствующий белковый продукт не составляет труда.

**РИС. I.23.**

Основные этапы транскрипции. Нуклеотидная последовательность РНК такая же, как у нематричной цепи ДНК.



**РИС. 1.24.**

Основные этапы трансляции матричной РНК с образованием белка (полипептидная цепь и мРНК изображены без соблюдения масштаба). *А.* Первая рибосома уже заняла свое место на мРНК; она несет короткий полипептид – начальный участок кодируемого белка. *Б.* Картина, наблюдаемая через несколько секунд. Рибосома перемещается по мРНК, которая диктует последовательное присоединение аминокислот к полипептидной цепи. Такой же процесс осуществляется с участием второй, третьей и четвертой рибосом, в результате чего на мРНК образуется полирибосомный кластер. *В.* Первая рибосома завершила сборку полипептидной цепи. Полипептидная цепь отделяется от рибосомы, а последняя отделяется от мРНК. Укладка синтезируемых полипептидных цепей и образование вторичной и третичной структуры начинаются еще до завершения их синтеза.

Для экспрессии гена в виде белкового продукта сначала должна произойти транскрипция ДНК с образованием РНК (рис. 1.23). Этот процесс осуществляется с помощью **РНК-полимераз** – ферментов, катализирующих синтез цепи РНК путем копирования нуклеотидной последовательности одной цепи ДНК с помощью комплементарного спаривания оснований. Гены, кодирующие белки, детерминируют синтез молекулы «мессенджер», или матричной РНК (**мРНК**), называемой так по-

тому, что она несет генетическую информацию, закодированную в соответствующем сегменте ДНК, и непосредственно участвует в сборке белков. Некоторые гены не кодируют никаких белков. При их транскрипции образуются не мРНК, а молекулы РНК, необходимые для образования зрелых РНК разного типа и для трансляции мРНК в белки.

Исследование взаимодействия РНК-полимераз и других вспомогательных белков транскрипции с ДНК расширило наши знания о специфичности

и прочности межмолекулярных взаимодействий. Так, было показано, что осуществляются очень точные молекулярные контакты между белками и специфичными группами нуклеотидов в ДНК, а это в свою очередь открыло новые перспективы в исследовании проблем экспрессии и регуляции генов. Мы вкратце прокомментируем, как такие взаимодействия опосредуют регуляцию работы генов.

В рамках вводной главы невозможно описать такой совершенный процесс, как трансляция последовательности нуклеотидов матричной РНК в белковую цепь. Он действительно очень сложен и состоит из множества повторяющихся этапов (рис. 1.24). Трансляцию молекул мРНК в белки катализируют рибонуклеопротеиновые частицы (**рибосомы**), содержащие более 50 различных белков и три вида молекул РНК. Синтез белковой цепи начинается с присоединения рибосом к матричной РНК. Белковая цепь удлиняется на одну аминокислоту, когда рибосома продвигается вдоль молекулы мРНК на один кодон. Ключевой момент трансляции—перевод генетической информации, закодированной в триплетных кодонах матричной РНК, в специфические аминокислоты—зависит от комплементарного спаривания оснований. Каждая аминокислота присоединяется к особой, родственной ей транспортной РНК (**тРНК**), содержащей триплет (**антикодон**), комплементарный кодоновому триплету в матричной РНК. Благодаря спариванию оснований между кодоном мРНК и антикодоном тРНК нужная аминокислота занимает свое место в растущей полипептидной цепи. За один цикл перемещения рибосомы по всей длине молекулы мРНК, кодирующей данный белок, образуется одна молекула этого белка.

Изучение экспрессии генов—только один из аспектов исследования механизма их действия. Другой связан с регуляторными процессами, контролирующими время и степень экспрессии при разных условиях. Неудивительно, что прогресс в понимании механизма транскрипции и трансляции позволил прояснить и проблему регуляции. Так, было показано, что у бактерий регуляция экспрессии

генов происходит дифференцированно. Действительно, при некоторых условиях многие гены не экспрессируются вовсе, а степень экспрессии других различается на порядки. Однако изменение условий может приводить к активации молчавших ранее генов и, напротив, к репрессии активных. Это предоставляет клеткам широкие возможности для изменчивости, обеспечивающей приспособленность их фенотипов к условиям среды.

Экспрессия генов обычно регулируется на уровне образования РНК. Как правило, инициация транскрипции регулируется либо **репрессорными** белками, блокирующими транскрипцию, либо **активаторными**, необходимыми для ее запуска. В первом случае экспрессия начинается после снятия репрессии в результате модификации белка-репрессора. Во втором ген транскрибируется только в том случае, если активаторный белок находится в соответствующем функциональном состоянии. Репрессорные и активаторные белки—не единственные средства регуляции транскрипции. В некоторых случаях белки—продукты генной экспрессии—сами служат регуляторами транскрипции собственных генов. Известны также случаи, когда на эффективность транскрипции влияют структурные изменения в ДНК. Образование РНК может регулироваться и путем контроля скорости элонгации или места ее окончания, т.е. транскрибироваться может весь ген или какая-то его часть при наличии специфического стоп-сигнала. Экспрессия генов может также регулироваться на уровне трансляции матричной РНК в белки. В этом случае специфическая регуляция тоже обычно осуществляется на начальных этапах процесса декодирования.

В этом вводном разделе мы кратко описали зарождение генетики и ее эволюцию, в ходе которой были созданы молекулярные основы этой науки. Материал, позволяющий понять эти основы, представлен главным образом в гл. 1—3, составляющих часть I книги. Данные, которые были получены при использовании микроорганизмов (бактерий и их вирусов) в качестве объектов исследования, создали фундамент для изучения более сложных геномов эукариот.

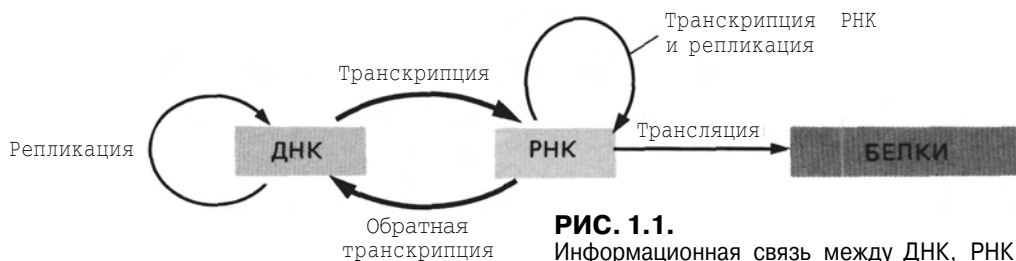
# МОЛЕКУЛЫ ГЕНЕТИЧЕСКОГО АППАРАТА

Генетическая информация у всех клеток закодирована в виде последовательности нуклеотидов в дезоксирибонуклеиновой кислоте (ДНК). Первый этап реализации этой информации состоит в образовании родственной ДНК молекулы—рибонуклеиновой кислоты (РНК), которая в свою очередь участвует в синтезе специфических белков. Фенотипические признаки любого организма в конечном счете проявляются в разнообразии и количестве белков, кодируемых ДНК. Информационная связь между молекулами генетического аппарата — ДНК, РНК и белками—представлена на рис. 1.1.

Чтобы генетическая информация могла передаваться от одного поколения клеток к другому, должна происходить **репликация ДНК**—процесс, в ходе которого родительские молекулы ДНК удваиваются и затем распределяются между потомками. Этот процесс должен осуществляться с большой точностью, а повреждения или случайные ошибки, возникшие в ДНК во время циклов репликации или между ними, необходимо исправить прежде, чем они попадут в геномы потомков. Кроме того, для формирования фенотипа генетическая информация должна экспрессироваться. У всех клеточных организмов экспрессия генов включает копирование ДНК с образованием РНК (**транскрипцию**) и последующую **трансляцию** РНК в белки. Как будет показано в гл. 3, при транскрипции образуется несколько типов РНК. Одни из них, **матричные РНК (мРНК)**, кодируют белки, другие участвуют в различных процессах, необходимых для сборки полноценного белка. ДНК не только кодирует ферментативный аппарат клетки; она участвует в процессах репарации, а при определенных условиях в ней могут происходить перестройки. Репликация, репарация и перестройки ДНК—ключевые процессы, с помощью которых организмы поддерживают свойственный им фенотип и изменяют его.

У многих вирусов генетическая информация также закодирована в ДНК. Механизмы репликации, репарации, перестройки и экспрессии вирусной ДНК аналогичны механизмам, используемым клетками других организмов. Геном некоторых вирусов представлен не ДНК, а РНК. Геномная РНК таких вирусов либо непосредственно транслируется в белки, либо обладает генетической информацией, необходимой для синтеза молекул РНК, которые в свою очередь транслируются в белки. Те вирусы, у которых геном представлен РНК в течение всего жизненного цикла, должны сами реплицировать родительскую РНК для получения потомства вирусных частиц. Существует класс **ретровирусов**, репродуктивный цикл которых начинается с того, что их генетическая информация в ходе так называемой **обратной транскрипции** переводится на язык ДНК. Полученные копии ДНК, или провирусы, способны к репликации и экспрессии только после интеграции в хромосомную ДНК клетки. В такой интегрированной форме вирусные геномы реплицируются вместе с ДНК клетки-хозяина, и для образования нового поколения вирусных геномов и мРНК, нужной для синтеза вирусных белков, они используют транскрипционный аппарат клетки.

Ключевым моментом в передаче генетической информации между нуклеиновыми кислотами, будь то репликация, транскрипция или обратная транскрипция, является то, что молекула нуклеиновой кислоты используется в качестве матрицы в направленной сборке идентичных или родственных структур. Насколько известно, информация, хранящаяся в белках, не используется для сборки соответствующих нуклеиновых кислот, т.е. обратная трансляция не обнаружена. Тем не менее белки играют ключевую роль в процессах передачи информации как между нуклеиновыми кислотами, так и от нуклеиновых кислот к белкам.



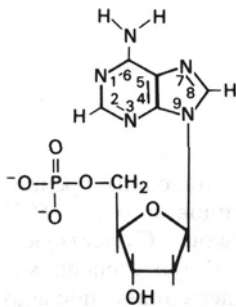
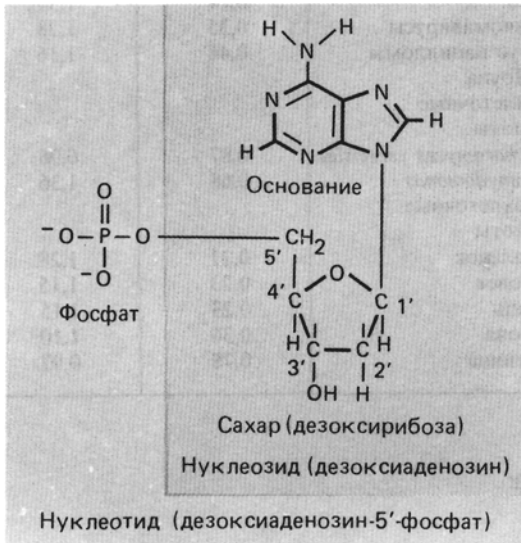
**РИС. 1.1.** Информационная связь между ДНК, РНК и белками.

В этой главе мы рассмотрим химические и физические характеристики ДНК, РНК и белков, поскольку секреты генотипических и фенотипических функций этих молекул скрыты в их молекулярной структуре и свойствах. Механизмы репликации, репарации и рекомбинации нуклеиновых кислот подробно обсуждаются в гл. 2, а механизмы транскрипции и трансляции в процессе экспрессии генетической информации — в гл. 3.

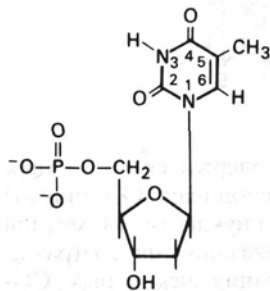
## 1.1. СТРУКТУРА И ПОВЕДЕНИЕ ДНК

### а. Компоненты молекулы ДНК и соединяющие их химические связи

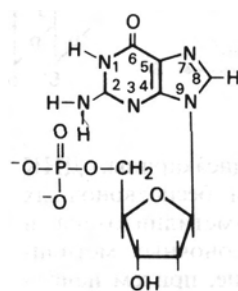
С помощью химических и физических методов установлено, что ДНК — это полимер, состоящий из четырех разных, но родственных мономеров. Каждый мономер — **нуклеотид** — содержит одно из четырех гетероциклических азотистых оснований: аденин (А), гуанин (G), цитозин (С) или тимин (Т), связанное с дезоксирибозофосфатом (рис. 1.2). Длин-



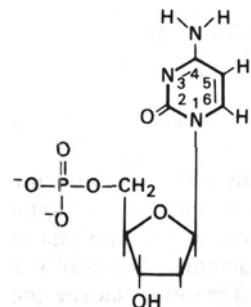
Дезоксиаденозин-монофосфат, dAMP



Дезокситимидин-монофосфат, dTMP



Дезоксигуанозин-монофосфат, dGMP



Дезоксицитидин-монофосфат, dCMP

ные полинуклеотидные цепи образуются путем соединения дезоксирибозных остатков соседних нуклеотидов с помощью фосфодиэфирных связей (рис. 1.3). Каждый фосфат соединяет гидроксильную группу (ОН) при 3'-углеродном атоме дезоксирибозы одного нуклеотида с ОН-группой при 5'-углеродном атоме дезоксирибозы соседнего нуклеотида.

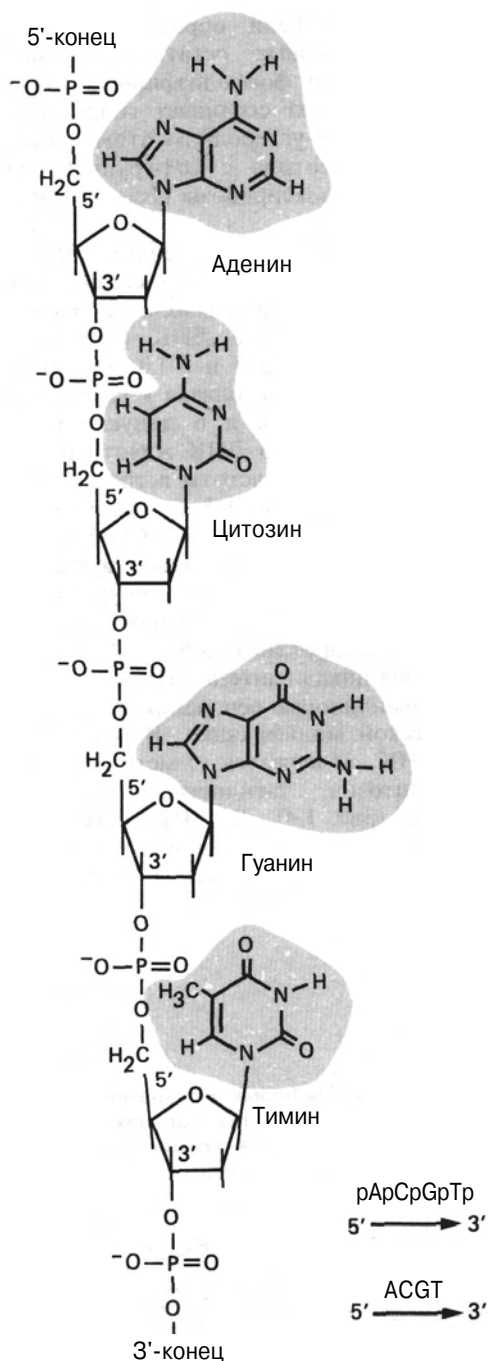
Частота встречаемости в определенном соседстве любых двух оснований в ДНК бактерий, бактериофагов и дрожжей зависит от количественного содержания этих оснований в ДНК (табл. 1.1). Частота встречаемости 5'-CG-3' и 5'-GC-3' в ДНК прокариот почти одинакова и близка к случайной; то же самое можно сказать и о динуклеотидах 5'-GA-3' и 5'-AG-3'. Однако в ДНК животных, вирусов животных и растений частоты встречаемости 5'-CG-3' составляют от 1/2 до 1/5 частот 5'-GC-3'. Таким образом, последовательность 5'-CG-3' встречается в ДНК высших эукариот довольно редко; это связано со способностью данного динуклеотида служить мишенью при метилировании и с его ролью в регуляции экспрессии генов.

После окончания цикла синтеза ДНК некоторые пуриновые и пиримидиновые основания могут подвергаться химической модификации. В результате в некоторых ДНК содержатся 5-метилцитозин, 5-гидроксиметилцитозин, 5-гидроксиметилурацил и N-метиладенин (рис. 1.4). В ДНК некоторых бактериофагов к гидроксиметильной группе гидроксиметилцитозина присоединены с помощью

### РИС. 1.2.

Дезоксирибонуклеотиды. Цифрами обозначены положения атомов в гетероциклических кольцах пуринов (аденин и гуанин) и пиримидинов (тимин и цитозин) и углеродных атомов в дезоксирибозе.

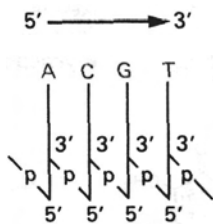


**РИС. 1.3.**

Связь между соседними дезоксирибонуклеотидами в полинуклеотидной цепи. В правом нижнем углу рисунка показаны некоторые способы схематического изображения дезоксирибонуклеотидной последовательности. Нуклеотидную последовательность принято изображать слева направо, от 5'- к 3'-концу.

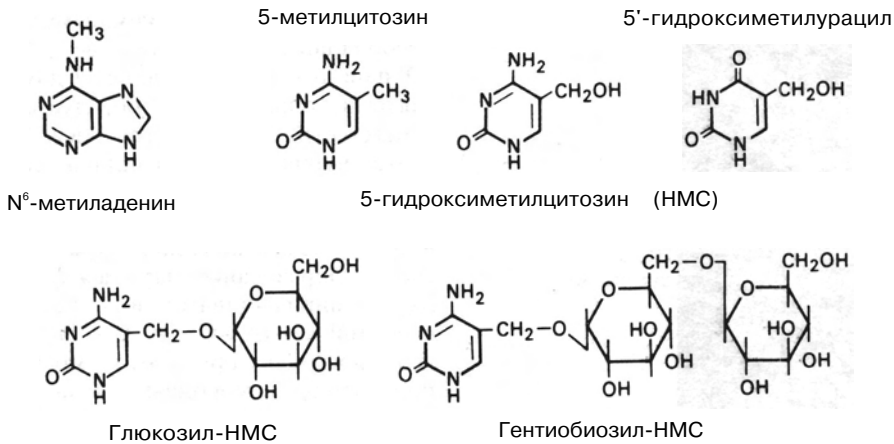
Таблица 1.1. Частоты встречаемости некоторых ближайших соседей в различных ДНК

Источник ДНК	Отношение	
	5'-CG-3'/5'-GC-3'	5'-AG-3'/5'-GA-3'
Вирусы бактерий		
λ	1,02	0,90
T2	0,82	0,96
Бактерии		
<i>Escherichia coli</i>	0,95	1,0
<i>Micrococcus lysodeikticus</i>	1,15	0,75
<i>Bacillus subtilis</i>	0,82	0,87
Вирусы животных		
SV40	0,14	1,35
Полиомавирусы	0,35	1,28
Вирус папилломы	0,44	1,16
Шоупа		
Одноклеточные эукариоты		
<i>Saccharomyces cerevisiae</i>	0,87	0,96
<i>Chlamydomonas</i>	0,68	1,36
Многоклеточные эукариоты		
Цыпленок	0,21	1,28
Человек	0,23	1,15
Мышь	0,23	1,15
Корова	0,36	1,10
Пшеница	0,78	0,97



гликозидной связи моно- или дисахариды. ДНК большинства низших эукариот и беспозвоночных содержат относительно мало 5-метилцитозина и N<sup>6</sup>-метиладенина. Однако у позвоночных метилирование оснований — частое явление, причем наиболее распространен 5-метилцитозин. Показано, что более 95% метильных групп в ДНК позвоночных

содержится в остатках цитозина редко встречающихся CG-динуклеотидов и более 50% таких динуклеотидов метилировано. Существуют четкие указания на то (разд. 8.7), что степень метилирования некоторых CG-содержащих последовательностей является важным фактором регуляции экспрессии определенных генов. У растений 5-ме-

**РИС. 1.4.**

Структурные формулы модифицированных пуринов и пиримидинов, обнаруженных в ДНК.

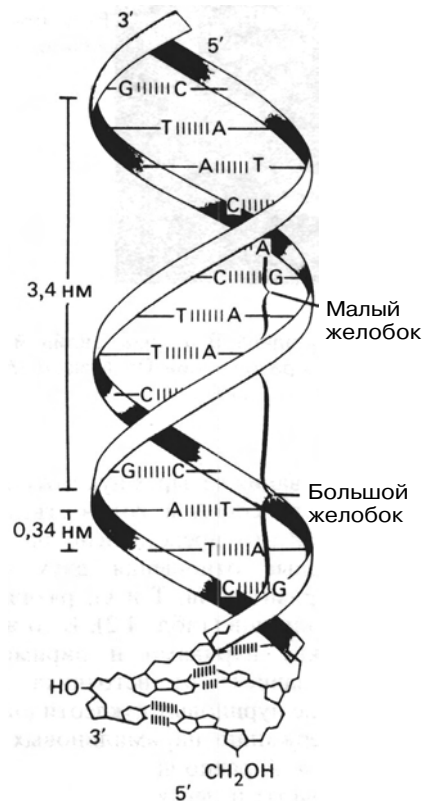
тилцитозин можно обнаружить в динуклеотидах CG и тринуклеотидах CNG (N–C, A или T).

## б. Спиральная структура ДНК

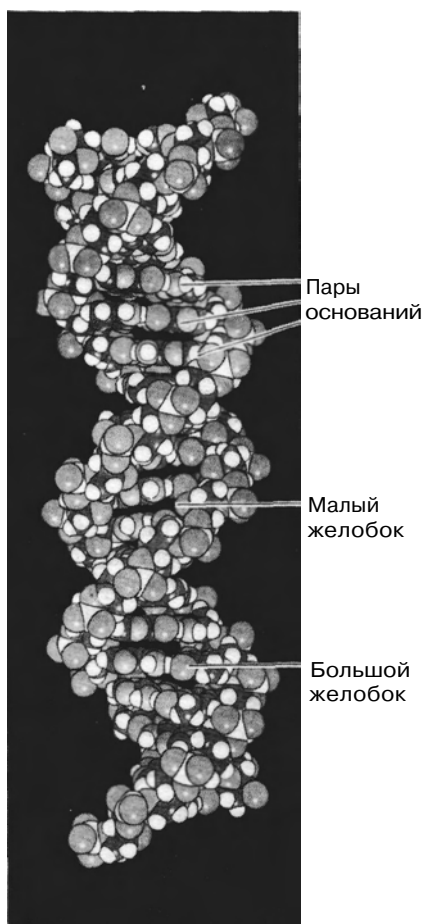
С помощью физико-химических, электронно-микроскопических и рентгеноструктурных методов показано, что большинство молекул ДНК представляют собой протяженные, гибкие, нитевидные структуры. Этими же методами установлено, что молекула ДНК имеет почти постоянный диаметр и состоит из регулярно расположенных повторяющихся звеньев, причем ее структура не зависит от нуклеотидного состава. Таким образом, в отличие от белков, двух- и трехмерная структура которых обязательно зависит от состава и порядка расположения аминокислот (разд. 1.3), молекула ДНК в обычных условиях при любом нуклеотидном составе и порядке расположения четырех нуклеотидов представляет собой абсолютно регулярную практически идентичную по всей длине структуру. Такие в какой-то степени парадоксальные химические и физические свойства ДНК порождаются особенностями ее структуры.

Молекула ДНК обычно находится в форме двойной спирали, образуемой двумя полинуклеотидными цепями, обвивающимися одна вокруг другой. Два дезоксирибозофосфатных остова, расположенные по периферии молекулы, имеют антипараллельную ориентацию (рис. 1.5 и 1.6). В наиболее часто встречающейся структурной форме пуриновые и пиримидиновые основания в каждой цепи уложены в стопки с интервалом 0,34 нм и направлены внутрь спирали; плоскости колец при-

мерно перпендикулярны оптической оси спирали. Спираль делает полный оборот каждые 3,4 нм, т.е. через каждые 10 оснований. На наружной ее поверхности имеются два желобка—большой и малый.

**РИС. 1.5.**

Схематическое изображение В-формы двойной спирали ДНК. Видны большой и малый желобки. Указаны расстояние между ближайшими парами оснований и шаг спирали. [A. Kornberg, DNA Replication (San Francisco: W. H. Freeman, 1980).]



**РИС. 1.6.** Пространственная модель В-формы двойной спирали ДНК. (С любезного разрешения Dr. Melson Max, Lawrence Livermore National Labs.)

Азотистые основания четырех нуклеотидов ДНК не находятся между собой в количественном соотношении 1:1, как это представлено на рис. 1.3. Напротив, молярные отношения двух пуринов, А и G, и двух пиримидинов, Т и С, различны для ДНК разных организмов (табл. 1.2). В то же время соотношение между пуринами и пиримидинами постоянно и не зависит от источника ДНК, а именно: содержание пуриновых нуклеотидов (А + G) всегда равно содержанию пиримидиновых нуклеотидов (Т + С); число А равно числу Т, и аналогично для G и С. Эти факты и легли в основу предположения, что пуриновые и пиримидиновые нуклеотиды в ДНК спарены, а двойная спираль стабилизируется с помощью водородных связей между пуринами одной цепи ДНК и пиримидинами другой (рис. 1.5 и 1.6).

Два указанных типа пар оснований, АТ и GC, обычно называемых **комплементарными парами**, преобладают в большинстве ДНК (рис. 1.7). В АТ-паре основания соединены двумя водородными связями: одна из них образуется между амини- и кето-группами, а другая—между двумя атомами азота пурина и пиримидина соответственно. В GC-паре имеются три водородные связи: две из них образуются между амини- и кето-группами соответствующих оснований, а третья—между атомами азота. Образование пар между двумя пуринами, двумя пиримидинами или некомплементарными основаниями (А + С или G + Т) стерически затруднено, поскольку при этом не могут образовываться подходящие водородные связи и, следовательно, нарушается геометрия спирали. Модифицированные пурины и пиримидины, с небольшой частотой встречающиеся в ДНК (рис. 1.4), образуют такие же водородные связи, что и их немодифицированные аналоги; тем самым правило спаривания не нарушается. Согласно этим правилам, последовательность оснований в одной цепи определяет их последовательность в другой. Комплементарность последовательности оснований в двух полинуклеотидных цепях—ключевое свойство ДНК.

Дополнительная стабилизация двойной спирали обеспечивается межплоскостными взаимодействиями ароматических колец соседних оснований. Размеры комплементарных пар оснований практически одинаковы; примерно одинаковы также угол и направление связи дезоксирибоза-основание. Расстояние между соседними основаниями равно 0,34 нм, а угол, на который они повернуты друг относительно друга,— 36°. Из всех этих данных следует, что диаметр спирали постоянен, а число пар оснований на виток спирали равно 10 (рис. 1.5). Точные данные о расположении, ориентации в пространстве и размерах различных составляющих ДНК были получены методом рентгеноструктурного анализа волокон ДНК.

## в. Альтернативные формы двойной спирали ДНК

Все, о чем мы говорили, касалось наиболее распространенной, так называемой В-формы двойной спирали ДНК. Известны также два других изомерных типа двойной спирали. Они образуются благодаря тому, что валентные углы между основаниями и сахаром могут меняться, а дезоксирибозное кольцо и сахарофосфатный остов достаточно гибки, чтобы могли сформироваться альтернативные конфигурации. Редко встречающаяся А-форма, существующая только при пониженной влажности, отличается от В-формы тем, что плоскости оснований составляют с перпендикуляром

Таблица 1.2. Нуклеотидный состав различных ДНК

Источник	A <sup>1)</sup>	G	C	T	$\frac{A+T}{G+C}$	$\frac{A+G}{T+C}$	G + C, мол. %
Бактериофаг $\lambda$	26,0	23,8	24,3	25,8	1,08	0,99	48
Бактериофаг T2	32,5	18,2	16,7 <sup>2)</sup>	32,6	1,86 <sup>2)</sup>	1,03 <sup>2)</sup>	35 <sup>2)</sup>
<i>Escherichia coli</i>	23,8	26,0	26,4	23,8	0,91	0,99	52
<i>Bacillus subtilis</i>	29,0	20,7	21,3	29,0	1,38	0,99	42
Вирус папилломы Шоупа	26,6	24,5	24,2	24,7	1,05	1,04	49
<i>Saccharomyces cerevisiae</i>	31,3	18,7	17,1	32,9	1,79	1,00	36
<i>Chlamydomonas</i>	19,6	30,2	30,0 <sup>3)</sup>	19,7	0,65 <sup>3)</sup>	0,99 <sup>3)</sup>	60 <sup>3)</sup>
Цыпленок	27,9	21,2	21,5 <sup>3)</sup>	29,4	1,34 <sup>3)</sup>	0,96 <sup>3)</sup>	43 <sup>3)</sup>
Мышь	28,9	21,1	20,3 <sup>3)</sup>	30,0	1,44 <sup>3)</sup>	1,00 <sup>3)</sup>	41 <sup>3)</sup>
Корова	27,3	22,5	22,5 <sup>3)</sup>	27,7	1,22 <sup>3)</sup>	0,99 <sup>3)</sup>	44 <sup>3)</sup>
Пшеница	27,2	22,6	22,8 <sup>3)</sup>	27,4	1,20 <sup>3)</sup>	0,99 <sup>3)</sup>	45 <sup>3)</sup>

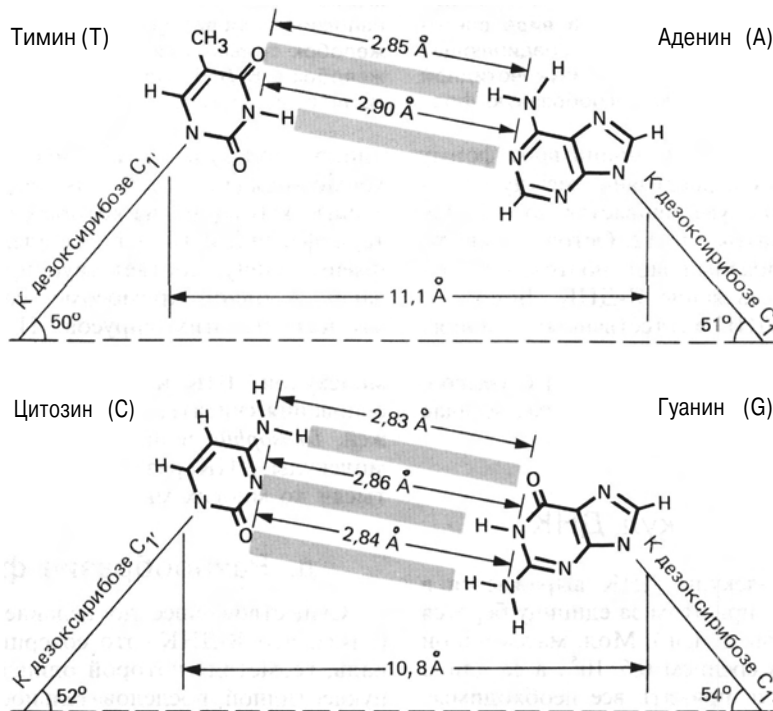
<sup>1)</sup> Используются общепринятые сокращенные обозначения пуринов и пиримидинов.

<sup>2)</sup> 5-гидроксиметилцитозин.

<sup>3)</sup> Включая 5-метилцитозин.

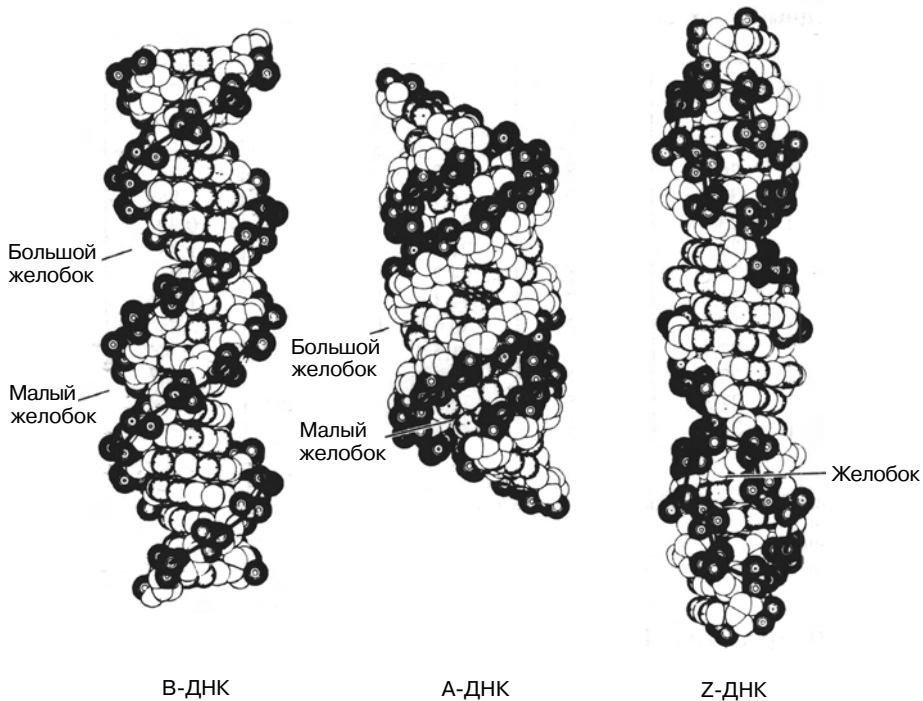
к оси спирали угол 20° (рис. 1.8). Поэтому расстояние между парами оснований по вертикали уменьшается до 0,29 нм, а число пар на виток увеличивается до 11–12. Какова биологическая функция А-формы ДНК—пока неясно.

Характерной особенностью В-формы ДНК является то, что сахарофосфатные остовы обеих цепей образуют правую спираль (рис. 1.5 и 1.6). Однако при определенных условиях участки ДНК, для которых характерно чередование пуриновых

**РИС. 1.7.**

Водородные связи между комплементарными основаниями в ДНК. Видно, что геометрия двух типов пар

оснований практически одинакова. [A. Kornberg, DNA Replication (San Francisco: W. H. Freeman, 1980.)]

**РИС. 1.8.**

Пространственные модели В-, А- и Z-ДНК (каждая из моделей содержит 20 пар оснований). Атомы фосфора и связанные с ними атомы кислорода изображены в виде темных шариков, атомы азота – в виде слегка затененных шариков. Сплошная линия, соединяющая фосфатные группы, показывает ход полинуклеотидной спирали. Обратите внимание на зигзагообразную фор-

му остова Z-ДНК. А-ДНК короче и толще, а Z-ДНК чуть длиннее и тоньше, чем В-ДНК. В А-ДНК большой желобок более глубокий, но зато малый уплощен и напоминает обвитую по поверхности ленту. У Z-ДНК желобок только один, он чуть глубже, чем большой желобок у В-ДНК, но не такой глубокий, как у А-ДНК. (С любезного разрешения А. Rich.)

и пиримидиновых нуклеотидов, принимают форму левой спирали. При этом расстояние между соседними парами оснований увеличивается до 0,77 нм, а число пар на один виток – до 12. Остов молекулы ДНК имеет зигзагообразный вид, поэтому подобная форма получила название Z-ДНК. Вопрос о том, существует ли Z-ДНК в естественных условиях и образуется ли она в определенных участках В-спирали под действием специфических белков, способных переводить В-форму в Z-форму, сейчас интенсивно исследуется (гл. 8, разд. 8.7).

длиной молекулы ДНК и массой одной небольшой хромосомы (табл. 1.3). Так, молекулы ДНК единственных хромосом, из которых состоят геномы бактериофагов  $\lambda$  и T4, а также адено- и герпесвирусов, имеют длину, соответствующую числу пар оснований в одной хромосоме, составляющей геном каждого из этих вирусов. Полный геном *E.coli* ( $\sim 4 \cdot 10^6$  п.н.) также представлен единственной молекулой ДНК и имеет длину 1,4 мм. Есть основания считать, что каждая из хромосом дрожжей, *Drosophila* и даже человека состоит из одной молекулы ДНК размером от нескольких десятков тысяч до многих миллионов пар нуклеотидов.

### г. Размер молекул ДНК

Обычно размер молекулы ДНК выражается в числе пар нуклеотидов, при этом за единицу берется тысяча пар нуклеотидов (т.п.н.). Мол. масса одной т.п.н. В-ДНК равна в среднем  $6,6 \cdot 10^5$ , а ее длина составляет 340 нм. Если принять все необходимые меры, чтобы не разрушить ДНК при выделении, и использовать мягкие методы измерения длины, то обнаружится удивительное соответствие между

### д. Разнообразие форм ДНК

Существовавшее до недавнего времени мнение о том, что В-ДНК – это совершенная двойная спираль, геометрия которой одинакова независимо от нуклеотидной последовательности, в действительности не совсем корректно. Детальный рентгеноструктурный анализ, построение моделей и термодинамические расчеты показали, что плоскости со-

Таблица 1.3. Молекулярная масса, длина и тип структуры ДНК различного происхождения

Источник	Мол. масса	Длина	Число пар оснований	Тип структуры
Бактериофаг $\phi$ X174	$1,6 \cdot 10^6$	1,6 мкм	$5 \cdot 10^{31}$	Кольцевая одноцепочечная
SV40	$3,5 \cdot 10^6$	1,1 мкм	$5,2 \cdot 10^3$	Кольцевая двухцепочечная
Бактериофаг T2	$1,2 \cdot 10^8$	50 мкм	$2 \cdot 10^5$	Линейная двухцепочечная
Хромосома <i>Hemophilus influenzae</i>	$7,9 \cdot 10^8$	300 мкм	$1,2 \cdot 10^6$	Неизвестен
Хромосома <i>Escherichia coli</i>	$2,6 \cdot 10^9$	1 мм	$4 \cdot 10^6$	Кольцевая двухцепочечная
<i>Saccharomyces cerevisiae</i>				
Хромосома 1	$1,4 \cdot 10^8$	50 мкм	$2,1 \cdot 10^5$	Линейная двухцепочечная
Хромосома 12	$1,5 \cdot 10^9$	500 мкм	$2,2 \cdot 10^6$	То же
<i>Drosophila melanogaster</i>				
Хромосома 2	$4 \cdot 10^{10}$	15 мм	$6,0 \cdot 10^7$	» »
Хромосома 3	$4,2 \cdot 10^{10}$	16 мм	$6,3 \cdot 10^7$	» »
Хромосома 4	$4 \cdot 10^9$	1,5 мм	$6 \cdot 10^6$	» »

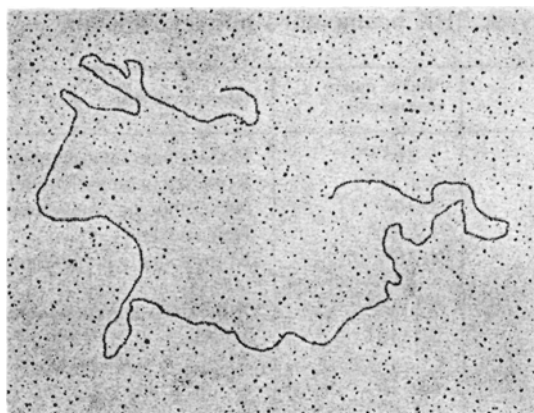
<sup>1)</sup> В данном случае указано число оснований, а не пар оснований.

седних пар оснований не строго параллельны. Каждая комплементарная пара оснований является как бы клином, отклоняющим ось спирали в одном или в другом направлении. Наибольший «крен» наблюдается тогда, когда два соседних аденина в одной цепи спарены с двумя тиминами другой. В этом месте происходит локальное искривление спирали. Если такие пары встречаются с периодичностью примерно один раз на 10 пар (т.е. один раз на каждый виток спирали), то молекула ДНК приобретает заметное искривленную форму. Искривленная, или изогнутая, структура была, например, обнаружена в линейных фрагментах ДНК кинетопластов трипаносомы *Leishmania tarentolae* по аномально малой подвижности этих фрагментов при электрофорезе в полиакриламидном геле. Изгибы в молекуле ДНК наблюдаются в тех участках последовательности, где с необычно высокой частотой встречаются повторы  $(A \cdot T)_{5-6}$ , разделенные GC-богатыми участками из четырех-шести нуклеотидов. Биологическая роль искривления ДНК окончательно не установлена. Предрасположенность к такому изгибанию, зависящая от последовательности оснований, может иметь значение при наматывании молекулы ДНК на гистоновые октамеры в хроматине (разд. 1.1.ж). Возможно, изгибание ДНК существенно и при специфическом связывании ДНК с белками в процессе регуляции экспрессии генов.

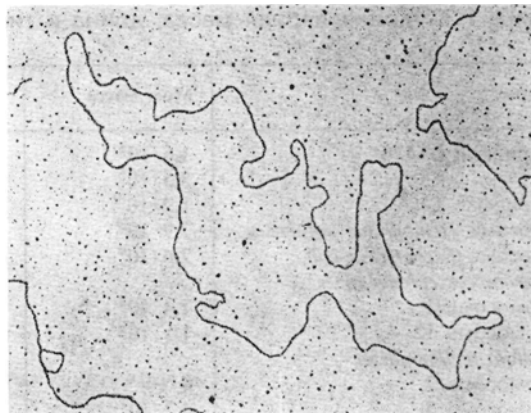
ДНК может находиться в линейной или кольцевой форме (рис. 1.9). Бактериальные плазмиды, хромосомы некоторых бактерий, большинство митохондриальных и хлоропластных ДНК, геномы вирусов млекопитающих представлены единственной ковалентно замкнутой кольцевой дуплексной молекулой ДНК. Хромосома бактериофага  $\lambda$  на разных стадиях жизненного цикла существует то как линейная молекула, то как замкнутая кольцевая

структура, то как кольцо с разрывами. По-видимому, никакого верхнего предела для размера кольцевой двухцепочечной молекулы ДНК не существует.

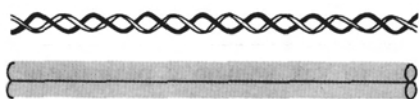
ДНК в клетке обычно находится в комплексе с белками (см. разд. 1.1.ж). Связанный белок слегка раскручивает спираль ДНК, соответственно и число витков спирали на единицу длины становится меньше, чем у свободной В-ДНК. При удалении белка восстанавливается обычное число правозакрученных (положительных) витков спирали. В линейной молекуле ДНК это происходит достаточно легко, поскольку обе цепи свободно вращаются одна вокруг другой. В замкнутой же кольцевой молекуле общее число витков спирали топологически фиксировано, и число оборотов одной цепи вокруг другой не может быть изменено без компенсаторного образования витков противоположного знака где-нибудь в другом месте молекулы. Итак, когда естественные кольцевые дуплексы освобождаются от белков, с которыми они часто бывают связаны *in vivo*, происходит следующее: 1) число правозакрученных (положительных) витков спирали возрастает до величины, характерной для В-ДНК; 2) в самом дуплексе образуется столько же витков противоположного знака, чтобы компенсировать увеличение скрученности спирали. О таких молекулах говорят, что они обладают отрицательной сверхспиральностью (рис. 1.10). При внесении одного разрыва в сверхспиральную кольцевую ДНК сверхспиральность снимается и кольцевая структура переходит в релаксированное состояние, при котором топологические ограничения отсутствуют. Любые химические или физические изменения, приводящие к уменьшению числа витков спирали на молекулу, уменьшают или вообще снимают отрицательную сверхспиральность в замкнутой кольцевой ДНК.



А



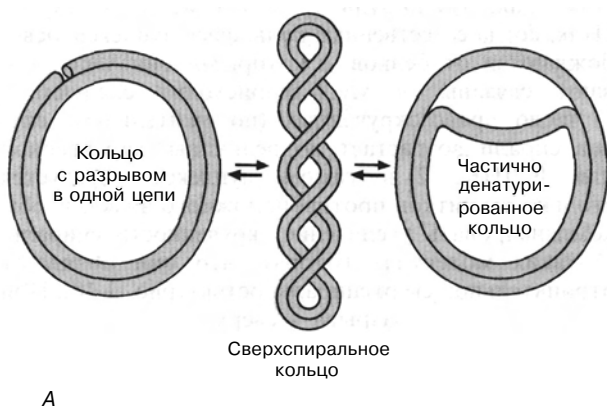
Б

**РИС. 1.9.**

Схематическое представление и электронные микрофотографии линейной (А) и кольцевой (Б) двухцепочечной ДНК фага  $\lambda$ . (Электронные микрофотографии любезно предоставлены L. Chow.) ДНК изображена в виде двойной спирали; такое же изображение для двухцепочечной ДНК будет использоваться повсюду в этой книге.

Не все ДНК *in vivo* являются двухцепочечными. Геномы некоторых мелких вирусов бактерий, растений и животных представляют собой ковалентно замкнутые кольца, состоящие только из одной цепи (рис. 1.11). Все известные одноцепочечные кольцевые ДНК относительно малы: ДНК бактериофагов фХ174 и М13 содержат примерно 5300 и 6000 нук-

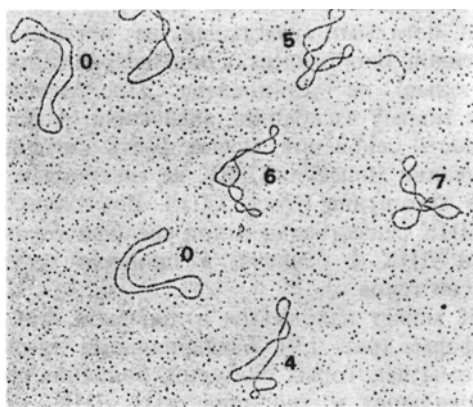
леотидов соответственно и имеют длину 1,5–2 мкм; длина молекул ДНК парвовирусов животных и некоторых вирусов растений составляет 2/3 и 1/2 указанных величин соответственно. Однако для репликации любой из этих вирусных ДНК совершенно необходимо превращение одноцепочечного кольца в соответствующее двухцепочечное, из ко-



А

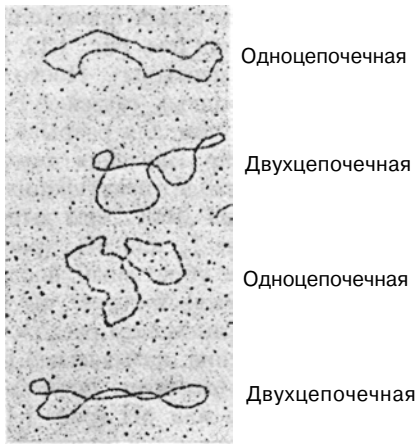
**РИС. 1.10.**

А. Схематическое изображение сверхспиральной кольцевой ДНК и релаксированных кольцевых форм, полученных либо в результате разрыва одной из двух цепей, либо в результате локального расхождения двух цепей.



Б

Б. Двухцепочечная кольцевая ДНК фага М13 с разной степенью сверхспиральности. Цифрами обозначено число сверхвитков в каждой молекуле. (С любезного разрешения L. Chow.)



**РИС. 1.11.**

Электронные микрофотографии одно- и двухцепочечных кольцевых ДНК фага М13. Двухцепочечная ДНК выглядит более гладкой и вытянутой и ее легче визуализировать, чем одноцепочечную. (С любезного разрешения L. Chow.)

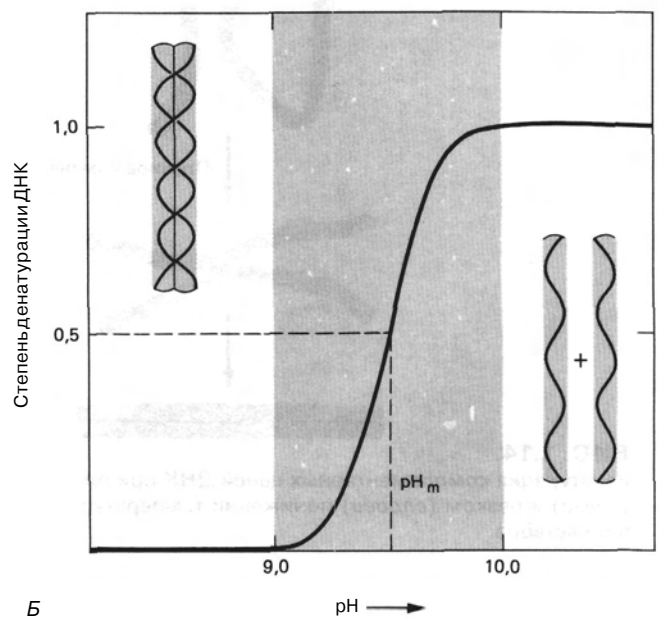
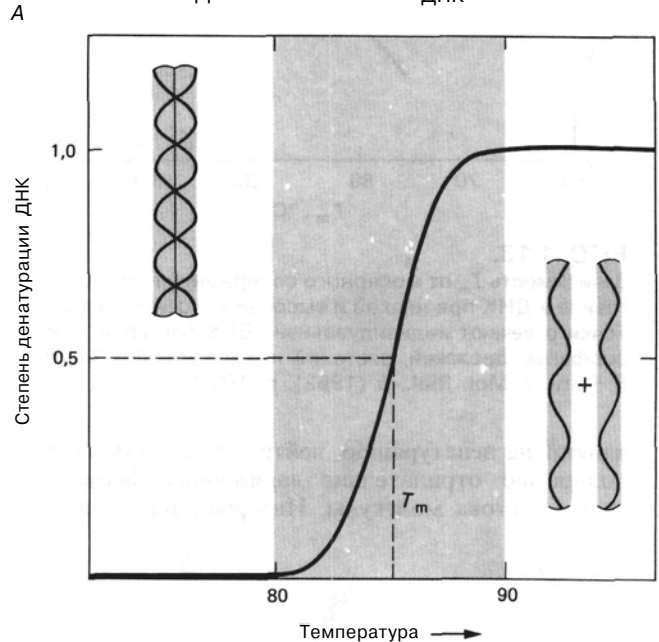
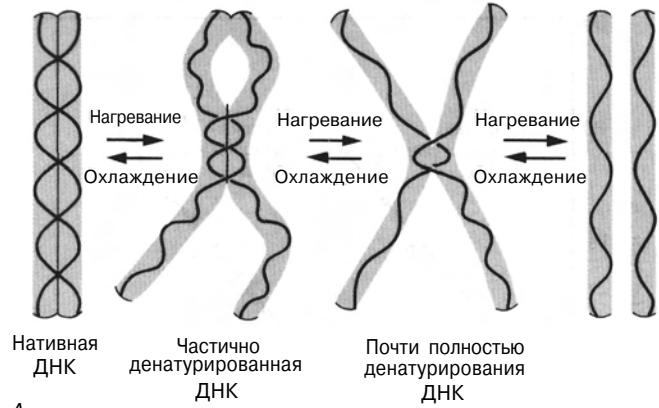
тогого затем образуются одноцепочечные кольцевые ДНК вирусного потомства (гл. 2). Более того, экспрессия генетической информации в таких геномах всегда осуществляется в фазе двухцепочечной ДНК, поскольку именно она является субстратом для транскрипции последовательности ДНК в РНК.

**е. Денатурация и ренатурация ДНК**

Водородные связи и межплоскостные взаимодействия, стабилизирующие двойную спираль, достаточно слабы, и при относительно небольших воздействиях происходит разделение цепей – процесс, именуемый **денатурацией**, или **плавлением** (рис. 1.12, А). Двухцепочечная спиральная ДНК в растворе легко разрушается при нагревании до температур, близких к 100°C. Денатурация происходит также при увеличении рН раствора до уровня, при котором разрушаются водородные связи между основаниями. Многие факторы (например, одно- и двухвалентные катионы, полиамины и белки)

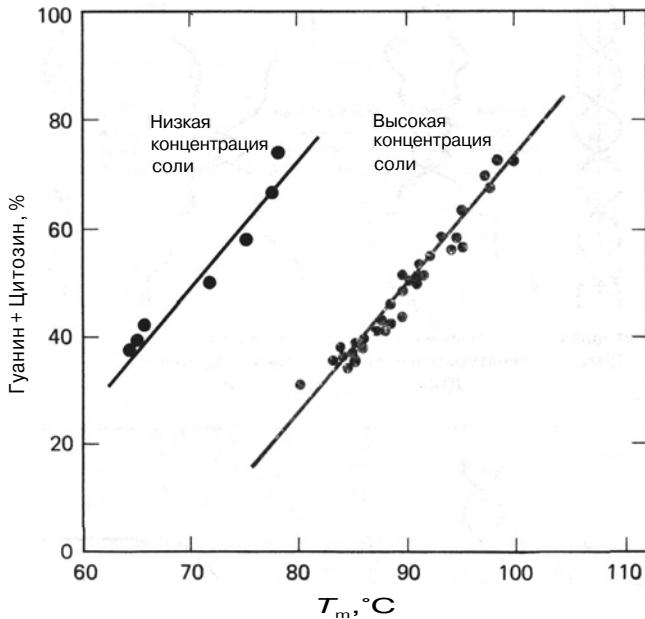
**РИС. 1.12.**

А. Денатурация (диссоциация) двухцепочечной ДНК при повышении температуры раствора и ренатурация (реассоциация) двух комплементарных цепей при охлаждении. Б. Кривые денатурации типичной двухцепочечной ДНК, получаемые при повышении температуры и рН.  $T_m$  и  $pH_m$  - это значения температуры и рН соответственно, при которых ДНК денатурирована (или ренатурирована) наполовину.



Б





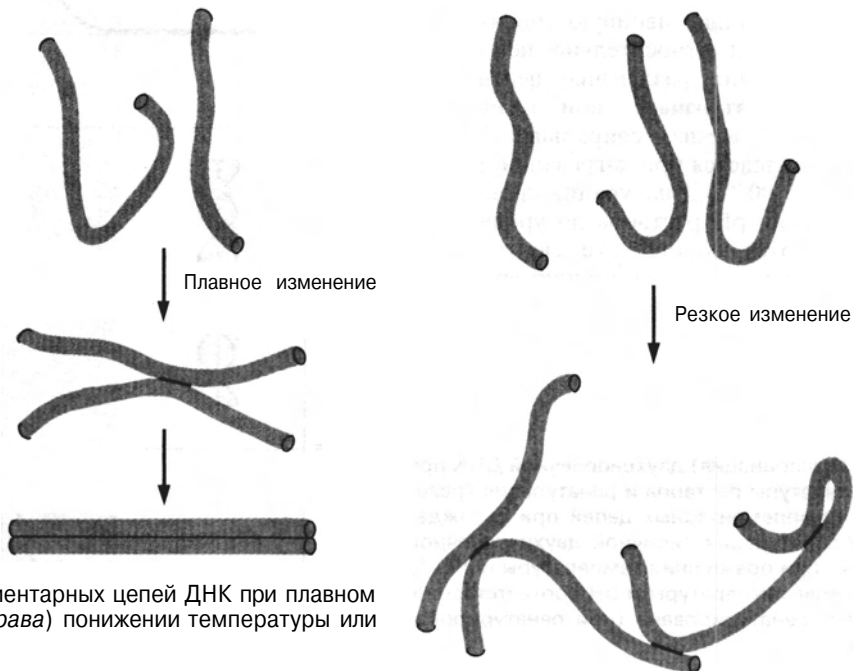
**РИС. 1.13.**

Зависимость  $T_m$  от молярного содержания гуанина и цитозина в ДНК при низкой и высокой концентрации соли. Точки отвечают индивидуальным ДНК бактерий, бактериофагов, дрожжей, растений и животных. [J. Marmur, P. Doty, J. Mol. Biol., 5 (1962), p. 109.]

вливают на денатурацию, нейтрализуя частично или полностью отрицательно заряженные фосфатные группы остова молекулы. Интервал значений тем-

пературы или рН, при которых происходит разделение цепей, очень невелик (см. рис. 1.12, Б, где  $T_m$  и  $pH_m$  — средняя точка перехода). Поскольку для разрушения двух водородных связей АТ-пар требуется меньше энергии, чем для разрыва трех водородных связей GC-пар, значения температуры и рН, при которых происходит денатурация, зависят от нуклеотидного состава ДНК. Чем выше содержание GC-пар, тем выше  $T_m$  или  $pH_m$  (рис. 1.13).

Денатурация — процесс обратимый, последующее восстановление двухцепочечной структуры ДНК может происходить даже при полном расхождении цепей. Процесс воссоединения, называемый **ренатурацией, реассоциацией** или **отжигом**, происходит при понижении температуры или рН (рис. 1.12). Если температура или рН понижаются постепенно, то цепи соединяются правильно, с восстановлением всех исходных пар оснований. При резком понижении температуры или рН правильное воссоединение комплементарных цепей затрудняется из-за спаривания оснований локально комплементарных участков в пределах одной или разных цепей (рис. 1.14). Диссоциация (денатурация) и реассоциация (ренатурация) ДНК в растворе являются по сути искусственным воссозданием процессов, играющих ключевую роль в реализации разнообразных биологических функций *in vivo*. Очень важным для дальнейшего изложения представляется то, что способность двух отдельных комплементарных цепей нуклеиновой кислоты воссоединиться с образованием исходной структуры является ключевым моментом для проведения соответствующих опы-



**РИС. 1.14.**

Ренатурация комплементарных цепей ДНК при плавном (слева) и резком (справа) понижении температуры или рН раствора.

тов *in vitro*, а также для выделения, сравнения и идентификации специфических нуклеиновых кислот. Уникальная способность нуклеиновой кислоты образовывать двойные спирали путем ассоциации одиночных комплементарных цепей имеет огромное значение для самых разных областей генетики.

### ж. Упаковка ДНК в хромосомах

В клетках или вирусах ДНК, по-видимому, никогда не находится в свободной, вытянутой форме. Она связана с низкомолекулярными катионами — ионами двухвалентных металлов либо с ди- и полиаминами или белками, а возможно, с теми и с другими. Взаимодействие осуществляется с помощью электростатических сил — отрицательно заряженные фосфатные группы частично нейтрализуются положительно заряженными ионами металлов и полиаминами или основными аминокислотными остатками белков. В результате таких взаимодействий происходит конденсация ДНК с уменьшением объема, занимаемого молекулой, иногда в тысячу раз. Кольцевая ДНК *E. coli* длиной 1,4 мм заключена в клетку, имеющую форму палочки диаметром 1 мкм и длиной 2 мкм; у эукариотических клеток ядерная ДНК длиной почти 2 м в стадии интерфазы заключена в ядре диаметром менее 10 мкм. Ядерная ДНК в клетках, находящихся в стадии митоза, конденсирована еще больше и в световом микроскопе имеет вид очень компактной структуры (рис. 1.7).

**Хромосомы эукариот.** Хромосомы эукариотических клеток состоят в основном из хроматина — комплекса двухцепочечной ДНК и пяти гистоновых белков, обозначаемых **H1**, **H2A**, **H2B**, **H3** и **H4** (табл. 1.4). Гистоны могут быть ацетилированы, метилированы, фосфорилированы, poly(ADP)-рибозилированы, а гистоны H2A и H2B — ковалентно связаны с белком, называемым убиквитином. Какова роль воздействия указанных компонентов на структуру и функции гистонов — до конца не выяснено. Гистон H1 млекопитающих состоит из примерно 215 аминокислот; размеры других гистонов варьируют от 100 до 135 аминокислот. Все они содержат необычно большое количество положи-

тельно заряженной аминокислоты лизина; H3 и H4 отличаются от других тем, что у них достаточно высок уровень положительно заряженной аминокислоты аргинина. Соотношение между H2A, H2B, H3 и H4, содержащимися в хроматине низших эукариот (дрожжи, плесневые грибы), такое же, как в хроматине млекопитающих.

На электронно-микроскопических фотографиях в зависимости от условий выделения и степени растяжения хроматин выглядит либо как длинное волокно диаметром 10 нм, либо чаще как более вытянутое волокно с утолщениями — «бусинками» диаметром 10 нм, нанизанными по всей длине волокна с определенными интервалами (рис. 1.15). Каждая из этих бусинок представляет собой **нуклеосомный кор**, на который намотан сегмент хромосомной ДНК длиной 145 пар оснований. Кор — это гистоновый октамер, состоящий из гистонов H2A, H2B, H3 и H4, по две молекулы каждого вида (рис. 1.16). Молекула ДНК, обвиваясь  $1\frac{3}{4}$  раза вокруг нуклеосомного кора, образует сверхспираль.

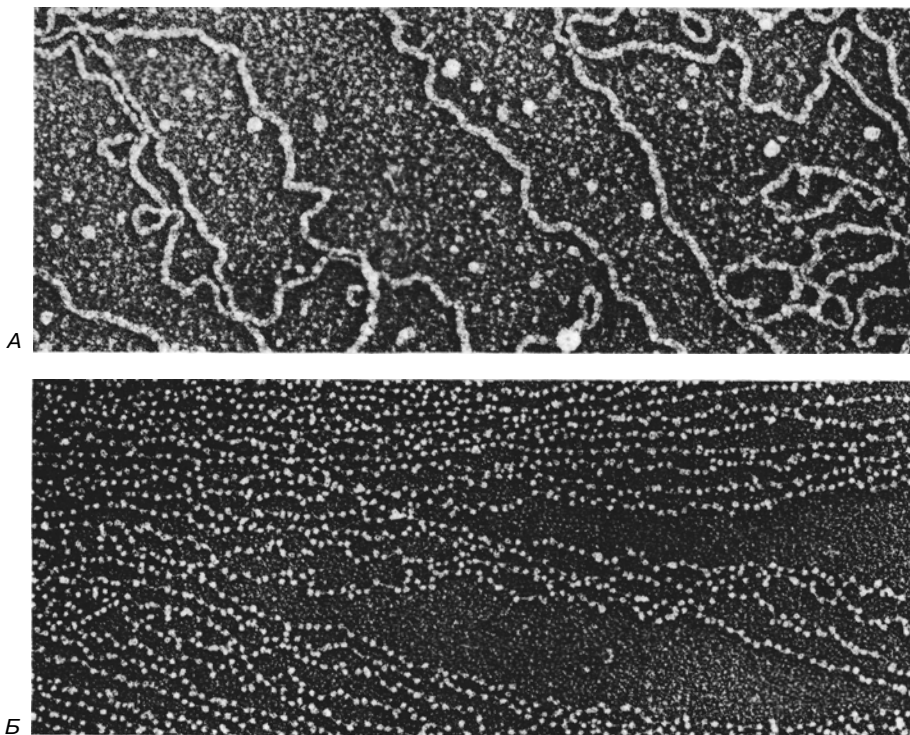
Пятый гистон, H1, не входит в состав нуклеосомного кора и не участвует в процессе наматывания ДНК на гистоновый октамер. Он контактирует с ДНК в тех местах, где двойная спираль входит и выходит из нуклеосомного кора (рис. 1.17). В такой структуре с одним гистоновым октамером и молекулой гистона H1 ассоциированы 168 пар оснований спиральной ДНК. Как мы уже отмечали, на электронно-микроскопических фотографиях хроматин часто обнаруживается в двух альтернативных формах: в форме волокна с четко разделенными нуклеосомами (нуклеосомы имеют вид бусинок, нанизанных на нитку) или в форме волокна диаметром 10 нм, в котором нуклеосомы упакованы бок о бок по всей его длине (рис. 1.15). Волокно диаметром 10 нм может подвергаться дальнейшей конденсации с образованием структур более высокого порядка. При этом нуклеосомы, по всей видимости, образуют соленид — структуру диаметром 30 нм (рис. 1.18).

В результате взаимодействия ДНК с гистонами сегмент двойной спирали ДНК из 168 пар оснований со средним диаметром 2 нм и длиной 57 нм превращается в спираль диаметром 10 нм и длиной 5 нм (рис. 1.18). При последующем сжатии этой спирали до волокна диаметром 30 нм степень конденсации увеличивается еще в шесть раз. Таким образом, упаковка дуплекса ДНК с пятью гистонами приводит к 50-кратной конденсации ДНК. Однако даже столь высокая степень конденсации не может объяснить почти 5000-кратное уплотнение ДНК в метафазной хромосоме.

Эукариотический хроматин содержит и другие белки, которые обычно называют негистоновыми. Некоторые из них, например ферменты, необхо-

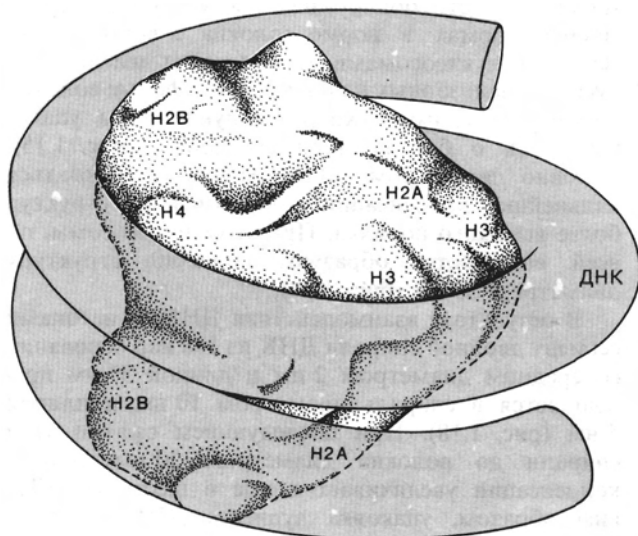
Таблица 1.4. Типичные характеристики гистонов млекопитающих

Тип	Число аминокислот	Мол. масса, кДа	Число основных аминокислот	Отношение Lys/Arg	Число кислых аминокислот
H1 (кролик)	213	23,0	65	21	12
H2A (корова)	129	14,0	26	1,2	20
H2B (корова)	125	13,8	28	2,5	16
H3 (корова)	135	15,3	32	0,7	18
H4 (корова)	102	11,3	26	0,8	10

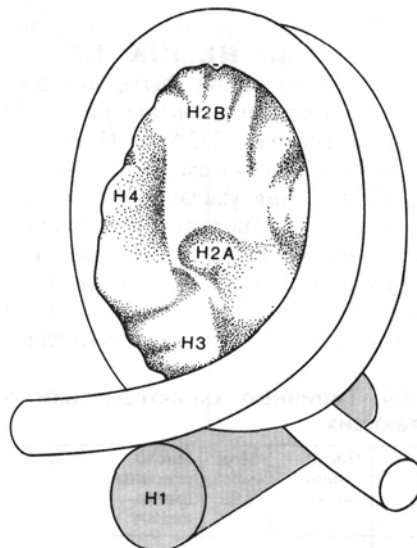
**РИС. 1.15.**

Электронные микрофотографии хроматина. А. Волокно хроматина диаметром 10 нм из почечных клеток CV1 обезьяны. (С любезного разрешения J. Griffith.) Б. Хро-

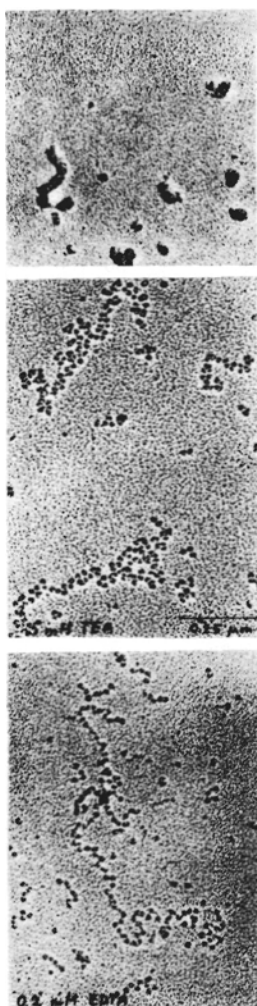
матин из эритроцитов цыпленка, имеющий вид нити с нанизанными на нее бусинками. (С любезного разрешения H. Zentgraf.)

**РИС. 1.16.**

Модель нуклеосомного кора, построенная по данным кристаллографического анализа низкого и высокого разрешения. Сегмент ДНК (145 пар оснований), изображенный в виде трубки, обвивает гистоновый октамер, делая вокруг него  $1\frac{3}{4}$  оборота. [R. D. Kornberg, A. Klug, *Sci. Amer.*, **244** (2) (1981), p. 52.]

**РИС. 1.17.**

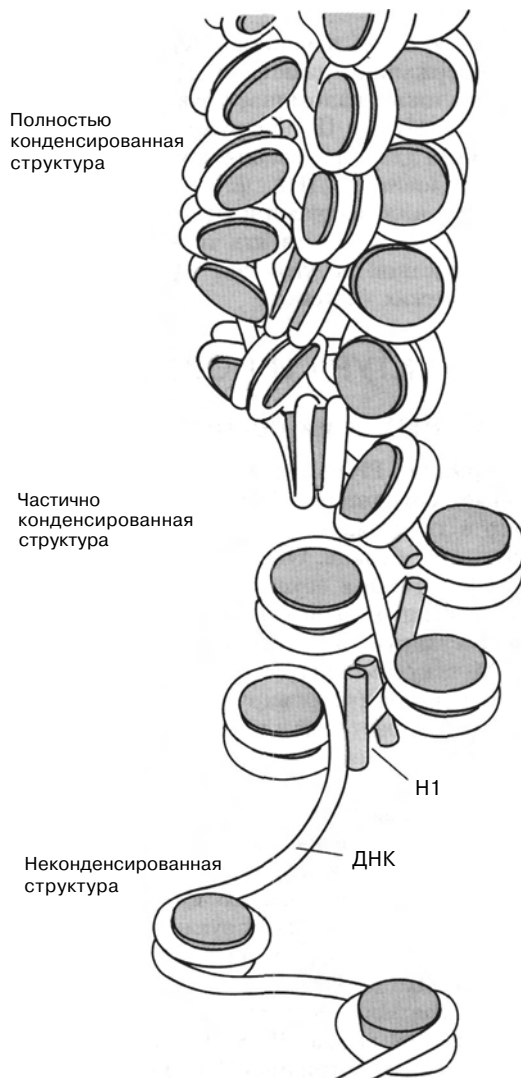
Гистон H1 «сшивает» ДНК в местах, где она начинает и прекращает наматываться на нуклеосомный кор. [A. Klug, *Les Prix Nobel* (Stockholm, Sweden, Nobel Foundation, 1982), p. 93.]

**РИС. 1.18.**

Структура хроматина с разной степенью конденсации. В нижней части рисунка представлен хроматин, находящийся в растянутой форме; он имеет вид нити с нанизанными на нее бусинками. Далее изображен хроматин в частично конденсированной форме, представляющий собой волокно диаметром 10 нм. В верхней части рисунка представлен хроматин в наиболее конденсированном состоянии, когда волокно диаметром 10 нм

димые для репликации и экспрессии ДНК, могут связываться с хроматином временно. Белки, принимающие участие в различных процессах регуляции, связываются с ДНК только в специфических тканях или на определенных стадиях дифференциации. Все эти вопросы, а также роль альтернативных способов организации хроматина в процессах репликации и экспрессии мы детально рассмотрим в последующих главах.

*Хромосомы прокариот.* Насколько известно, в



образует соленоид диаметром 30 нм. Обратите внимание на взаимодействие молекул гистона Н1, связанных с каждой нуклеосомой, которое способствует конденсации волокна диаметром 10 нм в более плотную структуру. [Электронные микрофотографии любезно предоставлены А. Klug. Рисунок выполнен по R.D. Kornberg, A. Klug, *Sci. Amer.*, **244** (2) (1981), p. 52.]

упаковке прокариотической геномной ДНК участвуют только два или три белка. О природе взаимодействия этих белков с ДНК и о структуре конденсированного комплекса белок-нуклеиновая кислота известно немного. У *E. coli*, по-видимому, существует лишь один белок или один класс ДНК-связывающих белков, называемых НУ-белками; по своему размеру, содержанию лизина и аргинина, антигенным свойствам они сходны с эукариотическим гистоном Н2А. Другой белок, белок II, обна-

руженный у *E. coli* и цианобактерий, по повышенному содержанию лизина и ДНК-связывающим свойствам также напоминает эукариотический гистон. Белки HU и II обнаружены в количествах, достаточных для образования комплекса по крайней мере с половиной ДНК *E. coli* и, по-видимому, совместно с полиаминами и еще неизвестными нам белками могут осуществлять те же самые функции при конденсации и упаковке ДНК, что и пять эукариотических гистонов.

## 1.2. СТРУКТУРА И ПОВЕДЕНИЕ РНК

### а. Типы РНК и их распространенность

Содержание РНК в любых клетках в 5–10 раз превышает содержание ДНК. Основная роль РНК состоит в трансляции генетической информации с образованием белков. Однако молекулы РНК принимают участие и в осуществлении некоторых специализированных эндонуклеазных функций, возможно регулирующих различные этапы экспрессии генов. Молекулами РНК представлены геномы некоторых вирусов (ретровирусов и множества вирусов животных, растений и насекомых с одно- и двухцепочечным геномом).

Во всех клетках присутствуют следующие виды РНК: **рибосомная РНК (рРНК)**, **транспортная РНК (тРНК)** и **информационная, или матричная, РНК (мРНК)**. Большинство (если не все) клеток содержат также много других малых цитоплазматических РНК (мцРНК), а в клетках эукариот присутствует еще и множество малых ядерных РНК (мяРНК) (табл. 1.5). Около 80% массы клеточных РНК составляют три или четыре вида рРНК, а около 15% – почти 100 видов тРНК. На долю нескольких тысяч различных матричных РНК приходится менее 5% клеточной РНК, а на долю малых ядерной и цитоплазматической РНК, число видов которых пока неизвестно, – менее 2% от общего количества.

### б. Компоненты молекулы РНК и соединяющие их химические связи

РНК – это полинуклеотид длиной от 70 мономерных единиц у некоторых тРНК до 10000 и более у некоторых мРНК. На рис. 1.19 приведены структурные формулы четырех нуклеотидов, обычно встречающихся в РНК. Два пурина (аденин и гуанин) и один пиримидин (цитозин) входят также в состав ДНК. А вместо тимина (5-метилдикетопиримидина) в РНК входит урацил, у которого 5-метильная группа отсутствует. Нуклеотиды в молекуле РНК соединены в цепочку такими же 5'–3'-фосфодиэфирными связями, как и в ДНК (рис. 1.20). Из-за наличия 2'-ОН-группы связь Р–О

Таблица 1.5. Некоторые основные РНК

Тип РНК	Приблизительное число разных видов в клетках	Приблизительная длина (число нуклеотидов)	Распространенность <sup>1)</sup>
Транспортная РНК (тРНК)	80-100	75-90	П, Э
Рибосомная 5S-РНК (рРНК)	1-2	120	П, Э
Рибосомная 5,8S-РНК (рРНК)	1	155	Э
Рибосомная 16S-РНК (рРНК)	1	1600	П
Рибосомная 23S-РНК (рРНК)	1	3200	П
Рибосомная 18S-РНК (рРНК)	1	1900	Э
Рибосомная 28S-РНК (рРНК)	1	5000	Э
Матричная РНК (мРНК)	Тысячи	Варьирует	П, Э
Гетерогенная ядерная РНК (гяРНК)	»	»	Э
Малая цитоплазматическая РНК (мцРНК)	Десятки	90-330	П, Э
Малая ядерная РНК (мяРНК)	»	58-220	Э

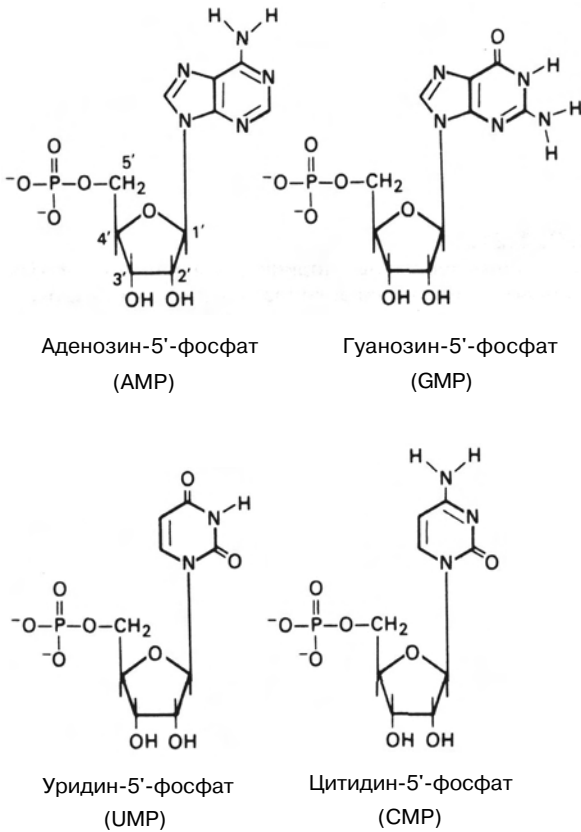
<sup>1)</sup> П - прокариоты, Э - эукариоты.

чувствительна к действию щелочей и ферментов, расщепляющих РНК.

Некоторые пурины и пиримидины в РНК модифицированы: они содержат метил-, тиол-, водород- и изопентенил-заместители. В РНК присутствуют 2'-О-метилнуклеотиды с модифицированным остатком рибозы, а также может наблюдаться иной способ связывания между урацилом и рибозой (так называемая псевдоуридиловая кислота) (рис. 1.21). Такие модифицированные нуклеотиды довольно редко встречаются в рРНК и мРНК и достаточно часто – в тРНК. Как правило, модификация оснований и рибозных остатков происходит после завершения синтеза РНК, а не на стадии биосинтетических предшественников. Функциональное значение этого явления установлено лишь отчасти.

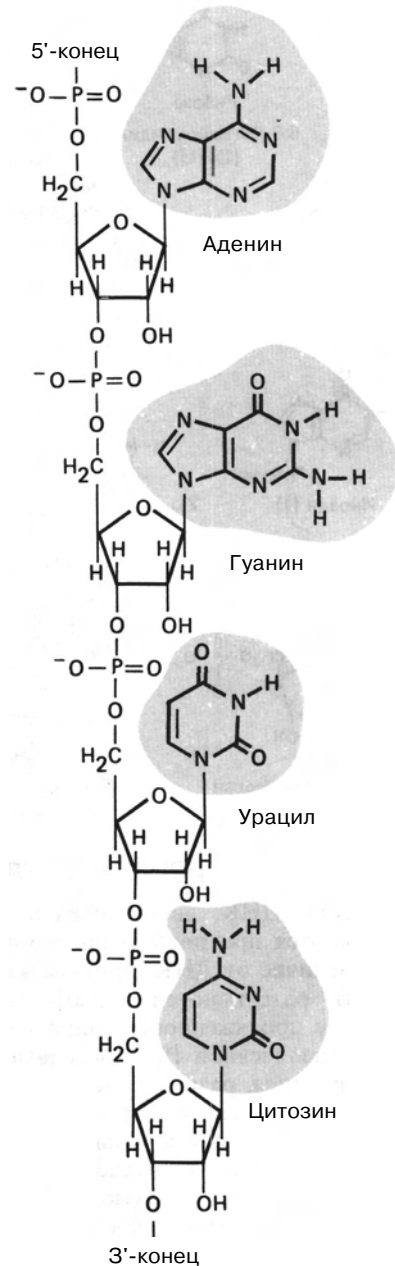
### в. Структура РНК

Большинство клеточных РНК – одноцепочечные молекулы, хотя некоторые вирусные геномы (в том числе геномы ретровирусов) представлены двухцепочечными РНК, напоминающими А-форму ДНК. В одиночных цепях все время образуются короткие

**РИС. 1.19.**

Структурные формулы четырех рибонуклеотидов, обычно встречающихся в РНК.

внутримолекулярные двухцепочечные участки. Это связано с тем, что в большинстве РНК имеются небольшие комплементарные последовательности, которые спариваются и образуют петли (рис. 1.22). В таких двухцепочечных участках А спаривается с U, а G с C; G может образовать пару и с U, но GU-пара менее стабильна, чем стандартная пара GC, поскольку ее компоненты соединены двумя, а не тремя водородными связями. Двухцепочечные области, образованные подобным образом, обычно непротяженны и прерывисты, поскольку спариваемые участки редко бывают абсолютно комплементарными. Укладка большинства РНК может происходить более чем одним способом, однако биологическое значение образующихся при этом изомеров установлено только в некоторых случаях. Например, известно, что адекватная укладка некоторых вирусных РНК чрезвычайно важна для экспрессии генов, поскольку ответ на ключевые регуляторные сигналы зависит от конфигурации молекулы (разд. 3.11.е). Подобная зависимость функции от упаковки молекулы наилучшим образом про-

**РИС. 1.20.**

Участок рибонуклеотидной цепи.

демонстрирована на примере тРНК. Несмотря на разную нуклеотидную последовательность, третичная структура разнообразных тРНК весьма сходна, и ее стабилизация, по-видимому, имеет огромное значение для функционирования этих молекул (рис. 1.23). Правильная терминация синтеза РНК и созревание (процессинг) транскрипта тоже часто зависят от характера укладки РНК.

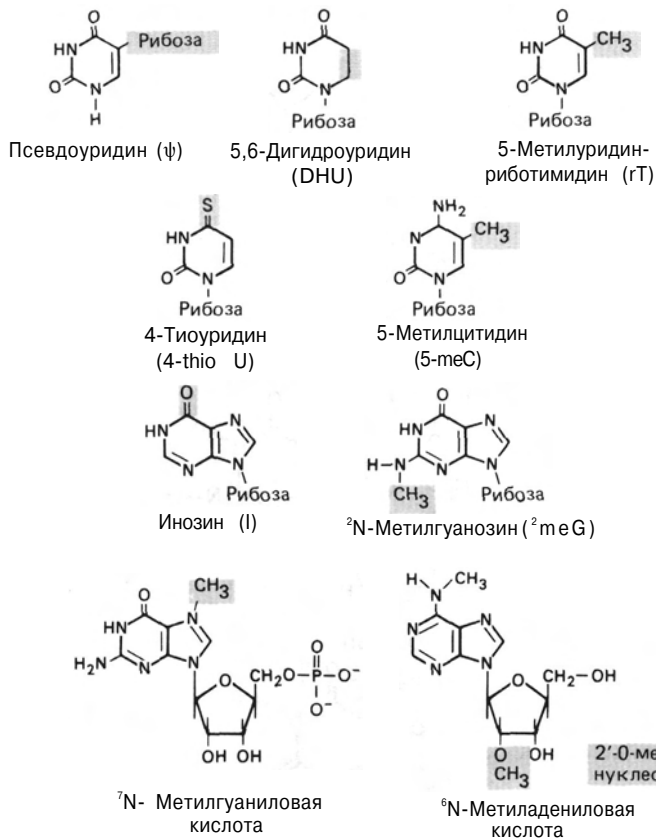


РИС. 1.21.

Структурные формулы модифицированных пуриновых и пиримидиновых нуклеозидов, в небольших количествах встречающихся в некоторых РНК.

## г. Денатурация и ренатурация РНК

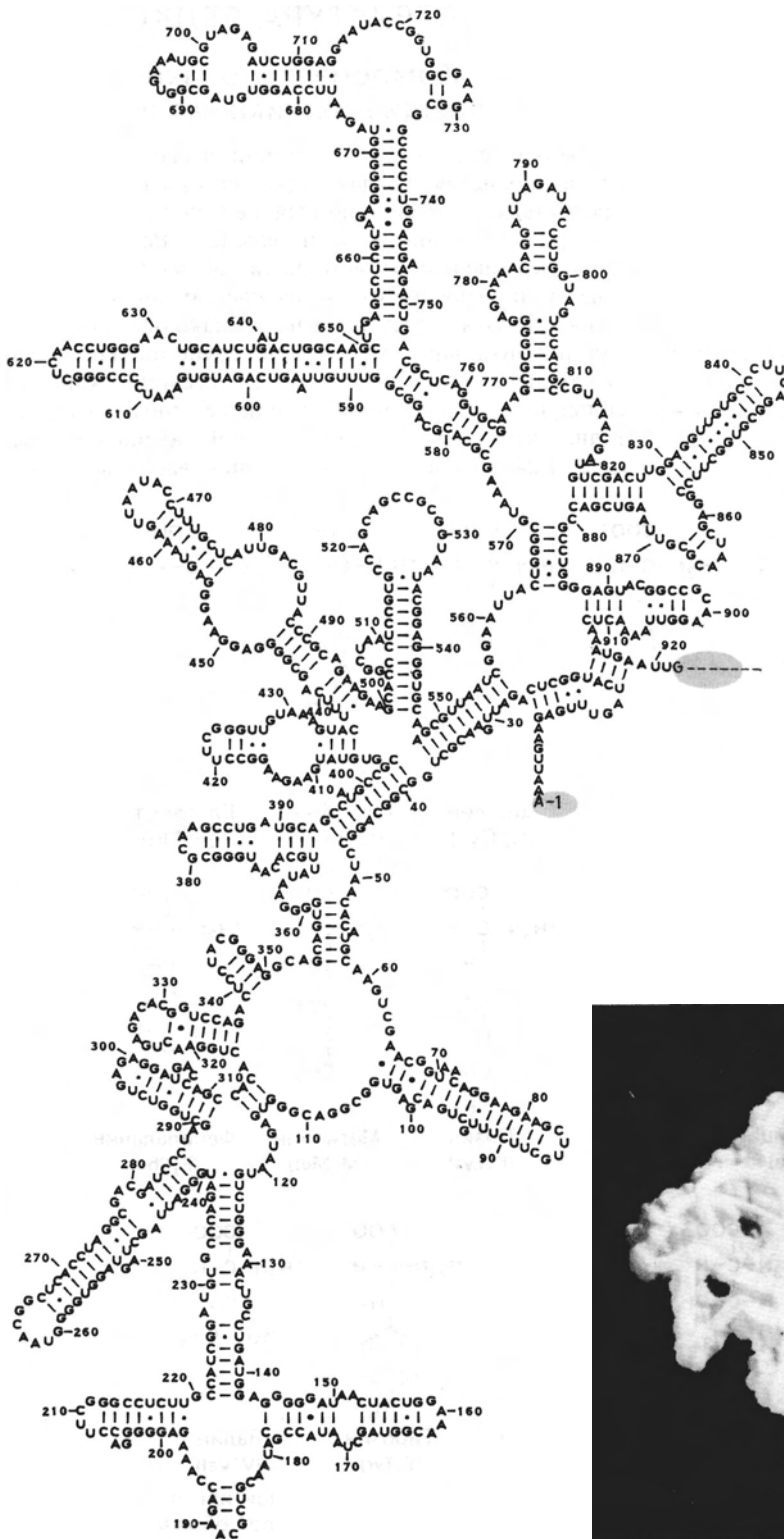
Как и в случае ДНК, двухцепочечные участки в РНК разрушаются при повышении температуры или рН, но, в отличие от ДНК, при высоких значениях рН в РНК разрушаются и фосфодиэфирные связи. Поскольку протяженность спирализованных участков в одноцепочечной РНК невелика, а сами спирали несовершенны, разрушаются они довольно легко. Однако полностью комплементарная двухцепочечная РНК плавится в довольно узком температурном интервале, как и двухцепочечная ДНК. В результате денатурации образуются две комплементарные одиночные цепи, способные к последующему воссоединению при плавном понижении температуры. После денатурации двухцепочечных участков одноцепочечной РНК восстановление тех же спаренных областей оказывается затрудненным, и в результате ренатурации могут образоваться структуры, отличные от исходной.

## д. Гибридные спирали ДНК-РНК

Полинуклеотидные цепочки РНК и ДНК, имеющие комплементарные последовательности, могут образовывать двойную спираль РНК-ДНК

с антипараллельными цепями. Структура таких комплексов напоминает А-ДНК (рис. 1.8). Спаривание оснований в них отвечает правилам Уотсона-Крика для ДНК: dA спаривается с rU, rA—с dT, dG—с rC и dC—с rG. В дуплексах РНК-ДНК пара оснований rG•dC более стабильна, чем пара dG•rC, и обе они стабильнее пар dG•dC в ДНК; обе пары, dT•rA и rU•dA, менее стабильны, чем пара dA•dT в ДНК. Гибридные спирали, или гетеродуплексы, способны к денатурации и ренатурации, как и дуплексы ДНК. В водных растворах с умеренной концентрацией солей двухцепочечные комплексы РНК-ДНК денатурируют легче, чем соответствующие спирали, состоящие исключительно из цепей РНК и ДНК. В растворах же, содержащих формамид и соли в определенных концентрациях, дуплекс РНК-ДНК стабильнее дуплекса ДНК. Эти условия могут использоваться для образования и специфической стабилизации гибридных дуплексов.

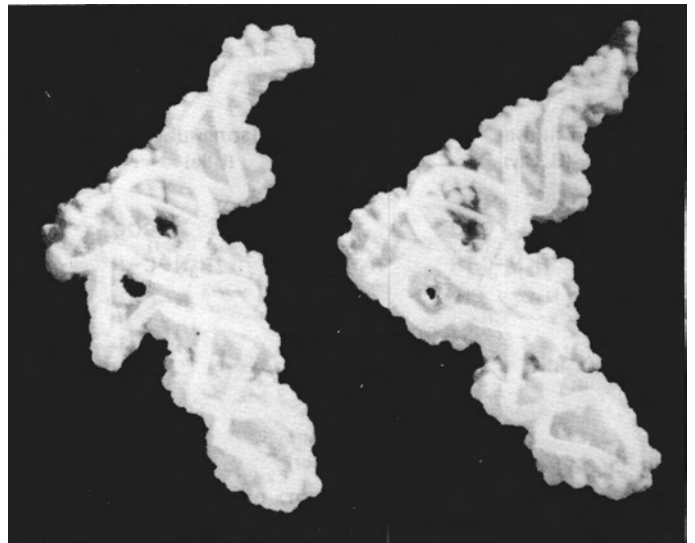
Полинуклеотидные цепи РНК или ДНК считаются комплементарными, если они способны образовать протяженную двойную спираль с уотсон-криковскими парами оснований. Две нуклеиновые кислоты можно назвать **гомологичными**, если их нуклеотидные последовательности идентичны или очень близки. Гомологичность двух нуклеино-

**РИС. 1.22.**

5'-концевой сегмент 16S-rPHK *E. coli*. Видно, что благодаря внутримолекулярному спариванию оснований происходит укладка цепи. Заметим, что в стабилизации дуплексной структуры РНК участвуют G•U-пары. Несколько оснований могут образовывать петлю, если в противоположной цепи отсутствуют комплементарные основания. Цветом выделены первое основание цепи и место продолжения цепи в 3'-направлении. (С любезного разрешения Н. Noller.)

**РИС. 1.23.**

Структура tPHK<sup>Asp</sup> и tPHK<sup>Phe</sup>, воссозданная с помощью компьютерной графики. Остов каждой цепи РНК имеет вид сплошной трубки. Такая компактная структура стабилизируется благодаря спариванию оснований, далеко отстоящих друг от друга в нуклеотидной последовательности. 5'- и 3'-концы цепей РНК находятся в правом верхнем углу структуры. (Модели и фото предоставлены Arthur J. Olson, Ph. D. Copyright 1987, Research Institute of the Scripps Clinic.)



asp-tPHK

phe-tPHK

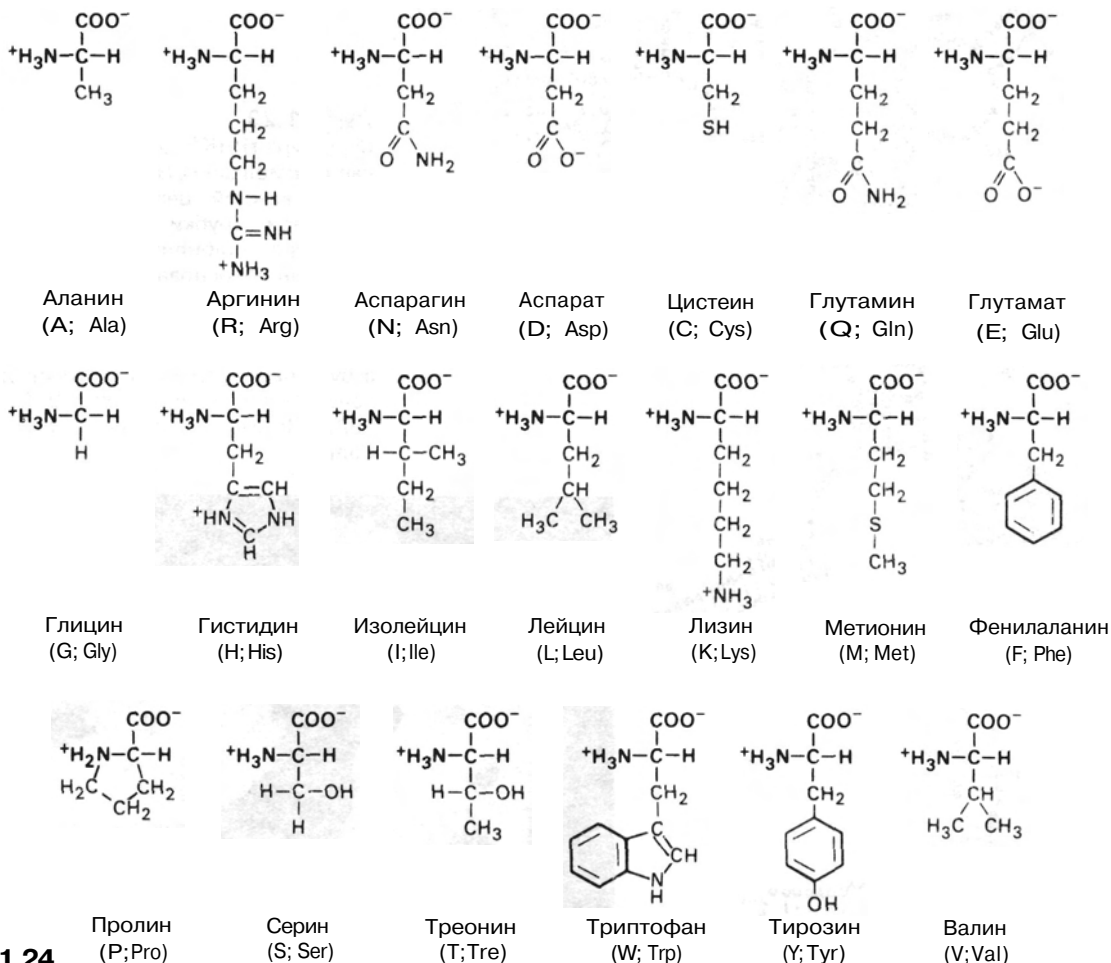


вых кислот можно определить по степени образования двойных спиралей соответствующими одинокими цепями РНК, ДНК или гибридов РНК–ДНК. За образованием ДНК-, РНК- и гибридных РНК–ДНК-спиралей можно следить разными способами. В качестве зонда для выявления и количественного анализа гомологичных молекул РНК часто используется одноцепочечная ДНК. Так, степень превращения одноцепочечной РНК в двухцепочечную форму в присутствии другой нуклеиновой кислоты является мерой гомологичности их последовательностей. Используя соответствующего партнера по **гибридизации**, можно установить протяженность специфической нуклеотидной последовательности и гомологичность двух ДНК, РНК и ДНК, а также двух РНК.

## 1.3. СТРУКТУРА БЕЛКОВ

### а. Компоненты белков и соединяющие их химические связи

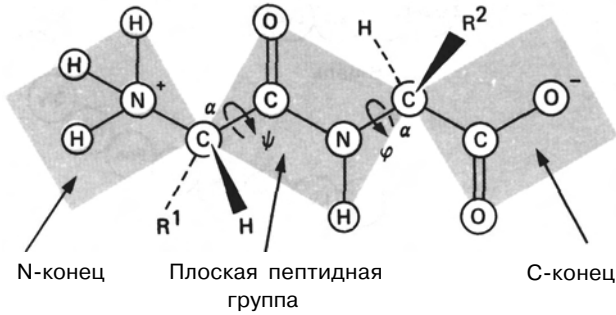
Белки состоят из одной или нескольких **полипептидных цепей**, каждая из которых в свою очередь представляет собой длинный неразветвленный полимер, состоящий из аминокислот. Все полипептиды независимо от источника—от вирусов до человека—построены из 20 разных аминокислот. В аминокислотах имеются как одинаковые для всех, так и уникальные химические группировки: атом углерода ( $\alpha$ -углеродный атом), несущий карбоксильную и амино-группы, и определенные заместители, характерные для каждой аминокислоты (рис. 1.24). Такие «боковые цепи», часто называе-



**РИС. 1.24.**

Структурные формулы аминокислот, обычно встречающихся в белках. Вариативная часть аминокислот выделена цветом. Под формулами даны полное название аминокислот, одно- и трехбуквенные обозначения. Очень

часто при изображении пептидной или полипептидной цепи боковые группы аминокислот обозначают буквой R.



**РИС. 1.25.**

Структура дипептида. В верхней части рисунка показано, как образуется пептидная связь. Представлены заряженные концевые амино- и карбоксильная группы и  $\alpha$ -углеродные атомы, к которым присоединены боковые группы. В нижней части рисунка указаны углы вращения ( $\psi$  и  $\phi$ ), характерные для пептидных групп. R-группы и атомы водорода, связанные с  $\alpha$ -углеродными атомами, выступают из плоскости рисунка вверх или вниз.

мые R-группами, различаются по размеру, форме, заряду и химической активности.

Остов любой белковой цепи образуется с помощью амидных (или пептидных) связей, соединяющих аминогруппу одной аминокислоты с карбоксильной группой другой, соседней аминокислоты. На рис. 1.25 показано простейшее полипептидное звено—дипептид, а на рис. 1.26 представлены геометрия полипептидной цепи и положение различных заместителей гептапептида. Таким образом, полипептиды—это длинные цепи, образуемые с помощью регулярно повторяющихся пептидных связей и содержащие набор боковых групп, распо-

ложенных вдоль остова. Полипептидная цепь имеет определенное направление. На одном ее конце находится свободная аминогруппа (в нейтральных условиях эта группа протонируется с образованием  $\text{NH}_3^+$ ); соответствующая аминокислота называется N-концевой, такое же название—и у данного конца цепи. На другом конце цепи находится карбоксильная группа, обычно существующая в виде  $\text{COO}^-$ -аниона; аминокислота в этой позиции называется C-концевой, такое же название имеет и соответствующий конец цепи. В некоторых белках между цистеиновыми остатками одной цепи образуются дисульфидные связи, объединяя разные участки цепи (рис. 1.27). Такие дисульфидные мостики могут объединять и разные полипептидные цепи независимо от того, идентичны они или нет. Полипептидные субъединицы некоторых олигомерных белков объединены именно таким образом.

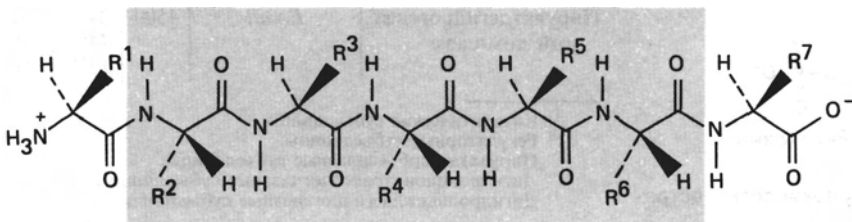
В некоторых белках встречаются в небольших количествах модифицированные формы природных аминокислот (рис. 1.28). Практически во всех случаях изменения в структуре аминокислот происходят лишь после образования пептидных связей: в результате гидроксирования уже включенных в белок пролина и лизина получают гидроксипролин и гидроксилизин соответственно; при карбоксилровании глутамата образуется  $\gamma$ -карбоксиглутамат, а при фосфорилировании гидроксильных групп серина и треонина или фенольной группы тирозина—фосфоаминокислоты.

Важными компонентами эукариотических клеток и многих вирусов являются белки, относящиеся к группе гликопротеинов. Они содержат сложные углеводы, ковалентно связанные с входящими в состав белка аспарагиновым, гидроксилизиновым, сериновым и треониновым остатками. Как и в случае модификации аминокислот, описанном выше, на определенных этапах, следующих за сборкой полипептидной цепи, к ней могут присоединяться различные сахара (разд. 3.10). Реакции гликозилирования играют важную роль в процессах транспортировки белков от места их синтеза в специфические клеточные органеллы и к поверхностным структурам клетки; гликопротеины также придают определенный рисунок поверхности клеток.

N-концевой остаток

Пептидные связи

C-концевой остаток



**РИС. 1.26.**

Структура гептапептида. Обозначения такие же, как на рис. 1.25.

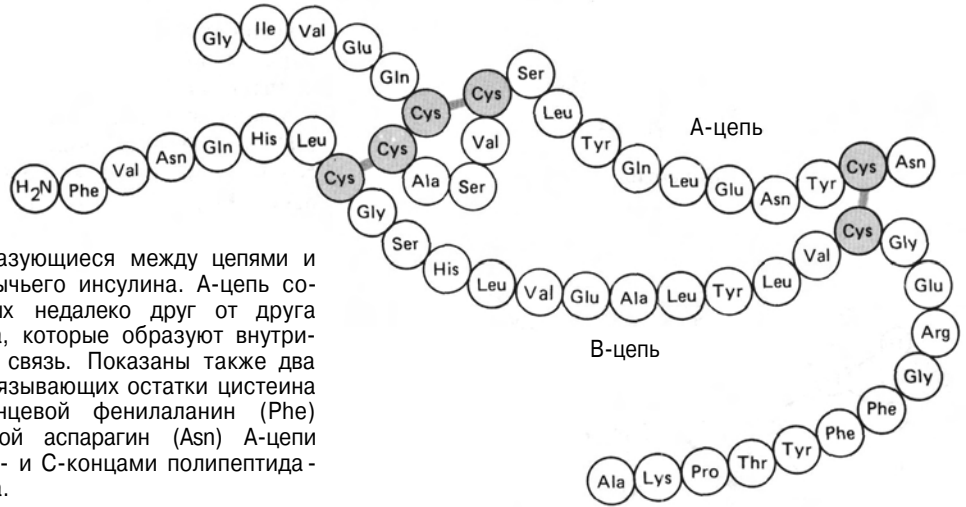


РИС. 1.27.

Дисульфидные связи, образующиеся между цепями и в пределах одной цепи бычьего инсулина. А-цепь содержит два расположенных недалеко друг от друга цистеиновых (Cys) остатка, которые образуют внутрицепочечную дисульфидную связь. Показаны также два дисульфидных мостика, связывающих остатки цистеина А- и В-цепей. Аминоконцевой фенилаланин (Phe) В-цепи и карбоксиконцевой аспарагин (Asn) А-цепи являются соответственно N- и С-концами полипептида-предшественника инсулина.

## б. Размер и форма белков

Размеры полипептидов очень сильно различаются: число составляющих их аминокислот колеблется от 50 до нескольких тысяч. Поскольку с каждой аминокислотой масса белка увеличивается примерно на 110 дальтон, мол. масса белков может варьировать от  $10^4$  до  $10^6$  дальтон. Одни белки состоят только из одной полипептидной цепи, другие — из нескольких одинаковых цепей, третьи — из нескольких цепей разного типа (табл. 1.6). При-

Таблица 1.6. Размер и субъединичный состав некоторых глобулярных белков

Белок	Источник	Мол. масса белка, кДа	Число субъединиц	Мол. масса субъединиц, кДа
ДНК-лигаза	<i>E. coli</i>	75	1	—
ДНК-полимераза I	<i>E. coli</i>	109	1	—
Щелочная фосфатаза	<i>E. coli</i>	86	2	43
Лас-репрессор	<i>E. coli</i>	160	4	40
$\beta$ -Галактозидаза	<i>E. coli</i>	544	4	135
Глутаминсинтетаза	<i>E. coli</i>	592	12	49
Гемоглобин	Млекопитающие	64	2 $\alpha$	16
			2 $\beta$	16
Триптофансинтаза	<i>E. coli</i>	148	2 $\alpha$	29
			2 $\beta$	45
Аспартат-транскарбамилаза	<i>E. coli</i>	310	6(C) <sup>1)</sup>	33
			6(R) <sup>2)</sup>	17
РНК-полимераза (основной фермент)	<i>E. coli</i>	400	2( $\alpha$ )	40
			1( $\beta$ )	155
			1( $\beta'$ )	165
Пируватдегидрогеназный комплекс	<i>E. coli</i>	4500	24 <sup>3)</sup>	91
			24 <sup>4)</sup>	70
			12 <sup>5)</sup>	56

<sup>1)</sup> Каталитические субъединицы.

<sup>2)</sup> Регуляторные субъединицы

<sup>3)</sup> Пируватдекарбоксилазные субъединицы.

<sup>4)</sup> Дигидролипоилтрансферазные субъединицы

<sup>5)</sup> Дигидролипоилдегидрогеназные субъединицы.

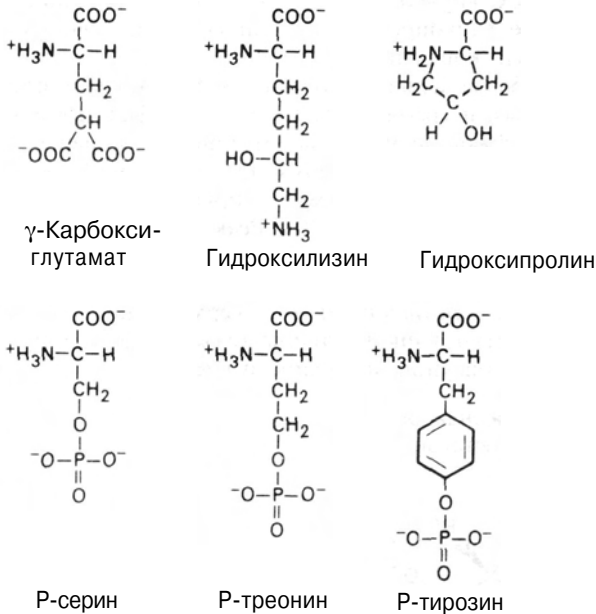


РИС. 1.28.

Некоторые модифицированные аминокислоты, встречающиеся в белках.

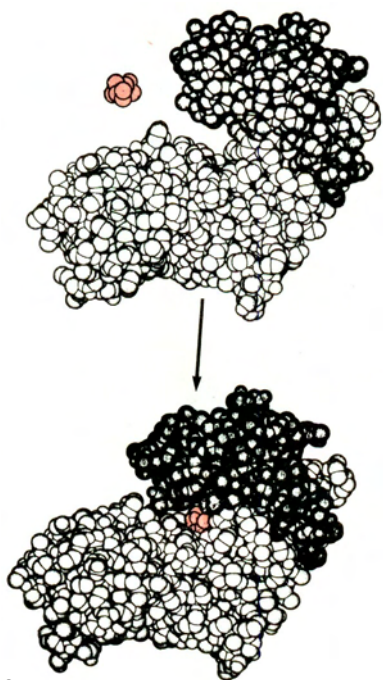


РИС. 1.29.

Пространственная модель глобулярного белка гексокиназы. Вверху показаны молекула гексокиназы и один из ее субстратов, D-глюкоза, в свободном состоянии. После связывания глюкозы домены белка (темный и светлый цвета) сближаются и закрывают карман, в котором находится центр связывания. [W.S. Bennett, Jr., T.A. Steitz, J. Mol. Biol., **140** (1980), p. 211.]

мером белка, представленного одной полипептидной цепью, служит миоглобин; глутаматсинтетазы бактерий содержит 12 идентичных полипептидных цепей. Гемоглобин—тетрамер, состоящий из цепей двух разных видов, а функциональная форма ДНК-полимеразы *E. coli* состоит из полипептидных цепей по крайней мере четырех разных видов.

В растворе белки имеют строго определенную конформацию, или трехмерную структуру. Биологическая активность почти всех без исключения белков, будь то белки-катализаторы, структурные белки, белки, ответственные за транспортные процессы, белки, участвующие в формировании опорно-двигательного аппарата, или белки-регуляторы, зависит от сохранения их природной, или активной, конформации. Белки в соответствии с их конформацией можно разделить на две категории. **Глобулярные** белки имеют компактную, примерно сферическую форму, образующуюся в результате нерегулярной укладки полипептидных цепей (рис. 1.29). В **фибриллярных** белках полипептидные цепи располагаются параллельно друг другу, образуя длинные нити или слои (рис. 1.30). Большинство ферментов

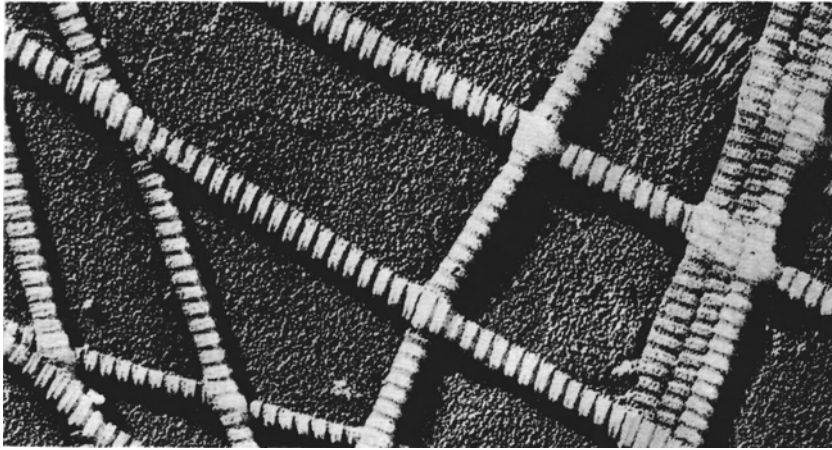
и растворимых белков имеют глобулярную структуру; такие белки, как коллаген и кератин, обнаруживаемые в структурных и соединительных тканях, принимают фибриллярную конформацию. Фибриноген и мышечный миозин могут встречаться как в одной, так и в другой форме.

### в. Чем определяется конформация белка

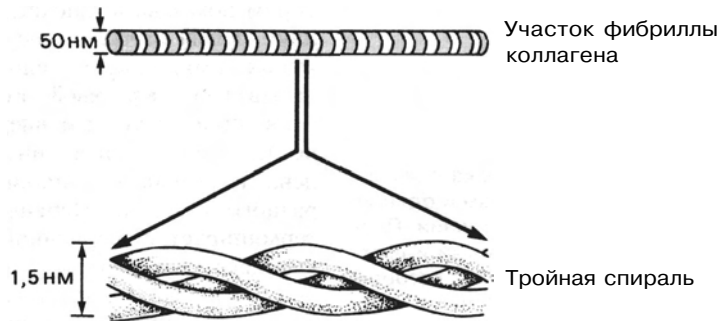
Структуру белка можно рассматривать на разных уровнях организации—на уровне **первичной**, **вторичной**, **третичной** или **четвертичной** структуры. Первые три относятся к структурным характеристикам полипептидных цепей, четвертый отражает структуру олигомерных белков, состоящих из двух или более полипептидных цепей.

Каждый белок обладает уникальной аминокислотной последовательностью (см. рис. 1.27, на котором показана аминокислотная последовательность двух цепей бычьего инсулина); порядок расположения аминокислот вдоль полипептидной цепи называется **первичной структурой**. Уникальность первичной структуры впервые была выявлена при исследовании бычьего инсулина, а затем подтверждена в результате анализа тысяч других белков разного размера. Первичная структура белка детерминируется первичной структурой соответствующего гена. Поэтому при изменении нуклеотидной последовательности гена, кодирующего данный белок, изменяется и первичная структура белка.

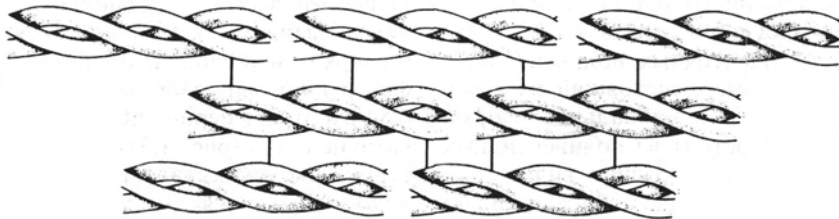
Полипептидные цепи могут укладываться в регулярные структуры, называемые **вторичными**. Наиболее часто встречающимися периодическими конформациями белков являются правозакрученная  $\alpha$ -спираль и  $\beta$ -слой. В  $\alpha$ -спирали остов имеет конфигурацию винтовой спирали с периодичностью 0,54 нм и примерно 3,6 аминокислотными остатками на виток (рис. 1.31). Стабилизация спиральной структуры осуществляется благодаря образованию водородных связей между атомом водорода NH-группы одной аминокислоты и CO-группой четвертой вдоль цепи аминокислоты. Боковые группы аминокислот располагаются на наружной стороне спирали (рис. 1.32). Длина участка данной полипептидной цепи, который может принимать  $\alpha$ -спиральную конфигурацию, зависит от аминокислотного состава и аминокислотной последовательности цепи. Некоторые аминокислоты или последовательности дестабилизируют  $\alpha$ -спираль, а если в цепи встречается пролин или гидроксипролин, то  $\alpha$ -спираль прерывается из-за ограничения вращения вокруг пептидной связи и отсутствия атома водорода для образования водородной связи. Как правило,  $\alpha$ -спиральные участки относительно непротяженны и состоят в среднем из 10–20 аминокислот.



А



Б



В

**РИС. 1.30.**

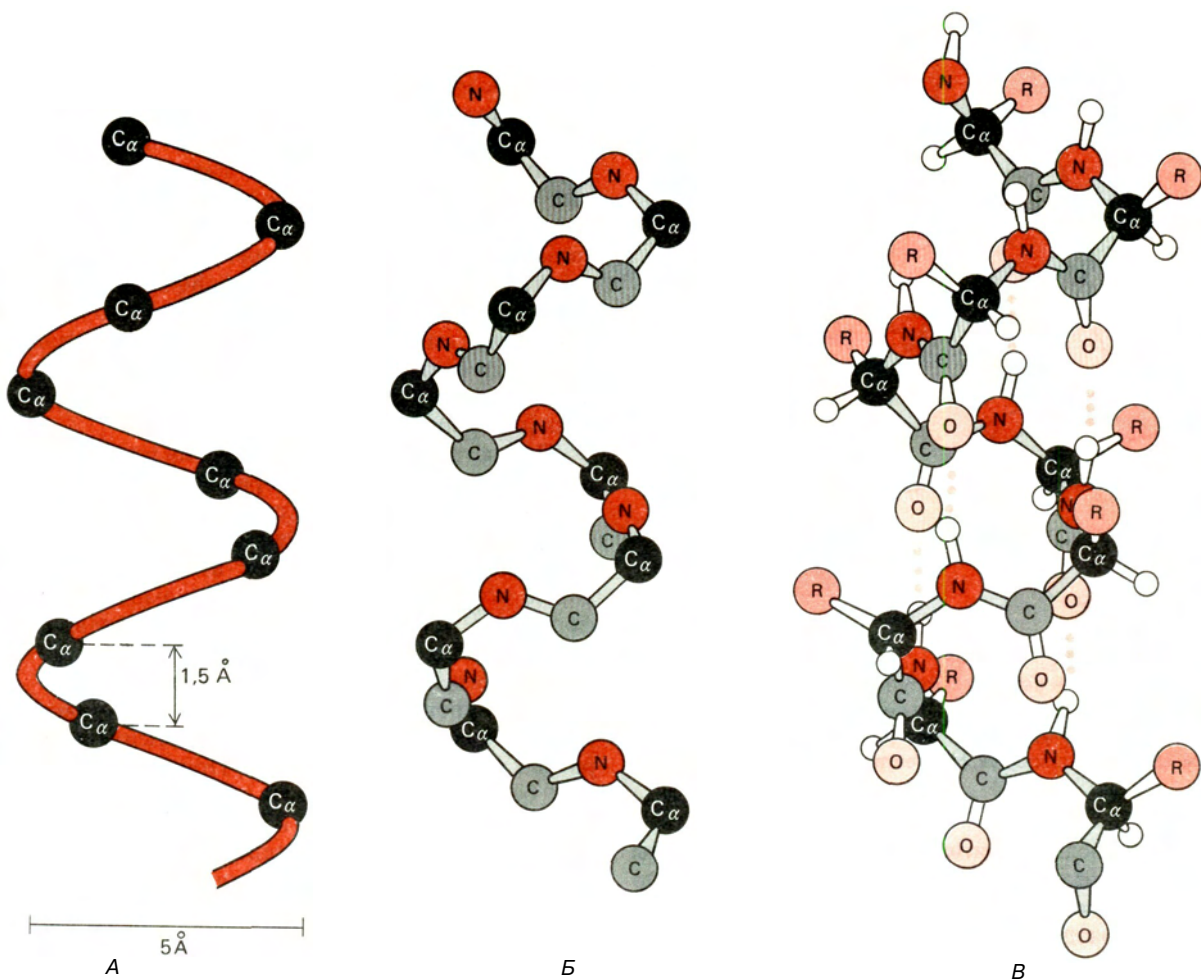
Структуры коллагена. А. Электронная микрофотография intactных фибрилл коллагена кожи. (С любезного разрешения J. Cross.) Б. Тропоколлагеновые волокна, состоящие из трех полипептидных цепей, сплетенных

в тройную спираль. В. Фибриллы образуются путем взаимодействия между несколькими тропоколлагеновыми тройными спиралями.

кислот. Иногда водородные связи, образующиеся между боковыми группами аминокислот в двух  $\alpha$ -спиральных сегментах, соединяют расположенные бок о бок витки спирали одной и той же или разных полипептидных цепей, и тем самым структура еще больше стабилизируется. Известны случаи (примером может служить фибриноген), когда множест-

венные  $\alpha$ -спиральные сегменты одной полипептидной цепи образуют беспорядочные пучки. У некоторых белков, в частности у фермента химотрипсина,  $\alpha$ -спиральные области отсутствуют; у других, например у  $\alpha$ - и  $\beta$ -субъединиц гемоглобина,  $\alpha$ -спиралы встречаются вдоль цепи довольно часто.

Другой тип вторичной структуры белков —



**РИС. 1.31.**

Различные модели правозакрученной  $\alpha$ -спирали. А. Показаны только  $\alpha$ -углеродные атомы. Б. Показаны только скелетные атомы азота (N),  $\alpha$ -углеродные атомы ( $C_\alpha$ ) и атомы углерода карбонильных групп (C). В.

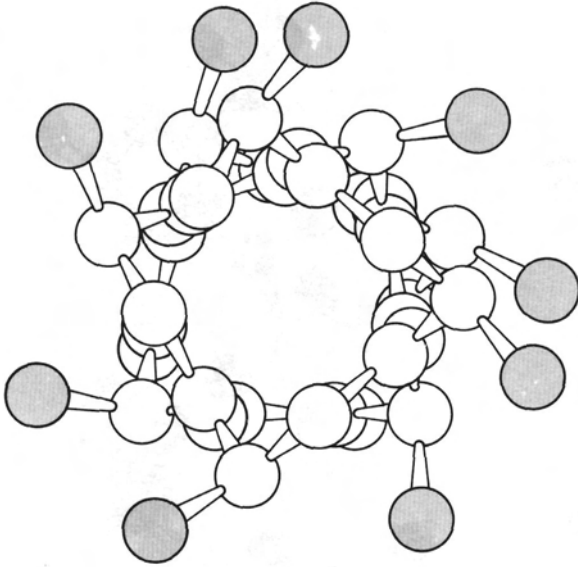
Показана спираль в целом. Ее стабилизируют водородные связи между NH- и CO-группами. [L. Stryer, Biochemistry (San Francisco, W.H. Freeman, 1981).]

$\beta$ -слой, названный так потому, что структура напоминает полотно. В  $\beta$ -слое в отличие от  $\alpha$ -спирали, в которой полипептидная цепь скручена и образует структуру, имеющую форму стержня, полипептидная цепь растянута. Водородные связи в  $\beta$ -слое образуются между отдаленными друг от друга аминокислотами одной полипептидной цепи или между аминокислотами из разных цепей (рис. 1.33). Ряды из двух-пяти цепей, скрепленных водородными связями, находящиеся в параллельной друг другу или антипараллельной ориентации, образуют структуру, напоминающую гофрированную ткань.

Некоторые белки (в частности, коллаген) содержат необычайно много глицина и пролина—тех

двух аминокислот, которые дестабилизируют или разрушают  $\alpha$ -спирали, а также две необычные аминокислоты—гидроксипролин и гидроксизин. Подобные белки образуют вторичную структуру третьего типа; она состоит из трех длинных левозакрученных спиралей, сплетенных в виде плотного каната (рис. 1.34). Структура также стабилизируется водородными связями, образуемыми между боковыми остатками прилегающих друг к другу нитей каната.

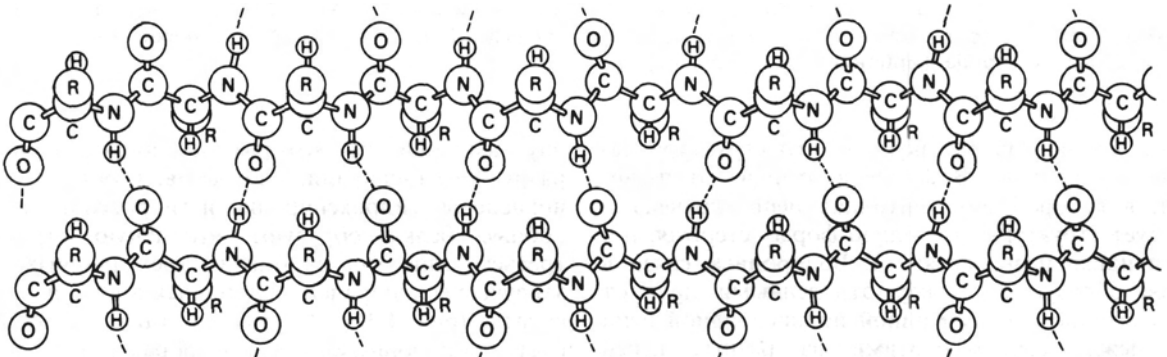
Почти все ферменты и регуляторные белки имеют глобулярную форму; это и есть их **третичная структура**. Боковые цепи полярных аминокислот локализируются на поверхности глобулы, контактирующей с растворителем, а боковые цепи неполяр-



**РИС. 1.32.**

Вид на  $\alpha$ -спираль сверху. R-группы аминокислот выступают из спирали. Для наглядности внутренняя часть спирали дана в увеличенном виде. [L. Stryer, *Biochemistry* (San Francisco: W.H. Freeman, 1981).]

ных аминокислот упрятаны внутри и экранированы от водной среды. Плотно скрученные полипептидные цепи содержат варьирующее число  $\alpha$ -спиралей и  $\beta$ -слоев, разделенных неструктурированными областями, как правило находящимися в местах изгибов остова (рис. 1.29 и 1.35). Биологически актив-



**РИС. 1.33.**

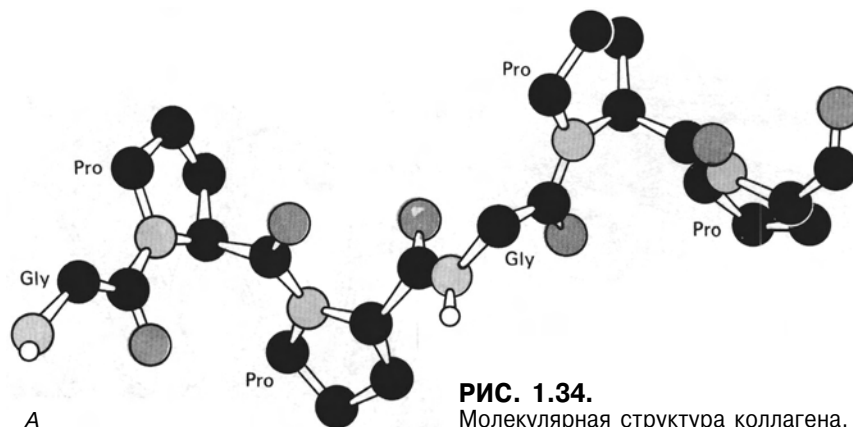
Структура  $\beta$ -слоя.  $\beta$ -Слой образуется из нескольких полипептидных цепей. Его стабилизируют водородные связи между атомами водорода амидных групп (NH) одной цепи и атомами кислорода карбонильных групп (CO) параллельной цепи. Обратите внимание, что вместо внутрицепочечных водородных связей, характерных

для  $\alpha$ -спирали, в  $\beta$ -слое образуются межцепочечные водородные связи. В представленном здесь примере полипептидные цепи параллельны, однако для некоторых  $\beta$ -слоев характерно антипараллельное расположение цепей.

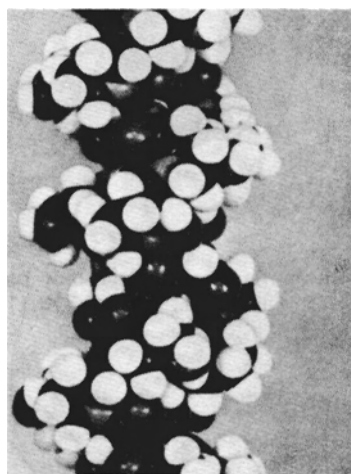
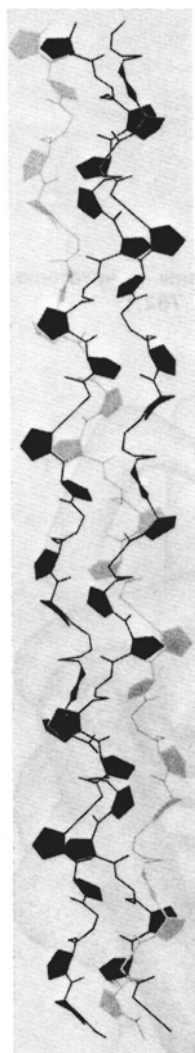
ная четвертичная структура белка поддерживается с помощью разнообразных взаимодействий между аминокислотами. К ним относятся: 1) взаимодействия, которые ответственны за формирование  $\alpha$ -спиралей и  $\beta$ -слоев; 2) водородные связи между некоторыми боковыми группами аминокислот; 3) ионные связи между противоположно заряженными боковыми группами; 4) ковалентные дисульфидные связи между отдаленными друг от друга вдоль цепи остатками цистеина; 5) гидрофобные взаимодействия между боковыми группами, более прочные, чем взаимодействия с водной фазой на наружной поверхности белковой глобулы. Относительный вклад всех этих сил в стабилизацию структуры у разных белков варьирует.

При нагревании либо повышении или понижении pH нативная глобула разворачивается (происходит **денатурация**). Этот процесс обратим, т.е. третичная структура может восстанавливаться с образованием тех химических и физических взаимодействий, которые стабилизируют нативную компактную конформацию (**ренатурация**). Механизм правильного свертывания полипептидной цепи и промежуточные этапы этого процесса интенсивно изучаются.

Большинство глобулярных белков—**олигомеры** (т.е. состоят из двух или более идентичных или разных полипептидных цепей). Примером глобулярных белков могут служить гемоглобин (рис. 1.36) и иммуноглобулин (рис. 1.37). **Четвертичная структура** подобных белков определяется тем, как взаимодействуют между собой в олигомерной структуре отдельные свернутые полипептидные цепи. В гемоглобине, например, благодаря взаимодействию аминокислот, входящих в состав  $\alpha$ -спиралей и  $\beta$ -слоев, образуется  $\alpha\beta$ -димер, а димеры в свою

**РИС. 1.34.**

Молекулярная структура коллагена. А. Основной повторяющийся структурный элемент, состоящий из остатков глицина и двух остатков пролина. Б. Скелетная модель трехцепочечной коллагеновой спирали. Каждая цепочка состоит из нескольких повторяющихся структурных элементов, указанных на рис. А. [L. Stryer, *Biochemistry* (San Francisco: W.H. Freeman, 1981).] В. Пространственная модель коллагена II. Каждая цепочка состоит из повторяющихся элементов глицин-пролин-гидрокси-пролин. Шаг спирали равен 2,86 нм. [A. Rich, F.H.C. Crick, *J. Mol. Biol.*, **3** (1961), p. 483.]



В

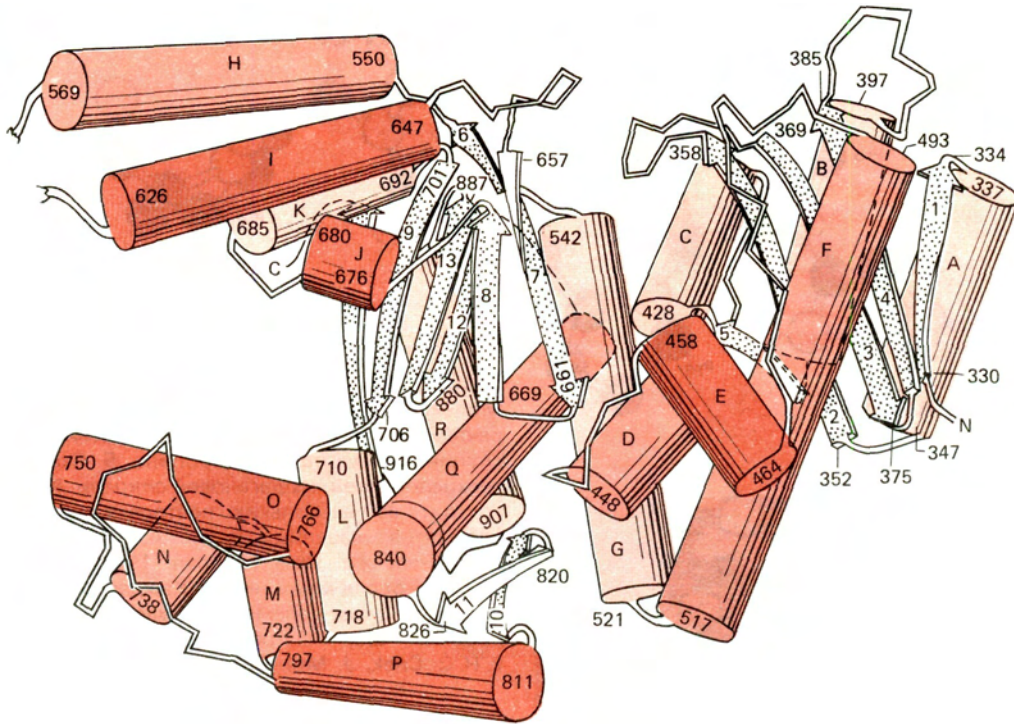
Б

очередь с помощью дополнительных взаимодействий объединяются в  $\alpha_2\beta_2$ -тетрамер (рис. 1.36). Детальные исследования структуры гемоглобина позволили установить, как изменяются соответствующие взаимодействия при связывании этим белком кислорода.

Вторичная, третичная и четвертичная структуры белков тесно связаны между собой и в конечном счете определяются первичной структурой одной или нескольких полипептидных цепей. Последствия такой взаимосвязи очень значительны: информация, определяющая укладку белковой молекулы и переход ее в биологически активное состояние, закодирована в его аминокислотной последовательности. Подтверждением этого принципиального положения служит то, что химические модификации и мутационные изменения аминокислотной последовательности полипептидов сильно влияют на их ренатурацию и способность формировать вторичную, третичную и четвертичную структуры с полноценной биологической активностью.

В качестве одного из многочисленных примеров зависимости структуры и функции белка от его аминокислотной последовательности можно привести серповидноклеточную анемию. Генетическое нарушение при этой болезни выражается в замене глутаминовой кислоты, шестой по счету от N-конца  $\beta$ -цепи в нормальном гемоглобине, на валин. Изменение в первичной структуре  $\beta$ -глобина приводит

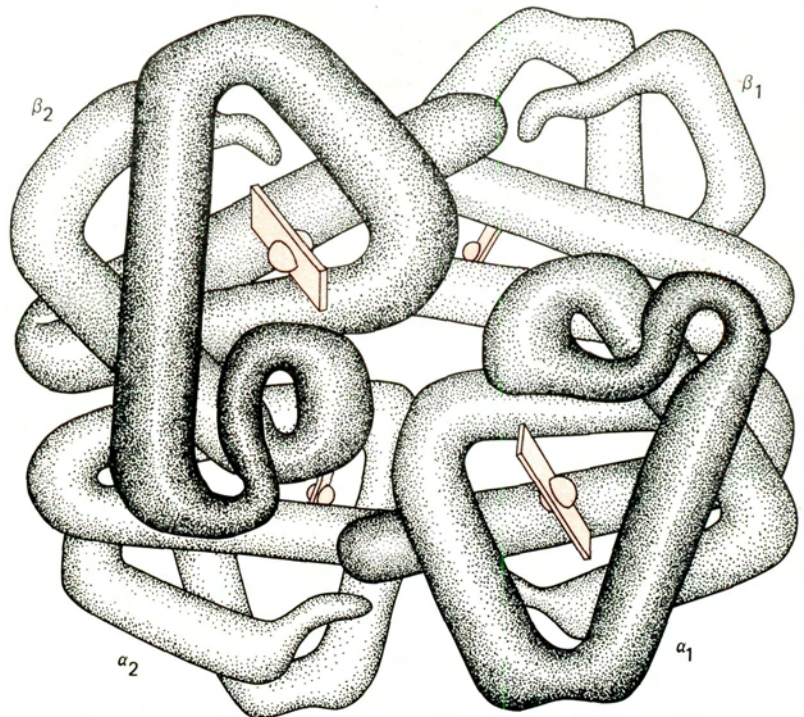




**РИС. 1.35.**

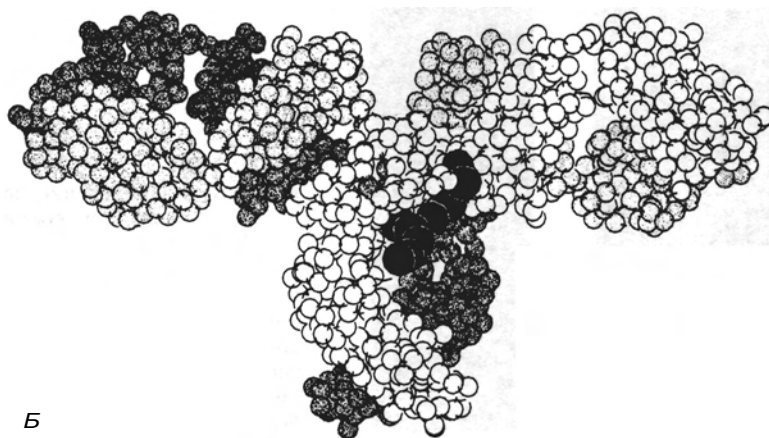
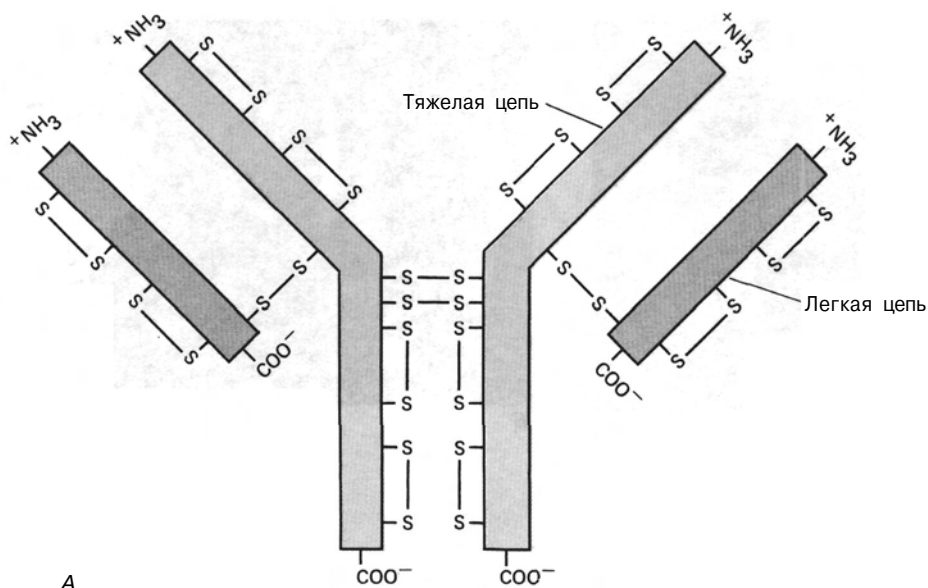
Схематическое изображение участка ДНК-полимеразы I *E. coli*.  $\alpha$ -Спирали изображены в виде цилиндров,  $\beta$ -слои - в виде стрелок. Чем интенсивнее окрашены эти

элементы, тем ближе они расположены к читателю. [D.L. Ollis et al., *Nature*, **313** (1985), p. 762.]



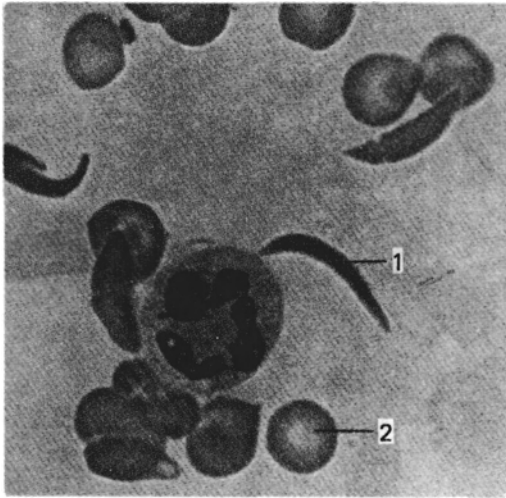
**РИС. 1.36.**

Модель гемоглобина, построенная при низком разрешении. Гемоглобин - это тетрамерный белок, состоящий из двух  $\alpha$ -цепей ( $\alpha_1$  и  $\alpha_2$ ) и двух  $\beta$ -цепей ( $\beta_1$  и  $\beta_2$ ). С каждой из цепей связана гемогруппа (выделена цветом; прямоугольник - гем, шарик - атом железа). (Copyright Irving Geis.)

**РИС. 1.37.**

Структура иммуноглобулина. А. Иммуноглобулин состоит из четырех полипептидных цепей: двух идентичных легких и двух идентичных тяжелых. Структура стабилизируется с помощью дисульфидных связей, образующихся между двумя тяжелыми цепями и тяжелыми и легкими. В самих тяжелых и легких цепях также

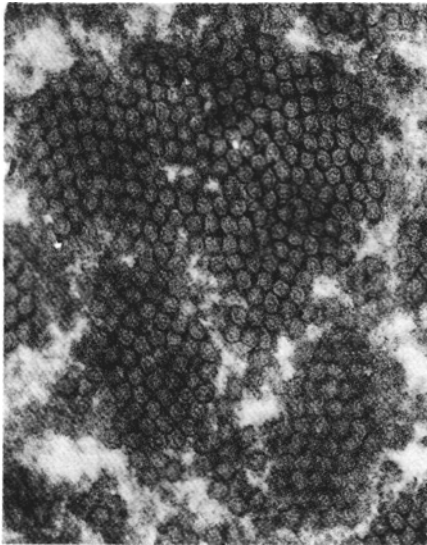
образуются дисульфидные связи. Б. Модель молекулы иммуноглобулина, иллюстрирующая переплетение легких и тяжелых полипептидных цепей. [E.W. Silvertown, M.A. Navia, D.R. Davies, Proc. Natl. Acad. Sci., USA, **74** (1977), p. 5140.]



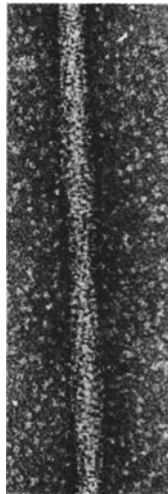
А



Б



В



Г

**РИС. 1.38.**

Дефекты, возникающие на клеточном и молекулярном уровнях при серповидноклеточной анемии. А. Микрофотография, на которой видны нормальный (2) и серповидный (1) эритроциты. (С любезного разрешения S. Schreier.) Б. Гемоглибиновые волокна в разрушенном серповидном эритроците. В. Поперечный разрез через гемоглибиновые волокна полученные при оксигенации серповидного гемоглобина. (Фото Б и В любезно предоставлены S.J. Edelstein.) Г. Электронная микрофотография неокрашенного тяжа серповидного гемоглобина. [G. Dykes, R.H. Crepeau, S.J. Edelstein, *Nature*, **272** (1978), p. 509.]

к тому, что на поверхности белковой глобулы оказывается аномальная гидрофобная аминокислота, из-за чего происходит агрегация дезоксигенированного гемоглобина с образованием олигомерных структур более высокого уровня организации (рис. 1.38). В результате форма и пластичность эритроцитов изменяются, и кровотоку через капилляры и мелкие вены затрудняется или вовсе прекращается. Основной вывод, который можно сделать

из этого классического случая, состоит в том, что одна-единственная мутация—замена нуклеотида в последовательности ДНК, приводящая к замене одной аминокислоты на другую в специфическом сайте полипептидной цепи,— может оказывать столь драматическое влияние на конформацию белка и его физиологическую функцию. В более общем смысле, таков механизм, связывающий генотипы всех организмов с их фенотипами.

# РЕПЛИКАЦИЯ, СОХРАНЕНИЕ И МОДИФИКАЦИЯ ГЕНОМА

Генетическая программа клеточных организмов записана в нуклеотидной последовательности ДНК. Следовательно, для сохранения уникальных свойств организма необходимо точное воспроизведение этой последовательности в каждом последующем поколении. *E. coli*, например, должна дублировать практически без ошибок полный геном размером  $4 \cdot 10^6$  нуклеотидных пар при образовании каждого последующего поколения; точно так же должны быть скопированы почти  $4 \cdot 10^9$  пар оснований в 23 парах хромосом человека при каждом акте деления клеток.

Одной из чудесных особенностей ДНК является то, что в ней закодирована информация о механизме ее собственного удвоения: одни гены кодируют ферменты, синтезирующие нуклеотидные предшественники ДНК, другие – белки, осуществляющие сборку активированных нуклеотидов в полинуклеотидные цепочки. Есть гены, координирующие процесс репликации с другими клеточными событиями, а также гены, кодирующие белки, которые упаковывают ДНК в хроматин. Еще одним необычным свойством ДНК является то, что она служит матрицей и определяет порядок, в котором нуклеотиды выстраиваются в новые нуклеотидные цепочки. Обладая одинаковым аппаратом синтеза, различные ДНК осуществляют образование только подобных себе реплик.

В генетической программе предусмотрены ферментативный механизм, который исправляет ошибки, иногда происходящие при репликации ДНК, и механизм репарации повреждений, затрагивающих основания или спиральную структуру при облучении рентгеновскими лучами и ультрафиолетовым светом или при воздействии различных химических агентов, а также механизм устранения дефектов, связанных с некоторыми заболеваниями. Генетическая программа обеспечивает создание геномных вариантов и возможность эволюционных изменений. Определенные гены кодируют белки, способствующие обмену цепями, принадлежащими разным молекулам ДНК, и тем самым созданию новых комбинаций генетического материала, передаваемых потомству. Известны белки, вызывающие геномные перестройки путем катализа транслокаций небольших сегментов или даже протяженных участков в пределах одной молекулы ДНК и между молекулами. С одной стороны, ре-

комбинации и транслокации подобного рода создают основу для эволюционных экспериментов, а с другой – некоторые перестройки могут быть причиной заболеваний. Как это ни странно, оптимальное функционирование некоторых генетических программ действительно зависит от специфических перестроек в ДНК (гл. 10).

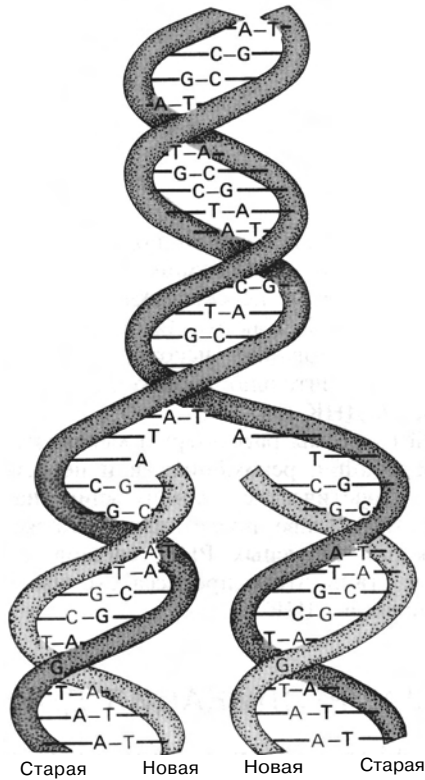
В этой главе мы рассмотрим механизмы репликации, репарации, рекомбинации и перестроек – по существу, генетические манипуляции на уровне ДНК. В этой главе представлены также данные о репликации вирусных РНК-геномов, поскольку некоторые этапы этого процесса сходны с процессом репликации ДНК.

## 2.1. РЕПЛИКАЦИЯ ДНК

### а. Матричная функция ДНК при репликации

Обнаружив свою модель структуры ДНК в 1953 г., Джеймс Уотсон и Фрэнсис Крик писали: «Мы не могли не осознавать, что специфическое спаривание (оснований), постулированное нами, подразумевает наличие какого-то механизма копирования генетического материала». И через месяц они углубили эту мысль: «Если известен точный порядок оснований в одной из цепей, то можно записать и порядок оснований в другой, поскольку спаривание оснований специфично. Таким образом, одна цепь является комплементом другой; именно это свойство наводит на мысль, что ДНК может удваивать саму себя».

Уотсон и Крик предположили, что для удвоения ДНК должны произойти разрыв водородных связей, удерживающих вместе спиральный дуплекс, и расхождение цепей. Они также высказали мысль, что каждая цепь дуплекса служит **матрицей** при синтезе комплементарной цепи и в результате образуются две пары цепей, в каждой из которых только одна является родительской (рис. 2.1). Таков механизм точного воспроизведения последовательности нуклеотидных пар в двойной спирали. Уотсон и Крик полагали, что репликация ДНК осуществляется спонтанно, без участия ферментов, но это оказалось неверно. Тем не менее идея о том, что удвоение ДНК происходит путем последовательного соеди-



**РИС. 2.1.**

Перед дубликацией ДНК должны произойти разрыв водородных связей, стабилизирующих двойную спираль, и расплетание цепей. Каждая цепь служит матрицей при синтезе другой, комплементарной цепи. [J.D. Watson, F.H.C. Crick, Cold Spring Harbor Symposium, **18** (1953), p. 123.]

нения нуклеотидов в соответствии с правилом комплементарности, заданным каждой цепью спирали, разрешила концептуальную проблему точного воспроизведения генов.

С того времени, как было высказано это предположение, матричная природа механизма репликации была подтверждена многочисленными данными, полученными как *in vivo*, так и *in vitro* для различных организмов. Согласно модели, репликация всех двухцепочечных ДНК **полуконсервативна** (рис. 2.2). Существуют ли в природе альтернативные способы репликации двухцепочечной ДНК (например, консервативный или дисперсный) — неизвестно. Итак, после одного раунда репликации одна цепь в каждой из двух дочерних молекул является родительской, т.е. консервативной, а другая — синтезированной заново. Если геном представлен одноцепочечной ДНК (как в некоторых вирусах), то эта единственная цепь служит матрицей для образования комплементарной цепи, с которой она об-

разуется дуплекс, а затем на этом дуплексе синтезируются либо дочерние дуплексы, либо одноцепочечные копии одной из матричных цепей (рис. 2.3).

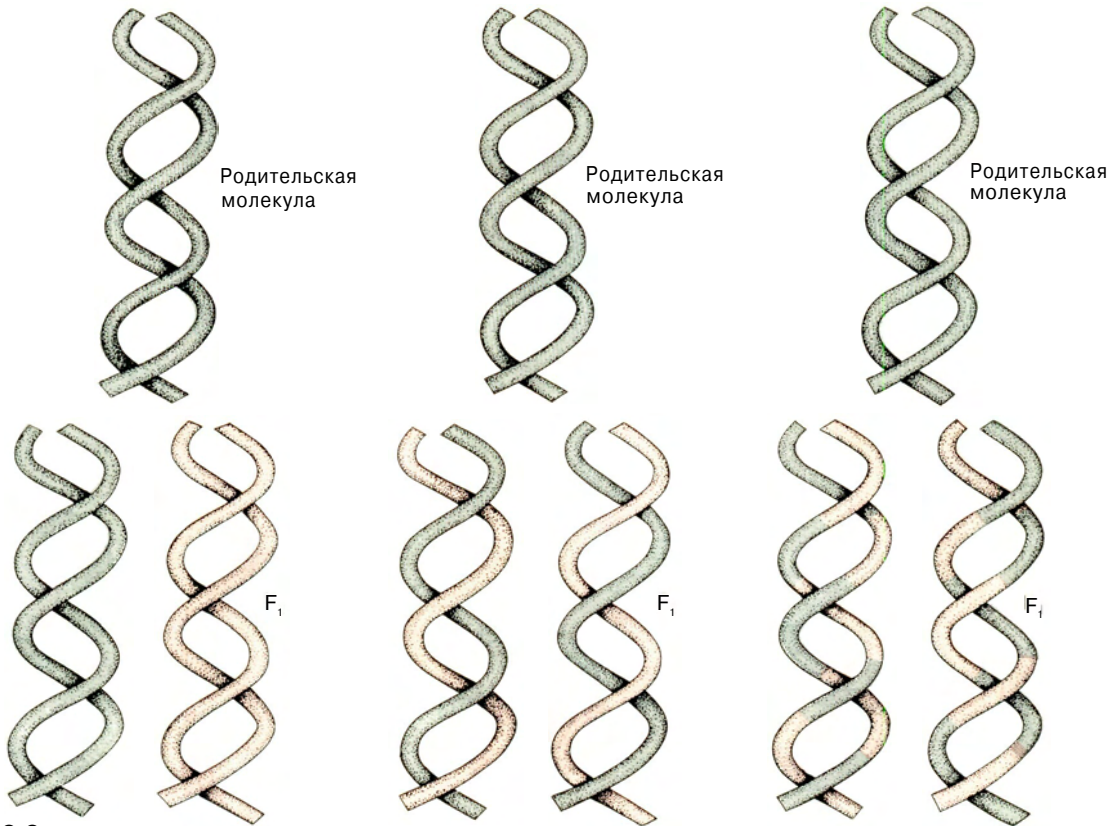
## б. Репликация начинается в определенных точках

*Начало репликации (origin).* Репликация ДНК начинается не в любой случайной точке молекулы, а в специфических местах, называемых **точками начала репликации**. Процесс копирования продолжается через образование **репликативных вилок** в одном или обоих направлениях до тех пор, пока ДНК полностью не удвоится (рис. 2.4). В замкнутых кольцевых дуплексах ДНК новосинтезированные цепи ковалентно соединяются в местах встречи увеличивающихся в размере репликативных вилок или в том месте, где единственная вилка возвращается к точке начала репликации. Дочерние молекулы, как правило, расходятся еще до начала нового раунда репликации. Такие различия относятся по размеру геномы, как геном вируса SV40 (5,2 т.п.н.), бактериофага  $\lambda$  (48,5 т.п.н.) и *E. coli* ( $4 \cdot 10^3$  т.п.н.), воспроизводятся в результате одного инициирующего события, происходящего в определенной точке.

У про- и эукариот можно встретить различные вариации на эту тему. Так, каждая из цепей родительской спирали митохондриальной ДНК животных (15 т.п.н.) и ДНК плазмиды Col E1 (6 т.п.н.) имеет свою точку начала репликации (рис. 2.5). Синтез комплементарной цепи некоторых небольших одноцепочечных фаговых геномов начинается вблизи одной специфической последовательности, а репликация полученного дуплекса может инициироваться совсем в другой точке (рис. 2.6). Репликация линейных дуплексных ДНК также инициируется в особых сайтах. Например, ДНК бактериофага T7 (40 т.п.н.) реплицируется в двух противоположных направлениях к разным концам молекулы, начиная от одной точки (рис. 2.7), а каждая из двух цепей ДНК аденовируса человека (30-38 т.п.н.) реплицируется последовательно всегда от 3'-конца (рис. 2.8).

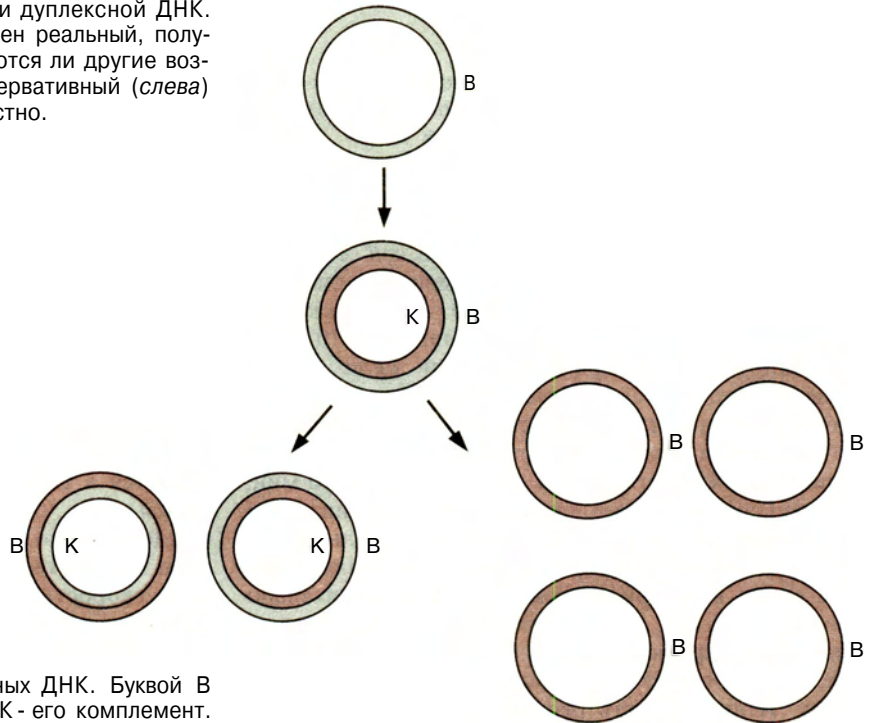
Для геномов эукариотических клеток (рис. 2.9) обычно характерно наличие множественных точек начала репликации, разбросанных по хромосоме на расстоянии 20 т.п.н. После инициации репликация продолжается в двух направлениях от каждой точки до тех пор, пока репликативные вилки двух соседних точек начала репликации не сольются. Полноразмерные ДНК каждой дочерней хромосомы получают путем соединения более коротких, независимо инициированных новосинтезированных цепей.

*Скорость репликации генома.* Скорость репликации генома регулируется в основном частотой



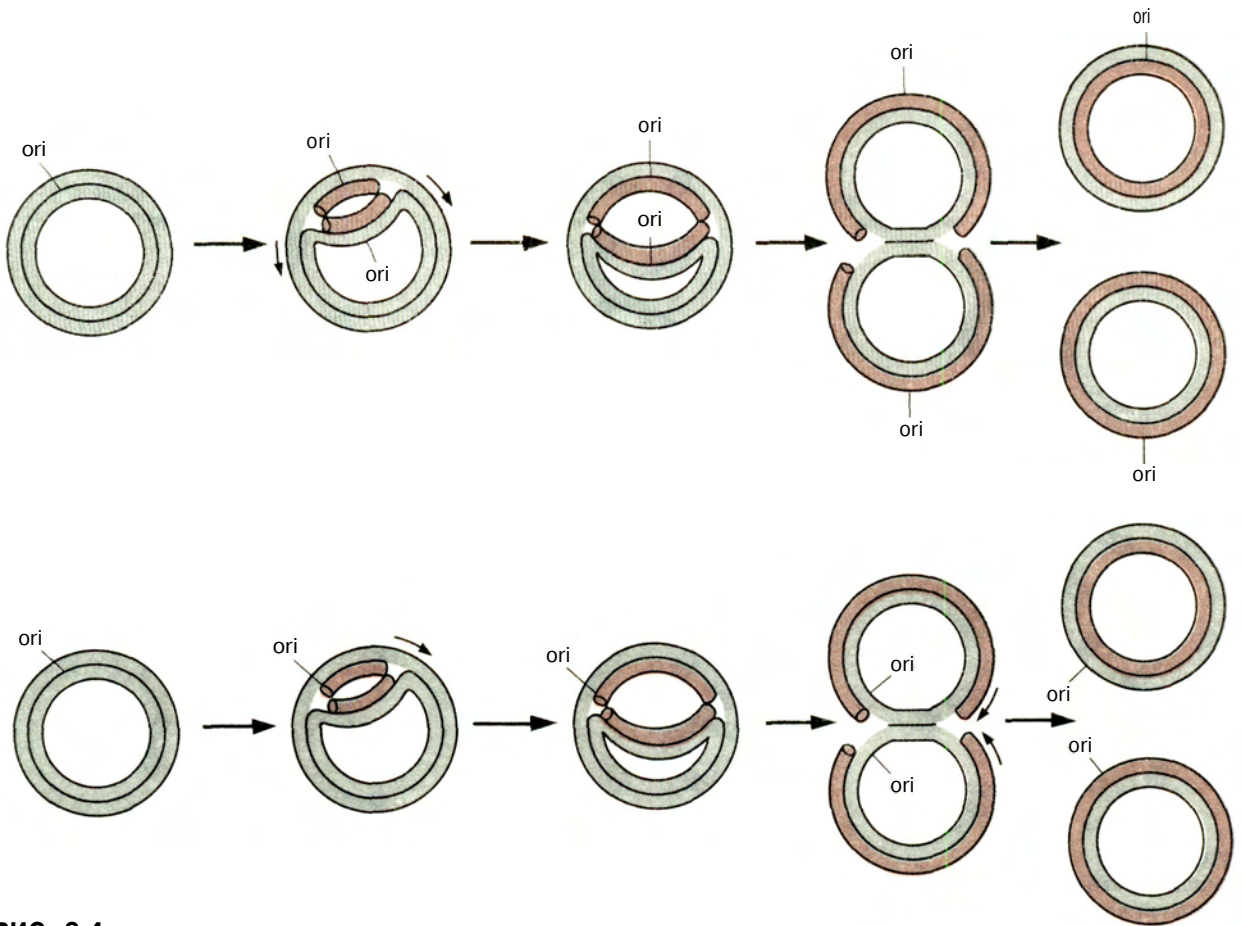
**РИС. 2.2.**

Предполагаемые модели репликации дуплексной ДНК. В средней части рисунка представлен реальный, полу-консервативный механизм. Реализуются ли другие возможные механизмы, а именно - консервативный (*слева*) и дисперсионный (*справа*), - неизвестно.



**РИС. 2.3.**

Репликация одноцепочечных вирусных ДНК. Буквой В обозначен вирусный геном, буквой К - его комплемент.

**РИС. 2.4.**

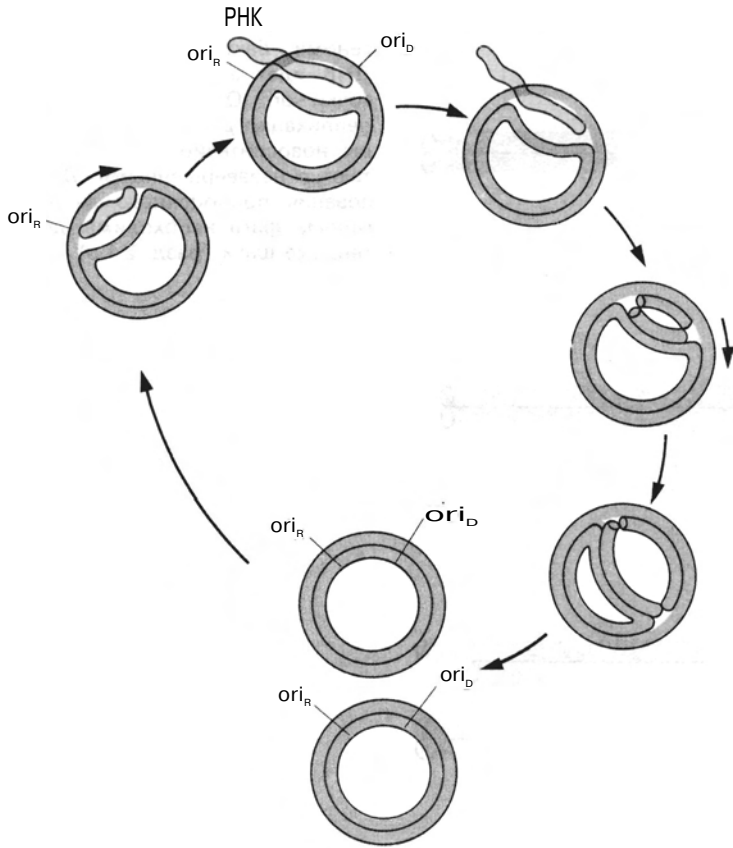
Репликация инициируется в специфическом участке ДНК, называемом точкой начала репликации (*ori*). Растущие цепи образуют репликативные вилки. Синтез новых дуплексов ДНК происходит в репликативных

вилках, перемещающихся либо в двух (*вверху*), либо в одном (*внизу*) направлении в зависимости от природы точки начала репликации.

инициирующих событий. Так, у *E. coli* скорость копирования в каждой репликативной вилке постоянна и равна примерно 1500 п.н. в секунду; следовательно, полный геном длиной  $4 \cdot 10^6$  пар реплицируется примерно за 40 мин. Если хромосома реплицируется быстрее, это значит, что увеличивается частота актов инициации в той же самой точке начала репликации при прежней скорости копирования. Клетки *E. coli* делятся каждые 20 мин; это означает, что репликация ДНК инициируется в хромосомах, еще не закончивших предыдущий раунд репликации. Скорость движения репликативной вилки в эукариотических клетках значительно меньше (10–100 п.н. в секунду), но завершение репликации хромосомы в разумное время обеспечивается одновременной инициацией во множестве точек. Итак, скорость репликации хромосом конт-

ролируется числом и расположением точек начала репликации. Например, в ранних эмбрионах дрозофилы репликация хромосомы осуществляется каждые 3 мин благодаря почти одновременной инициации событий в точках, отстоящих друг от друга на 7000–8000 п.н. В культуре клеток *Drosophila* наблюдается значительно более медленная скорость дупликации хромосомы, поскольку репликация начинается в гораздо меньшем числе точек, находящихся друг от друга на расстоянии 40000 п.н. Следовательно, при фиксированной скорости роста цепи множественная инициация обеспечивает большую скорость процесса и уменьшает время, необходимое для дупликации протяженных участков хромосом.

*Структура точек начала репликации.* Фрагменты ДНК, несущие точку начала репликации, выделены

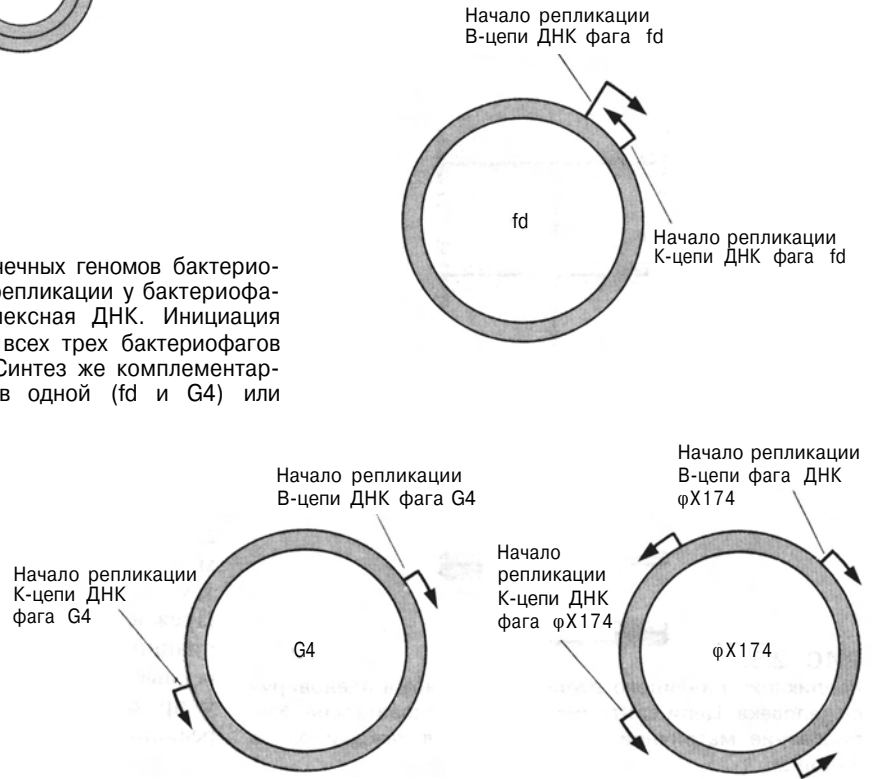


**РИС. 2.5.**

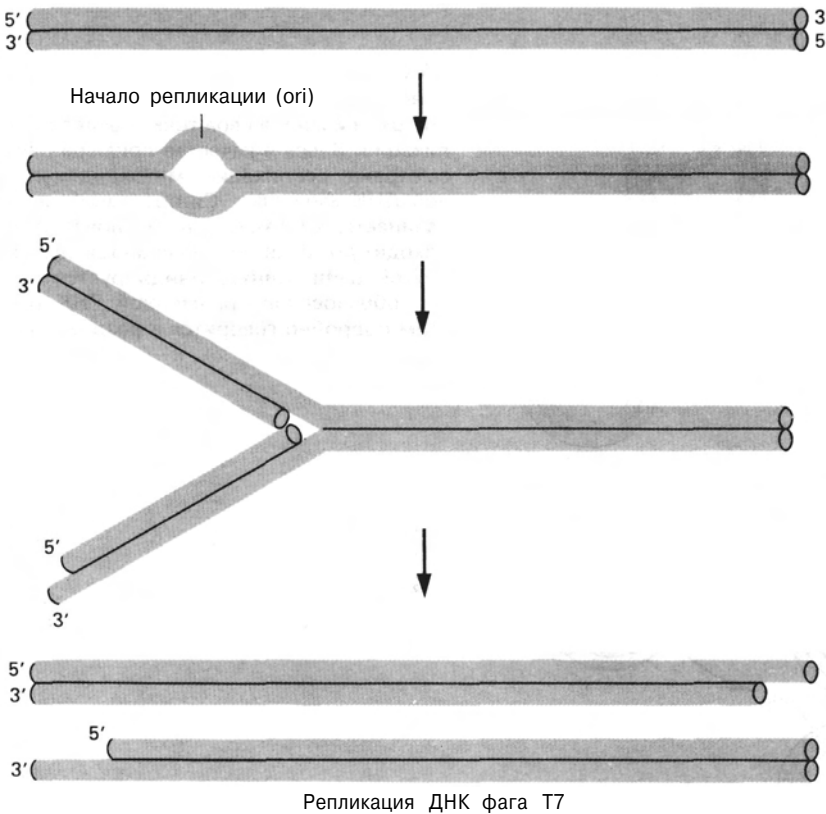
В некоторых кольцевых геномах в каждой цепи имеется своя точка начала репликации. Здесь в качестве примера рассмотрена репликация митохондриальной ДНК животных. Синтез одной цепи начинается в точке  $ori_R$ . Когда новая цепь доходит до точки  $ori_D$ , начинается синтез другой цепи. Синтез инициируется путем образования праймерной РНК (об этом подробно говорится в разд. 2.1.ж).

**РИС. 2.6.**

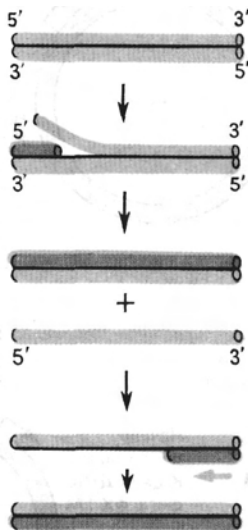
Точки начала репликации одноцепочечных геномов бактериофагов. Промежуточным продуктом репликации у бактериофагов fd, G4 и φX174 является дуплексная ДНК. Инициация синтеза геномной ДНК (В-цепь) у всех трех бактериофагов происходит в единственной точке. Синтез же комплементарной цепи (К-цепь) инициируется в одной (fd и G4) или нескольких (φX174) точках.





**РИС. 2.7.**

Репликация линейного дуплексного генома бактериофага T7. Синтез ДНК идет в двух направлениях от точки ori. Особенности механизма репликации ДНК таковы, что 5'-концы новосинтезированных цепей остаются незавершенными. Для образования полноценного дуплексного генома фага необходимы дополнительные шаги (разд. 2.1.3).

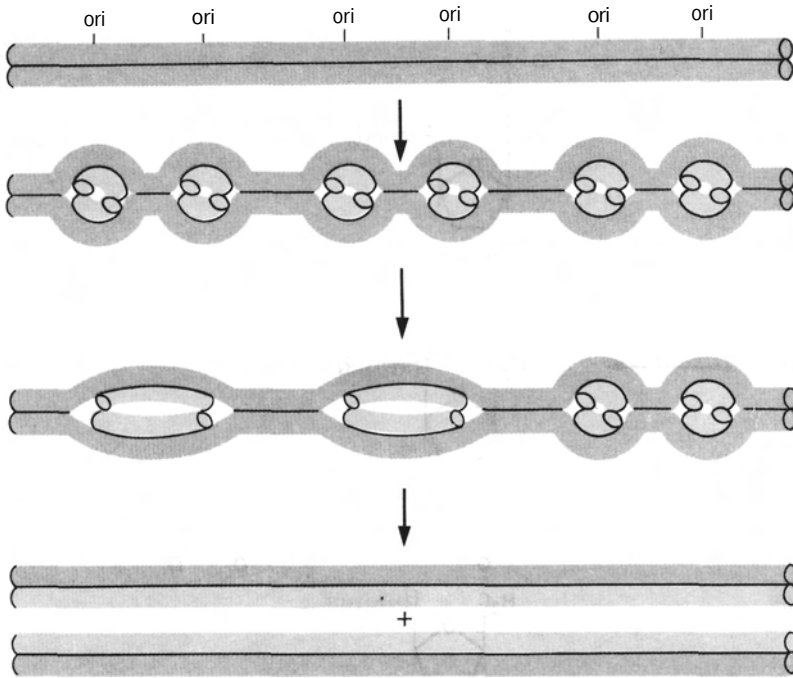
**РИС. 2.8.**

Репликация линейного дуплексного генома аденовируса человека. Цепи синтезируются последовательно. Копирование матричной цепи начинается каждый раз с 3'-конца.

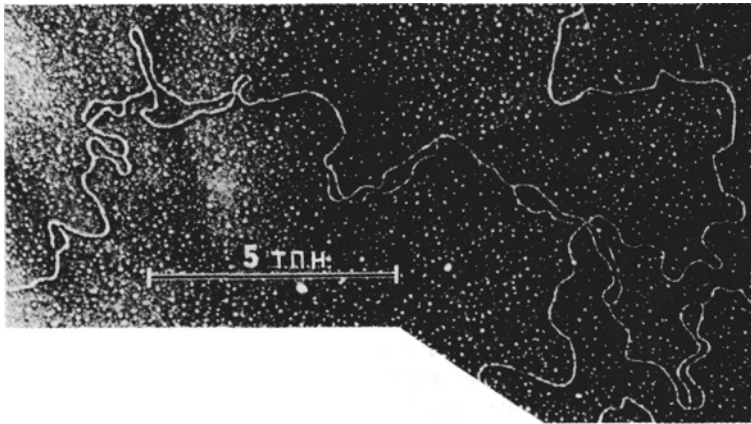
из *E. coli* и некоторых колифагов и плазмид, а также из дрожжей и ряда эукариотических вирусов. В некоторых случаях место начала репликации имеет такую нуклеотидную последовательность, что дуплекс принимает необычную конфигурацию, которую распознают белки, участвующие в инициации. Природа взаимодействия между точкой начала репликации и белками и механизм инициации в целом исследованы очень мало (разд. 2.1.ж), однако можно сказать, что, по-видимому, они в разных случаях различны.

## в. Репликация ДНК полуконсервативна

После начала репликации репликативные вилки движутся в одном или обоих направлениях вдоль молекулы ДНК. Чем дальше продвинулась вилка, тем длиннее новосинтезированный сегмент ДНК. Прежде чем обсуждать данные о механизмах инициации репликативной вилки и ее движения, полученные с помощью модельных систем репликации у прокариот (разд. 2.1.ж), рассмотрим основную реакцию, протекающую в репликативной вилке в процессе копирования двухцепочечной ДНК.

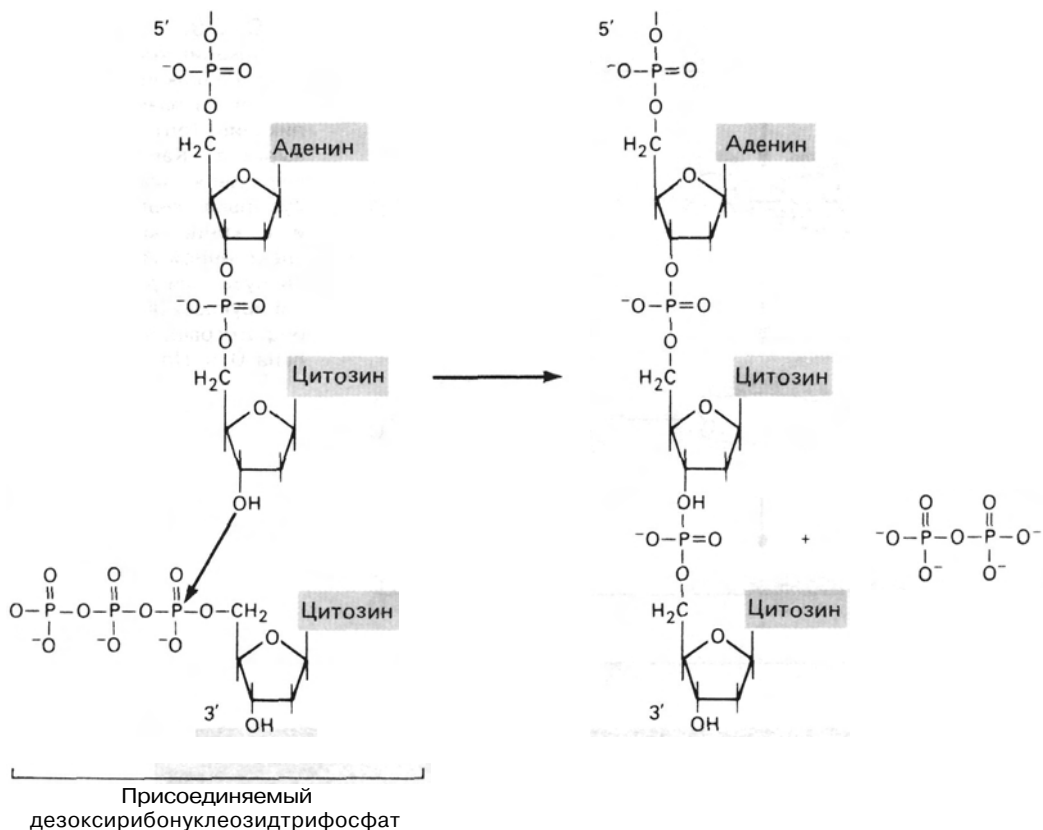
**РИС. 2.9.**

Репликация хромосомной ДНК эукариот. Репликация идет в двух направлениях из разных точек начала репликации (*ori*) с образованием «пузырьков». Как видно из рисунка, по мере движения репликативных вилок пузырьки увеличиваются в размере и в конце концов сливаются. На электронной микрофотографии видны пузырьки разных размеров в реплицирующейся ДНК *Drosophila*. (Микрофотография любезно предоставлена D.S. Hogness.)



Цепи ДНК синтезируются в результате присоединения 5'-дезоксинуклеотидильных единиц дезоксирибонуклеозидтрифосфатов к 3'-гидроксильному концу уже имеющейся цепи (**праймера**) (рис. 2.10). Следовательно, они синтезируются в направлении 5'→3' вдоль матричной цепи, ориентированной в противоположном, 3'→5', направлении (рис. 2.11). Синтез цепей в обратном направлении не происходит никогда, поэтому синтезируемые цепи в каждой репликативной вилке должны расти в противоположных направлениях (рис. 2.12). Синтез одной цепи (**ведущей, лидирующей**) происходит непрерывно, а другой (**отстающей**) — импульсами. Такой механизм репликации называется полунепрерывным. Ве-

дущая цепь растет от 5'- к 3'-концу в направлении движения репликативной вилки и нуждается только в одном акте инициации. Рост отстающей цепи также идет от 5'- к 3'-концу, но в направлении, противоположном движению репликативной вилки. Для синтеза отстающей цепи должно произойти несколько актов инициации, в результате чего образуется множество коротких цепей. Эти цепи, называемые **фрагментами Оказаки** в честь открывшего их ученого, имеют размеры 1000–2000 нуклеотидов у прокариот и 100–200 в реплицирующейся ДНК эукариот. По мере движения репликативной вилки концы соседних фрагментов Оказаки соединяются с образованием непрерывной отстающей цепи.

**РИС. 2.10.**

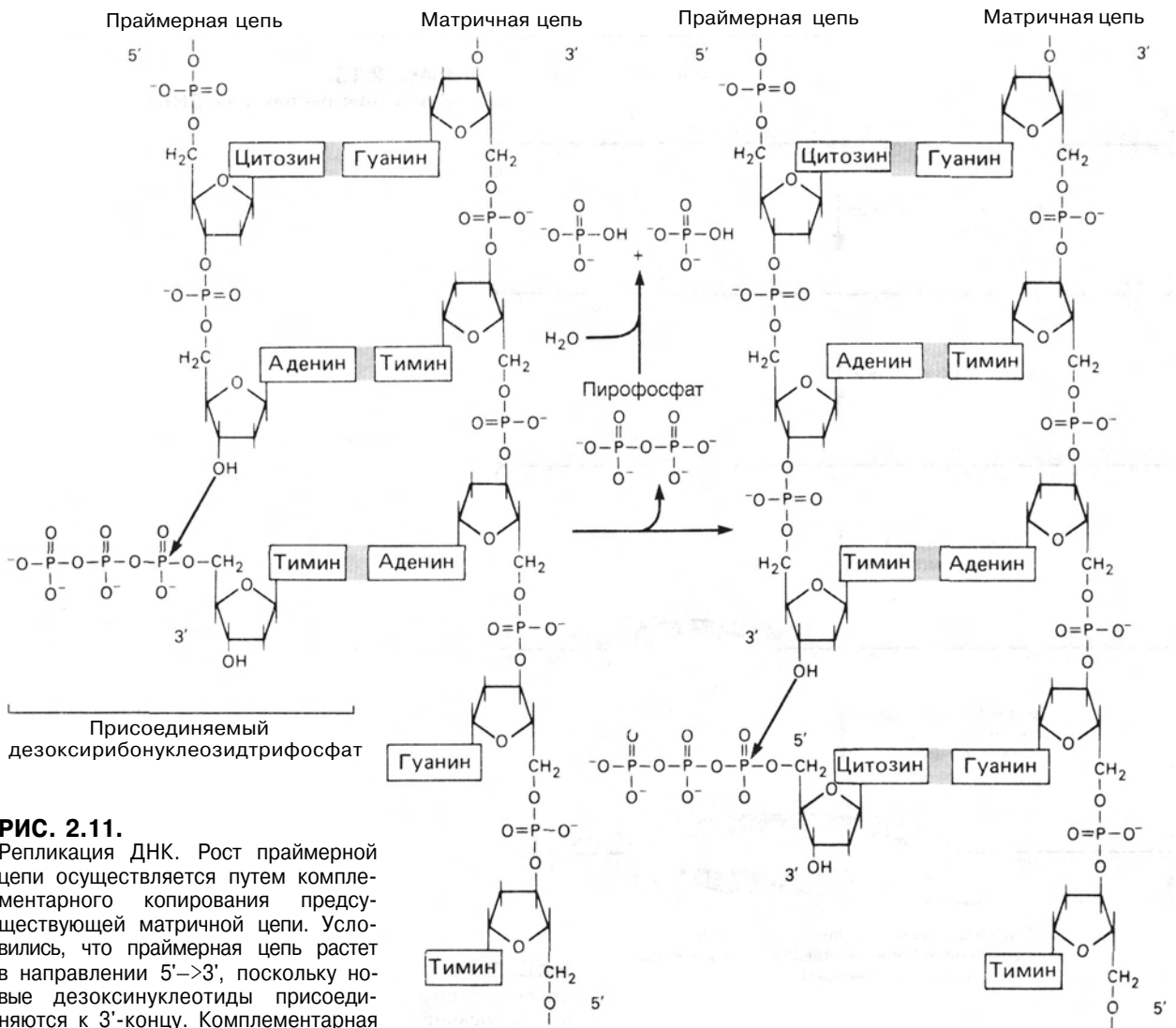
Цепи ДНК удлиняются путем последовательного присоединения дезоксинуклеотидов к концевой 3'-ОН-группе каждого предыдущего нуклеотида. Донором новых дезоксинуклеотидных единиц служат 5'-дезоксирибонуклеозид-

трифосфаты. Присоединение каждого нуклеотида сопровождается высвобождением одной молекулы пирофосфата.

Механизмы инициации репликации в точке начала репликации и при образовании фрагментов Оказаки в отстающей цепи в принципе аналогичны, хотя имеются некоторые тонкие различия (разд. 2.1.ж). В обоих случаях происходит образование коротких РНК-затравок (**праймеров**), комплементарных матричной ДНК, в виде продолжения которых синтезируется новая цепь ДНК (рис. 2.13). В дальнейшем короткие вставки РНК замещаются сегментами ДНК, которые затем объединяются с образованием непрерывной отстающей цепи.

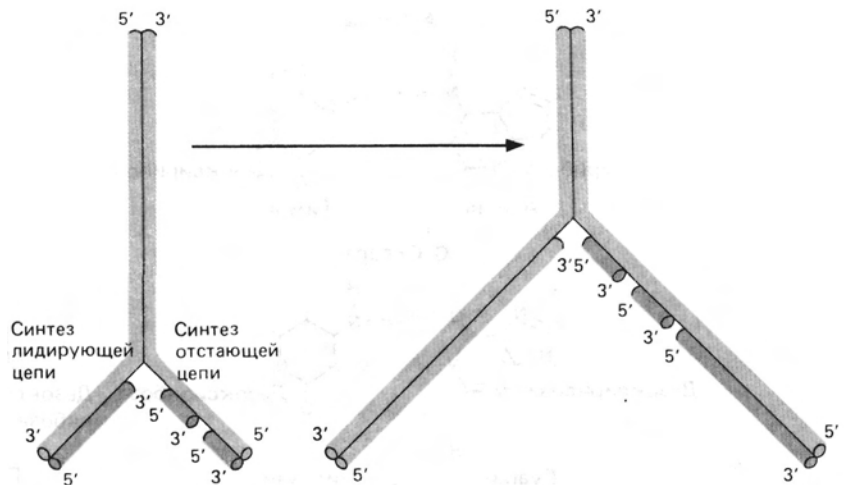
#### г. Комплементарное копирование оснований, перенос дезоксинуклеотидов и лигирование ДНК при репликации

Нуклеотидная последовательность матричной цепи задает порядок расположения нуклеотидов в новосинтезируемых цепях ДНК (рис. 2.11). Несмотря на то что правило комплементарности обеспечивает большую надежность процесса копирования, последний небезошибочен. Из-за случайных флуктуации в структуре присоединяемых оснований и оснований, входящих в состав матричной цепи, могут происходить ошибочное спаривание и включение неправильных нуклеотидов (рис. 2.14). При включении неправильного нуклеотида дальнейший рост цепи обычно останавливается. Ферменты, катализирующие присоединение нуклеоти-



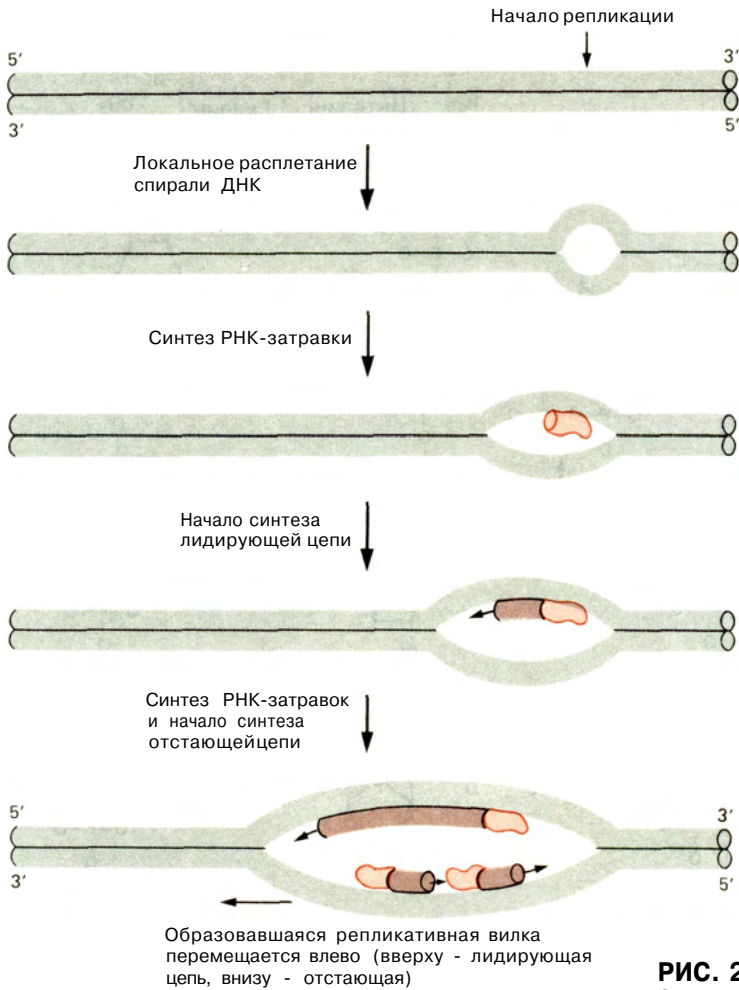
**РИС. 2.11.**

Репликация ДНК. Рост праимерной цепи осуществляется путем комплементарного копирования предсуществующей матричной цепи. Условились, что праимерная цепь растет в направлении 5'→3', поскольку новые дезоксирибонуклеотиды присоединяются к 3'-концу. Комплементарная матричная цепь имеет направление 3'→5'. Необходимую для протекания этой реакции энергию система получает за счет гидролиза пирофосфата.

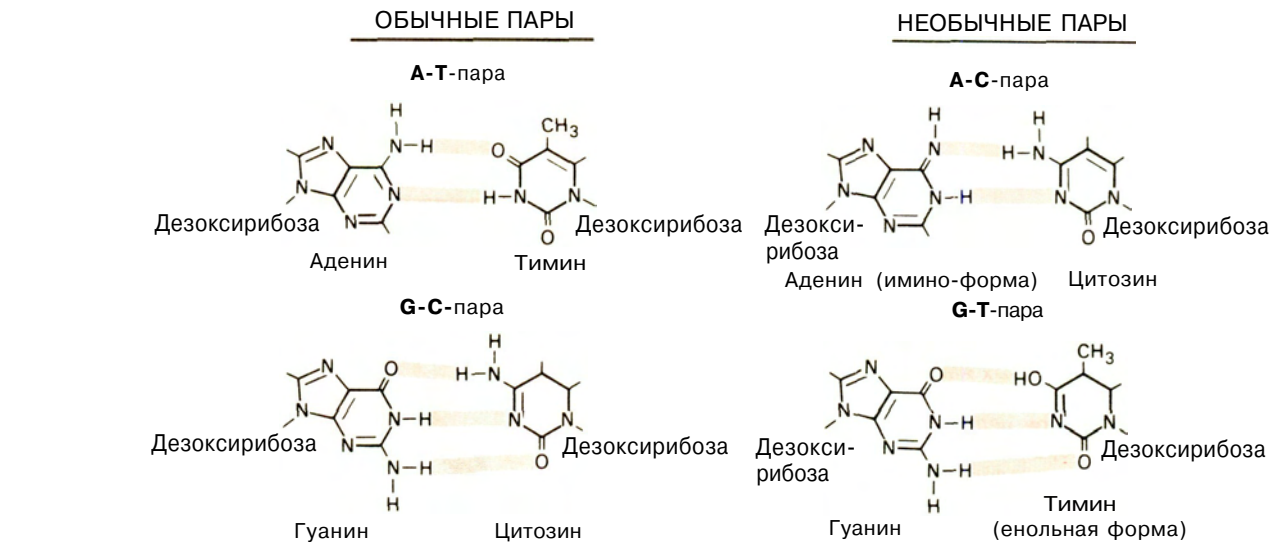


**РИС. 2.12.**

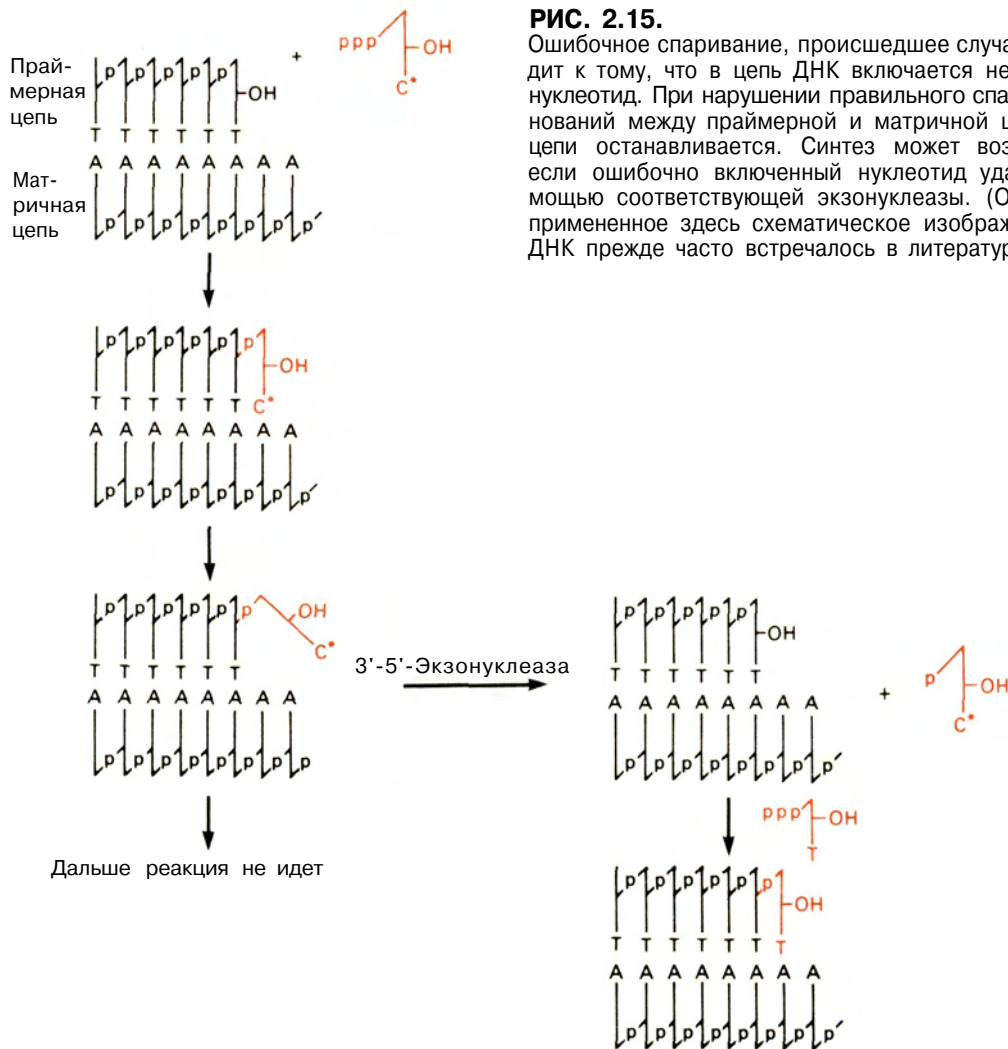
Синтез обеих цепей ДНК в репликативной вилке идет в направлении 5'→3', т.е. удлинение цепей происходит в противоположных направлениях. Одна цепь - лидирующая - синтезируется как единое целое, а другая - отстающая - короткими фрагментами, которые впоследствии соединяются с образованием непрерывной цепи.



**РИС. 2.13.**  
Инициация репликации ДНК.

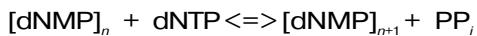


**РИС. 2.14.**  
Сравнение обычных и необычных, редко встречающихся пар оснований.

**РИС. 2.15.**

Ошибочное спаривание, происшедшее случайно, приводит к тому, что в цепь ДНК включается неправильный нуклеотид. При нарушении правильного спаривания оснований между праймерной и матричной цепями рост цепи останавливается. Синтез может возобновиться, если ошибочно включенный нуклеотид удалить с помощью соответствующей экзонуклеазы. (Отметим, что примененное здесь схематическое изображение цепей ДНК прежде часто встречалось в литературе.)

дов, обладают эффективной системой коррекции — они удаляют включенный, но неспаренный нуклеотид почти сразу после присоединения его к концу растущей цепи (рис. 2.15). В общем виде реакцию присоединения 5'-дезоксинуклеотидной группы к 3'-ОН-группе конечного нуклеотида праймерной цепи можно представить следующим образом:



где dNMP — любой из четырех обычных нуклеотидов. За один акт репликации праймерная цепь удлиняется на один остаток, при этом одновременно происходит удаление пирофосфата. Реакция присоединения нуклеотида обратима, но в целом она смещена в сторону синтеза, поскольку неорганический фосфат в клетках быстро разрушается. Ферменты, катализирующие праймер-зависимую, де-

терминируемую ДНК-матрицей реакцию присоединения дезоксинуклеотидов, называются **ДНК-полимеразами**. Многие ДНК-полимеразы выделены и охарактеризованы, и в разд. 2.1.д детально описаны их свойства и реакции, которые они катализируют.

Фрагменты Оказаки, образующиеся при импульсном копировании цепи ДНК, должны впоследствии объединиться в непрерывную отстающую цепь (рис. 2.12). Эта реакция осуществляется **ДНК-лигазами** — ферментами, катализирующими ковалентное сшивание цепей ДНК в дуплексе; ДНК-лигазы играют ключевую роль не только в репликации, но и в репарации повреждений (разд. 2.3) и рекомбинации ДНК (разд. 2.4).

## д. Ключевые ферменты, участвующие в синтезе ДНК

Многие известные теперь детали процесса репликации ДНК удалось установить благодаря исследованию поведения и активности ферментов, обеспечивающих работу аппарата репликации. Наиболее полно изучен механизм репликации бактериальной ДНК, особенно ДНК *E. coli* и бактериофагов, которые в ней размножаются. Довольно хорошо известны и ферменты репликации дрожжей, *Drosophila*, клеток и вирусов млекопитающих. Здесь мы обсудим механизм действия ДНК-полимераз и ДНК-лигаз, поскольку при синтезе длинных цепей ДНК эти два фермента работают согласованно.

**ДНК-полимеразы.** ДНК-полимеразы присутствуют во всех прокариотических и эукариотических клетках. Более того, многие вирусы бактерий и животных индуцируют образование вирус-специфических ДНК-полимераз или белков, способствующих эффективному участию ДНК-полимераз клеток-хозяев в репликации вирусной ДНК. Некоторые прокариотические и эукариотические ДНК-полимеразы выделены в чистом виде, а их физические и ферментативные свойства охарактеризованы. И хотя эти свойства не совсем идентичны, механизм катализа для всех указанных ферментов в общих чертах одинаков.

Наиболее полно изучена ДНК-полимераза I (Pol I) *E. coli*. Она представляет собой одиночный полипептид с мультифункциональными активностями. В качестве ДНК-полимеразы Pol I катализирует перенос 5'-дезоксинуклеотидильных единиц дезоксинуклеозид-5'-трифосфатов к 3'-ОН-группе в цепи ДНК или РНК, после чего происходит спаривание перенесенного основания с соответствующим основанием комплементарной цепи ДНК (разд. 2.1.г и рис. 2.11). Таким образом, для полимеризации ферменту необходимы праймер в качестве дезоксинуклеотидного акцептора и матрица, детерминирующая присоединение нужного нуклеотида. Помимо полимеризации нуклеотидов, Pol I катализирует две другие реакции, биологическая роль которых очень важна (рис. 2.16). В одной из них происходит гидролиз фосфодиэфирных связей в одной цепи ДНК или на неспаренном конце дуплексной ДНК, причем за один акт удаляется один нуклеотид, начиная с 3'-конца цепи (**3'-5'-эксонуклеаза**). Вторая реакция также состоит в отщеплении нуклеотидов, но гидролиз начинается с 5'-конца дуплексной ДНК в направлении к 3'-концу (**5'-3'-эксонуклеаза**). Эти различные активности присущи разным сайтам полипептидной цепи Pol I. Если *in vitro* обработать Pol I трипсином, то полипептидная цепь расщепится на большой и малый

фрагменты. Большой, С-концевой фрагмент («фрагмент Кленова») сохраняет ДНК-полимеразную и 3'-5'-эксонуклеазную активности; малый, N-концевой фрагмент обладает только 5'-3'-эксонуклеазной активностью.

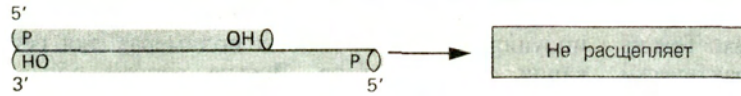
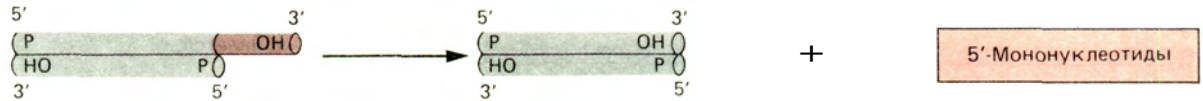
Pol I и присущие ей эксонуклеазные активности играют очень большую роль в репликации и репарации хромосомной ДНК *E. coli*. 3'-5'-эксонуклеазная активность обеспечивает контроль за присоединением каждого нуклеотида и удаление ошибочных нуклеотидов с растущего конца цепи (рис. 2.15). Если эта активность подавлена в результате каких-то мутаций в гене, кодирующем Pol I, то при репликации генома часто происходят мутации — замены оснований.

Способность ДНК-полимеразы удлинять 3'-конец цепи, спаренной с матричной цепью, позволяет ей заполнять пробелы между сегментами отстающей цепи. Pol I удлиняет фрагменты Оказаки с 3'-концов и удаляет рибонуклеотиды, с которых начинаются 5'-концы соседних фрагментов, что является необходимой предпосылкой для формирования непрерывной отстающей цепи (рис. 2.13). Поскольку Pol I способна удлинять 3'-конец одной из цепей в месте разрыва в двухцепочечной ДНК и удалять нуклеотиды с 5'-конца того же разрыва (процесс, называемый **ник-трансляцией**), этот фермент играет ключевую роль в репарации поврежденной ДНК (рис. 2.16 и разд. 2.3). Ник-трансляция широко используется *in vitro* для синтеза радиоактивно меченной ДНК (разд. 4.6.б).

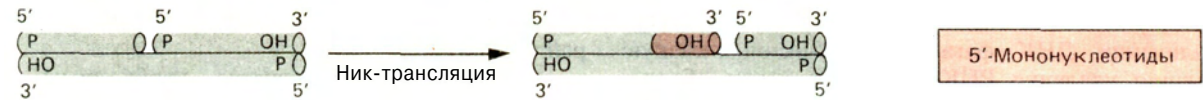
У *E. coli* имеются и две другие ДНК-полимеразы, но они присутствуют в клетке в меньших количествах. Pol II присоединяет нуклеотиды значительно менее эффективно, чем Pol I, и не обладает 5'-3'-эксонуклеазной активностью. Следовательно, Pol II может заполнять пробелы между фрагментами ДНК, спаренными с матричной цепью, но не способна отщеплять РНК-нуклеотиды от фрагментов Оказаки или осуществлять ник-трансляцию. Роль Pol II в репликации и сохранении хромосомной ДНК *E. coli* до настоящего момента неясна.

Pol III-холофермент — это ключевой фермент, ответственный за репликацию хромосомной ДНК *E. coli*. В каждой клетке содержится только 10–20 копий Pol III-холофермента, и тем не менее он является основным компонентом мультиферментного комплекса, инициирующего формирование репликативных вилок в точках начала репликации, участвующего в элонгации лидирующей цепи и вилке и удлиняющего РНК-праймеры с образованием фрагментов Оказаки. Но поскольку Pol III-холофермент не обладает 5'-3'-эксонуклеазной активностью, для репликации отстающей цепи необходимо участие Pol I, чтобы произошло удлинение продукта, образовавшегося при участии Pol III, и

## 3'→5'-Экзонуклеаза



## 5'→3'-Экзонуклеаза

**РИС. 2.16.**

ДНК-полимераза катализирует 3'→5'- и 5'→3'-экзонуклеазные реакции. 3'→5'-экзонуклеаза расщепляет одноцепочечную ДНК с 3'-ОН-конца, а 5'→3'-экзонуклеаза разрушает дуплексную ДНК с 5'-конца. 5'→3'-экзонуклеазная и полимеразная активности катализируют процесс, при котором происходят последовательное

отщепление нуклеотидов с 5'-конца одноцепочечного разрыва в дуплексе и удлинение цепи с 3'-конца. В результате место разрыва перемещается по цепи в направлении от 5'- к 3'-концу (так называемая ник-транслация). Для осуществления ник-транслации необходимы дезоксирибонуклеозидтрифосфаты.

удаление РНК-праймеров на 5'-конце фрагментов Оказаки.

Обнаружены изменения в полипептидной цепи основного (кор) фермента Pol III, известны аминокислотные замены, которые позволяют приписать определенные виды ферментативной активности конкретным субъединицам ферментного комплекса. Так, α-субъединица обладает полимеразной активностью, а ε-субъединица — 3'-5'-экзонуклеазной. Однако комплекс α- и ε-субъединиц обладает значительно более высокой полимеразной и экзонуклеазной активностями, чем каждая из соответствующих субъединиц в отдельности. Функция третьей, θ-субъединицы пока неясна.

Помимо субъединиц, составляющих Pol III-кор, Pol III-холофермент содержит еще семь субъединиц: τ, γ, β, δ, δ', χ и ψ. Перечисленные полипептиды также существуют во множестве копий, так что в результате мол. масса комплекса составляет примерно 10<sup>3</sup> кДа. Роль β-субъединицы заключается в том, чтобы свести к минимуму вероятность отделения фермента от матрицы до завершения процесса копирования; точная же функция других субъединиц неизвестна. Вполне возможно, что

Pol III-холофермент существует в двух формах, каждая из которых содержит определенный набор вспомогательных субъединиц, придающих ферменту определенные свойства. В одной форме фермент катализирует синтез непрерывной ведущей цепи, а в другой — прерывистой отстающей.

Pol III-холофермент катализирует те же реакции синтеза, что и Pol I, но работает примерно в 60 раз быстрее. Более того, Pol III-холофермент обладает повышенным сродством к матрице и обеспечивает более высокую эффективность копирования. Pol III-холофермент может связываться и с другими белками, увеличивая эффективность процесса копирования благодаря координации некоторых важных ферментативных этапов репликации. На этом более высоком уровне организации в комплексы могут включаться белки, расплетающие спираль ДНК в точках начала репликации и в репликативных вилках (расплетазы или геликазы), иницирующие образование праймерных РНК (праймазы), обеспечивающие последовательное наращивание цепей ДНК, терминирующие процесс репликации и разделяющие дочерние спирали ДНК.

ДНК-полимеразы, синтезируемые другими бак-



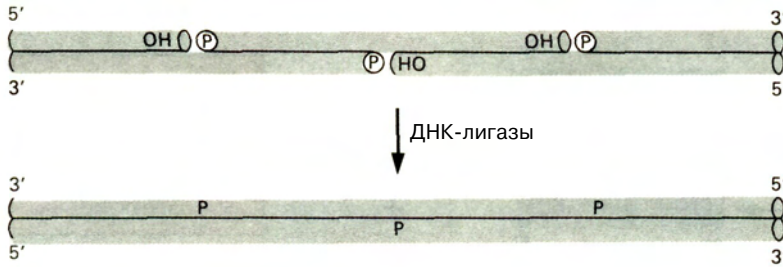


РИС. 2.17.

Все ДНК-лигазы соединяют 5'-фосфорильную и 3'-гидроксильную группы нуклеотидов, находящихся на противоположных концах одноцепочечного разрыва в дуплексе ДНК. Образуется новая фосфодиэфирная связь, разрыв сшивается.

териями и многими бактериофагами, различаются по своей физической структуре и свойствам. Тем не менее катализируемые ими реакции практически идентичны реакциям, изученным у *E. coli*. У всех ДНК-полимераз есть корректирующая 3'-5'-экзонуклеаза, однако 5'-3'-экзонуклеаза у некоторых ферментов отсутствует. Например, ДНК-полимераза фага Т4 может осуществлять 3'-5'-экзонуклеазную реакцию коррекции ошибок, но не способна катализировать 5'-3'-экзонуклеазную реакцию и поэтому не может обеспечить ник-трансляцию. При репликации ДНК фага Т4 5'-3'-экзонуклеазную реакцию удаления РНК-праймеров перед объединением фрагментов Оказки катализирует другой кодируемый фагом белок. В процессе прерывистого синтеза отстающих цепей и репарации повреждений ДНК фага Т4 этот фермент работает согласованно с фаговой ДНК-полимеразой.

В эукариотических клетках идентифицировано множество ДНК-полимераз, но их физические и функциональные свойства изучены менее детально, чем у соответствующих ферментов прокариот. Из клеток млекопитающих выделены четыре ДНК-полимеразы:  $\alpha$ ,  $\beta$  и  $\delta$  содержатся в ядрах, а  $\gamma$  — в митохондриях. ДНК-полимераза  $\alpha$  участвует в репликации хромосомной ДНК. Ее полимеразная активность связана с большим полипептидом, но она существует и, возможно, функционирует как мультисубъединичный белок, аналогично Pol III-холоферменту *E. coli*.  $\beta$ -Полимераза — это одиночный полипептид, функцией которого является заполнение пробелов при репарации повреждений ДНК. Митохондриальная полимеразы  $\gamma$ , состоящая из четырех идентичных полипептидов, ответственна за репликацию митохондриального генома.  $\delta$ -Полимераза похожа на полимеразу  $\alpha$  по своим молекулярным и синтетическим свойствам и также участвует в репликации хромосомной ДНК. Поскольку  $\alpha$ -,  $\beta$ - и  $\gamma$ -ДНК-полимеразы млекопитающих лишены 3'-5'- и 5'-3'-экзонуклеазных активностей, присущих ферментам *E. coli*, остается неясно, как в процессе репликации ДНК у этих организмов удаляются случайно включенные ошибочные нуклеотиды и РНК-затравки на концах фрагментов Оказки.

Некоторые вирусы животных (например, гер-

песвирус, вирус коровьей оспы и вирус гепатита) индуцируют синтез особых полимераз для репликации своих геномов. Другие вирусы образуют белки, которые стимулируют системы репликации клеточной ДНК или участвуют в репликации вирусной ДНК. Например, паповавирусы синтезируют белки, необходимые для инициации репликации. Аденовирусы человека кодируют белки, «запускающие» инициацию синтеза обеих цепей линейной вирусной ДНК (разд. 2.1.ж). Они продуцируют также особые ДНК-связывающие белки, облегчающие репликацию.

**ДНК-лигазы.** ДНК-лигазы необходимы для соединения цепей ДНК при репликации, репарации и рекомбинации. Все известные лигазы способны образовывать фосфодиэфирные мостики между 5'-фосфорильной и 3'-гидроксильной группами соседних дезоксирибонуклеотидов в местах разрывов ДНК (рис. 2.17). ДНК-лигаза, индуцируемая в *E. coli* после заражения клетки фагом Т4 (Т4-лигаза), уникальна по своей способности соединять двухцепочечные фрагменты ДНК по концам разрыва (рис. 2.18). Физиологическая роль этой реакции неизвестна, но практическое ее значение в манипуляциях с рекомбинантной ДНК неоспоримо (разд. 4.5 и 6.2).

ДНК-лигазы некоторых бактерий (*E. coli* и *B. subtilis*), бактериофагов (Т4 и Т7) и млекопитающих выделены и их структура и механизм каталитической активности установлены. ДНК-лигазы *E. coli*, Т4 и Т7 — это одиночные полипептидные цепи,

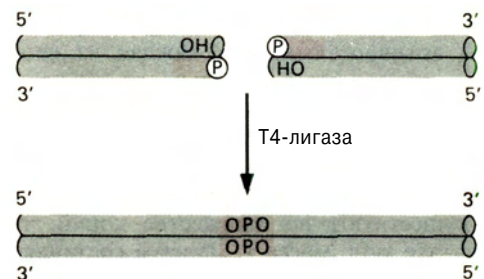


РИС. 2.18.

ДНК-лигаза, кодируемая ДНК фага Т4, соединяет два фрагмента дуплексной ДНК конец в конец. В каждой из цепей образуется новая фосфодиэфирная связь.

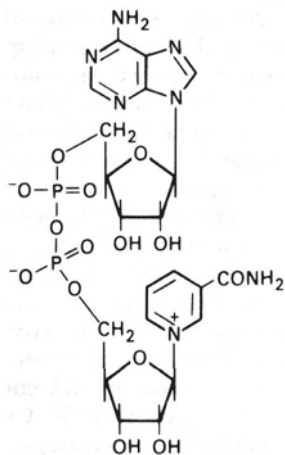


РИС. 2.19.

Никотинамидадениндинуклеотид (NAD).

а структура двух видов ДНК-лигазы млекопитающих неизвестна. Для образования фосфодиэфирных связей между соответствующими концами нуклеотидных цепей лигазы используют энергию гидролиза

либо АТФ, либо его производного — никотинамид-адениндинуклеотида (NAD) (рис. 2.19). Реакция протекает в несколько этапов (рис. 2.20): 1) аденилильная единица NAD (лигазы *E. coli* и *B. subtilis*) или АТФ (лигазы фагов Т4, Т7 и млекопитающих) переносится на ε-аминогруппу лизинового остатка лигазы с одновременным высвобождением никотинамидмононуклеотида или неорганического фосфата соответственно; 2) аденилильная группа переносится от белка на 5'-фосфорильную группу концевого остатка цепи ДНК с образованием пиррофосфорильного производного, аденилил-ДНК; 3) аденилильная группа, связанная с 5'-фосфорильной группой, замещается 3'-гидроксильной группой прилегающего конца ДНК. В результате этих реакций происходит образование фосфодиэфирных связей в цепи ДНК за счет энергии гидролиза пиррофосфорильной связи NAD или АТФ. Для образования фосфодиэфирной связи во всех случаях, кроме лигазы фага Т4, должно произойти соединение нуклеотида, содержащего акцепторную 3'-гидроксильную группу, с соседним нуклеотидом, несущим активированную 5'-фосфорильную группу. ДНК-лигазы *E. coli* и фага Т4 могут соединять

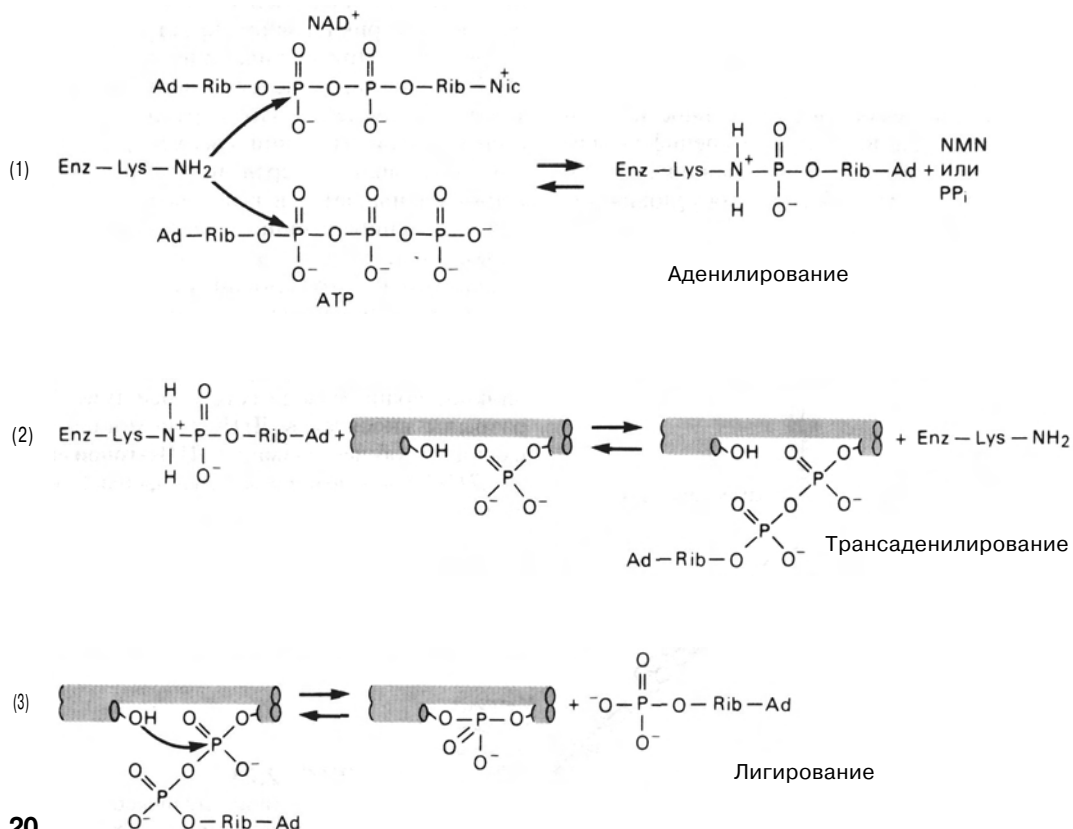


РИС. 2.20.

Механизм действия ДНК-лигазы (Enz). Сокращения: Lys - лизин, входящий в состав лигазы, Rib - рибоза,

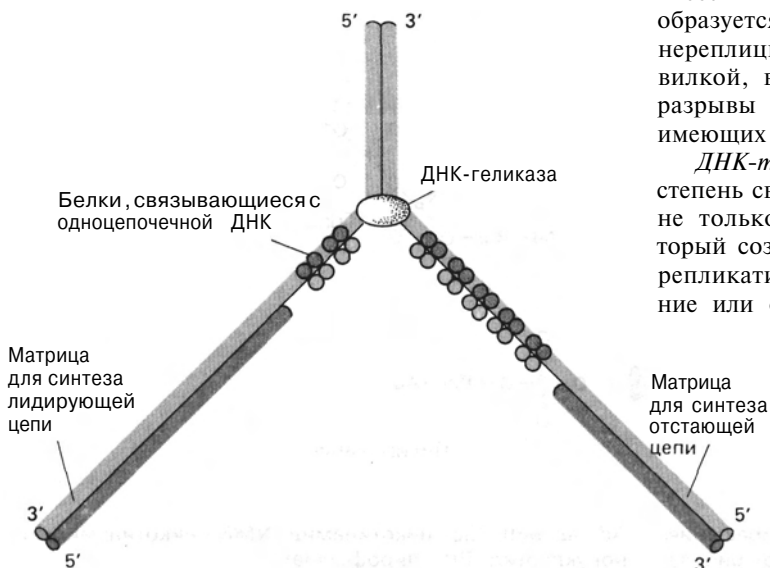
Ad - аденин, Nic - никотинамид, NMN - никотинамидмононуклеотид,  $\text{PP}_i$  - пиррофосфат.

концы двух разных дуплексных фрагментов или разорванные концы цепей линейной или кольцевой ДНК. Таким образом, с помощью ДНК-лигаз могут образовываться и линейные, и кольцевые дуплексные молекулы ДНК.

### е. Для репликации необходимо раскручивание спирали

Для осуществления комплементарного копирования цепей двухцепочечная ДНК должна постепенно раскручиваться. Раскручивание, или расплетание, спирали происходит только в локальном участке репликативной вилки. Расплетание—это не спонтанный процесс, в нем участвуют белки двух типов (рис. 2.21). Одни из них, называемые **ДНК-геликазами**, используют для разделения цепей энергию, высвобождающуюся при гидролизе АТФ до АДФ. Геликазы часто функционируют в составе комплекса, осуществляющего перемещение репликативной вилки и репликацию расплетенных цепей (разд. 2.1). Вообще говоря, для расплетания достаточно одного геликазного белка, но для того, чтобы максимизировать скорость раскручивания, несколько геликаз могут действовать совместно. Белки второго типа, дестабилизирующие спираль,— **это белки, связывающиеся с одноцепочечными участками** и тем самым стабилизирующие расплетенный дуплекс. Итак, геликазы вызывают локальное раскручивание двойной спирали, а другие специфические белки тотчас связываются с образовавшимися одноцепочечными участками, обеспечивая условия для комплементарного спаривания.

При репликации хроматиновой ДНК (разд. 1.2.ж) эукариотических организмов возникают допол-



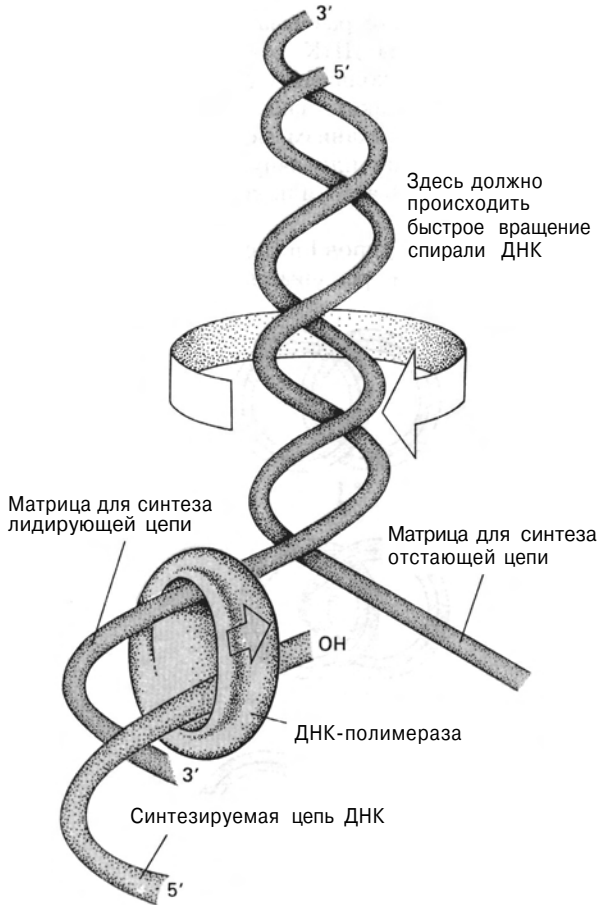
нительные сложности. Чтобы расплести ДНК в составе хроматина, необходимо разрушить сильно конденсированный комплекс гистонов и ДНК, а по завершении репликации вновь упаковать две дочерние спирали в сложные хроматиновые структуры. Каким образом раскручиваются эукариотические хромосомы при подготовке к репликации? Об этом известно очень мало. Развернутые, предсуществующие и новые нуклеосомы (собранные из новосинтезированных гистонов) должны ассоциировать с синтезированной дуплексной ДНК. Чтобы предотвратить образование слишком протяженных участков несвязанной ДНК в период репликации, сборка хроматина должна происходить параллельно образованию дуплексной ДНК. Связываются ли родительские гистоновые октамеры с одной или обоими дочерними спиральями—пока неизвестно.

**Топологические проблемы раскручивания и репликации ДНК.** Процесс раскручивания двойной спирали в репликативной вилке порождает механические и топологические проблемы. В принципе расплетание линейной дуплексной ДНК может происходить благодаря вращению родительской спирали вокруг собственной оси (рис. 2.22). Однако вращение очень длинных цепей ДНК вокруг длинных же осей во внутриклеточном пространстве механически затруднено. При репликации замкнутых кольцевых ДНК расплетание цепей в вилке создает дополнительные проблемы. По мере раскручивания цепей степень отрицательной сверхспиральности сегментов, находящихся перед вилкой репликации, постепенно уменьшается и в них возникает положительная сверхспирализация. Дальнейшее перемещение вилки вдоль кольца затрудняется и в конце концов блокируется. Это блокирование снимается путем внесения одноцепочечного разрыва. Тем самым образуется «шарнир», который дает возможность нереплицированному дуплексу, находящемуся перед вилкой, вращаться вместе с ней (рис. 2.23). Такие разрывы вносятся в ДНК с помощью ферментов, имеющих общее название **ДНК-топоизомеразы**.

**ДНК-топоизомеразы.** Эти ферменты изменяют степень сверхспиральности и тип сверхспирали. Они не только приводят к образованию шарнира, который создает условия для непрерывного движения репликативной вилки, но и обеспечивают разделение или образование **катенов**—сцепленных коль-

**РИС. 2.21.**

Расплетание дуплексов ДНК облегчается в присутствии двух белков: геликазы и белка, связывающегося с одноцепочечной ДНК.

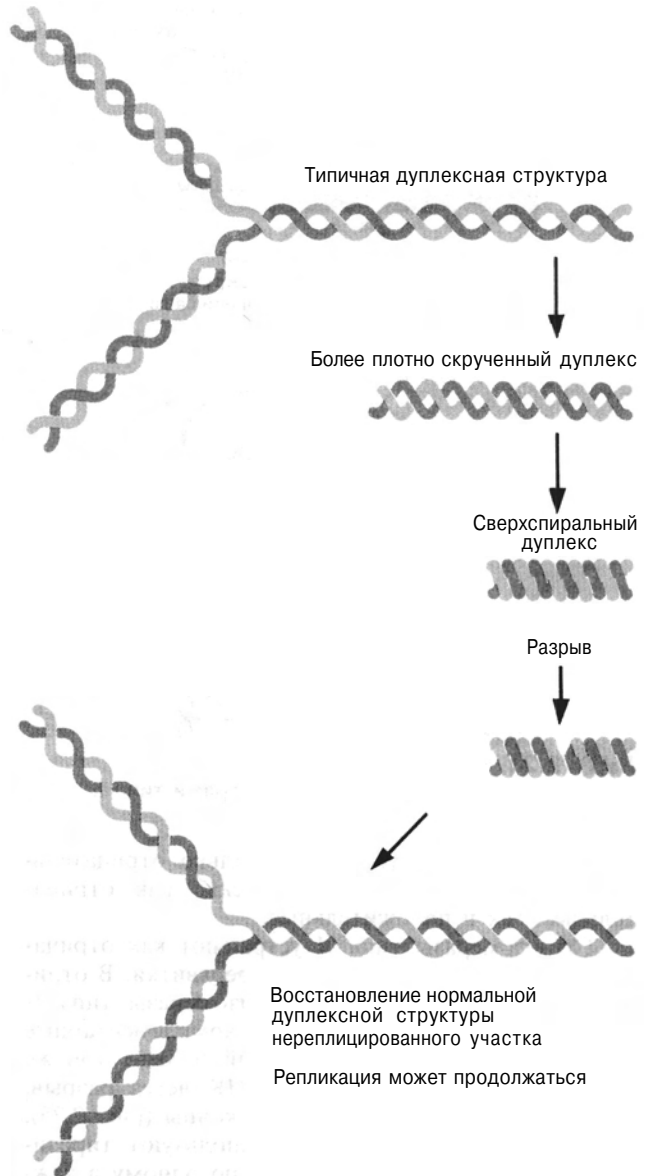
**РИС. 2.22.**

Вращение двойной спирали ДНК может снять механическое напряжение, возникающее при раскручивании двух цепей в репликативной вилке. [В. Alberts et al., *Molecular Biology of the Cell* (New York: Garland, 1983), p. 228.]

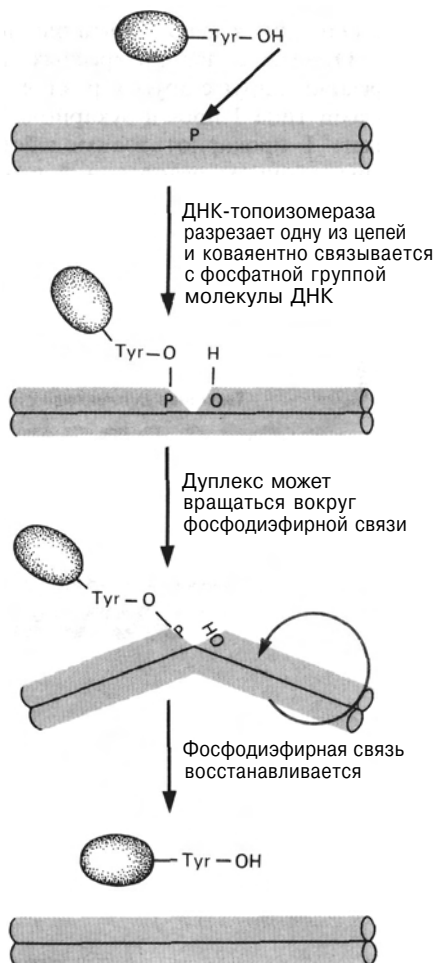
цепей ДНК, а также устранение узлов и спутанных клубков из длинной линейной ДНК. Топоизомеразы являются также неотъемлемыми участниками некоторых рекомбинационных процессов (разд. 2,4).

В различных организмах идентифицированы топоизомеразы двух основных типов. Одни ферменты, называемые топоизомеразами типа I, уменьшают число сверхвитков в ДНК на единицу за один акт. Эти топоизомеразы надрезают одну из двух цепей, в результате чего фланкирующие дуплексные области могут повернуться вокруг интактной цепи, и затем воссоединяют концы разрезанной цепи (рис. 2.24). Эта реакция не требует энергии АТФ, поскольку энергия фосфодиэфирной связи сохраняется благодаря тому, что тирозиновый остаток в молекуле фермента выступает то в роли акцептора, то в роли донора фосфорильного конца разре-

занной цепи. Одиночная цепь спонтанно проходит через разрез. Отмечены два интересных, но, возможно, не связанных друг с другом различия между топоизомеразами типа I про- и эукариот: 1) топоизомеразы типа I прокариот взаимодействуют с 5'-фосфорильным концом разорванной цепи, а эукариот — с 3'-фосфорильным концом; 2) топоизо-

**РИС. 2.23.**

Снятие механического и топологического напряжений в замкнутом кольцевом дуплексе перед репликативной вилкой. При внесении разрыва нереплицированная сверхспиральная часть ДНК может вращаться вместе с репликативной вилкой. Напряжение, возникающее в длинных цепях ДНК во время репликации *in vivo*, снимается именно таким способом.



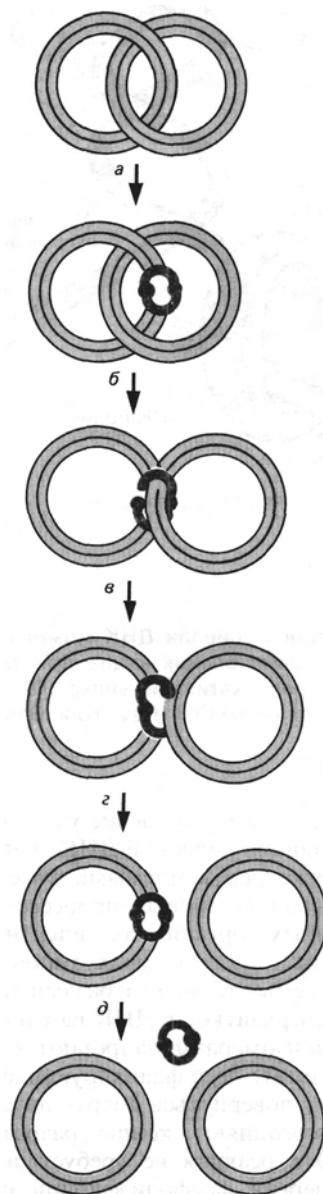
**РИС. 2.24.**  
Реакции, катализируемые топоизомерами типа I.

меразы прокариот устраняют только отрицательные сверхвитки, а эукариотические—как отрицательные, так и положительные.

Топоизомеразы типа II устраняют как отрицательные, так и положительные сверхвитки. В отличие от ферментов типа I топоизомеразы типа II вносят временные разрывы в обе комплементарные цепи, пропускают двухцепочечный сегмент той же самой или другой молекулы ДНК через разрыв, а затем соединяют разорванные концы (рис. 2.25). Топоизомеразы типа II тоже используют тирозиновые остатки (присутствующие по одному в каждой из субъединиц ферментов) для связывания 5'-конца каждой разорванной цепи в то время, когда другой дуплекс проходит через место разрыва. В результате внесения двухцепочечного разрыва и прохождения через него другого дуплекса за один акт снимаются два отрицательных или положительных сверхвитка. В некоторых случаях дуплексом, про-

ходящим через место разрыва, оказывается другая замкнутая молекула ДНК; это приводит к разделению сцепленных кольцевых ДНК или, напротив, к образованию таких сцепленных комплексов (катенанов). Этот механизм может использоваться и для распутывания или запутывания клубков, а также для раскручивания или конденсации крупных дуплексных ДНК.

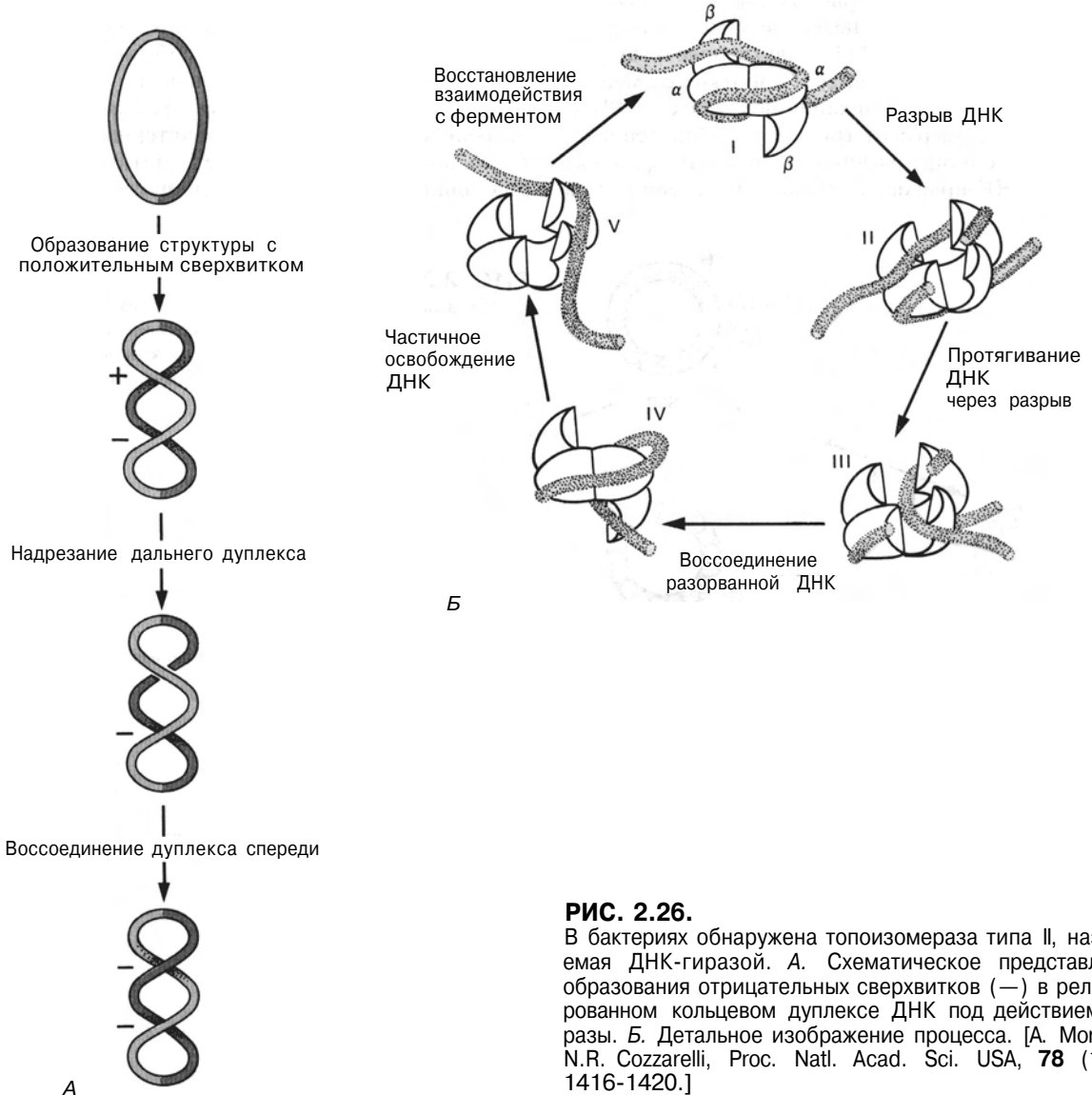
Топоизомеразы типов I и II снимают сверхвитки, образующиеся при репликации кольцевой ДНК.



**РИС. 2.25.**  
Топоизомеразы типа II катализируют размыкание двух кольцевых дуплексов. [B. Alberts et al., *Molecular Biology of the Cell* (New York: Garland, 1983), p. 229.]

Однако существует особая топоизомераза II, называемая **гиразой** и обнаруженная пока только у бактерий, которая индуцирует образование отрицательных сверхвитков в релаксированных кольцевых ДНК. Для этого гираза делает двухцепочечные надрезы и затем особым способом воссоединяет концы (рис. 2.26, А). Итак, гираза снимает положительные сверхвитки и вносит отрицательные в релаксированную ДНК. Сбалансированное действие топоизомеразы I и гиразы — по крайней мере у бактерий — по-видимому, регулирует степень сверхспиральности ДНК и таким образом влияет на скорость движения репликативной вилки. Механизм, с помощью которого гираза катализирует образование отрицательных сверхвитков в кольцевой

или другой ДНК с топологическими ограничениями, до конца не установлен. Гираза *E. coli* представляет собой тетрамер, состоящий из субъединиц двух типов ( $\alpha\beta_2$ ), при этом  $\alpha$ -субъединицы содержат сайты ковалентного связывания концов молекулы ДНК. Гираза катализирует образование отрицательных сверхвитков, создавая сначала положительные сверхвитки в определенных областях ДНК, связанных с ферментом (рис. 2.26, Б). Ориентация двух спиральных сегментов в этих областях меняется на противоположную при протягивании одного сегмента через «ворота», образовавшиеся в результате двухцепочечного разрыва в другом. В конце концов образуются два сверхвитка за один каталитический цикл. Для реализации всех этих



**РИС. 2.26.**

В бактериях обнаружена топоизомераза типа II, называемая ДНК-гиразой. А. Схематическое представление образования отрицательных сверхвитков (—) в релаксированном кольцевом дуплексе ДНК под действием гиразы. Б. Детальное изображение процесса. [A. Morrison, N.R. Cozzarelli, Proc. Natl. Acad. Sci. USA, **78** (1981), 1416-1420.]

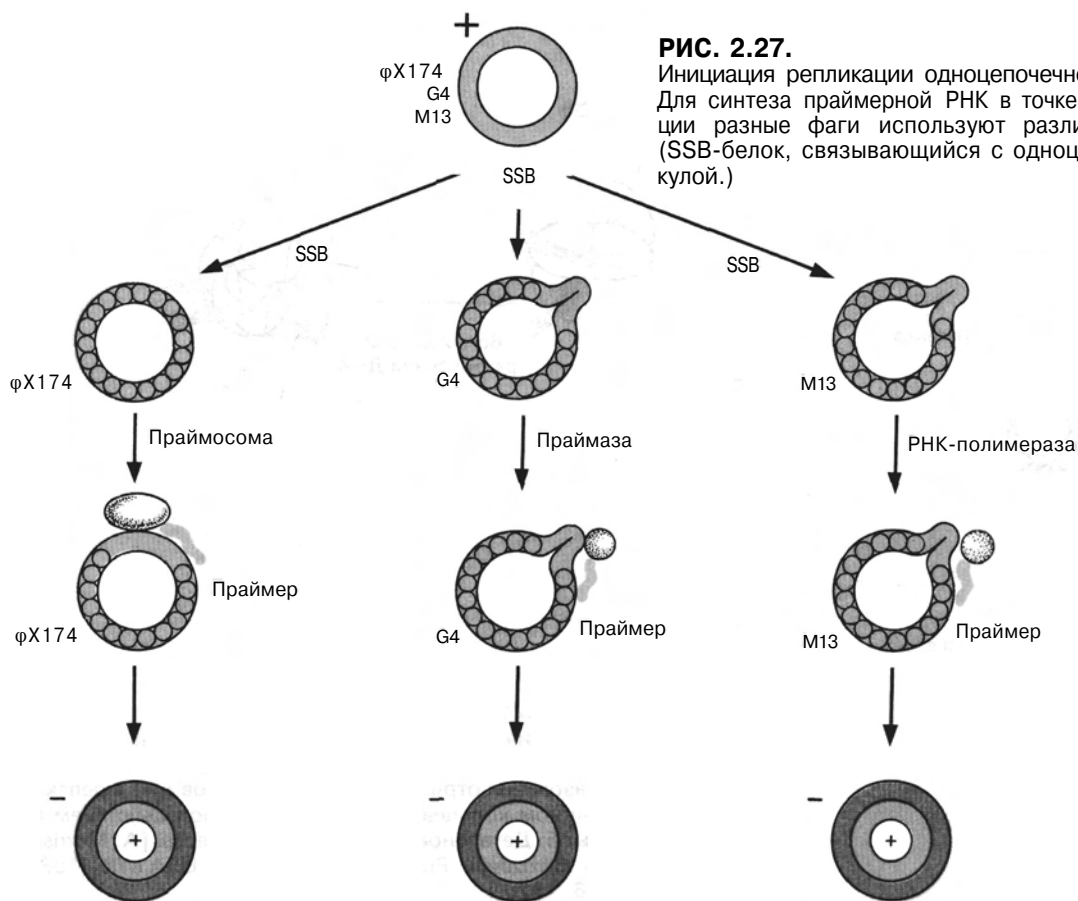
процессов необходима энергия гидролиза АТФ, поскольку релаксированная кольцевая ДНК должна быть переведена на более высокий энергетический уровень, характерный для сверхспиральной конформации.

### ж. Инициация образования новых цепей ДНК и их рост в репликативных вилках

*Инициация образования новых цепей ДНК.* В отличие от РНК, синтез которой иницируется РНК-репликазами (разд. 2.5) и РНК-полимеразами (часть III), для инициации синтеза ДНК требуется затравка (праймер). Сначала синтезируются короткие РНК-праймеры, которые затем наращиваются в виде цепей ДНК. При переводе кольцевой одноцепочечной ДНК фага М13 в двухцепочечную репликативную форму для синтеза коротких сегментов РНК в точке начала репликации используется РНК-полимераза-холофермент (рис. 2.27). При репликации других одноцепочечных кольцевых фаговых ДНК (G4) РНК-праймер синтезируется специальной

РНК-полимеразой, называемой **праймазой**. У другого фага с одноцепочечной кольцевой ДНК (φX174) мультиферментный комплекс, содержащий от 15 до 20 полипептидных цепей (**праймосома**), активирует матричную цепь, в результате чего праймаза синтезирует РНК-затравку. Праймосома-праймазный способ инициации применяется и для инициации синтеза ведущей цепи в точке начала репликации у *E. coli*, а также при образовании фрагментов Оказаки при синтезе отстающей цепи.

По-видимому, синтез цепей ДНК иницируется сходным образом (с незначительными вариациями) у про- и эукариот, в случае плазмид и вирусов. Различия могут касаться типов полимераз, синтезирующих РНК-затравку, вспомогательных белков, обеспечивающих инициацию РНК-транскриптов, и либо нуклеотидных последовательностей *reg se*, либо структурных особенностей, определяющих положение точки начала транскрипции. Длина РНК-праймеров варьирует в зависимости от особенностей инициации. У бактериофага φX174 и *E. coli* праймаза катализирует синтез РНК-праймеров длиной от двух до пяти нуклеотидов, а при



**РИС. 2.27.**

Инициация репликации одноцепочечной фаговой ДНК. Для синтеза праймерной РНК в точке начала репликации разные фаги используют различные ферменты. (SSB-белок, связывающийся с одноцепочечной молекулой.)

репликации ДНК эукариот праймерные РНК приблизительно в два раза длиннее. У бактериофагов М13 и G4 синтез РНК-затравок катализируют РНК-полимераза и G4-праймаза соответственно; длина этих затравок колеблется от 20 до 30 нуклеотидов, и они комплементарны сегментам фаговых ДНК, образующим петли (рис. 2.27). Точная природа сигналов, определяющих сайты синтеза праймерной РНК для инициации репликации или для синтеза отставшей цепи, неизвестна.

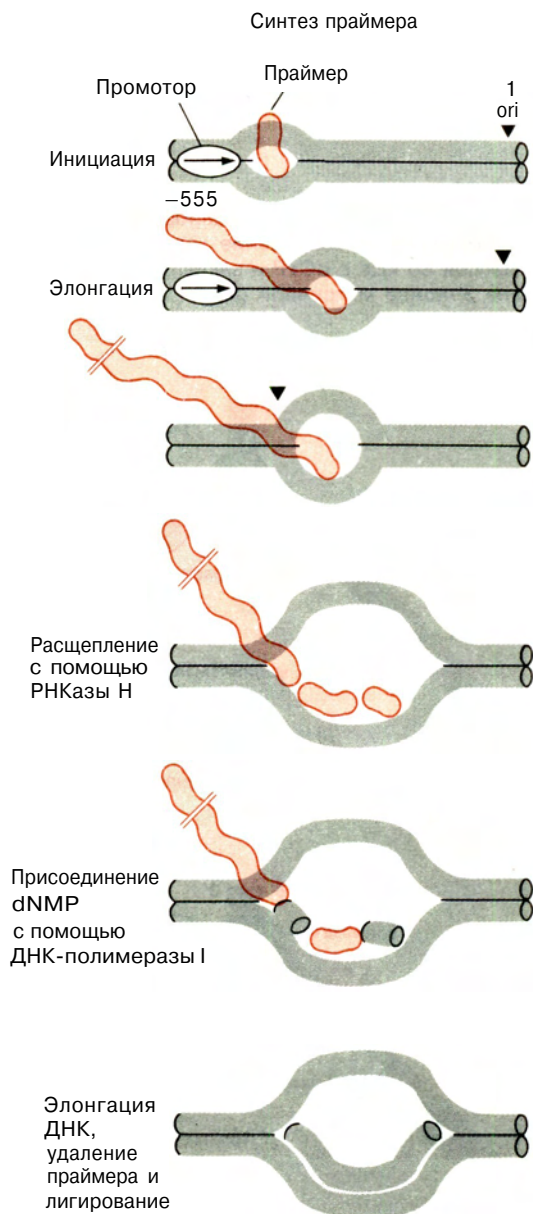
Репликация плазмидной ДНК Col E1 также начинается с транскрипции РНК (в гл. 5 описано использование таких плазмид в качестве векторов для клонирования). Однако в этом случае синтез праймерной РНК инициируется РНК-полимеразой в сайте, отстоящем на 555 пар оснований от точки начала репликации плазмидной ДНК (рис. 2.28). Транскрипция проходит точку начала репликации, при этом синтезируется цепь РНК длиной 500–600 нуклеотидов. Затем РНКаза-Н расщепляет большую часть цепи РНК. Оставшийся 3'-конец РНК служит затравкой при синтезе цепи ДНК с помощью Pol I-полимеразы. Эта цепь становится ведущей при репликации всей плазмидной ДНК.

У аденовирусов человека синтез новых цепей ДНК инициируется неким белком, а не РНК (рис. 2.29). Аденовирус типа 2 содержит белок с мол. массой 55 кДа, ковалентно связанный с 5'-фосфатом каждой цепи ДНК. Первый этап инициации синтеза новых цепей состоит в том, что кодируемый вирусом белок мол. массой 80 кДа связывается с дезоксирибонуклеозидтрифосфатом, соответствующим первому дезоксирибонуклеотидному остатку в ДНК (dСТР), с образованием комплекса белок-dСМР. 3'-гидроксильная группа связанного с белком дезоксицитидинового остатка служит праймером для удлинения цепей ДНК с использованием комплементарной цепи ДНК в качестве матрицы. В конечном счете вирусный 80 кДа-белок отщепляется и с новосинтезированной вирусной ДНК остается связанным только терминальный белок мол. массой 55 кДа. Аналогичный механизм инициации синтеза ДНК используется и другими аденовирусами и некоторыми бактериофагами.

*Одновременная репликация обеих цепей в репликативной вилке.* Изучение механизма репликации фаговой ДНК с использованием очищенных ферментов из *E. coli* позволило построить усовершенствованную модель полуконсервативной прерывистой репликации ДНК в репликативной вилке (рис. 2.30). Для расплетания спирали при образовании вилки необходимо, чтобы в ДНК образовались отрицательные сверхвитки (разд. 2.1.е). Расплетание осуществляется АТР-зависимой геликазой и облегчается в присутствии белка, связывающегося с одноцепочечной ДНК. После инициации репликации

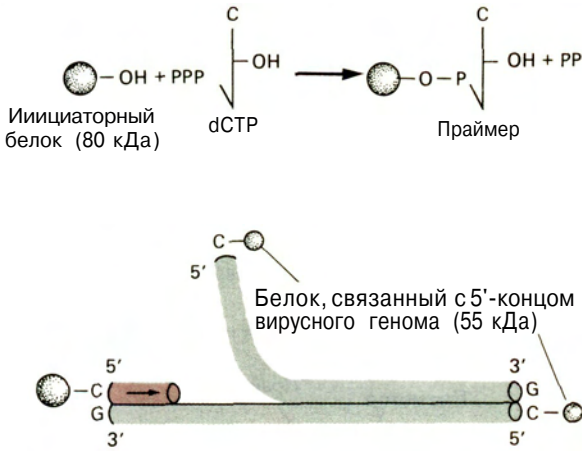
Pol III-холофермент удлиняет лидирующую цепь ДНК до размеров, равных длине хромосомы, путем последовательного присоединения к 3'-гидроксильному концу дезоксирибонуклеотидных единиц.

Синтез отставшей цепи инициируется праймосо-праймазным комплексом (рис. 2.31). Праймосома, обладающая геликазной активностью, расплетает спираль и, возможно, способствует формированию подходящих структур, позволяющих ассоциированной праймазе синтезировать праймер. Прай-



**РИС. 2.28.** Инициация репликации плазмид типа Col E1.

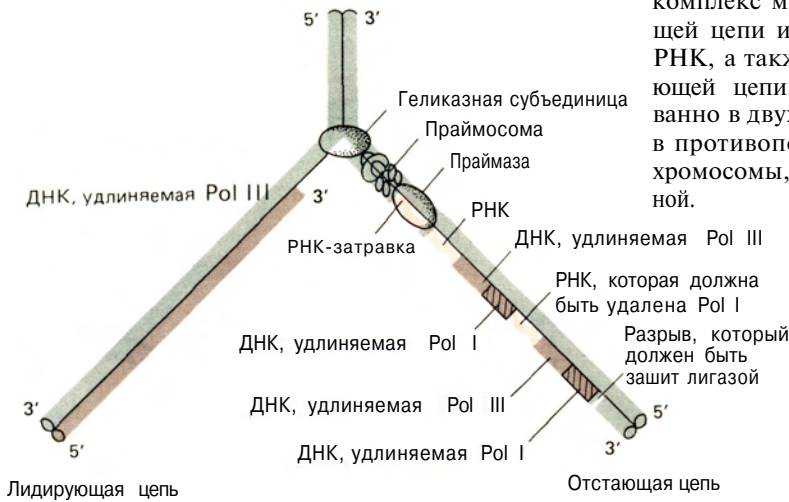




**РИС. 2.29.**

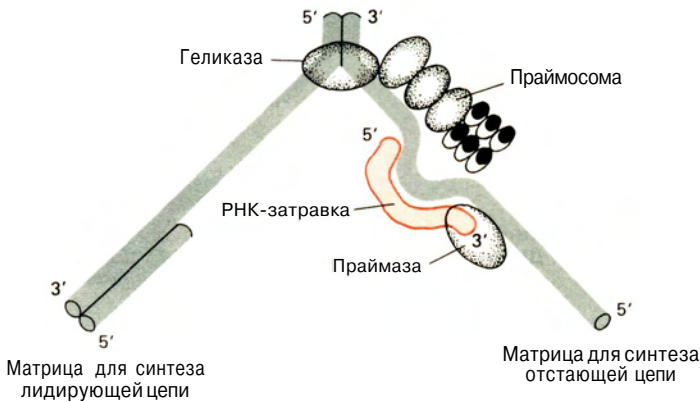
Инициация репликации ДНК аденовируса. Ключевым моментом является связывание инициаторного белка мол. массой 80 кДа с dCTP и закрепление этого комплекса на 3'-конце матричной цепи.

мосомо-праймазный комплекс движется в том же направлении, что и вилка, останавливаясь только для того, чтобы расплести дуплекс и синтезировать РНК-праймеры. Эти короткие фрагменты РНК достраиваются затем Pol III-холоферментом путем добавления дезоксинуклеотидтрифосфатов с образованием фрагментов Оказаки длиной от 1000 до 2000 нуклеотидов (рис. 2.30). Удаление сегментов РНК с 5'-конца каждого фрагмента Оказаки и заполнение пробелов между ними катализируется Pol I, способной удлинять цепи и осуществлять ник-трансляцию. Когда растущий 3'-гидроксильный конец каждого фрагмента Оказаки доходит до 5'-дезокси-нуклеотидного конца соседнего фрагмента, вступает в действие ДНК-лигаза и образуется непрерывная отстающая цепь. Более сложная, но все же гипотетическая модель (рис. 2.32) движения репликативной вилки предполагает формирование **реплисомы** — мультиферментного комплекса более высокого уровня организации, состоящего из функционального праймосомо-праймазного комплекса, геликазы, Pol III-холофермента и, возможно, гиразы. Такой комплекс может обеспечивать удлинение лидирующей цепи и одновременно инициацию праймерной РНК, а также достраивание ДНК при синтезе отстающей цепи. Две реплисомы, работающие согласованно в двух вилках репликации, которые движутся в противоположных направлениях вдоль кольцевой хромосомы, сделали бы эту модель еще более изящной.



**РИС. 2.30.**

Полуконсервативный механизм репликации ДНК *E.coli*.



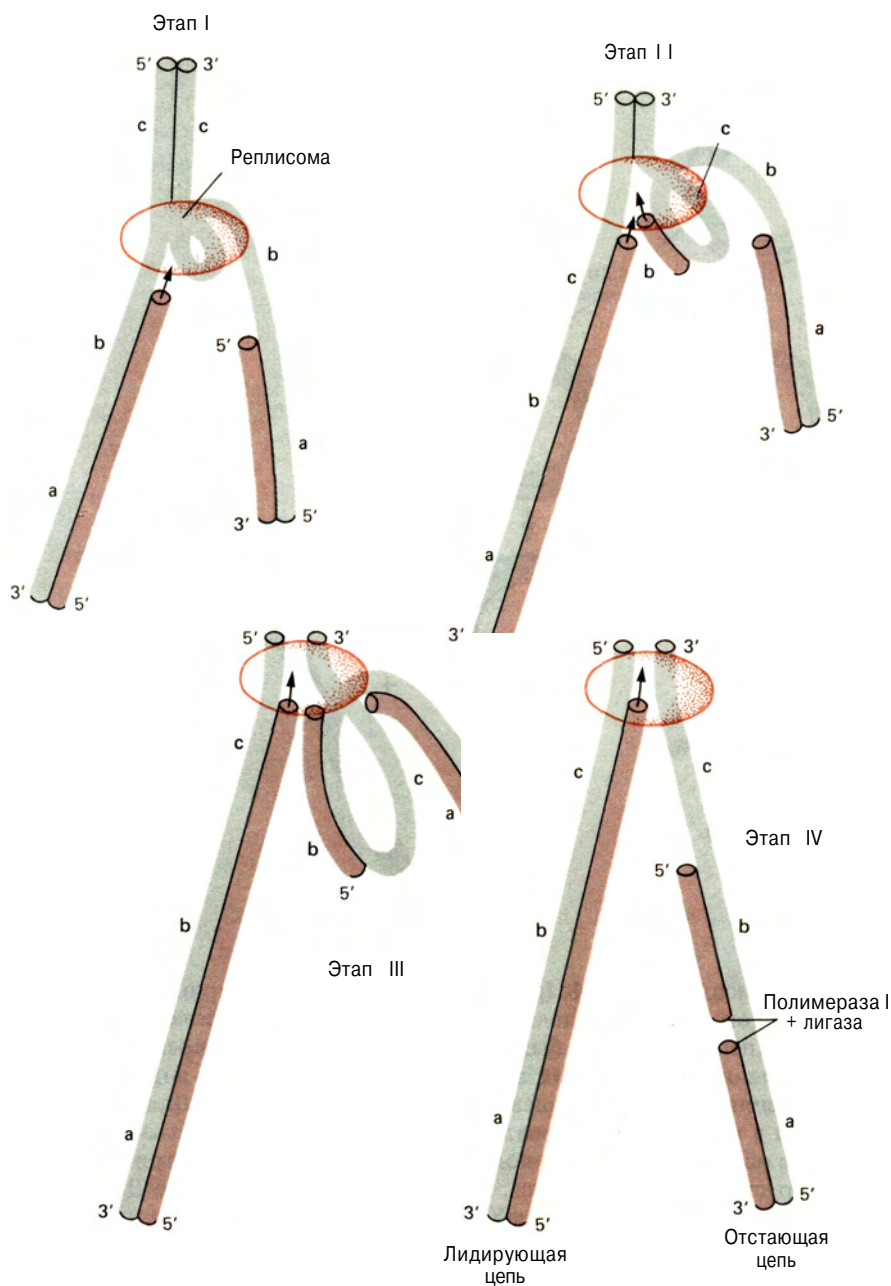
**РИС. 2.31.**

Инициация синтеза отстающей цепи. [Из работы А. Kornberg, DNA Replication, Supplement (San Francisco: W.H. Freeman, 1982), p. S113, с изменениями.]

*Репликация по типу катящегося кольца.* Некоторые двухцепочечные кольцевые хромосомы реплицируются альтернативным способом, называемым **репликацией по типу катящегося кольца** (рис. 2.33). В этом случае двухцепочечная кольцевая ДНК надрезается специфическим ферментом в уникальном сайте одной цепи (точке начала катящегося кольца), и к образовавшемуся в результате надреза 3'-гидроксильному концу с помощью Pol III-холофермента

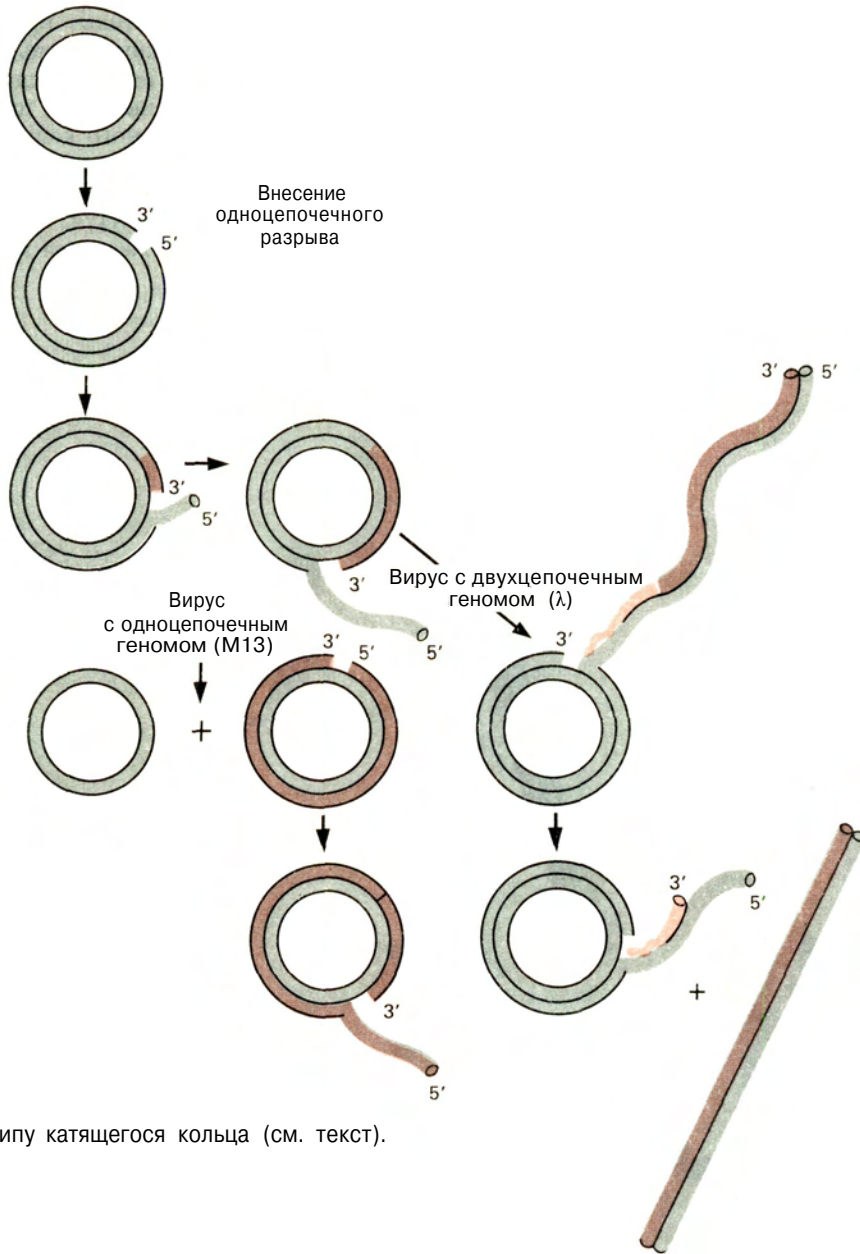
присоединяются нуклеотиды; при этом матрицей служит интактная замкнутая цепь. Таким образом, в вилке синтезируется только лидирующая цепь. По мере синтеза лидирующей цепи происходит вытеснение 5'-конца надрезанного кольца как одиночной цепи. В результате длина лидирующей цепи может превышать длину матрицы в два-пять раз.

У фагов M13 или φX174, чьи зрелые геномы представлены одноцепочечными кольцевыми ДНК,



**РИС. 2.32.**

Синтез обеих цепей ДНК в репликативной вилке с участием мультиферментного комплекса - реплисома. [Из работы А. Kornberg, DNA Replication, Supplement (San Francisco: W.H. Freeman, 1982), p. S125.]

**РИС. 2.33.**

Репликация по типу катящегося кольца (см. текст).

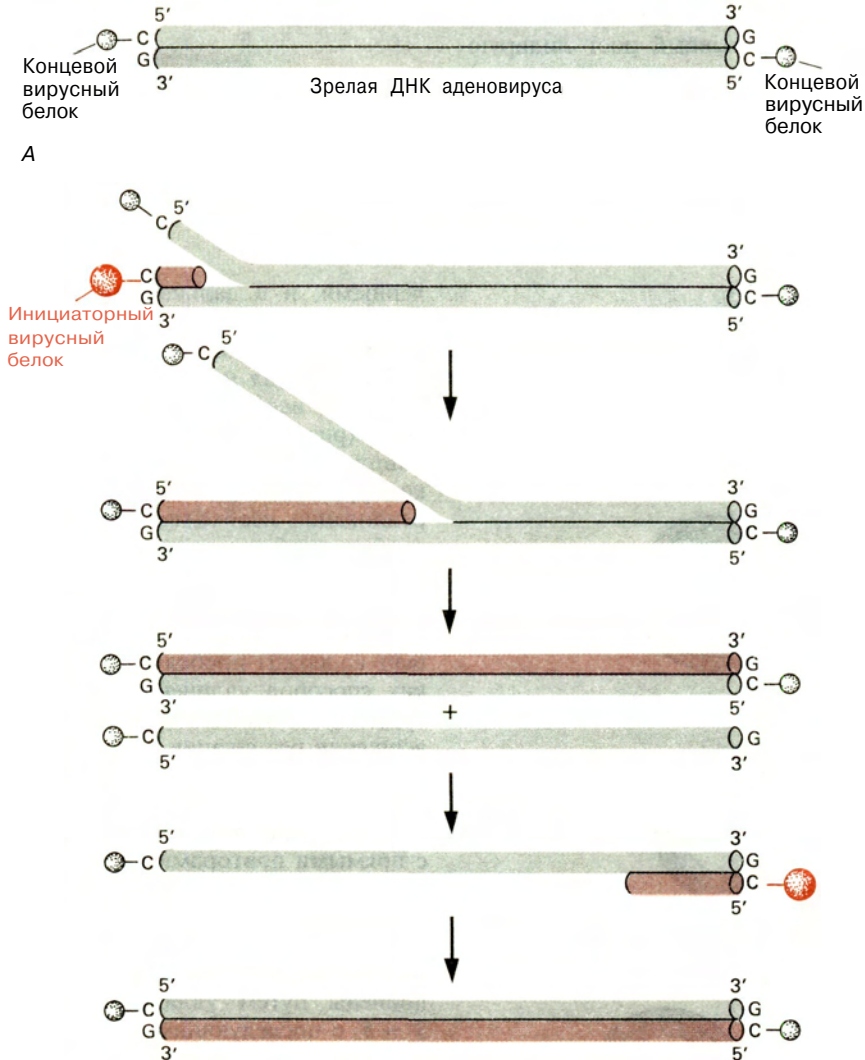
такой способ репликации используется на поздних стадиях инфекционного процесса, после того как инфицирующая ДНК превращается в двухцепочечную кольцевую форму. Постоянно отделяющиеся одиночные цепи ДНК, образуемые при репликации по типу катящегося кольца, надрезаются в каждой точке начала репликации и замыкаются с образованием зрелых форм, упаковываемых в вирусные частицы.

Фаг  $\lambda$  использует такой способ репликации при образовании двухцепочечной линейной вирусной

ДНК (рис. 2.33). Субстратной матрицей в этом случае является двухцепочечная кольцевая ДНК, которая была реплицирована после превращения на ранних этапах инфекции линейной вирусной ДНК в кольцевую репликативную форму. Для того чтобы образовалась линейная ДНК потомства, двухцепочечные кольца надрезаются и асимметрично реплицируются, как это происходит в случае ДНК фагов M13 и  $\phi$ X174. Однако отделяющиеся одиночные цепи превращаются в двухцепочечные структуры. Сначала во множестве сайтов вдоль одиночной цепи

праймаза синтезирует короткие сегменты РНК. Затем эти сегменты достраиваются Pol III-холоферментом, пробелы заполняются, РНК удаляется с помощью Pol I и наконец короткие фрагменты ДНК соединяются ДНК-лигазой. При упаковке ДНК в фаговые частицы в специальных участках, называемых **cos-сайтами** и отстоящих друг от друга на длину вирусного генома, образуются надрезы. В результате длинные дуплексы многократно повторенной ДНК фага  $\lambda$  расчлняются на фрагменты, соответствующие по размерам зрелой ДНК, обнаруживаемой в вирионах.

*Поочередная репликация цепей.* Репликация ДНК аденовируса происходит без синтеза отстающей цепи и поэтому без образования множественных сайтов инициации и фрагментов Оказаки (рис. 2.34). Вместо этого цепи линейного дуплексного генома реплицируются попеременно. Сначала через образование белкового праймера происходит инициация одной цепи, которая непрерывно удлиняется вплоть до завершения репликации. Вытесненная новосинтезированной цепью, вторая цепь дуплекса служит матрицей при синтезе таким же способом следующего дуплекса. С какого конца родительского дуп-



**РИС. 2.34.**

**Б**

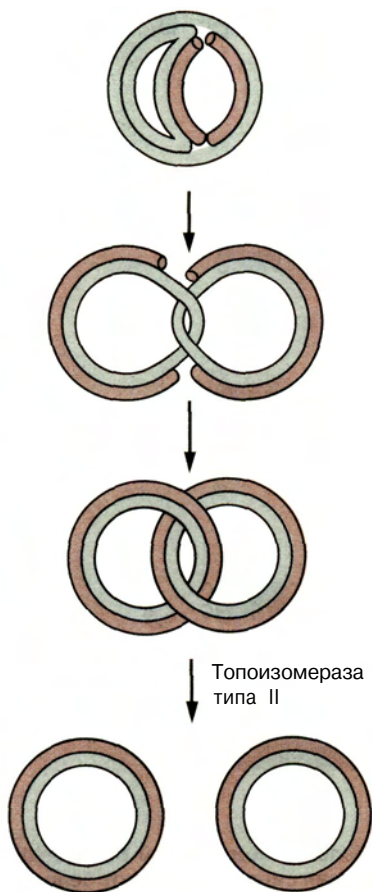
Поочередная репликация цепей ДНК аденовируса. А. Двухцепочечная ДНК аденовируса, у которой с каждым 5'-концевым остатком цитозина связан белок мол. массой 55 кДа. Б. Синтез цепей ДНК последовательно иницируется с конца каждой исходной цепи с помощью

комплекса «белок с мол. массой 80 кДа - dСТР». К моменту завершения синтеза цепи этот белок расщепляется, и часть его с мол. массой 55 кДа остается связанной с 5'-концом цепи.

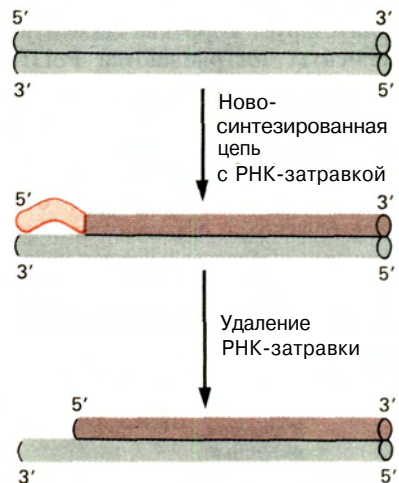
лекса начнется процесс, по-видимому, неважно. Репликация ДНК аденовируса необычна в двух отношениях: 1) использование белка для запуска синтеза цепи ДНК с конца линейного дуплекса; 2) отсутствие прерывистой репликации, опосредуемой РНК-транскриптами.

### 3. Терминация репликации ДНК и расхождение дочерних спиралей

*Терминация и расхождение в кольцевых геномах.* Замкнутость структуры многих геномных ДНК упрощает завершение репликации всей нуклеотидной последовательности. Непрерывный рост лидирующей и отстающей цепей вдоль кольцевой матрицы неизбежно приводит к совмещению 3'-гидрокси- и 5'-фосфорильного концов одной цепи либо в точке начала репликации, либо—при двунаправленной репликации—в середине кольца (рис. 2.35). Кольца в этих местах встречи соединяются ДНК-лигазой, при этом обычно они оказываются попарно сцеп-



**РИС. 2.35.** Репликация замкнутой кольцевой дуплексной ДНК и расхождение двух дочерних кольцевых молекул.



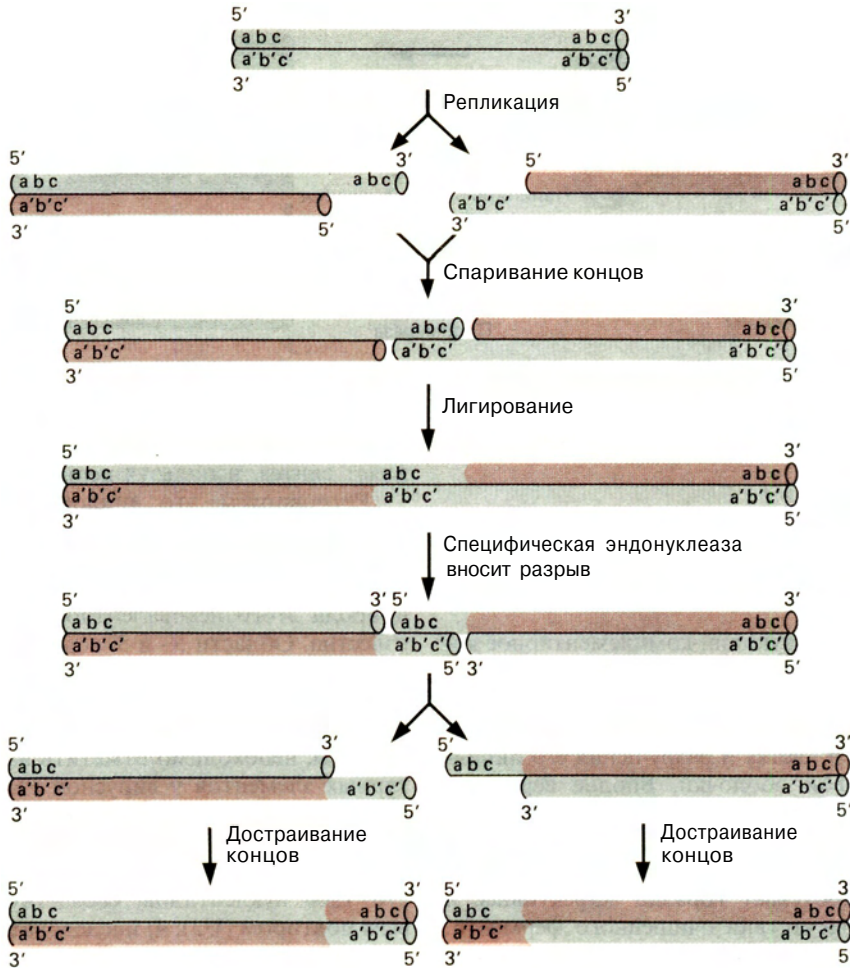
**РИС. 2.36.**

При удалении РНК-затравки с 5'-конца новосинтезированной цепи ДНК на одном конце дуплекса остается пробел.

ленными, и в дальнейшем должно произойти их разделение на отдельные геномы. Это происходит с помощью топоизомеразы типа II (рис. 2.25).

*Терминация и завершение репликации в линейных ДНК.* За исключением репликации аденовирусной ДНК (рис. 2.34), где синтез новых цепей ДНК иницируется белковым праймером и матричная цепь копируется полностью, во всех других случаях для репликации необходим РНК-праймер, что создает особые проблемы при завершении репликации линейной дуплексной ДНК (рис. 2.36). Дело в том, что после инициации синтеза новой цепи и последующего удаления РНК-праймера новосинтезированная цепь содержит пробел на 5'-конце. Поскольку никаких способов удлинения 5'-концов цепей ДНК не существует, необходимы какие-то иные методы завершения репликации.

Были предложены два способа, с помощью которых процесс репликации мог бы завершаться. Один из них предполагает, что существуют цепи ДНК с прямыми повторами на концах (рис. 2.37). После репликации два комплементарных конца обоих незавершенных дуплексов могут спариться и образовать линейные конкатемеры с одноцепочечными разрывами. Остающиеся пробелы могут быть заполнены путем удлинения цепей в направлении 3'→5' с последующим соединением их ДНК-лигазой либо путем прямого соединения стыкующихся концов с помощью ДНК-лигазы с образованием конкатемеров. После надрезания конкатемера специфической эндонуклеазой образуются выступающие 5'-концы, и ДНК-полимераза может наращивать более короткие цепи с 3'-конца. Другой способ предполагает наличие на конце каждой цепи ДНК

**РИС. 2.37.**

Предполагаемая модель завершения репликации линейной двухцепочечной ДНК. [По данным работы J.D. Watson, *Nature, New Biology*, **239** (1972), p. 197.]

коротких инвертированных повторов, благодаря которым образуются небольшие петли (рис. 2.38). 3'-конец петли служит праймером для копирования нереплицированного участка. Благодаря специфическому разрыву в начале инвертированного повтора получается структура, которую можно достроить с 3'-конца до восстановления исходной двухцепочечной концевой последовательности.

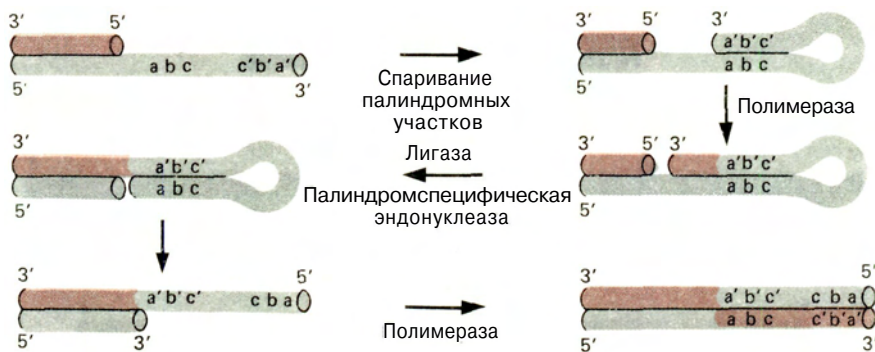
Недавно полученные данные свидетельствуют о том, что концевые области эукариотических хромосом — теломеры — реплицируются с помощью особого механизма, отличного от представленных выше. Концы хромосом дрожжей, беспозвоночных, растений и позвоночных имеют сходное, весьма необычное строение (разд. 9.6.б): они содержат шпилькообразные структуры, в которых 3'- и 5'-кон-

цы дуплекса ДНК оказываются рядом, и много коротких тандемных повторов. Около петли в одной из цепей в области повторов имеются множественные одноцепочечные разрывы. Вопрос о том, как подобная структурная организация может способствовать репликации концов дуплексных участков, выясняется. Если учесть сходство структурных особенностей конечных областей хромосом, то можно предположить, что механизм репликации всех эукариотических хромосом одинаков (разд. 9.6.б).

## 2.2. РЕПЛИКАЦИЯ РНК С ОБРАЗОВАНИЕМ ДНК

### а. Репликация геномов ретровирусов

Геном ретровирусов представлен единственной молекулой одноцепочечной РНК. После проникновения РНК в клетку хозяина вирусный геном подвергается **обратной транскрипции**, при этом сначала



**РИС. 2.38.**

Другая предполагаемая модель завершения репликации линейной двухцепочечной ДНК. [Т. Cavalier-Smith, *Nature*, **250** (1974), p. 467.]

образуется дуплекс РНК-ДНК, а затем двухцепочечная ДНК. Эти этапы предшествуют экспрессии вирусных генов на уровне белков и образованию РНК-геномов.

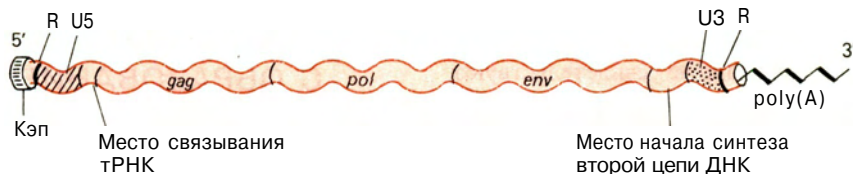
Фермент, катализирующий комплементарное копирование РНК с образованием ДНК, называется **обратной транскриптазой**. Он содержится в ретровирусных частицах (вирионах) и активируется после попадания вируса в клетку и разрушения его липидно-гликопротеиновой оболочки. Вполне вероятно, что обратной транскрипции способствуют какие-то вспомогательные белки, находящиеся внутри вирионов, поскольку в присутствии последних ферментативная реакция протекает гораздо эффективнее и быстрее, чем в присутствии очищенного фермента. Появляется все больше данных о том, что обратная транскрипция происходит в самых разных эукариотических клетках, а обратная транскриптаза играет важную роль в процессах перестройки генома (разд. 10.3 и 10.4).

Обратные транскриптазы ретровирусов по существу являются ДНК-полимеразами, и *in vitro* могут использовать в качестве матрицы ДНК. Однако гораздо эффективнее они работают, если матрицей является РНК. Как и все ДНК-полимеразы, обратные транскриптазы не способны инициировать синтез новых цепей ДНК. Но если синтез уже иници-

рован с помощью праймерной РНК или 3'-концевого участка ДНК, то фермент эффективно осуществляет синтез, используя цепь ДНК как матрицу.

Ретровирусы—это диплоидные организмы, поскольку каждый вирион содержит две идентичные цепи РНК размером от 8000 до 10000 нуклеотидов. Цепи соединены вблизи своих 5'-концов, однако природа этого нековалентного взаимодействия неизвестна. Области 5'- и 3'-концов обеих цепей модифицированы, как и у всех эукариотических мРНК (5'-кэпы и 3'-полиадениловые хвосты; разд. 3.8.a) (рис. 2.39). Рассматривая механизм обратной транскрипции, необходимо отметить наличие пяти структурных элементов у вирусной РНК: 1) прямые повторы на 5'- и 3'-концах РНК (R); 2) последовательность из 80–120 нуклеотидов, соседствующая с 5'-концевым повтором (U5); 3) последовательность из 170–1200 нуклеотидов, соседствующая с 3'-концевым повтором (U3); 4) последовательность из 15–20 нуклеотидов (P), в пределах которой клеточная тРНК спаривается с ретровирусной РНК, что создает праймер для синтеза первой цепи ДНК; 5) сегмент  $P_{\text{H}}$ , находящийся непосредственно перед повтором U3 и являющийся сайтом для праймирования второй цепи ДНК; такой сегмент одинаков у РНК всех ретровирусов определенного типа (т.е. его последовательность практически идентична у всех вирусов птиц и отличается от подобной последовательности у вирусов мышей).

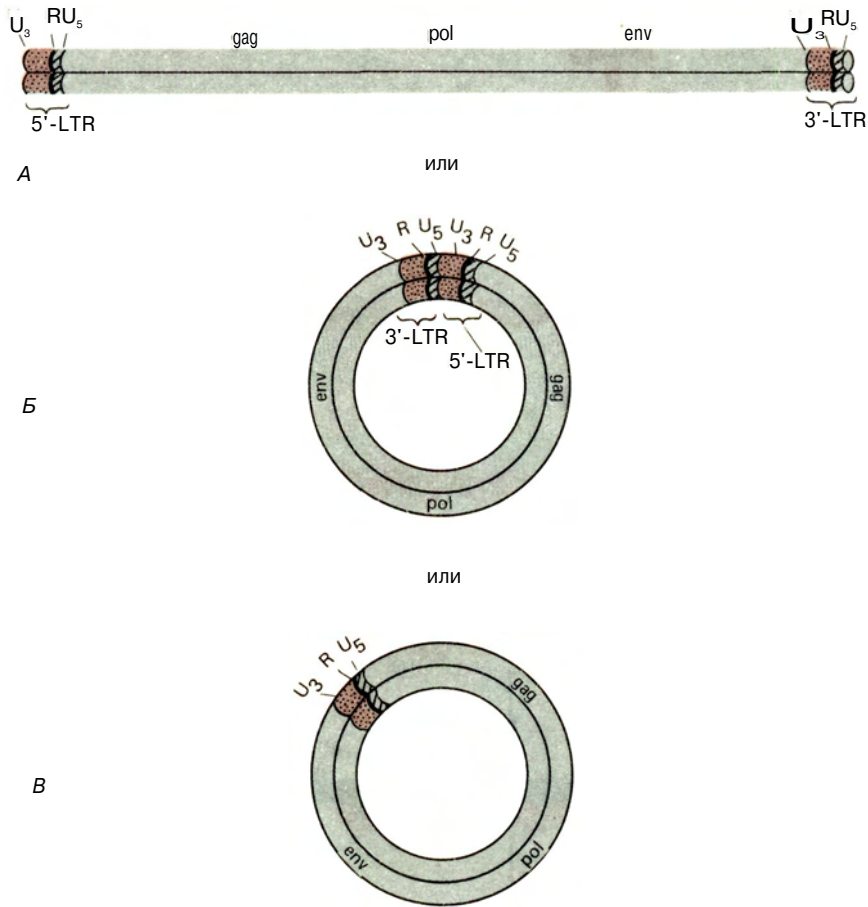
Известны три продукта, образующиеся в результате обратной транскрипции: форма А—линейный



**РИС. 2.39.**

РНК-геном ретровируса. Каждый вирион содержит две идентичные копии РНК. Обычно в геноме ретровирусов присутствуют три гена - *gag*, *pol* и *env*; ген *pol* кодирует обратную транскриптазу. Имеются также 5'-кэп и 3'-по-

лиадениловый хвост (см. разд. 3.8.a). Другие особенности генома (последовательности R, U5 и т.д.) описаны в тексте.

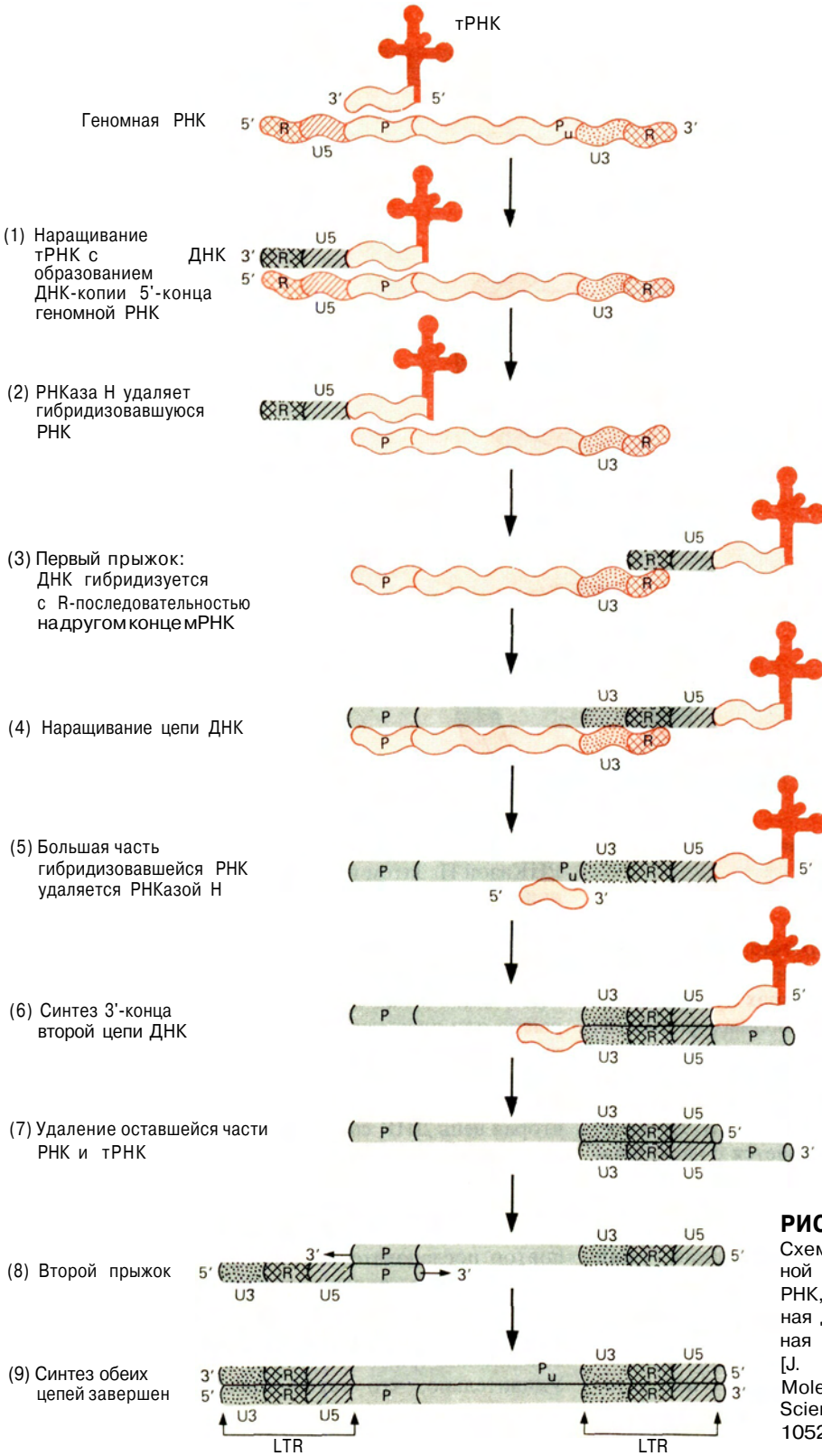
**РИС. 2.40.**

Двухцепочечная ДНК, образующаяся при обратной транскрипции ретровирусной РНК.

дуплекс ДНК с последовательностью U<sub>3</sub>RU<sub>5</sub> [длинный концевой повтор LTR (от англ. long terminal repeat)], имеющийся на обоих концах дуплекса; два кольцевых дуплекса ДНК, производных формы А; форма В с LTR-повторами на обоих концах и форма С только с одним LTR (рис. 2.40). Объяснить образование структур А, В и С при обратной транскрипции вирусной ДНК весьма непросто. Процесс, изображенный на рис. 2.41, начинается с наращивания тРНК-праймера на матрицах U<sub>5</sub> и R в направлении 3'→5' (этап 1). Затем РНКазы Н, специфичная к РНК в составе гибридного РНК-ДНК-дуплекса, расщепляет сегмент РНК этого дуплекса (этап 2). Поскольку на 3'-конце РНК имеется повтор R, новосинтезированная короткая цепь ДНК «перепрыгивает» на этот конец молекулы мРНК и спаривается там с комплементарным ей участком (этап 3). Далее происходит удлинение цепи ДНК с использованием в качестве матрицы остальной части мРНК (этап 4). К моменту завершения синтеза первой цепи

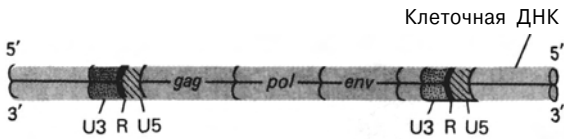
ДНК большая часть вирусной ДНК разрушается РНКазой Н. Затем в предполагаемом сайте связывания праймера (Ru) вблизи повтора U<sub>3</sub> инициируется синтез второй цепи ДНК с использованием новосинтезированной первой цепи в качестве матрицы (этап 5). Праймером для синтеза второй цепи может быть РНК, однако как идет синтез второй цепи—непрерывно или прерывисто—неизвестно. После репликации тРНК-связывающей последовательности на 5'-конце первой цепи ДНК (этап 6) тРНК, по-видимому, удаляется (этап 7). Затем новосинтезированная вторая цепь ДНК спаривается с тРНК-связывающей последовательностью первой цепи (этап 8). После удлинения 3'-концов каждой цепи завершается образование дуплекса ДНК (этап 9). Обратите внимание, что на каждом конце ДНК-дуплекса имеется прямой повтор последовательности U<sub>3</sub>RU<sub>5</sub>—LTR. Кольцевые ДНК, представленные на рис. 2.40, по-видимому, образуются либо путем лигирования концов линейной ДНК (Б), либо путем гомологичной рекомбинации (разд. 2.4) между сегментами LTR (В). Удивительно, что такая сложная последовательность реакций протекает без явного участия ферментов репликации клетки-хозяина (геликазы, прай-





**РИС. 2.41.**

Схематическое изображение обратной транскрипции ретровирусной РНК, при которой образуется линейная двухцепочечная ДНК, ограниченная LTR-последовательностями. [J. Darnell, H. Lodish, D. Baltimore, Molecular Cell Biology (New York: Scientific American Books, 1986), p. 1052.]

**РИС. 2.42.**

Структура ретровирусной ДНК, интегрированной в клеточный геном: провирус.

мазы, ДНК-связывающего белка, лигазы, топоизомеразы и т.д.).

Репликация двухцепочечной формы ретровирусной ДНК не начинается до тех пор, пока она не встроится в клеточную ДНК. Субстратом для такого интеграционного события является продукт обратной транскрипции—линейная дуплексная ДНК, представленная на рис. 2.40, А. Механизм рекомбинационного встраивания пока полностью не установлен. В результате интеграции образуется структура, представленная на рис. 2.42. При интеграции на обоих концах интегрированной вирусной ДНК утрачивается по несколько нуклеотидов в пределах LTR-последовательностей и происходит дупликация 3–10 нуклеотидов клеточной ДНК. После интеграции ретровирусная ДНК реплицируется как часть клеточной ДНК. РНК дочерних вирионов образуется в результате транскрипции интегрированных копий вирусной ДНК. Инициация синтеза РНК происходит в крайних левых точках стыковки U3R, а терминация—в крайних правых точках стыковки RU5 (рис. 2.42).

## б. Некоторые ДНК-содержащие вирусы используют для репликации обратную транскрипцию

Геном вируса гепатита В человека представлен кольцевой двухцепочечной ДНК с пробелами. Если вирус находится внутри клетки, то пробелы заполняются вирионной ДНК-полимеразой. Такая почти полноразмерная цепь ДНК играет роль матрицы, на которой синтезируется РНК. Длина этой РНК равна длине генома вируса, и она играет роль мРНК в процессе экспрессии вирусных белков и роль матрицы при обратной транскрипции с образованием дочерней вирусной ДНК. Аналогичный механизм реализуется и при репликации вируса мозаики цветной капусты и других вирусов растений.

## 2.3. РЕПАРАЦИЯ ДНК

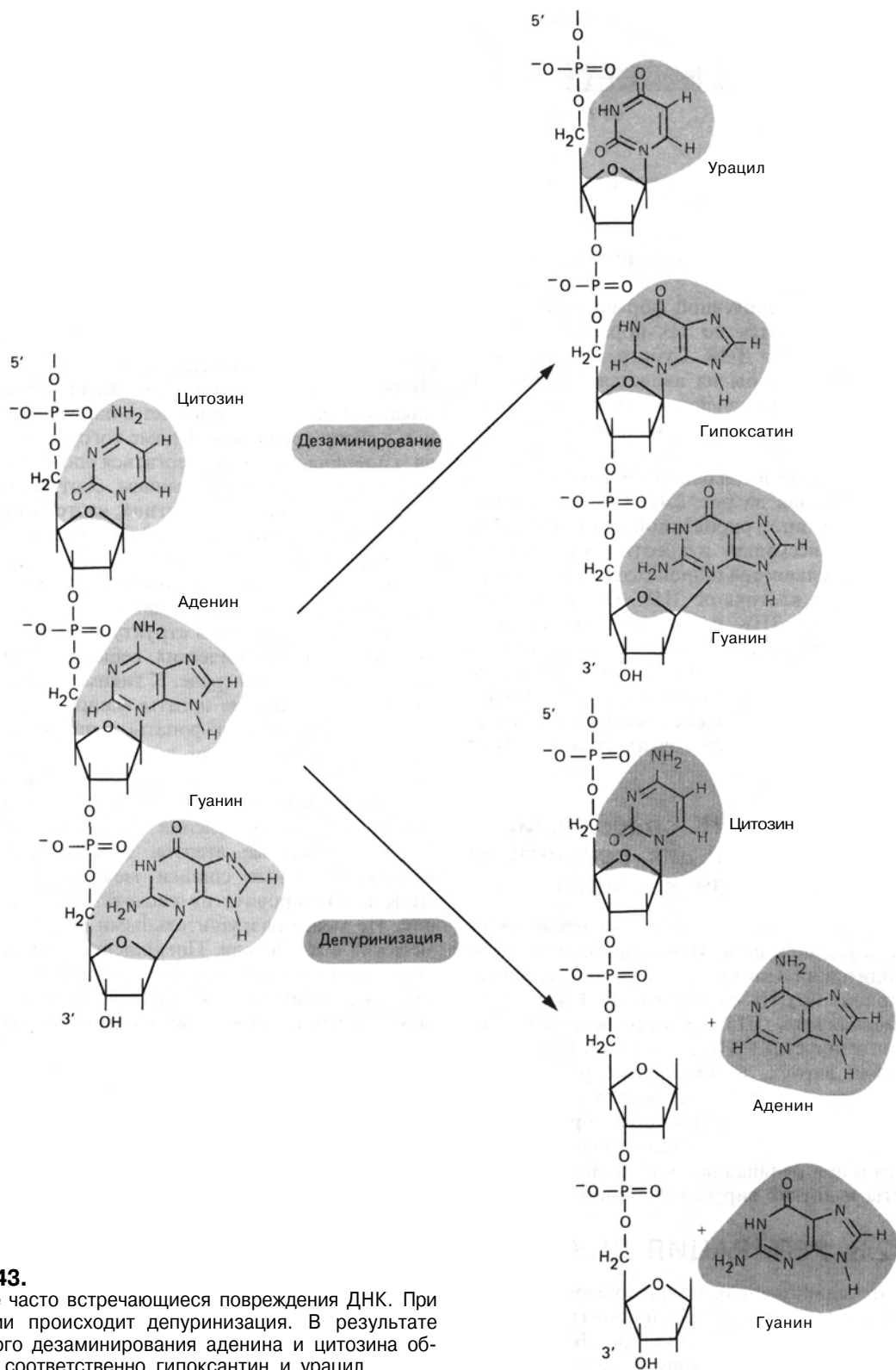
ДНК—это единственная макромолекула клетки, которая способна устранять (репарировать) повреждения, возникающие в ее структуре. Более того, в ней закодирована информация о механизмах са-

мых разнообразных репарационных процессов. Комплементарное спаривание лежит в основе не только репликации ДНК, но и процесса восстановления исходной структуры ДНК при репарации повреждений, затрагивающих остов молекулы, модификаций того или иного основания или ошибочного спаривания при рекомбинации (разд. 2.4). Одновременное повреждение обеих цепей в одном месте и двухцепочечные разрывы часто оказываются летальными для ДНК, поскольку такие дефекты репарируются лишь в редких случаях.

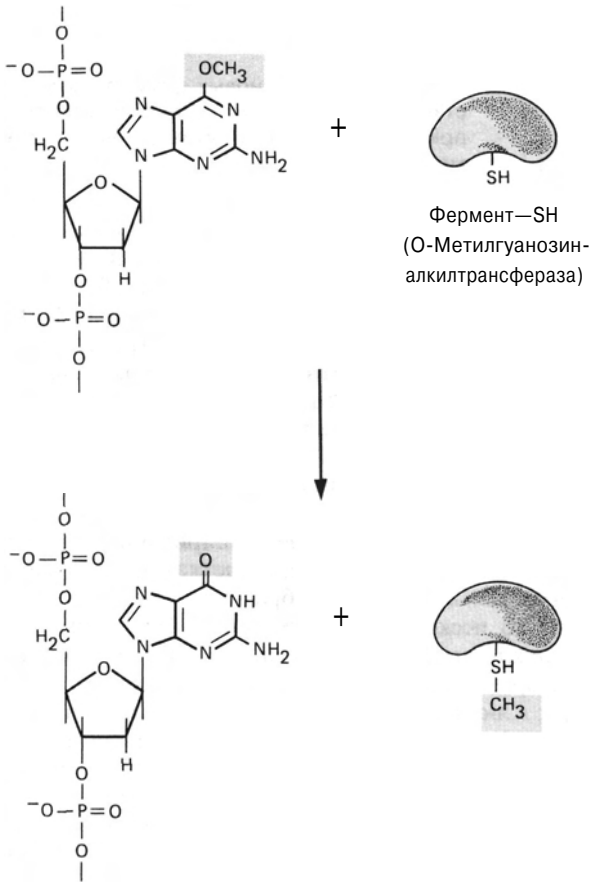
Наиболее часто происходит разрыв гликозидных связей между пурином и дезоксирибозой N (депуринизация) при повышении температуры (рис. 2.43). За сутки в клетке человека совершается от 5000 до 10000 актов депуринизации. Если не принимать никаких мер, то это приведет к нарушению репликации и экспрессии генов. Кроме того, остатки цитозина и аденина могут подвергаться спонтанному дезаминированию с образованием соответственно остатков урацила и гипоксантина; частота таких событий составляет примерно 100 на геном в сутки. Если подобные нарушения в ДНК не будут устранены до следующего раунда репликации, то они могут послужить источником мутаций.

Многие изменения в структуре ДНК происходят под действием химических веществ, присутствующих в окружающей среде. К таким веществам относятся алкилирующие агенты (например, азотистые соединения, алкилсульфонаты и нитрозомочевина), которые модифицируют предпочтительно гуаниновые остатки; соединения, встраивающиеся между соседними парами оснований и приводящие к появлению вставок и делеций во время репликации; бифункциональные агенты, способные образовывать ковалентные сшивки между двумя цепями ДНК и блокировать их расхождение при репликации. Не менее разрушительными могут быть и физические воздействия. Поглощение тиминовым или цитозиновым основанием ультрафиолетового света может приводить к образованию циклобутановых димеров между соседними пиримидинами (разд. 2.3.а); под действием ионизирующей радиации, например космических лучей, могут образовываться высокорекреационноспособные свободные радикалы, оказывающие на ДНК самые разнообразные воздействия; при облучении рентгеновскими лучами в медицинских целях в ДНК могут возникать одно- и двухцепочечные разрывы, а также другие повреждения, характерные для воздействия на ДНК свободных радикалов.

Как же ДНК противостоит таким разрушительным воздействиям? Благодаря какому механизму восстанавливается нормальная структура нуклеотидов и их последовательность прежде, чем эффект воздействия закрепится и проявится в виде мутации?

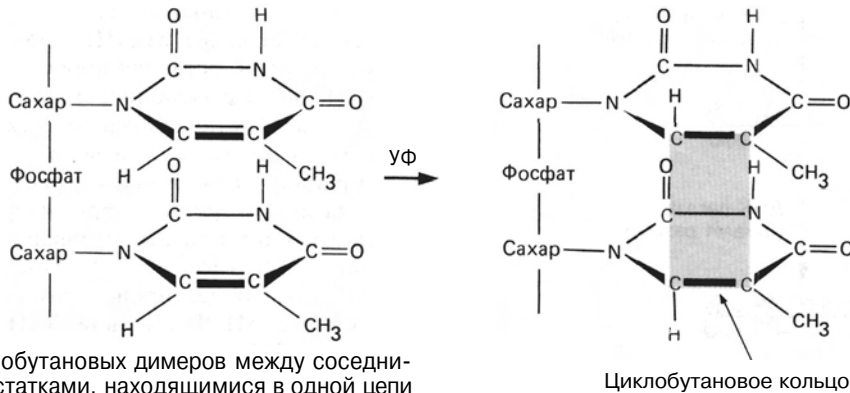
**РИС. 2.43.**

Наиболее часто встречающиеся повреждения ДНК. При нагревании происходит депуризация. В результате спонтанного дезаминирования аденина и цитозина образуются соответственно гипоксатин и урацил.



**РИС. 2.44.** Дезалкилирование  $O^6$ -метилгуаниновых остатков катализируется специфической ДНК-алкилтрансферазой.

Известны два основных типа репарационных процессов: 1) непосредственное исправление модификаций или неправильных спариваний, не требующее репликации для восстановления исходной структу-



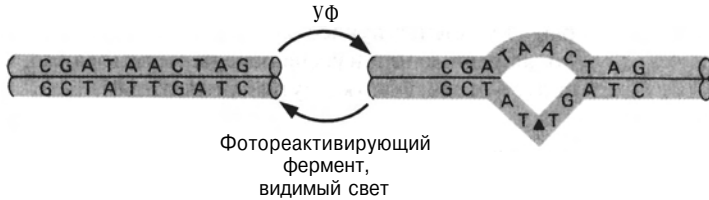
**РИС. 2.45.** Образование циклобутановых димеров между соседними тиминовыми остатками, находящимися в одной цепи ДНК.

ры; 2) удаление нуклеотидов, окружающих ошибочно спаренные или измененные пары оснований, и ресинтез этого участка путем репликации.

### а. Репарация путем прямого восстановления исходной структуры

Под действием алкилирующих агентов, например  $N$ -метил- $N$ -нитрозомочевины или  $N^1,N$ -диметилнитрозогуанидина, в ДНК образуются  $O^6$ -метил- или  $O^6$ -алкиламещенные гуаниновые остатки. Такие модифицированные остатки гуанина могут дезалкилироваться при участии ферментов, присутствующих в клетках бактерий и млекопитающих.  $O^6$ -метилгуанин-ДНК-алкилтрансфераза катализирует перенос алкильных групп на сульфгидрильные группы цистеиновых остатков фермента, при этом акцепторный белок инактивируется (рис. 2.44). Содержание алкилтрансферазы в клетках *E. coli* увеличивается в присутствии  $O^6$ -алкилгуанина, но в клетках млекопитающих аналогичной индуцибельности не наблюдается.

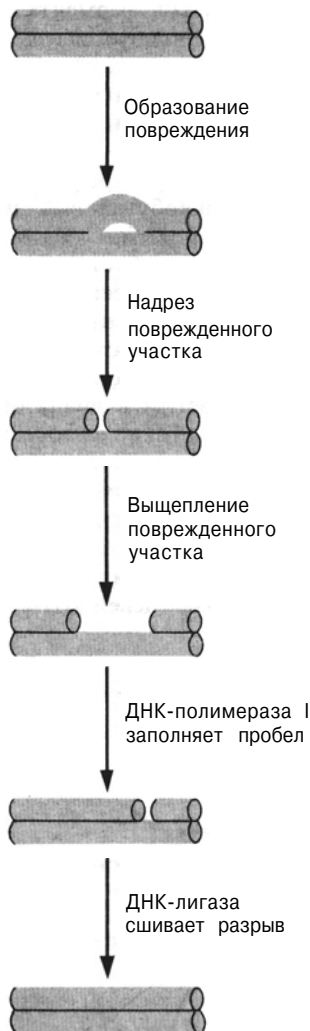
При облучении ДНК ультрафиолетовым светом в ней образуются циклобутановые димеры между соседними пиримидиновыми основаниями (рис. 2.45). Такие соединения блокируют репликацию ДНК, и для сохранения жизнеспособности клетки их необходимо удалить. Один из способов удаления пиримидиновых димеров состоит в ферментативном превращении их в мономеры при освещении раствора видимым светом в диапазоне длин волн 300–600 нм (рис. 2.46). Такие фотореактивирующие ферменты (фотолиазы) имеются у бактерий и низших эукариотических организмов, но в клетках млекопитающих они не обнаружены. Фермент образует стабильный комплекс с пиримидиновым димером и, используя энергию поглощенного им света, разрушает димер без разрыва цепей ДНК.

**РИС. 2.46.**

Образование тиминных димеров под действием УФ-света и их разрушение на свету при помощи фотореактивирующего фермента.

## б. Репарация путем замены модифицированных остатков

Замена модифицированного нуклеотида обычно происходит в четыре этапа. Во-первых, фермент распознает этот нуклеотид и надрезает полинуклеотидную цепь вблизи него либо разрывает гликозидную связь между модифицированным основани-

**РИС. 2.47.**

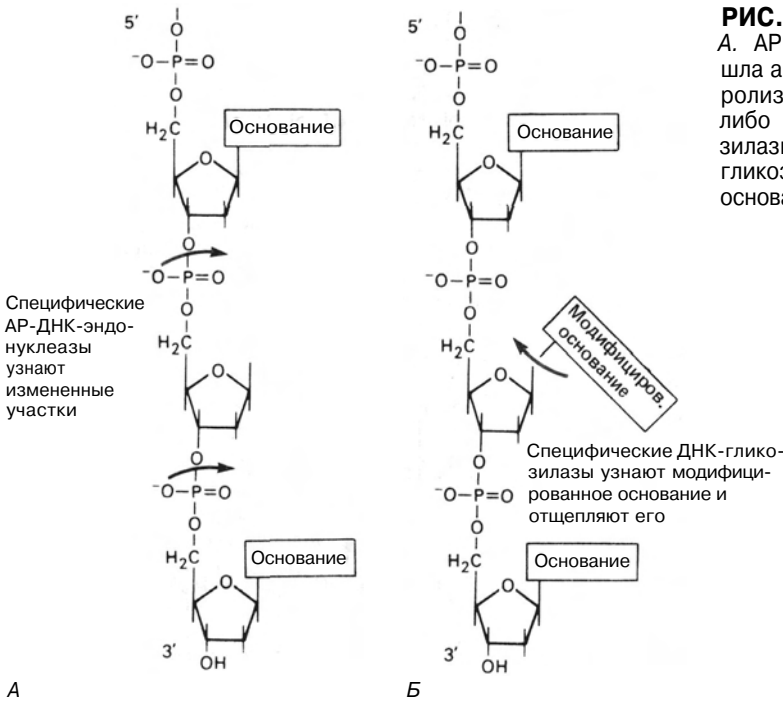
Репарация поврежденной ДНК путем замены модифицированных нуклеотидных остатков.

ем и дезоксирибозой. Во-вторых, экзонуклеаза удаляет модифицированный нуклеотид и/или соседние нуклеотиды, оставляя небольшую брешь. В-третьих, удаленный участок синтезируется заново с 3'-ОН-конца с использованием в качестве матрицы противоположной цепи. В-четвертых, концы разрыва, образовавшиеся в результате репарации, соединяются с восстановлением ковалентной целостности репарированной цепи (рис. 2.47).

Сайты, в которых произошла депуринизация или депиримидинизация, выщепляются ферментами, называемыми AP(апуриновые и апиримидиновые)-эндонуклеазами (рис. 2.48, А). В клетках про- и эукариот имеется много разнообразных AP-эндонуклеаз, часто по несколько разных типов. Некоторые AP-эндонуклеазы надрезают цепь с 3'-стороны AP-сайта, а другие расщепляют диэфирную связь с 3'-стороны; в любом случае образуются 3'-гидроксильный и 5'-фосфорильный концы. Разрыв фосфодиэфирной связи по одну или другую сторону от места нарушения (а иногда по обе стороны) позволяет экзонуклеазе удалить прилегающие остатки по обе стороны сайта расщепления и затем ресинтезировать удаленную последовательность.

В репарации N-алкилированных пуринов и других модифицированных оснований ключевую роль играют специфические N-гликозилазы – ферменты, расщепляющие гликозидную связь между модифицированными основаниями и дезоксирибозой (рис. 2.48, Б). N-гликозилазы используются и при коррекции нарушений, состоящих в спонтанном дезаминировании цитозина или аденина и превращении их соответственно в урацил или гипоксантин. Ферменты урацил-N-гликозилаза и гипоксантин-N-гликозилаза выщепляют из ДНК соответственно урацил и гипоксантин, оставляя пробелы в месте выщепления. И вновь с помощью механизма выщепления-ресинтеза восстанавливается исходная последовательность поврежденной цепи.

Урацил-N-гликозилаза играет очень важную антимутагенную роль и при исправлении ошибок, происходящих при использовании во время репликации dUTP вместо dTTP. Поскольку спаривание dUMP с матрицей происходит почти так же хорошо, как и dTMP, Pol III (или Pol I) его не распознает, и если dUMP не подвергается специфическому выщеплению, то при последующем раунде репликации в новую цепь может включиться dGMP. Таким



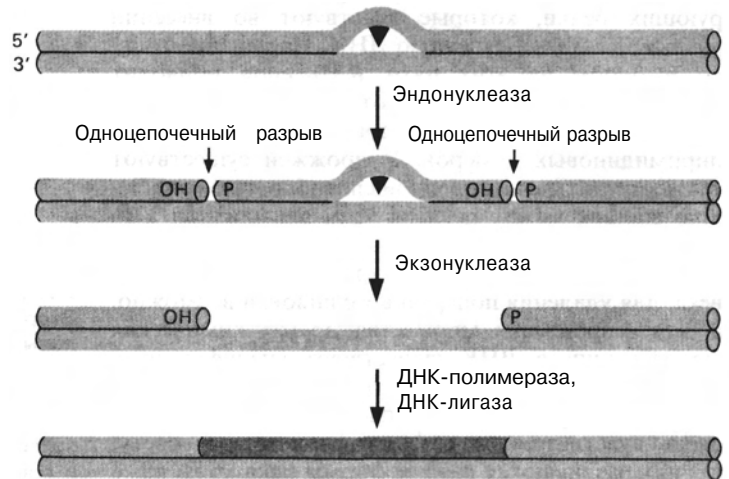
**РИС. 2.48.**

А. AP-эндонуклеазы узнают сайты, где произошла апуринизация или апириmidинизация, и гидролизуют фосфодиэфирную связь либо с 3'-, либо с 5'-стороны от этого сайта. Б. N-гликозилазы осуществляют репарацию, гидролизуя гликозидную связь между модифицированным основанием и дезоксирибозой.

образом, N-гликозилаза защищает клетки от возможных мутагенных последствий включения в цепь ДНК dUMP.

Репарация тиминных димеров может происходить и в темноте одним из следующих двух способов. В клетках *E. coli* синтезируются три полипептида, кодируемые генами *uvrA*, *uvrB* и *uvrC*, которые образуют ферментный комплекс *uvrABC*-эндонуклеазу. Этот фермент разрезает цепь ДНК, содержащую пиримидиновый димер, на расстоянии восьми фосфодиэфирных связей с 5'-стороны и четырех или пяти связей с 3'-стороны пиримидинового димера (рис. 2.49). Удаление поврежденного участка, по-

дому, катализируется геликазой, кодируемой еще одним геном, *uvrD*. В результате удаляется тиминный димер и еще примерно 12 окружающих его нуклеотидов. Образовавшийся пробел заполняется с помощью Pol I, а ДНК-лигаза завершает репарацию, соединяя два соседних основания цепи. Некоторые организмы используют для репарации поврежденных, возникающих при образовании пиримидиновых димеров, другой механизм. Репарация осуществляется при участии пиримидиновый димер-N-гликозилазы, которая создает апириmidиновый сайт. Гликозидная связь между одним из тиминных и дезоксирибозой разрезается так, что тиминный



**РИС. 2.49.**

Первый способ ферментативной репарации циклобутановых тиминных димеров. Ферментный комплекс *uvrABC* *E. coli* инициирует репарацию, гидролизуя цепь на расстоянии восьми фосфодиэфирных связей от димера с 5'-стороны и четырех-пяти связей с 3'-стороны.

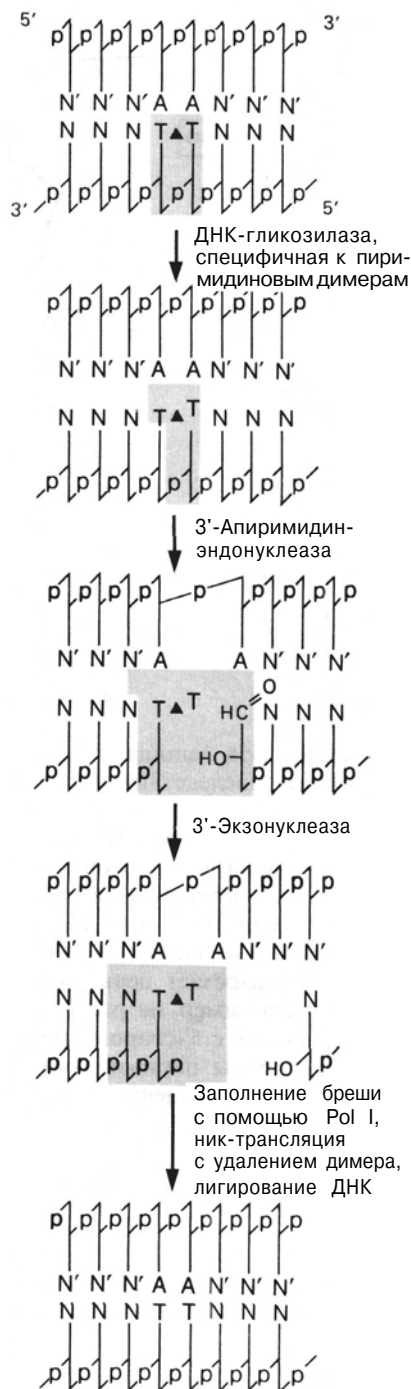
димер остается на 5'-конце разорванной цепи, а 3'-ОН-группа — на дезоксирибозе (рис. 2.50). Образовавшийся 3'-АР-конец отщепляется с помощью 3'-5'-экзонуклеазной активности ДНК-полимеразы, а затем путем ник-трансляции и лигирования удаляется нуклеотид вместе с прилежащим тиминным димером, заполняется образовавшийся пробел и сшиваются концы цепи.

## в. Значение репарации ДНК

У клеток в процессе эволюции выработался сложный механизм устранения повреждений, возникающих в ДНК под действием самых разнообразных химических и физических факторов, а также вследствие ошибок при репликации или рекомбинации. И это вполне понятно: большая часть повреждений блокирует передачу генетической информации последующему поколению, а остальные, если их не устранить, сохраняются в геномах потомков и приведут к драматическим изменениям в молекулах белков, а том числе и ферментов, необходимых для поддержания жизнедеятельности клетки. При повреждении определенных звеньев системы репарации клетки становятся особенно уязвимыми для некоторых химических и физических агентов. Например, клетки *E. coli*, у которых нарушена система внесения разрывов в ДНК при выщеплении тиминных димеров, очень чувствительны к УФ-свету. Клетки, неспособные осуществить ту или иную N-гликозилазную реакцию, гораздо больше, чем нормальные, подвержены мутагенному или летальному эффекту алкилирующих агентов или ионизирующей радиации. У клеток *E. coli*, дефектных по Pol I, существенно снижена выживаемость при облучении низкими дозами УФ-света.

У представителя низших эукариот *Saccharomyces cerevisiae* имеется по крайней мере пять генов, кодирующих белки, которые участвуют во внесении разрывов в УФ-облученную ДНК. Нарушение только в одном из этих пяти *RAD*-генов приводит к тому, что клетки утрачивают способность к внесению разрывов в ДНК и, следовательно, к удалению пиримидиновых димеров. У дрожжей существуют также мутанты с нарушенной способностью к удалению сшивок между цепями, хотя элиминация УФ-индуцированных повреждений проходит нормально. Это предполагает, что у дрожжей, как и у человека, для удаления поперечных сшивок, а возможно, и для исправления множества других химических модификаций в ДНК существуют специфические, весьма сложные механизмы репарации.

Люди, страдающие пигментной ксеродермой, очень чувствительны к ультрафиолетовому свету, и у них развиваются разные формы рака кожи даже



**РИС. 2.50.**

Второй способ ферментативной репарации циклобутановых тиминных димеров. N-гликозилаза, специфичная к пиридиноновым димерам, инициирует репарацию, разрывая гликозидную связь между одним из тиминных остатков, входящих в состав димера, и дезоксирибозой. Некоторые гликозилазы катализируют и последующее эндонуклеазное расщепление.

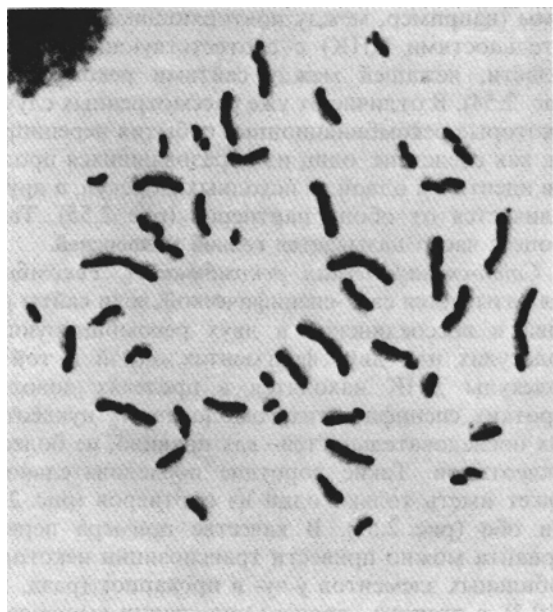
при очень слабом воздействии солнечного света. Клетки таких людей несут мутацию, сходную с *RAD*-мутацией дрожжей и проявляющуюся в том, что у них нарушена способность к выщеплению пиримидиновых димеров из УФ-облученной ДНК. Заболевание может быть обусловлено мутацией в одном из по крайней мере девяти генов, что говорит о достаточно сложном механизме репарации ДНК, содержащей тиминовые димеры, у человека. Как правило, заболевание бывает связано с неспособностью к выщеплению тиминовых димеров. Если к облученным клеткам в культуре добавить фермент, обладающий тиминдимергликозилазной и АР-эндонуклеазной активностями, то УФ-повреждения могут быть устранены.

## 2.4. РЕКОМБИНАЦИЯ ДНК

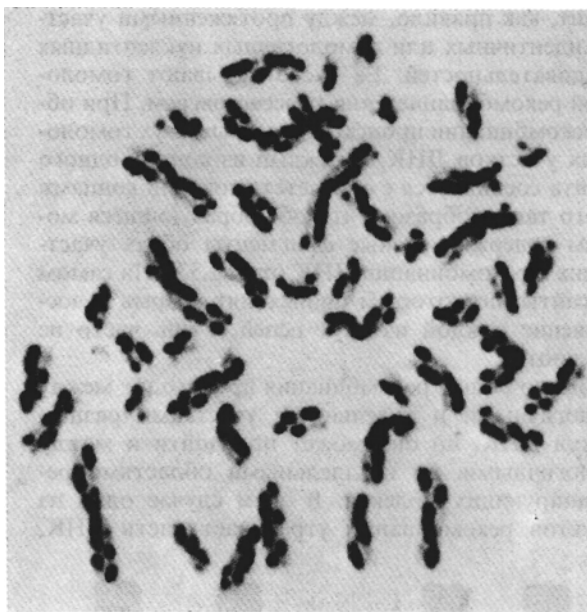
Генетическая рекомбинация включает несколько связанных между собой процессов, в результате которых в клетках или организмах, где они происходят, создаются новые комбинации элементов-

носителей генетической информации. Рекомбинация между близко расположенными гомологичными хромосомами приводит к интенсивной перетасовке отцовских и материнских генов в ходе мейоза и тем самым создает предпосылки для эволюционной проверки новых комбинаций этих генов в потомстве. Как правило, рекомбинационные события, происходящие в соматических клетках либо во время репликации ДНК, либо после нее и проявляющиеся в виде обмена сестринских хроматид (рис. 2.51), не приводят к изменению генотипа или фенотипа клетки. Однако нередко они порождают различные генотипные перестройки. Это, например, утрата, приобретение или амплификация генетических элементов и установление новых взаимосвязей между уже имеющимися, но по-новому расположенными элементами.

Если использовать молекулярные термины, то можно сказать, что генетическая рекомбинация состоит в образовании ковалентных связей между нуклеотидными последовательностями из разных областей одной и той же или разных молекул ДНК.



А



Б

**РИС. 2.51.**

Обмен сестринских хроматид. Если использовать специальную методику, то можно увидеть, как в метафазных хромосомах происходит обмен между двумя продуктами дубликации хромосом (сестринскими хроматидами). Методика вкратце состоит в следующем: клетки проходят одно деление в присутствии бромдезоксипуридина, а второе - в его отсутствие. Краситель Hoechst 33258 в комплексе с ДНК, содержащей бромдезоксипуридин, флуоресцирует значительно слабее, чем в комплексе с обычной тиминсодержащей ДНК. Поэтому

сестринские хроматиды, образовавшиеся во втором клеточном цикле, имеют более интенсивную окраску. Обмен сестринских хроматид выявляется по неравномерности интенсивности флуоресценции. А. Хромосомы из нормальных периферических лимфоцитов человека. Б. Хромосомы из периферических лимфоцитов больных с синдромом Блума. [S.A. Latt, R. Schreck, Amer. J. Hum. Genet., 32 (1980), p. 294.] Микрофотографии любезно предоставлены S.A. Latt.



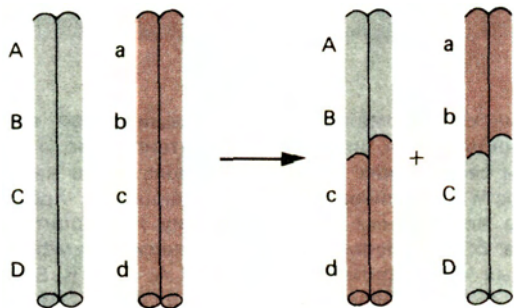
Все клетки и многие вирусы содержат информацию о синтезе ферментов, предназначенных не только для репарации повреждений в собственной ДНК, но и ферментов, осуществляющих рекомбинацию. На самом деле некоторые ферменты, участвующие в репликации и репарации ДНК, играют ключевую роль и при рекомбинации. В этом разделе мы рассмотрим механизмы некоторых рекомбинационных процессов и ферменты, которые их катализируют. Особое внимание будет обращено на рекомбинацию у бактерий и фагов, поскольку у них эти процессы довольно хорошо изучены. Несмотря на то что генетические и морфологические аспекты рекомбинации в эукариотических клетках известны, на молекулярном уровне здесь многое остается неясным.

### а. Типы рекомбинации

Существуют три типа рекомбинации: **общая**, или **гомологичная**, **сайт-специфическая** и **случайная**, или **негомологичная**.

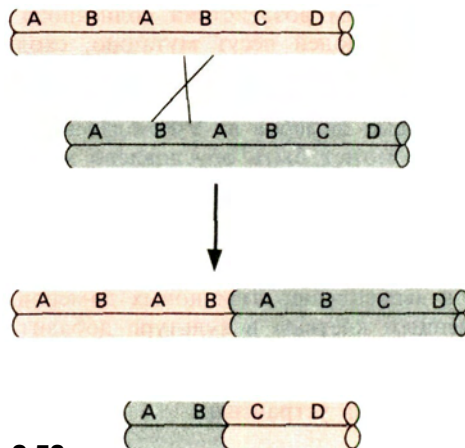
**Общая рекомбинация.** Общая рекомбинация происходит, как правило, между протяженными участками идентичных или гомологичных нуклеотидных последовательностей. Ее часто называют гомологичной рекомбинацией или **кроссинговером**. При общей рекомбинации происходит разрыв двух гомологичных участков ДНК, и каждый из концов одного сегмента соединяется с соответствующими концами другого таким образом, что обе образующиеся молекулы содержат разные фрагменты обоих участвующих в рекомбинации ДНК (рис. 2.52). На самом деле сайты, по которым происходит разрыв и воссоединение каждой из двух цепей, очень часто не совпадают.

Обычно общая рекомбинация происходит между гомологичными и аллельными участками разных молекул ДНК, но она может произойти и между гомологичными, но неаллельными областями рекомбинирующих молекул. В этом случае один из продуктов рекомбинации утрачивает часть ДНК,



**РИС. 2.52.**

Гомологичная рекомбинация, или кроссинговер.

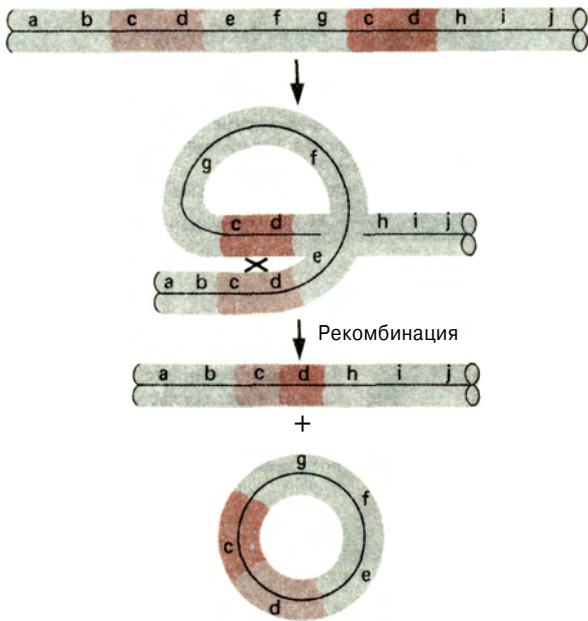


**РИС. 2.53.**

Неравный кроссинговер.

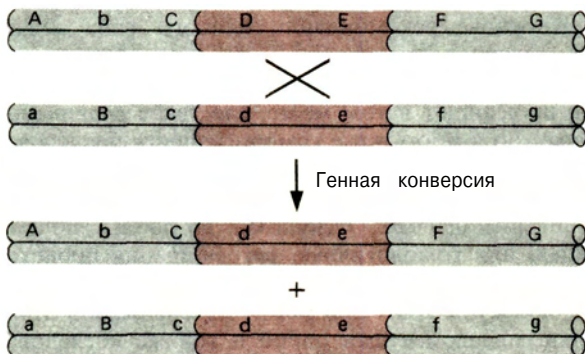
а другой приобретает «лишний» сегмент. Такой процесс получил название **неравного кроссинговера** (рис. 2.53). Иногда рекомбинация происходит между неаллельными участками одной и той же хромосомы (например, между повторяющимися последовательностями ДНК) с соответствующей потерей области, лежащей между сайтами рекомбинации (рис. 2.54). В отличие от уже рассмотренных случаев некоторые рекомбинационные события нереципрочны; как следствие, один из образовавшихся продуктов идентичен одной из исходных молекул, а другой отличается от обоих партнеров (рис. 2.55). Такой процесс часто называется **генной конверсией**.

**Сайт-специфическая рекомбинация.** Рекомбинация называется сайт-специфической, если сайты разрыва и воссоединения в двух рекомбинирующих молекулах или двух фрагментах одной и той же молекулы ДНК находятся в пределах довольно коротких специфических гомологичных нуклеотидных последовательностей—как правило, не более 25 нуклеотидов. Такие короткие последовательности могут иметь только один из партнеров (рис. 2.56) или оба (рис. 2.57). В качестве примера первого варианта можно привести транспозиции некоторых мобильных элементов у эу- и прокариот (разд. 10.2 и 10.3), а второго—процесс интеграции-выщепления ДНК фага  $\lambda$  из хромосомы *E. coli* (разд. 2.4.г). С помощью сайт-специфической рекомбинации происходят запрограммированные перестройки хромосомной ДНК при смене типов спаривания у дрожжей; она ответственна также за разнообразие антител (разд. 10.6). По-видимому, общая рекомбинация между любыми парами гомологичных последовательностей осуществляется с помощью одного и того же комплекса ферментов; с другой стороны, для каждого случая сайт-специфической рекомбинации необходим свой набор ферментов.

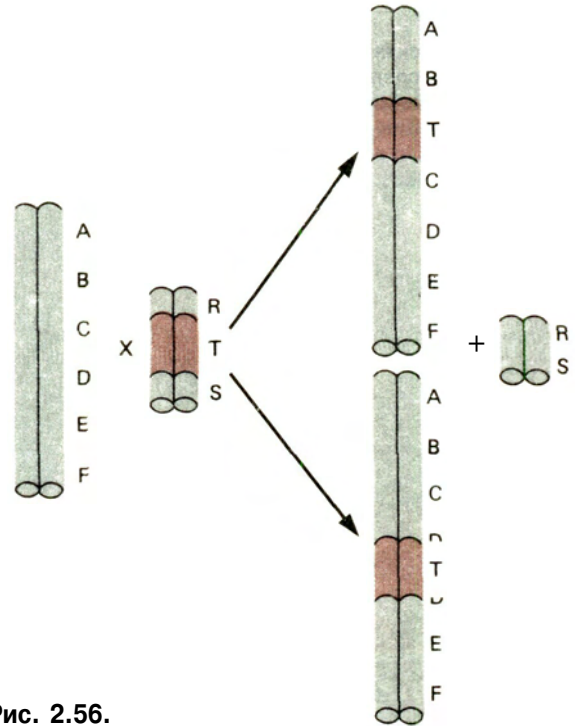


**РИС. 2.54.**  
Внутримолекулярная рекомбинация.

*Негомологичная рекомбинация.* Рекомбинация между негомологичными нуклеотидными последовательностями происходит в клетках прокариот и дрожжей достаточно редко, а в клетках млекопитающих — весьма часто. К негомологичной рекомбинации можно отнести процесс случайного встраивания вирусной или плазмидной ДНК в ДНК клеток животных, в результате чего в реплицирующихся геномах паповавирусов появляется множество делеций и дупликаций. Концы разорванной ДНК могут соединиться, даже если они негомологичны. В некоторых случаях рекомбинация происходит между последовательностями, содержащими несколько гомологичных пар оснований, или между короткими



**РИС. 2.55.**  
Нереципрокная рекомбинация, или генная конверсия.



**Рис. 2.56.**  
Сайт-специфическая рекомбинация, при которой специфический сайт имеется только в одном участвующем в рекомбинации фрагменте ДНК. Специфическими сайтами рекомбинации являются концы участка Т, а сам участок может рекомбинировать со случайными последовательностями ДНК, расположенными, например, между участками В и С (*вверху*) или в пределах участка D (*внизу*).

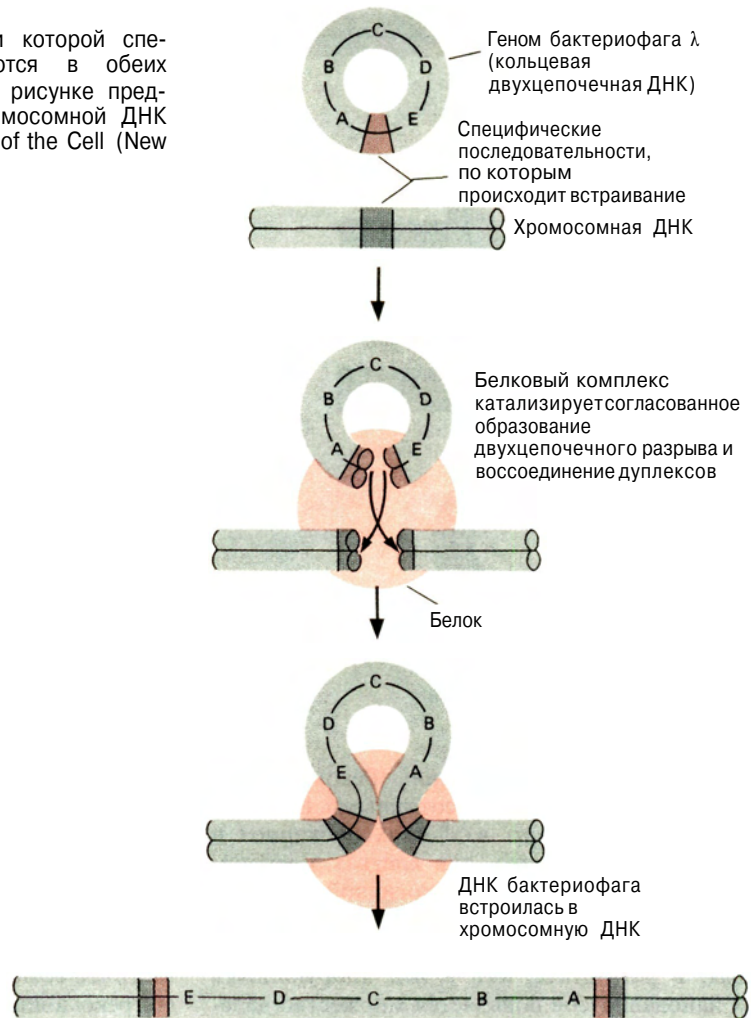
частично гомологичными участками. Но, как правило, рекомбинирующие сегменты не имеют гомологичных последовательностей.

### б. Общая рекомбинация между гомологичными молекулами ДНК

*Общая рекомбинация при согласованном внесении разрывов и воссоединении цепей двух спиралей ДНК с образованием протяженных гетеродуплексных областей.* Чтобы могла произойти рекомбинация между двойными спиральями, представленная на рис. 2.52, каждая из четырех цепей должна быть разорвана и затем соединена с новым партнером. Этапы этого процесса показаны на рис. 2.58. Соответствующие цепи обоих линейных гомологичных дуплексов ДНК надрезаются и свободные концы одной спирали спариваются с комплементарными участками другой (*a-g*). Перекрест стабилизируется сшиванием концов донорных цепей со свободными концами реципиентных спиралей (*d*). Точка перекреста обменивающихся цепей перемещается вдоль

**РИС. 2.57.**

Сайт-специфическая рекомбинация, при которой специфические последовательности имеются в обеих участвующих в рекомбинации ДНК. На рисунке представлена интеграция ДНК фага  $\lambda$  с хромосомной ДНК *E. coli*. [B. Alberts et al., Molecular Biology of the Cell (New York: Garland, 1983), p. 248.]



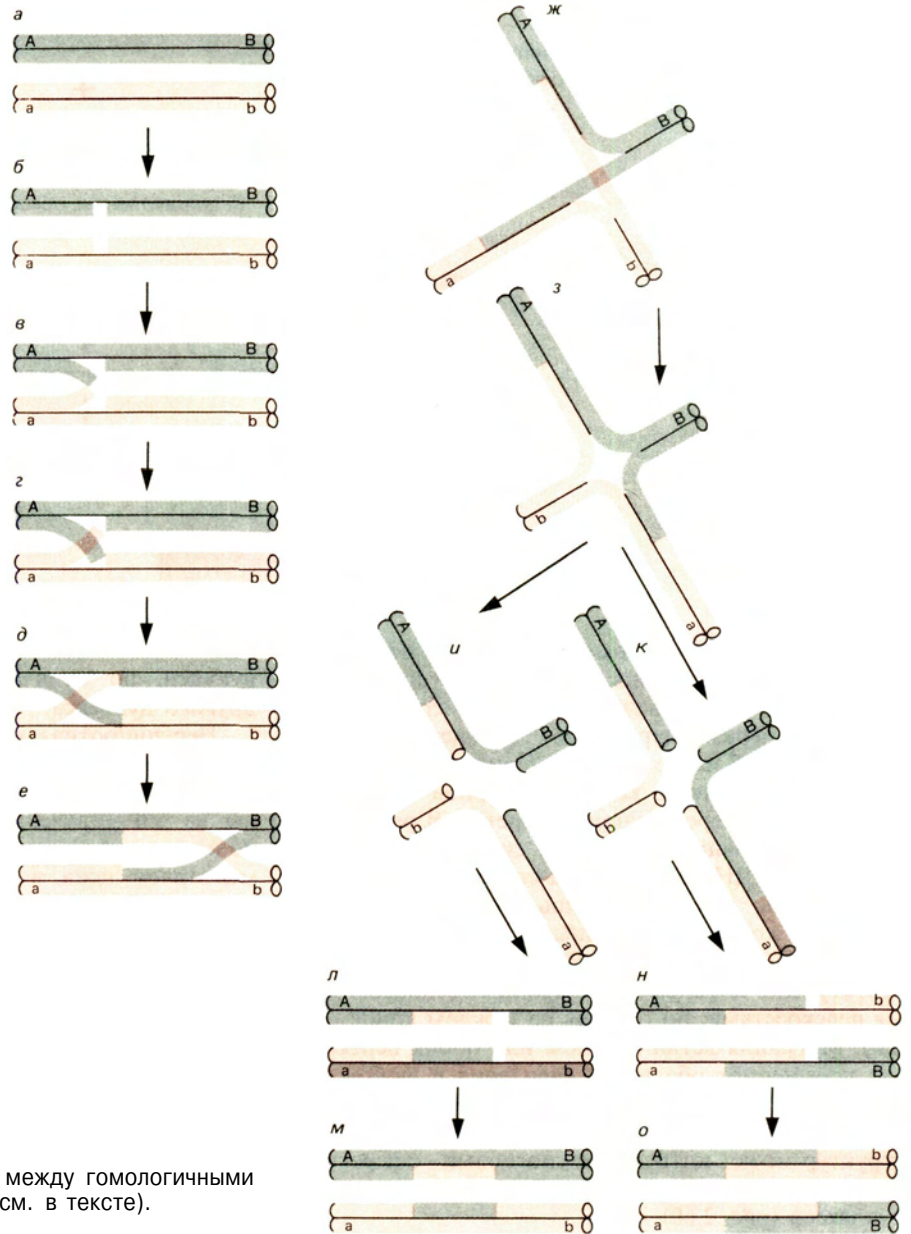
спиралей — процесс, называемый **миграцией ветви** (*e*). При этом происходит одновременное расхождение цепей исходных спиралей и их реассоциация с новыми партнерами с образованием дочерних дуплексов. Структуры *d* и *e*, а также *ж* называются **структурами Холлидея** по имени исследователя, впервые их предложившего.

Структуры Холлидея могут переходить в рекомбинантные двойные спирали путем внесения разрыва и воссоединения цепей двумя альтернативными способами. Один способ состоит в разрезании и воссоединении перекрещивающихся цепей. Два реципрокных продукта *л* и *м* могут образоваться, если разрыв и последующее воссоединение цепей произойдут в точке перекреста в структурах *e* и *д* или по линии пересечения четырех цепей в изомерной структуре Холлидея *и*. Размер обменивающихся фрагментов зависит от расстояния, на которое про-

изошла миграция ветви до акта рекомбинации. Альтернативные продукты *н* и *о* образуются в том случае, если структура Холлидея *з* переходит в результате разрыва в *к*.

В основе рекомбинации данного типа лежит гомологичное спаривание цепей, принадлежащих двум разным спиральям ДНК, поэтому скорее всего она произойдет в том месте, где такое спаривание возможно а priori и где гомологичность последовательностей достаточно велика, чтобы могла произойти миграция ветви в рамках структуры со скрестившимися цепями. Отсюда можно понять, почему общая, или гомологичная, рекомбинация происходит также между двумя повторами в пределах одной молекулы ДНК или между аллельными и неаллельными элементами одной и той же последовательности в двух разных хромосомах (рис. 2.53 и 2.54).

В ходе миграции ветви при спаривании цепей,

**РИС. 2.58.**

Этапы процесса рекомбинации между гомологичными дуплексными ДНК (объяснения см. в тексте).

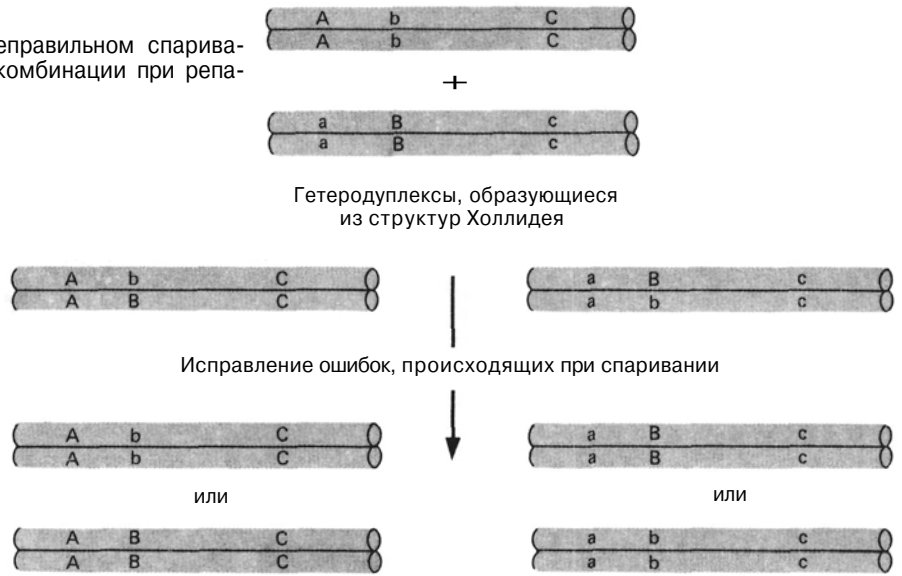
принадлежащих разным спиральям, образуются **гетеродуплексы**. В таких гетеродуплексах в пределах сегмента между сайтом начала образования структуры Холлидея и сайтом кроссинговера может содержаться по одному или более ошибочно спаренных оснований (рис. 2.58). Они удаляются так же, как любые модифицированные основания при репарации ДНК (разд. 2.3.6). Однако, поскольку удалено может быть любое из ошибочно спаренных оснований, в обеих рекомбинантных спиральях в данном сайте могут оказаться одинаковые пары осно-

ваний, т.е. рекомбинация для этого сайта окажется нераспространенной (рис. 2.59). Таким образом, каждая из рекомбинантных спиралей может быть похожа на любой из начальных дуплексов в тех позициях, где исходно они различались.

*Общая рекомбинация с образованием двухцепочечного разрыва.* Альтернативный механизм общей рекомбинации включает образование двухцепочечного разрыва в одном из дуплексов-партнеров (рис. 2.60). Далее с помощью экзонуклеаз в месте разрыва образуется брешь. При спаривании 3'-одноцепочеч-

**РИС. 2.59.**

Продукты, образующиеся при неправильном спаривании в процессе гомологичной рекомбинации при репарации гетеродуплексов.



ного конца бреши с комплементарной цепью интактной спирали в последней образуется петля. Размер этой петли увеличивается по мере того, как ДНК-полимераза наращивает 3'-конец «вклинившейся» цепи. В итоге другой одноцепочечный конец бреши спаривается с комплементарной последовательностью в перемещающейся петле. В результате такого спаривания образуется система «праймер-матрица», и ДНК-полимераза синтезирует недостающую цепь, заполняя брешь. Лигирование двух растущих концов с исходными цепями приводит к образованию двойной структуры Холлидея (т.е. структуры, в которой две спирали объединены двумя перекрестами, по одному на каждом конце бреши). Миграция ветви в одном или обоих перекрестках передвигает оба места сцепления в любом направлении, при этом в участках, фланкирующих брешь, могут возникать ошибки. Разделение таких структур может идти двумя способами — с перекрестом и без него (рис. 2.58), с образованием четырех дуплексов (рис. 2.60).

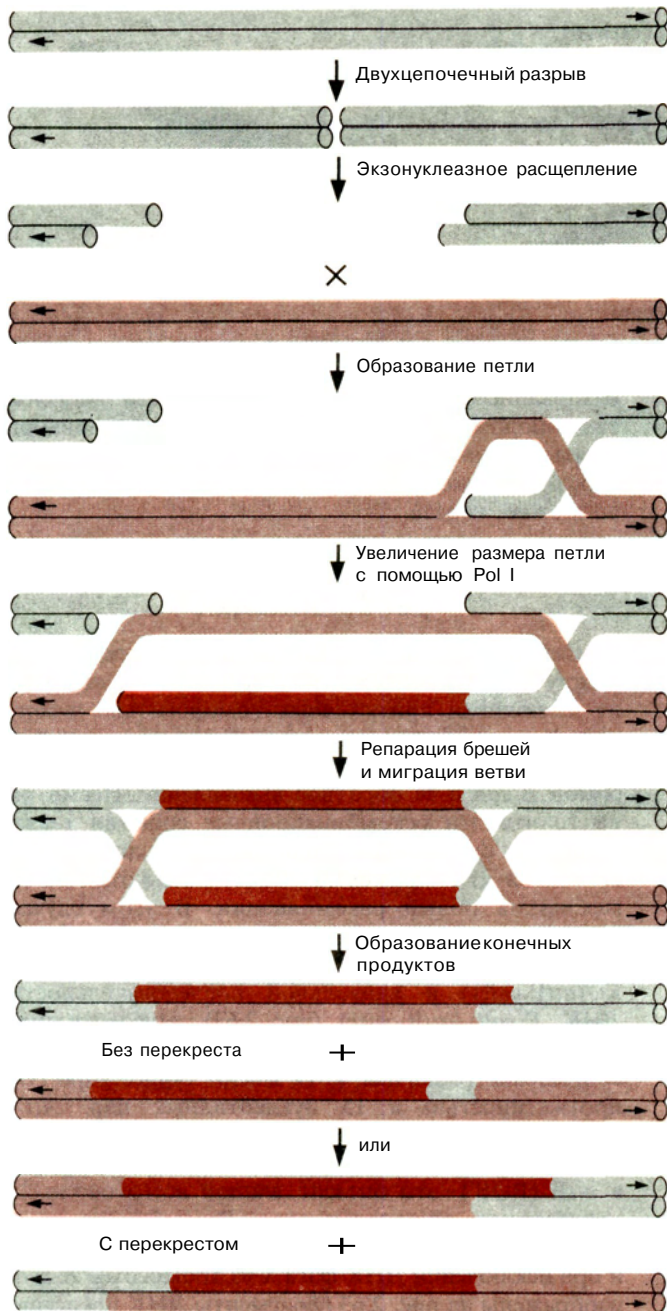
Необходимо отметить некоторые особенности этого механизма. Образование ошибочных пар (гетеродуплексов) в районах, фланкирующих брешь, обуславливает получение как реципрокных, так и нерципрокных рекомбинаций между генетическими маркерами. Если двухцепочечный разрыв происходит вблизи (или в пределах) участка, где между спиралью имеются различия (замены оснований, делеции, вставки, инверсии и т.п.), то рекомбинанты унаследуют нуклеотидную последовательность партнера, у которого разрыва не происходило. Этот механизм объясняет многие случаи генной конверсии, особенно те, в которых протяженная последо-

вательность одного дуплекса замещается соответствующей, но отличающейся последовательностью другого дуплекса.

Нереципрокная общая рекомбинация используется и при репарации некоторых повреждений ДНК. Например, если тиминные димеры не были удалены из УФ-облученной ДНК до того, как к ним подошла репликативная вилка, то синтез комплементарной цепи в этом участке не может быть завершен (рис. 2.61). Поскольку тиминные димеры, находящиеся напротив бреши, не могут быть выщелены, остается один путь для спасения хроматиды — использовать генетическую информацию гомологичной сестринской хроматиды и заполнить брешь. Для этого применяется такой же механизм, как для репарации брешей (рис. 2.60).

## в. Ферменты, участвующие в общей рекомбинации

В общей рекомбинации участвуют два специфических фермента и еще несколько ферментов, катализирующих также процессы репликации и репарации ДНК. Энзимология общей рекомбинации изучена только для некоторых прокариотических организмов, в частности *E. coli* и ее фагов. Один из специфических ферментов, необходимых для успешной гомологичной рекомбинации, называется *recA*-белком. Он катализирует обмен одиночными цепями, используя энергию гидролиза АТФ до АДФ и неорганического фосфата (рис. 2.62). *RecA*-зависимое внедрение одноцепочечных ДНК в дуплекс — первый этап рекомбинационного процесса в рамках обеих схем Холлидея (рис. 2.58) и механизма с обра-

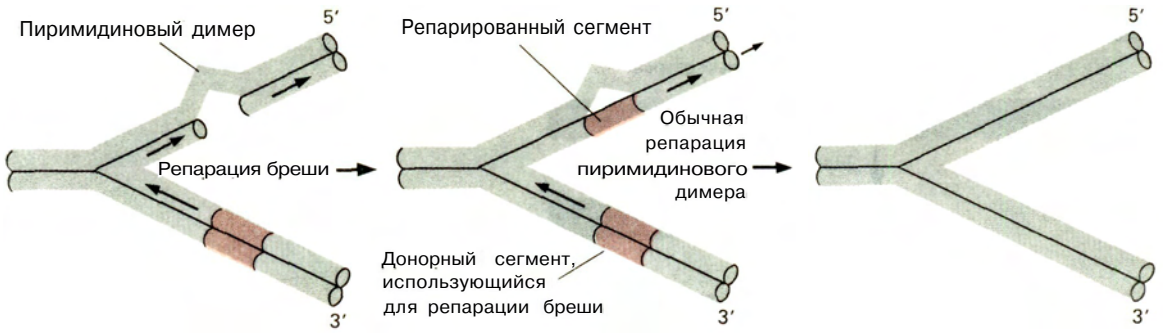
**РИС. 2.60.**

Общая рекомбинация с образованием двухцепочечного разрыва.

зованием двухцепочечных разрывов (рис. 2.60). Второй фермент, состоящий из трех отдельных субъединиц (B, C и D) и поэтому называемый гесBCD-нуклеазой, обладает эндо- и экзонуклеазной, а также геликазной активностями. Механизм его действия до конца не установлен, однако известно, что гесBCD-нуклеаза индуцирует разрывы в дуплексах ДНК и благодаря присущей ей геликазной активности вместе с гесA инициирует рекомбинационный

процесс. Идентифицирован также фермент, разрезающий узлы в структурах Холлидея; при его участии образуются липкие концы, соединяемые лигазой.

В общей рекомбинации участвуют также геликазы и белки, связывающиеся с одноцепочечной ДНК (SSB; от англ. single strand binding); оба они необходимы для обеспечения процесса миграции ветви. Как известно, перемещению цепей во время миграции ветви способствует Pol I, а в воссоединении

**РИС. 2.61.**

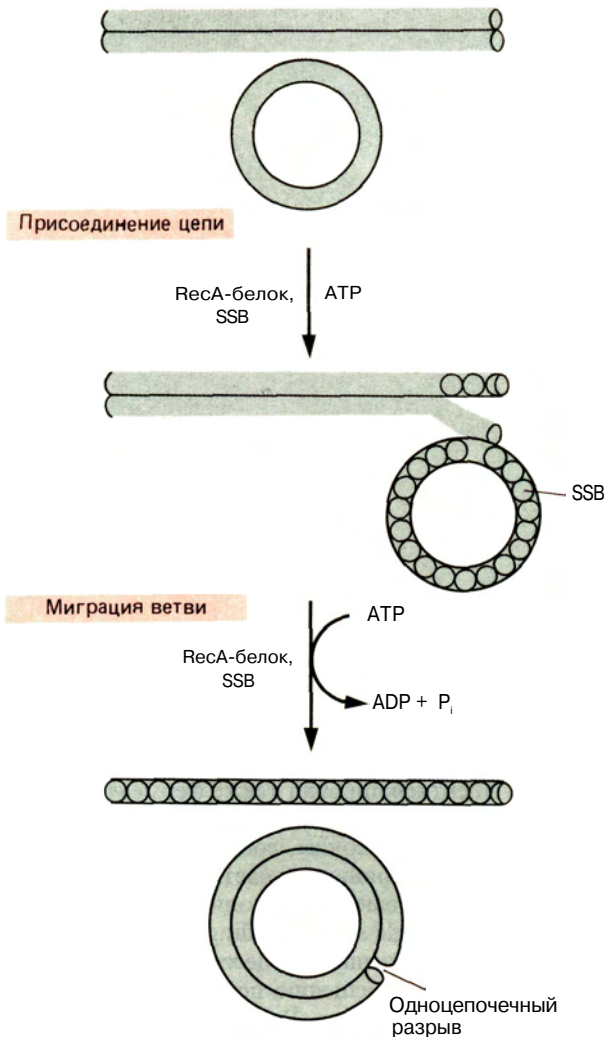
Нереципрокная гомологичная рекомбинация как один из этапов репарации ДНК в месте образования пиримидиновых димеров.

разорванных цепей участвует ДНК-лигаза. Для снятия топологических ограничений при раскручивании спирали и для распутывания перекрученных структур, по-видимому, нужны топоизомераза типа I и, возможно, гиразы.

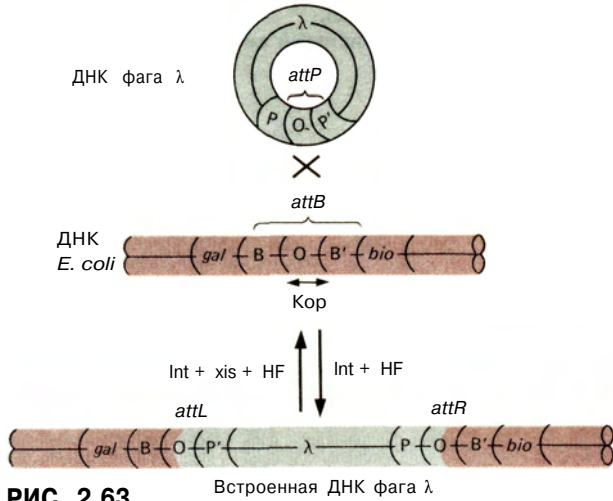
### г. Сайт-специфическая рекомбинация

Сайт-специфическая рекомбинация происходит между специфическими сегментами дуплексов ДНК, не имеющими протяженных гомологичных участков. Характерным примером такой рекомбинации служит интеграция кольцевой ДНК фага  $\lambda$  с хромосомой *E. coli* и ее обратное выщепление. Несмотря на то что эти рекомбинационные события также включают разрыв и воссоединение двух спиральных сегментов ДНК, их механизм абсолютно отличен от механизма общей рекомбинации. В этом случае рекомбинация происходит в пределах специфической нуклеотидной последовательности ДНК фага  $\lambda$  (*attP*-сайт) и уникальной последовательности ДНК *E. coli* (*attB*-сайт) (рис. 2.63). Нуклеотидные последовательности *attP*- и *attB*-сайтов (рис. 2.64) совершенно различны, хотя имеют общее ядро (O) протяженностью в 15 нуклеотидных пар. *AttP* (POP') простирается на 150 нуклеотидов влево (P) и на 75 нуклеотидов вправо (P') от общего ядра, а *attB* (BOB')—это сегмент длиной всего около 25 нуклеотидов, включая и ядро. Рекомбинационные события, происходящие как при интеграции, так и при исключении ДНК фага  $\lambda$  из хромосомы *E. coli*, изображены на рис. 2.63 (фрагменты ДНК представлены в виде линейных структур).

Поскольку нуклеотидные последовательности,

**РИС. 2.62.**

RecA-зависимая общая рекомбинация ДНК *E. coli*. Процесс ускоряется присоединением белка, связывающегося с одноцепочечной ДНК (SSB). [M. Cox, I.R. Lehman, Proc. Natl. Acad. Sci. USA, **78** (1981), p. 3433.]



**РИС. 2.63.**

Встраивание кольцевой ДНК фага  $\lambda$  в хромосому *E. coli*. Сайт *attB* расположен между генами *gal* и *bio* ДНК *E. coli*. Для встраивания необходимы белок, кодируемый фаговым геном *int*, и бактериальный белок HF. Для выщепления ДНК фага  $\lambda$  необходимы эти два белка и еще один фаговый белок, кодируемый геном *xis*.

фланкирующие *attP*- и *attB*-сайты слева (*attL*) и справа (*attR*), для этих сайтов различаются, механизм рекомбинационного выщепления ДНК фага  $\lambda$  из ДНК *E. coli* должен отличаться от механизма их рекомбинационной интеграции (рис. 2.63). И действительно, для рекомбинации между *attL* и *attR* при

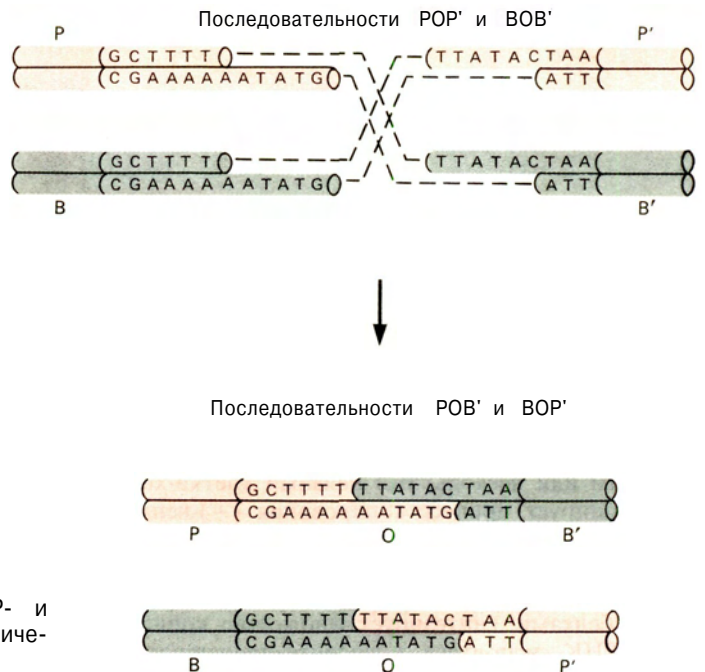
исключении фаговой ДНК помимо белка *Int* необходимы фаговый белок *xis* и клеточный белок HF. Процесс рекомбинационного выщепления, по-видимому, имеет некоторое сходство с процессом интеграции, но роль указанных трех белков, особенно белка *xis*, все еще изучается.

## 2.5. РЕПЛИКАЦИЯ

Образование РНК-содержащих вирусов происходит путем репликации их РНК, тогда как все клеточные РНК образуются в результате транскрипции ДНК (гл. 3). Здесь мы рассмотрим лишь процесс репликации вирусной РНК как таковой, не касаясь связи репликации с жизненным циклом соответствующих вирусов.

За исключением ретровирусов (разд. 2.2), репликация РНК в основном повторяет процесс репликации ДНК. Цепи РНК также удлиняются на один нуклеотид за один акт в направлении 5'→3' путем присоединения рибонуклеотидтрифосфата к 3'-ОН-концу растущей цепи. Как и при репликации ДНК, порядок расположения нуклеотидов определяется комплементарным копированием матрицы, в данном случае обязательно цепи РНК (рис. 2.65). Ферменты, осуществляющие этот процесс, называются РНК-зависимыми репликазами.

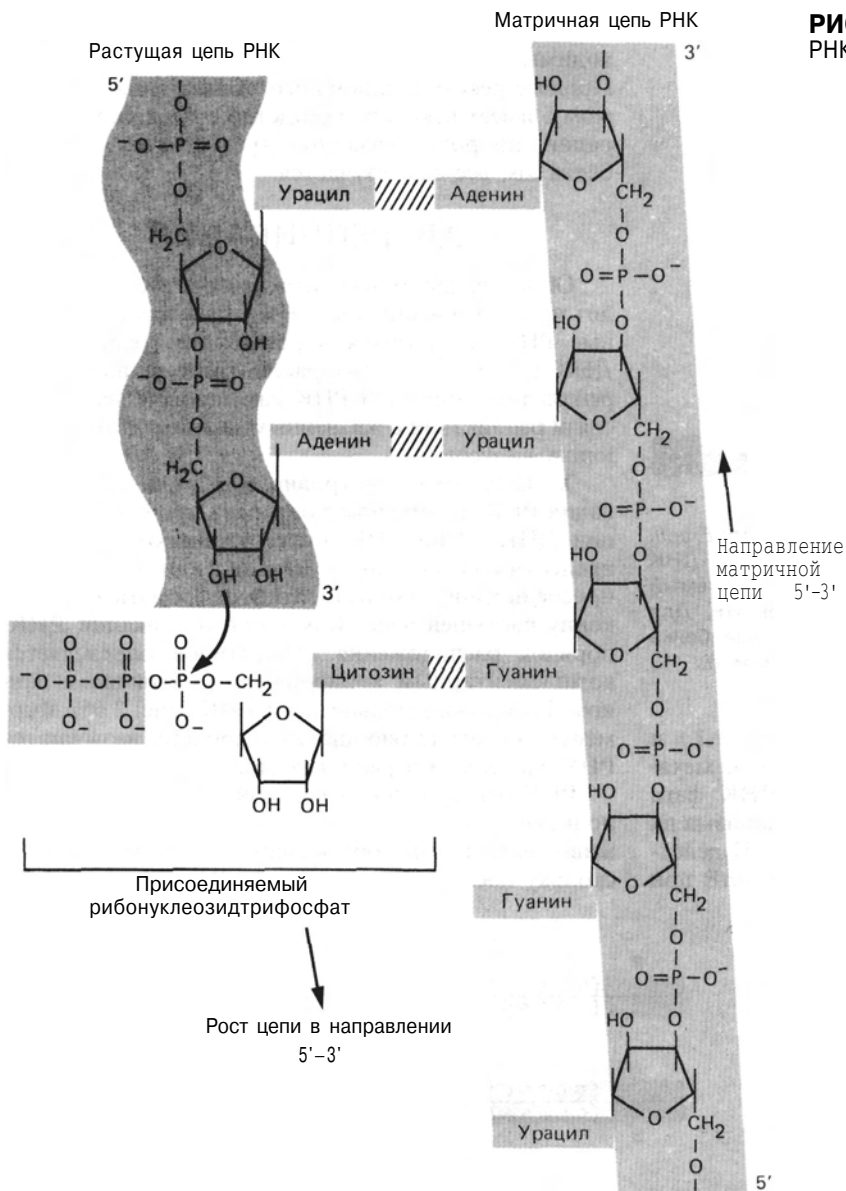
РНК бактериальных вирусов R17 и MS2, а также полиовирусов и вируса Синдбис, инфицирующих животных, всегда обозначается знаком (+), поскольку последовательность их РНК-геномов иден-



**РИС. 2.64.**

Обычные нуклеотидные последовательности *attP*- и *attB*-сайтов, по которым происходит сайт-специфическая рекомбинация.



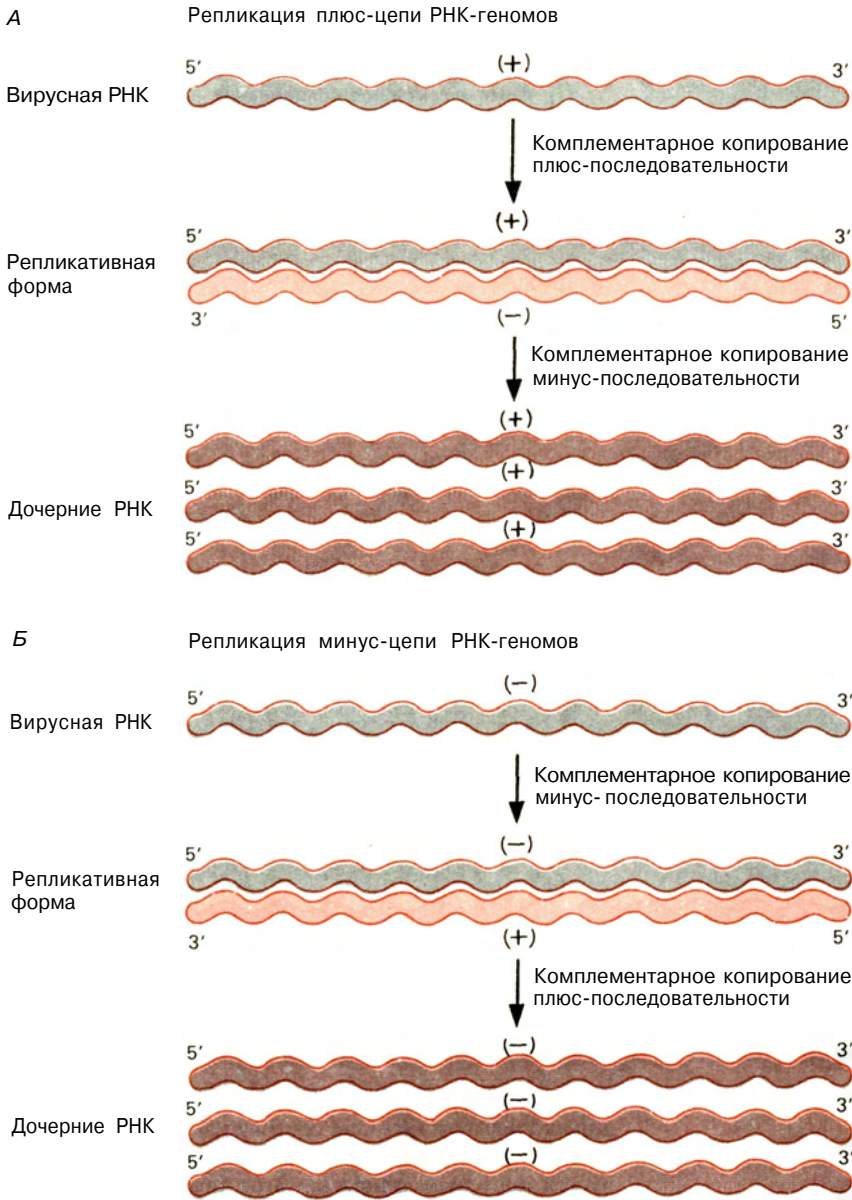


**РИС. 2.65.**  
РНК-зависимый синтез РНК.

тична последовательности мРНК. Таким образом, геном инфицирующего вируса может служить в качестве мРНК и содержит информацию о синтезе некоторых, если не всех, вирусных белков. Специфическая репликаза, кодируемая геномом вируса и образующаяся вскоре после инфекции, связывается с одним или несколькими белками клетки-хозяина и инициирует процесс копирования (+)-цепи с ее 3'-конца с образованием полной (-)-цепи, ассоциированной с (+)-цепью-матрицей (рис. 2.66, А). Затем та же репликаза, возможно вместе с другими вирусными белками, синтезирует множество копий (+)-цепи РНК, используя новосинтезированную (-)-

цепь в качестве матрицы. С единственной (-)-матричной цепи РНК синтезируется одновременно несколько новых (+)-цепей (рис. 2.67). У многих вирусов копирование (+)- или (-)-цепей инициируется путем спаривания рибонуклеозидтрифосфата с первым нуклеотидом на 3'-конце матричной цепи.

(+)-Цепь РНК полиовируса реплицируется способом, аналогичным представленному на рис. 2.66, А, но его геномная РНК имеет одну особенность — к ее 5'-концу с помощью фосфоэфириной связи присоединен белок через тирозиновый остаток. У (-)-цепи, синтезируемой в первом раунде репликации, этот белок отсутствует. Однако все

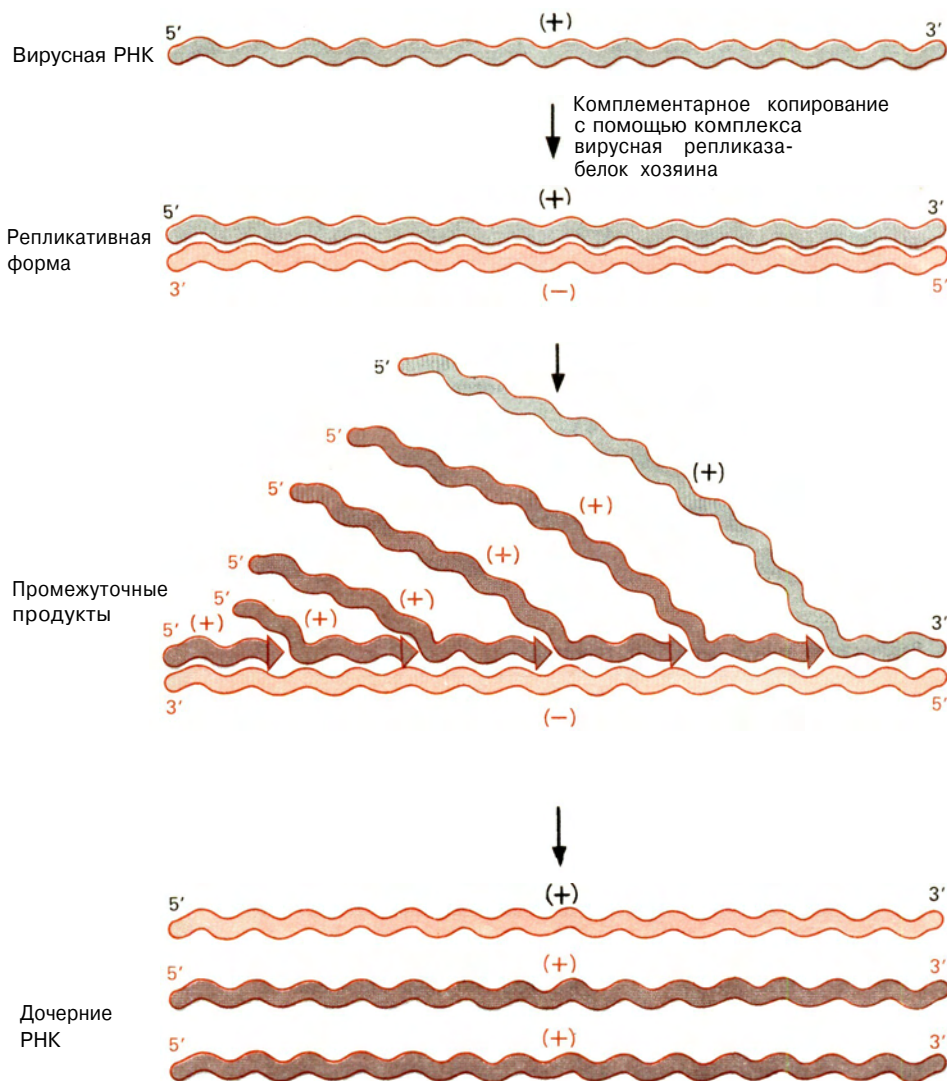
**РИС. 2.66.**

Репликация плюс- (А) и минус- (Б) цепей вирусной РНК по механизму РНК-зависимого синтеза РНК.

(+)-цепи, образующиеся на (-)-цепи как на матрице, содержат 5'-концевой белок, даже когда их длина еще очень мала. Это позволяет предположить, что репликаза полиовирусов использует гидроксильную группу тирозина в качестве праймера при инициации новых (+)-цепей, в то время как (-)-цепи фермент может инициировать *de novo*.

Геном некоторых вирусов (в том числе вируса везикулярного стоматита) представлен одной цепью РНК, нуклеотидная последовательность которой комплементарна, а не идентична последовательности мРНК. Такие геномы обозначают знаком (-). Существуют также вирусы (вирусы гриппа), геном

которых состоит из множественных (-)-цепей РНК, при этом каждый такой сегмент соответствует одному или двум вирусным генам. Если геном вируса представлен одной или более (-)-цепями РНК, то в самом вирионе содержится комплекс геномной РНК со специфической репликазой. После проникновения в клетку комплекс активируется, и на (-)-цепях как на матрицах синтезируются (+)-цепи (рис. 2.66, Б). Новосинтезированные (+)-цепи играют роль мРНК для синтеза новых или дополнительных репликационных белков и роль матрицы при синтезе дочерних (-)-цепей. Новые (-)-цепи синтезируются по механизму, представленному на рис.

**РИС. 2.67.**

Одновременный синтез нескольких плюс-цепей РНК на единственной РНК-матрице.

2.67 для синтеза (+)-цепей вирусов.

Геном некоторых вирусов (в частности, реовирусов) представлен двухцепочечной РНК, содержащей и (+)-, и (-)-цепи. Репликация у таких вирусов также инициируется ферментами, содержащимися в самом вирионе. Эти ферменты копируют (-)-цепь дуплекса РНК консервативным путем (т.е. исходная двухцепочечная РНК сохраняется, а новые (+)-цепи отделяются). Белки-репликазы, транслируемые с этих (+)-цепей мРНК, синтезируют затем комплементарные (-)-цепи, которые, объединяясь с (+)-цепями, образуют дочерние двухцепочечные РНК вируса.

Имеющиеся данные о структуре и функции РНК-репликаз весьма ограничены. Наиболее детально

изучена РНК-репликаза, синтезируемая РНК-содержащим колифагом Q $\beta$ . Она состоит из четырех идентичных полипептидов, причем только один из них кодируется вирусным геномом. Этот полипептид не способен реплицировать (+)-цепи вируса, он может катализировать только ограниченную реакцию присоединения рибонуклеотидов. Остальные три белка, входящие в состав РНК-репликазы фага Q $\beta$ , являются компонентами хозяйского аппарата синтеза белка, и их точная функция в процессе репликации РНК неизвестна. Инициация и репликация, осуществляемые РНК-репликазой, кодируемой вирусами животных, также зависит от активности белков, образуемых клеткой-хозяином. Отличительная особенность РНК-репликаз состоит в их необычной способности специфически узнавать нуклеотидные последовательности на 3'-концах как (+)-, так и (-)-цепей, что гарантирует инициацию синтеза новых цепей.

# АППАРАТ ЭКСПРЕССИИ ГЕНОВ И ЕГО ЛОГИКА

В основу концепции взаимосвязи генотипа и фенотипа была положена теория «один ген—один фермент». Однако эта теория не учитывала молекулярную природу носителей генетической информации и способ передачи этой информации от генов к белкам. Не содержала она и никаких предположений о механизме регуляции экспрессии генов. Прогресс в этих областях наметился сразу после того, как были установлены следующие ключевые положения: 1) показано, что гены—это участки ДНК; 2) расшифрована молекулярная структура ДНК; 3) установлено, что структура и функция белков определяются их уникальной аминокислотной последовательностью; 4) обнаружено, что передача информации от ДНК к белкам осуществляется с помощью РНК; 5) разработаны относительно простые бактериальные генетические системы, позволяющие связать мутационные изменения в генах со структурными изменениями в соответствующих белках; 6) разработаны системы для изучения синтеза РНК и сборки белков *in vitro*.

ДНК и информационно связанные с ней молекулы РНК и белков можно представить как одномерные структуры, состоящие из множества мономерных единиц. В ДНК информация записана в виде последовательно расположенных дезоксирибонуклеотидных пар, образующих длинные цепи, а в РНК—в виде последовательности рибонуклеотидов. Уникальность белков определяется линейной последовательностью аминокислот в их полипептидной цепи. Природу информационной связи между ДНК и белками удалось понять, проводя генетические и биохимические исследования мутаций в данном гене и сопоставляя их со специфическими изменениями в аминокислотной последовательности соответствующего белка. Благодаря этим исследованиям была выявлена также коллинеарность последовательностей нуклеотидов в ДНК и аминокислот в белках. Наличие такой корреляции подразумевало существование **генетического кода**, связывающего нуклеотидные и аминокислотные последовательности обоих полимеров. Но какова природа этого кода? И в частности—как последовательности ДНК, состоящие всего из четырех нуклеотидов, могут детерминировать белковые последовательности, состоящие не менее чем из 20 аминокислот? Какие химические процессы управляют трансляцией генетического кода и как они регулируются при формировании

свойственных разным клеткам и организмам фенотипов?

Сейчас природа генетического кода известна, составлен словарь, переводящий нуклеотидную последовательность в аминокислотную. Установлены и основные особенности различных этапов экспрессии генов и их регуляции, хотя многие молекулярные детали еще ждут своего разъяснения. В этой главе мы рассмотрим, как устроен генетический словарь, и опишем механизмы, используемые клеткой для экспрессии генетической информации. Основное внимание будет уделено прокариотам, но мы вкратце остановимся на соответствующих процессах и у эукариот. Последние данные об этих процессах у эукариот приведены в гл. 8.

## 3.1. ОСНОВНЫЕ ПОЛОЖЕНИЯ ПРОЦЕССА ЭКСПРЕССИИ ГЕНОВ

**Экспрессия генов**—это процесс реализации информации, закодированной в структуре ДНК, на уровне РНК и белков. Прежде чем переходить к детальному описанию и анализу этих процессов, мы вкратце рассмотрим суть экспрессии генов—ее механизм и регуляцию.

### а. Транскрипция ДНК в РНК

Экспрессия всех генов начинается с транскрипции их нуклеотидной последовательности, т.е. перевода ее на язык РНК. При этом определенный участок одной из двух цепей ДНК используется как матрица для синтеза РНК путем комплементарного спаривания оснований. В результате транскрипции генов, в которых закодирована структурная информация о белках, образуются молекулы мРНК; другие гены кодируют молекулы РНК, являющиеся частью аппарата, необходимого для трансляции мРНК с образованием белков. У прокариот, например *E. coli*, ДНК транскрибируется с помощью одного фермента—**ДНК-зависимой РНК-полимеразы**, который участвует в синтезе всех типов РНК. В отличие от прокариот эукариоты имеют три разные ДНК-зависимые РНК-полимеразы, каждая из которых ответственна за транскрипцию генов, кодирующих разные типы клеточных РНК (гл. 8). Несмотря на то что механизмы синтеза РНК и матричного копирования для всех РНК-полимераз идентичны,

каждый фермент узнает в матрице ДНК свои характерные особенности, определяющие сайты инициации, терминации и регуляции транскрипции.

## б. Соответствие между нуклеотидными триплетами и аминокислотами

Генетический код устанавливает соответствие между нуклеотидной последовательностью данной мРНК и аминокислотной последовательностью синтезируемой на ней полипептидной цепи. Размер единиц кодирования и сами эти единицы, однозначно задающие ту или иную аминокислоту (**кодоны**), практически одинаковы у всех живых организмов. Более того, основные принципы и механизмы перевода генетических посланий также универсальны.

Генетический словарь содержит 64 кодона, каждый из которых образован тремя последовательными нуклеотидами (триплетами). 61 из 64 кодонов детерминируют 20 аминокислот, обнаруженных в белках, один определяет начало большинства последовательностей, кодирующих белки, и три обозначают окончания этих последовательностей.

Отличительной особенностью генетического кода является то, что каждый кодон кодирует только одну аминокислоту, т.е. код однозначен. Следовательно, зная словарь и правила пользования им, можно перевести нуклеотидную последовательность мРНК в определенную аминокислотную последовательность. Но генетический код является вырожденным. Это означает, что одной аминокислоте могут соответствовать несколько кодонов. Вырожденность генетического кода приводит к тому, что нельзя однозначно перевести аминокислотную последовательность данного белка в нуклеотидную последовательность соответствующей мРНК.

## в. Расшифровка кода с помощью тРНК

Аминокислоты не взаимодействуют с соответствующими им кодонами непосредственно. Каждая аминокислота вначале связывается с адаптером — родственной тРНК, и образуемая при этом аминоацил-тРНК узнает «родственный» кодон путем комплементарного спаривания оснований. Таким образом, декодирование осуществляется с помощью спаривания оснований триплетных кодонов мРНК с триплетными **антикодонами** в аминоацил-тРНК.

Присоединение аминокислот через карбоксильные группы к родственным тРНК катализируют ферменты, называемые **аминоацил-тРНК-синтетазами**. При связывании тРНК с аминокислотой карбоксильная группа последней активируется, и в ре-

зультате образование пептидных связей становится энергетически выгодным. Энергия же, необходимая для активации аминокислоты при присоединении ее к тРНК, поступает от гидролиза АТФ.

Присоединение аминокислот к родственным тРНК осуществляется с помощью специфических ферментов. Так, тирозил-тРНК-синтетаза присоединяет L-тирозин только к тем тРНК, которые могут спариться с тирозиновым кодоном. Аналогично лейцил-тРНК-синтетаза катализирует присоединение лейцина к молекулам тРНК, которые узнают кодоны лейцина. Таким образом, специфичность декодирования обеспечивается двумя реакциями: точным присоединением каждой аминокислоты к родственной ей тРНК и комплементарным спариванием антикодонов аминоацил-тРНК с соответствующими им кодонами в мРНК.

## г. Правильная инициация трансляции

Имеются три «рамки считывания», при которых может осуществляться перевод последовательных нуклеотидных триплетов мРНК в аминокислоты. Правильная инициация трансляции чрезвычайно важна для точной расшифровки генетического кода. Выбор рамки считывания зависит от того, какое сочетание из трех последовательных нуклеотидов выбрано в качестве первого кодона. Ниже приведены три возможные рамки считывания для последовательности GUACGUAAGUAAGUAUGGACGUA:

Рамка считывания 1	GUA CGU AAG UAA GUA UGG ACG
Рамка считывания 2	.G UAC GUA AGU AAG UAU GGA CGU
Рамка считывания 3	.GU ACG UAA GUA AGU AUG GAC GUA

Обычно аминокислотной последовательности кодируемой полипептидной цепи соответствует только одна из рамок. Следовательно, должен существовать какой-то способ инициации трансляции с правильной рамкой считывания. У всех организмов, изученных к настоящему времени, — бактерий, вирусов и эукариот — правильная рамка считывания определяется с помощью механизма, распознающего специфический кодон, который детерминирует концевую аминокислоту синтезируемого белка. Почти всегда таким кодоном является триплет AUG, отвечающий метионину. Поэтому образующийся полипептид неизменно содержит на N-конце метионин, но при последующем удалении аминоконцевой последовательности на N-конце конечного белкового продукта оказывается аминокислота, находящаяся изначально внутри синтезированной полипептидной последовательности. В рассмотренном выше примере кодон AUG, с которого может начаться транскрипция, содержит рамка считывания 3.

#### д. Трансляция кодонов и соединение аминокислот

Последовательное спаривание разных аминоксил-тРНК с кодонами мРНК и рост полипептидной цепи осуществляются с помощью целой серии взаимно согласованных реакций. Одним из главных участников этого в высшей степени скоординированного процесса является **рибосома** – особый мультиферментный комплекс, состоящий из нескольких видов РНК и множества белков. Кроме того, целая армия ферментов и различных факторов катализирует мириады химических событий, необходимых для успешного синтеза белка.

Рибосомы, несущие особую инициаторную метионил-тРНК, находят инициаторный кодон в мРНК, AUG, и связываются с ним. Затем с рибосомой связывается аминоксил-тРНК, соответствующая второму кодону, и при участии рибосомной ферментативной активности остаток метионина соединяется со второй аминокислотой, все еще связанной со «своей» тРНК. В результате образуется дипептидил-тРНК. По мере продвижения рибосомы по цепи мРНК и считывания каждого последующего кодона полипептидная цепь удлиняется на одну аминокислоту за один шаг. Элонгация прекращается в тот момент, когда рибосома достигает одного из трех терминирующих кодонов. Завершенная полипептидная цепь тотчас же высвобождает последнюю тРНК, и происходит разделение рибосомы и мРНК.

#### е. Регуляция экспрессии генов на разных этапах образования РНК и белка

Клетки про- и эукариот обладают способностью к дифференциальной регуляции экспрессии генов. Так, при определенных условиях многие гены вообще не экспрессируются, а степень экспрессии других различается на несколько порядков. Изменение условий может привести к активации «молчавших» ранее генов и репрессии активно работавших. Подобная способность позволяет клеткам приспособить свои фенотипы к самым разнообразным условиям окружающей среды и физиологическим воздействиям. Дифференцированная экспрессия одного генома у многоклеточных организмов обуславливает развитие огромного множества типов клеток, имеющих специфические функции, из одной или нескольких зародышевых клеток.

Экспрессия генов, как правило, регулируется на уровне образования РНК. Обычно регулируемым этапом является инициация транскрипции, при этом регуляция осуществляется либо с помощью **репрессорных белков**, предотвращающих транскрипцию,

либо с помощью **активаторных**, необходимых для ее начала. В первом случае транскрипция начинается только после того, как инактивируется репрессорный белок. Во втором ген транскрибируется лишь тогда, когда белок-активатор находится в соответствующем функциональном состоянии. В регуляции транскрипции генов участвуют не только репрессорные и активаторные белки. В некоторых случаях сами белки – продукты генной экспрессии – оказываются регуляторами транскрипции собственных генов. Эффективность транскрипции зависит также от конформационного состояния ДНК или РНК. Кроме того, регуляция синтеза РНК может осуществляться путем контроля скорости ее элонгации или с помощью «стоп-сигнала» в транскрибируемой последовательности, который может остановить транскрипцию гена. Модификация и/или процессинг, которые могут предшествовать образованию зрелой функциональной РНК, также регулируются.

Экспрессия генов может регулироваться и на уровне трансляции мРНК с образованием белков. И в этом случае специфическая регуляция, как правило, осуществляется на начальном этапе декодирования. Однако контроль может осуществляться и на разных этапах сборки полипептидной цепи. Более того, синтез тех белков, которые претерпевают посттрансляционные модификации или транспортируются к местам своего назначения внутри клетки, может регулироваться на каждом из этих этапов.

Позднее, когда мы проанализируем эти процессы подробнее, мы увидим, что механизмы регуляции экспрессии генов весьма разнообразны, многочисленны и очень сложны. И хотя многим из них присущи общие черты, тонкие механизмы регуляции всегда уникальны для данного гена, определенного физиологического состояния организма и условий окружающей среды. Анализ регуляторных механизмов бактериальных систем позволил выявить широкий спектр способов регуляции и координации экспрессии генов. Однако исследование механизма контроля экспрессии генов в клетках эукариот только начинается, а процессы, ответственные за дифференцировку многоклеточных организмов, пока остаются невыясненными.

### 3.2. ТРАНСКРИПЦИЯ: ПЕРЕДАЧА ИНФОРМАЦИИ О НУКЛЕОТИДНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ДНК НА УРОВЕНЬ РНК

В этом и следующем разделах мы рассмотрим некоторые аспекты переноса информации о нуклеотидной последовательности ДНК на уровень РНК – процесса, ответственного за синтез всех типов кле-

точных РНК как у про-, так и у эукариот. Подавляющее число пионерских работ, в которых изучалась транскрипция,— природа соответствующих реакций и их субстраты, ферментативный аппарат, сигнальные нуклеотидные последовательности, определяющие, какие области ДНК должны транскрибироваться, некоторые способы процессинга, превращающего **первичные транскрипты** в зрелые молекулы РНК,—было выполнено на прокариотических системах. Параллельное проведение генетических и биохимических экспериментов позволило исследовать ферменты, участвующие в транскрипции, и механизм самого этого процесса. Предпринимаемые в то же время усилия по изучению транскрипционного и регуляторного аппаратов у эукариот были сильно затруднены и гораздо менее успешны главным образом из-за того, что компоненты их транскрипционного аппарата—ДНК в форме хроматина и РНК-полимеразы—были слабо охарактеризованы. Кроме того, была неизвестна природа транскрипционных единиц, а применение генетических подходов для их определения было невозможно.

Ситуация резко изменилась с появлением методов молекулярного клонирования. Сейчас многие гены, составляющие различные типы транскрипционных единиц, выделены, секвенированы и даже соответствующим образом модифицированы с целью исследования их функций. Более того, использование некоторых современных подходов позволило по-новому посмотреть на транскрипционный аппарат самых разных организмов—от дрожжей до человека. Имеются в виду методы введения ДНК в культуры клеток млекопитающих (гл. 5) и даже в клетки целого организма животных, применяемые наряду с традиционными методами исследования очищенных транскрипционных систем *in vitro*. Одновременно был достигнут прогресс в установлении структуры хроматина и свойств эукариотических РНК-полимераз. В этой главе мы лишь вкратце рассмотрим системы транскрипции и соответствующие сигналы у эукариот, в первую очередь останавливаясь на их отличии от прокариотических систем. Подробные данные о структуре генов, механизме транскрипции и непосредственно связанных с ней посттранскрипционных событиях, происходящих при синтезе РНК у эукариот, приведены в гл. 8.

### а. Синтез РНК на ДНК-матрице

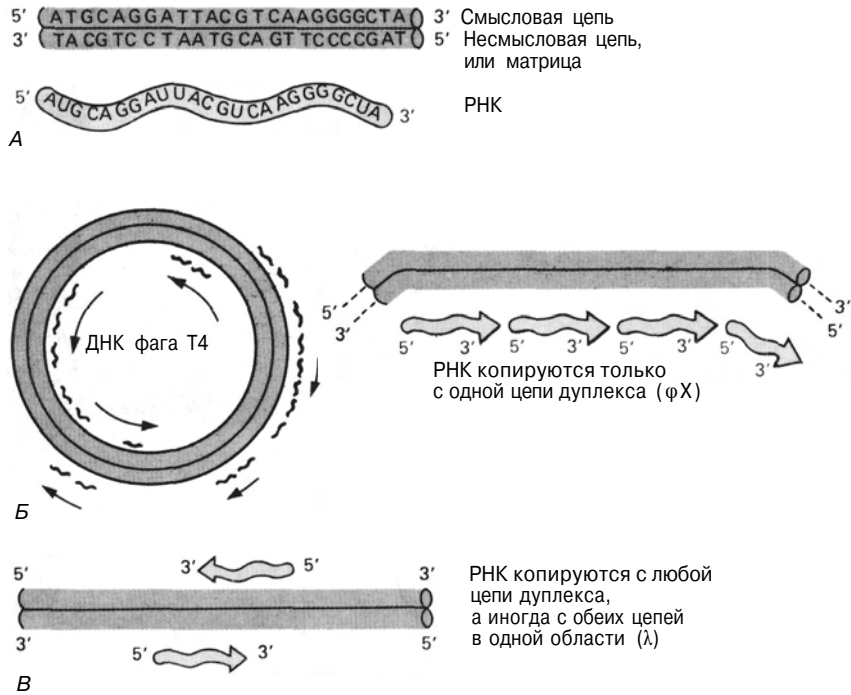
Двухцепочечная молекула ДНК—это физиологическая матрица для синтеза всех клеточных РНК. Даже если геном, как у некоторых вирусов, представлен одноцепочечной ДНК, последняя перед транскрипцией обязательно переходит в двухцепочечную репликативную форму. Транскрибирована

может быть любая из двух цепей геномной ДНК, однако матрицей при транскрипции отдельного гена обычно служит только какая-то одна из них (рис. 3.1). Впрочем, в некоторых случаях (примером может служить бактериофаг  $\phi$ X174) все мРНК транскрибируются с одной и той же цепи. Очень редко транскрипция идет на обеих цепях в одном и том же месте, так что образующиеся цепи РНК оказываются комплементарны друг другу; возможно, подобный способ транскрипции имеет особое регуляторное значение.

Нуклеотидными предшественниками для синтеза РНК являются четыре рибонуклеозид-5'-трифосфата: АТР, ГТР, УТР и СТР (рис. 3.2). Многие РНК содержат модифицированные нуклеотиды, но изменения в основаниях и рибозных остатках происходят после полимеризации, т.е. посттранскрипционно. Тем не менее РНК-полимеразы могут использовать рибонуклеозид-5'-трифосфаты, отличные от указанных четырех, при условии, что модифицированные основания обладают способностью к спариванию, сравнимой с таковой для аденина, гуанина, цитозина и урацила.

РНК-полимеразы катализируют реакцию присоединения 3'-ОН-группы нуклеотида, находящегося на растущем конце цепи, к  $\alpha$ -фосфату следующего рибонуклеозид-5'-трифосфата (рис. 3.3). Многократное повторение этой реакции приводит к постепенному удлинению цепи РНК. Образование каждой новой фосфодиэфирной связи сопровождается высвобождением неорганического пирофосфата; быстрый гидролиз пирофосфата до неорганического фосфата *in vivo* делает реакцию образования фосфодиэфирной связи энергетически выгодной.

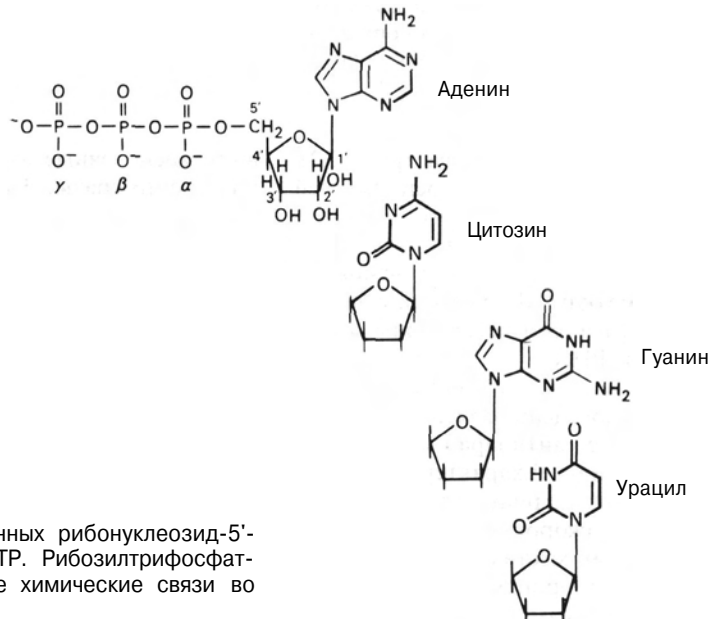
Транскрипция аналогична репликации в том смысле, что для ее осуществления также нужна ДНК-матрица (рис. 3.3). Порядок присоединения нуклеотидов определяется комплементарным спариванием оснований. Чтобы могло происходить комплементарное спаривание каждого следующего нуклеозидтрифосфата с матричным транскрибируемым основанием, спираль ДНК во время транскрипции должна раскручиваться с помощью ДНК-полимеразы (рис. 3.4). Растущая цепь РНК остается связанной с ферментом и спаренной своим растущим концом с участком матричной цепи длиной 20–30 нуклеотидов; остальная часть образовавшейся цепи не связана ни с ферментом, ни с ДНК. По мере продолжения транскрипции временно разошедшиеся цепи ДНК воссоединяются и восстанавливается исходная дуплексная структура. Таким образом, транскрипция—процесс консервативный, в котором сохраняется двойная спираль ДНК, а синтезированная цепь РНК отделяется. В противоположность этому репликация ДНК полуконсервативна, поскольку обе цепи исходного дуплекса распре-



**РИС. 3.1.**

РНК-транскрипты и соответствующие матричные ДНК. **А.** РНК-транскрипты комплементарны и антипараллельны матричной цепи и имеют такую же нуклеотидную последовательность (за исключением того, что место Т занимает U), как и смысловая цепь. **Б.** Индивидуальные мРНК обычно закодированы в одной из цепей дуплексного генома, однако для разных мРНК матрицей могут служить разные цепи (как у бактериофага

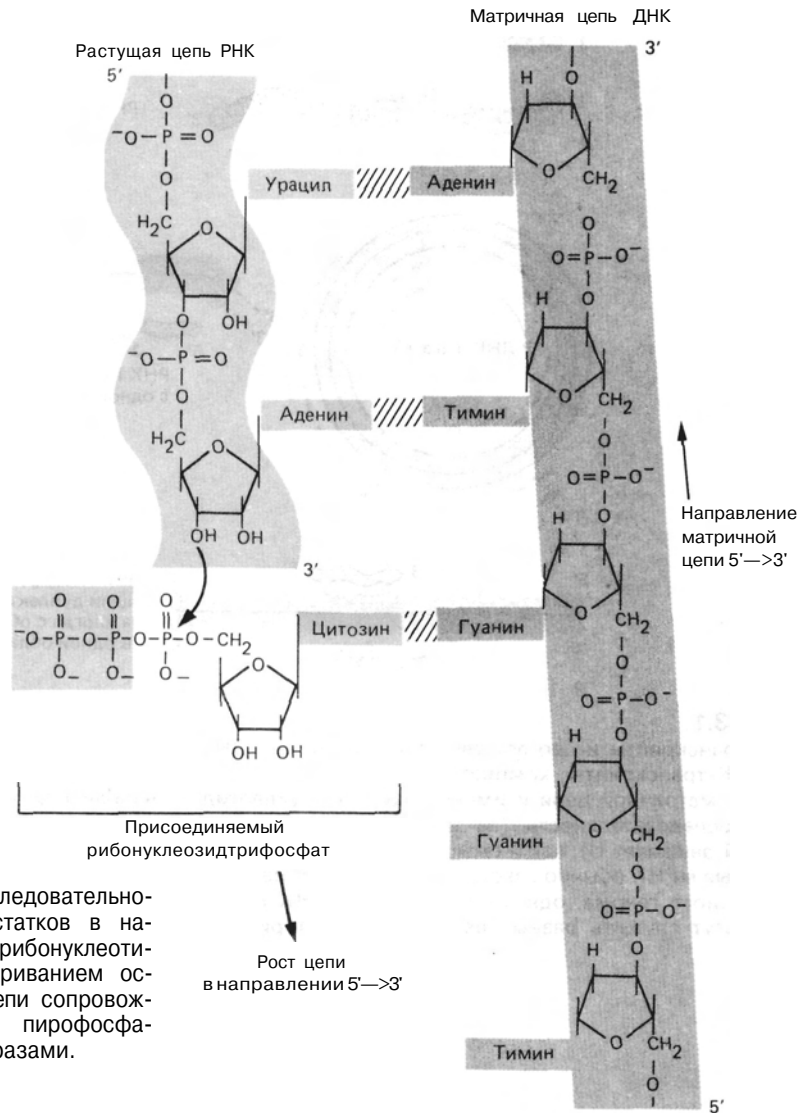
Т4). Таким образом, транскрипция, которая всегда осуществляется от 5'- к 3'-концу, может идти в обоих направлениях. Иногда все транскрипты образуются на одной цепи генома (как у бактериофага φX174). **В.** В тех редких случаях, когда гены перекрываются, обе цепи в одной и той же области могут служить матрицами для молекул РНК, транскрибируемых в противоположных направлениях.



**РИС. 3.2.**

Четыре наиболее распространенных рибонуклеозид-5'-трифосфата: АТР, СТР, ГТР и УТР. Рибозилтрифосфатные остатки и соответствующие химические связи во всех случаях идентичны.





### РИС. 3.3

Рост цепи РНК осуществляется путем последовательного присоединения рибонуклеотидных остатков в направлении 5'→3'. Последовательность рибонуклеотидов определяется комплементарным спариванием оснований. Присоединение нуклеотида к цепи сопровождается высвобождением одной молекулы пирофосфата. Реакция катализируется РНК-полимеразой.

деляются по двум дочерним спиральям (рис. 2.2). Другое существенное различие между репликацией и транскрипцией ДНК состоит в том, что репликация не может начаться без затравки—праймера (разд. 2.1), а инициация синтеза РНК с помощью РНК-полимеразы происходит *de novo*, начинаясь с рибонуклеозидтрифосфата, соответствующего первому нуклеотиду в цепи РНК.

Наращивание РНК идет в направлении от 5'-к 3'-концу вдоль матричной цепи, ориентированной в направлении 3'→5', т.е. антипараллельно (рис. 3.3). Несмотря на процессивный характер элонгации (фермент не отделяется от матрицы в течение всего раунда транскрипции), ее скорость вдоль матрицы не постоянна. В некоторых местах фермент делает остановки; возможно, это происходит там, где в одноцепочечной ДНК или в самой РНК образуются

внутрицепочечные дуплексы, мешающие продвижению полимеразы. Такие паузы могут при определенных обстоятельствах приводить к преждевременной термации транскрипции. Как мы вскоре увидим, сигналами для нормальной термации и отделения синтезированной РНК и полимеразы от матрицы являются особые структуры РНК — шпильки.

Каков механизм однонаправленного движения РНК-полимеразы вдоль матричной ДНК, остается неясным. Не знаем мы пока и того, как расплетается и заплетается вновь во время транскрипции дуплекс ДНК и почему восстановление этого дуплекса более выгодно, чем образование дуплекса ДНК-РНК. Можно лишь отметить, что, поскольку РНК-полимераза одна осуществляет все эти функции *in vitro* даже в случае ковалентно замкнутых кольцевых

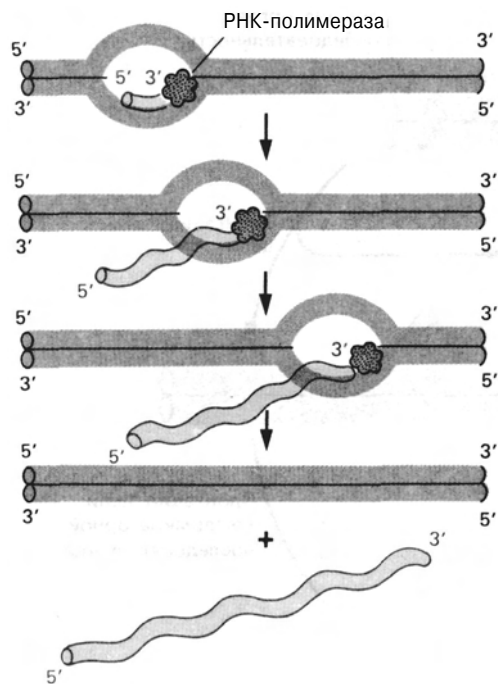


РИС. 3.4.

РНК-полимераза расплетает дуплекс ДНК в том месте, где происходит транскрипция. В ходе транскрипции фермент и локально расплетенный участок перемещаются вдоль ДНК и происходит удлинение цепи РНК.

матричных ДНК, все секреты, по-видимому, кроются в самом этом ферменте. Для сравнения вспомним, что ДНК-полимеразы не способны к инициации синтеза новых цепей *de novo* и что в процессах расплетания и восстановления дуплексов при репликации двухцепочечной ДНК участвуют геликазы и топоизомеразы.

### б. ДНК-зависимые РНК-полимеразы

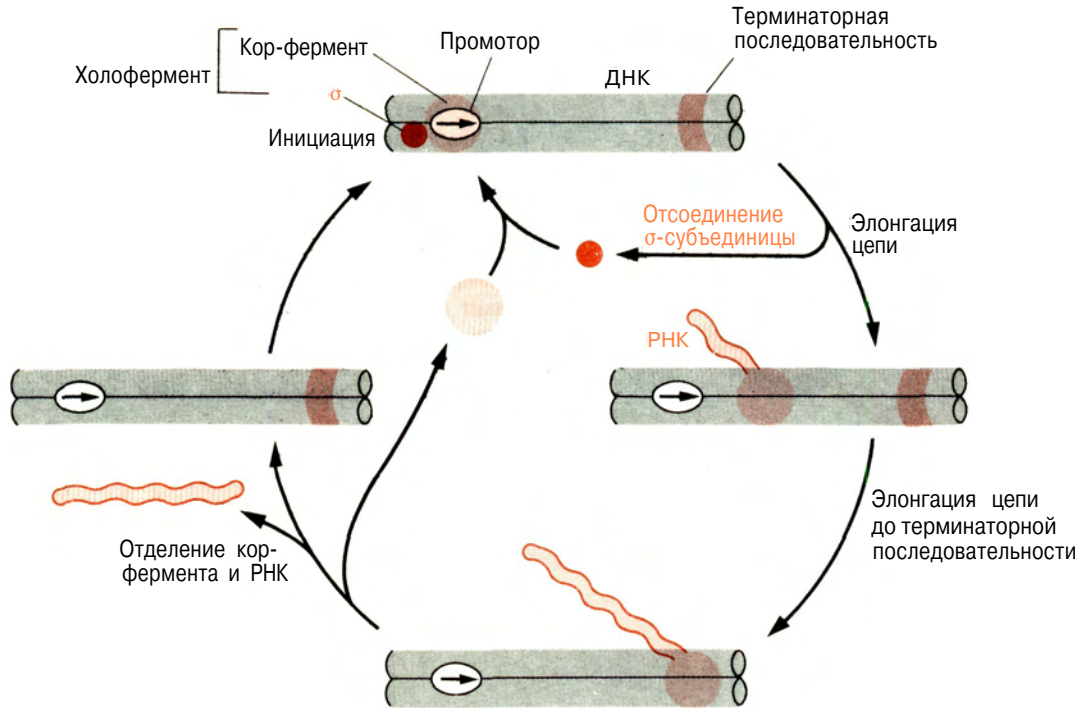
У прокариот синтез всех видов РНК—мРНК, рРНК и тРНК, а также более специализированных РНК, участвующих в процессинге РНК (разд. 3.3.в),—катализируется единственной ДНК-зависимой РНК-полимеразой (РНК-праймеры для синтеза ДНК образуются с помощью специальных РНК-полимераз, называемых праймазами; разд. 2.1.д). Бактериальные РНК-полимеразы—это сложные белки, состоящие из нескольких разных субъединиц. Наиболее изученный фермент—**холофермент РНК-полимераза *E. coli***—содержит пять разных полипептидных субъединиц (табл. 3.1): две  $\alpha$ -цепи, одну  $\beta$ - и одну  $\beta'$ -цепи,  $\sigma$ - и  $\omega$ -цепи ( $\alpha_2\beta\beta'\sigma\omega$ ). Альтернативная, сосуществующая с первой форма фермента, называемая **кором**, лишена  $\sigma$ -субъединицы.

Исследования с применением препаратов, ингибирующих РНК-полимеразу, и с ферментами, субъединицы которых были изменены в результате мутации, несколько прояснили роль кор-субъединиц.  $\beta$ -Субъединица скорее всего участвует в связывании рибонуклеозидтрифосфатов в реакциях инициации и элонгации. Комплекс  $\alpha$ - и  $\beta'$ -субъединиц ( $\alpha\beta'$ ) участвует в неспецифическом прочном связывании с ДНК и в специфическом взаимодействии холофермента с **промоторами**—сайтами, детерминирующими инициацию транскрипции (разд. 3.2.в). Наиболее полно по сравнению со всеми другими субъ-

Таблица 3.1. Субъединицы РНК-полимеразы *E. coli*

Субъединицы	Число копий в холоферменте	Мол. масса, дальтоны	Функция
Альфа ( $\alpha$ )	2	36500	Неизвестна
Бета ( $\beta$ )	1	150000	Связывание рибонуклеозидтрифосфатов и полимеризация РНК
Бета прим ( $\beta'$ )	1	155000	Связывание ДНК
Сигма ( $\sigma$ )	1	70-80000	Узнавание промотора и инициация
Омега ( $\omega$ )	1	11000	Неизвестна

единицами холофермента изучена роль  $\sigma$ -субъединицы. Кор-фермент, лишенный  $\sigma$ -субъединицы, катализирует большинство реакций, необходимых для транскрипции ДНК с образованием РНК, а именно комплементарное копирование матричной цепи, образование фосфодиэфирных связей и терминацию цепи РНК. Однако он не может иницировать синтез РНК в нужном месте, поскольку не способен узнавать промоторные сайты (разд. 3.2.в). Точное связывание и инициация в промоторах происходят только после добавления к кор-ферменту  $\sigma$ -субъединицы и образования холофермента (рис. 3.5). Такое поведение можно объяснить, например, тем, что кор-фермент очень прочно, но неспецифично связывается с ДНК, а поэтому редко оказывается в том месте, где находится промотор. И напротив, холофермент связывается с неспецифическими участками ДНК непрочно и, последовательно связываясь с разными областями ДНК, находит промотор и прочно связывается с ним. После образования нескольких первых фосфодиэфирных связей (возможно, от 5 до 10)  $\sigma$ -субъединица отделяется от иницирующего комплекса, и дальнейшая транскрипция осуществляется с помощью кор-фермента. Транскрипция непрерывно продолжается до тех пор, пока фермент не достигнет сайта терминации тран-



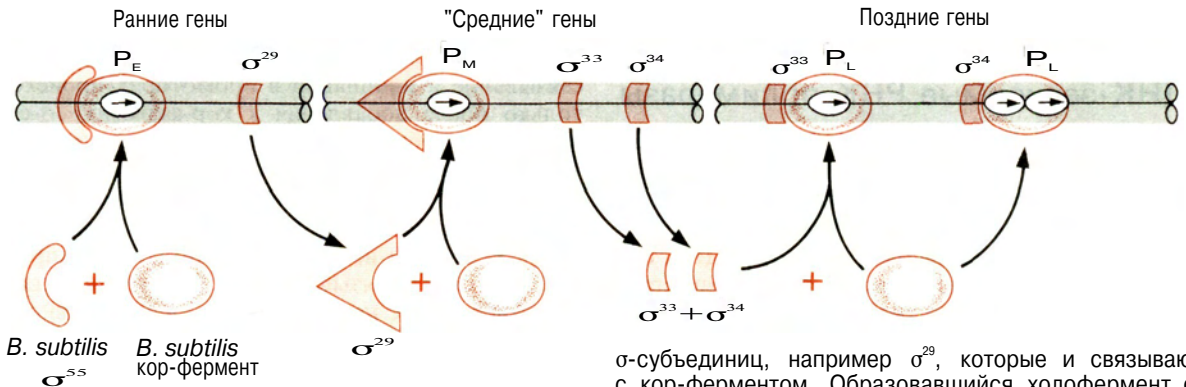
**РИС. 3.5.**

«Рабочий цикл»  $\sigma$ -субъединицы. Кор-фермент связывается с  $\sigma$ -субъединицей, образуя холофермент, который прочно связывается с промотором. После запуска

транскрипции  $\sigma$ -субъединица отделяется, а кор-фермент катализирует элонгацию цепи.

скрипции. Итак,  $\sigma$ -субъединица обеспечивает эффективное связывание холофермента с промотором, а при ее отсоединении полимеразы переключается на элонгацию.  $\sigma$ -Субъединица может снова стимулировать инициацию, специфически связавшись с другой молекулой РНК-полимеразы.

Имеются данные о том, что способность РНК-полимеразы узнавать промотор может изменяться при связывании с разными  $\sigma$ -субъединицами (рис. 3.6). Так, после заражения *Bacillus subtilis* определенными бактериофагами или на ранних стадиях споруляции экспрессируются разные  $\sigma$ -субъединицы и



**РИС. 3.6.**

Схема, иллюстрирующая участие различных  $\sigma$ -субъединиц в споруляции у *B. subtilis*. Некоторые  $\sigma$ -субъединицы, в частности  $\sigma^{55}$ , присутствуют и в вегетативных клетках. В условиях, стимулирующих споруляцию (в частности, голодание), появляется несколько новых

$\sigma$ -субъединиц, например  $\sigma^{29}$ , которые и связываются с кор-ферментом. Образовавшийся холофермент связывается с промоторами специфических генов споруляции, названных здесь «средними» генами. Затем экспрессируются другие специфические  $\sigma$ -субъединицы (в том числе  $\sigma^{33}$  и  $\sigma^{34}$ ); в комплексе с кор-полимеразой они иницируют транскрипцию генов, продукты которых необходимы для поздних стадий споруляции.

в результате изменяется порядок транскрипции клеточных и вирусных генов. Использует ли РНК-полимераза других прокариот различные  $\sigma$ -субъединицы для регуляции промоторной специфичности, пока неизвестно.

Не все РНК-полимеразы прокариот представляют собой мультисубъединичные ферменты. РНК-полимеразы, кодируемые бактериофагами T3 и T7 *E. coli*, — это одиночные полипептидные цепи средней длины (~ 100 кДа). Эти ферменты высокоспецифичны в отношении промоторных сайтов, используемых для транскрипции определенного набора вирусных генов. Такие «упрощенные» ферменты обладают всеми активностями мультисубъединичных РНК-полимераз. Они катализируют синтез РНК на ДНК-матрицах и осуществляют правильную терминацию цепей РНК.

### в. Транскрипция инициируется в особых нуклеотидных последовательностях

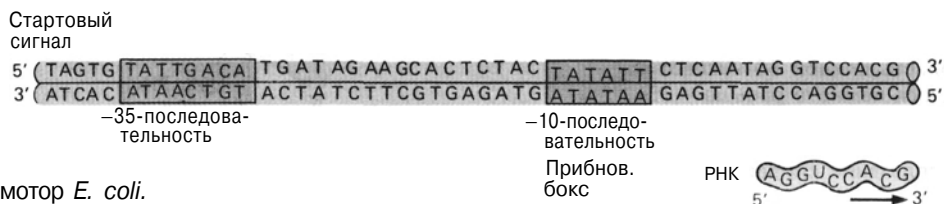
Транскрипция инициируется при образовании стабильного комплекса между холоферментом и специфической последовательностью, называемой промотором и располагающейся в начале всех транскрипционных единиц (рис. 3.7). Изучение нуклеотидной последовательности более чем 50 разных промоторных сайтов прокариот и мутационный анализ выявили только два консервативных участка, по-видимому играющих ключевую роль в узнавании и функционировании промотора. Одна из этих последовательностей состоит из шести или семи пар оснований и расположена на расстоянии примерно 10 оснований до того нуклеотида, с которого начинается транскрипция (+1); этот сигнал обычно обозначают как -10-последовательность, или **Прибнов-бокс** в честь ее открывателя. Сравнительный анализ -10-последовательностей примерно 50 промоторов прокариот показал, что все они немного отличаются от консенсус-последовательности ТАТААТ (рис. 3.8). Подчеркнутый Т присутствует почти во всех промоторах, тогда как по другим позициям в каждом промоторе может наблюдаться от одного до нескольких вариантов.

Вторая последовательность, длина которой обычно равна девяти нуклеотидам, расположена на

расстоянии ~ 35 оснований до сайта инициации (-35-последовательность) и также встречается в большинстве промоторов прокариот. Нуклеотидная последовательность сегмента между -35- и -10-участками не является критической, важно лишь расстояние между этими участками. -35-последовательность участвует в связывании РНК-полимеразы, которое предшествует перемещению фермента в Прибнов-бокс. Возможно, РНК-полимераза вызывает локальное раскручивание спирали, начиная этот процесс с Прибнов-бокса, и создает условия для инициации синтеза РНК (рис. 3.9).

Однако остается открытым вопрос, достаточно ли простого связывания РНК-полимеразы с промотором для локального расхождения цепей вблизи сайта инициации синтеза РНК или РНК-полимераза расплетает спираль в стартовом сайте. Независимо от механизма образование «открытого» промоторного комплекса позволяет РНК-полимеразе осуществить спаривание первого и второго рибонуклеозидтрифосфатов с матричной цепью и катализировать образование первой фосфодиэфирной связи.

Различия в эффективности транскрипции индивидуальных генов отчасти зависят от структуры их промоторов. Как прочность взаимодействия РНК-полимеразы с промоторной последовательностью, так и эффективность образования «открытого» промоторного комплекса определяются конкретными нуклеотидными последовательностями -35- и -10-участков соответственно. Мутации в этих участках приводят к значительным изменениям способности многих промоторов обеспечивать инициацию транскрипции (рис. 3.10). В некоторых случаях инициацию транскрипции в малоэффективных промоторах облегчают вспомогательные белки, связывающиеся с ДНК вблизи -35-последовательностей (разд. 3.11.в). Суть подобного феномена до конца не выяснена, но известно, что иногда белки, связывающиеся с -35-последовательностью, увеличивают вероятность того, что она будет обнаружена РНК-полимеразой и свяжется с ней. Связывание белков-активаторов может привести также к изменению структуры ДНК и тем самым способствовать инициации транскрипции. Изменение топологической структуры ДНК — особенно увеличение числа отрицательных сверхвитков — также может приводить к повышению или снижению эффективности некоторых промоторов.

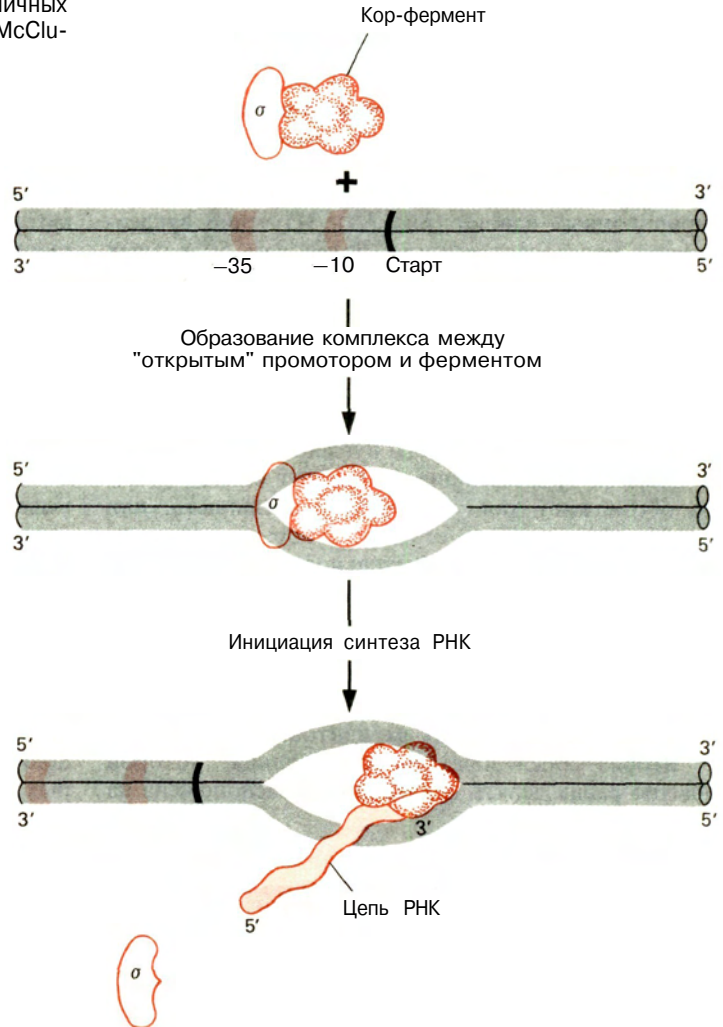


**РИС. 3.7.**  
Типичный промотор *E. coli*.

	Консенсусная -35-последо- вательность	11-15	Консенсусная -10-последо- вательность	5-8	+1
	TTGACA	←	TATAAT	←	→
araBAD	TTAGCGGATCCTACCTGACGCTTTT		TATCGCAACTCTCTACTGT		TTCTCCATACCCGTTTTT
lacP1	TAGGCACCCCAGGC TTTACA CTTTA		TGCTTCCGGCTCG TATGTT GTG		TGGAATTGTGAGC
ma1T	AAAAACGTCATCGC TTGCAT TAGAA		AGGTTTCTGGCC GACCTT ATA		ACCATTAATTACG
trp	TCTGAAATGAGCTG TTGACA ATTAA		TCATCGAACTAG TTAACT AGT		ACGCAAGTTCACGT
ilvGEDA	GGCCAAAAAATATC TTGTAC TATTT		ACAAAACCTATGG TAACTC TTT		AGGCATTCTTTCGA
argCBH	TTTGT TTTTCATTG TTGACA CACCT		CTGGTCATGA TAGTAT CAATATTCAT		TGCAGTATT
bioB	TTGT CATAATCGAC TTGTAA ACCAA		ATTGAAAAGATT TAGGTT TACAAGCT		TACACCGAAT
recA	TTTCTACAAAACAC TTGATA CTGTA		TGAGCATACAG TATAAT TGC		TTCACAGAACAT
rrnAB P1	TTTTAAATTTCTCT TTGTCA GGCCG		GAATAACTCCC TATAAT GCGCCACC		ACTGACACGG
rrnG P1	TTTATATTTTTCGC TTGTCA GGCCG		GAATAACTCCC TATAAT GCGCCACC		ACTGACACGG
str	TCGTTGTATATTTT TTGACACCTTT		TGGGCATCGCCC TAAAAT TCG		GCGTCTCATAT
rpoA	TTCGCATATTTTTC TTGCAAAGTTG		GGTTGAGCTGGC TAGATT AGC		CAGCCAACTTT
λPR	TAACACCGTGCGTG TTGACTATTTT		ACCTCTGGCGGT GATAAT GGT		TGCATGTACTAAG
λPL	TATCTCTGGCGGTG TTGACATAAAT		ACCACTGGCGGT GACTACT GAG		CACATCAGCAGGA
φX A	AATAACCGTCAGGA TTGACACCCCTC		CCAATTGTATGT TTTCAT GCC		TCCAAATCTTGGA
pBR tet	AAGAATTCTCATGT TTGACAGCTTA		TCATCGATAAGC TTGATT GCG		GTAGTTTATCACA
ColEI P1	GGAAGTCCACAGCT TTGACAGGGAA		AATGCAGCGGCG TAACTT TTA		TGCTGTATATAAAA

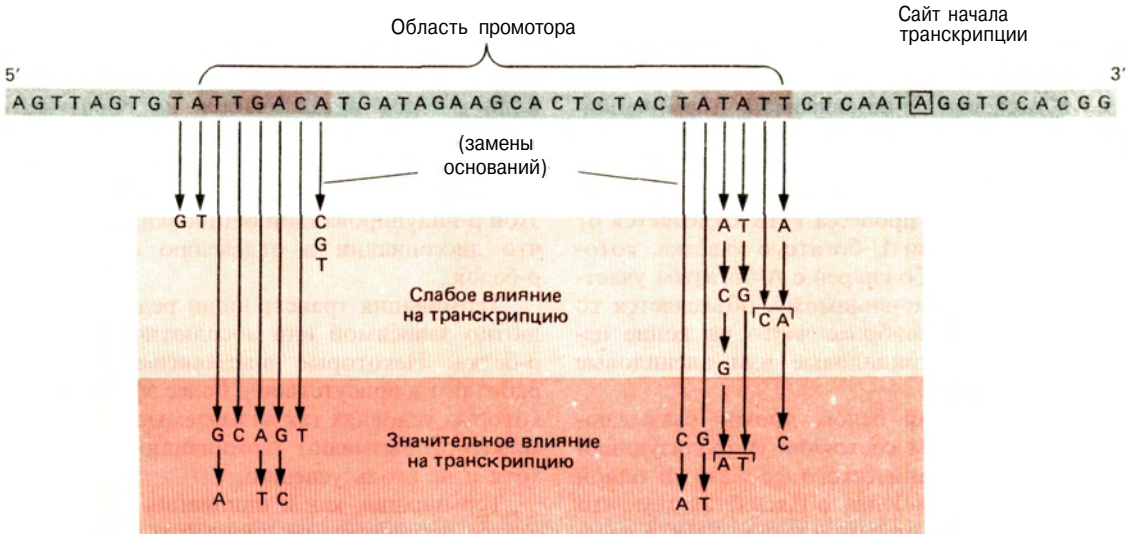
**РИС. 3.8.**

Нуклеотидные последовательности смысловой цепи ДНК до сайта инициации транскрипции для различных генов *E. coli* и бактериофагов. [D.K. Hawley, W.R. McClure, *Nucleic Acid Research*, **11** (1983), p. 2237.]



**РИС. 3.9.**

Связывание холоферментной формы РНК-полимеразы (показана схематически) с областью промотора. Считается, что белок перекрывает область длиной 50 нуклеотидов, начинающуюся немного раньше -35-последовательности и кончающуюся внутри транскрибируемого участка. Связывание стимулирует расплетание спирали и транскрипцию.



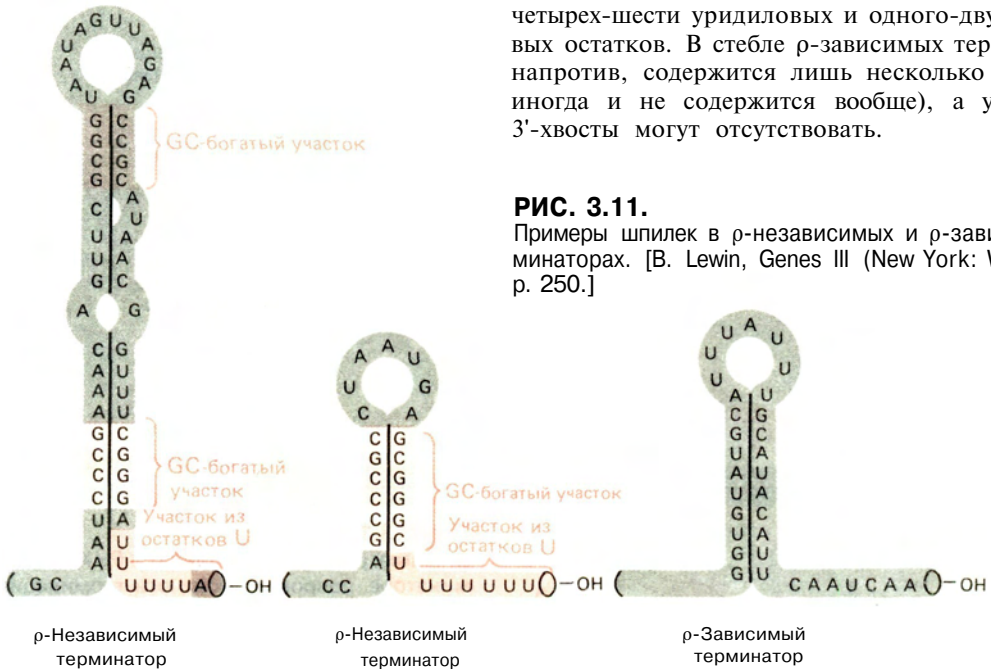
**РИС. 3.10.** Влияние мутаций в  $-10$ - и  $-35$ -последовательностях на транскрипцию. В экспериментах использовался промотор генома фага P22. Все представленные здесь

мутации подавляют транскрипцию. Некоторые мутации состояли в замене двух оснований. [P. Youderian, S. Bouvier, M.M. Suskind, *Cell*, **30** (1982), p. 843.]

### г. Терминация транскрипции и отделение цепей РНК

Последовательности ДНК, являющиеся сигналами остановки транскрипции, называются **транскрипционными терминаторами**. Обнаружены два типа сигналов терминации —  $\rho$ -зависимый и  $\rho$ -независимый

терминаторы. Оба они имеют некоторые общие признаки (рис. 3.11). И тот и другой содержат инвертированные повторы, благодаря чему 3'-концы РНК-транскриптов складываются с образованием **шпилек** разной длины. Стебли шпилек  $\rho$ -независимых терминаторов обычно содержат GC-богатые участки; один из них находится вблизи основания стебля, и к нему примыкает участок, состоящий из четырех-шести уридиловых и одного-двух адениловых остатков. В стебле  $\rho$ -зависимых терминаторов, напротив, содержится лишь несколько GC-пар (а иногда и не содержится вообще), а уридиновые 3'-хвосты могут отсутствовать.



**РИС. 3.11.** Примеры шпилек в  $\rho$ -независимых и  $\rho$ -зависимых терминаторах. [B. Lewin, *Genes III* (New York: Wiley, 1987), p. 250.]

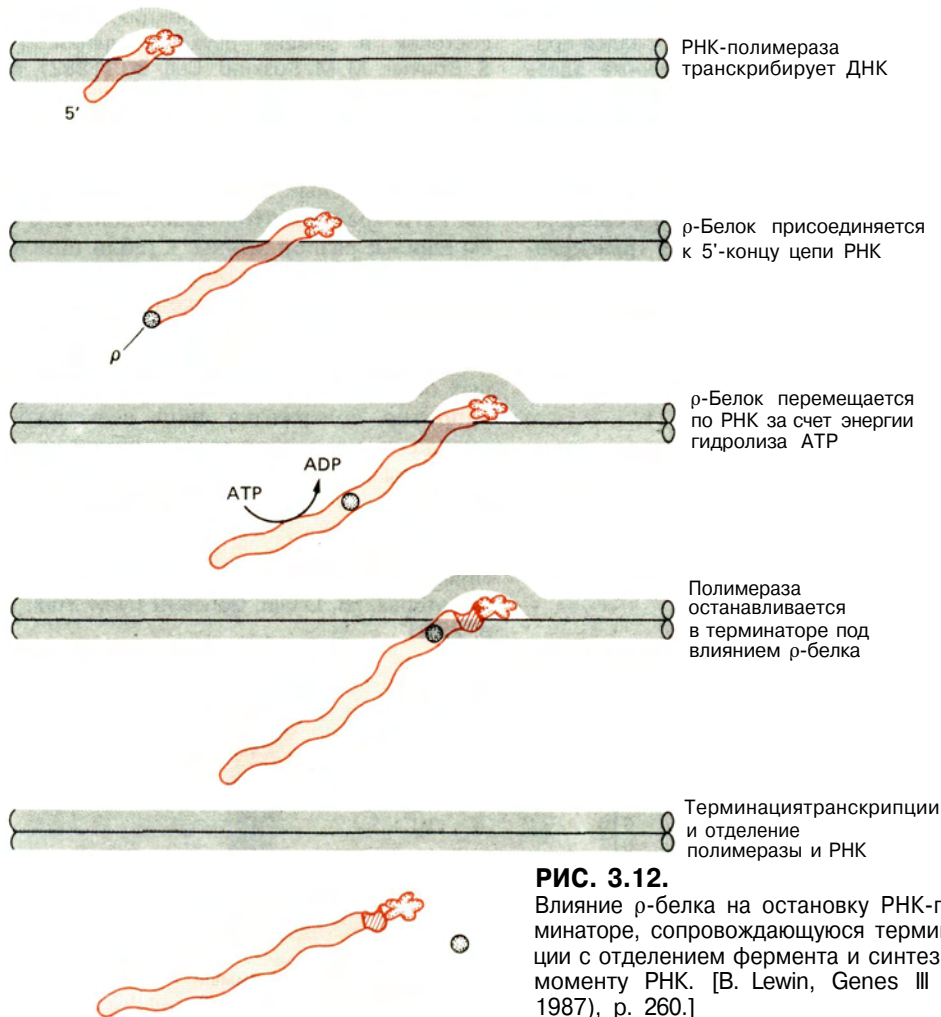
Точный механизм  $\rho$ -независимой и  $\rho$ -зависимой терминации транскрипции является пока предметом дискуссий. Наиболее вероятным представляется следующее объяснение  $\rho$ -независимой терминации: РНК-полимераза останавливается после транскрипции инвертированного повтора, потому что шпильчатая структура оказывается помехой. В результате сразу после остановки процесса РНК отделяется от матричной цепи вблизи U-богатого участка, который относительно слабо спарен с A-богатым участком матрицы. Этим, по-видимому, объясняется то обстоятельство, что наиболее часто на конце цепи РНК находятся уридилловые или адениловые остатки.

$\rho$  — это олигомерный белок, прочно связывающийся с РНК и в этом состоянии гидролизующий АТФ до АДФ и неорганического фосфата. В одной из моделей действие  $\rho$ -белка объясняется тем, что он связывается с синтезируемой цепью РНК и пере-

мещается вдоль нее в направлении  $5' \rightarrow 3'$  к месту синтеза РНК; необходимая для его перемещения энергия выделяется при гидролизе АТФ (рис. 3.12). Если  $\rho$ -белок наталкивается на образующуюся в РНК шпильку, он останавливает полимеразу, которая могла бы продолжать транскрипцию. Возможно, для отсоединения РНК от матрицы достаточно этой  $\rho$ -индуцированной остановки, но не исключено, что диссоциации и отделению способствует сам  $\rho$ -белок.

Терминация транскрипции редко является абсолютно зависимой или абсолютно независимой от  $\rho$ -белка. Некоторые  $\rho$ -независимые терминаторы работают в присутствии  $\rho$  более эффективно. А в некоторых условиях так называемые  $\rho$ -зависимые терминаторы вызывают терминацию и в отсутствие  $\rho$ , хотя и не столь успешно.

Терминация, как и инициация синтеза РНК, является регулируемым процессом. Существуют бел-



**РИС. 3.12.**

Влияние  $\rho$ -белка на остановку РНК-полимеразы в терминаторе, сопровождающуюся терминацией транскрипции с отделением фермента и синтезированной к этому моменту РНК. [B. Lewin, Genes III (New York: Wiley, 1987), p. 260.]

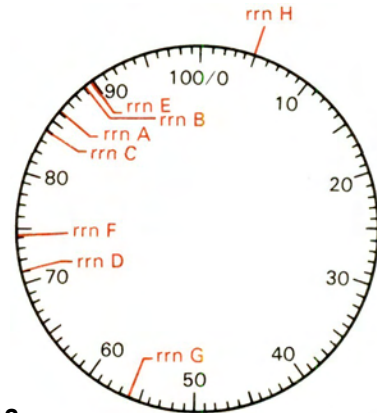
ки, функционирующие как **антитерминаторы** и предотвращающие терминацию в  $\rho$ -независимых терминаторах; известны также другие белки, ингибирующие  $\rho$  и тем самым обеспечивающие удлинение цепи после сигналов терминации. При определенных обстоятельствах молекулы РНК могут образовывать альтернативные шпильчатые структуры на особых последовательностях определенных участков ДНК. Одни из таких структур могут приводить к abortивной терминации, а другие — к продолжению транскрипции. Регулируемая терминация как способ контроля за образованием мРНК обсуждается в разд. 3.11.г.

### 3.3. ПРОЦЕССИНГ РНК У ПРОКАРИОТ

Первичные транскрипты, образующиеся при транскрипции прокариотических генов, кодирующих белки, функционируют в качестве мРНК без последующей модификации или процессинга. Действительно, трансляция мРНК часто начинается даже до завершения синтеза 3'-конца транскрипта (разд. 3.7.а). Совсем иная ситуация наблюдается для молекул рРНК и тРНК. В этом случае кластеры рРНК-или тРНК-генов или даже перемежающиеся участки этих генов часто транскрибируются с образованием единой цепи РНК. И хотя транскрипция этих генов всегда начинается на определенных промоторах и заканчивается на определенных терминаторах, для образования зрелых функциональных форм должны произойти специфическое надрезание первичных РНК-транскриптов и модификация. Подобные молекулярные события называют общим термином **посттранскрипционные модификации** или просто **процессинг РНК**. Механизмы процессинга рРНК и тРНК и ферменты, с помощью которых он осуществляется, наиболее полно изучены у *E. coli*, и для иллюстрации особенностей посттранскрипционного процессинга РНК мы используем эту систему. Аналогичные модификации эукариотических РНК обсуждаются в гл. 8; в этом случае помимо процессинга рРНК и тРНК используются более сложные системы созревания транскриптов с образованием мРНК.

#### а. Группы генов, кодирующих рРНК и тРНК

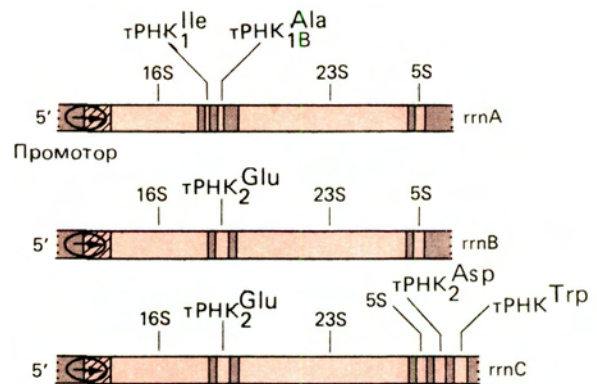
В геноме *E. coli* идентифицированы и картированы семь дискретных транскрипционных единиц, кодирующих рРНК; на рис. 3.13 они обозначены *rrnA–H*. Каждая транскрипционная единица — это молекула РНК, которая состоит из ~ 5000 нуклеотидов и содержит по одной копии кодирующих последо-



**РИС. 3.13.**

Локализация на хромосомной карте *E. coli* семи транскрипционных единиц рРНК.

вательностей для 5S-, 16S- и 23S-рРНК. Транскрипция в этой области осуществляется в направлении 16S → 23S → 5S. Помимо этих трех последовательностей, кодирующих рРНК, транскрипты содержат вставки разной длины (спейсеры) и одну или более копий тРНК-генов (рис. 3.14). Спейсеры могут находиться перед последовательностями для рРНК, между ними и после них, а тРНК-гены обычно лежат в пределах вкрапленных или 3'-концевых спейсерных сегментов. Для образования функционально зрелых молекул РНК должен произойти процессинг таких транскриптов. До процессинга или во время него происходит модификация специфических оснований в спейсерах, а также в рРНК- и тРНК-генах.



**РИС. 3.14.**

Схематическое представление трех типичных транскрипционных единиц рРНК *E. coli*. Указаны кодирующие участки для 5S-, 16S- и 23S-рРНК, молекул тРНК и спейсерные сегменты.



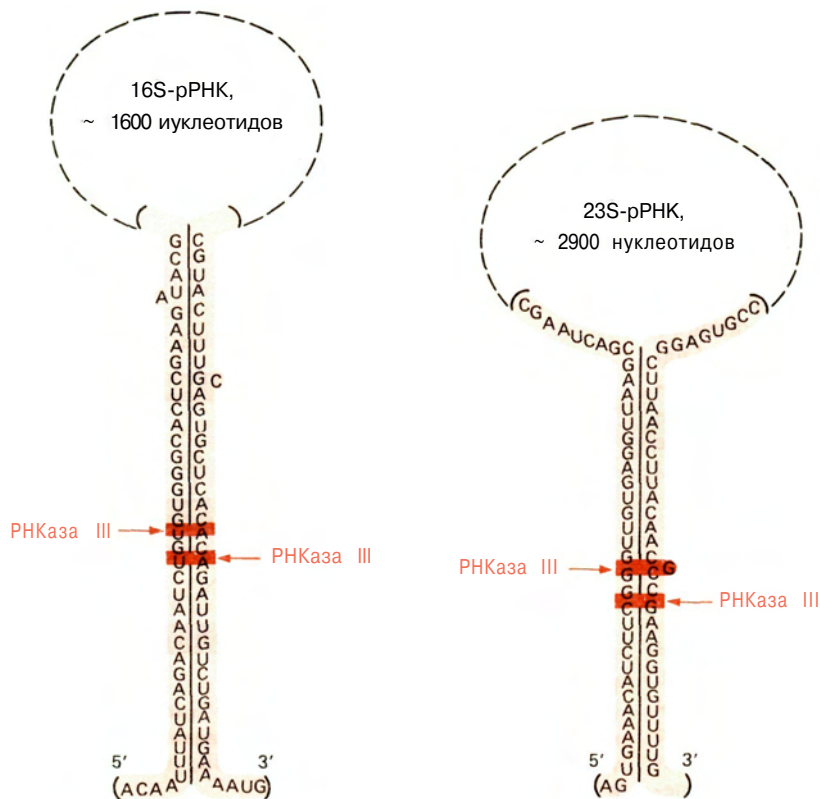
## б. Разрезание рРНК-тРНК-котранскриптов

Начальное расщепление первичных транскриптов на фрагменты, содержащие либо тРНК, либо 16S-, 23S- или 5S-рРНК-последовательности, осуществляет эндонуклеаза РНКаза III. Ее мишенями служат короткие дуплексы РНК, образующиеся при внутримолекулярном спаривании оснований в последовательностях, фланкирующих каждый из рРНК-сегментов (рис. 3.15). Например, комплементарные участки в спейсерных областях, фланкирующих последовательность 16S-рРНК, образуют стебель шпильки, в петле которой находится последовательность 16S-рРНК. Аналогичные шпильки образуют и последовательности 23S- и 5S-рРНК. РНКаза III вносит разрывы в двухцепочечный стебель, в результате образуется цепь РНК, содержащая последовательность той или иной рРНК, фланкированную короткими спейсерными участками с 5'-фосфатным и 3'-гидроксильным концами. Затем лишние нуклеотиды спейсерных последовательностей удаляются, возможно с помощью той же самой РНК-экзонуклеазы, которая катализирует и последние этапы процессинга тРНК (разд. 3.3.в). В принципе для того, чтобы произошло ферментативное расщепление, должны быть транскрибированы толь-

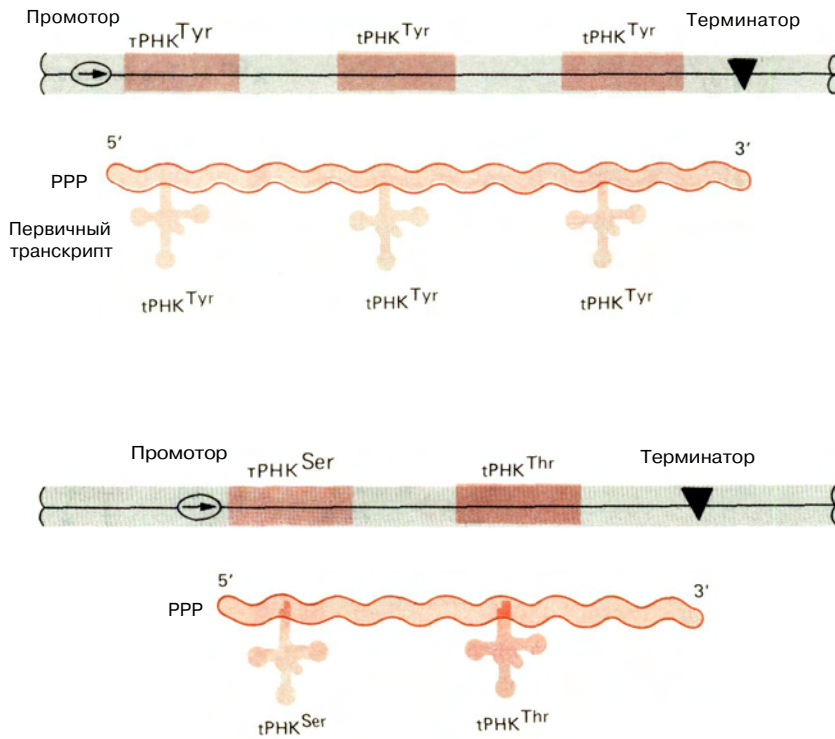
ко те нуклеотидные последовательности, которые образуют шпильки. Однако процессинг происходит лишь после завершения синтеза всего первичного транскрипта, поскольку, по-видимому, для правильной укладки целого РНК-транскрипта, который и распознается эндонуклеазой III, необходимы рибосомные или какие-либо другие белки. Процессинг тРНК-сегментов, выщепляющихся из мультигенных транскриптов, осуществляется так же, как и процессинг тРНК из транскрипционных единиц одиночных генов (разд. 3.3.в).

## в. Образование зрелых тРНК из более крупных транскриптов

Несмотря на то что некоторые кодирующие тРНК гены находятся внутри транскрипционных единиц рРНК и экспрессируются совместно с генами рРНК, основная часть тРНК-генов представлена одиночными генами или объединена в кластеры. Одни кластеры содержат множественные повторы одних и тех же генов (например, три идентичные копии гена тРНК<sup>Tyr</sup>), другие — различные и неродственные тРНК-гены (рис. 3.16). В некоторых случаях каждый кластер транскрибируется как одна большая молекула РНК, которая подвергается процессингу с последовательным выщеплением зрелых



**РИС. 3.15.** Сайты расщепления для РНКазы III в предшественниках рРНК. Эти сайты находятся в стебле, образуемом при внутримолекулярном спаривании оснований спейсерных участков, фланкирующих 16S- и 23S-рРНК-последовательности. [B. Lewin, Genes III (New York: Wiley, 1987), p. 453.]

**РИС. 3.16.**

Типичные кластеры генов тРНК *E. coli* и их первичные транскрипты.

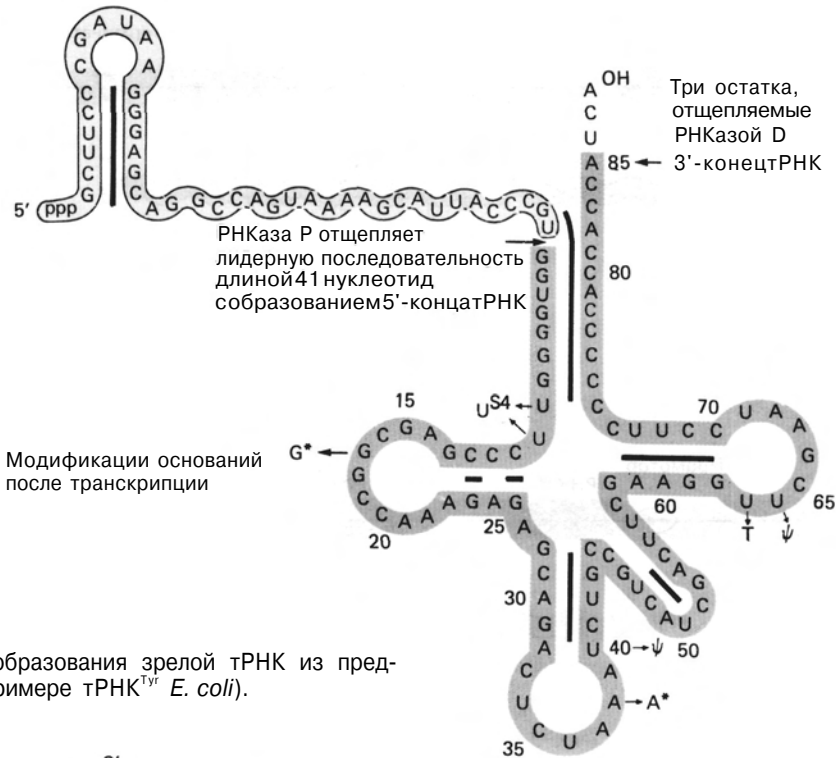
тРНК-фрагментов (рис. 3.17). Для образования зрелой функциональной тРНК, по-видимому, должны произойти специфическая модификация оснований и присоединение одного, двух или всех трех нуклеотидов 3'-ССА-конца.

Независимо от того, содержит ли первичный транскрипт одну или более тРНК-последовательностей или эти последовательности внедрены в спейсерные участки рРНК, 5'-концы всех тРНК образуются при участии одной эндонуклеазы, называемой РНКазой Р. По-видимому, РНКаза Р узнает характерную свернутую структуру тРНК в полинуклеотиде-предшественнике и отщепляет лидерную или спейсерную последовательности, расположенные перед 5'-концом зрелой последовательности тРНК. 3'-концы тРНК образуются с помощью нескольких активностей. До сих пор неидентифицированная эндонуклеаза расщепляет предшественник в том месте шпильки, где находится 3'-конец зрелой тРНК, а затем другая эндонуклеаза, РНКаза D, завершает образование правильного 3'-конца. В некоторых случаях экзонуклеазное расщепление прекращается точно у 3'-ССА-конца зрелой тРНК (рис. 3.17), а в других случаях под действием экзонуклеазы образуется конец, служащий затравкой, к которому тРНК-нуклеотидилтрансфераза добавляет один или более

инвариантных концевых нуклеотидов (С и/или А).

Отличительной особенностью РНКазы Р является то, что сайт расщепления для нее формируется в результате правильной укладки молекулы тРНК. Изменения в нуклеотидной последовательности, не приводящие к нарушению этой укладки, не сказываются и на процессинге 5'-конца. Другим необычным свойством РНКазы Р является то, что она состоит из белка и РНК. Эта РНК имеет специфическую последовательность из 377 нуклеотидов и сама транскрибируется РНК-полимеразой с гена чуть большего размера и затем подвергается процессингу до размера зрелой молекулы. Удивительной особенностью этой РНК оказалось то, что она одна может катализировать такую же эндонуклеазную реакцию, что и целый рибонуклеопротеин; белок же не обладает самостоятельной эндонуклеазной активностью. Таким образом, эндонуклеазная активность может быть присуща самой РНК, а белок, по-видимому, необходим для сохранения структуры РНК в максимально активной конфигурации.

Зрелые тРНК не только имеют характерную конформацию, но и содержат модифицированные нуклеотиды (рис. 3.18). Многие из таких модификаций оказываются существенными для выполнения некоторых физиологических функций тРНК. Сегодня охарактеризованы лишь немногие из целой армии ферментов, катализирующих огромное количество реакций модификации. Однако ясно, что моди-

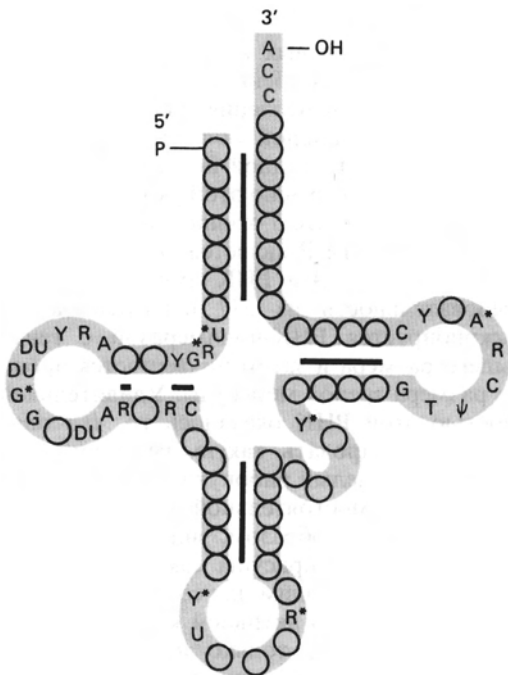


**РИС. 3.17.** Различные этапы образования зрелой тРНК из предшественника (на примере тРНК<sup>Tyr</sup> *E. coli*).

Модификации происходят в основном на стадии РНК-предшественника и в полностью процессированной тРНК. Такие модифицирующие ферменты представляют особый интерес благодаря своей необычной специфичности в отношении определенных последовательностей: например, только отдельные урациловые остатки превращаются в тиюрацил, метилируются до тимина или восстанавливаются до дигидроурацила. Еще более загадочным представляется образование псевдоуридата при модификации обычной связи между урацилом и рибозой (рис. 1.21).

### 3.4. ГЕНЕТИЧЕСКИЙ КОД

Последовательность аминокислот в белках определяется порядком расположения дезоксирибонуклеотидов в генах, кодирующих белки, точнее — последовательностью рибонуклеотидов в мРНК-транскриптах. Информационная связь между нуклеотидными



- Y - Пиримидин
- R - часто пурин
- DU - часто дигидроуридин
- \* - модифицированное основание
- T - тимидин
- ψ - псевдоуридин

**РИС. 3.18.** Схематическое изображение тРНК в форме клеверного листа. В определенных позициях у всех тРНК находятся модифицированные основания.

и аминокислотными последовательностями осуществляется с помощью **генетического кода**. Для составления генетического словаря было проведено множество специальных генетических и биохимических экспериментов. Он включает также и знаки препинания—начало и конец участков, кодирующих белки. За исключением незначительных вариаций в использовании нескольких нуклеотидов для кодирования особых аминокислот у митохондрий и некоторых инфузорий, генетический словарь универсален, т.е. конкретная последовательность нуклеотидов задает одинаковую для всех живых организмов аминокислотную последовательность.

Наличие подобной системы кодирования подразумевает существование некоего механизма для перевода информации с языка нуклеотидов на язык аминокислот. Как и следовало ожидать, этот механизм и реакции, осуществляющие перевод (трансляцию), очень сложны. Несмотря на различия между про- и эукариотами как в том, что касается структуры мРНК (разд. 3.8.а), так и в физическом взаимоотношении генов и аппарата трансляции, оба типа организмов используют весьма сходные механизмы для расшифровки генетических посланий.

В этом разделе описываются свойства генетического кода. Аппарат трансляции и процесс сборки белка мы обсудим в разд. 3.5-3.8.

### а. Аминокислотная последовательность белков соответствует нуклеотидной последовательности кодирующих их генов

Предположение о коллинеарности нуклеотидных и аминокислотных последовательностей было высказано в числе первых в дискуссии о природе генетического кода. Догадки такого рода возникли после того, как было показано, что многие мутации проявляются в замене одной-единственной аминокислоты в белках бактерий, растений или животных. Но гипотеза коллинеарности оставалась неподтвержденной до тех пор, пока не были проведены тщательные генетические и биохимические исследования хорошо охарактеризованных генно-белковых систем. Например, было показано, что относительное положение аминокислотных замен в  $\alpha$ -субъединице триптофансинтетазы *E. coli* согласуется с относительной локализацией на карте соответствующих мутаций в *trpA*-гене этого микроорганизма. Однако выделить и охарактеризовать мутантную ДНК в то время не удалось, поэтому нельзя было установить соответствие между нуклеотидными последовательностями и аминокислотными.

Теоретический анализ привел к предположению, что наиболее подходящей по размеру для генетической кодирующей единицы, или **кодона**, является последовательность из трех нуклеотидов. В основе



**РИС. 3.19.**

Изучение мутаций, сдвигающих рамку считывания, показало, что кодоны состоят из трех нуклеотидов. Использовались мутагены (в частности, профлавин), которые индуцировали делеции или вставки одного или двух оснований. Образующиеся мутанты синтезировали неактивные или укороченные полипептиды, поскольку сдвиг рамки приводит к появлению кодонов, детерминирующих неправильные аминокислоты, или стоп-кодона. Рекомбинация между двумя мутантами, один из которых содержит делецию, а другой в этой же или близкой позиции - вставку (например, -1 x +1 или -2 x +2), может привести к восстановлению функционального гена, если несколько концевых, находя-

щихся за рамкой считывания кодонов приходятся на несущественную часть полипептида. Рекомбинация между мутантами, которые оба имеют по одной вставке (+1 x +1) или по делеции из двух оснований (-2 x -2), не способна исправить мутацию. Напротив, рекомбинация между генами, имеющими в сумме три близко расположенных вставочных основания (например, +1 x +2) или три близко расположенных делегированных основания (например, -1 x -2), может привести к восстановлению функциональных генов, поскольку за этими измененными участками восстанавливается нормальная рамка считывания.

этого вывода лежали три соображения. Во-первых, четыре нуклеотида, взятые по одному, могут кодировать только четыре разные аминокислоты. Сочетания из двух нуклеотидов могут кодировать только  $4^2$ , или 16 аминокислот, а это меньше, чем те 20 аминокислот, которые, как было известно, присутствуют в белках. И только совокупности трех нуклеотидов дают 64 возможных кодона ( $4^3$ ), т.е. число, более чем достаточное для кодирования 20 разных аминокислот. Генетические эксперименты, выполненные на мутантах с делециями или вставками длиной один, два или три нуклеотида в генах, кодирующих белки, позволили доказать, что наиболее подходящий размер для кодона—три нуклеотида (рис. 3.19). Более того, из этих исследований был сделан вывод, что нуклеотидная последовательность считывается расположенными один за другим триплетами с фиксированной точки. Все эти выводы наряду с данными о том, что полипептидные цепи синтезируются последовательно путем соединения аминогруппы одной аминокислоты с карбоксильной группой другой, послужили краеугольным камнем в расшифровке генетического кода.

## б. Соответствие между аминокислотами и их кодонами

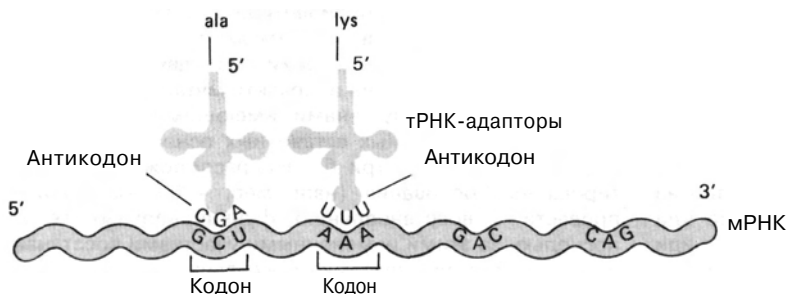
Одним из загадочных моментов кодирования было отсутствие структурной комплементарности между нуклеиновыми кислотами, с одной стороны, и аминокислотными цепочками—с другой. Выход из этого концептуального тупика—ответ на вопрос, как аминокислоты спариваются с соответствующими кодонами,—был найден, когда появилась идея о существовании **адаптора**. Согласно этой идее, аминокислоты сначала связываются с молекулами РНК, а затем такие гибриды выстраиваются вдоль мРНК, соединяясь с ней путем комплементарного спаривания нескольких оснований в адапторной молекуле РНК с соответствующим кодоном в мРНК (рис. 3.20). Адапторная гипотеза получила строгое экспериментальное подтверждение после того, как были обнаружены тРНК и ферменты, ответственные за связывание аминокислот и тРНК, и показано, что присоединенные к тРНК аминокислоты являются

прямыми предшественниками при сборке полипептида.

Если тРНК—это адапторы, то каждая аминокислота должна присоединяться только к специфической тРНК, а каждая тРНК—спариваться только с одним, соответствующим ей кодоном. Правильность первого из этих положений была доказана с открытием особых ферментов—аминоацил-тРНК-синтетаз, каждый из которых связывает одну-единственную аминокислоту с одной или несколькими родственными тРНК (рис. 3.21). Эти ферменты и катализируемые ими реакции более детально будут рассмотрены в разд. 3.5.а. Здесь достаточно сказать, что связывание аминокислот с молекулами тРНК—это первый шаг в процессе расшифровки. Правильность второго положения адапторной гипотезы—о том, что тРНК сама определяет место своей аминокислоты в полипептидной цепи—была подтверждена с помощью простого эксперимента. Одна из аминокислот, будучи связанной с соответствующей тРНК, была химически превращена в другую. После включения этой модифицированной аминокислоты в белок *in vitro* была установлена ее локализация в белковой цепи (рис. 3.22). Оказалось, что после превращения цистеинил-тРНК<sup>Cys</sup> в аланил-тРНК<sup>Cys</sup> остаток аланина, связанный с тРНК<sup>Cys</sup>, обнаруживается в тех сайтах белковой цепи, которые обычно занимает цистеин, а не в сайтах, где обычно находится аланин. Стало ясно, что именно тРНК с присоединенной к ней аминокислотой, а не сама аминокислота определяет, с каким кодоном должно произойти спаривание.

## в. Расшифровка генетического кода

Предпосылками для расшифровки кода послужили два открытия. Во-первых, было установлено, что мРНК—это информационный посредник между генами и белками. Во-вторых, обнаружилось, что мРНК, введенная в бактериальные экстракты, транслируется с образованием соответствующих белков. Прорыв в этой области произошел, когда с помощью экстрактов из клеток *E. coli* была осуществлена трансляция синтетических РНК—полиуридилата (poly U), полиаденилата (poly A) и полицитидилата



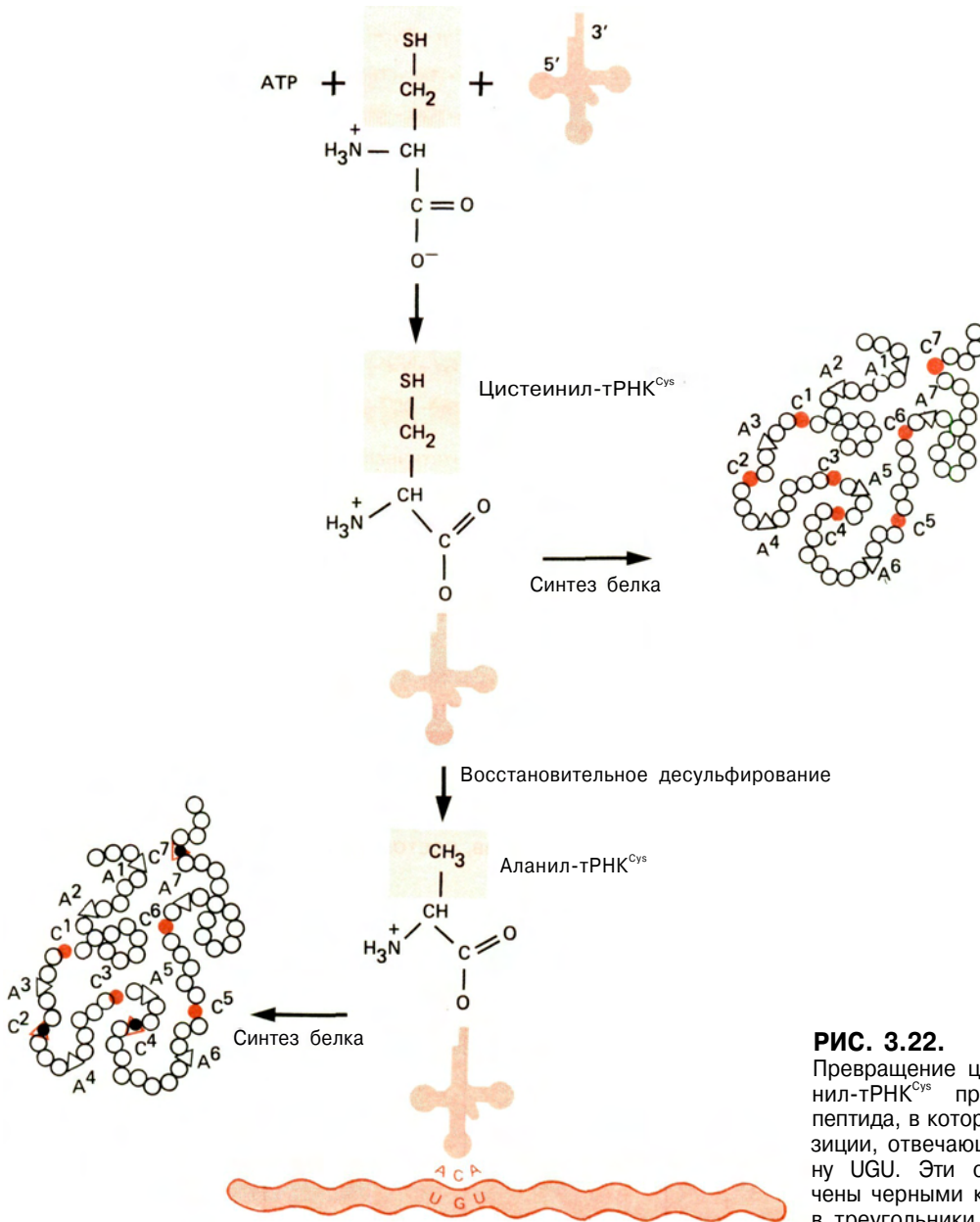
**РИС. 3.20.**

Адапторная гипотеза. Аминокислоты присоединяются к специфическим адапторным РНК (тРНК), которые затем выстраиваются вдоль мРНК благодаря комплементарному спариванию между кодонами мРНК и антикодонами адаптора.



**РИС. 3.21.**

Каждая аминоацил-тРНК-синтетаза присоединяет специфическую аминокислоту к одной или нескольким родственным тРНК.



**РИС. 3.22.**

Превращение цистеинил-тРНК<sup>Cys</sup> в аланил-тРНК<sup>Cys</sup> приводит к синтезу полипептида, в котором аланин занимает позиции, отвечающие цистеиновому кодону UGU. Эти остатки аланина обозначены черными кружками, помещенными в треугольники.

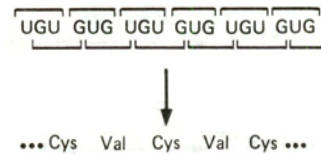
та (poly C)—с образованием полифенилаланина, полилизина и полипролина соответственно. Это привело к заключению, что триплеты, состоящие только из U, A и C, кодируют соответственно фенилаланин, лизин и пролин. Затем были проделаны эксперименты с применением смешанных полимеров с варьирующим соотношением двух и трех нуклеотидов; в результате был определен состав кодонов (табл. 3.2). Однако эти данные позволили установить лишь нуклеотидный состав кодонов, но не порядок следования нуклеотидов в них.

Все кодоны были в конце концов идентифицированы при помощи следующих двух экспериментов. В экспериментах первого типа сравнивали аминокислотную последовательность полипептидов, полученных *in vitro* с использованием синтетических мРНК, содержащих определенные повторы из двух или трех нуклеотидов (рис. 3.23 и табл. 3.3). В экспериментах другого типа определяли, какая именно из аминоксил-тРНК связывается с рибосомами в присутствии каждого из возможных тринуклеотидов (рис. 3.24). Эти эксперименты позволили составить непротиворечивый словарь, в котором 61 тринуклеотидный кодон соответствует 20 аминокислотам, а три кодона—окончанию кодирующей последовательности (рис. 3.25).

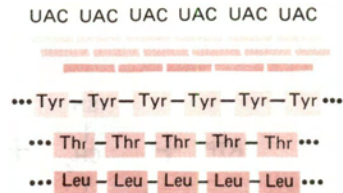
В коде заложена некая неоднозначность, которая связана с точкой начала трансляции, а не с соответствием кодон-аминокислота. Эта неоднозначность обусловлена наличием альтернативных наборов триплетов, или рамок считывания, для любой полинуклеотидной последовательности (рис. 3.26). Большинство прокариотических генов транслируется при одной непрерывной рамке считывания; при альтернативных рамках на каждые 20 нуклеотидов приходится в среднем по одному терминирующему кодону.

В экспериментах по расшифровке кода, описанных выше, синтетические полинуклеотиды транслировались в условиях, не требующих точной инициации. Однако *in vivo* и в соответствующих условиях

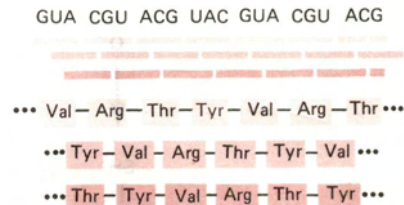
Полимер, состоящий из повторяющихся динуклеотидов



Полимер, состоящий из повторяющихся тринуклеотидов



Полимер, состоящий из четырех повторяющихся нуклеотидов



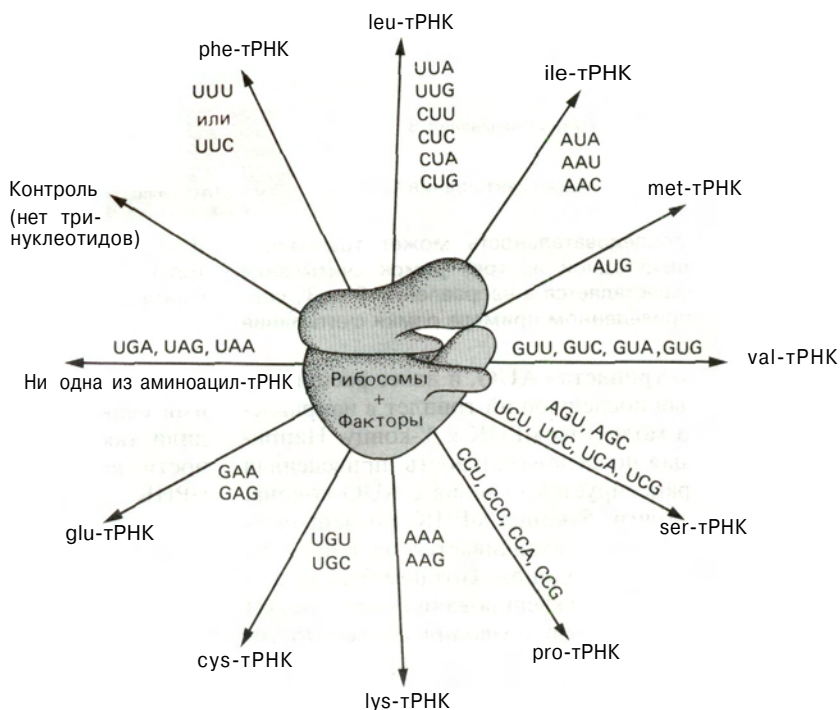
### РИС. 3.23.

Некоторые синтетические мРНК, кодирующие определенные полипептиды. Синтез этих специфических полипептидов может осуществляться в клеточных экстрактах, которые катализируют синтез белков. Заметим, что аминокислотный состав полипептидного продукта зависит от того, какая из трех возможных комбинаций оснований в  $\text{poly}(\text{UAC})_n$  реализуется: UAC, ACU или CUA.

*in vitro* инициация происходит только с правильной рамкой считывания. Однозначность прочтения белок-кодирующей последовательности обеспечивается тем, что трансляция мРНК начинается только со

Таблица 3.2. Кодирующие свойства случайных сополимеров, состоящих из U (0,76) и G (0,24) в различных комбинациях

Триплет	Расчетная частота триплета	Относительная частота	Аминокислота	Относительная встречаемость в полипептидах	Возможный состав кодона
UUU	0,44	100	Фенилаланин	100	3U
UUG	0,14	32	Валин	37	2U1G
UGU	0,14	32	Лейцин	36	2U1G
GUU	0,14	32	Цистеин	35	2U1G
UGG	0,044	10	Триптофан	14	1U2G
GUG	0,044	10			
GGU	0,044	10	Глицин	12	1U2G
GGG	0,014	3			



**РИС. 3.24.**

Идентификация кодонов, основанная на изучении связывания специфических аминокислот-тРНК с рибосомами в присутствии синтетических тринуклеотидов. Рибосомы и необходимые растворимые факторы инкубировали с определенными тринуклеотидами и смесью аминокислот-тРНК, в которой радиоактивно меченной была только одна аминокислотная единица. С каждым из комплексов рибосома-тринуклеотид связывалась только одна специфическая меченная аминокислот-тРНК (например, с комплексом рибосома-UCA связывалась радиоактивно меченная пролил-тРНК).

Первая позиция (5')      Вторая позиция      Третья позиция (3')

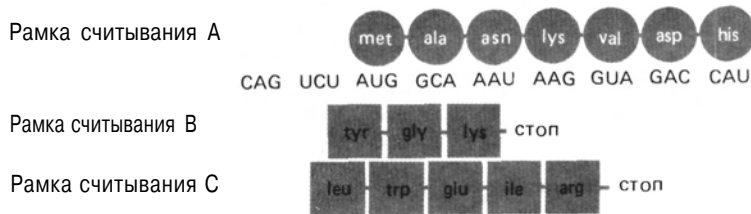
	U	C	A	G	
U	PHE	SER	TYR	CYS	U
	PHE	SER	TYR	CYS	C
	LEU	SER	Stop	Stop	A
	LEU	SER	Stop	TRP	G
C	LEU	PRO	HIS	ARG	U
	LEU	PRO	HIS	ARG	C
	LEU	PRO	GLN	ARG	A
	LEU	PRO	GLN	ARG	G
A	ILE	THR	ASN	SER	U
	ILE	THR	ASN	SER	C
	ILE	THR	LYS	ARG	A
	MET	THR	LYS	ARG	G
G	VAL	ALA	ASP	GLY	U
	VAL	ALA	ASP	GLY	C
	VAL	ALA	GLU	GLY	A
	VAL	ALA	GLU	GLY	G

**РИС. 3.25.**

Генетический код. Верхняя схема построена так, что мы можем идентифицировать аминокислоту, кодируемую тем или иным триплетом. Так, триплет CAG можно расшифровать, найдя точку пересечения C из левой колонки, A - из верхнего ряда и G - из правой колонки. Нижняя схема показывает число и состав кодонов, детерминирующих каждую из аминокислот.

AGA	AGG	GCA	CGA	GCG	CGG	GCC	CGC	GAC	AAC	UGC	GAA	CAA	GGC	CAC	AUC	CUC	AAA	UUC	CCC	UCC	ACC	UAC	GUC	UAG
GCU	CGU	GAU	AAU	UGU	GAG	CAG	GGU	CAU	AUU	CUU	AAG	AUG	UUU	CCU	UCU	ACU	UGG	UAU	GUU	UGA	UAA	GUG	UUA	
ALA	ARG	ASP	ASN	CYS	GLU	GLN	GLY	HIS	ILE	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL	Stop	Stop	Stop	Stop	



**РИС. 3.26.**

Нуклеотидная последовательность может транслироваться с помощью одной из трех рамок считывания (трансляция осуществляется в направлении 5'→3', слева направо). В приведенном примере рамки считывания

В и С прерываются стоп-кодонами. «Открытой» является только рамка считывания А, начинающаяся со стартового кодона AUG.

специфического триплета—AUG, и далее расшифровывается каждый последующий триплет в направлении от 5'-конца молекулы мРНК к 3'-концу. Например, кодирующая последовательность, приведенная на рис. 3.27, транслируется начиная с AUG-кодона, находящегося вблизи 5'-конца мРНК, по направлению к 3'-концу, где и заканчивается на одном из трех терминирующих кодонов. Позднее был разработан метод быстрого секвенирования нуклеиновых кислот и белков, который позволил проверить систему кодирования непосредственно, путем сравне-

ния последовательностей ДНК, РНК и кодируемых ими белков. Сравнительные исследования подтвердили также и то, что кодирующие последовательности действительно читаются от 5'- к 3'-концу мРНК.

## г. Избыточность генетического кода

Удивительной особенностью кода оказалось то, что все аминокислоты, кроме двух, кодируются более чем одним кодоном. Эти две составляющие исключения аминокислоты, метионин и триптофан, встречаются в белках достаточно редко. Наибольшее число кодонов имеют серин и лейцин, которыми белки изобилуют. Такие достаточно часто встречающиеся аминокислоты, как цистеин, аланин, глицин, валин, а также дикарбоновые кислоты и их амиды, кодируются двумя-четырьмя кодонами каждая. Из-за такой избыточности разные нуклеотидные последовательности могут при трансляции давать одну и ту же аминокислотную последовательность (рис. 3.28). Итак, если мы знаем нуклеотидную последовательность, то можем однозначно определить последовательность белка, обратное же проделать невозможно.

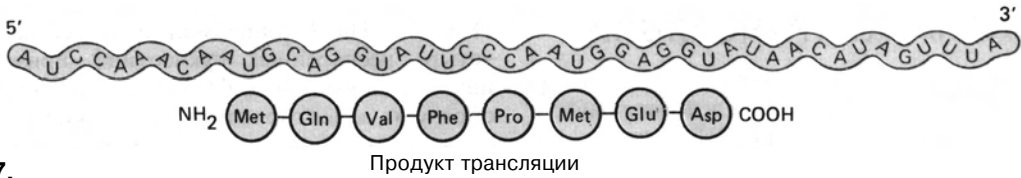
Сигналом для остановки синтеза белка служит любой из трех кодонов: UAA, UAG или UGA. Кодон AUG выполняет двойную функцию: он детерминирует аминокислоту метионин и в определенных последовательностях обозначает начало сегмента, кодирующего белок.

Избыточность кода имеет одну интересную особенность: наибольшее число вариаций в кодонах, детерминирующих данную аминокислоту, приходится на третью позицию (3'-конец триплета). Например, аминокислоты глицин, валин, пролин, аланин и треонин кодируются четырьмя кодонами каждая, и в каждом случае эти четыре кодона различаются только нуклеотидами в третьей позиции (рис. 3.25). Если какая-то аминокислота кодируется двумя кодонами, то последние различаются только пуринами или пиримидинами, находящимися в третьей позиции. И только кодоны для лейцина,

Таблица 3.3. Включение аминокислот в полипептидную цепь при использовании в качестве матрицы синтетических полинуклеотидов с определенными повторами<sup>1)</sup>

Повторяющаяся последовательность	Кодоны	Включаемые аминокислоты
UC	UCU-CUC	Ser-Leu
AG	AGA-GAG	Arg-Glu
UG	UGU-GUG	Cys-Val
AC	ACA-CAC	Thr-His
UUC	UUC; UCU; CUU	Phe; Ser; Leu
AAG	AAG; AGA; GAA	Lys; Arg; Glu
UUG	UUG; UGU; GUU	Leu; Cys; Val
CAA	CAA; AAC; ACA	Gln; Thr; Asp
GUA	GUA; UAG; AGU	Val; Ser
UAC	UAC; ACU; CUA	Tyr; Thr; Leu
AUC	AUC; UCA; CAU	Ile; Ser; His
GAU	GAU; AUG; UGA	Asp; Met
UAUC	UAU-CUA-UCU- -AUC	Tyr-Leu-Ser-Ile
GAUA	GAU-AGA-UAG- -AUA	Нет
UUAC	UUA-CUU-ACU- -UAC	Leu-Thr-Tyr
GUAA	GUA-AGU-AAG- -UAA	Нет

<sup>1)</sup> H.G. Khorana et. al., Cold Spring Harbor Symp. Quant. Biol., 31, 39 (1966). Черточки разделяют кодоны в одной и той же рамке считывания, а точки с запятой в разных рамках.

**РИС. 3.27.**

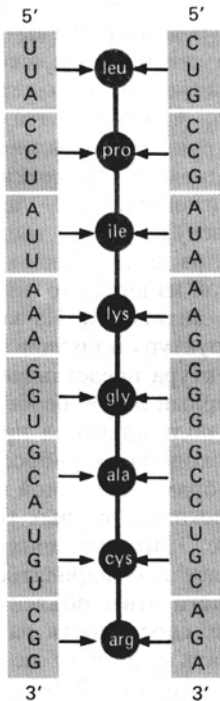
Кодирующая последовательность начинается с кодона AUG и заканчивается терминаторным (стоп) кодоном. Она фланкируется 5'- (5'-лидерная последовательность)

и 3'- (3'-трейлерная последовательность) нетранслируемыми участками.

серина и аргинина различаются нуклеотидами, находящимися в первой, второй или обеих позициях. Поэтому мутации, приводящие к заменам нуклеотидов в третьей позиции, часто не сопровождаются изменением аминокислотной последовательности. Кроме того, код устроен так, что при замене нуклеотидов даже в первой или второй позиции некоторых кодонов в полипептид включается структурно родственная аминокислота, сводя тем самым к минимуму нарушения во вторичной структуре белка. Кодоны для гидрофобных аминокислот, например фенилаланина, лейцина, изолейцина и валина, различаются только одним нуклеотидом. Аналогичная ситуация наблюдается и для кодонов серина и треонина или аланина и глицина.

#### д. Универсальность генетического кода

По-видимому, все прокариоты, а также большинство эукариот (по крайней мере в том, что касается ядерного генетического материала) пользуются одним и тем же словарем кодонов независимо от того, представлен ли их геном ДНК или РНК. В этом смысле нередко говорят, что код универсален. Тем не менее частота использования кодонов-синонимов варьирует как на уровне организмов, так и на уровне мРНК. Если действительно какие-то кодоны используются в большинстве кодирующих белки последовательностей чаще, чем другие, то это должно найти отражение в относительном содержании в клетке различных тРНК, расшифровывающих кодоны-синонимы. Так, если кодон AGA встречается в мРНК какого-либо организма сравнительно редко, то и содержание тРНК<sup>Arg</sup>, расшифровывающей AGA, должно быть невелико. В некоторых случаях выбор кодонов для конкретной кодирующей последовательности определяется задачами, не связанными с их трансляцией. Например, определенное взаимное расположение кодонов может благоприятствовать образованию специфической вторичной структуры мРНК; таким образом, присутствие конкретных кодонов может влиять на готовность мРНК к трансляции и тем самым играть регуляторную роль. В некоторых случаях экспериментальная замена существующих кодонов их синонимами приводит к изменению стабильности мРНК при полном сохранении ее способности к правильной трансляции.

**РИС. 3.28.**

Две разные нуклеотидные последовательности могут кодировать один и тот же полипептид.

Редкие исключения в стандартном словаре кодонов обнаружены у инфузорий и в митохондриальных генах. Например, в митохондриях млекопитающих кодон UGA читается как триптофан и митохондриальная ДНК кодирует тРНК<sup>Trp</sup>, чей антикодон UCA спаривается с UGA почти так же прочно, как с нормальным триптофановым кодоном UGG. В митохондриях млекопитающих кодоны AGA и AGG прочитываются как сигналы терминации. AUU, AUC, AUA и AUG служат иницилирующими кодонами, а AUA кодирует метионин вместо изолейцина. В митохондриях дрожжей триплеты CUU, CUC, CUA и CUG кодируют треонин, а не лейцин.

### 3.5. АППАРАТ ТРАНСЛЯЦИИ

Основными участниками процесса считывания информации, закодированной в последовательности мРНК, являются аминоацил-тРНК-синтетазы, тРНК, рибосомы, белки, связанные с рибосомами, и некоторые другие белки. Они ответственны за инициацию, элонгацию и терминацию сборки полипептида. В этом разделе мы опишем свойства каждого из перечисленных компонентов, а в следующем обсудим, как они функционируют на разных этапах процесса трансляции.

#### а. Присоединение аминокислот к «родственным» тРНК

Функционирование тРНК при трансляции сводится к двум уникальным процессам. Первый из них состоит в присоединении аминокислоты к 3'-концу родственной тРНК с помощью специфической **аминоацил-тРНК-синтетазы**, второй – в специфическом связывании аминоацил-тРНК с соответствующим кодоном мРНК, находящейся в комплексе с рибосомой. Ключевой особенностью обеих реакций является их специфичность, поскольку сбои в образовании аминоацил-тРНК или связывании аминоацил-тРНК с соответствующим кодоном приведут к ошибкам в экспрессии генов. Поэтому очень важно понять природу такой специфичности, механизм ее обеспечения и последствия нарушений точности на этих двух этапах трансляции.

*Основные особенности структуры тРНК.* В любой клетке присутствует очень много разных тРНК. Молекула тРНК состоит обычно из 75–85 нуклеотидов и содержит уникальный тринуклеотид, который определяет, какую аминокислоту эта тРНК присоединяет и с каким кодоном она может спариться. На основании данных о нуклеотидной последовательности более чем 150 отдельных видов тРНК, выделенных из клеток про- и эукариот, были построены компьютерные модели внутримолекулярного комплементарного спаривания оснований в молекуле тРНК. Был сделан вывод, что практически все тРНК, независимо от их нуклеотидной последовательности, обладают характерной вторичной структурой, которую называют структурой «**клеверного листа**» из-за наличия в ней трех шпилек (рис. 3.29). Реальность предсказанной структуры была подтверждена данными о разной химической чувствительности оснований, одни из которых спарены, а другие нет (последние образуют так называемые петли).

Большинство молекул тРНК, имеющих форму клеверного листа, содержат четыре области, каждая из которых обладает инвариантными свойствами независимо от аминокислотной специфичности

тРНК. 1. На 3'-конце молекулы всегда находятся четыре неспаренных нуклеотида, причем три из них – это обязательно ССА. 5'- и 3'-концы цепи РНК образуют акцепторный стебель. Цепи удерживаются вместе благодаря комплементарному спариванию семи нуклеотидов 5'-конца с семью нуклеотидами, находящимися вблизи 3'-конца. 2. У всех молекул имеется шпилька ТψС, обозначаемая так потому, что она содержит два необычных остатка: риботимидин (Т) и псевдоуридин (ψ). Шпилька состоит из двухцепочечного стебля из пяти спаренных оснований, включая пару G-C, и петли длиной семь нуклеотидов. Тринуклеотид ТψС всегда расположен в одном и том же месте петли. 3. В антикодонной шпильке стебель всегда представлен семью спаренными основаниями. Триплет, комплементарный родственному кодону, – **антикодон** – находится в петле, состоящей из семи нуклеотидов. С 5'-конца антикодон фланкируют инвариантный остаток урацила и модифицированный цитозин, а к его 3'-концу примыкает модифицированный пурин, как правило аденин. 4. Еще одна шпилька состоит из стебля длиной три-четыре пары нуклеотидов и петли варьирующего размера, часто содержащей урацил в востановленной форме – дигидроурацил (DU).

Наиболее сильно варьируют нуклеотидные последовательности стеблей, число нуклеотидов между антикодонным стеблем и стеблем ТψС (вариабельная петля), а также размер петли и локализация остатков дигидроурацила в DU-петле.

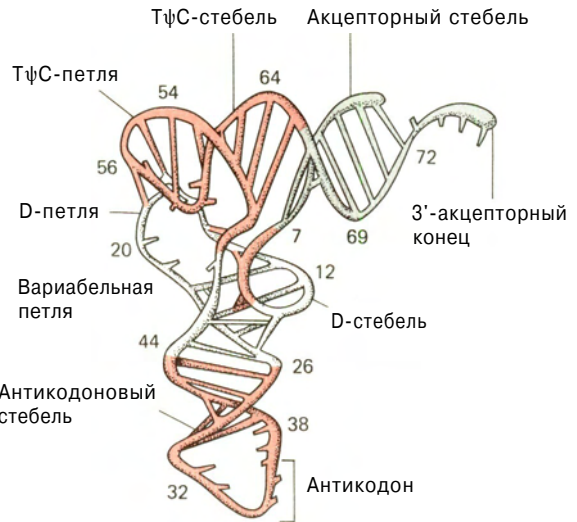
Рентгеноструктурный анализ некоторых молекул тРНК позволил выявить их характерную четвертичную структуру (рис. 3.29, Б–Г). Эта структура более компактна, чем структура «клеверного листа». Она образуется благодаря внутримолекулярным взаимодействиям, сближающим DU- и ТψС-шпильки. В результате молекула тРНК выглядит так, как будто она состоит из двух взаимно перпендикулярных частей – в одной из них находится акцепторный участок, в другой – антикодон. Из-за такого общего вида молекулы структура получила название L-конфигурации. L-структура представляется более адекватной, чем «клеверный лист», особенно если учесть, что тРНК играет роль адаптора при взаимодействии кодона и антикодона на рибосоме.

Обычно акцепторами для одной и той же аминокислоты служат несколько разных тРНК (изоакцепторные тРНК) (рис. 3.30), имеющих разные антикодоны, что позволяет им спариваться с кодонами-синонимами. Отчасти этим объясняется и вырожденность кода, т.е. способность разных антикодонов детерминировать одну и ту же аминокислоту.

*Этерификация молекул тРНК.* Для выполнения функции адаптора в процессе трансляции мРНК тРНК должна связаться с аминокислотой, соответствующей своему антикодону. Это происходит в ре-



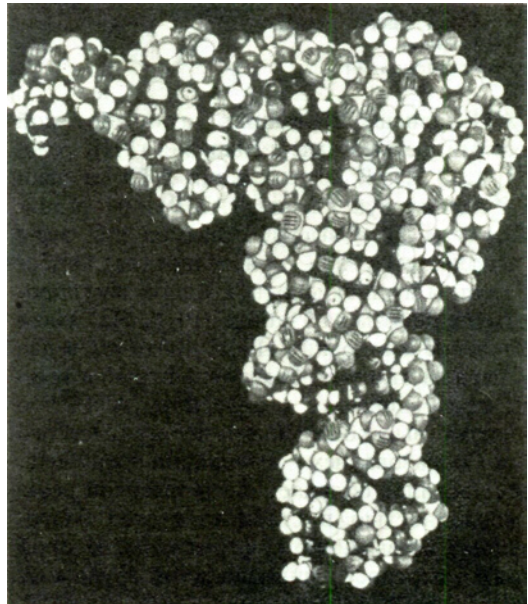
А



Б



В

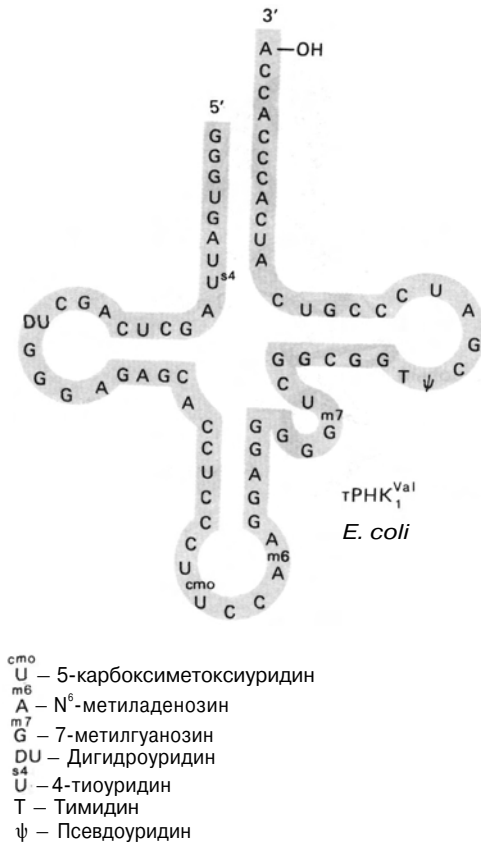


Г

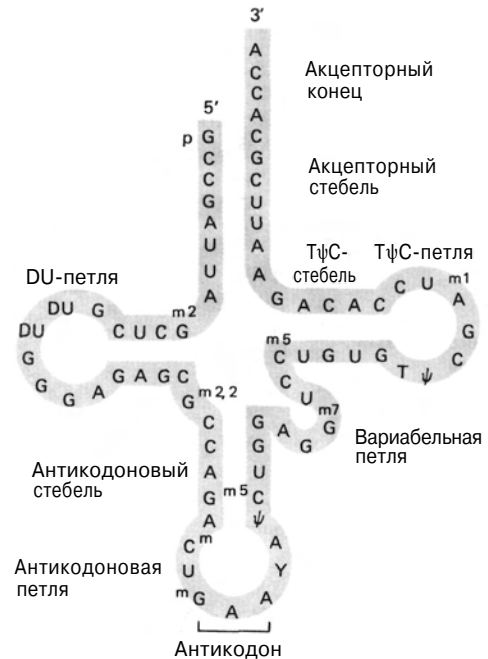
**РИС. 3.29.**

Типичная эукариотическая тРНК: дрожжевая тРНК<sup>Phe</sup>. А. Развернутая структура, имеющая форму клеверного листа. Б. Свернутый «клеверный лист»; рибозофосфатный остов изображен в виде непрерывной ленты, а водородные связи - в виде стержней. В и Г - проволочная и пространственная модели свернутого «клеверного лис-

та» соответственно. Структура тРНК на рис. Г повернута относительно структуры на рис. Б на 180°. Антикодон во всех случаях находится внизу структуры. [Рис. Б, В и Г любезно предоставлены S.M. Kim. См.: S.M. Kim et al., Science, **185** (1974), p. 435.]

**РИС. 3.30.**

Изоакцепторные тРНК для валина у *E. coli*. Эти две тРНК различаются своими нуклеотидными последовательностями, антикодонами, типом и положением модифицированных оснований.

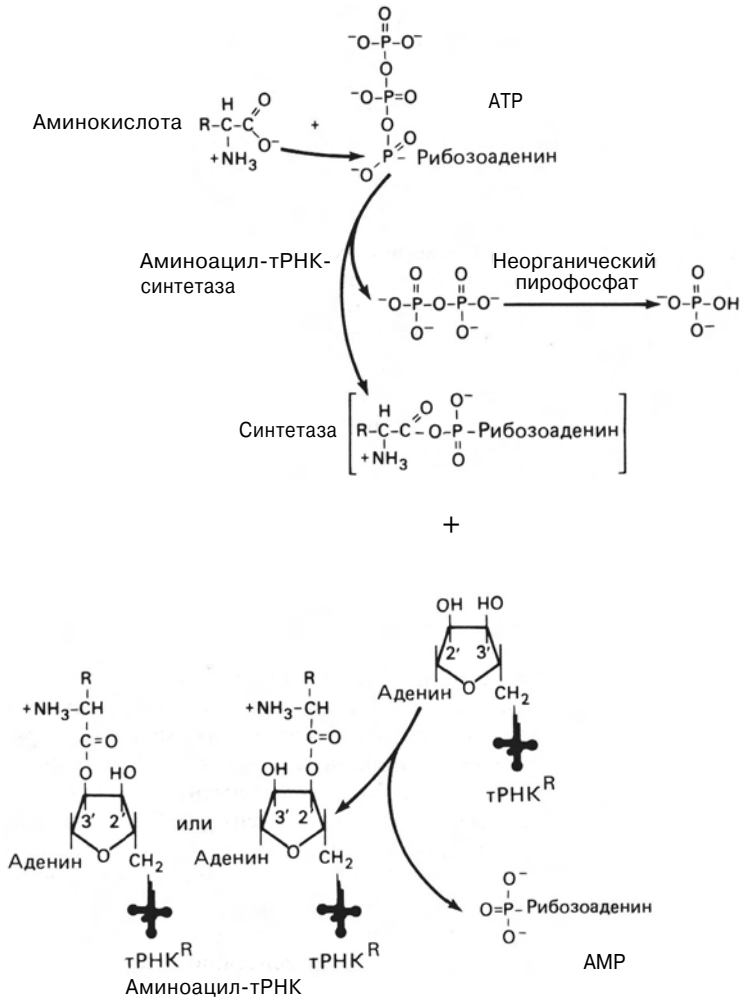


зультате АТР-зависимой реакции, катализируемой специфическими ферментами аминокил-тРНК-синтетазы (рис. 3.31). В ходе реакции АТР расщепляется на 5'-адениловую кислоту (АМР) и неорганический фосфат (РР), а высвобождаемая при этом энергия используется для присоединения карбоксильной группы аминокислоты к одной из гидроксильных групп рибозы на 3'-конце тРНК. На самом деле образование аминокил-тРНК проходит в два этапа. На первом этапе карбоксильная группа аминокислоты присоединяется к α-фосфату АТР, что сопровождается высвобождением неорганического фосфата и образованием аминокил-аденилата. Аминокил-аденилат обладает очень высокой реакционной способностью и стабилизируется благодаря прочному связыванию с ферментом. Второй этап состоит в переносе аминокильной группы от связанного с ферментом аминокил-аденилата на 2'- или 3'-ОН-группу концевой рибозы тРНК. Потенциала переноса ацильной группы аминокил-тРНК более чем достаточно для образования пептидной связи без дополнительного поступления энергии.

Ключевой особенностью реакции, приводящей к аминокетилированию тРНК, является специфичность участвующих в ней ферментов. Присоедине-

ние к тРНК каждой из 20 аминокислот, встречающихся в белках, катализируется определенной аминокил-тРНК-синтетазой. Фермент должен отличить одну аминокислоту от 19 других и перенести ее к одной или нескольким изоакцепторным тРНК из имеющихся примерно 75 других тРНК. Вспомним при этом, что многие аминокислоты очень сходны по структуре: лейцин, валин и изолейцин; валин и треонин; аспарагиновая и глутаминовая кислоты. Аминокил-тРНК-синтетазы должны отличить «свои» тРНК от всех других, несмотря на удивительное сходство их вторичной и третичной структур. Поэтому необходимо, чтобы ферменты обладали очень высокой специфичностью, позволяющей им сделать правильный выбор из столь родственных структур и избежать ошибок при синтезе белка.

*Комментарий по поводу структур аминокил-тРНК-синтетаз и их способности к узнаванию аминокислот и родственных тРНК.* Многие аминокил-тРНК-синтетазы удалось очистить. Некоторые из них состоят из одной полипептидной цепи, другие — из двух или четырех идентичных цепей, каждая мол. массой от 35 до 115 кДа. Некоторые димерные и тетрамерные ферменты состоят из субъединиц двух типов. Четкой корреляции между раз-

**РИС. 3.31.**

Образование аминоксил-тРНК. Аминоацил-тРНК-синтетаза катализирует образование ферментсвязанного промежуточного продукта аминоксил-аденилата, с помощью которого аминоксилная единица переносится на тРНК.

мером молекулы фермента или характером его субъединичной структуры и специфичностью не существует.

Исследования взаимодействия между аминоксил-тРНК-синтетазами и родственными им тРНК не позволили выяснить природу их высокой специфичности. Большинство работ показало, что специфичность фермента определяется его прочным связыванием с акцепторным концом тРНК, DU-участком и варибельной петлей. Некоторые ферменты, по-видимому, не распознают антикодоновый триплет и катализируют реакцию аминоксилации даже при измененном антикодоне. Однако отдельные ферменты проявляют пониженную активность по отношению к таким модифицированным тРНК и при замене антикодона присоединяют не ту аминоксилоту. Следовательно, в некоторых случаях существенным оказывается и взаимодействие с антикодоновой петлей. В любом случае акцепторный конец тРНК должен быть ориентирован так, чтобы

каталитический центр фермента смог перенести связанный аминоксил-аденилат к концевому нуклеотиду тРНК.

В какой-то степени способность фермента присоединять нужную аминоксилоту к родственной тРНК зависит от специфического связывания аминоксилоты. Однако, если безошибочное распознавание родственных аминоксилот невозможно, синтетазы могут исправлять ошибки, происходящие при присоединении. Например, нельзя полностью исключить возможность связывания валина изолейцил-тРНК-синтетазой из-за сходства размера и структуры изолейцина и валина. Дефект в специфичности обнаруживается в первой же реакции: изолейцил-тРНК-синтетаза образует ферментсвязанный валил-аденилат, хотя и с меньшей эффективностью, чем изолейцил-аденилат, однако такой активированный валин не связывается ни с тРНК<sup>Val</sup>, ни с тРНК<sup>Ile</sup>. Вместо этого ферментсвязанный валил-АМР быстро гидролизуется в присутствии тРНК<sup>Ile</sup>,

и образование валил-тРНК<sup>lle</sup> предотвращается. Подобный механизм позволяет валил-тРНК-синтетазе различать валин и треонин, а метионил-тРНК-синтетазе отличать треонин от метионина. Очевидно, что аминоксил-тРНК-синтазы пользуются механизмом коррекции с целью предотвращения неизбежных ошибок в аминокислотировании тРНК. И напротив, механизм, с помощью которого удалялась бы уже присоединенная к тРНК неправильная аминокислота, отсутствует. В таких случаях аминокислота занимает неправильную позицию в белке. Частота таких ошибок очень низка. В гемоглобине кролика, например, валин оказывается в местах, обычно занимаемых изолейцином, только в одном из 25000–50000 возможных случаев (число остатков изолейцина в белке). Таким образом, точность первого шага на сложном пути считывания генетической информации обеспечивается четкой работой разных аминоксил-тРНК.

### б. На рибосомах осуществляются спаривание аминоксил-тРНК с кодонами и сборка белковых цепей

Во всех клетках имеются рибосомы, играющие ключевую роль в синтезе белка; их число колеблется от 20000 до 50000 в зависимости от белоксинтезирующей активности клеток. Рибосомы индифферентны в отношении синтезируемых ими белков или тех клеточных мишеней, к которым они направляют синтезированные продукты. Тип синтезируемого рибосомой белка в каждом синтетическом цикле диктуется мРНК, с которой рибосома оказалась связанной. Внутри- или внеклеточная локализация белков определяется их структурными особенностями и — в зависимости от этих особенностей — характером взаимодействия со специализированными мембранами и органеллами.

Рибосомы про- и эукариот обладают в общем очень сходными структурой и функциями. Тем не менее из-за различий в структуре и организации про- и эукариотических мРНК и из-за того, что процессы транскрипции и трансляции у эукариот являются сопряженными во времени и в пространстве, тонкие различия между рибосомами про- и эукариот имеются. Типичными прокариотическими рибосомами являются рибосомы *E. coli*, и поскольку их структура и функции изучены лучше остальных, мы используем эту модель в последующем обсуждении. Для сравнения мы остановимся на некоторых структурных особенностях рибосом эукариот.

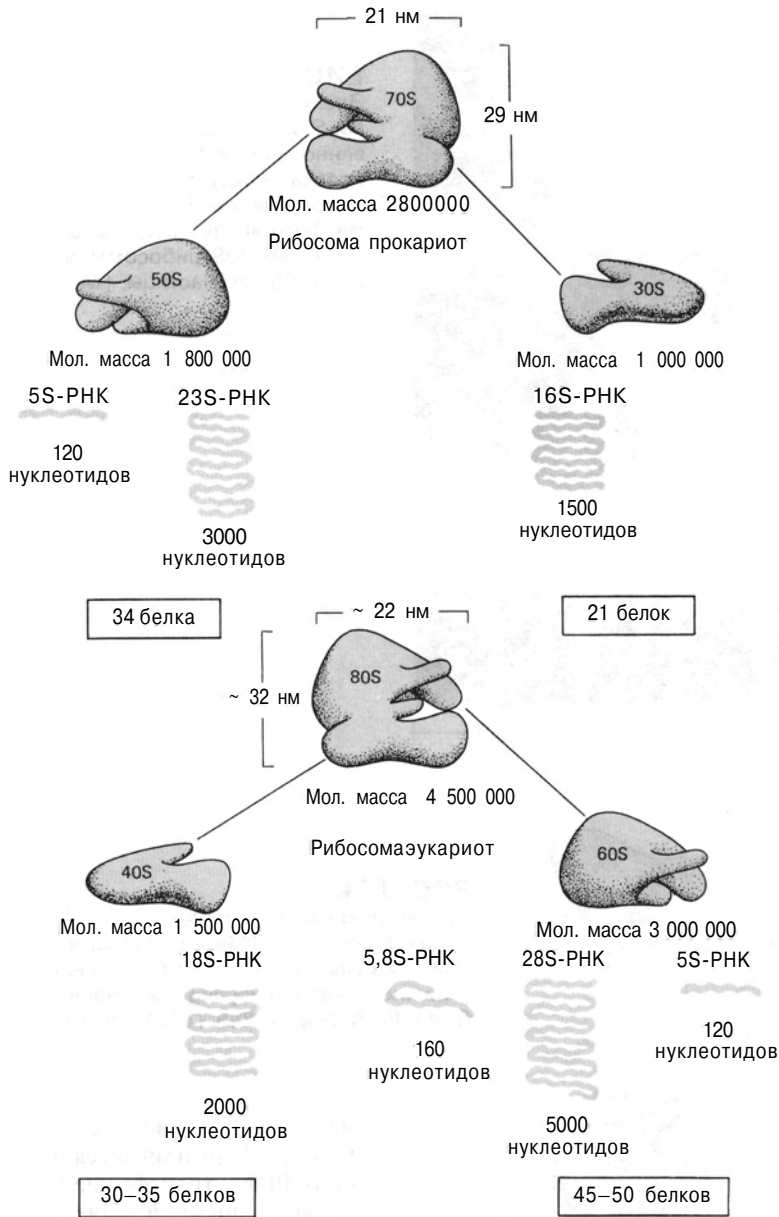
**Строение рибосомных частиц.** Рибосомы прокариот (называемые 70S-частицами из-за их седиментационных свойств) состоят из малых (30S) и больших (50S) субчастиц (рис. 3.32). 30S-субчастицы состоят из единственной молекулы рРНК (16S) разме-

ром 1542 нуклеотида и 21 белка с разной мол. массой (S1–S21). 50S-субчастицы содержат две рРНК — большую (23S), состоящую из 2904 нуклеотидов, и более мелкие (5S), из 120 нуклеотидов; они связаны с 34 разными белками (L1–L34). Нуклеотидные и аминокислотные последовательности всех рРНК и белков известны. Электронно-микроскопические исследования 70S-рибосом (рис. 3.33) и построения их трехмерных моделей (рис. 3.34) показали, что малая и большая субчастицы соприкасаются в нескольких точках, но самой характерной особенностью является наличие бороздки между ними, необходимой, по-видимому, для размещения в ней мРНК на время трансляции.

И малая, и большая рибосомные субчастицы могут диссоциировать на составляющие молекулы РНК и белка. Более того, даже после отделения друг от друга молекулы всех РНК и белков способны восстанавливать исходную функционально активную рибосомную субчастицу, если их смешать в соответствующих условиях. Это означает, что вся информация о сборке мультимерного комплекса заключена в структуре его компонентов. Эксперименты по реконструированию рибосом позволяют лучше понять характер взаимодействия между этими компонентами и определить возможный порядок, в котором собираются белки и РНК *in vivo*. Кроме того, в подобных экспериментах можно проверить совместимость эквивалентных РНК или белковых субчастиц из различных источников (например, из отдаленно родственных прокариот или из прокариот и эукариот). Далее, с помощью этого метода можно оценить способность мутантных РНК или белков к взаимодействию с восстановлением структуры рибосом и проявлению различных видов активности, присущих реконструированным рибосомам.

Рибосомы эукариот, находящиеся в цитозоле, также состоят из малых (40S) и больших (60S) субчастиц (рис. 3.32). Малые субчастицы содержат одну молекулу РНК (18S) размером 1900 нуклеотидов и 30–35 белков; большие — три цепи РНК длиной 120 (5S), 160 (5,8S) и 4800 (28S) нуклеотидов и 45–50 белков. Рибосомы митохондрий и хлоропластов отличаются от цитозольных рибосом. Как правило, они меньше и содержат меньшее количество белков и различных рРНК. Данные по физическому и химическому реконструированию более сложных эукариотических хромосом значительно уступают аналогичным данным, полученным для *E. coli*.

Необходимо выделить два важных в функциональном отношении участка, образующихся при ассоциации субчастиц в процессе формирования 70S-рибосомы (рис. 3.35). Это участки, в которых происходит связывание двух тРНК — одной, присоединенной к растущей белковой цепи (P-участок),

**РИС. 3.32.**

Состав типичных рибосом про- и эукариот.

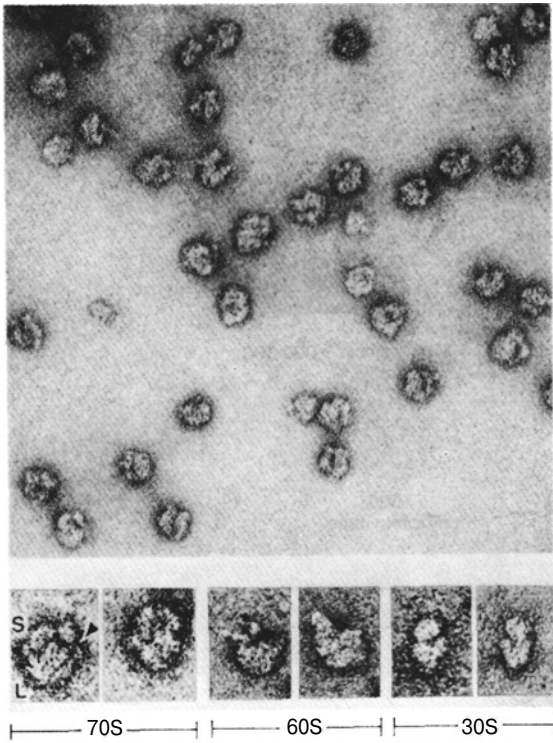
и другой, несущей следующую добавляемую к цепи аминокислоту (А-участок) (разд. 3.6. б).

*Особые тРНК и некоторые вспомогательные белки, участвующие в трансляции.* Как у про-, так и у эукариот имеются два вида тРНК, которые связывают метионин. У прокариот они обозначаются как  $tRNA_{F}^{Met}$  и  $tRNA_{M}^{Met}$ , а у эукариот — соответственно  $tRNA_{1}^{Met}$  и  $tRNA_{M}^{Met}$ . Каждая из обоих видов тРНК<sup>Met</sup> как у про-, так и у эукариот аминоацети-

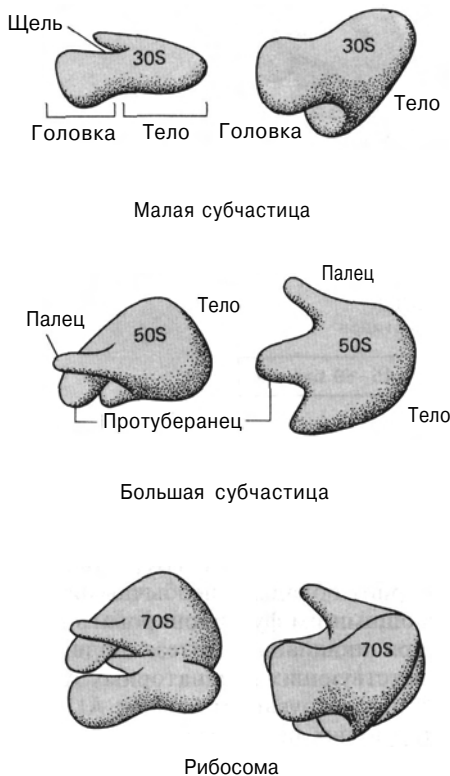
руется метионином с помощью соответствующих аминоацил-тРНК-синтетаз.  $tRNA_{F}^{Met}$  прокариот и  $tRNA_{1}^{Met}$  эукариот обладают необычными свойствами, позволяющими им функционировать в качестве адапторов при инициации синтеза полипептидной цепи в соответствующих инициаторных AUG-кодонах.  $tRNA_{M}^{Met}$  про- и эукариот узнают AUG-кодона в белок-кодирующих последовательностях.

У прокариот аминогруппа метионил-тРНК<sub>F</sub><sup>Met</sup>



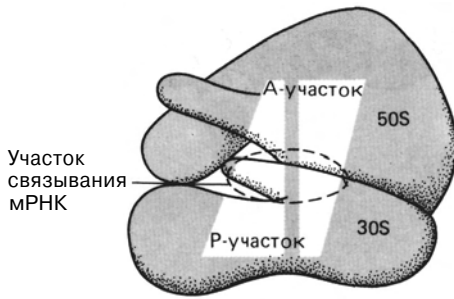
**РИС. 3.33.**

Электронные микрофотографии рибосом *E. coli*. Буквами S и L на нижнем левом фото обозначены соответственно 30S- и 50S-субчастицы. Стрелкой отмечена бороздка между двумя субчастицами. На верхней фотографии видны 70S-моносомы при увеличении 350000. На фото внизу представлены при большем увеличении (560000) 70S-рибосомы, а также большая (50S) и малая (30S) субчастицы. [M. Boublik, *Cytologie* **14** (1977), p. 203.]

**РИС. 3.34.**

Реконструированные изображения 30S-, 50S- и 70S-рибосом *E. coli*, полученные с помощью данных электронной дифракции. Субчастицы представлены в двух взаимно перпендикулярных проекциях. [J.A. Lake, *Scientific American*, **245** (1981), p. 84.]

( $\text{met-ТРНК}_F^{\text{Met}}$ ), но не метионил- $\text{ТРНК}_M^{\text{Met}}$  ( $\text{met-ТРНК}_M^{\text{Met}}$ ) формируется особым ферментом (метионил- $\text{ТРНК-трансформилазой}$ ) до  $\text{Fmet-ТРНК}_F^{\text{Met}}$  с использованием в качестве донора формильной группы  $\text{N}^{10}$ -формилтетрагидрофолата (рис. 3.36). Очевидно, трансформилаза отличает  $\text{met-ТРНК}_F^{\text{Met}}$  от  $\text{met-ТРНК}_M^{\text{Met}}$ .  $\text{Fmet-ТРНК}_F^{\text{Met}}$  используется исключительно для инициации белковых цепей, а  $\text{met-ТРНК}_M^{\text{Met}}$  — только для декодирования внутренних метиониновых кодонов. Несмотря на то что  $\text{ТРНК}_I^{\text{Met}}$  эукариот также используется только для инициации, ее метионильная группа не подвергается формилированию. Очевидно, некие особые свойства, присущие  $\text{ТРНК}_I^{\text{Met}}$  и необходимые для выполнения ею специальной инициаторной функции, связаны исключительно с ее нуклеотидной последовательностью и/или трехмерной структурой.

**РИС. 3.35.**

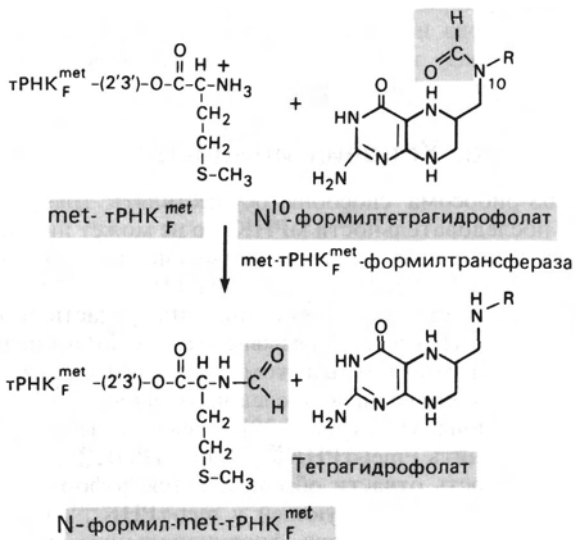
Схема, показывающая расположение Р- и А-участков в 70S-рибосоме *E. coli*. В формирование этих участков, по-видимому, вовлечены как 30S-, так и 50S-субчастицы. Показана предполагаемая область связывания мРНК.

Известны белки, которые только временно, на период трансляции, связываются с рибосомами. Они играют важную роль при инициации, элонгации и терминации синтеза белковой цепи. Прежде чем подробно обсудить эти процессы (разд. 3.6), мы познакомим читателя с такими белками и кратко опишем их свойства и роль в трансляции.

Эти белки, называемые факторами инициации и обозначаемые IF-1, IF-2 и IF-3, необходимы для инициации трансляции мРНК с образованием белков. IF-1 и IF-3 связываются с 30S-субчастицей, что обеспечивает ее взаимодействие с комплексом, образуемым IF-2 и уникальной инициаторной аминоксил-тРНК, Fmet-тРНК<sup>Met</sup>. Именно комплекс 30S-субчастицы со всеми факторами инициации и Fmet-тРНК<sup>Met</sup> узнает соответствующий AUG-кодон в

мРНК. Когда формирование инициаторного комплекса завершается присоединением 50S-субчастицы, все три инициаторных белка отсоединяются. Для доставки следующей аминоксил-тРНК к рибосоме нужны два белка, называемые факторами элонгации, EF-Tu и EF-Ts. EF-Tu с присоединенным к нему GTP связывает все аминоксил-тРНК, кроме Fmet-тРНК<sup>Met</sup>, и доставляет их к А-участку комплекса 70S-рибосома-мРНК. По ходу процесса образуется EF-Tu-GDP, и для восстановления EF-Tu-GTP нужен EF-Ts. В элонгации цепи участвует еще один белок, EF-G; он обеспечивает перемещение рибосомы по мРНК при трансляции кодонов. Для терминации синтеза полипептидной цепи в одном из трех стоп-кодонов также нужны особые белки, только временно присоединяющиеся к трансляционному комплексу. Известны три таких белка: RF-1 вызывает отделение полипептидной цепи при считывании кодонов UAA и UAG; RF-2 действует аналогичным образом при считывании UAA и UGA, а EF-3 может облегчить работу двух других факторов.

Все многообразие вспомогательных белков, необходимых на различных этапах сборки белков в клетках эукариот, изучено значительно слабее, а их функции пока недостаточно ясны. Для инициации нужно не менее пяти таких белков; некоторые из них по своим функциям, по-видимому, аналогичны прокариотическим белкам IF-2 и IF-3. Более сложно устроен и аппарат элонгации; он снабжен значительно большим количеством белков, участвующих как в связывании аминоксил-тРНК с рибосомами, так и в перемещении рибосом. Терминация трансляции у эукариот также изучена слабо; тем не менее здесь, по-видимому, необходим только один белок (eRF). Важным этапом терминации или отделения белковой цепи от мРНК является гидролиз GTP.

**РИС. 3.36.**

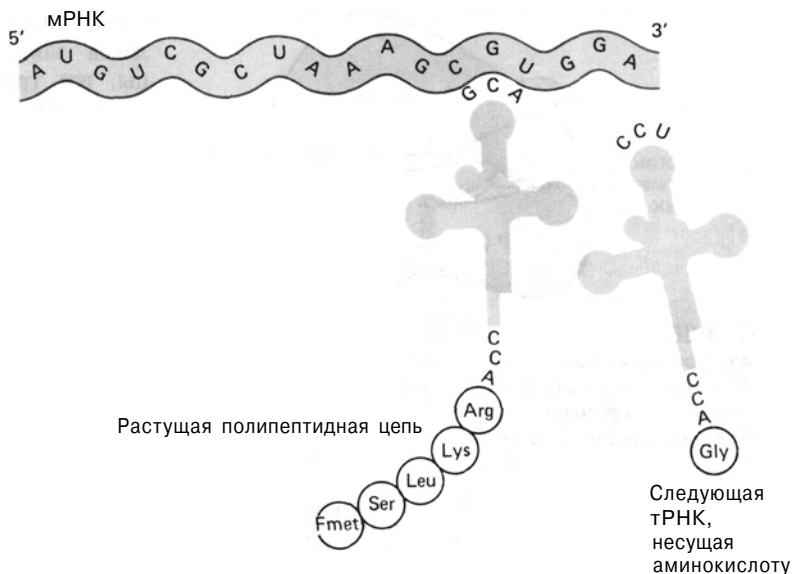
Формирование метионил-тРНК<sup>Met</sup> у *E. coli*.

### 3.6. ТРАНСЛЯЦИЯ мРНК У ПРОКАРИОТ

Зная всех участников процесса, мы можем теперь приступить к рассмотрению химических реакций, протекающих при синтезе полипептидов, т.е. реакций, участвующих в собственно трансляции. Несмотря на то что этот процесс протекает непрерывно от старта к финишу, обычно выделяют три его этапа: инициацию, элонгацию и терминацию. Рассматривая каждый из этапов направляемого мРНК синтеза полипептидной цепи, мы должны учитывать два основных свойства этого процесса. Во-первых, полипептидные цепи синтезируются однонаправленно: с amino-конца к карбокси-концу (рис. 3.37). При этом карбоксильная группа уже образовавшегося участка полипептидной цепи соединяется с amino-группой следующей присоединяемой аминокислоты с помощью пептидной связи. Это может произойти,

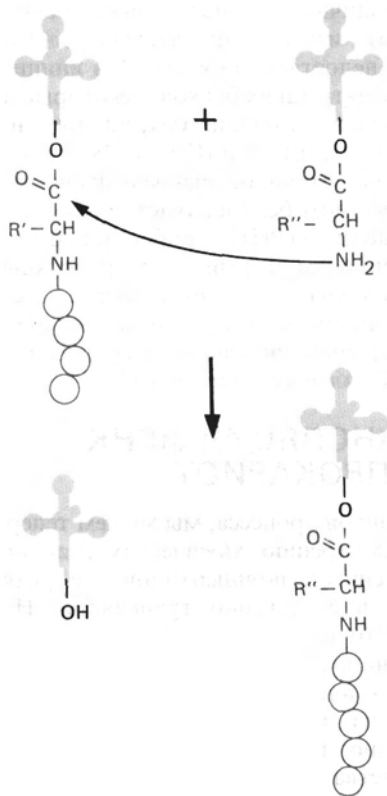
**РИС. 3.37.**

Сборка полипептидной цепи. Кодоны мРНК считываются один за другим в направлении 5'→3', начиная с инициаторного кодона AUG, который связывается с N-формилметионил-тРНК<sup>Met</sup>. Полипептид синтезируется от N- к С-концу.



Р-участок

А-участок

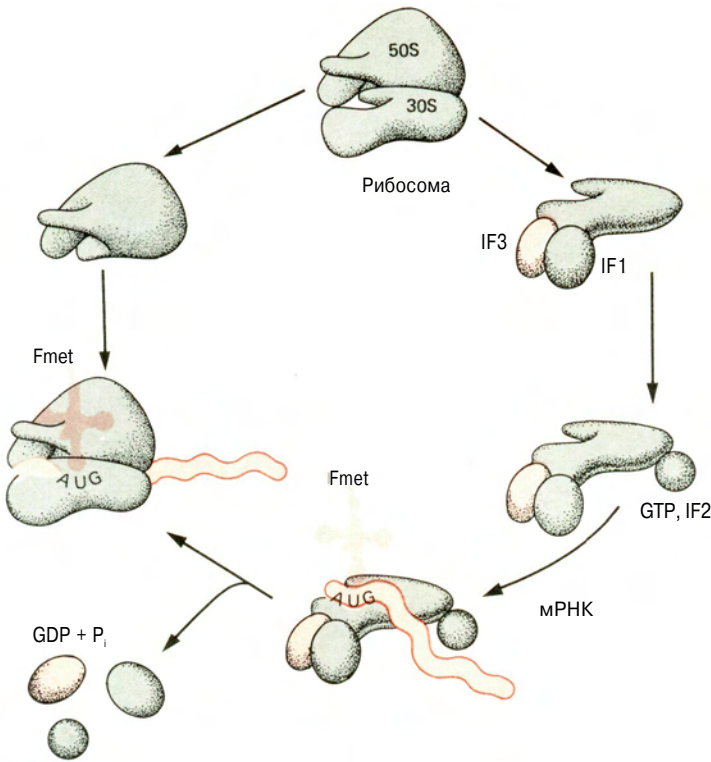
**РИС. 3.38.**

Этапы пептидилтрансферазной реакции. Азот аминогруппы присоединяющейся аминокислоты осуществляет нуклеофильную атаку активированной карбоксильной группы растущей цепи, в результате чего образуется новая пептидная связь.

лишь если карбоксильный конец растущей полипептидной цепи находится в активированном состоянии. Как мы уже отмечали, необходимая для этого энергия поступает в результате присоединения карбоксильной группы растущей полипептидной цепи и каждой присоединяемой аминокислоты к тРНК (рис. 3.38). Во-вторых, считывание мРНК начинается с кодона AUG, который обозначает 5'-конец кодирующей последовательности и детерминирует N-концевую аминокислоту синтезируемого полипептида (рис. 3.37). При инициации первая и вторая молекулы аминоацил-тРНК спариваются с первыми двумя кодонами мРНК. Далее трансляция продолжается в направлении 5'→3' кодон за кодоном до тех пор, пока не достигнет стоп-сигнала, расположенного сразу же за кодоном, детерминирующим С-концевую аминокислоту.

### а. Условия инициации

70S-рибосома способна осуществлять трансляцию последовательности мРНК, но не может иницировать этот процесс. При связывании инициаторных белков IF-1 и IF-2 с 30S-субчастицей происходит диссоциация 70S-рибосомы. 30S-субчастица в комплексе с IF-1 и IF-3 связывает IF-2, GTP и Fmet-тРНК<sup>Met</sup>. Такой полный комплекс связывается с 5'-концом кодирующей последовательности мРНК вблизи кодона AUG (рис. 3.39). Очевидно, IF-2 способен отличить Fmet-тРНК<sup>Met</sup> от met-тРНК<sup>Met</sup>, и эта специфичность отчасти обеспечивается N-формильной группой, отсутствующей у met-тРНК<sup>Met</sup>. Формирование полноценного функционального комплекса инициации завершается ассоциацией 50S-субчастицы с преинициаторным комплексом. С образо-



**РИС. 3.39.**

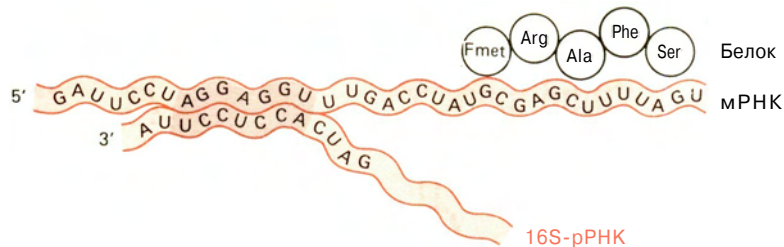
События, предшествующие инициации трансляции. При связывании IF-1 и IF-3 с 30S-субчастицей происходит диссоциация 70S-рибосомы. Затем образуется преиницирующий комплекс, состоящий из 30S-субчастицы, N-формилметирил-тРНК<sup>Met</sup> и мРНК. Последующая ассоциация этого комплекса с 50S-субчастицей сопровождается отделением IF-1, IF-2 и IF-3 и гидролизом одной молекулы GTP до GDP.

ванием функциональной 70S-субчастицы отделяются все три белка инициации.

Как узнается первый кодон? Связывание 30S-субчастицы с мРНК находится под строгим контролем нуклеотидной последовательности, расположенной примерно за 10 нуклеотидов до 5'-конца инициаторного кодона. Взаимодействию способствует комплементарное спаривание этой богатой пуринами последовательности из пяти-восьми нуклеотидов, называемой **последовательностью Шайна-Дальгарно**,

с полипиримидиновым участком, находящимся вблизи 3'-конца 16S-pPHK (рис. 3.40). Эффективность инициации существенно зависит от степени комплементарности между последовательностями Шайна-Дальгарно и 16S-pPHK и от расстояния пурин-богатого участка до кодона AUG. Эта особенность наряду с другими, о которых будет сказано позднее, и объясняет различия в эффективности трансляции различных мРНК.

Процесс инициации зависит также от вторичной



5' A A U C U U G G A G G C U U U U U U A U G G U U C G U U C U 3' фX174: A-белок  
 5' U A A C U A A G G A U G A A A U G C A U G U C U A A G A C A 3' Qβ: репликаза  
 5' U C C U A G G A G G U U U G A C C U A U G C G A G C U U U 3' RI7: A-белок  
 5' A U G U A C U A A G G A G G U U G U A U G G A A C A A C G C 3' λ: cro-белок

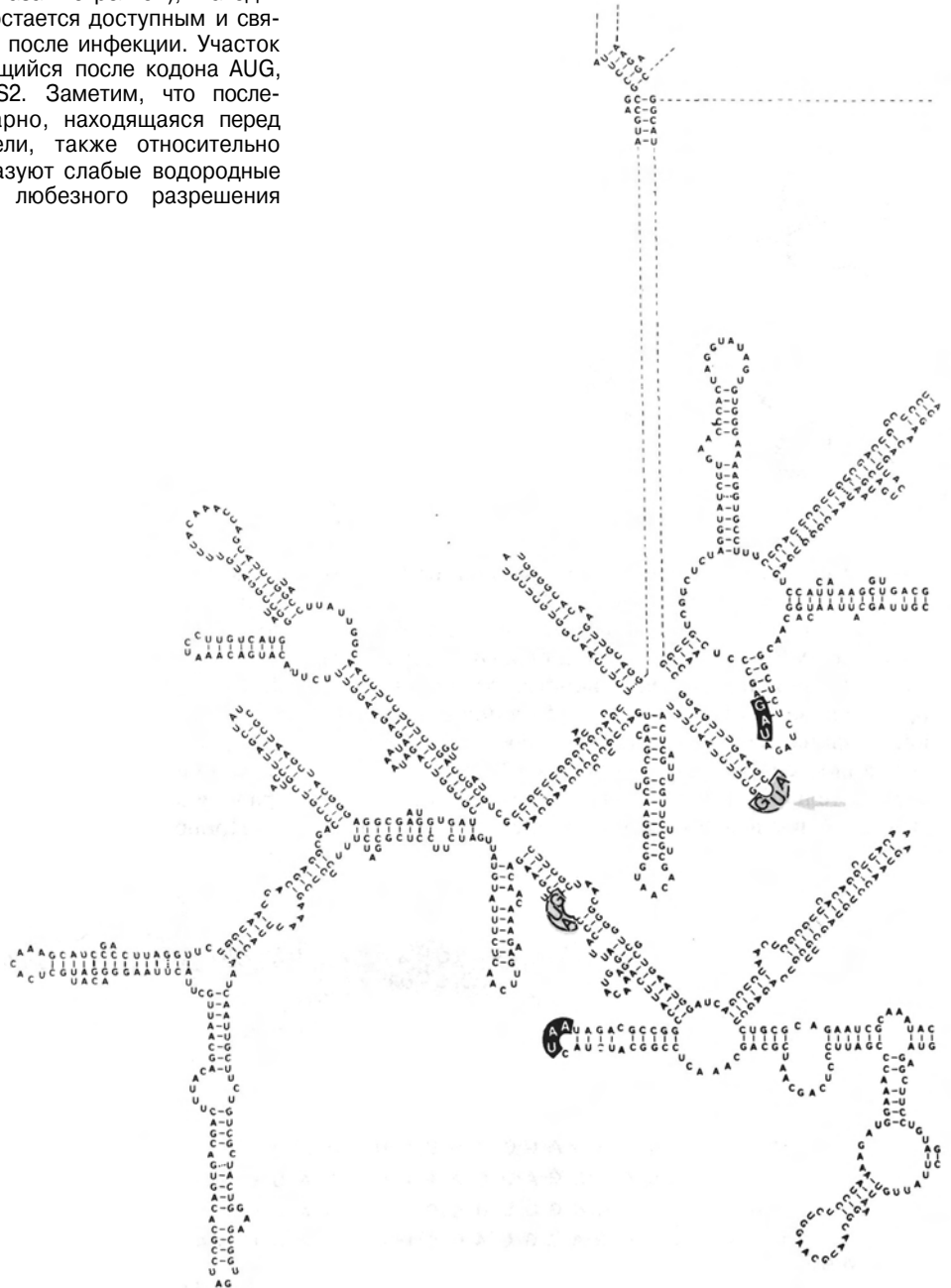
**РИС. 3.40.**

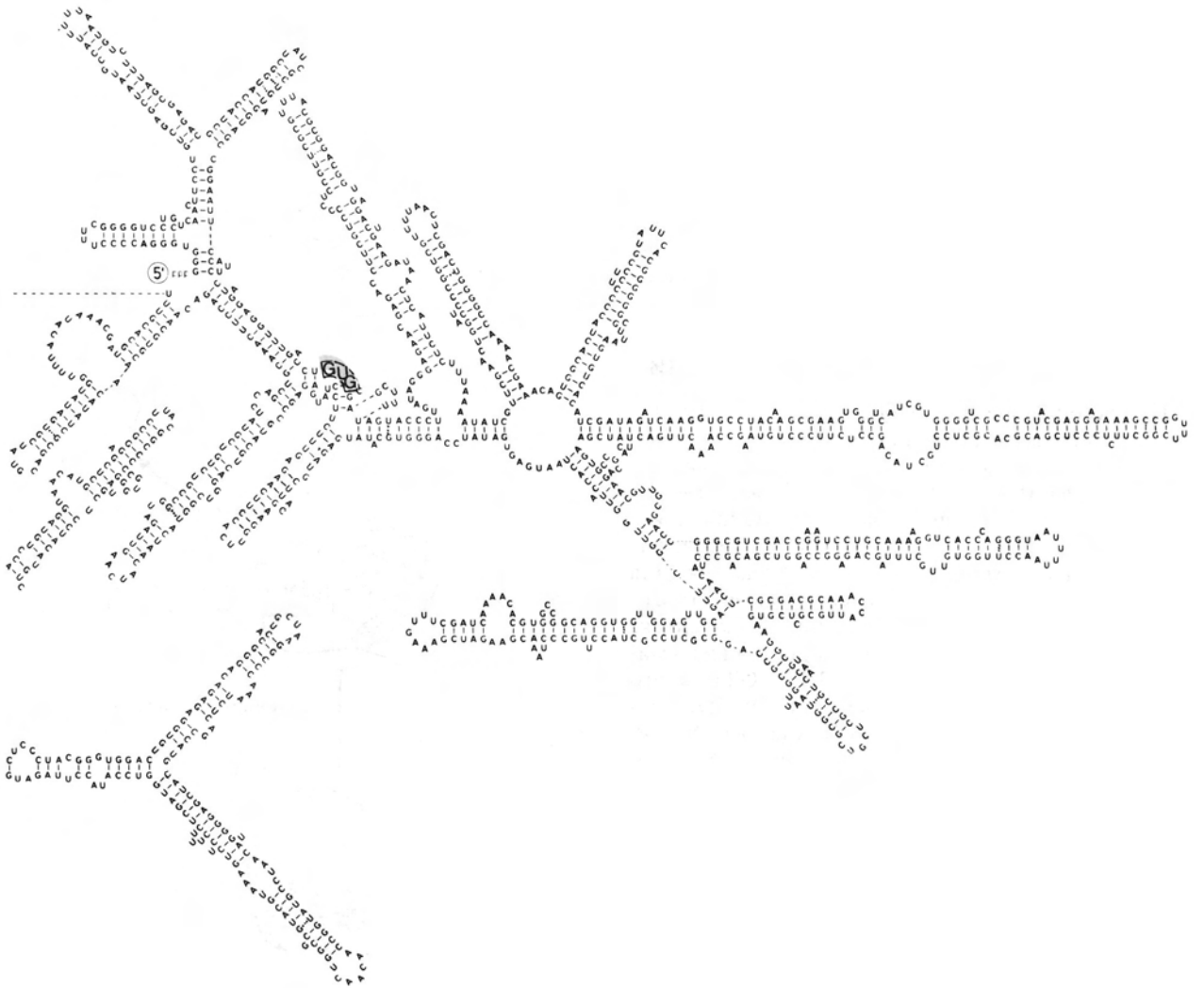
Богатая пуринами последовательность Шайна-Дальгарно, расположенная до иницирующего кодона AUG, спаривается с участком, находящимся вблизи 3'-конца

16S-pPHK *E. coli*. В нижней части рисунка приведены типичные последовательности Шайна-Дальгарно в мРНК некоторых бактериофагов.

**РИС. 3.41.**

Первичная и предполагаемая вторичная структуры 5'-концевой части генома бактериофага MS2. Несколько иницирующих кодонов (в том числе необычный кодон GUG вблизи 5'-конца) оказываются экранированными при внутримолекулярном спаривании оснований. Только кодон AUG (показан стрелкой), находящийся на вершине шпильки, остается доступным и связывается с рибосомами сразу после инфекции. Участок последовательности, начинающийся после кодона AUG, кодирует белок оболочки MS2. Заметим, что последовательность Шайна-Дальгарно, находящаяся перед кодоном AUG в этой модели, также относительно доступна; два остатка G образуют слабые водородные связи с остатками U. (С любезного разрешения W. Fiers.)





ACC AUC A A A G U C G G G U G C C U A A  
U G G U G G U U U C A G C C C A C G G U A

C C C C A U G G C G U V C U A C U U A A A U U G G A A C U A  
G C C G U C C U A A G C A U G C U U U A A C C U U A C C

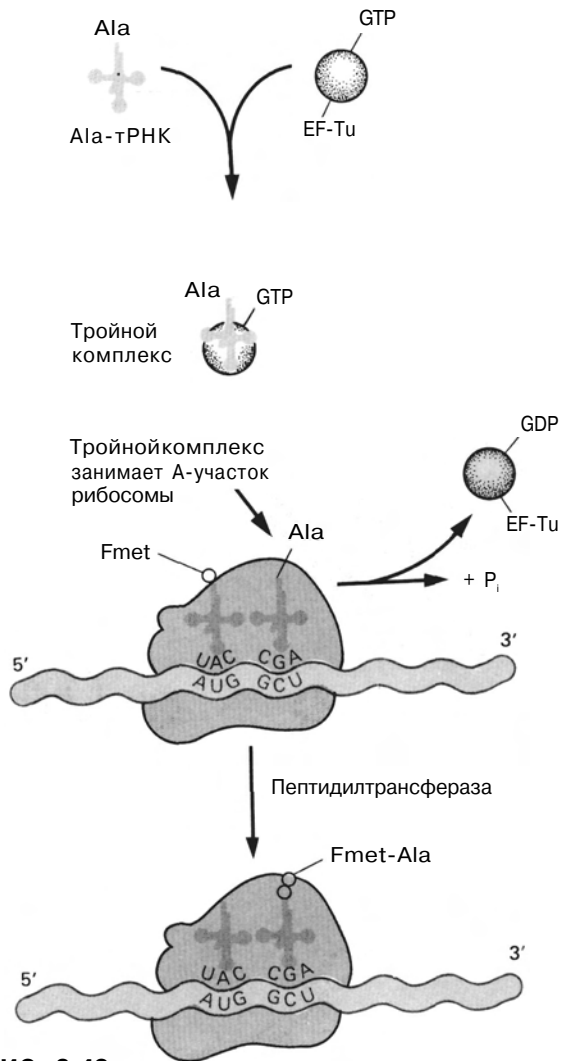
структуры того участка молекулы мРНК, в котором находится инициаторный кодон AUG. Если этот кодон окажется внутри двухцепочечного участка, то инициация будет неэффективна или вовсе блокируется (рис. 3.41). Именно таким образом может регулироваться доступность инициаторного AUG-кодона для 30S-рибосомы. AUG становится недоступным, если он оказывается спаренным при образовании конденсированной формы зрелой мРНК, и, напротив, доступным для инициации во время транскрипции мРНК или во время трансляции других кодирующих последовательностей на той же молекуле мРНК.

## б. Элонгация полипептидной цепи

При ассоциации двух рибосомных субчастиц перед инициацией трансляции образуются два функциональных участка, необходимых для сборки белка: Р- и А-участки. Fmet-ТРНК<sub>Met</sub> занимает Р-участок, а для образования первой пептидной связи необходимо, чтобы аминоацил-тРНК, соответствующая следующему кодону, заняла А-участок. Для этого аминоацил-тРНК должна сначала связать EF-Tu и GTP. Образовавшийся тройной комплекс (аминоацил-тРНК-[EF-Tu-GTP]) и доставляет аминоксил-тРНК к А-участку (рис. 3.42). GTP в это время гидролизуется, и комплекс (EF-Tu-GDP) отделяется от рибосомы. Когда оба участка, А и Р, заняты, пептидилтрансферазная активность 50S-субчастицы катализирует перенос группы Fmet с ее тРНК на аминокислотную группу аминоацил-тРНК, находящейся в А-участке. В результате в А-участке оказывается дипептидил-тРНК, а в Р-свободная тРНК.

Для прочтения следующего кодона и удлинения полипептидной цепи еще на одну аминокислоту вся серия реакций должна повториться. Однако, прежде чем это произойдет, тРНК должна освободить Р-участок, образовавшаяся дипептидил-тРНК должна переместиться на него, а новый кодон должен быть готов к тому, чтобы занять освободившийся А-участок (рис. 3.43). Все эти процессы осуществляются с помощью EF-G при GTP-зависимой транслокации рибосомы. Источником энергии для перемещения рибосомы к следующему триплету кодирующей последовательности и удаления свободной тРНК из Р-сайта служит реакция гидролиза GTP до GDP. Теперь новый кодон, занявший А-сайт, готов к спариванию с родственной аминоацил-тРНК.

Сразу после связывания аминоацил-тРНК с А-участком высвобождается комплекс EF-Tu-GDP и происходит регенерация функционально активного EF-Tu-GTP (рис. 3.44). При этом EF-Tu-GDP взаимодействует с белком EF-Ts, что приводит к отделению GDP и образованию комплекса EF-Tu•EF-Ts. Далее EF-Tu•EF-Ts взаимодействует с GTP, проис-

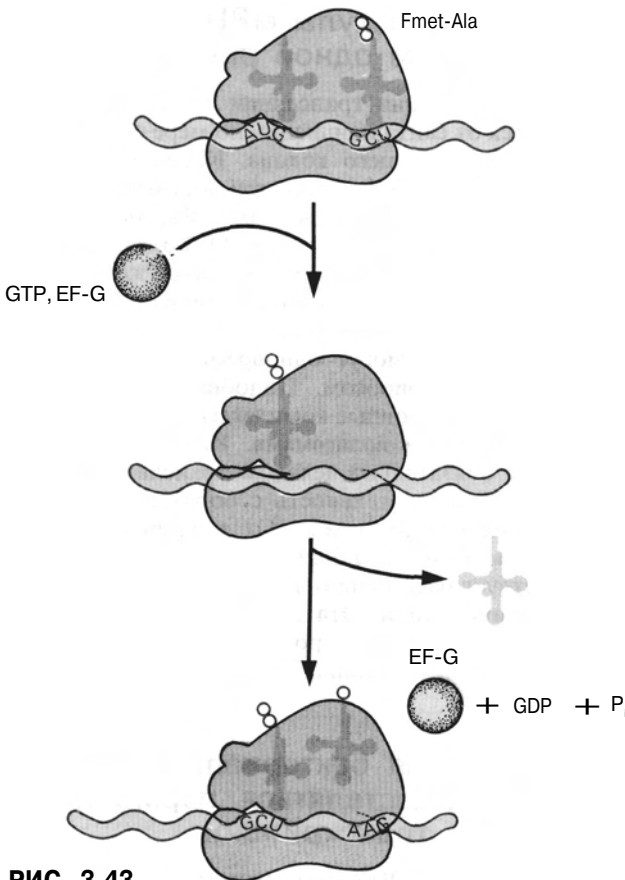


**РИС. 3.42.**

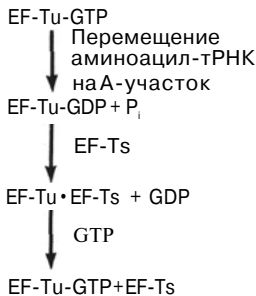
Образование первой пептидной связи. Преиницирующий комплекс 70S-рибосомы с мРНК и N-формилметионил-тРНК<sub>Met</sub>, находящийся в Р-участке, связывает тройной комплекс аланил-тРНК<sup>Ala</sup>-EF-Tu-GTP. Аланил-тРНК<sup>Ala</sup> перемещается к А-участку и связывается со вторым кодоном, GCU, с помощью антикодона 5'-AGC-3'. GTP гидролизуется, и комплекс EF-Tu-GDP отделяется. В результате пептидилтрансферазной реакции дипептидил-тРНК занимает А-участок.

ходит регенерация EF-Tu-GTP и отделение EF-Ts, и оба соединения оказываются готовыми к следующему циклу.

Необходимо отметить несколько особенностей процесса элонгации. 1. При образовании каждой пептидной связи расходуется энергия, равная четырем энергетическим эквивалентам (если за один эквивалент принять энергию образования фосфат-

**РИС. 3.43.**

Элонгация полипептидной цепи. Фактор EF-G осуществляет GTP-зависимое перемещение 70S-рибосомы вдоль мРНК. Далее GTP гидролизуется, тРНК удаляется из Р-участка и пептидил-тРНК перемещается из А-участка на освободившийся Р-участок. Теперь А-участок и следующий кодон готовы принять «свою» аминоксил-тРНК.

**РИС. 3.44.**

Регенерация EF-Tu-GTP после высвобождения EF-Tu-GDP из комплекса с аминоксил-тРНК.

ной связи): два эквивалента АТФ потребляются при аминокислотировании тРНК и два эквивалента GTP-в каждом цикле элонгации. 2. При инициации трансляции IF-2 узнает Fmet-тРНК<sub>F<sup>Met</sup></sub> среди всех других аминоксил-тРНК, а EF-Tu отличает met-тРНК<sub>F<sup>Met</sup></sub> от Fmet-тРНК<sub>M<sup>Met</sup></sub> при внедрении в А-участок. 3. Факторы элонгации EF-Tu и EF-G то присоединяются, то отделяются от рибосомы в зависимости от того, связаны ли они с GTP или с GDP соответственно. 4. Растущая полипептидная цепь всегда соединена своим карбоксильным концом с тРНК, которая соответствует С-концевой аминокислоте в растущей полипептидной цепи. 5. Пептидилтрансфераза катализирует формирование пептидных связей между карбоксильным концом растущей цепи и аминокислотной группой аминоксил-тРНК.

### в. Терминация элонгации полипептидной цепи

Процесс последовательной трансляции кодонов в конце концов доходит до того момента, когда в А-участке оказывается один из трех терминирующих кодонов—UAG, UAA или UGA (рис. 3.45). Из-за отсутствия тРНК, отвечающих этим кодоном, пептидил-тРНК остается связанной с Р-участком. Здесь вступают в действие специфические факторы RF-1 и RF-2, которые катализируют отсоединение полипептидной цепи от тРНК, отделение их обоих от рибосомы, а 70S-рибосомы — от мРНК. RF-1 узнает в А-участке кодон UAA или UAG; RF-2 включается в том случае, когда в А-участке оказывается UAA или UGA; RF-3 облегчает работу двух других факторов. Если терминирующим кодоном является UAA, то эффективность процесса терминации оказывается наибольшей, поскольку этот кодон узнают оба фактора—RF-1 и RF-2. Однако, каким бы из стоп-кодонов ни обеспечивалась терминация, ее эффективность зависит от фланкирующих эти кодоны последовательностей в мРНК. Хотя общие черты и даже некоторые детали процесса терминации известны, точный его механизм и каталитическая роль факторов RF-1 и RF-2 пока неясны.

### 3.7. НЕКОТОРЫЕ ОБЩИЕ ОСОБЕННОСТИ ПРОЦЕССА ТРАНСЛЯЦИИ

В предыдущем разделе мы подробно рассмотрели события одного цикла трансляции мРНК на рибосоме. Здесь акцент будет сделан на некоторых общих принципах процесса в целом и на нарушении или блокировании его отдельных стадий.



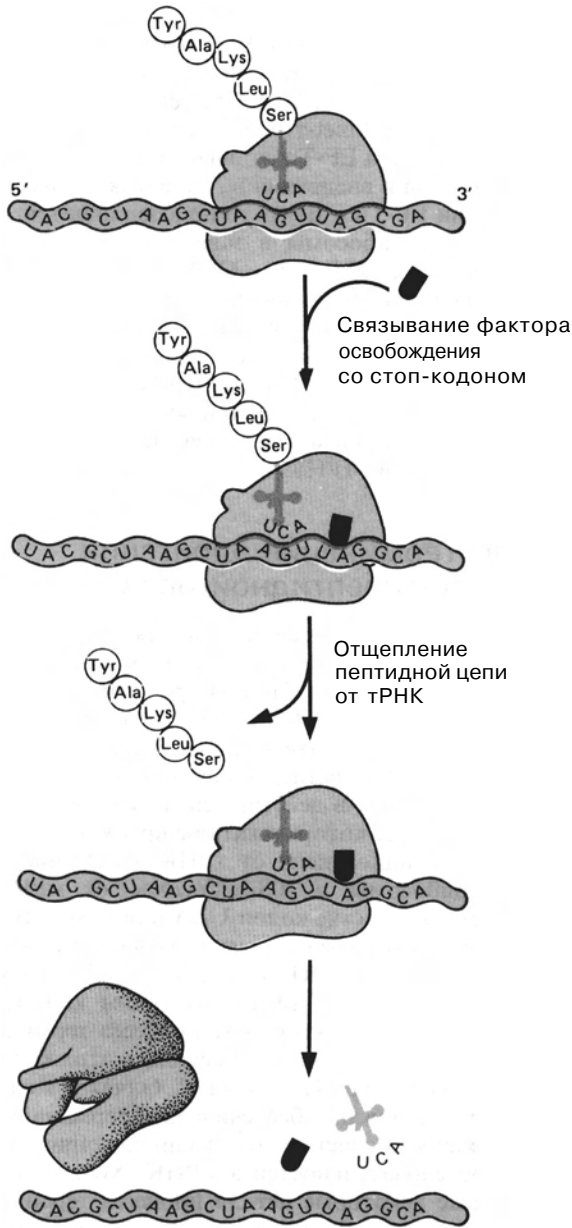


РИС. 3.45.

Терминация трансляции. Когда в А-участке оказывается один из стоп-кодонов, факторы освобождения инициируют серию реакций, приводящую к отделению синтезированной полипептидной цепи, тРНК и мРНК от 70S-рибосомы.

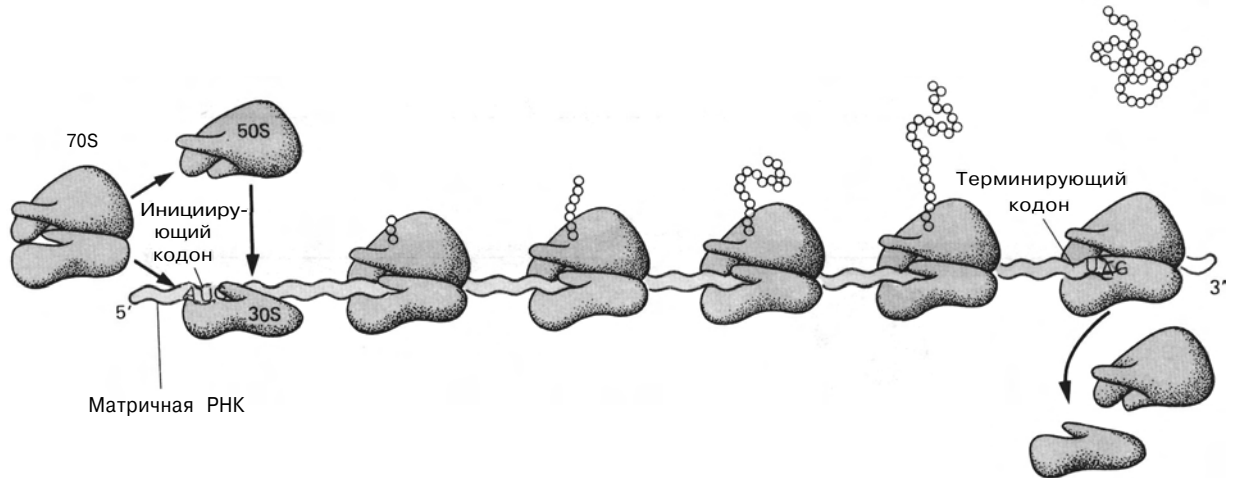
### а. Одновременная трансляция молекулы мРНК более чем одной рибосомой

После инициации трансляции 70S-рибосома перемещается от сайта инициации по мере считывания каждого последующего кодона. Когда расстояние от рибосомы до сайта инициации достигнет величины 100–200 нуклеотидов, в этом сайте может произойти новая инициация (рис. 3.46). Более того, как только вторая рибосома пройдет такое же расстояние, может произойти третья инициация, и т.д. Итак, одну и ту же белок-кодирующую последовательность мРНК могут одновременно транслировать несколько рибосом. Подобные мультирибосомные трансляционные комплексы называются **полирибосомами** или **полисомами**. Каждая рибосома полисомы обязательно целиком транслирует кодирующую последовательность с образованием полноразмерного полипептида. От каждой рибосомы в полисоме отходит полипептид, длина которого соответствует расстоянию, пройденному рибосомой от сайта инициации. Эта длина пропорционально увеличивается по мере продвижения рибосомы по мРНК, начиная с 5'-конца кодирующей последовательности.

### б. Трансляция бактериальных мРНК может осуществляться параллельно транскрипции

Образование мРНК при транскрипции гена или кластера генов начинается с 5'-конца в направлении к 3'-концу (разд. 3.2). Следовательно, формирование комплекса инициации трансляции может произойти сразу же после того, как будет транскрибирована последовательность, в пределах которой находится инициирующий кодон (рис. 3.47). И в самом деле, синтез полипептидной цепи обычно начинается до завершения транскрипции 3'-концевой части мРНК.

Если бактериальная транскрипционная единица содержит более одной кодирующей белок последовательности, как, например, в случае *trp*- или *lac*-оперонов (разд. 3.11), то на рибосомах может начаться и даже завершиться трансляция первой из этих кодирующих последовательностей еще до окончания транскрипции остальных. Матричные РНК, состоящие из многих белок-кодирующих участков, часто транслируются последовательно, т.е. инициация, элонгация и терминация трансляции первой кодирующей последовательности сопровождаются такими же событиями на втором, третьем и последующих кодирующих сегментах. После того как рибосомы доходят до сигнала терминации в первой или любой другой кодирующей последовательности, они отделяются от мРНК, и со следую-

**РИС. 3.46.**

Образование полирибосомы. После того как первая рибосома удалится от сайта инициации на определенное расстояние, вторая образует инициаторный комплекс в том же сайте и начинает трансляцию. То же самое

затем проделывает третья рибосома и т.д. В результате на одной и той же мРНК могут одновременно синтезироваться сразу несколько полипептидов.

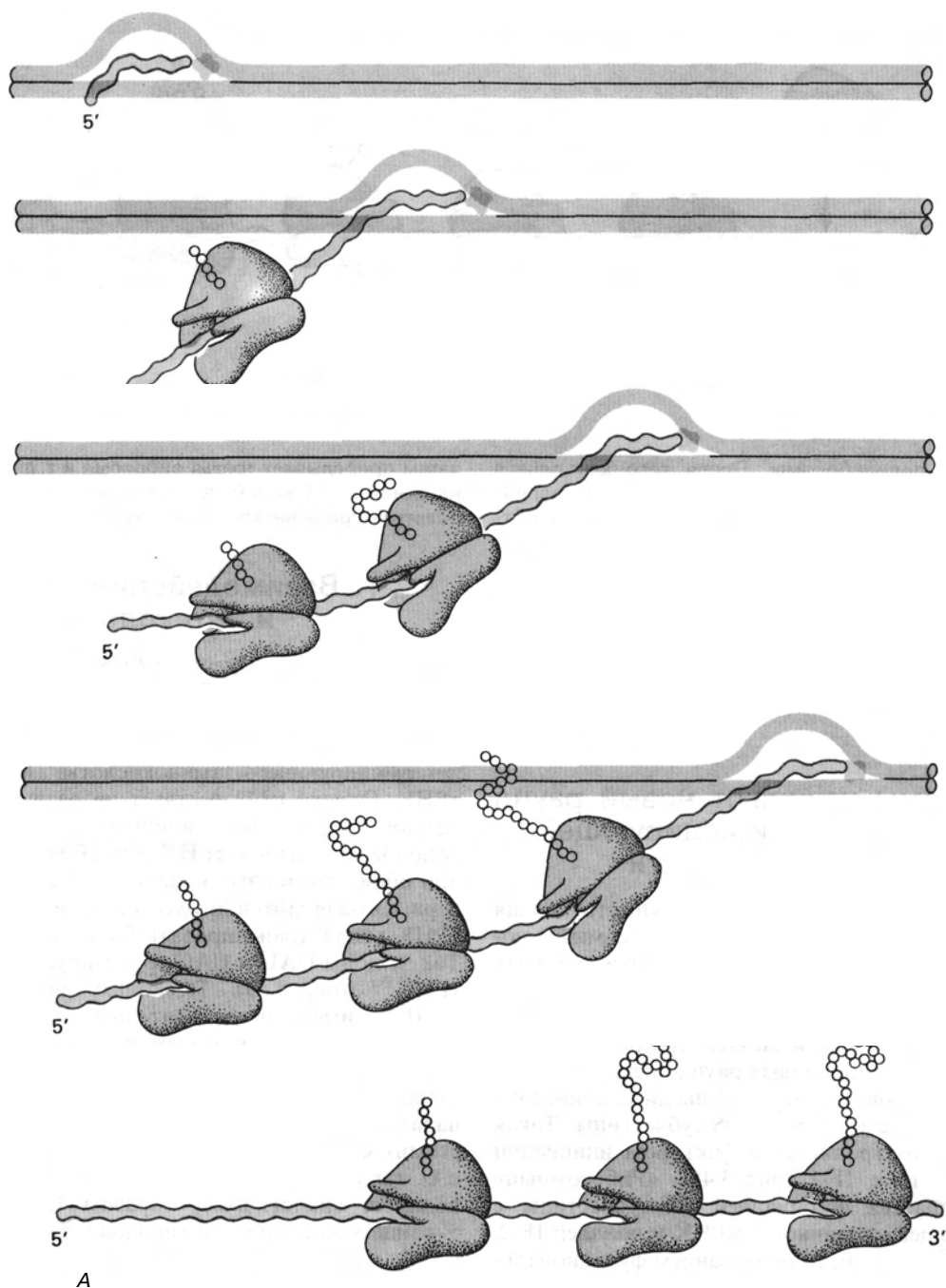
щим инициаторным участком связывается новый комплекс. Однако в некоторых случаях рибосомы не отделяются от мРНК, а перемещаются вдоль молекулы, образуя новые комплексы инициации в других сайтах.

#### в. Рибосомы начинают новый раунд после трансляции кодирующей последовательности

Как мы уже говорили, когда осуществляющая трансляцию 70S-рибосома доходит до терминирующего кодона, полноразмерная полипептидная цепь отделяется от мРНК, а обе они — полипептид и тРНК — отделяются от рибосомы, и происходит разделение 70S-рибосомы и мРНК. 70S-рибосомы не способны к инициации новых раундов синтеза полипептидов и, следовательно, должны диссоциировать на составляющие их 50S- и 30S-субчастицы. Такая диссоциация контролируется фактором инициации IF-3 совместно с IF-1 (рис. 3.48). «Рибосомный» цикл завершается ассоциацией 50S-субчастицы с 30S-субчастицей, связанной с мРНК и несущей IF-2, Fmet-тРНК<sub>F<sup>Met</sup></sub> и GTP, с образованием функционального трансляционного 70S-аппарата (рис. 3.39). Таким образом, изменяя количество 30S- и 50S-субчастиц по отношению к их предшественнику — 70S-рибосоме, — фактор IF-3 осуществляет общий контроль уровня белкового синтеза.

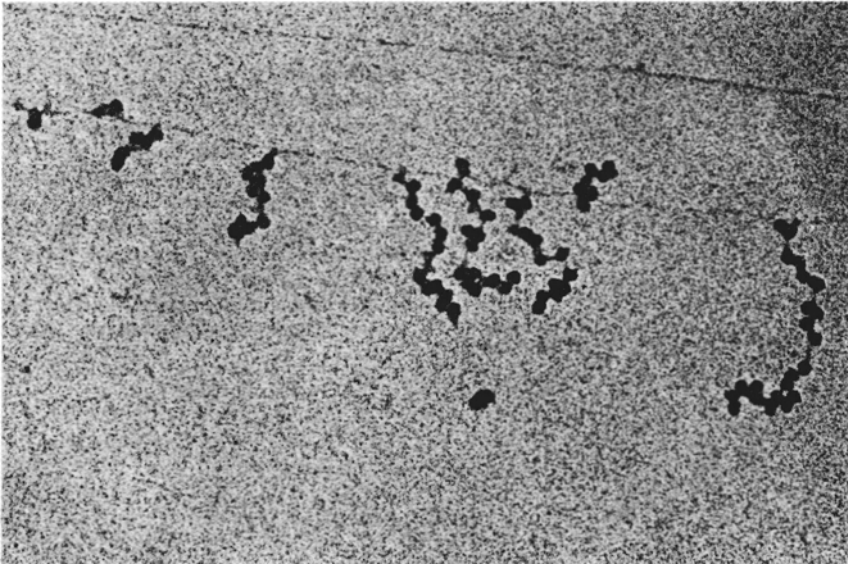
#### г. Взаимодействие кодона и антикодона

Большинство молекул тРНК спаривается более чем с одним кодоном. Поскольку кодоны транслируются при участии антикодонов разных тРНК, можно было бы ожидать, что для каждого из 61 кодона, детерминирующего аминокислоты, имеется своя тРНК. Однако не существует ни разных тРНК для каждого из четырех валиновых или глициновых кодонов, ни разных тРНК для обоих тирозиновых или обоих лизиновых кодонов. Действительно, эксперименты *in vitro* и *in vivo* показали, что некоторые тРНК могут транслировать более чем один кодон. Так, кодоны UAU и UAC транслирует единственная тРНК<sup>Tyr</sup> (рис. 3.49). Поскольку антикодон этой тРНК<sup>Tyr</sup> имеет последовательность 5'-GUA-3', он может образовать комплементарные пары с первыми двумя основаниями любого из этих кодонов (напомним, что спаривание происходит между антипараллельно ориентированными последовательностями). Соответственно G способен спариваться как с U, так и с C, находящимися в третьем положении кодонов; аналогичным образом U, находясь на 5'-конце какого-либо антикодона, может спариться и с A, и с G, находящимися на 3'-конце соответствующих кодонов. На самом деле трансляция всех пар кодонов, у которых в третьей (3') позиции стоит U или C, может осуществляться одной и той же тРНК, у которой первым основанием в антикодоне оказывается G или какое-то модифицированное основание. По-видимому, при спаривании кодонов и антикодонов в A- и P-участках включаются какие-то стабилизирующие взаимодействия, отличные от

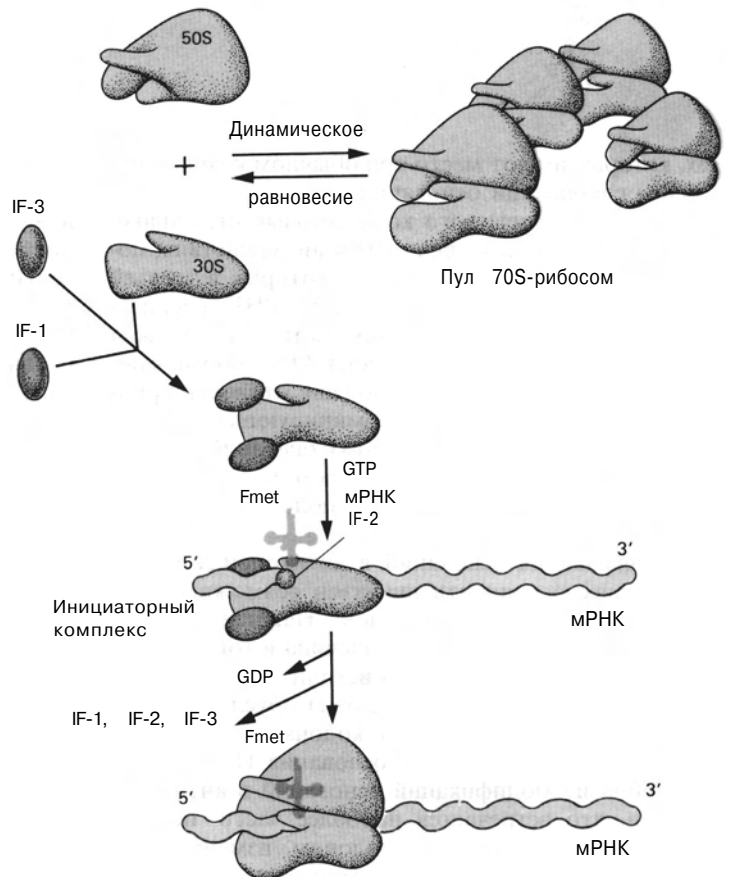
**РИС. 3.47.**

Транскрипция и трансляция у прокариот идут параллельно. А. Схематическое представление транскрипции ДНК и параллельной трансляции мРНК полирибосомами. Б. Электронная микрофотография, на которой видны рибосомы, прикрепившиеся к молекулам РНК, кото-

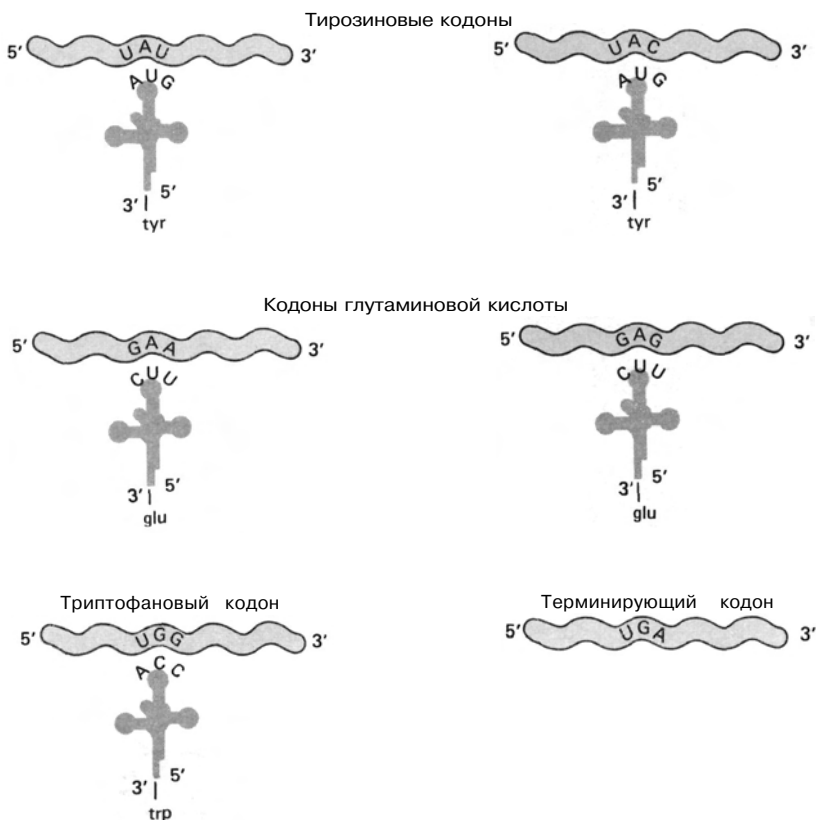
рые транскрибируются с ДНК *E. coli*. К более длинным цепям РНК, находящимся на большем расстоянии от сайта инициации транскрипции, прикреплено и больше рибосом (142000 x). [O.L. Miller, Jr., В.А. Hamkalo, С.А. Thomas, Jr., *Science*, **169** (1970), p. 392.]



Б



**РИС. 3.48.**  
«Рабочий цикл» 70S-рибосом при синтезе белка.

**РИС. 3.49.**

Большинство молекул тРНК спаривается более чем с одним кодоном. Тирозиновые кодоны UAU и UAC транслируются с помощью одной тРНК<sup>Tyr</sup>. Аналогично кодоны GAA и GAG для глутаминовой кислоты транслируются с помощью одной тРНК<sup>Glu</sup>. В то же время тРНК<sup>Trp</sup> узнает триптофановый кодон UGG, но не узнает терминирующий кодон UGA.

тех, которые имеют место при обычном комплементарном спаривании оснований.

Анализ генетического кода показывает, однако, что существуют специфические взаимодействия, позволяющие различать кодоны, у которых в третьей позиции стоит А или G. Например, тРНК, расшифровывающая кодон AUG как метионин, должна отличать этот триплет от кодона AUA, обозначающего изолейцин, а тРНК<sup>Trp</sup> должна отличать триптофановый кодон UGG от терминирующего кодона UGA. Специфичность обеих этих операций декодирования определяется спариванием С антикодона с G, находящимся в третьем положении кодона (рис. 3.49).

Модификация оснований в антикодонах может еще сильнее ограничить диапазон возможных взаимодействий кодон-антикодон. Например, гипоксантин (Hx), занимая место аденина в той позиции антикодона, по которой происходит спаривание с третьим основанием кодона, может обусловить спаривание такого антикодона с кодонами, у которых в последней позиции стоят основания U, C или A. Разнообразие модификаций оснований в антикодонах или — что встречается наиболее часто — оснований, соседствующих с антикодоном, изменяет специфичность взаимодействия аминоксил-тРНК—ко-

дон. Таким способом обычно предотвращаются ошибки при считывании третьего основания кодонов и обеспечивается надежность процесса декодирования.

Правила спаривания оснований, согласно которым молекулы тРНК одного типа могут узнавать несколько разных кодонов, называются правилами неоднозначного соответствия (или «качания» от англ. "wobble" rules) (табл. 3.4). Следует отметить, однако, что термин «качание», используемый для описания некоторой свободы спаривания третьего

Таблица 3.4. Пары оснований, которые могут образовываться согласно теории неоднозначного соответствия

Основание антикодона, спаривающееся с третьим основанием кодона <sup>1)</sup>	Третье основание кодона
G	U или C
C	G
A	U
U	A или G
Hx	A, U или C

<sup>1)</sup> U, C, A, G - четыре обычных основания РНК; Hx - дезаминированная форма аденина - гипоксантин.

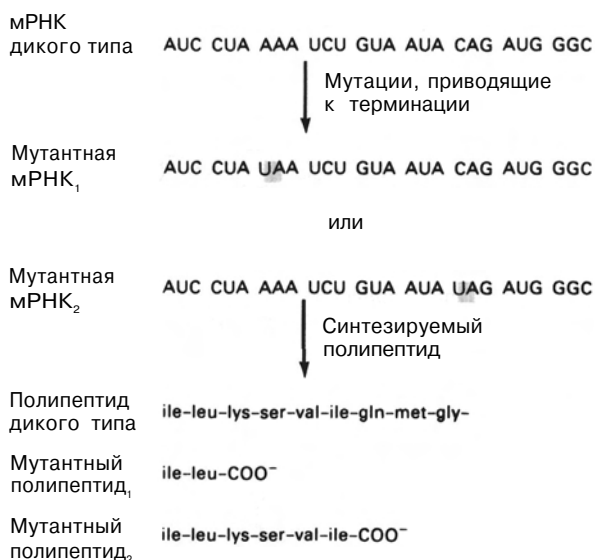
основания кодона, просто как бы затушевывает тот факт, что мы до конца не знаем, какие именно химические и структурные особенности обуславливают кодон-антикодоновые взаимодействия в Р- и А-участках рибосомы.

**Мутации в кодонах и антикодонах.** Мутации, затрагивающие различные компоненты трансляционного аппарата, могут изменить результат считывания кодирующей последовательности. Наиболее драматичные последствия вызывают те мутации в гене, кодирующем белок, которые превращают кодон, отвечающий какой-то аминокислоте, в терминирующий кодон и тем самым приводят к преждевременному завершению синтеза из-за досрочной терминации трансляции в мутировавшем сайте (рис. 3.50). Примером может служить превращение лизинового кодона AAA в UAA и глутаминового кодона CAG в UAG. Аналогично любая мутация, в результате которой происходит замена аминокислотного кодона на кодон UGA, тоже вызовет преждевременную остановку синтеза полипептидной цепи. Однако, если в результате второй мутации произойдет изменение соответствующего основания в антикодоне тРНК, терминация может быть предотвращена, или **супрессирована**, и образуется полно-размерный, хотя и измененный, белок (рис. 3.51). Например, если тРНК<sup>Tr</sup>, тРНК<sup>Leu</sup> или тРНК<sup>Ser</sup> изменятся подобным образом, то они смогут прочесть кодон UAG как аминокислотный. С помощью различных механизмов может произойти ошибочная трансляция и таких мутантных кодонов, как UAA и UGA. Мутации в тРНК-генах, затрагиваю-

щие основания, отличные от тех, которые составляют антикодон, могут привести к изменению специфичности или стабильности взаимодействий кодона и антикодона. Благодаря таким механизмам может быть предотвращена преждевременная терминация синтеза полипептида, если терминирующий кодон будет прочитан как смысловой. Подобная супрессия терминации, как правило, не очень эффективна, поэтому наряду с полноразмерными образуются и укороченные, преждевременно терминированные полипептидные цепи. Благодаря относительной неэффективности такой трансляционной супрессии не приносит большого вреда и случайные проскакивание терминирующих кодонов, находящихся на естественных концах кодирующих мРНК.

Миссенс-мутации, т.е. мутации, приводящие к аминокислотным заменам и соответственно к утрате белком его функции, также могут быть ревертированы благодаря супрессорным мутациям, вызывающим ошибочное считывание мутантного кодона (рис. 3.52). Это может произойти в том случае, если тРНК, несущая нужную аминокислоту или любую другую, которая может быть включена в данный сайт белковой цепи, имеет антикодон, способный к спариванию с мутантным кодоном. Мутации, вызывающие сдвиг рамки считывания кодирующей последовательности, также могут быть супрессированы, если мутантные тРНК или рибосомы случайно транслируют два или четыре основания вместо трех.

Итак, ошибки трансляции могут компенсировать последствия нарушений кодирующей последовательности. Мутационные изменения в антикодоне тРНК—это наиболее распространенный механизм супрессии; изменения в других участках молекулы тРНК могут привести к неправильной этерификации аминокислот аминокил-тРНК-синтетазами или ошибочному спариванию на рибосоме. Ошибки в трансляции могут возникать и в том случае, если в результате мутаций происходит изменение белков или РНК-компонент рибосом, участвующих в кодон-антикодоновом взаимодействии. Точность трансляции уменьшается и под действием некоторых химических соединений (например, стрептомицина), которые связываются с рибосомными белками в 30S-субчастице. Такие случаи нарушения процесса трансляции приводят к более тяжелым последствиям.

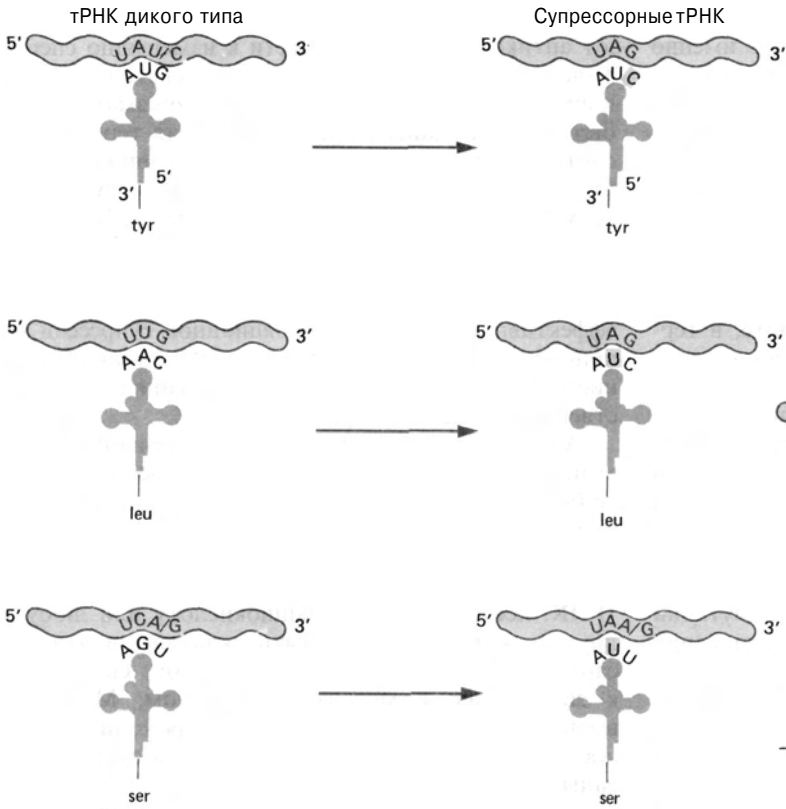


**РИС. 3.50.**

Мутации, приводящие к превращению внутренних кодонов в терминирующие (стоп) кодоны, обуславливают преждевременную терминацию трансляции.

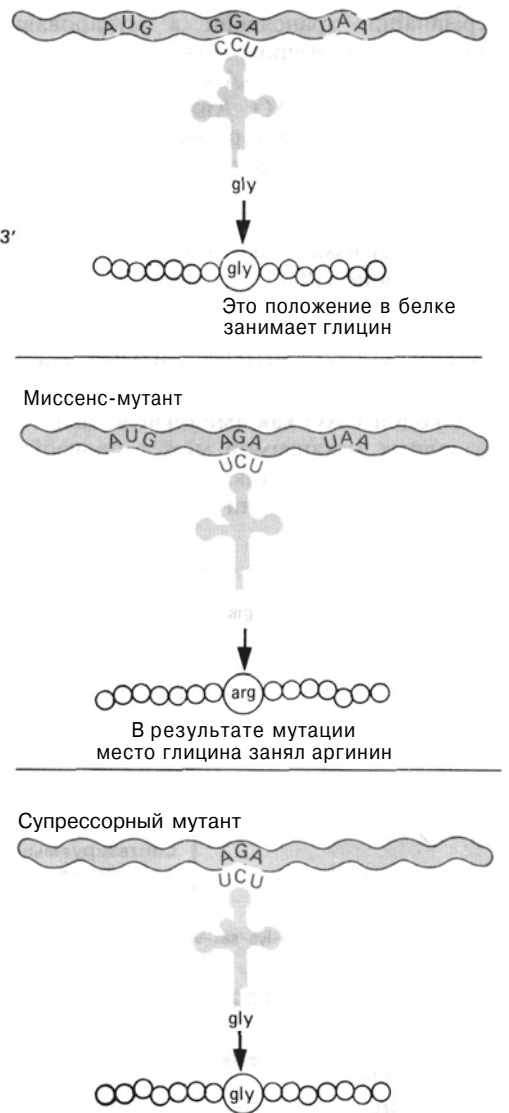
## 3.8. ТРАНСЛЯЦИЯ мРНК У ЭУКАРИОТ

Процесс трансляции эукариотической мРНК в основном аналогичен таковому прокариотической мРНК. За некоторыми отмеченными выше исключениями, генетический код универсален и кодоны по-



**РИС. 3.52.** Трансляционная супрессия миссенс-мутаций. В тРНК<sup>Gly</sup> произошло такое изменение антикодона, что аргининовый кодон был прочитан как глициновый.

**РИС. 3.51.** Трансляционная супрессия терминирующих кодонов. Мутации, вызывающие изменения в антикодонах тРНК<sup>Tyr</sup>, тРНК<sup>Leu</sup> или тРНК<sup>Ser</sup>, приводят к тому, что терминирующие кодоны считываются как кодоны, отвечающие определенным аминокислотам.



Супрессорная тРНК<sup>Gly</sup> с измененным антикодоном узнает аргининовый кодон

следовательно транслируются с помощью специфических аминоктил-тРНК-синтетаз на рибосомах. Есть, однако, и три явных различия, обусловленных определенными свойствами эукариотических клеток. Во-первых, аппараты транскрипции и трансляции у эукариот физически разобщены, поскольку транскрипция осуществляется в ядре, а трансляция — в цитоплазме. Во-вторых, на 5'- и 3'-концах эукариотических мРНК имеются особые структуры. И в-третьих, эукариотические мРНК, за исключением мРНК, транскрибируемых с ДНК геномов вирусов, обычно содержат только одну белок-кодирующую последовательность.

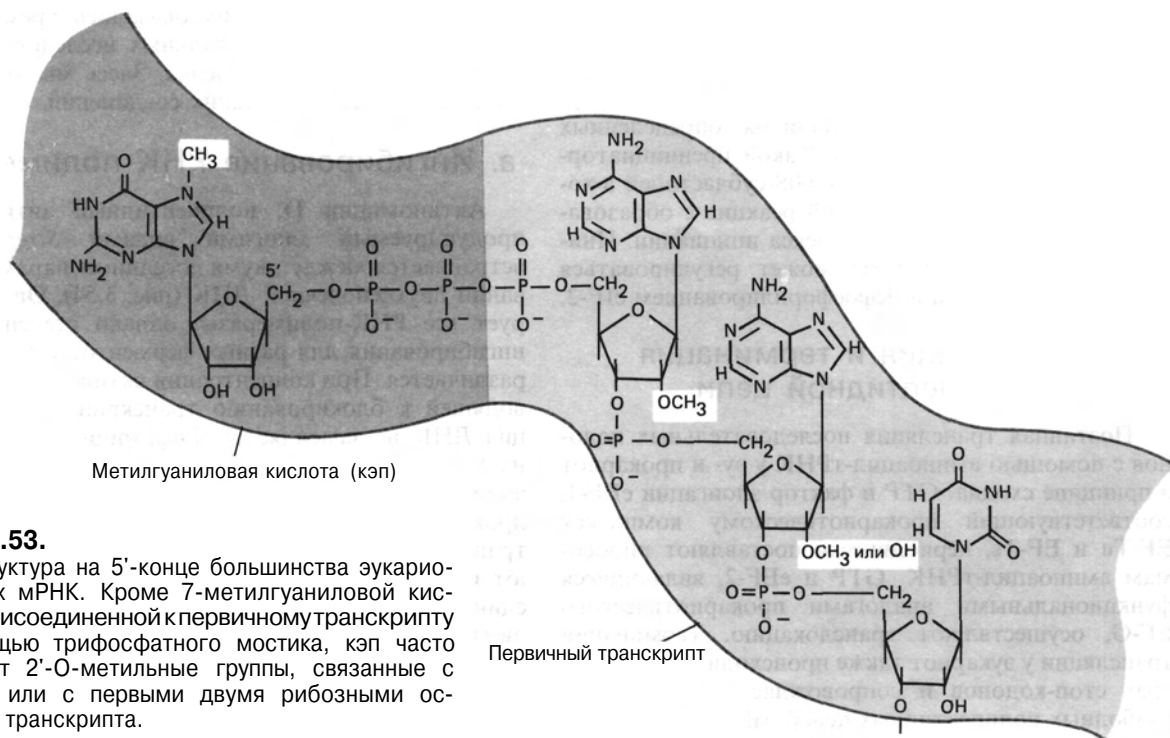
Структура и свойства участников трансляции эукариотической мРНК пока изучены гораздо хуже, чем у прокариот. И хотя у эукариот выделяют те же три стадии процесса — инициацию, элонгацию и терминацию, — на каждой из них требуется больше нерибосомных белковых факторов. Несмотря на эти различия, последовательности, кодирующие белки прокариот, нормально транслируются эукариотическими системами трансляции при условии соответствующей модификации их мРНК на 3'- и 5'-концах (рис. 3.8, А). И наоборот, кодирующие последовательности эукариот эффективно транслируются системами прокариот, если у них перед 5'-концом инициаторного кодона AUG имеется последовательность Шайна-Дальгарно. Это значит, что трансляционные аппараты обоих типов организмов

могут осуществлять свои функции, несмотря на особенности нуклеотидных последовательностей мРНК из разных источников.

### а. Особые модификации мРНК эукариот

У эукариотических мРНК, транскрибированных с ядерных или вирусных геномов РНК-полимеразой II, всегда модифицированы 5'-концы, которые в этом случае называют «кэпами» (рис. 3.53). РНК, транскрибируемые эукариотическими РНК-полимеразой I (в том числе рРНК) и III (5S- и тРНК), не кэпированы и имеют обычные 5'-фосфатные концы. У большинства мРНК, синтезируемых РНК-содержащими вирусами животных, также имеются кэпы, хотя они синтезируются вирусными РНК-транскриптазами. Многие некэпированные мРНК неэффективно транслируются эукариотическими белоксинтезирующими системами из-за слабого связывания рибосом с мРНК. Кэпирование происходит на 5'-нуклеозидтрифосфате вскоре после инициации синтеза РНК-транскриптов и задолго до его завершения. Детали процесса кэпирования изложены в разд. 8.3.в.

На 3'-концах эукариотических мРНК имеются еще и полиаденилатные последовательности. Такой 3'-«хвост» из 50–200 аденилатных остатков не кодируется смысловыми последовательностями соответ-



**РИС. 3.53.**

Кэп-структура на 5'-конце большинства эукариотических мРНК. Кроме 7-метилгуаниловой кислоты, присоединенной к первичному транскрипту с помощью трифосфатного мостика, кэп часто включает 2'-О-метильные группы, связанные с первым или с первыми двумя рибозными остатками транскрипта.



ствующих генов, а присоединяется посттранскрипционно, после разрезания транскрипта в специфическом месте за сигналом терминации трансляции (разд. 8.3.в).

### б. Инициация трансляции на 5'-кэпированных концах малыми рибосомными субчастицами

Как мы уже говорили, обязательным этапом при инициации трансляции прокариотической мРНК является диссоциация 70S-рибосомы; точно так же и 80S-рибосомы должны диссоциировать до начала трансляции эукариотических мРНК. Малая субчастица (40S) в комплексе со множеством белков-помощников (eIF), из которых один или несколько нужны для диссоциации рибосомы на составляющие субчастицы, связывает особую инициаторную мет-тРНК<sub>1</sub><sup>Met</sup>. И вновь для связывания инициаторной аминоксил-тРНК необходимы GTP и особый белок—eIF-2. У эукариот, однако, мет-тРНК<sub>1</sub><sup>Met</sup> не подвергается N-формилированию; но, как и у прокариот, структура тРНК<sub>1</sub><sup>Met</sup> отличается от структуры тРНК<sub>M</sub><sup>Met</sup>. Комплекс 40S, содержащий мет-тРНК<sub>1</sub><sup>Met</sup>, GTP и целую армию других факторов eIF, связывается с мРНК вблизи кэпированного 5'-конца или прямо с этим концом; по крайней мере один из этих факторов узнает кэп и связывается с ним. Малая субчастица перемещается при помощи неизвестного пока механизма от кэпированного конца до первого кодона AUG. Никакой последовательности вблизи кодона AUG, аналогичной последовательности Шайна-Дальгарно, пока не обнаружено. Однако эффективность работы AUG в качестве инициаторного кодона зависит от наличия определенных фланкирующих нуклеотидов. Такой преинициаторный комплекс объединяется с 60S-субчастицей в ходе энерго- и фактор-зависимой реакции с образованием функционального комплекса инициации. Инициация белкового синтеза может регулироваться фосфорилированием и дефосфорилированием eIF-2.

### в. Элонгация и терминация полипептидной цепи

Поэтапная трансляция последовательных кодонов с помощью аминоксил-тРНК у эу- и прокариот в принципе сходна. GTP и фактор элонгации eEF-1, соответствующий прокариотическому комплексу EF-Tu и EF-Ts, периодически поставляют рибосомам аминоксил-тРНК. GTP и eEF-2, являющиеся функциональными аналогами прокариотического EF-G, осуществляют транслокацию. Терминация трансляции у эукариот также происходит в одном из трех стоп-кодонов и сопровождается отделением свободных полипептидных цепей, мРНК и, возмож-

но, 80S-рибосом от мРНК. По-видимому, всю серию терминационных событий осуществляют один фактор eRF и GTP.

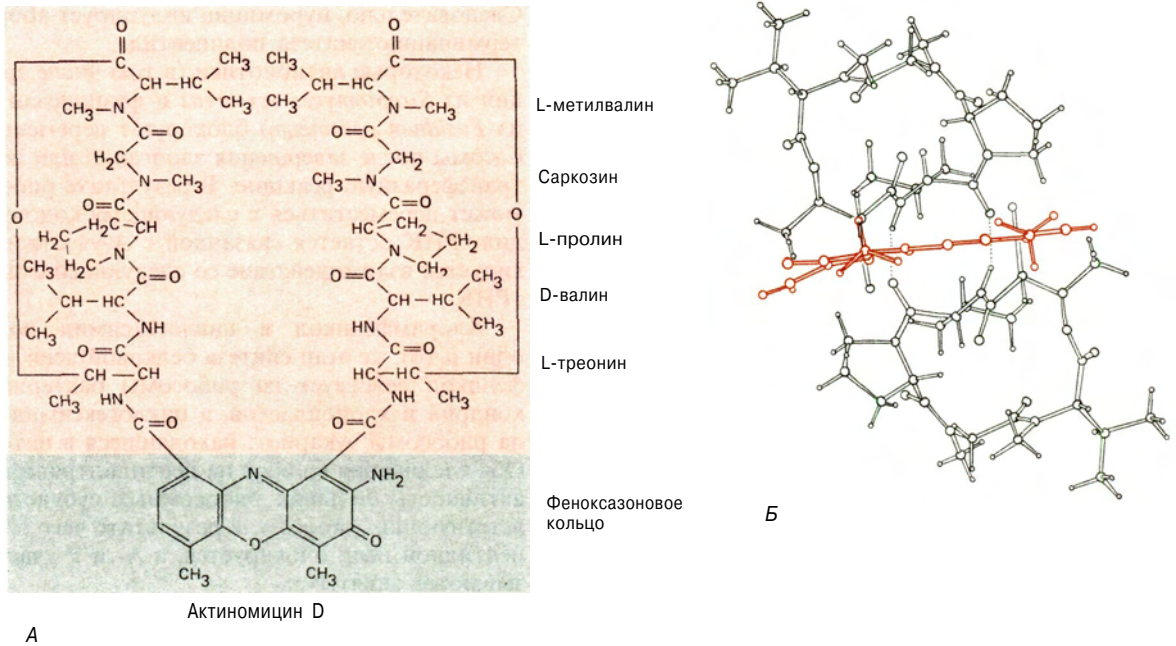
Биогенез эукариотической мРНК детально обсуждается в разд. 8.3. Здесь мы лишь отметим, что трансляция не начинается до тех пор, пока зрелая мРНК не достигнет цитоплазмы, где и происходит инициация. В результате последовательных актов инициации образуются полисомы и одновременно, как и у прокариот, начинается трансляция множества кодирующих белки последовательностей. Заметное отличие трансляции у эукариот состоит в том, что клеточные мРНК обычно содержат одну-единственную кодирующую последовательность. Если в мРНК имеется несколько таких областей, то последующие либо вовсе не транслируются, либо транслируются неэффективно.

## 3.9. ИНГИБИТОРЫ ТРАНСКРИПЦИИ И ТРАНСЛЯЦИИ

Было исследовано огромное количество природных и синтетических соединений на их способность оказывать противомикробное или противоопухолевое действие. Многие из них обладали высокой ингибирующей активностью в отношении как транскрипции, так и трансляции у про- и эукариот. Действие большинства таких соединений неспецифично, и они слишком токсичны для того, чтобы их можно было использовать в терапевтических целях. Однако некоторые из них оказались чрезвычайно полезны для экспериментальных исследований аппаратов синтеза РНК и белка. Здесь мы упомянем лишь о некоторых из таких соединений.

### а. Ингибирование РНК-полимеразы

Актиномицин D, полипептидный антибиотик, продуцируемый многими видами *Streptomyces*, встраивается между двумя соседними парами оснований двухцепочечной ДНК (рис. 3.54). Он ингибирует все РНК-полимеразы, однако степень этого ингибирования для разных ферментов существенно различается. При концентрации актиномицина, приводящей к блокированию транскрипции, репликация ДНК не замедляется. Рифамицин, антибиотик из *Streptomyces mediterranei*, или его синтетический аналог рифампицин блокируют транскрипцию у прокариот, но ни тот ни другой не влияют на транскрипцию у эукариот. Эти соединения блокируют инициацию синтеза РНК, связываясь с β-субъединицами РНК-полимераза-холофермента. Но если синтез РНК уже инициирован, то процесс элонгации оказывается нечувствительным к рифампицину. Другой антибиотик, применяемый для блокирования бактериальных РНК-полимераз,— стрептолиди-

**РИС. 3.54.**

Структура актиномицина D. А. Структурная формула актиномицина D. Б. Трехмерное изображение, в котором плоскость феноксазонового кольца перпендикулярна плоскости страницы. В. Структура, изображенная на рис. Б, интеркалирует (встраивается) между двумя соседними GC-парами дуплексной ДНК. [S.C. Jain, H.M. Sobell, J. Mol. Biol., **68** (1972), P. 1.]

гин из *Streptomyces lydicus*. Он блокирует элонгацию, но не инициацию РНК.

Широко используемый ингибитор транскрипции у эукариот—бициклический октапептид  $\alpha$ -аманитин, получаемый из гриба *Amanita phalloides*. Из трех эукариотических РНК-полимераз (разд. 3.1.а) наиболее чувствительна к  $\alpha$ -аманитину РНК-полимераза II. РНК-полимераза I не ингибируется совсем, а РНК-полимераза III ингибируется только при высоких концентрациях антибиотика. Это соединение очень удобно для разграничения ферментов, ответственных за транскрипцию конкретных видов РНК.

### б. Ингибирование трансляции

Здесь мы рассмотрим действие лишь некоторых широко используемых антибиотиков—ингибиторов трансляции как у про-, так и у эукариот (табл. 3.5 и 3.6).

Высокоосновный трисахарид стрептомицин, как и большинство других аминогликозидов, связывается только с S12-белком 30S-субчастицы бактериальной рибосом. Одним из результатов такого связывания является подавление образования правильного комплекса инициации и тем самым блокирование инициации синтеза полипептидной цепи. При связывании стрептомицина с 30S-субчастицей происходит

Таблица 3.5. Ингибиторы белкового синтеза у прокариот

Антибиотик	Действие
Стрептомицин	Связывается с 30S-субчастицей и подавляет связывание Fmet-ТРНК <sup>Met</sup> с Р-участком; вызывает ошибки в трансляции
Неомицин, канамицин	То же
Хлорамфеникол	Ингибирует пептидилтрансферазную активность 70S-рибосомы
Тетрациклин	Ингибирует присоединение аминокислот к 30S-субчастице
Эритромицин	Связывается со свободной 50S-субчастицей и предотвращает образование 70S-рибосомы; на активную 70S-рибосому не действует
Пуромицин	Вызывает преждевременную терминацию синтеза белка, действуя как аналог аминоацил-ТРНК
Фузидиевая кислота	Подавляет связывание аминоацил-ТРНК с А-участком, блокируя высвобождение EF-G после транслокации
Казугамицин	Подавляет связывание Fmet-ТРНК <sup>Met</sup> с Р-участком
Линкомицин	Ингибирует активность пептидилтрансферазы
Кирромицин	Связывается с EF-Tu, стимулирует образование (EF-Tu)-GTP и связывание тройного комплекса с рибосомой, но подавляет высвобождение EF-Tu
Тиострептон	Предотвращает транслокацию, ингибируя EF-G

также модификация структуры А-участка, что приводит к случайным неправильным спариваниям между аминокислот-ТРНК и кодонами. У мутантов бактерий, устойчивых к стрептомицину, 30S-субчастицы изменены таким образом, что они либо не способны к связыванию со стрептомицином, либо если связывают его, то мешают его ингибирующему действию. С рибосомами эукариот стрептомицин не связывается и поэтому на их функцию не влияет.

Пуромицин из *Streptomyces alboniger* блокирует синтез полипептидов на рибосомах как у про-, так и у эукариот, имитируя действие аминокислот-ТРНК (рис. 3.55). Попав в клетку, пуромицин занимает свободный А-участок, и его аминокислотная группа действует как акцептор растущей полипептидной цепи, почти аналогично аминокислотной группе связанной аминокислот-ТРНК. Поскольку взаимодействие с кодоном отсутствует, образовавшийся полипептидилпуромицин отделяется от рибосомы, а рибосома — от мРНК.

Следовательно, пуромицин индуцирует abortивную терминацию синтеза полипептида.

Некоторые антибиотики (в том числе эритромицин из *Streptomyces erythreus* и фузидиевая кислота из *Fusidium coccineum*) блокируют перемещение рибосомы после завершения элонгации или пептидилтрансферазную реакцию. В результате рибосома не может переместиться к следующему кодону, пептидил-ТРНК остается связанной с А-участком и блокируется взаимодействие со следующей аминокислот-ТРНК.

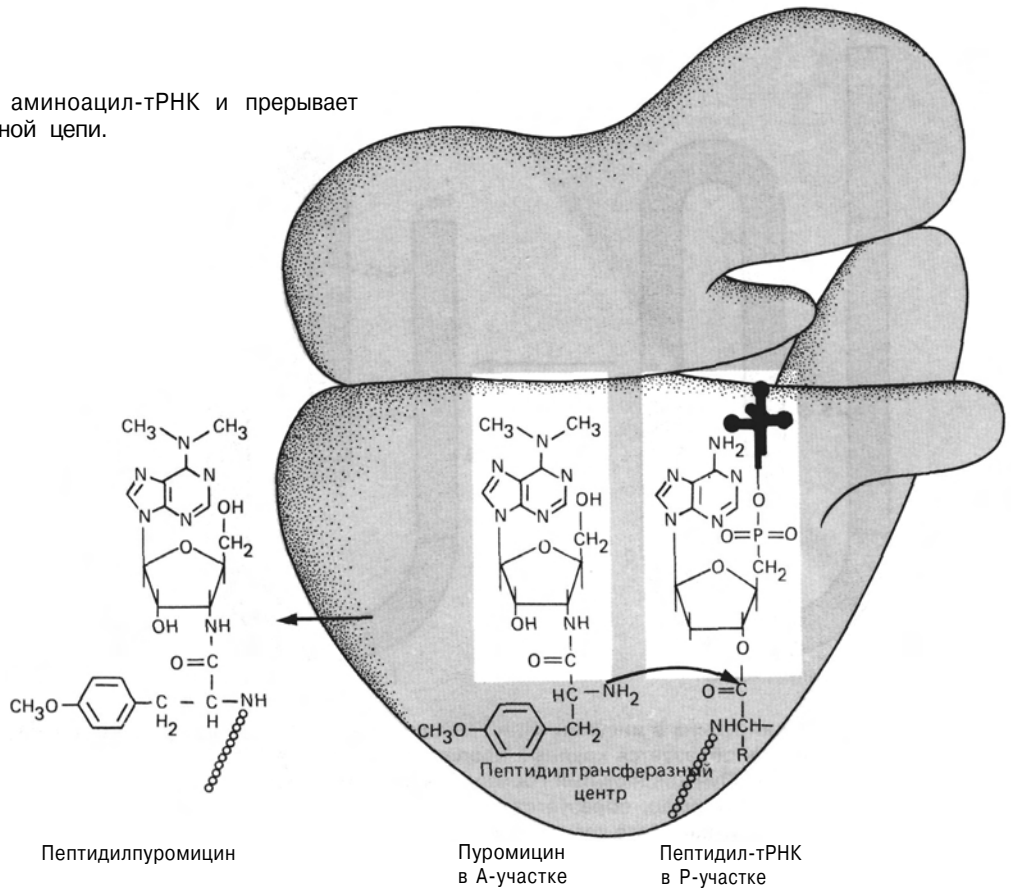
Хлорамфеникол и циклогексимид подавляют один и тот же этап синтеза белковой цепи. Хлорамфеникол действует на рибосомы бактерий, митохондрий и хлоропластов, а циклогексимид — только на рибосомы эукариот, находящиеся в цитоплазме. Оба соединения влияют на пептидилтрансферазную активность больших рибосомных субчастиц соответствующих рибосом, в результате чего элонгация пептидной цепи блокируется, а А- и Р-участки оказываются занятыми.

Таблица 3.6. Ингибиторы белкового синтеза у эукариот

Ингибитор	Действие
Абрин, ризин	Подавляет связывание аминокислот-ТРНК
Дифтерийный токсин	Катализирует реакцию между NAD и eIF-2 с образованием неактивного фактора; подавляет транслокацию
Хлорамфеникол	Ингибирует активность пептидилтрансферазы митохондриальных рибосом
Пуромицин	Вызывает преждевременную терминацию синтеза белка, действуя как аналог аминокислот-ТРНК
Фузидиевая кислота	Подавляет транслокацию, ингибируя eEF-2
Анизомидин	То же
Циклогексимид	Ингибирует активность пептидилтрансферазы
Пактамицин	Препятствует локализации met-ТРНК <sup>Met</sup> , на 40S-рибосоме
Шовдомицин	Подавляет образование комплекса eIF-2-met-ТРНК <sup>Met</sup> -GTP
Спарсомидин	Подавляет транслокацию
Интерферон	Блокирует инициацию, индуцируя фосфорилирование eIF-2

**РИС. 3.55.**

Пуromицин имитирует аминоксил-тРНК и прерывает элонгацию полипептидной цепи.



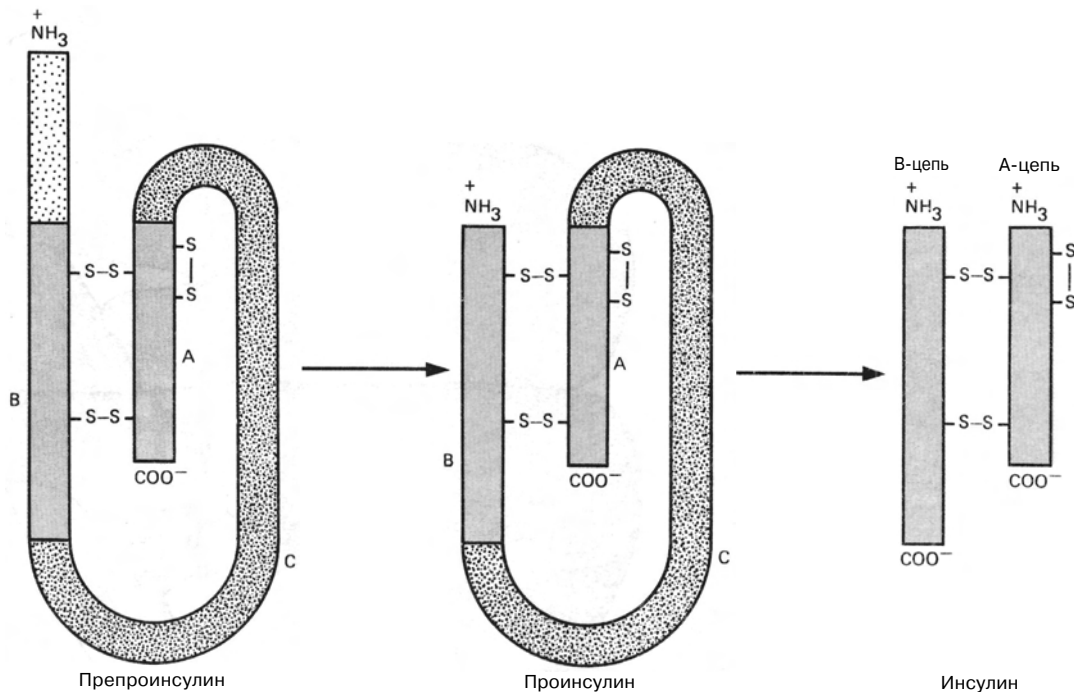
### 3.10. СУДЬБА СИНТЕЗИРОВАННЫХ БЕЛКОВ

Соединение аминокислот с образованием полипептидной цепи—это только первый шаг в процессе формирования многих белков. Полипептидная цепь должна быть уложена так, чтобы образовались правильные вторичная и третичная структуры, а в большинстве случаев отдельные полипептиды должны объединиться в функциональные олигомерные комплексы. Являются ли эти процессы полностью саморегулирующимися или протекают при участии ферментов—пока неясно. До образования функционально активных белков первичные продукты трансляции часто претерпевают разные сложные изменения. Приведем несколько примеров таких изменений.

#### а. Посттрансляционная модификация полипептидных цепей

Функционально активный гемоглобин образуется только после того, как  $\alpha$ - и  $\beta$ -цепи объединятся в  $\alpha_2\beta_2$ -структуру и с боковыми группами аминокис-

лот обеих субъединиц свяжется гемогруппа. Для того чтобы пируваткарбоксилаза и ацетил-СоА-карбоксилаза стали активными ферментами, с определенными боковыми цепями их аминокислот должен ковалентно связаться биотин. Некоторые белки, участвующие в процессе свертывания крови, должны претерпеть карбоксилирование специфических остатков глутаминовой кислоты, для того чтобы образовались центры связывания  $\text{Ca}^{2+}$ . Для образования коллагена должно произойти гидроксирование специфических пролиновых и лизиновых остатков. Очень важным и широко распространенным способом регуляции метаболизма являются фосфорилирование и дефосфорилирование определенных остатков серина, треонина и тирозина особыми протеинкиназами и протеинфосфатазами соответственно. Многие протеолитические белки, участвующие в процессах пищеварения и свертывания крови, синтезируются в виде крупных предшественников, которые далее активируются путем отщепления участка полипептидной цепи. Инсулин синтезируется в виде препроинсулинового полипептида и превращается в зрелый инсулин после расщепления цепи и удаления сначала N-концевого, а затем внутренне-

**РИС. 3.56.**

Превращение препроинсулина в инсулин. При трансляции инсулиновой мРНК образуется крупный предшественник - препроинсулин. Функциональный гормон, представляющий собой гетеродимер, образуется после отщепления N-концевого сигнального сегмента и расщепления молекулы проинсулина в двух местах, в результа-

те чего удаляется центральная часть (С) первичного трансляционного продукта. Остающиеся А- и В-цепи активного инсулина удерживаются вместе дисульфидными связями, образовавшимися в исходном полипептиде.

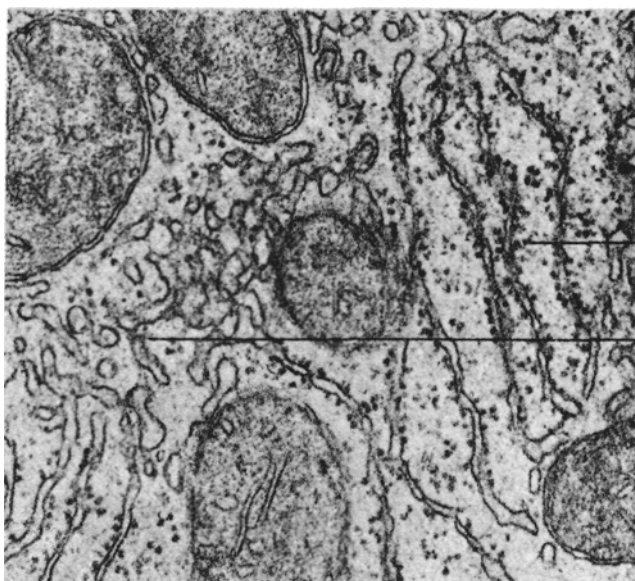
го сегмента (рис. 3.56). Многие вирусные белки, гормоны, нейропептиды образуются из первичных трансляционных продуктов полипротеинов—в результате расщепления по многим сайтам и образования нескольких зрелых белков и пептидов меньшего размера.

## б. Доставка эукариотических белков к клеточным мембранам и проникновение через них

В клетках эукариот помимо плазматических мембран имеются разнообразные внутриклеточные мембраны, ограничивающие различные клеточные органеллы: митохондрии, хлоропласты (у растений), **эндоплазматический ретикулум, аппарат Гольджи, пероксиомы, лизосомы и секреторные пузырьки (везикулы)** (рис. 1.1, А). Каким образом белки, предназначенные для столь многочисленных клеточных мембран и компартментов, попадают к месту своего назначения? Закодирована ли в белковой структуре информация об их внутри- или внеклеточной локализации? Если да, то какова химическая приро-

да такой информации и как она реализуется?

*Сложный транспортный аппарат эукариотических клеток.* Рибосомы в эукариотических клетках существуют в двух состояниях: в свободном и связанном с мембранами. Белки, предназначенные для некоторых органелл и **цитозоля**—водной фракции цитоплазмы,—синтезируются на свободных рибосомах, а белки, остающиеся в лизосомах, структурах Гольджи и плазматической мембране, образуются на рибосомах, ассоциированных с эндоплазматическим ретикулумом (ЭР)—непрерывной сети внутриклеточных мембран, окружающих пространство, называемое **просветом ЭР**. Мембраны ЭР, с которыми связаны рибосомы, называются **шероховатым ЭР**, а мембраны, не связанные с рибосомами,— **гладким ЭР** (рис. 3.57). Белки, синтезированные на рибосомах, связанных с ЭР, проходят через мембраны ЭР в просвет ЭР, а «прокладывает» им путь N-концевой участок (рис. 3.58). Затем этот участок, состоящий преимущественно из гидрофобных аминокислот,— **сигнальная последовательность**—отщепляется специфическими эндопептидазами, локализованными в просвете ЭР. Такой направленный перенос белков от рибосом в просвет ЭР начинается уже во время



Шероховатый  
эндоплазматический  
ретикулум

Гладкий  
эндоплазматический  
ретикулум

**РИС. 3.57.**

Электронная микрофотография среза печени крысы, на которой видны гладкий и шероховатый эндоплазматический ретикулум (ЭР). Мембраны гладкого ЭР (слева вверху) почти не содержат связанных рибосомных час-

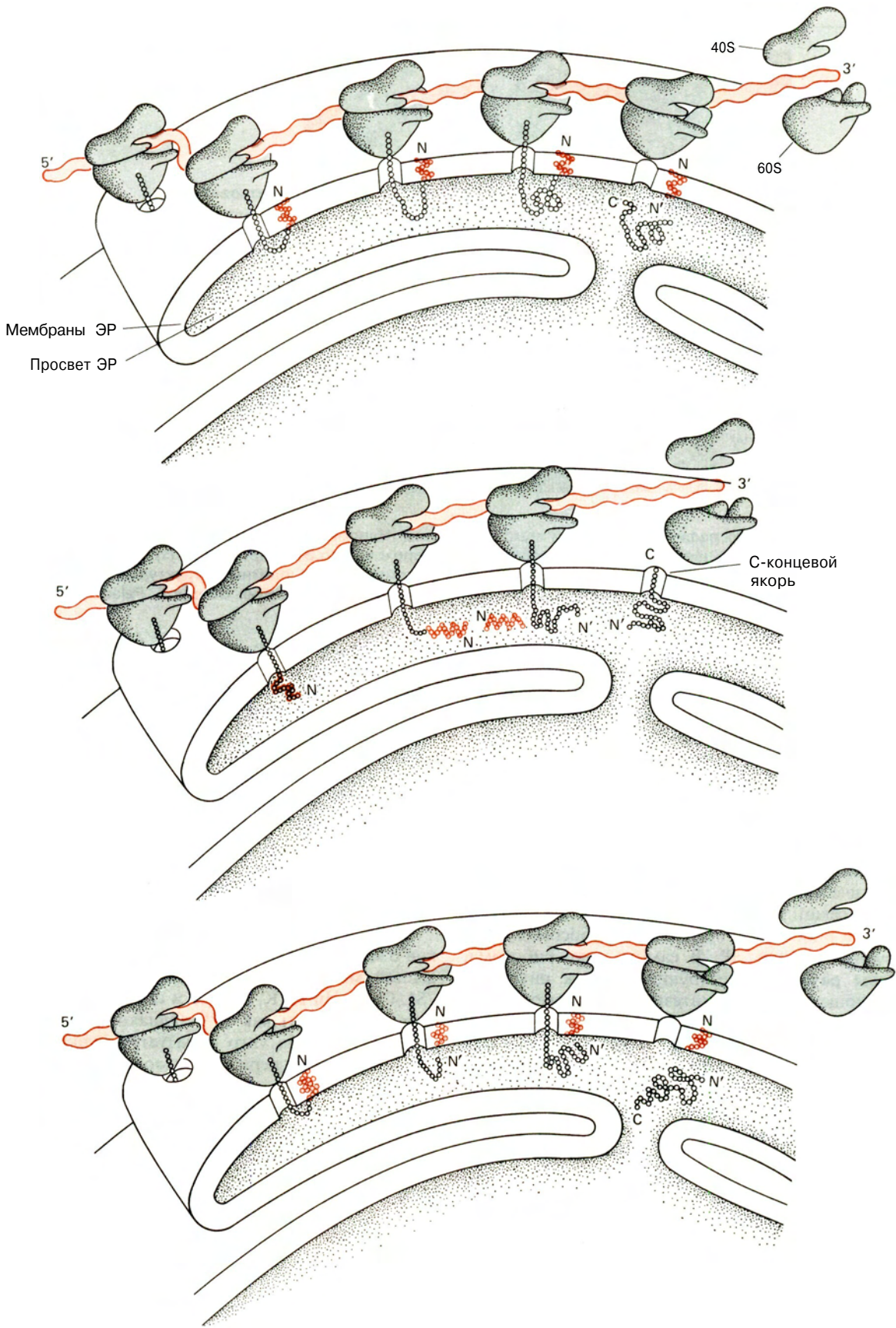
тиц. Напротив, протяженные мембранные структуры шероховатого ЭР (справа и внизу) усеяны рибосомами. Крупные, более плотные структуры - митохондрии. (С любезного разрешения George E. Palade.)

синтеза полипептидной цепи и поэтому называется **котрансляционным транспортом**.

*Направление синтезируемых полипептидных цепей в просвет ЭР.* Транспорт мембранных и секретрируемых белков в просвет ЭР опосредуется взаимодействием полипептидной сигнальной последовательности с **сигнал-распознающей частицей (СРЧ) и рецептором этой частицы (СРЧ-Р)**. Что заставляет рибосому направиться к ЭР, если белок, который она синтезирует, должен впоследствии секретироваться или транспортироваться в лизосомы или плазматические мембраны? Несомненно, это зависит не от самой рибосомы. Белки цитозоля синтезируются на рибосомах, идентичных рибосомам, синтезирующим белки ЭР. В самом деле, рибосома, в одном раунде синтезирующая белок цитозоля, в следующем может синтезировать мембранный белок. Попадание белка в ЭР происходит только в результате взаимодействия СРЧ с сигнальной последовательностью синтезируемой белковой цепи (рис. 3.59). Комплекс «СРЧ—сигнальная последовательность» взаимодействует впоследствии с рецептором СРЧ, локализованным в мембране ЭР. Таким образом, СРЧ служит адаптером между аппаратом синтеза белка в цитоплазме и аппаратом его доставки и отвечает за попадание белков в просвет ЭР. СРЧ связывается с сигнальной последовательностью сразу после выхода ее из рибосомы, т.е. после синтеза сегмента длиной примерно 70 аминокислот.

Элонгация полипептидной цепи замедляется до тех пор, пока комплекс СРЧ—сигнальная последовательность не свяжется с СРЧ-Р в мембране ЭР. Сразу после этого СРЧ отделяется, скорость синтеза полипептидной цепи увеличивается и растущая полипептидная цепь протягивается сквозь мембрану в просвет. Итак, ассоциация транслирующей рибосомы с СРЧ-Р приводит к однонаправленному переносу растущей полипептидной цепи в просвет ЭР. Особая пептидаза, локализованная в просвете ЭР, осуществляет специфическое отщепление сигнальной последовательности от новосинтезированного белка.

СРЧ представляет собой комплекс из шести белков с мол. массами от 10000 до 75000 Да и единственной молекулы РНК длиной 300 нуклеотидов — 7SL-РНК. Ни РНК, ни белки сами по себе не могут функционировать как СРЧ. Однако при смешивании РНК и белков *in vitro* образуется функциональная СРЧ. В таких экспериментах молекулы 7SL-РНК из клеток млекопитающих, амфибий и насекомых могут образовать СРЧ с белками млекопитающих. Этот факт наряду с данными о близости нуклеотидных последовательностей 7SL-РНК из этих источников говорит о том, что структура 7SL-РНК эволюционно консервативна. Состоит ли роль 7SL-РНК в прямом узнавании сигнальной последовательности или она служит каркасом для сборки разных белковых субъединиц в функциональной СРЧ, а воз-



**РИС. 3.58.**

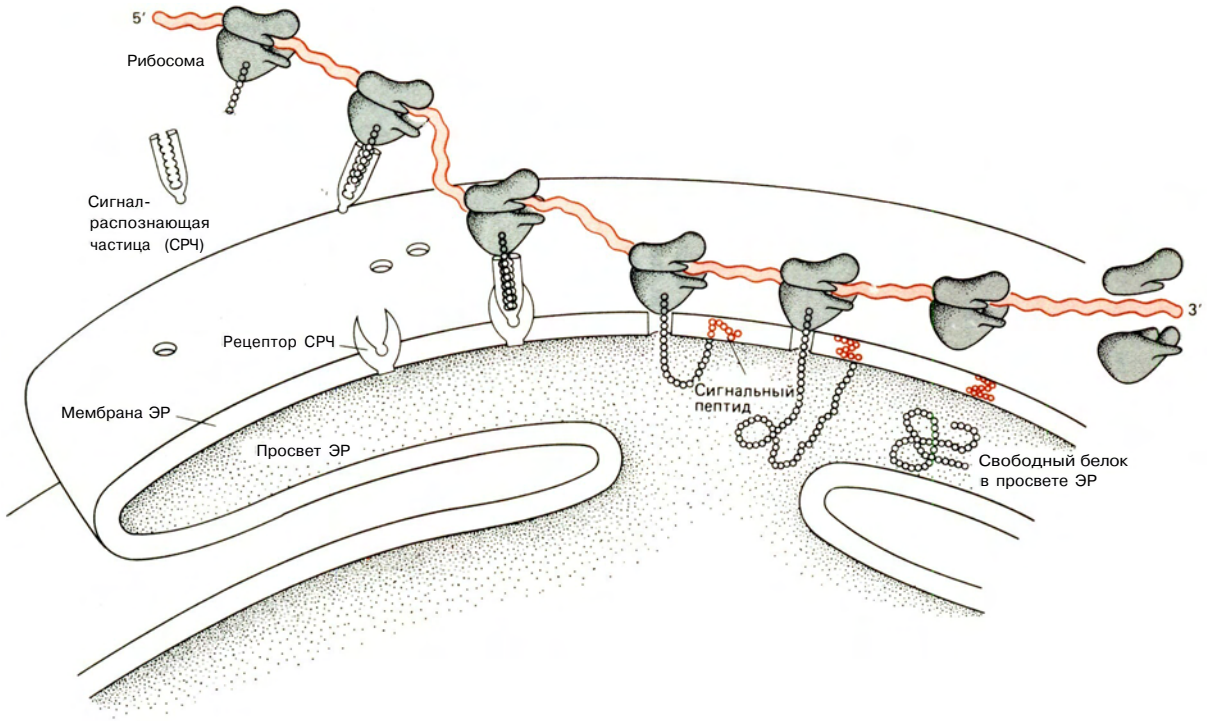
Котрансляционный транспорт полипептидных цепей. Полипептиды, синтезируемые на рибосомах шероховатого ЭР, направляются через мембраны в просвет ЭР. Гидрофобный N-концевой участок нового белка - сигнальная последовательность - проходит через мембрану вскоре после начала трансляции и по мере элонгации полипептида выходит в просвет ЭР. Синтезированные белковые молекулы подвергаются в ЭР различным модификациям в зависимости от их окончательной судьбы, однако все они утрачивают свою сигнальную последовательность. N' - это новый N-конец, образующийся после отщепления сигнальной последовательности. На верхнем рисунке сигнальная последовательность закреплена в мембране; она отщепляется по завершении трансляции, и полипептид с новым N-концом оказывается в просвете ЭР. На среднем рисунке сигнальная последовательность отщепляется еще до завершения трансляции; С-конец синтезированного полипептида закорен в мембране ЭР. На нижнем рисунке сигнальная последовательность отщепляется сразу после попадания полипептида в просвет ЭР.

можно, и то и другое, — пока неизвестно.

Сигнальная последовательность обычно находится на N-конце белков, предназначенных для экспорта к одной из клеточных мембран или внутриклеточным органеллам. В некоторых случаях, однако, аминокислотная последовательность, узнаваемая СРЧ, не является ни N-концевой, ни отщепляемой от транспортируемого белка, а отдельные секретируемые белки не содержат вообще никакой сигнальной последовательности.

Длина сигнальной последовательности колеблется от 15 до 35 аминокислот, и в первых трех ее четвертях преобладают гидрофобные остатки (рис. 3.60), однако никакой консервативной аминокислотной последовательности этот участок не имеет. Для узнавания сигнальной последовательности СРЧ скорее важна ее вторичная структура. Интересно, что сигнальные последовательности белков эукариот узнаются аппаратом транслокации бактерий, а сигнальные последовательности секретируемых бактериями белков могут проникать через мембранные компоненты эукариотических клеток.

В пределах сигнальной последовательности находится также и сайт, распознаваемый сигнальной пептидазой. Обычно расщепление происходит со

**РИС. 3.59.**

Перенос синтезированного полипептида в ЭР опосредуется сигналь-распознающей частицей (СРЧ). Показано, как СРЧ связывается с N-концевой сигнальной последо-

вательностью растущей полипептидной цепи и затем с СРЧ-рецептором в мембране ЭР, в результате чего пептид проникает в просвет ЭР.

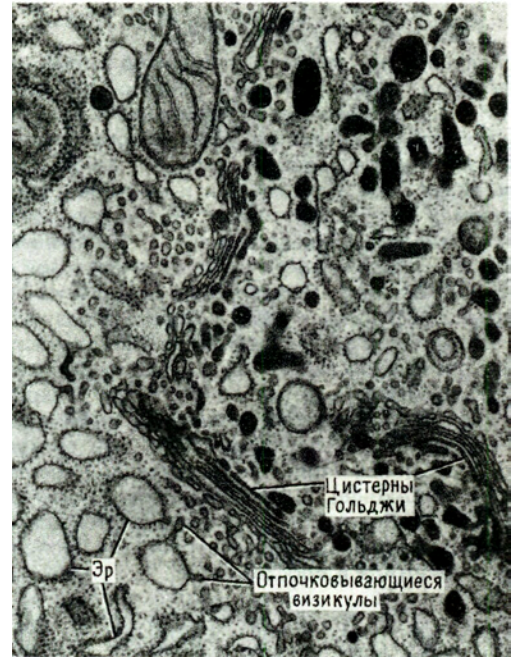


Белок	Заряженный участок										Гидрофобный участок												
	Met	Lys	Ala	Thr	Lys	Leu	Val	Leu	Gly	Ala	Val	Ile	Leu	Gly	Ser	Thr	Leu	Leu	Ala	Gly	Ser	Pro	Glu
Липопротейн	Met	Lys	Ala	Thr	Lys	Leu	Val	Leu	Gly	Ala	Val	Ile	Leu	Gly	Ser	Thr	Leu	Leu	Ala	Gly	Ser	Pro	Glu
β-Лактамаза рВR322	Met	Ser	Ile	Gln	His	Phe	Arg	Ala	Leu	Ile	Pro	Phe	Ala	Ala	Phe	Cys	Leu	Pro	Val	Phe	Ala	His	Pro
Основной белок оболочечки фага td	Met	Lys	Ser	Leu	Val	Leu	Lys	Ala	Ser	Val	Ala	Thr	Leu	Val	Pro	Met	Leu	Ser	Phe	Ala	Ala	Glu	Gly
Щелочная фосфатаза Мальтозосвязывающий белок	Met	Lys	Ile	Lys	Thr	Gly	Ala	Arg	Ile	Leu	Ala	Leu	Ala	Leu	Thr	Met	Met	Phe	Ser	Ala	Ser	Ala	Leu
Лейцинсвязывающий белок	Met	Lys	Ala	Asn	Ala	Lys	Thr	Ile	Ile	Ala	Gly	Met	Ile	Ala	Leu	Ala	Ile	Ser	His	Thr	Ala	Met	Ala
Пролактин	Met	Lys	Trp	Val	Thr	Phe	Leu	Leu	Leu	Phe	Ile	Ser	Gly	Ala	Phe	Ser	Arg	Pro	Gly	Thr	Arg	Cys	Ile
Легкая цепь IgG	Met	Arg	Ser	Leu	Ile	Leu	Val	Val	Phe	Gly	Phe	Leu	Leu	Leu	Phe	Pro	Gly	Thr	Arg	Cys	Lys	Leu	Leu
Лизоцим	Met	Asn	Ser	Gln	Val	Ser	Ala	Arg	Lys	Ala	Gly	Thr	Leu	Leu	Leu	Met	Met	Ser	Asn	Leu	Leu	Leu	Leu
Пролактин	Met	Lys	Cys	Leu	Leu	Tyr	Leu	Ala	Phe	Leu	Phe	Ile	His	Val	Asn	Cys	Lys	Met	Ser	Asn	Leu	Leu	Leu
Гликопротеин вируса везикулярного стоматита	Met	Lys	Cys	Leu	Leu	Tyr	Leu	Ala	Phe	Leu	Phe	Ile	His	Val	Asn	Cys	Lys	Met	Ser	Asn	Leu	Leu	Leu

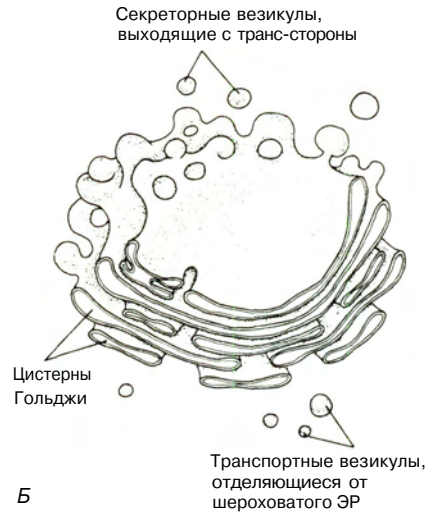
**РИС. 3.60.**

Характерные последовательности, находящиеся на N-конце некоторых прокариотических и эукариотических белков. В начале этих последовательностей имеется короткий участок из одной или более заряжен-

ных аминокислот, а за ним следует гидрофобный сегмент. [S. Michaelis, J. Beckwith, *Ann. Rev. Microbiol.*, **36** (1982), p. 435; D.P. Leader, *Trends Biochem. Sci.*, **4** (1979), p. 205.]



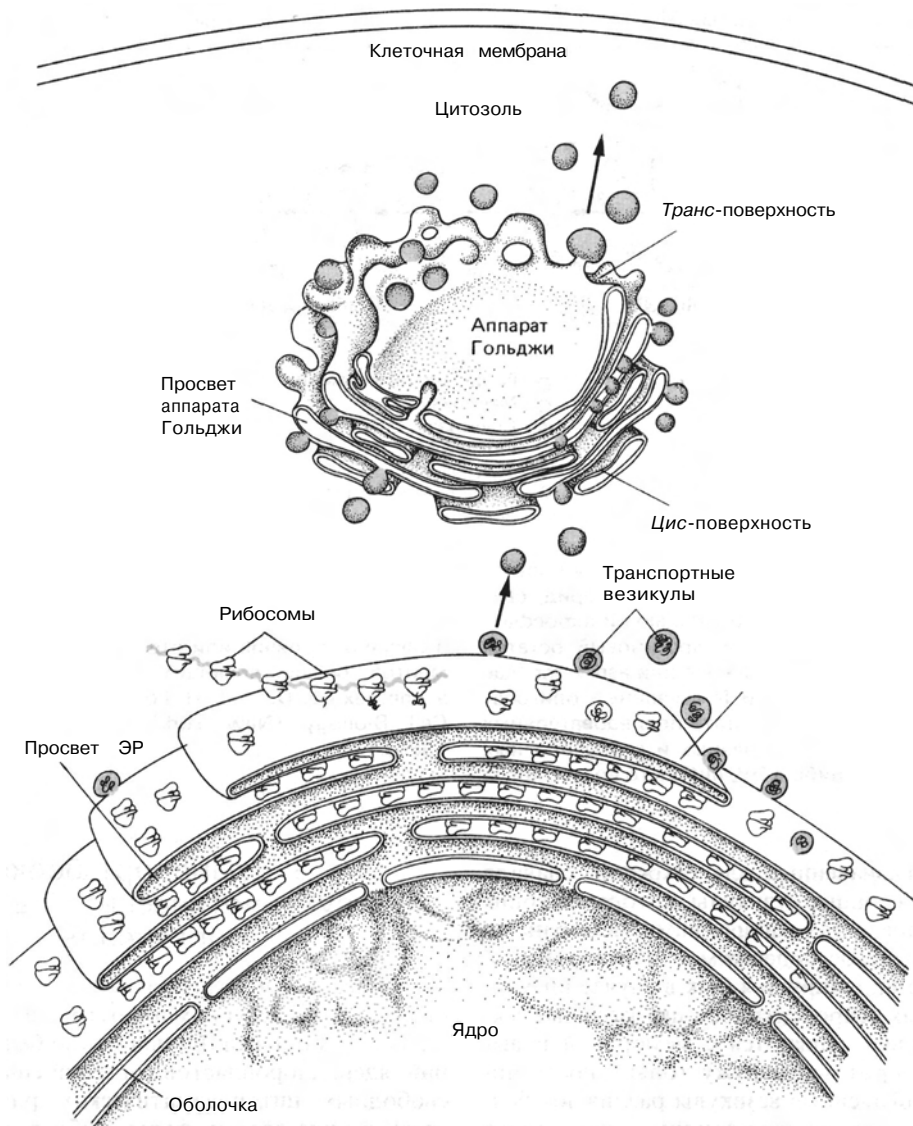
А



Б

**РИС. 3.61.**

Аппарат Гольджи. А. Электронная микрофотография экзокринных клеток поджелудочной железы крысы, на которой видны примыкающие к ЭР цистерны аппарата Гольджи. Цистерны, расположенные ближе к ЭР, называются *цис*-цистернами, а дальше - *транс*. Видны везикулы, отпочковывающиеся от элементов ЭР. [E.M. Merisko, M. Fletcher, G. Palade, *Pancreas*, **1** (1986), p. 95. С любезного разрешения M.G. Farquhar.] Б. Схематическое изображение цистерн и везикул аппарата Гольджи. [J. Darnell, H. Lodish, D. Baltimore, *Molecular Cell Biology* (New York: Scientific American Books, 1986) и в соответствии с моделью J. Kephart]

**РИС. 3.62.**

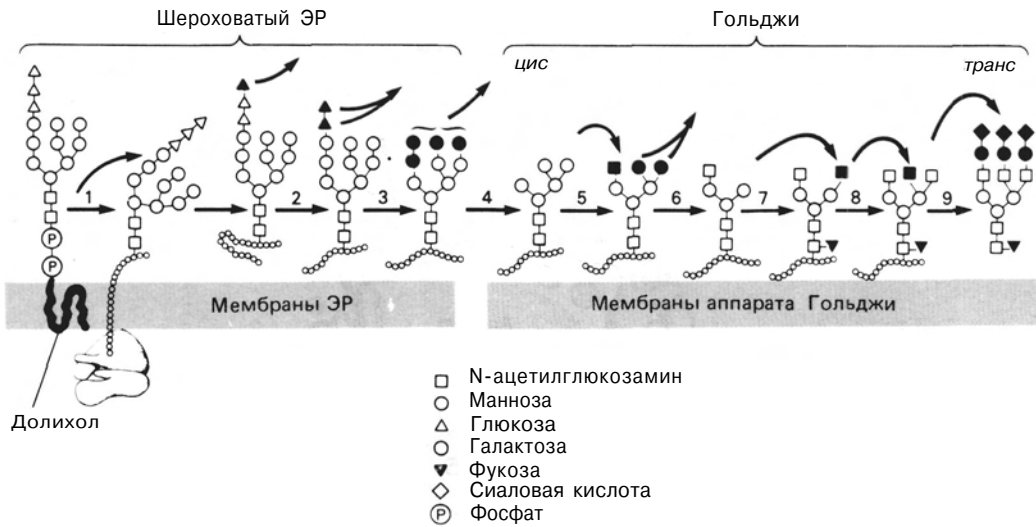
Транспортные везикулы, содержащие полипептиды, которые должны попасть в лизосомы, плазматические мембраны или секретироваться, отделяются от ЭР и сливаются с *цис*-поверхностью аппарата Гольджи. Полипептиды проходят через аппарат Гольджи в составе этих везикул и выходят в цитозоль через *транс*-поверх-

стороны С-конца остатков глицина, серина или аланина, поэтому N-конец многих зрелых секретруемых или мембранных белков соседствует с одной из этих аминокислот в сигнальной последовательности.

*Транспорт белков от ЭР к аппарату Гольджи и из него.* Белки направляются к лизосомам, плазматическим мембранам или секреторируются с помощью аппарата Гольджи — набора тесно упакован-

ность в составе транспортных везикул, отделяющихся от цистерн Гольджи. Во время прохождения полипептидов через структуры аппарата Гольджи они претерпевают такие посттрансляционные модификации, как протеолитическое расщепление и гликозилирование.

ных, взаимопроникающих, окруженных мембранами цистерн (рис. 3.61). Перенос белков к аппарату Гольджи осуществляется с помощью так называемых окаймленных пузырьков (везикул), отпочковывающихся от ЭР и сливающихся с цистернами Гольджи (рис. 3.62). Белки проходят через аппарат Гольджи от *цис*-цистерн к *транс*. Транспорт белков через цистерны также осуществляется с помощью окаймленных транспортных везикул.

**РИС. 3.63.**

Гликозилирование белков в шероховатом ЭР и цистернах Гольджи. В ЭР разветвленный олигосахарид, связанный с липидом долихолом с помощью пиррофосфатного мостика, переносится на аспарагиновый остаток белка в последовательности asn-x-ser или asn-x-thr (где x - любая аминокислота). Затем N-связанный олигосахарид модифицируется в результате последовательных специфических гликозилгидролазных и гликозиллазных реакций, протекающих либо в ЭР, либо в то время, когда

белок проходит через аппарат Гольджи. Разные белки содержат различные N-связанные олигосахариды в зависимости от первичной и вторичной структуры белка, а также от наличия или отсутствия специфических ферментов, осуществляющих процессинг олигосахаридов в клетках [J. Darnell, H. Lodish, D. Baltimore, *Molecular Cell Biology* (New York: Scientific American Books, 1986).]

Белки, предназначенные для секреции, сначала попадают в секреторные везикулы, которые в конце концов сливаются с плазматическими мембранами и высвобождают свое содержимое наружу. Когда секреция белков индуцируется (как в случае инсулина или некоторых нейропептидов), цитозольные секреторные пузырьки сливаются с мембраной и высвобождают содержимое наружу только после индукции. Как окаймленные везикулы различают белки, предназначенные для того или иного клеточного компартмента, неясно. Неизвестно также, различаются ли между собой везикулы, несущие разные белки, и если различаются, то чем именно.

**Гликозилирование.** Проходя через ЭР и аппарат Гольджи, белки подвергаются интенсивной модификации (рис. 3.63). Специальный фермент ЭР присоединяет разветвленный олигосахарид к специфическим аспарагиновым остаткам транспортируемых белков. Далее олигосахарид подвергается действию целого ряда специфических гликозилгидролазов и гликозиллаз; эти реакции протекают в ЭР и в то время, когда белки проходят через аппарат Гольджи.

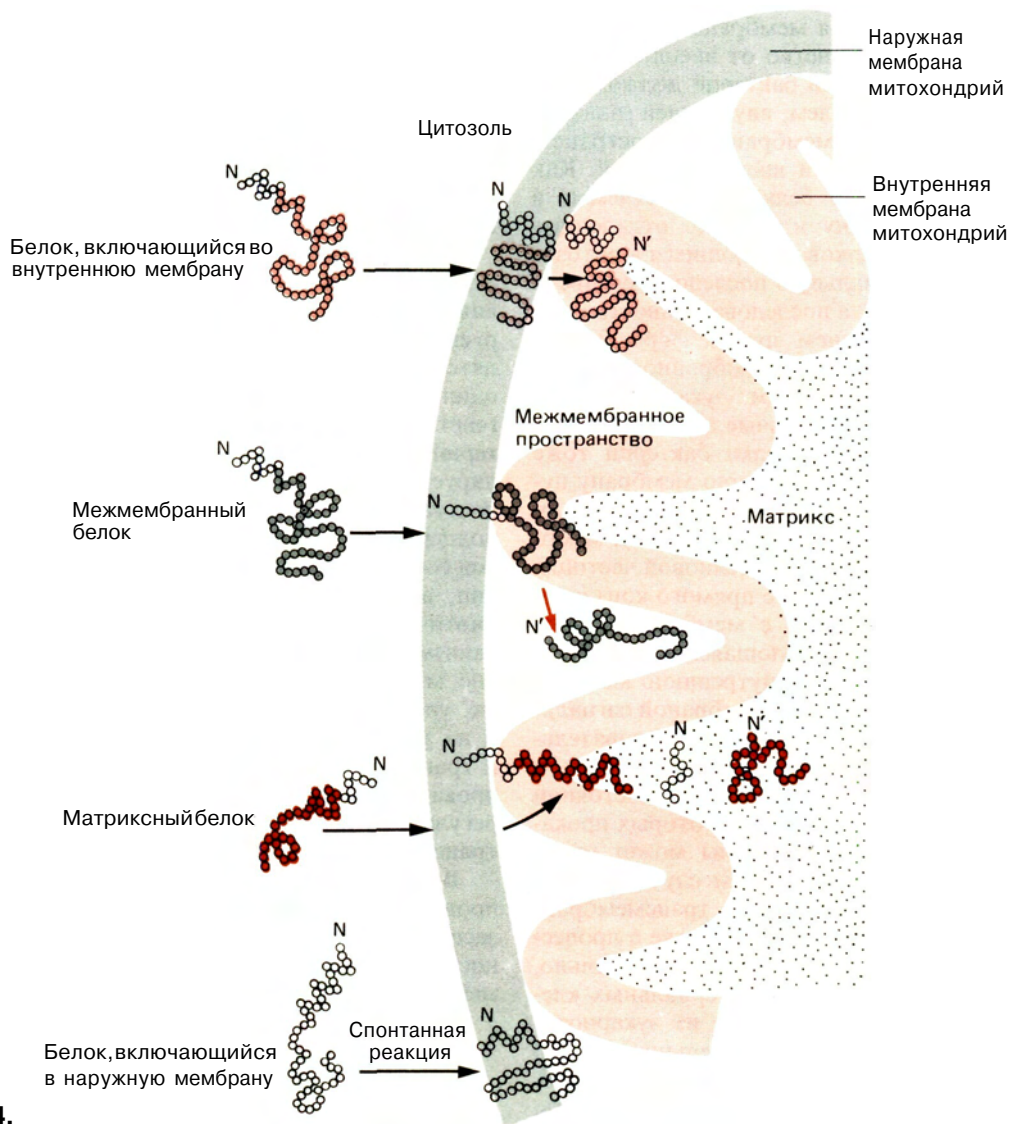
## в. Транспорт белков в эукариотические клеточные органеллы

Не все белки, направляемые к внутриклеточным органеллам, следуют маршрутом, включающим ЭР и аппарат Гольджи. Большинство белков митохондрий, ядер, хлоропластов растений синтезируются на свободных цитоплазматических рибосомах либо сразу в виде зрелых форм, либо в виде предшественников и затем подхватываются соответствующими органеллами. Остальные белки митохондрий и хлоропластов кодируются ДНК, содержащейся в этих органеллах, и синтезируются с помощью аппарата трансляции этих органелл аналогично тому, как это происходит в цитоплазме.

Митохондрии, например, имеют несколько компартментов: матрикс, наружную мембрану и межмембранное пространство (рис. 3.64). Белки, кодируемые ядерной ДНК, которые должны быть локализованы в матриксе, синтезируются в виде предшественника и содержат от 25 до 60 лишних аминокислот на N-конце. После связывания с рецептором на наружной мембране предшественник транспортируется в этот внутренний матрикс. Движущей силой процесса является электрохимический

потенциал на внутренней мембране, генерируемый с помощью энергетически выгодных реакций, протекающих в митохондриях. Одновременно с переносом или сразу после него лишние аминокислоты на N-конце удаляются специфическими протеазами, находящимися в матриксе. Некоторые белки остаются во внутренней мембране, а другие проходят сквозь мембрану и попадают в матрикс. Белки, синтезированные в цитозоле и направленные в межмембранное пространство митохондрий, переносятся несколькими способами (рис. 3.64). Некоторые предшественники содержат особый N-концевой сегмент и попадают в митохондрии так, как это было

описано выше. У других предшественников отщепление концевой последовательности приводит к тому, что белки остаются временно встроенными во внутреннюю мембрану; затем следующий процессирующий фермент расщепляет этот промежуточный продукт, и зрелый белок выходит в межмембранное пространство. Белки, локализуемые в наружной мембране, транспортируются энергонезависимым путем и без превращения предшественника в зрелый белок. В некоторых случаях переносу цитозольного белка в наружную мембрану способствуют конформационные изменения белковой молекулы при связывании протестической группы. Белки закрепляют-



**РИС. 3.64.**

Включение белков в митохондрии. В зависимости от конечной локализации митохондриальных белков (она указана слева) последние используют тот или иной

способ транспортировки и подвергаются разным модификациям.

ся в мембранах с помощью якорных последовательностей, которые проникают в липидный бислой; подобные последовательности могут находиться и на амино-, и на карбокси-конце, и даже в середине полипептидной цепи.

### г. Транспорт белков в клетках прокариот

И про-, и эукариотические клетки имеют пограничные липопротеиновые мембраны. Внутренняя мембрана бактериальных клеток отделяет цитоплазму от **межмембранного**, или **периплазматического**, пространства. Наружная мембрана отделяет периплазматическое пространство от внешней среды. Новосинтезированные белки бактерий должны распределяться между цитозолем, внутренней (плазматической) мембраной, межмембранным пространством, наружной мембраной и внешней средой. Как и у эукариот, у бактерий белки, направляемые в плазматическую мембрану или более отдаленные места, отличаются от белков, остающихся в цитозоле, тем, что имеют сигнальную последовательность на N-конце (рис. 3.60). Эта последовательность тоже отщепляется с образованием зрелых форм белков внутренней мембраны или межмембранного пространства. Однако в отличие от эукариотических мембранных белков бактериальные аналоги не гликозилируются. Мембранные белки бактерий тоже транспортируются через внутреннюю мембрану путем котрансляционного переноса. Направляется ли сигнальная последовательность к внутренней мембране с помощью некой нуклеопротеиновой частицы, как у эукариот, или в результате прямого контактирования гидрофобного конца с мембраной — неизвестно. В любом случае образующаяся полипептидная цепь протягивается через внутреннюю мембрану по мере синтеза. Связанная с мембраной сигнальная протеаза отщепляет сигнальную последовательность, и зрелый белок остается заякоренным в мембране или находится в свободном состоянии в периплазме. Векторный перенос некоторых прокариотических белков через мембраны может произойти и посттрансляционно; в этом случае движущей силой переноса также служит трансмембранный потенциал. При имеющемся сходстве в процессах транспорта у про- и эукариот неудивительно, что белки, экспортируемые из бактериальных клеток, могут экспортироваться также из эукариотических клеток, и наоборот. Неудивительно и то, что если к цитозольному белку искусственно пришить соответствующую сигнальную последовательность, то он сможет транспортироваться к мембране или даже пройти сквозь нее в зависимости от своей вторичной структуры.

## 3.11. РЕГУЛЯЦИЯ ГЕННОЙ ЭКСПРЕССИИ

Основным свойством как про-, так и эукариот является их способность осуществлять дифференциальную регуляцию экспрессии генов. Осуществляя контроль за тем, каким генам экспрессироваться, а каким нет, а также регулируя уровень экспрессии различных генов, клетки приспосабливают свой фенотип к определенным условиям внешней и внутренней среды. Часто гены экспрессируются последовательно: активация одного гена вызывает экспрессию другого или нескольких генов, приводя в конечном счете к каскаду событий. Некоторые гены или родственные группы генов экспрессируются координированно, т.е. отвечают на регуляторный сигнал одновременно и, как правило, в одинаковой степени. Вероятно, наиболее совершенные регуляторные системы — те, которые позволяют плюрипотентным **стволовым клеткам** дифференцироваться с образованием клеток разного типа в процессе развития сложного организма.

Фенотипические признаки клеток разных типов, а также одной и той же клетки в различных условиях зависят от количества и свойств продуцируемых ими структурных, каталитических и регуляторных белков. Регулироваться может какой-то один или несколько отдельных этапов считывания генетической информации при синтезе белка. У бактерий, например у *E. coli*, образование белков регулируется главным образом содержанием мРНК, доступной для трансляции. Дополнительный способ поддержания нужной концентрации клеточных белков состоит в регуляции различных этапов трансляции, а также скорости деградации белков. Эукариотические клетки обладают более сложными механизмами регуляции белкового состава. Содержание мРНК в цитоплазме регулируется не только на уровне инициации транскрипции в ядре, но и на уровне процессинга первичных транскриптов и транспорта зрелых РНК в цитоплазму. Подобно прокариотам, эукариотические клетки тоже могут регулировать как трансляцию, так и скорость транспорта и деградации белков.

В этом разделе мы на нескольких примерах проиллюстрируем механизмы регуляции генной экспрессии, используемые прокариотами, в частности *E. coli* и ее бактериофагами. Более сложные механизмы регуляции у эукариот обсуждаются в гл. 8.

### а. Регуляция содержания РНК в процессе биосинтеза

Образование РНК у прокариот чаще всего регулируется на уровне инициации транскрипции несколькими способами. Первый заключается в мо-

дификации структуры РНК-полимеразы. Так, бета- (или бета-прим) субъединица РНК-полимеразы *E. coli* изменяется при заражении клеток некоторыми бактериофагами. Другой пример — образование новой, сигма-субъединицы при споруляции определенных штаммов *Bacillus*. И в том и другом случае изменяется способность РНК-полимеразы к связыванию с промотором и скорость транскрипции соответствующих генов. Второй способ регуляции заключается в изменении пространственной структуры ДНК, что влияет на способность РНК-полимераз связываться с определенными промоторами и инициировать синтез РНК. И наконец, взаимодействие РНК-полимеразы с некоторыми промоторами может ингибироваться или стимулироваться белками, которые связываются с ДНК в месте присоединения полимеразы или вблизи него. На связывание таких регуляторных белков — **репрессоров** и **активаторов** часто влияют определенные метаболиты, играющие роль **коррепрессоров** и **коактиваторов**.

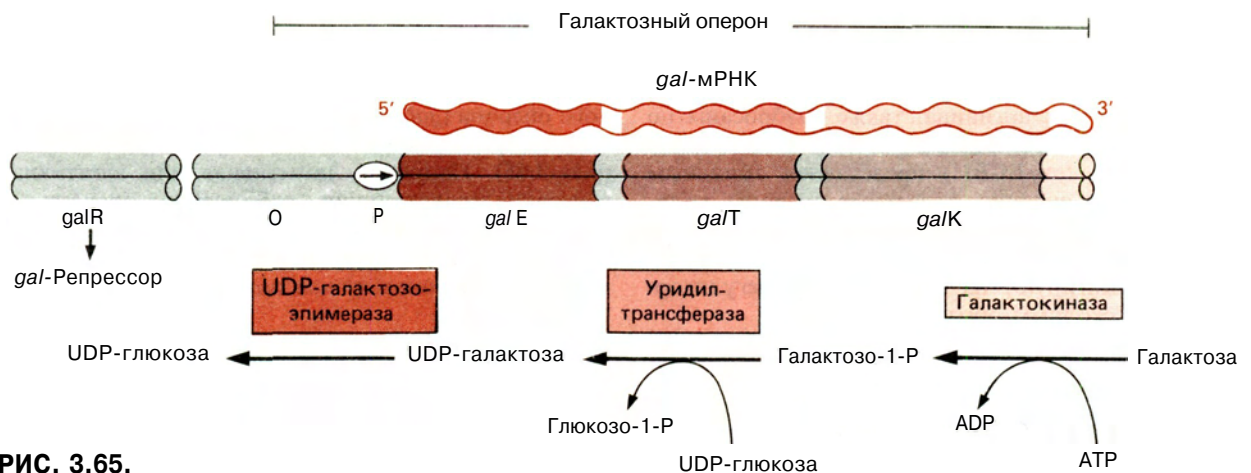
Скорость элонгации РНК зависит также от вторичной структуры мРНК. Сигналы, которые определяют, дойдет ли транскрипция до конца или закончится преждевременно, играют ключевую роль в регуляции уровня мРНК. В подобных случаях за прекращение или продолжение транскрипции отвечают определенные нуклеотидные последовательности и белки.

Вторичная структура транскриптов может также влиять на их стабильность, при этом особую роль в определении продолжительности жизни РНК в клетке играет структура 3'-конца. Одни мРНК очень стабильны и участвуют во многих циклах трансляции. Другие, наоборот, быстро разрушают-

ся уже в процессе трансляции. Контроль на уровне транскрипции особенно важен в случае наименее стабильных мРНК. Как уже указывалось, зрелые молекулы рРНК и тРНК — это продукты процессирования первичных транскриптов (разд. 3.3). Посттранскрипционные события в свою очередь тоже регулируются, и от них зависит содержание в клетке функциональных молекул РНК.

## б. Согласованная регуляция экспрессии прокариотических генов

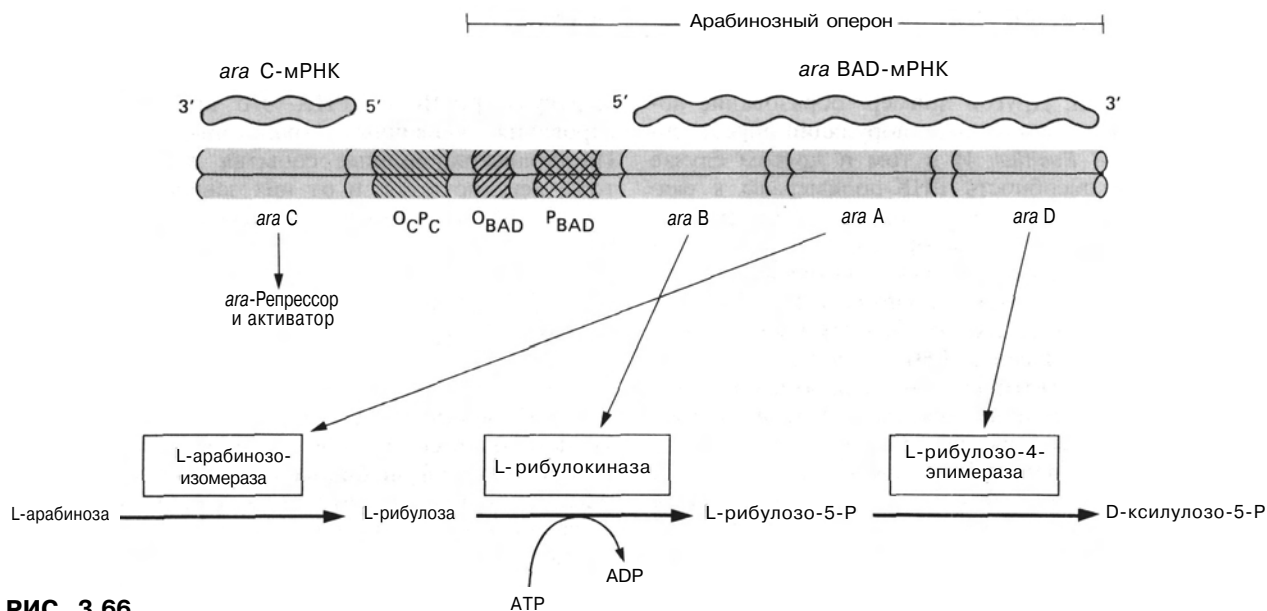
*Согласованная регуляция групп родственных генов.* У *E. coli* гены, кодирующие белки одного и того же метаболического пути или определяющие близкородственные функции, часто регулируются согласованно. Это значит, что их экспрессия начинается в ответ на один и тот же регуляторный сигнал. Гены, подчиняющиеся согласованной регуляции, в геноме часто бывают сцеплены и транскрибируются с промотора, находящегося на 5'-конце такой группы генов (кластера), в виде единственной молекулы РНК, называемой **полицистронным** (или полигенным) транскриптом. Группа координированно экспрессирующихся генов называется **опероном**. Три гена, кодирующие ферменты, ответственные за метаболизм галактозы у *E. coli*, организованы в оперон с промотором (P) и примыкающим к нему регуляторным сегментом — оператором (O) на 5'-конце транскрибируемой последовательности: *gal E-gal T-gal K* (рис. 3.65). Синтез ферментов, необходимых для утилизации арабинозы, регулируется двумя сопряженными транскрипционными единица-



**РИС. 3.65.**

Галактозный оперон (*gal*) *E. coli*. Представлены оператор (O), промотор (P) и три гена оперона — *gal E*, *gal T* и *gal K*. Внизу показаны реакции, катализируемые тремя генными продуктами — галактокиназой, уридилтрансфера-

зой и UDP-галактозоэпимеразой. Все три фермента транслируются с единственной полицистронной *gal*-мРНК. Ген *gal R*, кодирующий репрессор *gal*-оперона, не сцеплен с опероном.

**РИС. 3.66.**

Арабинозный оперон (*ara*) *E. coli* и примыкающий к нему ген *ara C*. Транскрипция этих двух единиц осуществляется в противоположных направлениях. Они

имеют собственные промоторы и операторы:  $O_C$  и  $P_C$  для *ara C* и  $O_{BAD}$  и  $P_{BAD}$  для *ara*-оперона.

ми (рис. 3.66). Первая—это *ara*-оперон, состоящий из трех генов, *ara B*, *ara A* и *ara D*, и некоего 5'-контролирующего участка. Вторая представлена геном *ara C*, кодирующим регуляторный белок, необходимый для транскрипции *ara*-оперона.

Гены, кодирующие несколько родственных функций, не всегда образуют единый оперон. Так, гены, кодирующие 30S- и 50S-рибосомные белки, организованы во множественные опероны, в чей состав иногда входят гены, кодирующие другие белки, которые участвуют в транскрипции и/или трансляции (рис. 3.67). Гены, продукты которых необходимы для синтеза аргинина, также разбросаны по нескольким генетическим локусам, и только один кластер генов (*arg* ЕСВН) представлен типичным опероном. Как правило, отдельные опероны, кодирующие родственные функции, имеют одинаковые или сходные регуляторные последовательности и поэтому реагируют на определенный регуляторный сигнал сходным образом.

**Позитивная и негативная регуляция.** Негативная регуляция инициации транскрипции, или репрессия, осуществляется белками-репрессорами, которые связываются с операторами. Поскольку последовательности оператора и промотора часто перекрываются, связывание репрессоров со своими операторами ограничивает доступ РНК-полимеразы к промотору, подавляя тем самым инициацию транскрипции. Позитивная регуляция может осу-

ществляться путем связывания специфических белков с нуклеотидными последовательностями, расположенными в области промотора. Считается, что связанный активаторный белок способствует ассоциации РНК-полимеразы с промотором и, следовательно, увеличивает вероятность инициации транскрипции.

Гены, кодирующие регуляторные белки, которые связываются с операторными или активаторными последовательностями, могут находиться как вблизи контролируемых ими генов, так и далеко от них. Например, ген, кодирующий репрессор галактозного оперона (*galR*), не сцеплен с транскрипционной единицей, состоящей из генов *galE*, *galT* и *galK* (рис. 3.65). Напротив, позитивная или негативная регуляция транскрипции арабинозного оперона зависит от того, образуется или нет комплекс между арабинозой и белком, кодируемым только сцепленным с опероном геном *araC* (рис. 3.66).

## в. Регуляция экспрессии лактозного оперона

**Негативная регуляция.** Бактерии *E. coli* могут использовать в качестве единственного источника углерода и энергии лактозу, поскольку они способны образовывать в большом количестве  $\beta$ -галактозидазу—фермент, расщепляющий лактозу на глюкозу и галактозу. Однако при росте на других источни-

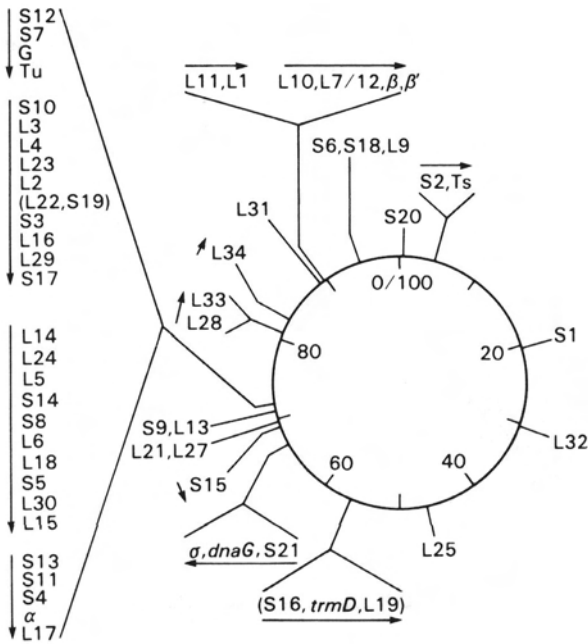


РИС. 3.67.

Положение на генетической карте *E. coli* генов, кодирующих рибосомные белки. Белки субчастиц 30S и 60S обозначены соответственно S и L. На карте указаны также позиции генов, кодирующих субъединицы  $\alpha$ ,  $\beta$ ,  $\beta'$  и  $\delta$  РНК-полимеразы; ДНК-праймазу (*dna G*) и тРНК-метилтрансферазу (*trm D*). G, Tu и Ts- факторы трансляции EF-G, EF-Tu и EF-Ts соответственно. Направление транскрипции (там, где оно известно) показано стрелками.

ках углерода в клетках *E. coli* образуется очень мало  $\beta$ -галактозидазы. Ген, ответственный за синтез  $\beta$ -галактозидазы (*lac Z*), называется индуцибельным, поскольку кодируемый им фермент синтезируется только тогда, когда в клетке присутствуют сахара, имеющие  $\beta$ -галактозилные остатки. Помимо  $\beta$ -галактозидазы,  $\beta$ -галактозиды индуцируют образование еще двух белков:  $\beta$ -галактозидпермеазы (кодируемой геном *lac Y*), необходимой для проникновения  $\beta$ -галактозидов в клетку, и  $\beta$ -галактозидтранс-ацетилазы (*lac A*), фермента с невыясненной пока функцией (рис. 3.68). В этих трех генах — *lac Z*, *lac Y* и *lac A* — содержится вся информация о белках, кодируемых *lac*-опероном. Они транскрибируются в единую полицистронную РНК, при трансляции которой образуются почти одинаковые количества соответствующих белков. Поэтому можно сказать, что три гена *lac*-оперона экспрессируются согласованно.

Со структурными генами *lac*-оперона связаны несколько типов регуляторных элементов, ответ-

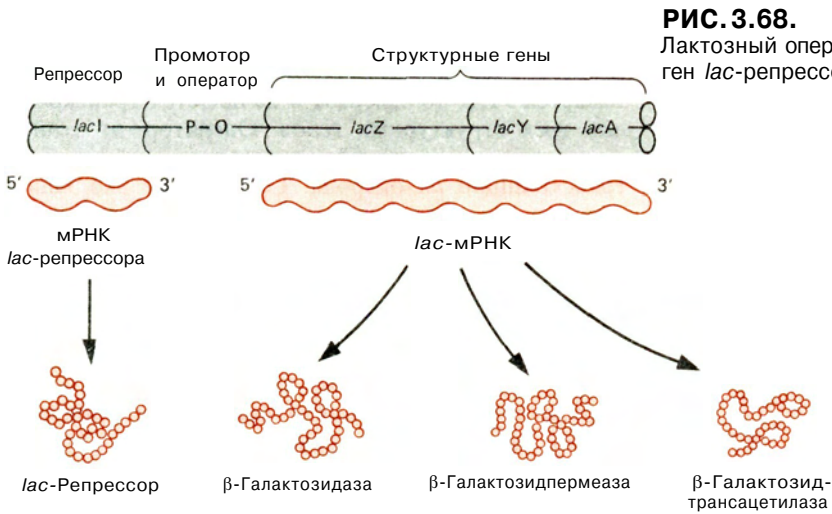
ственных за индуцибельность и координированную регуляцию этих генов (рис. 3.68). Промотор — это нуклеотидная последовательность, с которой связывается РНК-полимераза и начинается транскрипция трех структурных генов. Оператор — это сайт, с которым связывается *lac*-репрессор, подавляющий транскрипцию *lac*-оперона. Ген *lac I*, не входящий в состав *lac*-оперона, кодирует репрессор — полипептидную цепь с мол. массой 37000 Да. Репрессор прочно связывается с оператором, находясь в тетрамерной форме.

Поскольку промоторная и операторная последовательности перекрываются, связывание репрессора с оператором мешает связыванию РНК-полимеразы с промотором, что приводит к блокированию транскрипции структурных генов. Транскрипцию оперона можно индуцировать, если заблокировать связывание репрессора с оператором (рис. 3.69). Такое блокирование происходит при связывании одного из  $\beta$ -галактозидов с той или иной субъединицей репрессора, что уменьшает сродство последнего к оператору. После отсоединения репрессора от промотора полимераза может связаться с промотором и инициировать транскрипцию оперона.

Очень важно сохранение нуклеотидной последовательности домена *lac*-оператора, связывающего репрессор. Об этом свидетельствуют результаты исследования эффективности репрессии *in vivo* у мутантов с модифицированными репрессорами или операторами (рис. 3.70). Интересные данные были получены и при изучении связывания очищенных репрессорных белков с операторами дикого типа и мутантными формами *in vivo*. Мутации, уменьшающие сродство репрессора к оператору, приводили к **конститутивному** синтезу ферментов, кодируемых *lac*-опероном, т.е. к экспрессии *lac*-ферментов в отсутствие индуктора. Мутации, сопровождающиеся накоплением репрессора в клетках или увеличением сродства репрессора к оператору, делали *lac*-оперон неиндуцибельным.

**Позитивная регуляция.** Для экспрессии *lac*-оперона, как и других индуцибельных оперонов, которые осуществляют контроль синтеза ферментов, участвующих в метаболизме сахаров, необходимо не только снять репрессию оперона, но и получить некий сигнал. Таким сигналом служит комплекс **циклического АМР** (сАМР) с белком-активатором катаболизма (САР, от англ. catabolite activator protein), который связывается со специфической последовательностью, находящейся в самом начале *lac*-промотора (рис. 3.71). сАМР, принимающий участие во многих клеточных процессах, образуется из АТР (рис. 3.72) в ответ на самые разные вне- и внутриклеточные события. САР представляет собой димер из идентичных полипептидных цепей



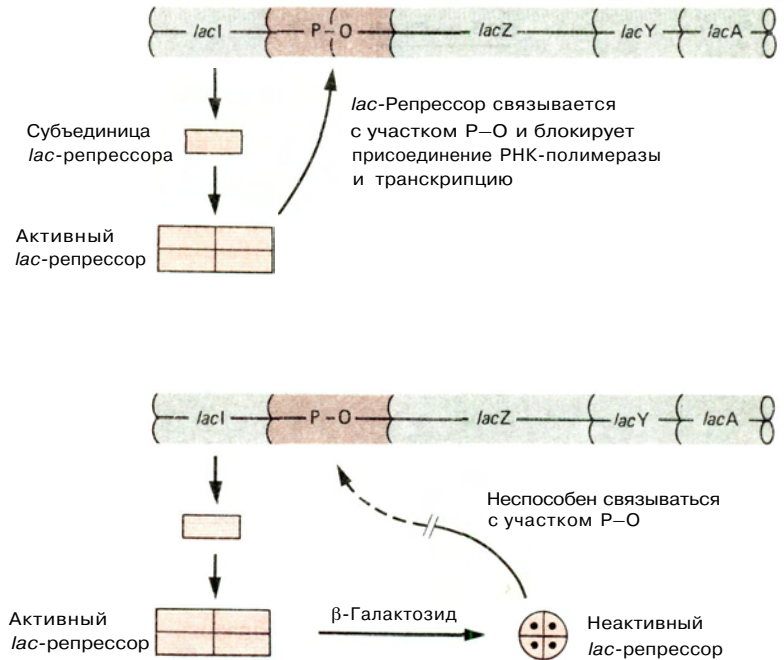


**РИС. 3.68.**

Лактозный оперон *E. coli* (*lac*) и тесно сцепленный с ним ген *lac*-репрессора (*lac I*).

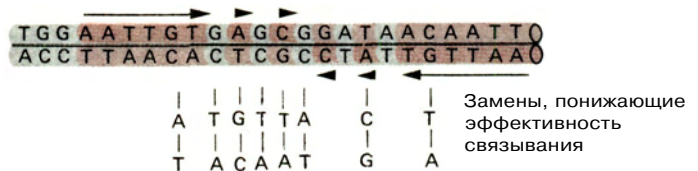
**РИС. 3.69.**

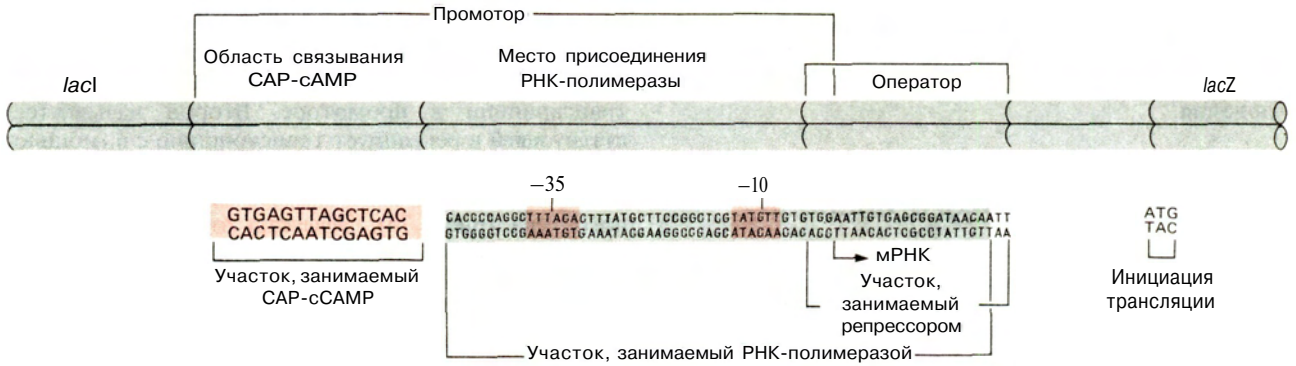
Репрессия *lac*-оперона гомотетрамерным *lac*-репрессором (вверху) и индукция *lac*-оперона после связывания с репрессором  $\beta$ -галактозидного индуктора (внизу).



**РИС. 3.70.**

Нуклеотидная последовательность *lac*-операторного участка, связывающегося с репрессором. Внизу указаны замены пар оснований, понижающие эффективность связывания репрессора с оператором. Стрелками показаны инвертированные повторы. Наличие таких повторов характерно для операторных последовательностей *E. coli*. [J.D. Watson et al., *Molecular Biology of the Gene*, 4 th ed. (Menlo Park, Calif.: Benjamin/Cummings, 1987).]





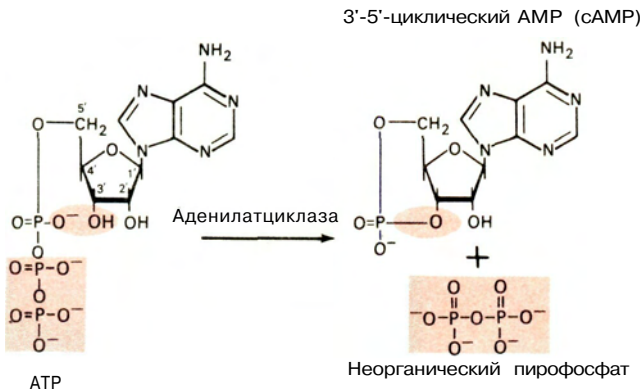
**РИС. 3.71.**

Нуклеотидные последовательности, принимающие участие в регуляции экспрессии *lac*-оперона: область связывания с репрессором, -10- и -35-участки промотора

и область связывания комплекса белка-активатора катаболизма и сAMP (CAP-cAMP).

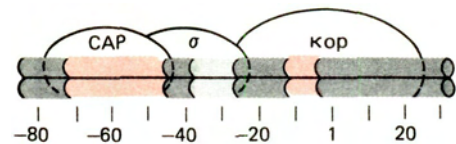
с мол. массой 22 кДа. Связывание комплекса CAP-cAMP со специфической последовательностью в начале промотора приводит к усилению транскрипции *lac*-оперона почти в 50 раз. Сам по себе CAP не способен к такому связыванию и стимуляции транскрипции. Усиление транскрипции с помощью комплекса CAP-cAMP можно объяснить тем, что, связываясь с ДНК в непосредственной близости от сайта присоединения РНК-полимеразы, он усиливает сродство этого фермента к промотору (рис. 3.73). Альтернативная гипотеза заключается в том, что связывание CAP-cAMP с CAP-сайтом предотвращает присоединение РНК-полимеразы к расположенному поблизости слабому промотору и увеличивает тем самым вероятность того, что полимеразы свяжется с «правильным» промоторным сайтом. Высказывавшееся ранее предположение, что комплекс CAP-cAMP изменяет топологию ДНК, так что облегчаются связывание полимеразы и транскрипция, кажется несостоятельным.

Комплекс CAP-cAMP является положительным сигналом при регуляции экспрессии и других оперонов, в частности тех, которые кодируют ферменты расщепления углеводов. Например, для экспрессии *ara*- и *gal*-оперонов должны произойти дерепрессия с помощью индукторов—арабинозы и галактозы соответственно—и связывание комплекса CAP-cAMP с областью промотора. Так, у бактерий, растущих на глюкозе, уровень внутриклеточного сAMP и соответственно комплекса CAP-cAMP очень низок. Поэтому, если даже в среде присутствует арабиноза или галактоза, в клетках не образуются ферменты, необходимые для утилизации этих сахаров. При уменьшении количества глюкозы уровни сAMP и CAP-cAMP увеличиваются и опероны в присутствии необходимых индукторов начинают экспрессироваться. Подобная комбинация позитивной и негативной систем регуляции очень важна, поскольку это предотвращает образование ферментов, потребность в которых в данный момент отсутствует. Сигнальная система CAP-cAMP регулирует также работу оперонов, кодирующих ферменты, которые участвуют в деградации аминокислот, пуринов и пиримидинов. Накопление в клетке CAP-cAMP служит сигналом «голодания»: в ответ на него снижается экспрессия оперонов, кодирующих ферменты расщепления аминокислот, пуринов



**РИС. 3.72.**

Синтез циклического AMP (сAMP) из АТР катализируется аденилатциклазой.



**РИС. 3.73.**

Модель, иллюстрирующая влияние связывания CAP-cAMP со своим участком на присоединение РНК-полимеразы к *lac*-промотору. Предполагается, что CAP и холополимераза взаимодействуют друг с другом.

и пиримидинов. В этом случае логика работы системы также состоит в том, чтобы предотвратить образование ненужных ферментов в период голодания.

### г. Регуляция экспрессии триптофанового оперона

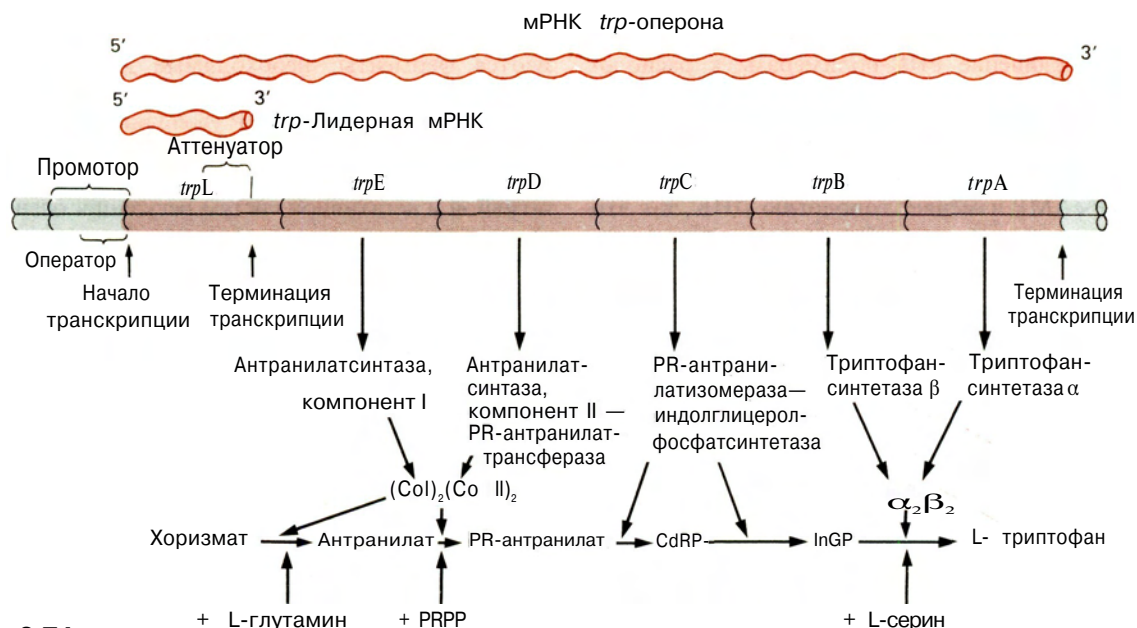
Триптофан синтезируется в *E. coli* из ароматического предшественника—хоризмовой кислоты—в ходе пяти последовательных реакций, катализируемых ферментным комплексом из пяти белков (рис. 3.74). Эти кодируемые *trp*-опероном белки образуются согласованно, примерно в равных количествах, при трансляции полицистронной мРНК длиной 7000 нуклеотидов, транскрибируемой с промотора. При наличии в среде достаточного для роста бактерий количества триптофана клетки *E. coli* образуют ферменты биосинтеза триптофана в очень малых количествах. Однако если клетки лишены экзогенного триптофана, то в них начинается довольно интенсивный синтез всех пяти ферментов. Содержание ферментов этой группы в клетках *E. coli* может различаться до 700 раз в зависимости от внутриклеточного уровня триптофана.

Подобный разброс в уровне экспрессии обусловлен наличием двух практически независимых механизмов регуляции, каждый из которых «подгоняет»

уровень синтеза *trp*-мРНК к концентрации внутриклеточного триптофана. Один механизм основан на репрессии, изменяющей эффективность инициации транскрипции в промоторе. Второй называется **аттенуацией** и регулирует транскрипцию с помощью сигнала терминации транскрипции, расположенного между промотором и началом первого структурного гена (рис. 3.74). Эффекты этих двух регуляторных механизмов комплементарны, мультипликативны и позволяют увеличивать степень экспрессии оперона. Мы обсудим в этом разделе оба способа регуляции и их вклад в гармоничный процесс биосинтеза триптофана.

**Репрессия триптофанового оперона.** Транскрипция *trp*-оперона блокируется, когда репрессор связывается с последовательностью *trp*-оператора. Ген *trp R*, кодирующий репрессорный белок мол. массой 58000 Да, находится далеко от *trp*-оперона. Для того чтобы *trp R*-белок мог связаться с оператором и действовать как репрессор, он должен образовать комплекс с триптофаном. Поскольку уровень экспрессии *trp R* очень низок и не зависит от триптофана, концентрация активного репрессора отражает концентрацию внутриклеточного триптофана.

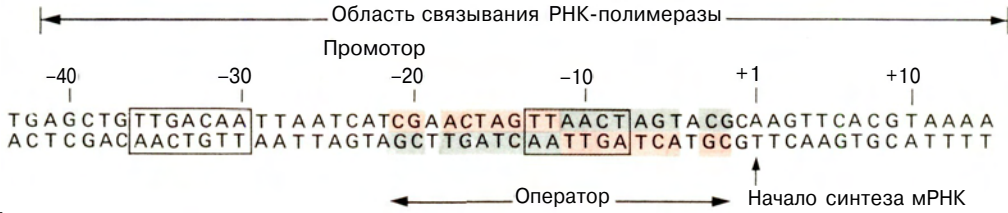
Нуклеотидные последовательности *trp*-оператора и промотора перекрываются (рис. 3.75), поэтому связывание комплекса репрессор-триптофан с оператором препятствует правильному взаимодействию



**РИС. 3.74.**

Триптофановый оперон *E. coli* (*trp*-оперон). Указаны генные продукты и катализируемые ими реакции по превращению хоризмовой кислоты в триптофан. Пять генных продуктов дают четыре фермента. Представлены лидерная последовательность *trp L* и ее транскрипт

(*trp*-лидерная РНК), а также аттенуатор и полицистронная мРНК *trp*-оперона. Сокращения: PR - фосфорибозил, PRPP - фосфорибозилпирофосфат, CdRP - карбоксифениламинодезоксирибозулофосфат и InGP - индол-глицерофосфат. (С любезного разрешения С. Yanofsky.)



**РИС. 3.75.**

Перекрывающиеся нуклеотидные последовательности промотора и оператора *trp*-оперона *E. coli*. Связывание комплекса *trp*-репрессор–триптофан с оператором

препятствует присоединению РНК-полимеразы к промотору.

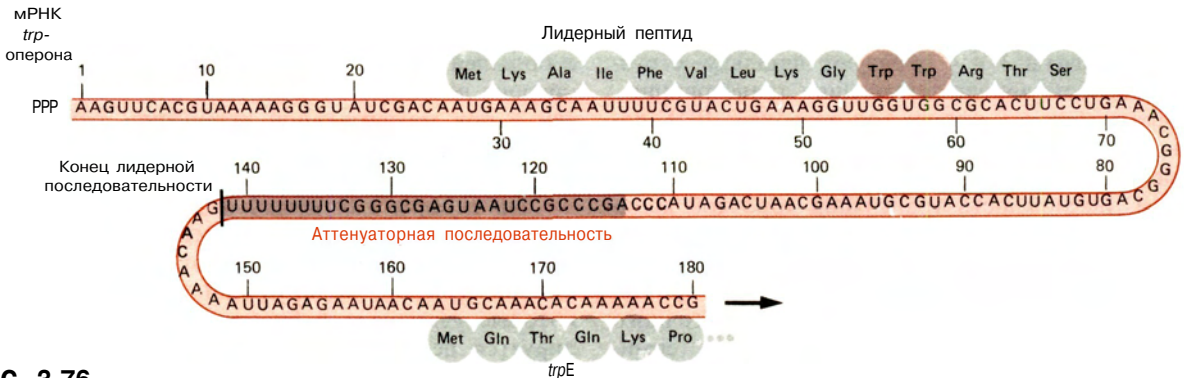
вию РНК-полимеразы с промотором. При этих условиях транскрипция оперона не осуществляется и ферменты биосинтеза не образуются. В отсутствие триптофана не происходит и связывания репрессора с оператором, что позволяет РНК-полимеразе без помех инициировать транскрипцию на промоторе и синтезировать мРНК.

**Аттенуация экспрессии триптофанового оперона.** Регуляция работы триптофанового оперона с помощью аттенуации осуществляется при участии последовательности, находящейся на расстоянии примерно 100–140 пар оснований от начала транскрипции (рис. 3.74). Этот так называемый лидерный сегмент, *trp L*, содержит **аттенуаторную** последовательность, которая вынуждает РНК-полимеразу прервать транскрипцию как раз перед началом *trp E*-гена с отделением фрагмента РНК длиной 141 нуклеотид – *trp*-лидерной мРНК. Подобная терминатия происходит при уровнях триптофана в клетках от среднего до высокого. При низком уровне триптофана терминатия транскрипции в пределах аттенуаторной последовательности в значительной мере блокируется, и транскрипция проходит через терминатор с образованием полноразмерной мРНК *trp*-оперона.

Каков механизм регуляции терминатии транскрипции в аттенуаторе с помощью изменения содер-

жания триптофана? Ответ на этот вопрос был получен после того, как определили нуклеотидную последовательность аттенуатора и установили, что на аттенуацию влияет трансляция, идущая одновременно с транскрипцией. Участок из первых 145 нуклеотидов транскрипта *trp*-оперона содержит три необычных по нуклеотидной последовательности сегмента (рис. 3.76). Один из них кодирует короткий полипептид из 14 аминокислот, две из которых являются tandemными триптофановыми остатками. Два других содержат инвертированные повторы, из-за чего происходит внутримолекулярное спаривание оснований РНК-транскрипта и образование двух альтернативных шпильчатых структур (рис. 3.77). Одна из этих структур, в которой спарены сегменты 1 и 2, 3 и 4 соответственно, опосредствует  $\rho$ -независимую терминатию транскрипции на участке из нескольких расположенных подряд за четвертым сегментом остатков урацила (разд. 3.2.г). Другая шпильчатая структура, в которой сегменты 2 и 3 спарены, а 1 и 4 остаются одноцепочечными, разрешает транскрипции дойти до конца оперона.

Какая именно шпильчатая структура образуется при транскрипции аттенуаторного участка – зависит от того, транслируется ли рибосомами полная последовательность, кодирующая полипептид длиной 14 аминокислот. Если синтез лидерного полипеп-



**РИС. 3.76.**

Лидерная *trp*-мРНК кодирует короткий пептид и содержит аттенуаторную последовательность. В лидерном

полипептиде имеются два расположенных рядом триптофановых остатка.

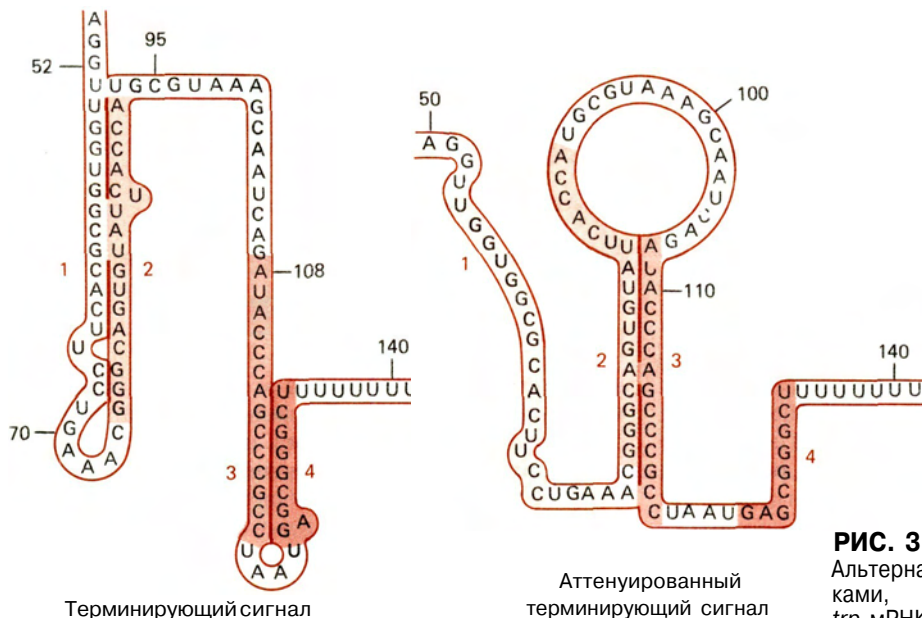


Рис. 3.77.

Альтернативные структуры со шпильками, которые образует лидерная *trp*-мРНК. Нумерация нуклеотидных остатков такая же, как на рис. 3.76.

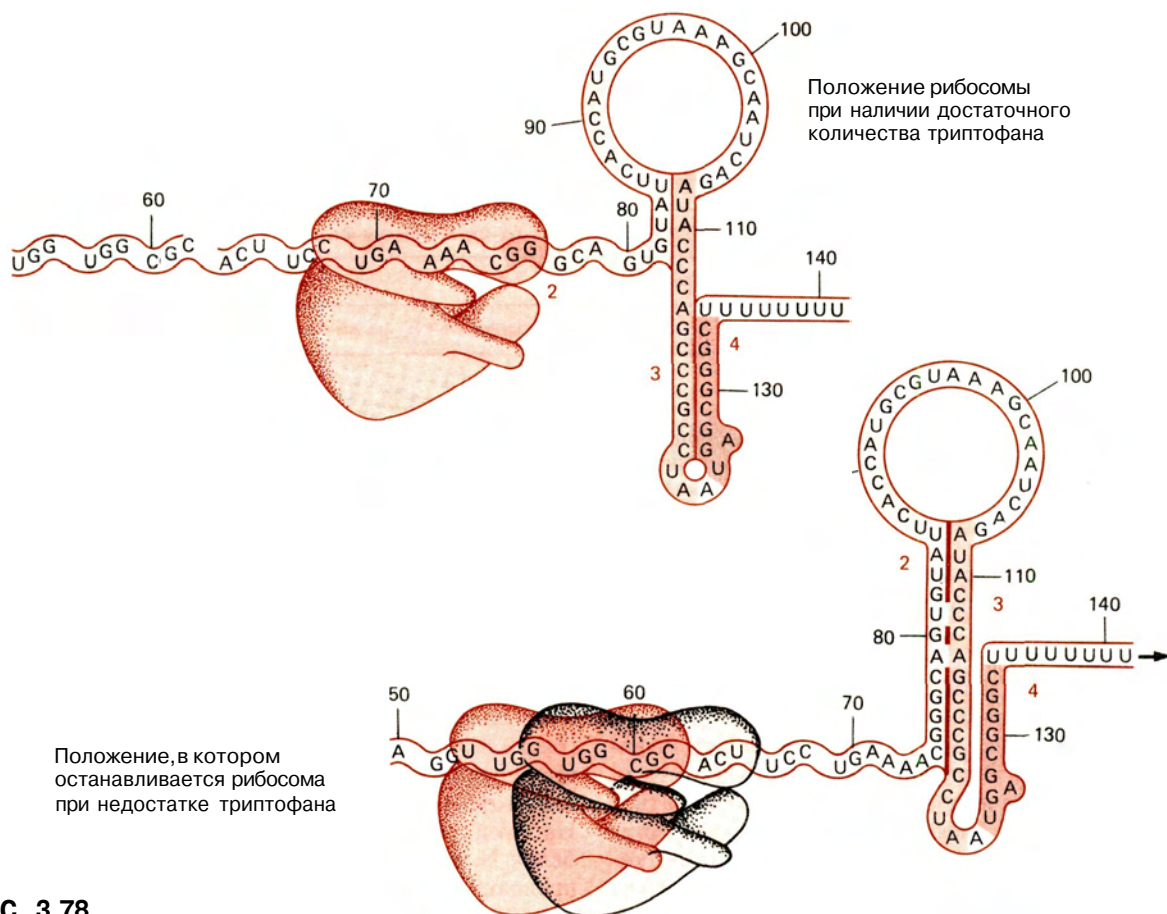
тида инициируется в кодоне AUG (рис. 3.76), то рибосомы транслируют следующие 13 кодонов при условии, что имеются все необходимые аминоксил-тРНК. При недостатке триптофанил-тРНК рибосомы останавливаются, когда доходят до tandemных триптофановых кодонов. Если рибосомы останавливаются в одном из триптофановых кодонов, расположенных в аттенуаторном сегменте 1, то последовательности сегментов 2 и 3 свободно спариваются и образуют шпилечную структуру, а сегмент 4 остается одноцепочечным (рис. 3.78). Эта структура не препятствует продолжению транскрипции с образованием полноразмерной *trp*-мРНК длиной 7000 нуклеотидов. Если триптофанил-тРНК имеется в количестве, достаточном для трансляции tandemных триптофановых кодонов, то рибосомы движутся до терминального кодона лидерного пептида и, поскольку сегмент 2 ассоциирован с рибосомой, сегменты 3 и 4 образуют шпильку—структуру, способствующую терминации транскрипции.

Таким образом, аттенуатор позволяет РНК-полимеразе опосредованно «почувствовать» концентрацию триптофанил-тРНК через расположение рибосомы. Возможность трансляции короткой кодирующей последовательности в лидерном сегменте определяет вторичную структуру транскрипта и, следовательно, выбор между терминацией и прочтением последовательности. Наличие достаточного количества триптофанил-тРНК стимулирует терминацию транскрипции, при ее недостатке транскрипция проходит весь оперон.

*Суммарный эффект аттенуации и репрессии.* Ат-

тенуация и репрессия вместе содействуют оптимальной экспрессии *trp*-оперона. При избытке триптофана благодаря репрессии блокируется инициация синтеза мРНК. Снижение концентрации триптофана приводит к уменьшению содержания функционального комплекса репрессор—триптофан, и транскрипция возобновляется. Однако при концентрации триптофана, допускающей инициацию транскрипции, количества триптофанил-тРНК еще вполне достаточно для того, чтобы в пределах лидерного участка в основном происходила терминация транскрипции; следовательно, в таких условиях уровень экспрессии оперона еще очень низок. По мере снижения концентрации триптофана репрессия полностью снимается: если концентрация триптофанил-тРНК становится ниже уровня, необходимого для продолжения синтеза, то терминация аттенуируется, а *trp*-мРНК и ферменты начинают синтезироваться с большей скоростью. Таким образом, экспрессия оперона достигает максимума в отсутствие репрессии и при максимальной аттенуации терминации; экспрессия минимальна при почти полной репрессии оперона и отсутствии аттенуации терминации.

*Аттенуация как общий механизм.* Аттенуация используется при регуляции экспрессии многих генов и оперонов как у *E. coli*, так и у других организмов. Оперон, кодирующий девять ферментов, которые участвуют в биосинтезе гистидина, регулируется исключительно с помощью аттенуации; аттенуатором служит полипептид-кодирующая последовательность с шестью tandemными гистидиновыми кодонами (рис. 3.79). У оперонов, функцио-

**РИС. 3.78.**

Образование разных шпильчатых структур лидерной *trp*-мРНК. Положение рибосомы определяет, какая именно из двух альтернативных шпилек образуется,

нирование которых зависит от содержания определенных аминокислот—треонина и изолейцина (или лейцина), валина и изолейцина,— лидерный участок буквально «напичкан» кодонами для этих аминокислот. Антитерминация, т.е. ослабленная терминация, не требует сопряжения трансляции и транскрипции. Она может осуществляться некоторыми белками, допускающими экспрессию особых кластеров генов только после того, как синтезированы белки, способствующие аттенуации (разд. 3.11.д).

#### д. Временная регуляция генной экспрессии в жизненном цикле бактериофага $\lambda$

У бактериофага  $\lambda$ , есть два альтернативных способа существования. При литическом пути все

и тем самым - сформируется ли структура, способствующая терминации транскрипции.

вирусные гены экспрессируются в определенной временной последовательности, в результате чего образуется примерно сотня фаговых частиц и происходит лизис инфицированной бактерии. Развитие вируса, вызвавшего инфекцию, может пойти по лизогенному пути, при котором его ДНК встраивается в ДНК клетки-хозяина в специфическом месте (разд. 2.4.г). Интегрированная форма вирусного генома называется **профагом**. В лизогенных клетках профаговая ДНК многократно реплицируется при помощи клеточного аппарата так, как будто она является частью клеточного генома. При этом, однако, все фаговые гены, кроме одного, выключены. Лизогенные клетки переключаются на литический путь либо спонтанно с низкой частотой, либо под действием различных химических или физических факторов. Независимо от природы индукции литического цикла «молчащие» ранее фаговые гены включаются, запуская процессы, приводящие к образованию инфекционных фаговых частиц.

Оперон	Лидерная пептидная последовательность	Аминокислоты-регуляторы
<i>his</i>	Met-Thr-Arg-Val-Gln-Phe-Lys-His-His-His-His-His-His-His-Pro-Asp	His
<i>pheA</i>	Met-Lys-His-Ile-Pro-Phe-Phe-Phe-Ala-Phe-Phe-Phe-Thr-Phe-Pro	Phe
<i>leu</i>	Met-Ser-His-Ile-Val-Arg-Phe-Thr-Gly-Leu-Leu-Leu-Leu-Asn-Ala-Phe-Ile-Val-Arg-Gly-Arg-Pro-Val-Gly-Gly-Ile-Gln-His	Leu
<i>thr</i>	Met-Lys-Arg-Ile-Ser-Thr-Thr-Ile-Thr-Thr-Thr-Ile-Thr-Ile-Thr-Thr-Gly-Asn-Gly-Ala-Gly	Thr, Ile
<i>ilv</i>	Met-Thr-Ala-Leu-Leu-Arg-Val-Ile-Ser-Leu-Val-Val-Ile-Ser-Val-Val-Val-Ile-Ile-Ile-Pro-Pro-Cys-Gly-Ala-Ala-Leu-Gly-Arg-Gly-Lys-Ala	Leu, Val, Ile

**РИС. 3.79.**

Некоторые опероны, лидерные мРНК которых содержат многократно повторенные кодоны для аминокислот, которые регулируют работу данного оперона.

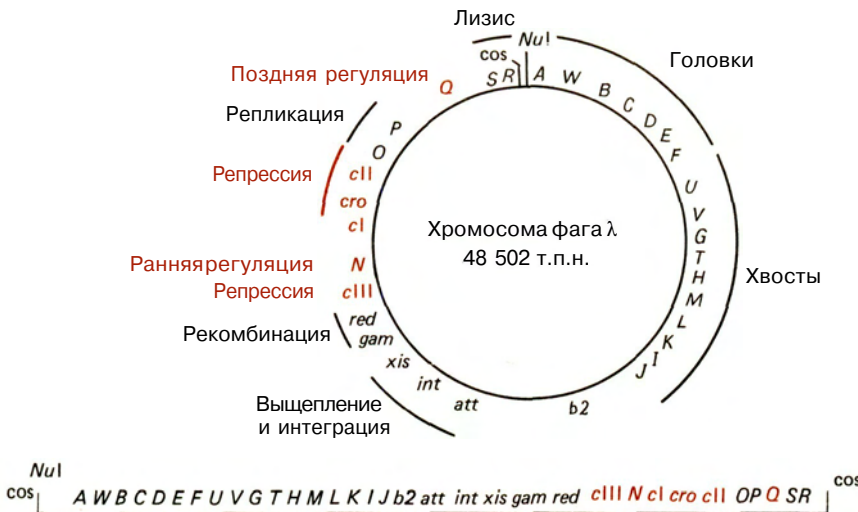
Каким образом осуществляется временная регуляция экспрессии вирусных генов, обеспечивающая строго определенный порядок репликации вирусной ДНК, образования вирусных белков и сборку частиц бактериофага при литической инфекции? Какие регуляторные механизмы определяют выбор литического или лизогенного пути развития для инфицирующих фагов? Эти вопросы заслуживают того, чтобы обсудить их подробно. Пытаясь разобраться в них, мы увидим, насколько элегантно осуществляется взаимодействие различных механизмов, которые используют прокариотические организмы для регуляции своих фенотипов.

**Литический путь.** Определим сначала порядок экспрессии генов при литической инфекции. На рис. 3.80 показано расположение основных структурных и регуляторных генов на кольцевой и линейной картах фагового генома. Гены, кодирующие

структурные белки (головки и хвосты), сконцентрированы в одной области ДНК; гены, кодирующие ферментативные (репликацию и рекомбинацию) и регуляторные (репрессию и антитерминацию) функции, сосредоточены в другой области генома.

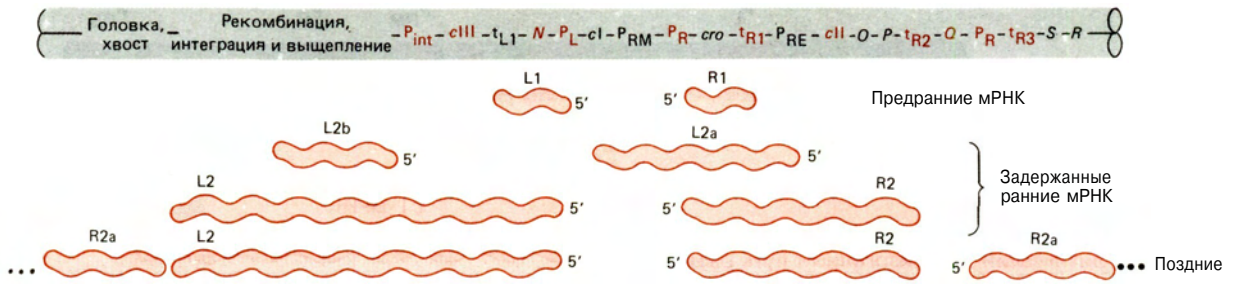
После инфекции фаговая ДНК замыкается в кольцо путем соединения липких концов (разд. 2.4.г). Затем РНК-полимераза *E. coli* транскрибирует те фаговые мРНК, которые кодируют белки, необходимые на самых ранних этапах жизненного цикла, — так называемые предранние мРНК (рис. 3.81). Одна из этих предранных мРНК транскрибируется справа налево с промотора  $P_L$ , а терминатором служит последовательность  $t_{L1}$ . На этой мРНК ( $L1$ ) синтезируется регуляторный белок N, работающий как антитерминатор. Другая предранняя мРНК транскрибируется слева направо с промотора  $P_R$  к терминатору  $t_{R1}$  и кодирует только белок Cro. Позднее мы обсудим роль этого белка в выборе литического, а не лизогенного пути развития фага.

По мере накопления белка N происходит аттенуация терминации в  $t_{L1}$  и  $t_{R1}$ , РНК-полимераза



**РИС. 3.80.**

Линейная и кольцевая формы генома бактериофага  $\lambda$  с указанием положения генов. Кольцевая форма образуется при отжиге линейной ДНК и лигировании одноцепочечных липких концов. Эти концы и места их соединения обозначены *cos*.

**РИС. 3.81.**

Очередность экспрессии мРНК фага  $\lambda$  при литическом пути развития.

продолжает транскрипцию через эти сайты с образованием мРНК второго типа – задержанных ранних транскриптов. Более крупный транскрипт, начинающийся с промотора  $P_L$  ( $L2$  на рис. 3.81), кодирует ферменты, участвующие в рекомбинации, и (что особенно важно) ферменты, катализирующие встраивание ДНК фага  $\lambda$  в ДНК клетки-хозяина. Задержанный ранний транскрипт, начинающийся с промотора  $P_R$  (транскрипт  $R2$ ), кодирует ферменты, ответственные за репликацию фаговой ДНК (белки  $O$  и  $P$ ), и еще один регуляторный белок  $Q$ . Белок  $Q$  вызывает аттенуацию терминации транскрипции в терминаторном сайте ( $t_{R3}$ ), расположенном сразу за промотором  $P_R$ . При транскрипции с промотора  $P_R$  транскрибируются гены ( $S$  и  $R$ ), ответственные за включение лизиса клеток. Кроме того, поскольку ДНК фага  $\lambda$  сразу после инфекции замыкается в кольцо (рис. 3.80),  $S$ - и  $R$ -гены оказываются рядом с генами, кодирующими белки головки и хвоста фага. Следовательно, в результате транскрипции, инициированной в  $P_R$  и продолженной через  $t_{R3}$ , образуется мРНК, кодирующая белки лизиса и все структурные вирусные белки. Итак, мы рассмотрели образование аппарата, необходимого для литической инфекции: ферментов репликации вирусной ДНК и вирусных белков, участвующих в формировании зрелых фаговых частиц.

**Лизогенный путь.** Как происходит переключение всей серии событий, настроенных на литический цикл развития, на путь лизогении? Для того чтобы понять это, нужно усложнить схему строения промоторов  $P_L$  и  $P_R$  и ввести несколько дополнительных генов, генных продуктов и промоторов. На самом деле  $P_L$  и  $P_R$  являются частью сложной регуляторной области, в которой промоторы перемежаются с операторными последовательностями  $O_L$  и  $O_R$  соответственно (рис. 3.82).  $O_L$  и  $O_R$  – это сайты связывания двух регуляторных белков: репрессора  $cI$  и антирепрессора  $Cro$ . Как мы уже говорили, белок  $Cro$  – это продукт трансляции предранней мРНК, транскрибируемой вправо с про-

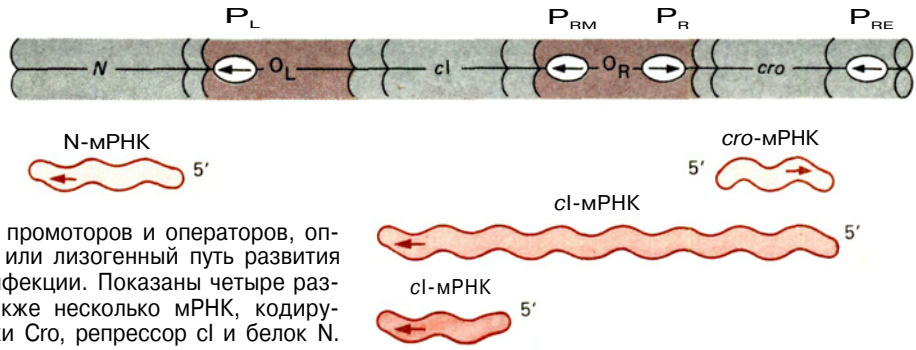
мотора  $P_R$ . Репрессор кодируется геном  $cI$ , локализованным между  $P_R$  и  $P_L/O_L$  (рис. 3.82).  $cI$ -мРНК транскрибируется в направлении влево от промотора  $P_{RE}$ , находящегося справа от  $Cro$ -гена.  $P_{RE}$  сам активируется двумя «положительными» регуляторными белками,  $cII$  и  $cIII$ , которые синтезируются после того, как благодаря действию белка  $N$  образуются транскрипты мРНК, инициированные в  $P_R$  и  $P_L$ , соответственно.

По мере накопления репрессора  $cI$  происходит его связывание с левым и правым операторными сайтами, в результате чего подавляется экспрессия всех генов, транскрибирующихся с  $P_R$  и  $P_L$ . В этом случае предпочтительным оказывается лизогенный путь, поскольку блокируется образование ферментов репликации и структурных вирусных белков, а небольшое количество интеграционного фермента  $Int$ , синтезированного с задержанной ранней мРНК, успевает катализировать рекомбинацию между фаговой и бактериальной ДНК до момента полной репрессии фаговых генов. После интеграции состояние лизогении поддерживается благодаря образованию в небольших количествах репрессора  $cI$  на мРНК, транскрибируемой с промотора  $P_{RM}$ , который находится непосредственно справа от структурного  $cI$ -гена.

Но может произойти переключение и на литический путь, что зависит от антирепрессорных и активаторных свойств  $Cro$ .  $Cro$  также связывается с  $O_R$  и  $O_L$ , предотвращая тем самым связывание с ними  $cI$  и сводя на нет репрессорное действие последнего. По мере накопления  $Cro$  синтез репрессора  $cI$  все больше блокируется и накапливаются белки репликации, лизиса и структурные фаговые белки.

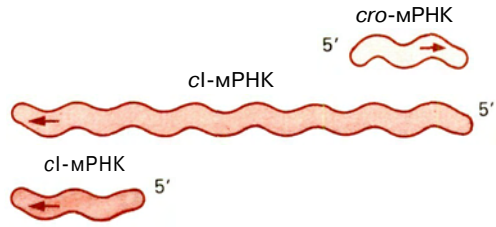
**Связывание  $Cro$  и  $cI$  с  $O_R$  и  $O_L$ .** Чтобы понять механизм очень тонкого антагонистического действия белков  $Cro$  и  $cI$ , нам нужно выяснить природу и расположение сайтов связывания, находящихся в пределах операторов  $O_R$  и  $O_L$ .  $O_R$  и  $O_L$  каждый состоит из тандемного кластера трех субоператоров:  $O_{L1}$ ,  $O_{L2}$ ,  $O_{L3}$  и  $O_{R1}$ ,  $O_{R2}$  и  $O_{R3}$  соответственно.  $O_L$  перекрывается с  $P_L$ ,  $O_R$  – с  $P_{R1}$  и  $O_{R3}$  – с  $P_{RM}$  (рис. 3.83). И  $Cro$ , и  $cI$  связываются со всеми субоператорными сегментами, но относительное средство





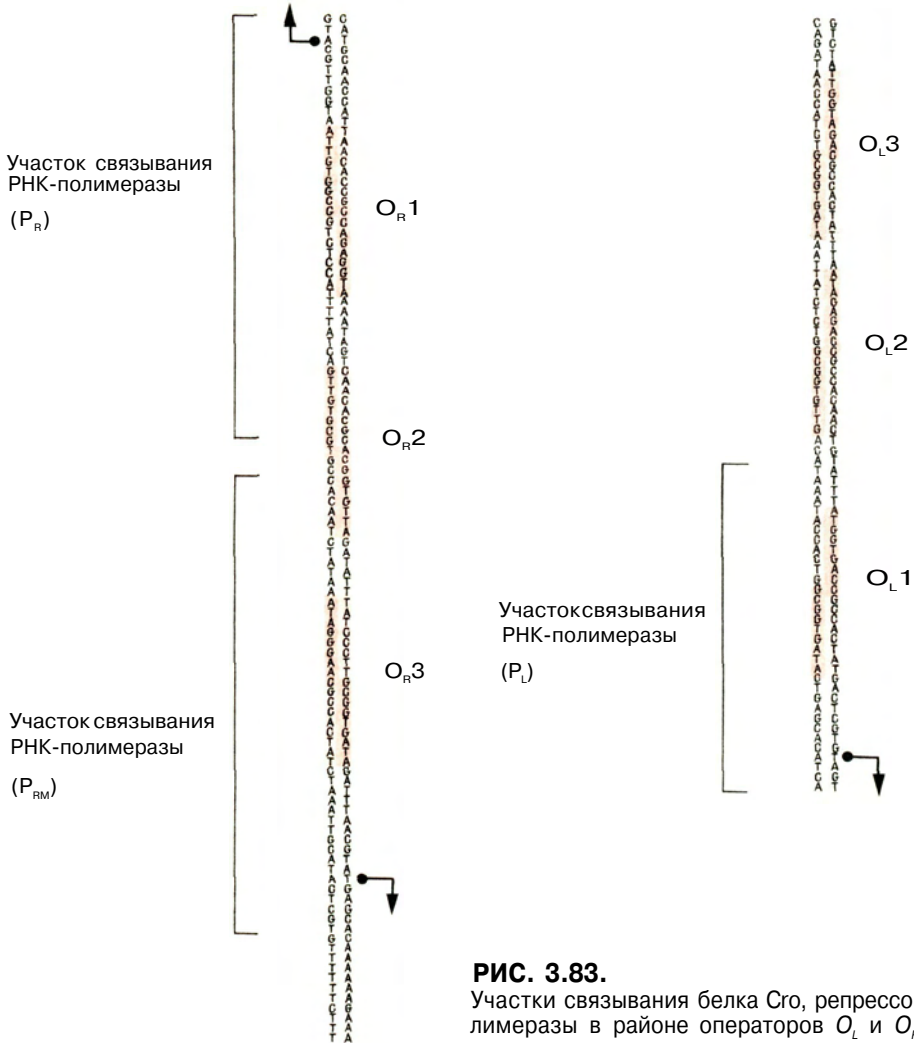
**РИС. 3.82.**

Взаимное расположение промоторов и операторов, определяющих литический или лизогенный путь развития бактериофага  $\lambda$  после инфекции. Показаны четыре различных промотора, а также несколько мРНК, кодирующих регуляторные белки Cro, репрессор *cI* и белок *N*.



их к этим сегментам различно. Так, Cro наиболее прочно связывается с  $O_{R3}$  и значительно слабее — с  $O_{R2}$  и  $O_{R1}$ . Связывание белка Cro с  $O_L$  уменьшается последовательно от  $O_{L3}$  к  $O_{L2}$  и далее к  $O_{L1}$ . Белок Cro накапливается уже на ранних стадиях

инфекции, поэтому он связывается с  $O_{R3}$ , предотвращая транскрипцию *cI*-гена с  $P_{RM}$ . Поскольку  $O_{R1}$  и  $O_{L1}$  остаются свободными до тех пор, пока концентрация белка Cro не достигнет достаточно высокого уровня, транскрипция с  $P_R$  и  $P_L$  не подавляется.



**РИС. 3.83.**

Участки связывания белка Cro, репрессора *cI* и РНК-полимеразы в районе операторов  $O_L$  и  $O_R$ .

Напротив, связывание белка *Сго* с  $O_{R3}$  и  $O_{L3}$  стимулирует транскрипцию с  $P_R$  и  $P_L$  соответственно. Белок *сI* также обладает разным сродством к субоператорам. Наиболее прочно он связывается с субоператорами  $O_{R1}$  и  $O_{L1}$ , поэтому транскрипция с  $P_R$  и  $P_L$  блокируется при низких его концентрациях. В этом случае связывание *сI*-репрессора с  $O_{R1}$  также стимулирует транскрипцию *сI*-гена с  $P_{RM}$ .

Таким образом, выбор между литическим и лизогенным путями развития зависит от относительной скорости накопления четырех регуляторных белков: репрессорного белка *сI*, антирепрессорного белка *Сго* и «положительных» регуляторных белков *сII* и *сIII*. Если доминирует антирепрессорная функция *Сго*-белка, то предпочтение отдается литическому пути. Если не удается установиться репрессия, то функции, необходимые для реализации литического пути, блокируются и происходит лизогенизация. Поскольку равновесие между этими белками очень неустойчивое, исход зависит от состояния бактерии-хозяина, от культуральной среды и от многих других факторов.

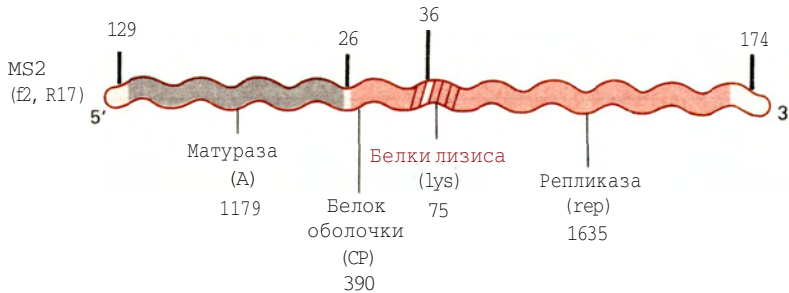
### е. Трансляционная регуляция экспрессии некоторых генных продуктов

Как мы уже говорили, экспрессия прокариотических генов, кодирующих РНК (например, рРНК, тРНК), регулируется на двух этапах—на уровне транскрипции и на уровне посттранскрипционных модификаций и превращения предшественников РНК в зрелые формы. Экспрессия некоторых генов, кодирующих белки, также регулируется двумя путями. Первый из них состоит в регуляции содержания мРНК (разд. 3.11.в-д), второй основан на репрессии трансляции. Рассмотрим два примера.

1. Геном каждого из близкородственных бактериофагов *E. coli*—MS2, R17 и Q $\beta$  — представлен единственной молекулой РНК, функционирующей

и в качестве генома, и в качестве мРНК, на которой синтезируются белки, необходимые для размножения вирусов. Геном бактериофага MS2, одного из представителей этого класса вирусов, состоит из четырех генов (рис. 3.84). Один из них, *rep*, кодирует субъединицу комплекса РНК-репликазы, которая амплифицирует фаговую РНК для того, чтобы, во-первых, обеспечить высокий уровень трансляции фаговых генов и, во-вторых, снабдить геномами фаговое потомство. Второй ген кодирует белок оболочки (CP), 180 копий которого образуют вирусный икосаэдрический капсид. Третий ген кодирует белок А, или матуразу, который является компонентом зрелой вирусной частицы и необходим для адсорбции и проникновения вируса в клетку. Четвертый ген детерминирует белок (*Lys*), обеспечивающий лизис инфицированных клеток и высвобождение вирусного потомства. Последовательность, кодирующая белок *Lys*, начинается на 3'-конце гена *CP* и продолжается до 5'-конца гена *rep*. Однако рамка считывания для белка *Lys* сдвинута на один нуклеотид относительно кодирующих последовательностей *CP* и *Rep*.

При нормальной инфекции *CP* присутствует в клетке в больших количествах, *Rep* и *Lys*—в значительно меньших, а А представлен только несколькими молекулами. Дифференциальная экспрессия вирусных генов осуществляется благодаря тому, что для рибосом доступен только инициаторный кодон AUG гена *CP*. Соответствующие последовательности генов А, *Lys* и *rep* находятся в двухцепочечных участках молекулы РНК (рис. 3.41). Рибосомы редко транслируют ген А, а если быть точными, то только один раз—в момент синтеза 5'-конца РНК. Напротив, ген *CP* транслируется рибосомами многократно. При трансляции *CP*-кодирующей последовательности рибосомы разворачивают РНК и экспонируют инициаторный кодон AUG гена *rep*, обеспечивая его трансляцию. Для того чтобы произошла трансляция *Lys*-кодирующей последовательности, рибосома иногда «неестественно» тран-



**РИС. 3.84.**

Расположение четырех генов РНК-генома бактериофага MS2 (3569 нуклеотидов). Длины 5'-лидерной последовательности, четырех генов и 3'-трейлерной последовательности, а также длины спейсеров между генами

А и *CP* и генами *CP* и *rep* выражены в числе нуклеотидов. Спейсер между генами *CP* и *rep*—это часть кодирующей области гена *Lys*. (См. рис. 3.41.)

слюцируется (меняет место) при трансляции последовательности гена *CP*, в результате чего вскоре происходит терминация трансляции. Это приводит к возобновлению трансляции в «неположенной» инициаторной последовательности в пределах гена *lys*.

По мере увеличения концентрации *CP* происходит его связывание с последовательностью, перекрывающей последовательность инициации трансляции гена *rep*. Это предотвращает синтез репликазы, но не препятствует образованию белков оболочки. Итак, трансляция гена *rep* сначала регулируется с помощью структурных ограничений, благодаря которым уменьшается доступность инициаторной последовательности для рибосом, и случайного сползания рибосом во время трансляции, а затем путем репрессии трансляции, обусловленной связыванием белка оболочки с его инициаторной последовательностью.

2. Синтез белков, составляющих собственно аппарат трансляции, регулируется на уровне трансляции. Гены, кодирующие белки больших (L) и малых (S) субчастиц рибосомы и некоторых белков, участвующих в процессе трансляции (в том числе EF-Tu и EF-G), рассеяны по нескольким оперонам (рис. 3.67). Это позволяет координированно регулировать синтез тех генных продуктов, которые должны функционировать согласованно. Экспрессия таких генов как на уровне транскрипции, так и на уровне трансляции происходит координированно. Как мы увидим далее, синтез рибосомных белков частично регулируется также путем изменения содержания трех рРНК и кинетических параметров процесса сборки рибосом.

Контроль трансляции некоторых оперонов рибосомных белков осуществляется по одинаковому механизму (табл. 3.7). Один из рибосомных белков, кодируемый полицистронной мРНК, связывается со специфической последовательностью, локализован-

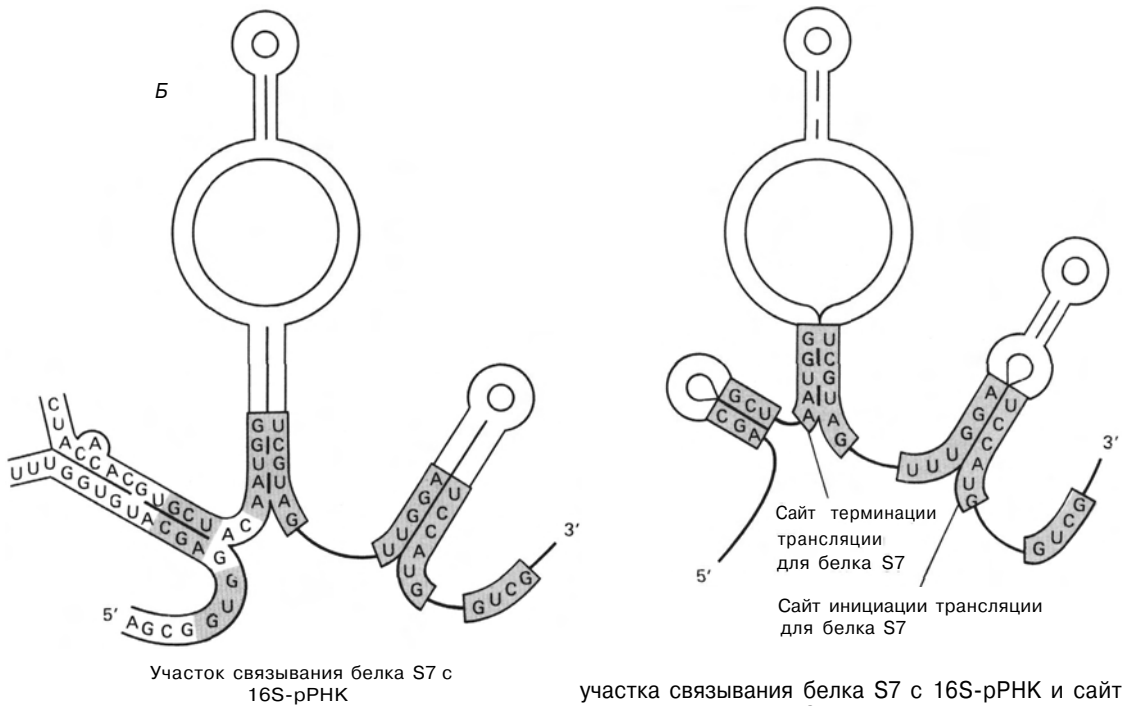
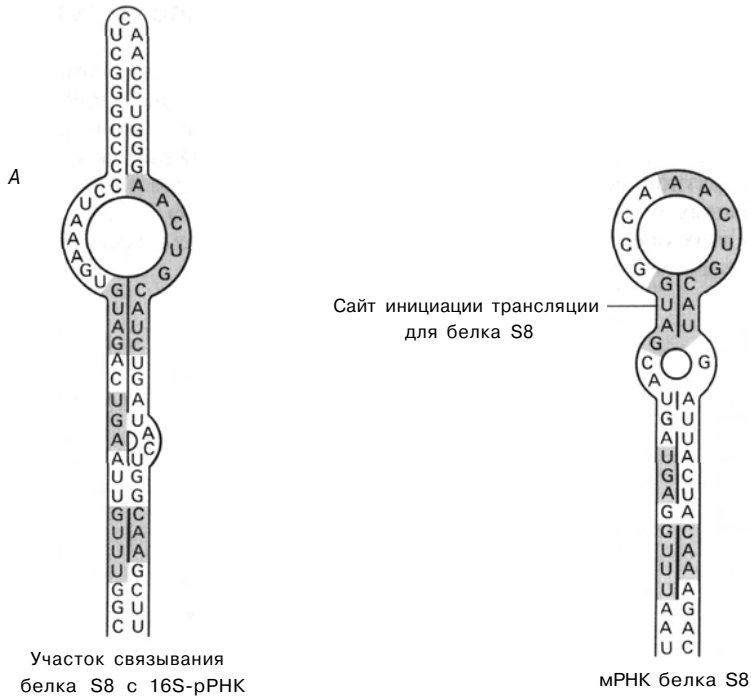
ной либо на 5'-конце мРНК, либо в начале одной из кодирующих последовательностей в середине мРНК. В обоих случаях это блокирует доступ рибосом к ближайшей инициаторной трансляционной последовательности. В зависимости от места нахождения сайта инициации трансляции мРНК—вблизи 5'-конца кодирующей последовательности или в одном из внутренних участков—блокируется трансляция всей мРНК или ее части. Контроль по типу обратной связи, при котором продукт регулирует экспрессию собственного гена, называется **аутогенной регуляцией**. Регуляция может осуществляться на уровне транскрипции (например, репрессорный белок *cI* фага  $\lambda$  регулирует транскрипцию соответствующего гена с  $P_{RM}$ ) и на уровне трансляции, как в приведенном примере.

При сборке рибосом некоторые из рибосомных белков связываются с рРНК. Это наводит на мысль, что, возможно, свободная рРНК регулирует синтез рибосомных белков. Как это может происходить? При наличии достаточного количества рРНК вновь синтезированные рибосомные белки ассоциируют с ней, чтобы инициировать сборку рибосом. При недостатке же рРНК накапливающиеся рибосомные белки связываются с собственной мРНК вместо рРНК и соответственно блокируют собственный синтез и синтез других родственных рибосомных белков. В результате предотвращается накопление свободных рибосомных белков. Таким образом, некоторые ключевые рибосомные белки—это репрессоры, блокирующие трансляцию кодирующей их мРНК. Одновременно они блокируют синтез и других белков, кодируемых той же мРНК. Способность рибосомных белков узнавать как рРНК, так и свою собственную мРНК связана с тем, что обе эти РНК обладают сходными нуклеотидными последовательностями (рис. 3.85). Так, последовательности, в которых рибосомные белки S8 и S7 связываются с 16S-РНК и своими собственными

Таблица 3.7. Порядок расположения генов, кодирующих рибосомные белки, факторы белкового синтеза и субъединицы РНК-полимеразы. Гены этих белков распределены по мРНК, транскрибируемым с различных оперонов<sup>1)</sup>

Оперон	Белки, кодируемые полицистронной мРНК (начиная с 5'-конца мРНК)	Белок-регулятор	Инициаторная последовательность, с которой связывается регулятор
<i>str</i>	S12 S7 EF-G EF-Tu	S7	S7
<i>spc</i>	L14 L24 L5 S14 S8 L6 L18 S5 L30 L15 L30	S8	L5
<i>S10</i>	S10 L3 L2 L4 L23 S19 L22 S3 S17 L16 L29	L4	S10
$\alpha$	S13 S11 S4 $\alpha$ L17	S4	S13
<i>L11</i>	L11 L1	L1	L11
<i>rif</i>	L10 L7/12 $\beta$ $\beta'$	L10	L10

<sup>1)</sup> мРНК представлены в виде набора кодируемых ими белков. Первый белок соответствует 5'-концу мРНК. Белки, гены которых регулируются координированно, подчеркнуты сплошной линией, а соответствующие белки-регуляторы указаны в 3-й колонке.



**РИС. 3.85.**

Иллюстрация положения, состоящего в том, что связывание одного из рибосомных белков с мРНК может подавлять трансляцию этой мРНК. А. Сходство первичной и вторичной структур S8-мРНК (*spc*-оперон полицистронной мРНК) и участка связывания S8 с 16S-рРНК. Б. Сходство первичной и вторичной структур

участка связывания белка S7 с 16S-рРНК и сайта инициации трансляции S7-мРНК (*str*-оперон полицистронной мРНК). Сходные нуклеотидные последовательности выделены цветом. [M. Nomura et al., Proc. Natl. Acad. Sci. USA, **77** (1980), p. 7084; J.D. Watson et al., Molecular Biology of the Gene, 4th ed. (Menlo Park, Calif.: Benjamin/Cummings, 1987). См. рис. 3.67 и табл. 3.7, где представлена более подробная информация об этих оперонах.]

мРНК, имеют сходную вторичную структуру, причем петли имеют идентичные последовательности.

Кроме регуляции на уровне трансляции при образовании рибосомных белков осуществляется также регуляция по типу обратной связи при транскрипции оперонов рибосомных белков. Мы не будем обсуждать регуляцию этих оперонов на уровне транскрипции, а лишь отметим, что могут происхо-

дить репрессия и аттенуация транскрипции рибосомными субчастицами или даже целыми рибосомами. Наиболее вероятно, однако, что ключевым пунктом в регуляции скорости синтеза рибосом является образование рРНК. Таким образом, синтез рибосомных белков и тем самым сборка рибосом регулируются путем изменения содержания рРНК.

# Литература

## Введение

Приведены классические статьи и книги по генетике и биохимии, в которых изложены основные концепции молекулярной генетики. Список дан в хронологическом порядке.

- G. Mendel. 1866. Versuche über Pflanzen-Hybriden. Verhandlungen der Naturforschenden Verein, Brünn, 4, 3-47 (перепечатано в Journal of Heredity, 42, 3-47, 1951). Эта и еще одна важная работа Менделя, „Über einige aus künstlicher Befruchtung gewonnene Hieracium-Bastarde“, напечатаны по-английски в кн.: C. Stern, E.R. Sherwood (eds.). 1966. The Origins of Genetics. A Mendel Source Book. W.H. Freeman, San Francisco.
- F. Miescher. 1871. Über die chemische Zusammensetzung der Eiterzellen. Hoppe-Seyler's Medizinisch-Chemischen Untersuchungen, 4, 441-460.
- A.E. Garrod. 1902. Inborn Errors of Metabolism. Lancet, 2, 1616-1620.
- W.S. Sutton. 1903. The Chromosomes in Heredity. Biol. Bull., 4, 231-251.
- A.E. Garrod. 1909. Inborn Errors of Metabolism. Frowde, Hodder, and Stoughton, London.
- W. Johannsen. 1909. Elemente der Exakten Erblichkeitslehre. Fischer. Jena.
- T.H. Morgan. 1910. Sex-Linked Inheritance in *Drosophila*. Science, 32, 120-122.
- T.H. Morgan, A.H. Sturtevant, H.J. Muller, C.B. Bridges. 1915. The Mechanism of Mendelian Heredity. Holt, Rinehart & Winston, New York.
- B. McClintock, H.B. Creighton. 1931. A Correlation of Cytological and Genetical Crossing Over in *Zea mays*. Proc. Natl. Acad. Sci. U.S.A., 17, 492-497.
- B. McClintock. 1934. The Relation of a Particular Chromosomal Element to the Development of the Nucleoli in *Zea mays*. A. Zellforsch. u. Mikr. Anat., 21, 294-328.
- G.W. Beadle, E.L. Tatum. 1941. Genetic Control of Biochemical Reactions in *Neurospora*. Proc. Natl. Acad. Sci. U.S.A., 27, 499-506.
- O.T. Avery, C.M. Macleod, M. McCarty. 1944. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. I. Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III. J. Exp. Med., 79, 137-158.
- E. Chargaff, R. Lipschitz, C. Green, M.E. Hodes. 1951. The Composition of the Deoxyribonucleic Acid of Salmon Sperm. J. Biol. Chem., 192, 223-230.
- D. M. Brown, A.R. Todd. 1952. Nucleotides. Part X. Some Observations on the Structure and Chemical Behavior of the Nucleic Acids. J. Chem. Soc., 52-58.
- A.D. Hershey, M. Chase. 1952. Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage. J. Gen. Physiol., 36, 39-56.
- J.D. Watson, F.H.C. Crick. 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxynucleic Acids. Nature, 171, 737-738.
- J.D. Watson, F.H.C. Crick. 1953. General Implications of the Structure of Deoxyribonucleic Acid. Nature, 171, 964-967.
- V.M. Ingram. 1957. Gene Mutations in Human Hemoglobin:

The Chemical Difference Between Normal and Sick Cell Hemoglobin. Nature, 180, 326-328.

- M.B. Hoagland, M.L. Stephenson, J.F. Scott, L.I. Hecht, P.C. Zamecnik. 1958. A Soluble Ribonucleic Acid Intermediate in Protein Synthesis. J. Biol. Chem., 231, 241-257.
- M. Meselson, F.W. Stahl. 1958. The Replication of DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A., 44, 671-682.
- A. Kornberg. 1960. Biological Synthesis of Deoxyribonucleic Acid. Science, 131, 1503-1508.
- S. Brenner, F. Jacob, M. Meselson. 1961. An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. Nature, 190, 576-581.
- F. Jacob, J. Monod. 1961. Genetic Regulatory Mechanisms in the Synthesis of Proteins. J. Mol. Biol., 3, 318-356.
- M.W. Nirenberg, H.J. Matthaei. 1961. The Dependence of Cell-Free Protein Synthesis in *E. coli* upon Naturally Occurring or Synthetic Polyribonucleotides. Proc. Natl. Acad. Sci. U.S.A., 47, 1588-1602.
- C. Yanofsky, B.C. Carlton, J.R. Guest, D.R. Helinski, U. Henning. 1964. On the Colinearity of Gene Structure and Protein Structure. Proc. Natl. Acad. Sci. U.S.A., 51, 266-272.
- The Genetic Code. 1966. Cold Spring Harbor Symp. Quant. Biol., 31, whole issue.

В следующих работах описано историческое развитие генетики и биохимии нуклеиновых кислот.

- J.S. Fruton. 1972. Molecules and Life. Wiley-Interscience, New York.
- H. Stubbe. 1972. History of Genetics: From Prehistoric Times to the Discovery of Mendel's Laws. Translated by T.R.W. Waters. MIT Press, Cambridge, Massachusetts.
- R. Olby. 1974. The Path to the Double Helix. University of Washington Press, Seattle.
- F.H. Portugal, J.S. Cohen. 1977. A Century of DNA. MIT Press, Cambridge, Massachusetts.
- H.F. Judson. 1979. The Eighth Day of Creation. Simon and Schuster, New York.

## Общие работы для всех глав части I

- C.K. Cantor, P.R. Schimmel. 1980. Biophysical Chemistry. W.H. Freeman, San Francisco. [Имеется перевод: Кантор Ч., Шиммель П. Биофизическая химия: В 3-х т. - М.: Мир, 1983-1985 гг.]
- B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J.D. Watson. 1983. Molecular Biology of the Cell. Garland. New York. [Имеется перевод: Албертс Б., Уотсон Дж., Моне Дж. и др. Молекулярная биология клетки: В 5 т. - М.: Мир, 1986-1987 гг.]
- C. Zubay. 1983. Biochemistry. Addison-Wesley, Reading, Massachusetts.
- J. Darnell, H. Lodish, D. Baltimore. 1986. Molecular Cell Biology. Scientific American Books, New York.
- D. Freifelder. 1987. Molecular Biology, 2nd ed. Jones and Bartlett, Boston.
- B. Lewin. 1987. Genes, 3rd ed. Wiley, New York. [Имеется

- перевод 2-го издания с дополнениями: Льюин Б. Гены.- М.: Мир, 1987.]
- J.D. Watson, N.H. Hopkins, J.W. Roberts, J.A. Steitz, A.M. Weiner. 1987. *Molecular Biology of the Gene*, 4th ed. Benjamin/Cummings, Menlo Park, California.
- J.D. Rawn. 1988. *Biochemistry*. Carolina Biological Supply Co., New York.
- L. Stryer. 1988. *Biochemistry*, 3rd ed. W.H. Freeman, San Francisco. [Имеется перевод 2-го издания: Страйер Л. Биохимия.- М.: Мир, 1985.]
- ## Литература к указанным разделам глав
- ### 1.1.
- F.H.C. Crick, J.D. Watson. 1954. The Complementary Structure of Deoxyribonucleic Acid. *Proc. Roy. Soc.*, **223** (4), 80-96.
- F.H.C. Crick. 1976. Linking Numbers and Nucleosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 2639-2643.
- W.R. Bauer, F.H.C. Crick, J.H. White. 1980. Supercoiled DNA. *Sci. American*, **243**, 118-133.
- A. Kornberg. 1980 and 1982 supplement. *DNA Replication*. W.H. Freeman, San Francisco. [Имеется перевод издания 1974 г.: Корнберг А. Синтез ДНК. - М.: Мир, 1977.]
- R.D. Kornberg, A. Klug. 1981. The Nucleosome. *Sci. American*, **244**, 52-79.
- R.E. Dickerson, H.R. Drew, B.N. Conner, R.M. Wing, A.V. Fratini, M.L. Kopka. 1982. The Anatomy of A-, B- and Z-DNA. *Science*, **216**, 475-485.
- A. Klug. 1982. From Macromolecules to Biological Assemblies. Nobel lecture. Nobel Foundation, Stockholm, Sweden.
- Structure of DNA. 1982. Cold Spring Harbor Symp. Quant. Biol., 47 whole issue.
- J.C. Wang. 1982. DNA Topoisomerases. *Sci. American*, **247**, 94-109.
- A. Rich, A. Nordheim, H.-J. Wang. 1984. The Chemistry and Biology of Left-Handed Z-DNA. *Annu. Rev. Biochem.*, **53**, 791-846.
- W. Sanger. 1984. *Principles of Nucleic Acid Structure*. Springer-Verlag, New York. [Имеется перевод: Зенгер В. Принципы структурной организации нуклеиновых кислот. - М.: Мир, 1987.]
- G. Felsenfeld. 1985. DNA. *Sci. American*, **253** (4), 58-67.
- D.E. Pettijohn. 1988. Histone-like Proteins and Bacterial Chromosome Structure. *J. Biol. Chem.*, **263**, 12793-12796.
- K.E. van Holde. 1988. *Chromatin*. Springer-Verlag, New York.
- R.D. Wells. 1988. Unusual DNA Structures. *J. Biol. Chem.*, **263**, 1095-1098 (краткий обзор).
- ### 1.2
- R.W. Holley. 1968. The Nucleotide Sequence of a Nucleic Acid. *Sci. American*, **214**, 30-39.
- A. Rich, U.L. Raj Bhandary. 1976. Transfer RNA: Molecular Structure, Sequence and Properties. *Annu. Rev. Biochem.*, **45**, 805-860.
- H.F. Noller. 1984. Structure of Ribosomal RNA. *Annu. Rev. Biochem.*, **55**, 119-162.
- J.E. Darnell, Jr. 1985. RNA. *Sci. American*, **253** (4), 68-78.
- ### 1.3
- R.E. Dickerson, I. Geis. 1969. *The Structure and Action of Proteins*. Benjamin/Cummings, Menlo Park, California.
- T.E. Creighton. 1983. *Proteins: Structure and Molecular Properties*. W.H. Freeman, San Francisco.
- R.E. Dickerson, I. Geis. 1983. *Hemoglobin: Structure, Function, Evolution and Pathology*. Benjamin/Cummings, Menlo Park, California.
- C. Chothia. 1984. Principles That Determine the Structure of Proteins. *Annu. Rev. Biochem.*, **53**, 537-572.
- R.F. Doolittle. 1985. Proteins. *Sci. American*, **253** (4), 88-99.
- R.J. Fletterick, T. Schroer, R.J. Matela. 1985. *Molecular Structure: Macromolecules in Three Dimensions*. Blackwell Scientific Publications, Oxford, England.
- M.E. Goldberg. 1985. The Second Translation of the Genetic Message: Protein Folding and Assembly. *Trends Biochem. Sci.*, **10**, 388-391.
- ### 2.1
- J. Cairns. 1966. The Bacterial Chromosome. *Sci. American*, **214**, 36-44.
- W. Gilbert, D. Dressier. 1968. DNA Replication: The Rolling Circle Model. *Cold Spring Harbor Symp. Quant. Biol.*, **33**, 473-484.
- J.S. Sussenbach, P.C. Van der Vliet, D.J. Ellens, H.S. Jansz. 1972. Linear Intermediates in the Replication of Adenovirus DNA. *Nature New Biol.*, **239**, 47-49.
- J.D. Watson. 1972. Origin of Concatameric T7 DNA. *Nature New Biol.*, **235**, 197-201.
- J.D. Wolfson, D. Dressier, M. Magazin. 1972. Bacteriophage T7 DNA Replication: A Linear Replicating Intermediate. *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 499-504.
- I.R. Lehman. 1974. DNA Ligase: Structure, Mechanism and Function. *Science*, **186**, 790-797.
- D. Bastia, M. Sueoka, E. Cox. 1975. Studies in the Late Replication of Phage  $\lambda$ : Rolling Circle Replication. *J. Mol. Biol.*, **98**, 305-320.
- D.M. Rekosch, W.C. Russel, A.J.D. Bellet, A.J. Robinson. 1977. Identification of a Protein Linked to the Ends of Adenovirus DNA. *Cell*, **11**, 283-295.
- DNA: Replication and Recombination. 1978. Cold Spring Harbor Symp. Quant. Biol. 43 whole issue.
- S.H. Wickner. 1978. DNA Replication Proteins of *Escherichia coli*. *Annu. Rev. Biochem.*, **47**, 1163-1191.
- T.J. Kelly, Jr., R.L. Lechner. 1976. The Structure of Replicating Adenovirus DNA Molecules: Characterization of DNA-Protein Complexes from Infected Cells. *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 721-728.
- N.R. Cozzarelli. 1980. DNA Gyrase and the Supercoiling of DNA. *Science*, **207**, 953-960.
- M. Gellert. 1981. DNA Topoisomerases. *Annu. Rev. Biochem.*, **50**, 879-910.
- Y. Hirota, M. Yamada, A. Nishimura, A. Oka, K. Sugimoto, K. Asada, M. Takamami. 1981. The DNA Replication Origin (ori) of *Escherichia coli*: Structure and Function of the ori-Containing DNA Fragment. *Prog. Nucleic. Acid. Res.*, **26**, 33-48.
- O. Sundin, A. Varshavsky. 1981. Arrest of Segregation Leads to Accumulation of Highly Intertwined Catenated Dimers: Dissection of the Final Stages of SV40 DNA Replication. *Cell*, **25**, 659-669.
- B.M. Alberts, B.P. Bedinger, T. Formosa, C.V. Jongeneel, K.N. Kreuzer. 1982. Studies on DNA Replication in the Bacteriophage T4 in vitro Systems. *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 655-668.
- B.M. Baroudy, S. Venkatesan, B. Moss. 1982. Structure and Replication of Vaccinia Virus Telomeres. *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 723-729.
- B.W. Stillman, F. Tamenoi. 1982. Adenoviral DNA Replication: DNA Sequences Required for Initiation in vitro. *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 741-750.
- A. Kornberg. 1983. Mechanism of Replication of the *E. coli* Chromosome. *Eur. J. Biochem.*, **137**, 377-382.
- N.G. Nossal. 1983. Prokaryotic DNA Replication Systems. *Annu. Rev. Biochem.*, **52**, 581-615.
- B.M. Alberts. 1984. The DNA Enzymology of Protein Machines. *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 1-12.
- S. Dinardo, K. Voelkel, R. Sternglanz. 1984. DNA Topoisomerase II Is Required for Segregation of Daughter Molecules at the Termination of DNA Replication. *Proc. Natl.*

- Acad. Sci. U.S.A., **81**, 2616-2620.
- M. Kaguni, A. Kornberg. 1984. Replication Initiated at the Origin (oriC) of the *E. coli* Chromosome Reconstituted with Purified Enzymes. *Cell*, **38**, 183-190.
- J. Tomizawa. 1984. Control of Col E1 Plasmid Replication: The Process of Binding of RNA I to the Primer Transcript. *Cell*, **38**, 861-870.
- T. Kirchhausen, J.C. Wang, S.C. Harrison. 1985. DNA Gyrase and Its Complexes with DNA: Direct Observation by Electron Microscopy. *Cell*, **41**, 933-943.
- C.S. McHenry. 1985. DNA Polymerase III Holoenzyme of *Escherichia coli*: Components and Function of a True Replicative Complex. *Mol. Cell. Biochem.*, **66**, 71-85.
- B.W. Stillman, Y. Gluzman. 1985. Replication and Supercoiling of SV40 DNA in Cell-Free Extracts from Human Cells. *Mol. Cell. Biol.*, **5**, 2051-2060.
- J.L. Campbell. 1986. Eukaryotic DNA Replication. *Annu. Rev. Biochem.*, **55**, 733-771.
- H. Echols. 1986. Multiple DNA-Protein Interactions Governing High-Precision DNA Transactions. *Science*, **233**, 1050-1056.
- A. Maxwell, M. Gellert. 1986. Mechanistic Aspects of DNA Topoisomerases. *Adv. Prot. Chem.*, **38**, 69-107.
- S.A. Wasserman, N.R. Cozarella. 1986. Biochemical Topology: Applications to DNA Recombination and Replication. *Science*, **232**, 951-960.
- C.M. Joyce, T. Steitz. 1987. DNA Polymerase I: From Crystal Structure to Function via Genetics. *Trends Biochem. Sci.*, **12**, 288-292.
- T. Kelly, B. Stillman (eds.). 1988. Eukaryotic DNA Replication. Cold Spring Harbor Laboratory, New York.
- A. Kornberg. 1988. DNA Replication. *J. Biol. Chem.*, **263**, 1-4.
- 2.2**
- D. Baltimore. 1970. Viral RNA-Dependent DNA Polymerase. *Nature*, **226**, 1209-1211.
- H.M. Temin, S. Mizutani. 1970. Viral RNA-Dependent DNA Polymerase. *Nature*, **226**, 1211-1213.
- E. Giboa, S.W. Mitra, S. Goff, D. Baltimore. 1979. A Detailed Model of Reverse Transcription and Tests of Crucial Aspects. *Cell*, **18**, 93-100.
- W.S. Mason, J.M. Taylor, R. Hull. 1987. Retrovirus Genome Replication. *Adv. Virus Res.*, **32**, 35-96.
- H. Varmus. 1987. Reverse Transcription. *Sci. American*, **257**, 56-64.
- 2.3**
- P. Howard-Flanders. 1981. Inducible Repair of DNA. *Sci. American*, **245**, 72-103.
- T. Lindahl. 1982. DNA Repair Enzymes. *Annu. Rev. Biochem.*, **51**, 61-87.
- L.A. Loeb, T.A. Kunkel. 1982. Fidelity of DNA Synthesis. *Annu. Rev. Biochem.*, **51**, 429-457.
- A. Sancar, W.D. Rupp. 1983. A Novel Repair Enzyme: UVRABC Excision Nuclease of *Escherichia coli* Cuts a DNA Strand on Both Sides of the Damaged Region. *Cell*, **33**, 249-260.
- A.T. Yeung, W.B. Mattes, E.Y. Oh, L. Grossman. 1983. Enzymatic Properties of Purified *Escherichia coli* uvrABC Proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 6157-6161.
- E.C. Friedberg. 1984. DNA Repair. W.H. Freeman, San Francisco.
- A. Sancar, G.B. Sancar. 1988. DNA Repair Enzymes. *Annu. Rev. Biochem.*, **57**, 29-67.
- 2.4**
- N. Sigal, B. Alberts. 1972. Genetic Recombination: The Nature of a Crossed Stranded Exchange Between Two Homologous DNA Molecules. *J. Mol. Biol.*, **71**, 789-793.
- S.A. Latt. 1979. Sister Chromatid Exchanges. *Genetics*, **92**, 583-595.
- F.W. Stahl. 1979. Genetic Recombination: Thinking About It in Phage and Fungi. W.H. Freeman, San Francisco.
- C. DasGupta, A. Wu, R. Kahn, R. Cunningham, C. Radding. 1981. Concerted Strand Exchange and Formation of Holliday Structures by *E. coli* RecA Protein. *Cell*, **25**, 507-516.
- H. Nash. 1981. Integration and Excision of Bacteriophage  $\lambda$ : The Mechanism of Conservative Site Specific Recombination. *Annu. Rev. Genet.*, **15**, 143-167.
- C. Radding. 1982. Homologous Pairing and Strand Exchange in Genetic Recombination. *Annu. Rev. Genet.*, **16**, 405-437.
- J. Szostak, T. Orr-Weaver, R. Rothstein, F. Stahl. 1983. The Double-Strand Break Repair Model for Recombination. *Cell*, **33**, 25-35.
- R. Weisberg, A. Landy. 1983. Site-Specific Recombination in Phage  $\lambda$ . In R. Hendrix, J. Roberts, F. Stahl and R. Weisberg (eds.), *Lambda II*, pp. 211-250. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- M.M. Cox, J.R. Lehman. 1987. Enzymes of General Recombination. *Annu. Rev. Biochem.*, **56**, 229-262.
- P. Modrich. 1987. DNA Mismatch Correction. *Annu. Rev. Biochem.*, **56**, 435-466.
- 2.5**
- S. Spiegelman, I. Haruna, I.B. Holland, B. Beaudreau, D. Mills. 1965. The Synthesis of a Self-Propagative and Infectious Nucleic Acid with a Purified Enzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **54**, 919-927.
- T. Blumenthal, G.C. Carmichael. 1979. RNA Replication: Function and Structure of QB-Replicase. *Annu. Rev. Biochem.*, **48**, 525-548.
- 3.1**
- F.H.C. Crick. 1958. On Protein Synthesis, Biological Replication of Macromolecules. *Symp. Exp. Biol.*, **12**, 138-163.
- J.D. Watson. 1963. The Involvement of RNA in the Synthesis of Proteins. *Science*, **140**, 17-26.
- H. Weissbach, S. Pestka (eds.). 1977. Molecular Mechanisms of Protein Biosynthesis. Academic Press, New York.
- 3.2**
- J.W. Roberts. 1969. Termination Factor for RNA Synthesis. *Nature*, **224**, 1168-1174.
- A.A. Travers, R.R. Burgess. 1969. Cyclic Reuse of the RNA Polymerase Sigma Factor. *Nature*, **222**, 537-540.
- C. Lowery-Goldhammer, J.P. Richardson. 1974. An RNA-Dependent Nucleoside Triphosphate Phosphohydrolase (ATPase) Associated with Rho Termination Factor. *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 2003-2007.
- M. Rosenberg, D. Court. 1979. Regulatory Sequences Involved in the Promotion and Termination of RNA Transcription. *Annu. Rev. Genet.*, **13**, 319-353.
- R. Losick, J. Pero. 1981. Cascades of Sigma Factors. *Cell*, **25**, 582-584.
- M. Chamberlin. 1982. Bacterial DNA-Dependent RNA Polymerases. In P.D. Boyer (ed.), *The Enzymes*, part B, vol. 15, pp. 61-108. Academic Press, New York.
- R. Rodriguez, M. Chamberlin (eds.). 1982. Promoters: Structure and Function. Praeger, New York.
- O.K. Hawley, W.R. McClure. 1983. Compilation and Analysis of *Escherichia coli* Promoter DNA Sequences. *Nucleic Acid Res.*, **11**, 2237-2255.
- P.H. von Hippel, D.G. Bear, W.D. Morgan, J.A. McSwiggen. 1984. Protein-Nucleic Acid Interactions in Transcription: A Molecular Analysis. *Annu. Rev. Biochem.*, **53**, 389-446.
- W.R. McClure. 1985. Mechanism and Control of Transcription Initiation in Prokaryotes. *Annu. Rev. Biochem.*, **54**, 171-204.
- T. Platt. 1986. Transcription Termination and the Regulation of Gene Expression. *Annu. Rev. Biochem.*, **55**, 339-372.



- C.A. Brennan, A.J. Dombrowski, T. Platt. 1987. Transcription Termination Factor Rho Is an RNA-DNA Helicase. *Cell*, **48**, 945-952.
- 3.3**
- E. Lund, J.E. Dahlberg, L. Lindahl, R. Jaskunas, P.P. Dennis, M. Nomura. 1976. Transfer RNA Genes Between 16S rRNA Gene in RNA Transcription Units. *Cell*, **7**, 165-177.
- J.D. Smith. 1976. Transcription and Processing of Transfer RNA Precursors. *Proc. Nucleic Acid. Res.*, **16**, 25-73.
- N. Nakajima, H. Ozeki, Y. Shimura. 1977. Organization and Structure of an *E. coli* tRNA Operon Containing Seven tRNA Genes. *Cell*, **23**, 239-249.
- P. Gegenheimer, D. Apirion. 1981. Processing of Prokaryotic Ribonucleic Acid. *Microbiol. Rev.*, **45**, 502-541.
- S. Altman, C. Guerrier-Takeda, H. Frankfort, H. Robertson. 1982. RNA Processing Nucleases. In S. Linn and R. Roberts (eds.), *Nucleases*, pp. 243-274. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- H.D. Robertson. 1982. *Escherichia coli* Ribonuclease III Cleavage Sites. *Cell*, **30**, 669-672.
- M. Deutscher. 1984. Processing of tRNA in Prokaryotes and Eukaryotes. *Crit. Rev. Biochem.*, **17**, 45-71.
- 3.4**
- F.H.C. Crick. 1958. On Protein Synthesis, Biological Replication of Macromolecules. *Symp. Exp. Biol.*, **12**, 158-163.
- F.H.C. Crick, L. Barnett, S. Brenner, R.J. Watts-Tobin. 1961. General Nature of the Genetic Code for Proteins. *Nature*, **192**, 1227-1232.
- M.W. Nirenberg, J.H. Matthaei. 1961. The Dependence of Cell-Free Protein Synthesis in *E. coli* upon Naturally Occurring or Synthetic Polyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.*, **47**, 1588-1682.
- J.F. Speyer, P. Lengyel, C. Basilica, A.J. Wahba, R.S. Gardner, S. Ochoa. 1963. Synthetic Polynucleotides and the Amino Acid Code. *Cold Spring Harbor Symp. Quant. Biol.*, **28**, 559-568.
- M. Nirenberg, P. Leder. 1964. The Effect of Trinucleotides upon the Binding of sRNA to Ribosomes. *Science*, **145**, 1399-1407.
- A.S. Sarabhai, A.O. Stretton, S. Brenner, A. Bolle. 1964. Colinearity of the Gene with the Polypeptide Chain. *Nature*, **201**, 13-17.
- C. Yanofsky, B.C. Carlton, J.R. Guest, D.R. Helinski, U. Henning. 1964. On the Colinearity of Gene Structure and Protein Structure. *Proc. Natl. Acad. Sci. U.S.A.*, **51**, 266-272.
- F.H.C. Crick. 1966. Codon-Anticodon Pairing: The Wobble Hypothesis. *J. Mol. Biol.*, **19**, 548-555.
- The Genetic Code. 1966. Cold Spring Harbor Symp. Quant. Biol., **31**, whole issue.
- M. Yeas. 1969. *The Biological Code*. Wiley-Interscience, New York.
- 3.5**
- P. Berg, E.J. Ofengand. 1958. An Enzymatic Mechanism for Linking Amino Acids to RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **44**, 78-86.
- M.B. Hoagland, M.L. Stephenson, J.F. Scott, L.I. Hecht, P.C. Zamecnik. 1958. A Soluble Ribonucleic Acid Intermediate in Protein Synthesis. *J. Biol. Chem.*, **231**, 241-257.
- F. Chapeville, F. Lipmann, G.V. Elivenstein, B. Weisblum, W.J. Ray, Jr. S. Benzer. 1962. On the Role of Soluble Ribonucleic Acid in Coding for Amino Acids. *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 1086-1092.
- R.W. Holley. 1966. The Nucleotide Sequence of a Nucleic Acid. *Sci. American*, **214**, 30-39.
- M. Nomura. 1973. Assembly of Bacterial Ribosomes. *Science*, **179**, 864-873.
- A. Rich, U.L. Raj Bhandary. 1976. Transfer RNA: Molecular Structure, Sequence and Properties. *Annu. Rev. Biochem.*, **45**, 805-860.
- A. Rich, S.H. Kim. 1978. The Three-Dimensional Structure of Transfer RNA. *Sci. American*, **238**, 56-62.
- P.R. Schimmel, D. Soil. 1979. Amino Acyl tRNA Synthetases: General Features and Recognition of tRNAs. *Annu. Rev. Biochem.*, **48**, 601-648.
- I. Wool. 1979. The Structure and Function of Eucaryotic Ribosomes. *Annu. Rev. Biochem.*, **48**, 719-754.
- J.A. Lake. 1981. The Ribosome. *Sci. American*, **245**, 84-97.
- K.H. Nierhaus. 1982. Structure, Assembly and Function of Ribosomes. *Current Topics in Microbiol. and Immun.*, **97**, 82-155.
- P. Schimmel, S. Putney, R. Starzyk. 1982. RNA and DNA Sequence Recognition and Structure-Function of Amino Acyl tRNA Synthetases. *Trends Biochem. Sci.*, **7**, 209-212.
- H.G. Wittmann. 1982. Components of Bacterial Ribosomes. *Annu. Rev. Biochem.*, **51**, 155-183.
- H.G. Wittmann. 1983. Architecture of Prokaryotic Ribosomes. *Annu. Rev. Biochem.*, **52**, 35-66.
- M. Yarus, R. Thompson. 1983. Precision of Protein Biosynthesis. In: J. Beckwith, J. Davies, J.A. Gallant (eds.), *Gene Function in Prokaryotes*, pp. 23-63. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- H.F. Noller. 1984. Structure of Ribosomal RNA. *Annu. Rev. Biochem.*, **53**, 119-162.
- D.M. Blow, P. Brick. 1985. Amino Acyl tRNA Synthetases. In: F.A. Jurnak, A. McPherson (eds.), *The Structure of Biological Macromolecules and Assemblies*, vol. 2, *Nucleic Acid Binding Proteins*, pp. 442-469. Wiley, New York.
- J.A. Lake. 1985. Evolving Ribosome Structure: Domains in Archaeobacteria, Eubacteria, Eocytes and Eukaryotes. *Annu. Rev. Biochem.*, **54**, 507-530.
- G.R. Björk, J.U. Ericson, C.E.D. Gustafsson, T.G. Hagerrall, Y.H. Jönsson, P.M. Wilkström. 1987. Transfer RNA Modification. *Annu. Rev. Biochem.*, **56**, 263-288.
- P. Schimmel. 1987. Amino Acyl tRNA Synthetases: General Scheme of Structure-Function Relationship in the Polypeptides and Recognition of Transfer RNAs. *Annu. Rev. Biochem.*, **56**, 125-158.
- 3.6**
- H.M. Dintzis. 1961. Assembly of the Peptide Chain of Hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.*, **47**, 247-261.
- J. Shine, L. Dalgarno. 1974. The 3'-Terminal Sequence of *E. coli* 16S rRNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites. *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 1342-1346.
- Y. Kaziro. 1978. The Role of Guanosine 5'-Triphosphate in Polypeptide Chain Elongation. *Biochem. Biophys. Acta*, **505**, 95-127.
- M. Grunberg-Manago. 1980. Initiation of Protein Synthesis as Seen in 1979. In: G. Chambliss, G.R. Craven, J. Davies, K. Davis, L. Kahan, and M. Nomura (eds.), *Ribosomes: Structure, Function and Genetics*, pp. 445-478. University Park Press, Baltimore.
- J. Ofengand. 1980. The Topography of tRNA Binding Sites on the Ribosome. In: G. Chambliss, G.R. Craven, J. Davies, K. Davis, L. Kahan, and M. Nomura (eds.), *Ribosomes: Structure, Function and Genetics*, pp. 497-530. University Park Press, Baltimore.
- H. Weissbach. 1980. Soluble Factors in Protein Synthesis. In: G. Chambliss, G.R. Craven, J. Davies, K. Davis, L. Kahan, M. Nomura (eds.), *Ribosomes: Structure, Function and Genetics*, pp. 445-478. University Park Press, Baltimore.
- L. Gold, D. Pribnow, T. Schneider, S. Shinedling, B.S. Singer, G. Starmo. 1981. Translational Initiation in Prokaryotes. *Annu. Rev. Microbiol.*, **35**, 365-403.
- R.A. Garrett, P. Wooley. 1982. Identifying the Peptidyl Trans-

- ferase Centre. Trends Biochem. Sci., **7**, 385-386.
- A. Johnson, H. Adkins, E. Mathews, C. Cantor. 1982. Distance Moved by Transfer RNA During Translocation from the A Site to the P Site on the Ribosome. J. Mol. Biol., **156**, 113-140.
- M. Kozak. 1983. Comparison of Initiation of Protein Synthesis in Prokaryotes, Eukaryotes and Organelles. Microbiol. Rev., **47**, 1-45.
- C.T. Caskey, W.S. Forrester, W. Tate. 1984. Peptide Chain Termination. In: B. Clark and H. Petersen (eds.), Alfred Benzon Symposium, vol. 19, pp. 457-466. Munksgaard, Copenhagen.
- K. Moldave. 1985. Eukaryotic Protein Synthesis. Annu. Rev. Biochem., **54**, 1109-1149.
- 3.7**
- K. Moldave. 1965. Nucleic Acids and Protein Biosynthesis. Annu. Rev. Biochem., **34**, 419-448.
- F.H.C. Crick. 1966. Codon-Anticodon Pairing: The Wobble Hypothesis. J. Mol. Biol., **19**, 548-555.
- Protein Synthesis. 1969. Cold Spring Harbor Symp. Quant. Biol., **34**, whole issue.
- B.A. Hamkalo, O.L. Miller, Jr. 1973. Electron Microscopy of Genetic Activity. Annu. Rev. Biochem., **42**, 379-396.
- R. Haselkorn, L.B. Rothman-Denes. 1973. Protein Synthesis. Annu. Rev. Biochem., **42**, 397-438.
- M. Nomura, A. Tissieres, P. Lengyel (eds.). 1974. Ribosomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- E. Beronak. 1978. Mechanisms in Polypeptide Chain Elongation on Ribosomes. Proc. Nucl. Acid Res. Mol. Biol., **21**, 63-100.
- M. Grunberg-Manago, R.H. Buckingham, B.S. Cooperman, J.W.B. Hershey. 1978. Structure and Function of the Translation Machinery. Symp. Soc. Gen. Microbiol., **28**, 27-110.
- D.A. Steege, D.G. Söll. 1979. Suppression. In: R.F. Goldberg (ed.), Biological Regulation and Development I, pp. 433-486. Plenum, New York.
- H. Ozeki, H. Inokuchi, F. Yamao, M. Kodaira, H. Sakano, T. Ikemura, Y. Shimura. 1980. Genetics of Nonsense Suppressor tRNAs in *E. coli*. In: D. Söll, J.M. Abelson, P.R. Schimmel (eds.), Transfer RNA, Biological Aspects, pp. 341-349. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- J.R. Roth. 1981. Frameshift Suppression. Cell, **24**, 601-602.
- M. Yarns, R. Thompson. 1983. Precision of Protein Biosynthesis. In: J. Beckwith, J. Davies, J.A. Gallant (eds.), Gene Function in Prokaryotes, pp. 23-63. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- 3.8**
- K. Moldave. 1985. Eukaryotic Protein Synthesis. Annu. Rev. Biochem., **54**, 1109-1149.
- 3.9**
- J.H. Goldberg, P.A. Friedman. 1971. Antibiotics and Nucleic Acids. Annu. Rev. Biochem., **40**, 775-810.
- H.M. Sobell. 1974. How Actinomycin Binds to RNA. Sci. American, **231**, 82-91.
- S. Pestka. 1977. Inhibitors of Protein Biosynthesis. In: H. Weissbach, S. Pestka (eds.), Molecular Mechanisms of Protein Biosynthesis, pp. 467-553. Academic Press, New York.
- R.J. Subadnik. 1979. Naturally Occurring Nucleoside and Nucleotide Antibiotics. Prog. Nucl. Acid Res. Mol. Biol., **22**, 193-291.
- E. Cundliffe. 1980. Antibiotics and Prokaryote Ribosomes: Action, Interaction and Resistance. In: G. Chambliss, G.R. Craven, J. Davies, K. Davis, L. Kahan, M. Nomura (eds.), Ribosomes: Structure, Function and Genetics, pp. 377-412. University Park Press, Baltimore.
- 3.10**
- S. Michaelis, J. Beckwith. 1982. Mechanism of Incorporation of Cell Envelope Proteins in *Escherichia coli*. Annu. Rev. Microbiol., **36**, 435-465.
- G. Schatz, R.A. Butow. 1983. How Are Proteins Imported into Mitochondria? Cell, **32**, 316-318 (краткий обзор).
- T.J. Silhavy, S.A. Benson, S.D. Emr. 1983. Mechanisms of Protein Localization. Microbiol. Rev., **47**, 313-344.
- L.L. Randall, S.J.S. Hardy. 1984. Export of Protein in Bacteria: Dogma and Data. In: B. Satir (ed.), Modern Cell Biology, vol. 3, pp. 1-20. Liss, New York.
- P. Walter, R. Gilmore, G. Blobel. 1984. Protein Translocation Across the Endoplasmic Reticulum. Cell, **38**, 5-8.
- J.E. Rothman. 1985. The Compartmental Organization of the Golgi Apparatus. Sci. American, **253** (3), 74-89.
- R. Scheckman. 1985. Protein Localization and Membrane Traffic. Annu. Rev. Cell Biol., **1**, 115-143.
- M. Schleyer, W. Neupert. 1985. Transport of Proteins into Mitochondria: Translocational Intermediates Spanning Contact Sites Between Outer and Inner Membranes. Cell, **43**, 339-350.
- S.R. Pfeffer, J.E. Rothman. 1987. Biosynthetic Protein Transport by the Endoplasmic Reticulum and Golgi. Annu. Rev. Biochem., **56**, 829-852.
- J.E. Rothman. 1987. Protein Sorting by Selective Retention in the Endoplasmic Reticulum and Golgi Stack. Cell, **50**, 521-522.
- H.F. Lodish. 1988. Transport of Secretory and Membrane Glycoproteins from the Rough Endoplasmic Reticulum to the Golgi. J. Biol. Chem., **263**, 2107-2110.
- D. Raise, G. Schatz. 1988. Mitochondrial Presequences. J. Biol. Chem., **263**, 4509-4511 (краткий обзор).
- 3.11**
- F. Jacob, J. Monod. 1961. Genetic Regulatory Mechanisms in the Synthesis of Proteins. J. Mol. Biol., **3**, 318-356.
- S. Adhya, M. Gottesman. 1978. Control of Transcription Termination. Annu. Rev. Biochem., **47**, 967-996.
- W. Fiers. 1979. Structure and Function of RNA Bacteriophages. Comp. Virology, **13**, 69-204.
- R.F. Goldberg (ed.). 1979. Biological Regulation and Development. I. Gene Expression. Plenum, New York.
- J. Miller, W. Reznikoff (eds.). 1980. The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- C. Yanofsky, R. Koller. 1982. Attenuation in Amino Acid Biosynthesis Operons. Annu. Rev. Genet., **16**, 113-134.
- C. Bauer, J. Carey, L. Kasper, S. Lynn, D. Woelcher, J. Gardner. 1983. Attenuation in Bacterial Operons. In: J. Beckwith, J. Davies, J. Gallant (eds.), Gene Function in Prokaryotes, pp. 65-89. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- K.M. Campbell, C.D. Starvo, L. Gold. 1983. Protein-Mediated Translation Repression. In: J. Beckwith, J. Davies, J. Gallant (eds.), Gene Function in Prokaryotes, pp. 185-187. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- R.W. Hendrix, J.W. Roberts, F.W. Stahl, R.A. Weisberg (eds.). 1983. Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- O.C. Uhlenbeck, J. Carey, P.J. Romaniok, P.T. Lowary, D. Beckett. 1983. Interaction of R17 Coat Protein with Its RNA Binding Site for Translational Repression. J. Biomol. Structure and Dynamics, **1**, 539-552.
- B. De Crombrughe, S. Busby, H. Buc. 1984. Cyclic AMP Receptor Protein: Role in Transcription Activation. Science, **224**, 831-838.
- S. Gottesman. 1984. Bacterial Regulation: Global Regulatory

- Networks. *Annu. Rev. Genet.*, **18**, 415-442.
- M. Nomura, R. Course, G. Baughman.* 1984. Regulation of the Synthesis of Ribosomes and Ribosomal Components. *Annu. Rev. Biochem.*, **53**, 75-117.
- C. Pabo, R. Sauer.* 1984. Protein-DNA Recognition. *Annu. Rev. Biochem.*, **58**, 293-321.
- D. Raibaud, M. Schwartz.* 1984. Positive Control of Transcription Initiation in Bacteria. *Annu. Rev. Genet.*, **18**, 173-206.
- M. B. Matthews* (ed.). 1986. Translational Control. *Current Communications in Molecular Biology*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- M. Ptashne.* 1986. A Genetic Switch, Gene Control and Phage λ. Cell Press and Blackwell Scientific Publications, Palo Alto, California. [Имеется перевод: Пташне М. Переключение генов. Регуляция генной активности и фаг λ. - М.: Мир, 1988.]
- C. Yanofsky.* 1988. Transcription Attenuation. *J. Biol. Chem.*, **263**, 609-612.

# ТРИУМФ РЕКОМБИНАНТНЫХ ДНК

## ВВЕДЕНИЕ

Примерно к 1970 г. стали известны основные свойства генетических систем. Несмотря на отсутствие многих важных деталей, удалось установить принципы репликации, рекомбинации и репарации и каждый из этих процессов был воспроизведен *in vitro*. Была сформулирована центральная догма, согласно которой генетическая информация передается от ДНК к РНК и далее к белку, что создало основу для определения генотипа и фенотипа организма на молекулярном уровне. Был идентифицирован основной посредник при переносе информации от ДНК к белку — информационная РНК (мРНК). Расшифрован генетический код, и в экспериментах с реконструированными клеточными компонентами в системе *in vitro* была получена информация о клеточном аппарате и основных механизмах трансляции мРНК в белок. Подтвердилось предположение о том, что процессы транскрипции ДНК в РНК и трансляции РНК в белок регулируются и что существуют позитивный и негативный способы контроля функций генов. С расшифровкой генетического кода разрешился имеющий долгую историю вопрос о связи между химической структурой гена и кодируемого им белка и стало ясно, что мутации есть следствие изменений в структуре ДНК. В этот период выдающихся открытий неожиданной наградой исследователям стала идентификация многих ферментов, для которых нуклеиновые кислоты являются субстратом. Получение их в очищенном виде и определение свойств в значительной мере облегчило анализ структуры и функций нуклеиновых кислот, а применение в дальнейших исследованиях привело к созданию новой области молекулярной биологии — технологии рекомбинантных ДНК.

Несмотря на широко распространенное мнение, что всем генетическим системам присущи одни и те же основные свойства, процессы, происходящие в клетках прокариот, изучены значительно глубже, чем процессы, протекающие в эукариотических организмах. Действительно, провести генетический

анализ небольших по размеру и менее сложно организованных бактериальных геномов значительно проще, чем геномов эукариот. Сравнительно легко удалось индуцировать и идентифицировать мутационные изменения в специфических генах. Случайный обмен генетической информацией между различными бактериями и некоторыми бактериями и их вирусами (бактериофагами) облегчил картирование этих генов, что в свою очередь позволило установить организацию бактериальных и фаговых геномов в целом. Еще более важное значение имело замечательное взаимопроникновение генетики и биохимии. Совместное использование генетических и биохимических методов способствовало разгадке сложного процесса репликации ДНК и даже позволило осуществить полноценную репликацию *in vitro* вирусных геномов. Благодаря объединению этих методов удалось получить отдельные гены в изолированном виде, что подготовило почву для изучения транскрипции и трансляции генов *in vitro* и идентификации молекулярных продуктов, участвующих в этих процессах. С помощью того же двустороннего подхода был установлен механизм регуляции экспрессии генов: было показано, что контроль осуществляется главным образом путем взаимодействия между специфичными белками и соответствующими регуляторными последовательностями в ДНК или информационной РНК.

В то же время успехи в расшифровке молекулярной структуры, организации и функций эукариотического генома были весьма скромными. Сложные генетические карты локусов, содержащих мутации, удалось составить лишь для тех немногих эукариотических организмов, с чьими генетическими системами можно было проводить манипуляции (к их числу относились некоторые виды дрожжей, *Neurospora crassa*, *D. melanogaster*). По сравнению с ними генетические карты млекопитающих, в частности мыши и человека, представлялись сплошными «белыми пятнами». Еще более загадочными

были молекулярная структура эукариотических генов и их организация в хромосомных ДНК, в частности наличие множественных повторов некоторых сегментов ДНК у большинства эукариот. Без более полного изучения молекулярной структуры геномов эукариот дальнейший прогресс в этой области был невозможен.

Биохимические исследования экспрессии и регуляции эукариотических генов также зашли в тупик из-за отсутствия информации о структуре клеточных генов. Было ясно, что ядерная ДНК эукариот транскрибируется в РНК и что мРНК транслируется в белки с помощью цитоплазматических комплексов рибосома–тРНК, во многом напоминающих прокариотические (гл. 3). Однако механизм транскрипции и последующая судьба транскриптов оставались совершенно загадочными. У многих эукариот только небольшая фракция ядерной РНК (<10%) переходит в цитоплазму в виде информационной, рибосомной и транспортной РНК. Некоторые молекулы существуют в стабильной форме в виде коротких цепей РНК в составе рибонуклеопротеиновых частиц, однако большинство из них быстро деградирует, не покидая ядра. Вопросы о природе быстро распадающейся РНК, ее происхождении и функции постоянно дебатировались. Проблема биогенеза информационной РНК также оставалась нерешенной из-за различий между мРНК про- и эукариот. У последних в начале и в конце молекул имеются особые структуры—так называемые экзы и poly(A)-хвосты соответственно. Это указывало на то, что информационная РНК эукариот в процессе биогенеза подвергается посттранскрипционным модификациям. Почему, где и как происходят такие модификации? Каково их значение? Каков путь превращения первичных транскриптов ДНК в зрелые молекулы информационной РНК? Далее, как осуществляется контроль транскрипции и превращений РНК, образованных на разных генах, в разных типах клеток одного и того же организма? Чем различаются механизмы экспрессии и регуляции генов у про- и эукариот? Из-за отсутствия молекулярно-генетической информации и методологии для ее получения эти вопросы первостепенной важности оставались нерешенными.

На фоне появления все новых данных об организации и экспрессии генетической информации у прокариот отсутствие таких данных для эукариот становилось все более ощутимым. Для преодоления такого отставания нужна была общая методология исследования клеточных геномов эукариот на молекулярном уровне. В идеале это позволило бы выделить дискретные гены и определить их молекулярную структуру и организацию геномов. При наличии таких изолированных генетических элементов можно было бы затем установить биохимичес-

кие основы механизмов транскрипции и трансляции. Объективные предпосылки к этому появились лишь в первой половине 70-х годов, когда была разработана технология получения рекомбинантных молекул ДНК. Прежде чем обсуждать эти достижения, мы рассмотрим концепции и методологию, которые подготовили почву для решающих экспериментов. Они ведут начало от экспериментов в сфере бактериальной генетики, и особенно большую роль здесь сыграла возможность введения молекул ДНК в клетки бактерий. Такая введенная ДНК независимо от того, произошла ли она из бактериофага или из других бактериальных клеток, изменяет генотип, а нередко и фенотип реципиентной клетки. Гены донорной ДНК способны экспрессироваться и могут рекомбинировать с хромосомной ДНК.

## ВВЕДЕНИЕ НОВОЙ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ В КЛЕТКИ БАКТЕРИЙ

Бактерии могут приобретать новый генетический материал несколькими способами. Это: 1) **трансформация**, при которой в клетки проникают молекулы ДНК, добавленные в культуральную среду; 2) **конъюгация**, в процессе которой ДНК непосредственно переносится от одной клетки к другой; 3) опосредуемая бактериофагами **трандукция**, при которой новая генетическая информация вводится в клетку с частицей бактериофага. Независимо от способа попадания в реципиентную клетку донорная ДНК рекомбинирует с гомологичными участками или специфическими сайтами в геноме реципиентного организма или сохраняется в виде автономной мини-хромосомы, изменяя таким образом генотип хозяина.

### Трансформация бактерий

Трансформация, т.е. изменение генотипа клетки путем внесения в нее молекул ДНК из культуральной среды, была первым из способов введения новых генов в клетки бактерий. Это послужило также первым доказательством роли ДНК как носителя генетической информации. Если в клетки организма с определенным генетическим нарушением (например, организма, не способного синтезировать триптофан, сбрасывать галактозу, образовывать какой-то полисахарид и т.п.) ввести ДНК, выделенную из нормальных клеток, то у этого организма нередко восстанавливаются утраченные функции. Такая трансформация является обычно наследственной и стабильной, поскольку в ее основе лежит

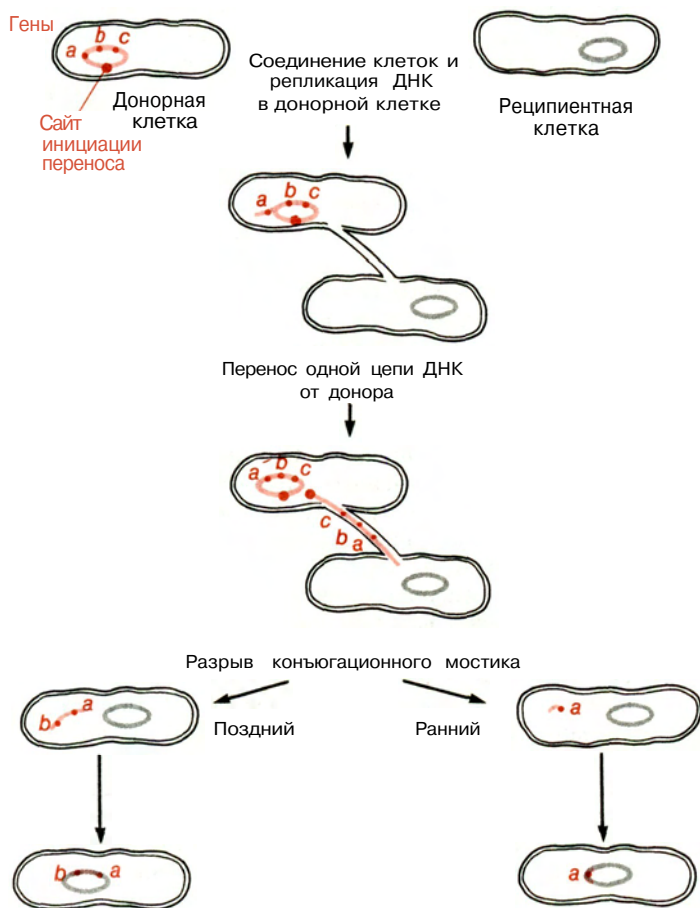
рекомбинация между функциональным геном (генами) донорной ДНК и дефектным геном (генами) реципиента. Однако осуществляемая с помощью ДНК трансформация оказалась полезной только при изучении молекулярной генетики прокариот. В других случаях возможности трансформации ограничивались трудностями выявления трансформирующего гена, что делало нереальным определение его структуры. Тем не менее принцип трансформации нашел применение в другой области. Например, получение трансформированных клеток является важным этапом во многих экспериментах с рекомбинантными молекулами ДНК. Термин «трансформация» используется в молекулярной биологии эукариот для обозначения стабильного изменения генотипа и фенотипа клетки.

## Конъюгация

При конъюгации осуществляется прямой перенос ДНК из одной клетки в другую при их контактировании. В тех случаях, когда конъюгация происходит между определенными штаммами *E. coli*, один из

них выполняет функции донора, другой—реципиента (рис. II.1). Хромосомная ДНК донорной клетки переходит в клетку реципиента по мостику, образуемому между двумя клетками. Полный перенос ДНК осуществляется примерно за 90 мин, однако в случае повреждения мостика конъюгация прерывается и перенос хромосомы оказывается неполным.

Разные штаммы *E. coli* начинают перенос с разных точек хромосомы. Следовательно, у различных доноров ранними или поздними генами во время конъюгации оказываются разные группы генов. Однако перенос генов всегда происходит только в одном из двух возможных противоположных направлений кольцевой хромосомы *E. coli* (рис. II.2). Эти эксперименты впервые показали, что все гены *E. coli* расположены на одной кольцевой молекуле ДНК (рис. II.3). Сравнивая время, необходимое для переноса различных генов во время конъюгации, можно построить генетическую карту хромосомы *E. coli*, т.е. установить порядок следования хромосомных генов.



**РИС. II.1.**

Перенос генетической информации между бактериями при конъюгации. Две клетки *E. coli*, из которых одна передает свою хромосому другой, вступают в контакт с помощью белкового конъюгационного мостика. Репликация донорной ДНК начинается в особом участке (отмечен точкой), и одна из цепей ДНК переносится в клетку реципиента. Перенос прекращается при разрыве конъюгационного мостика (при случайном движении клеток или встряхивании сосуда, в котором эти клетки находятся). Чем продолжительнее контакт двух клеток до разрыва мостика, тем большее число генов переносится. Перенесенная ДНК может с помощью рекомбинации заменить соответствующую часть генома реципиентной клетки. У разных штаммов *E. coli* перенос инициируется в различных хромосомных локусах.

Донорный штамм	Порядок переноса генов
1	● <i>thr-leu-azi-ton-pro-lac-pur-gal-trp-his-gly-str-mal-xyl-mtl-ile-met-thi</i>
2	● <i>met-thi-thr-leu-azi-ton-pro-lac-pur-gal-trp-his-gly-str-mal-xyl-mtl-ile</i>
3	● <i>str-mal-xyl-mtl-ile-met-thi-thr-leu-azi-ton-pro-lac-pur-gal-trp-his-gly</i>
4	● <i>leu-thr-thi-met-ile-mtl-xyl-mal-str-gly-his-trp-gal-pur-lac-pro-ton-azi</i>
5	● <i>his-trp-gal-pur-lac-pro-ton-azi-leu-thr-thi-met-ile-mtl-xyl-mal-str-gly</i>
6	● <i>mtl-xyl-mal-str-gly-his-trp-gal-pur-lac-pro-ton-azi-leu-thr-thi-met-ile</i>

### РИС. II.2.

Перенос ДНК при конъюгации начинается с разных генов в зависимости от штамма донора. Например, штамм 1 начинается перенос с гена *thr*, а штамм 2 с гена *met*. Группы генов всегда переносятся в одном из двух

взаимно противоположных направлений (например, в штаммах 1, 2 и 3 порядок переноса - *thr-leu-azi*, в штамме 4 перенос начинается между *azi* и *leu*; в штаммах 5 и 6- *azi-leu-thr*, и т.д.)

## Трансдукция

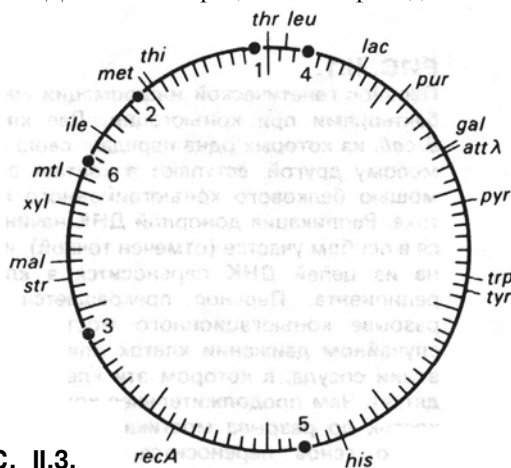
Охарактеризованы два типа трансдукции, осуществляемой с помощью бактериофагов, **общая** и **специфическая**. При общей трансдукции фаговые частицы, содержащие сегменты ДНК клетки-хозяина, переносят относительно протяженные участки геномной ДНК от одной бактериальной клетки к другой. Трансдуцирующие фаговые частицы образуются в ходе определенных инфекционных процессов, когда ДНК клетки эффективно деградирует и фрагменты, по размеру примерно соответствующие фаговому геному, случайно упаковываются в зрелые частицы бактериофага (рис. II.4). В результате последующего инфицирования клеток бактерий популяцией фаговых частиц, содержащей трансдуцирующие фаги, происходит передача ДНК донорных клеток этим инфицируемым клеткам. Рекомбинация между введенными фрагментами донорной ДНК и ДНК клетки-реципиента приводит к измене-

нию генотипа последней. Каждая трансдуцирующая фаговая частица обычно содержит только один случайный фрагмент исходной донорной хромосомы. Вероятность включения в такую частицу любой части этого генома примерно одинакова. Однако благодаря довольно большому размеру трансдуцируемых сегментов ДНК (для определенных бактериофагов он составляет около 100 т.п.н., или 2,5% всей хромосомы *E. coli*) обычно реципиентная клетка приобретает за один акт трансдукции целую группу генов. В результате гены, тесно сцепленные друг с другом в хромосоме донора, с высокой частотой котрансдуцируются, тогда как гены, удаленные друг от друга, трансдуцируются независимо (рис. II.5). Определение частоты котрансдукции генов помогает уточнить генетические карты, позволяя оценить относительные расстояния между тесно сцепленными генами.

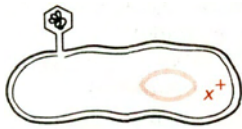
Трансдукция второго типа, **специфическая**, свойственна бактериофагам, инфекционный цикл которых прерывается в результате включения генома вируса в специфический хромосомный локус ДНК инфицированной клетки (гл. 2). Бактерии, содержащие такие интегрированные фаговые геномы, получили название **лизогенных**. Они несут вирусные геномы как наследственные элементы в своих собственных хромосомах (рис. II.6). В лизогенной клетке вирусные и клеточные геномы реплицируются как единое целое и являются взаимно совместимыми. Интеграция фагового генома с геномом клетки-хозяина лишает фаг возможности вызывать гибель клетки и продуцировать инфекционное потомство. По этой причине бактериофаг, способный лизогенизировать, в отличие от **вирулентного** фага получил название **умеренного**. При определенных условиях — индукции — лизогенное состояние прерывается и вирусный геном вырезается из хромосомы клетки-хозяина. Он реплицируется, образует множество вирусных частиц и убивает клетку. Обычно вырезание вирусного генома происходит очень точно, и образующийся фаг содержит вирусный геном, полно-

### РИС. II.3.

Кольцевая генетическая карта *E. coli*, на которой указано положение отдельных генов. Порядок генов в хромосоме соответствует указанному на рис. II.2. Точками отмечены области инициации переноса в донорных штаммах, представленных на рис. II.2.



Образование  
трандуцирующих  
частиц



Лизис



Жизнеспособный фаг

+



Трандуцирующий фаг

Фаг,  
несущий  
ген  $x^+$

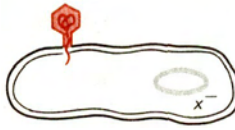
Бактерия,  
содержащая ген  $x^+$

Трандукция  
 $x^-$ -бактерии  
 $x^+$ -фагом

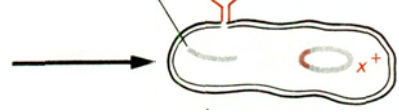


Бактерия,  
несущая ген  $x^-$

Инфицирование



Фрагмент  $x^-$   
(будет деградирован)



$x^-$ -Трансдуктант

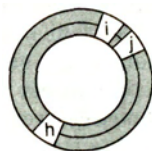
**РИС. II.4.**

Образование фаговых частиц, осуществляющих общую трандукцию, и введение донорной ДНК в инфицируемую клетку. Один из трандуцирующих фагов несет бактериальный ген  $x^+$  дикого типа, который при последующем инфицировании вводится в бактериальную клетку, несущую мутантный ген  $x^-$ . В результате рекомбинации эта клетка превращается в  $x^+$ -трансдуктант.

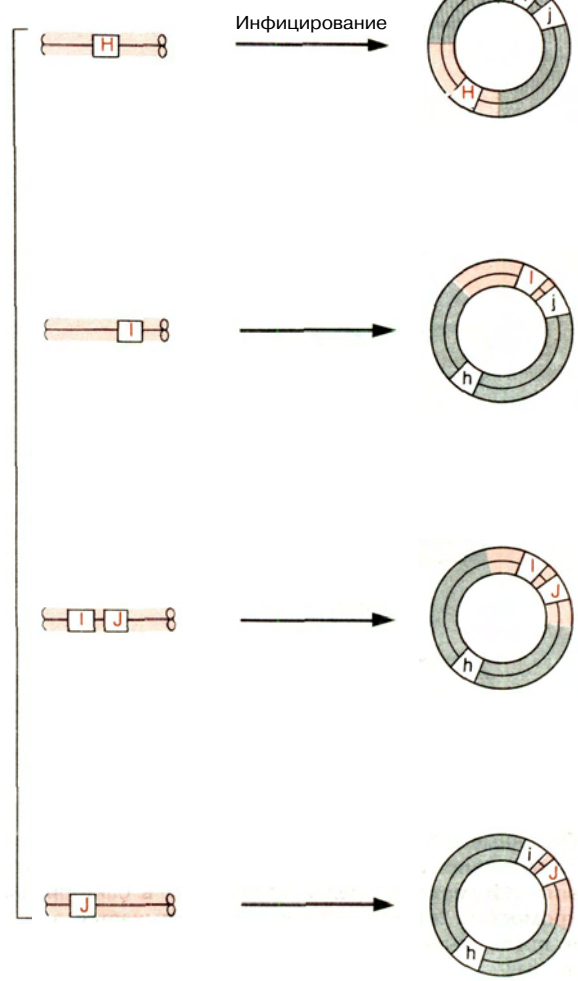
Геном  
бактерии-  
реципиента

Смесь фрагментов  
трандуцирующей  
фаговой ДНК

Бактериальные  
трансдуктанты



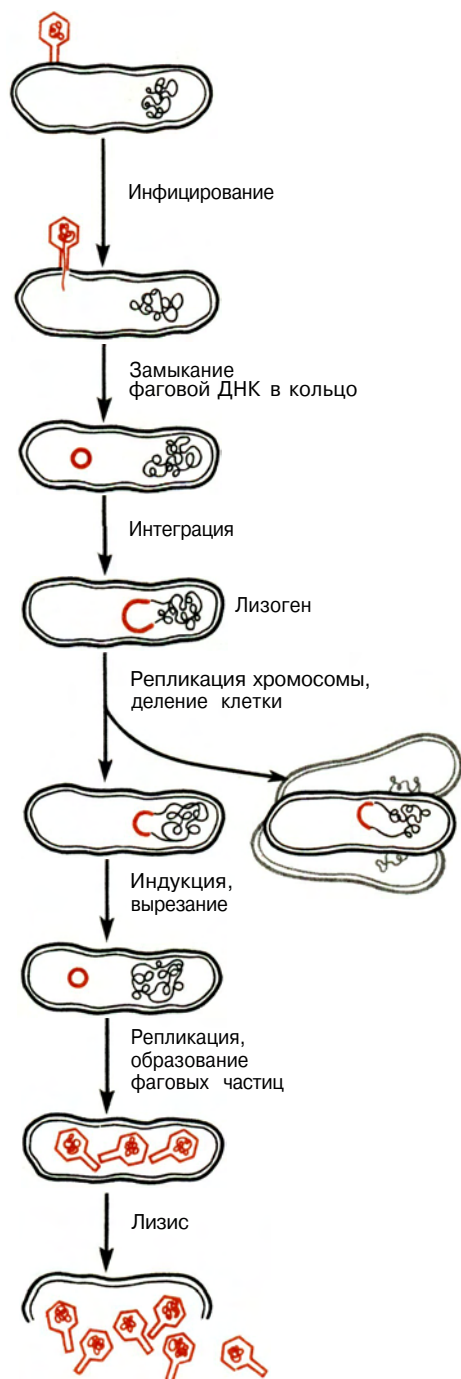
+



**РИС. II.5.**

Зависимость частоты котрансдукции генов от расстояния между ними. Гены  $i$  и  $j$  могут транслироваться как независимо, так и совместно, поскольку они тесно сцеплены друг с другом в бактериальной хромосоме. Ген  $h$  находится достаточно далеко от обоих генов и не может быть котранслирован с ними.



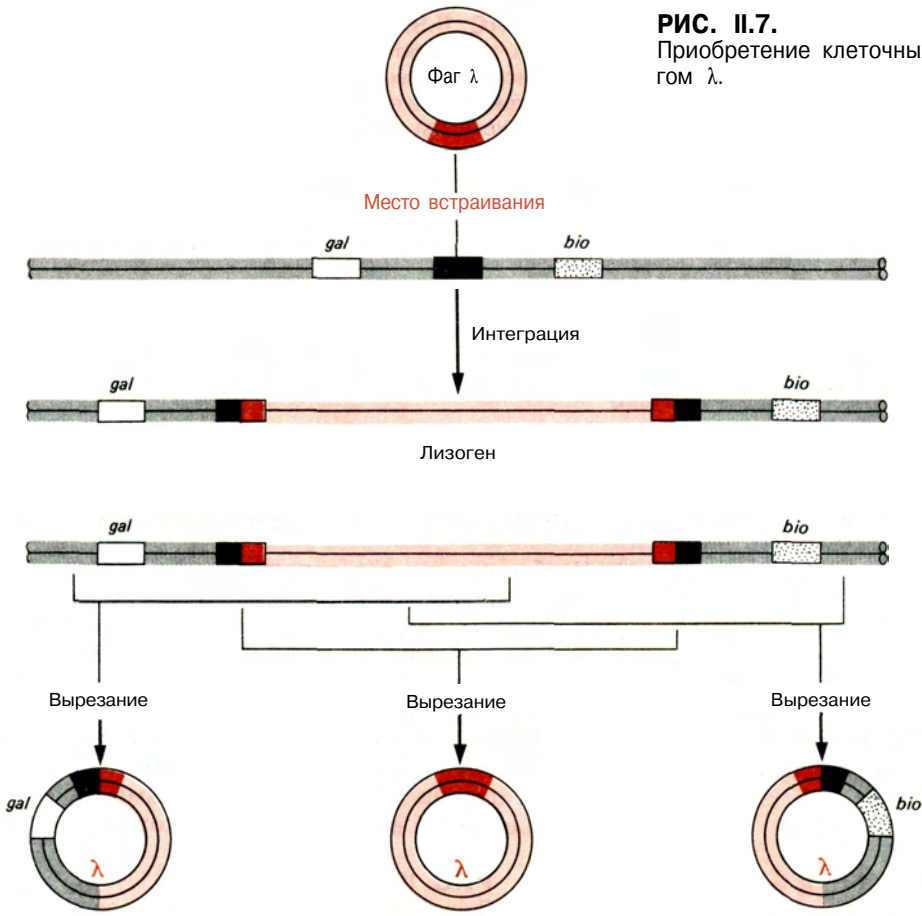
**РИС. II.6.**

Основные события, ведущие к установлению лизогенного состояния, индукции лизогенов и образованию фагового потомства. Перед встраиванием в бактериальную хромосому инфицирующая ДНК замыкается в кольцо и после вырезания тоже превращается в кольцевую форму.

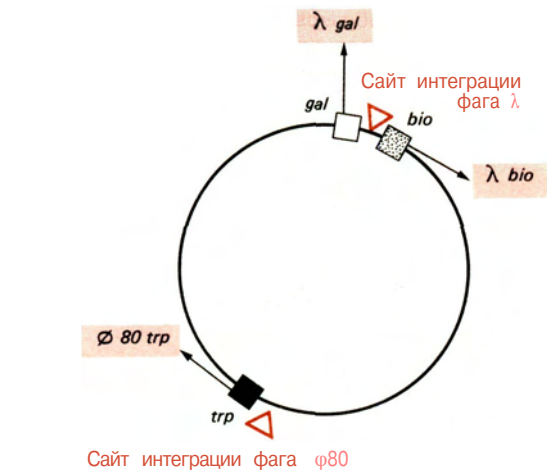
стью соответствующий исходному. Иногда, однако, фаговый геном вырезается неправильно и в дочерние фаговые геномы включаются хромосомные гены, прилегавшие к интегрированному вирусному геному. Эти гены включаются вместо некоторых вирусных генов (рис. II.7). Во время следующего цикла инфекции гены клетки-донора переходят вместе с фаговыми генами в реципиентные клетки. После включения ДНК трансдуцирующего фага в геном реципиента клетка приобретает наряду с фаговыми генами генетическую информацию предыдущего хозяина фага. Таким образом, при специфической трансдукции фаг служит вектором для переноса генов из одной клетки в другую. С помощью этого механизма трансдуцируются только те хромосомные гены клетки-хозяина, которые тесно сцеплены с сайтом интеграции вирусного генома.

Поскольку различные умеренные фаги встраиваются в разные хромосомные сайты, при их неправильном вырезании образуются фаги, которые трансдуцируют разные хромосомные гены. Так, фаги  $\lambda$  трансдуцируют гены, ответственные за метаболизм галактозы ( $\lambda gal$ ), или гены, контролирурующие синтез биотина ( $\lambda bio$ ), а фаги  $\phi 80$  — различное число генов, кодирующих ферменты биосинтеза триптофана (рис. II.8). Были разработаны определенные генетические приемы, способствующие приобретению этими и другими фагами различных генов *E. coli* или генов родственных организмов. Та же стратегия с небольшими модификациями позволяет получить трансдуцирующие фаги, содержащие мутантные бактериальные гены. Такие трансдуцирующие фаги легко идентифицировать, размножить и очистить, что позволяет получать значительные количества аллелей дикого или мутантного генов *E. coli* в высокоочищенном виде.

Обогащение бактериальными генами, сопровождающее их включение в геном трансдуцирующих фагов, имеет важные последствия. Рассмотрим, например, ген *E. coli*, кодирующий фермент  $\beta$ -галактозидазу (*lac Z*). Этот белок состоит из идентичных полипептидных цепей длиной 1173 аминокислотных остатка; следовательно, ген, кодирующий этот полипептид, содержит около 3600 пар нуклеотидов. Ген  $\beta$ -галактозидазы составляет одну тысячную генома *E. coli* ( $3,6 \cdot 10^3$  из  $4,0 \cdot 10^6$  п.н.), но в геноме трансдуцирующего вируса  $\lambda lac$  он занимает 1/15 часть ( $3,6 \cdot 10^3$  из  $5 \cdot 10^4$  п.н.). Таким образом, ДНК фага  $\lambda lac$  обогащена  $\beta$ -галактозидазным геном примерно в 100 раз больше, чем ДНК *E. coli*. Это упрощает выделение  $\beta$ -галактозидазного гена и позволяет идентифицировать его регуляторные участки и определить их нуклеотидную последовательность. Подобным же образом с получением трансдуцирующих фагов  $\phi 80 trp$  удалось выделить и охарактеризовать гены и регуляторные последо-



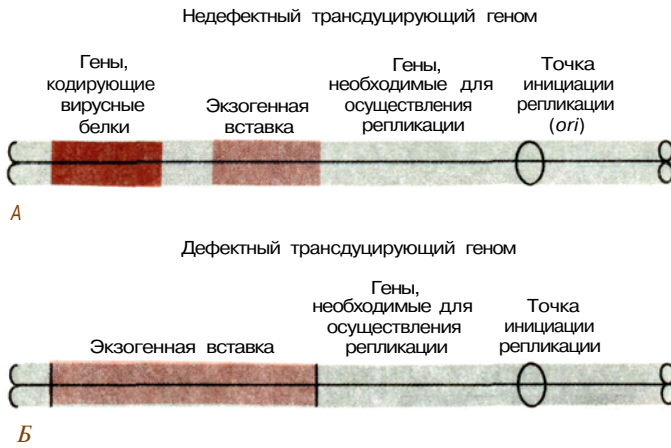
**РИС. II.7.** Приобретение клеточных генов трансдуцирующим фагом  $\lambda$ .



**РИС. II.8.** Различные фаги, осуществляющие специфическую трансдукцию, переносят только те бактериальные гены, которые расположены вблизи сайта их интеграции в хромосому. Показаны сайты интеграции фагов  $\lambda$  ( $\lambda att$ ) и  $\phi 80$  ( $\phi 80 att$ ).

вательности, которые образуют триптофановый оперон (разд. 3.11.г). Кроме того, благодаря трансдукции появилась возможность изучать влияние различных мутационных изменений на экспрессию и регуляцию генов *in vivo*.

Рассмотрим те свойства фагового генома, которые ответственны за его способность к специфической трансдукции (рис. II.9). Во-первых, геном должен быть способен реплицироваться после того, как произошла инфекция [т.е. в вирусной ДНК должны сохраняться область начала репликации (*ori*) и гены, необходимые для осуществления репликации]. Во-вторых, он должен приобрести ковалентно сцепленный сегмент невирусной ДНК, который будет трансдуцироваться. Этот сегмент ДНК обычно имеет клеточное происхождение, но в принципе он может быть из любого источника. Он может включиться в любое место вирусного генома, если это не влияет на репликацию вирусной ДНК в инфицированной клетке хозяина или на ее способность упаковываться в зрелые фаговые частицы. Будучи составной частью фагового генома, трансдуцируемый сегмент ДНК реплицируется вместе

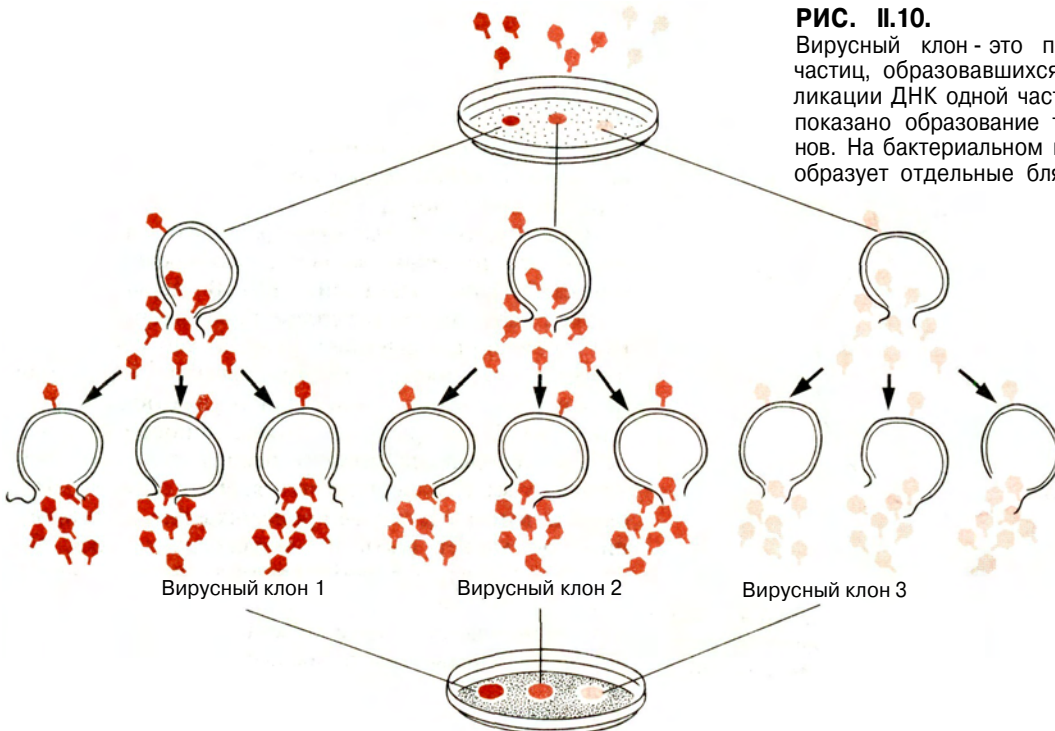
**РИС. II.9.**

А. Недефектный вирусный геном, содержащий все гены, необходимые для репликации и упаковки вирусной ДНК. Б. Дефектный геном, в котором отсутствуют гены, кодирующие вирусные белки. Такие белки должны кодироваться генами клетки-хозяина или коинфицирующего вируса.

с вирусной ДНК. В-третьих, гены, кодирующие структурные фаговые белки, должны быть функционально активными либо их роль должен выполнять коинфицирующий фаг или клетка-хозяин. И наконец, если мы хотим использовать трансдукцию, нам нужно найти способ разделения различных типов вирусных геномов и идентификации интересующего нас генома, поскольку трансдуцирующие вирусы часто инфицируют клетку совместно с вирусом дикого типа. Обычно для такого разделения используют **клонирование**.

## ПРИНЦИПЫ КЛОНИРОВАНИЯ

Чтобы понять, как использовались концепции трансдукции при разработке методов получения рекомбинантных ДНК, нам следует ознакомиться с тем, что представляет собой клонирование. Клон вируса или клеток — это некая популяция, каждый член которой ведет происхождение от одного репродуцирующегося вириона или от одной клетки соответственно. Все члены клона независимо от того, являются ли они вирусами или клетками, по существу идентичны вирусу или клетке, которые дали

**РИС. II.10.**

Вирусный клон — это популяция вирусных частиц, образовавшихся в результате репликации ДНК одной частицы. Схематически показано образование трех вирусных клонов. На бактериальном газоне каждый клон образует отдельные бляшки.

начало клону; они также идентичны друг другу. При клонировании вирусов необходимо, чтобы вирусное потомство из одной клетки, инфицированной одной вирусной частицей, размножалось в течение многих циклов инфекции, не смешиваясь с потомством из других инфицированных клеток (рис. II.10). Такие вирусные клоны образуют прозрачные зоны (или бляшки) на монослое (или газоне) неинфицированных клеток. Клонирование клеток можно осуществить только в том случае, если клетки при размножении остаются изолированными друг от друга (рис. II.11). Клоны бактериальных клеток или клеток млекопитающих легко образуются при разреженном посеве их на чашке таким образом, что они образуют отдельные колонии.

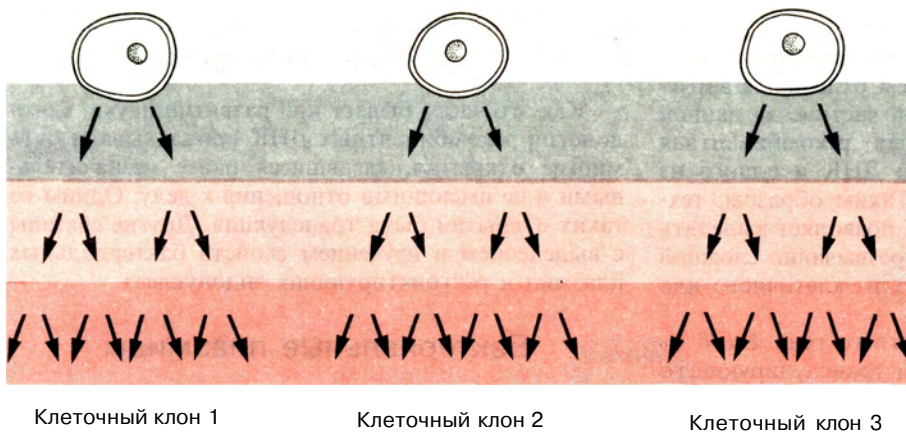
С помощью клонирования получают чистый препарат одного генома, поскольку все члены клона содержат идентичные ДНК. Эта концепция молекулярного клонирования используется и при получении чистых препаратов определенных молекул рекомбинантных ДНК.

## КОНЦЕПЦИЯ РЕКОМБИНАНТНОЙ ДНК

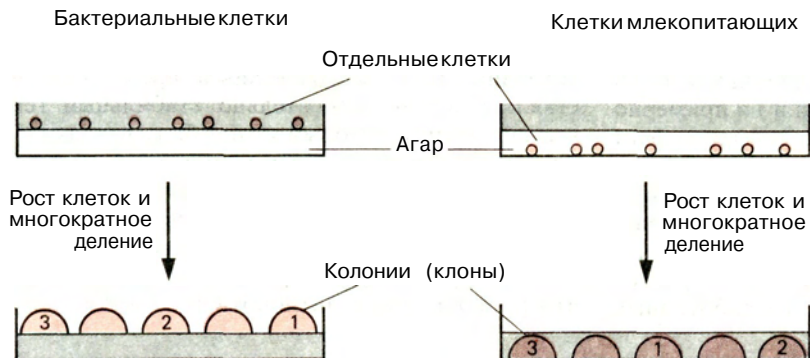
Методология получения рекомбинантных ДНК основана на тех же принципах, что и трансдукция.

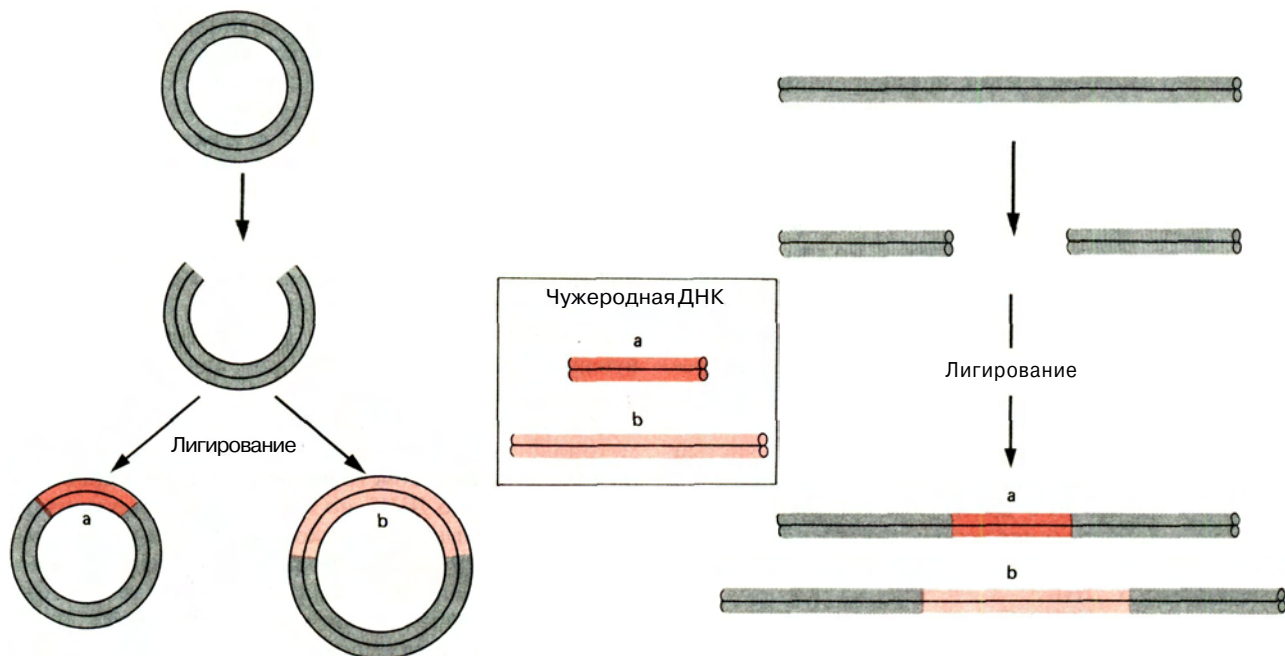
Молекулы ДНК, способные реплицироваться в соответствующих клетках, представленные вирусными геномами или плазмидами (см. ниже), служат переносчиками, или **векторами**, «чужеродных» сегментов ДНК, получивших название **вставок**. При этом, вместо того чтобы полагаться на клеточные процессы, ведущие к образованию рекомбинантных трансдуцирующих геномов, проводят объединение, или рекомбинацию, соответствующим образом модифицированных вставок и векторов *in vitro* с помощью фермента ДНК-лигазы (рис. II.12). Такие рекомбинантные ДНК вводят затем в соответствующие клетки, где они амплифицируются в результате репликации.

Громадные потенциальные возможности этой методологии обусловлены не только тем, что с ее помощью можно конструировать и реплицировать рекомбинантную ДНК, но и тем, что она позволяет клонировать отдельные рекомбинантные молекулы ДНК. Рассмотрим, например, что получается в результате объединения смеси случайных сегментов ДНК какого-либо организма с векторной ДНК (рис. II.13). Ассортимент всех возможных рекомбинантов чрезвычайно разнообразен; каждый рекомбинант содержит какой-то определенный сегмент



**РИС. II.11.** Клеточный клон - это популяция клеток, полученная в результате множественных клеточных делений от одной клетки. Каждый клон представлен отдельной колонией.





**РИС. II.12.**  
Встраивание чужеродных фрагментов ДНК (a и b) в кольцевую и линейную молекулы вирусной ДНК in vitro.

исходной ДНК. Однако при клонировании рекомбинантных ДНК с образованием отдельных вирусных бляшек в каждой вирусной частице из данной бляшки содержится уникальная рекомбинантная ДНК, состоящая из векторной ДНК и одного из сегментов исходного генома. Таким образом, технология рекомбинантных ДНК позволяет выделять отдельные сегменты ДНК из чрезвычайно сложной смеси сегментов, происходящих из клеточного или вирусного генома.

В результате переноса какого-то гена *E. coli* из бактериального генома в геном трансдуцирующего фага можно получить примерно 100-кратное обогащение по этому гену. По сравнению с тем уровнем обогащения, которого удастся достичь путем молекулярного клонирования сегментов ДНК сложных организмов, такое обогащение является незначительным. Например, ген млекопитающих длиной 5 т.п.н. составляет лишь одну миллионную часть всего генома (5 т.п.н. от  $\sim 3 \cdot 10^6$  т.п.н.) и примерно одну десятую часть рекомбинантного генома фага  $\lambda$ . Таким образом, молекулярное клонирование позволяет расщеплять даже самые большие и самые сложно организованные геномы и получать отдельные сегменты, содержащие один или несколько генов в чистом виде. Такое относительно простое применение принципов и методов, разработанных

при изучении молекулярной генетики прокариот, устранило барьер, который препятствовал проведению подобного анализа на эукариотических геномах.

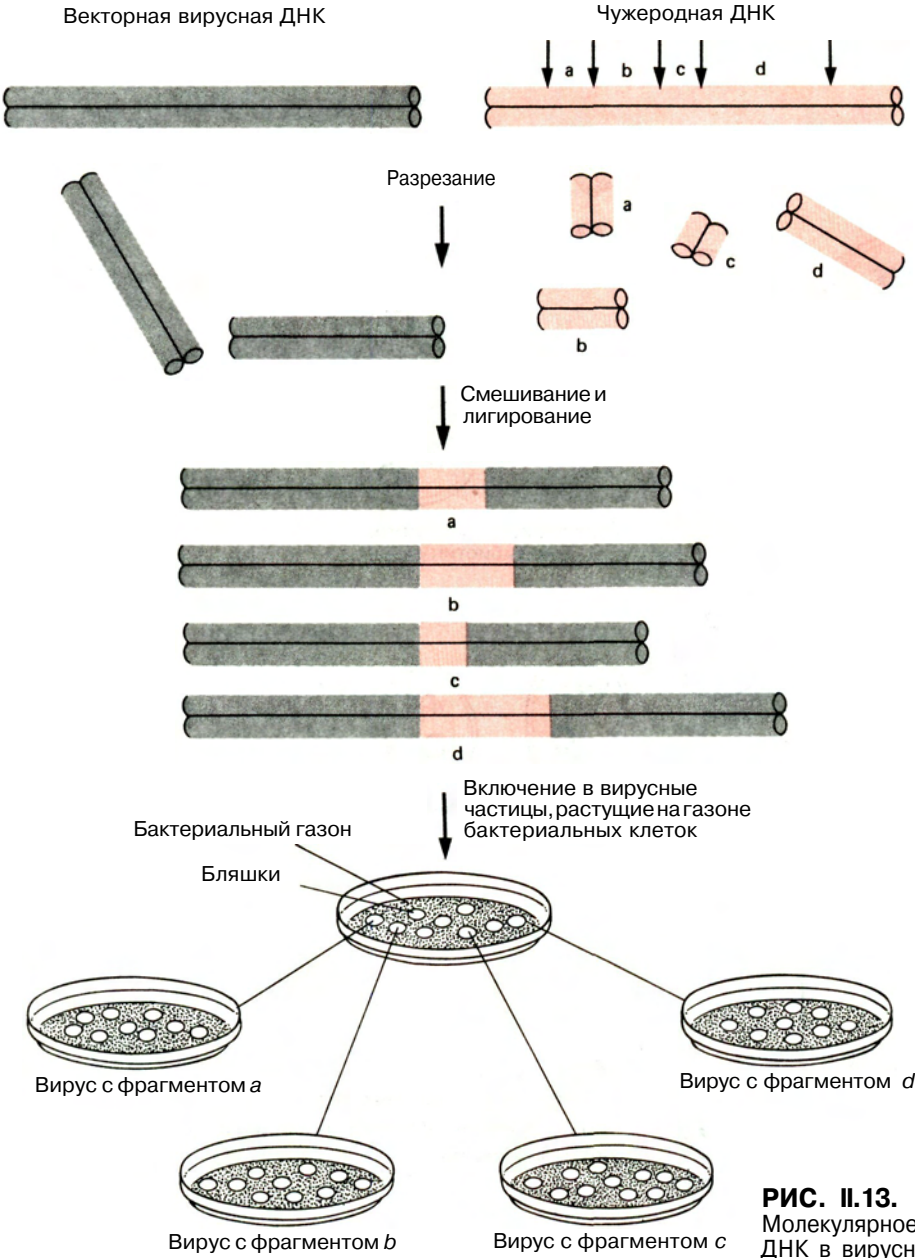
## ВАЖНЫЕ ОТКРЫТИЯ

Как это часто бывает при развитии науки, методологии рекомбинантных ДНК прокладывали путь многие открытия, казавшиеся ранее незначительными и не имеющими отношения к делу. Одним из таких открытий была трансдукция. Другие связаны с выделением и изучением свойств бактериальных плазмид и рестриктирующих эндонуклеаз.

### Бактериальные плазмиды

Одним из неожиданных следствий использования антибиотиков для лечения инфекционных заболеваний было появление устойчивых штаммов патогенных микроорганизмов. Эта очень важная проблема требовала своего решения и стимулировала интенсивные исследования. Вскоре стало ясно, что устойчивость к лекарственным препаратам представляет собой относительно стабильный генетический признак, который может быть передан чувствительным бактериальным клеткам способом, напоминающим инфекционный процесс. Позже было установлено, что распространение резистентности к антибиотикам происходит при контактировании клеток друг с другом (рис. II.14). Было показано, что резистентные к антибиотикам клетки содержат

ВВЕДЕНИЕ



**РИС. II.13.** Молекулярное клонирование чужеродной ДНК в вирусном (фаговом) векторе.

генетические элементы—плазмиды, не связанные с хромосомной ДНК, способные реплицироваться независимо от хромосомы и передаваться от клетки к клетке при их контактировании. Именно такие внехромосомные элементы и содержат гены, которые придают клеткам наследуемую устойчивость к одному или нескольким антибиотикам. Они получили название факторов резистентности, или **Р-факторов**.

Механизм распространения Р-факторов стал

более понятен после того, как был установлен способ переноса генетической информации от одной клетки к другой во время конъюгации бактерий. Из рис. II.1 видно, что перенос генетического материала осуществляется благодаря способности донорных клеток реплицироваться и переносить свою геномную ДНК через конъюгационный мостик в клетки реципиента. Штаммы *E. coli*, являющиеся донорами, содержат в составе своего генома ДНК плазмидного происхождения, названную фактором

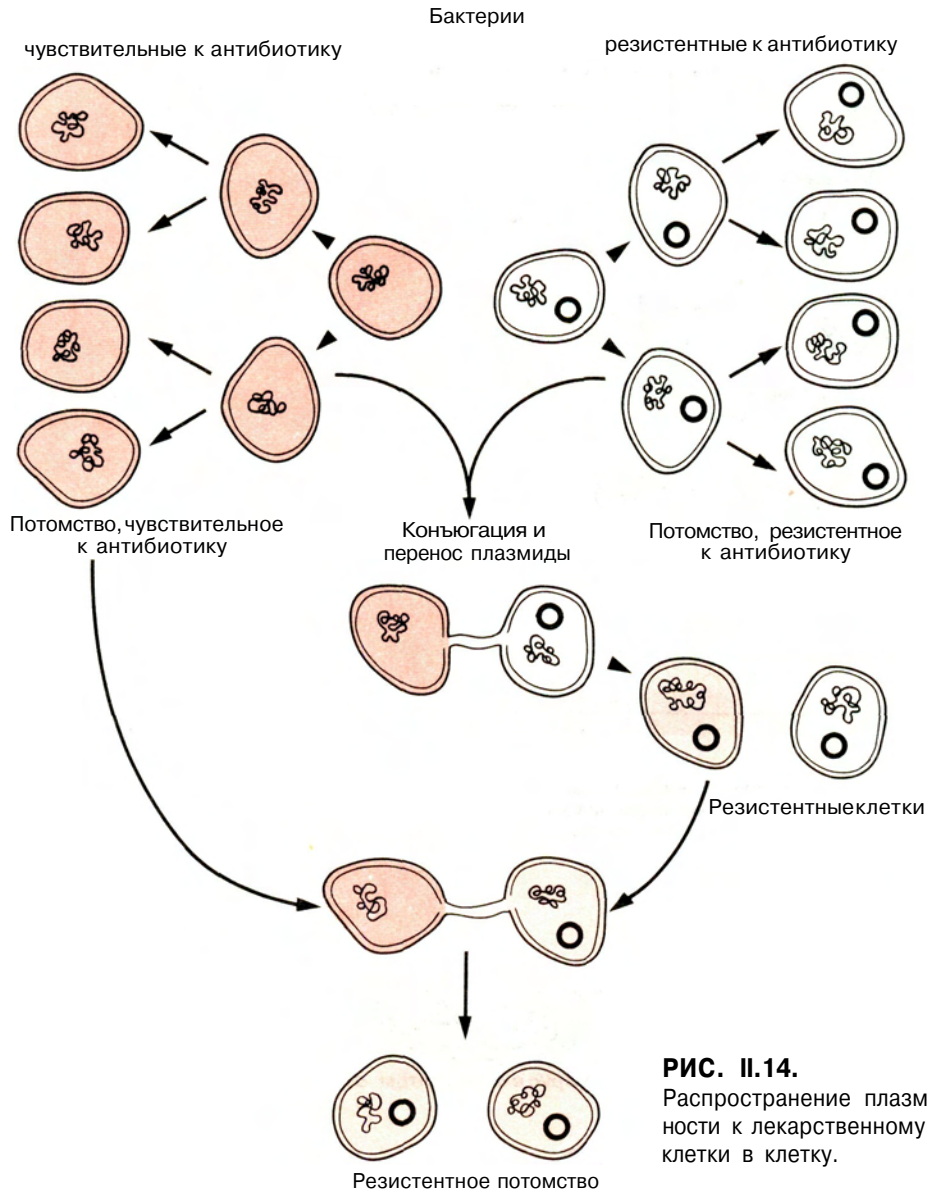


РИС. II.14.

Распространение плазмиды, несущей гены резистентности к лекарственному препарату, при переносе ее из клетки в клетку.

фертильности, или **F-фактором**. Такие донорные хромосомы образуются после приобретения клетками F-фактора и рекомбинации между этой плазмидой и хромосомой клетки. В присутствии интегрированного F-фактора облегчаются конъюгация и перенос хромосомы. Перенос инициируется в сайте интеграции F-фактора. Благодаря способности плазмиды встраиваться в ДНК во многих сайтах разные штаммы инициируют перенос в различных сайтах хромосомы *E. coli* (рис. II.1 и II.2). В некоторых случаях F-фактор остается в клетках в виде независимого внехромосомного элемента – F-плазмиды (рис. II.15). Такие клетки, обозначаемые  $F^+$ , перено-

сят плазмиду F в реципиентные клетки таким же способом, как в случае переноса R-факторов.

R- и F-факторы представляют собой ковалентно замкнутые кольцевые молекулы двухцепочечной ДНК. Оба они содержат гены, обеспечивающие их репликацию в виде автономных плазмид и их перенос при контакте с соответствующими реципиентами. R-факторы несут также гены резистентности к антибиотикам. Одни такие гены изменяют реакцию клетки на антибиотик, другие индуцируют образование белков, вызывающих деградацию или модификацию определенных антибиотиков. Например, одна из R-плазмид кодирует  $\beta$ -лактамазу, обес-



РИС. II.15.

Перенос F-фактора между бактериями при конъюгации.

печивающую устойчивость к ампициллину благодаря деградации этого антибиотика. Устойчивость к хлорамфениколу обуславливается ацелированием этого антибиотика. Ацелирование катализируется ферментом хлорамфеникол-ацетилтрансферазой, кодируемым плазмидой.

Плазмидные ДНК легко выделить в больших количествах и получить в очищенном виде, что позволяет исследовать их потенциальную способность служить векторами для введения в клетки новых сегментов ДНК. Соединение плазмидной ДНК с соответствующим образом модифицированными вставками можно осуществить *in vitro* аналогично тому, как это происходит в случае соединения вставок с фаговыми векторами (рис. II.12). В этих экспериментах большую роль сыграл опыт осуществления трансформации клеток с помощью ДНК. Были разработаны методы включения очищенной плазмидной ДНК в клетки соответствующих бактерий-хозяев. Благодаря наличию в таких плазидах генов резистентности к антибиотикам можно провести отбор клеток, содержащих R-плазмиду в качестве стабильно реплицирующейся структуры. Такие клетки живут и делятся в присутствии антибиотика, а клетки, утратившие эту плазмиду, погибают. Рекомбинантные плазмидные ДНК сохраняются в присутствии антибиотика как некие автономно реплицирующиеся геномы во время последующих клеточных делений. При клонировании образуются колонии клеток, содержащих уникальную рекомбинантную ДНК, потому что каждая клетка хозяина содержит только одну плазмиду. Поскольку размер многих плазмидных векторов не превышает нескольких тысяч пар нуклеотидов, обогащение по сегментам эукариотической ДНК в этом случае оказывается более высоким, чем при использовании фаговых векторов. Именно плазмиды оказались первыми векторами, с помощью которых было осуществлено молекулярное клонирование в клетках бактерий.

### Рестриктирующие эндонуклеазы

Одним из наиболее важных результатов генетического изучения бактерий и их бактериофагов явилось открытие рестриктирующих эндонуклеаз —

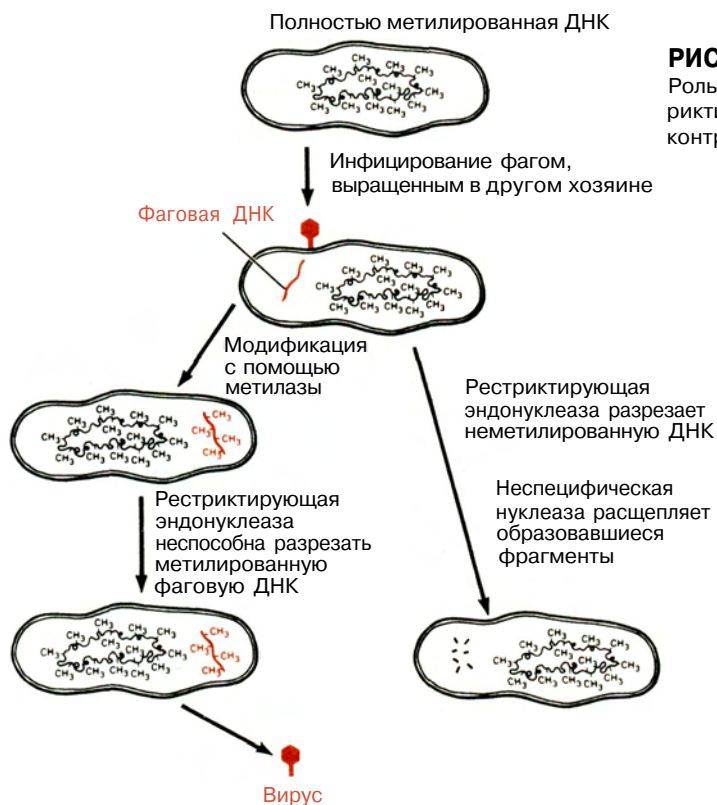
ферментов, которые узнают специфические короткие нуклеотидные последовательности и разрезают обе цепи двойной спирали ДНК в сайте узнавания или на некотором расстоянии от него. В открытии этих ферментов ключевую роль сыграло наблюдение, сделанное примерно 30 лет назад. Оно состояло в том, что бактериофаг, выращенный в клетках одного штамма, инфицируя клетки других штаммов этого же вида, часто растет очень плохо. Более того, выделенные после такой неэффективной инфекции бактериофаги в свою очередь плохо развиваются в исходных клетках (рис. II.16). Этот феномен не связан с генотипом фага и был объяснен тем, что в фаге происходят какие-то модификации, контролируемые хозяином. Было высказано предположение, что некий фаговый компонент, необходимый для репликации, специфическим образом модифи-



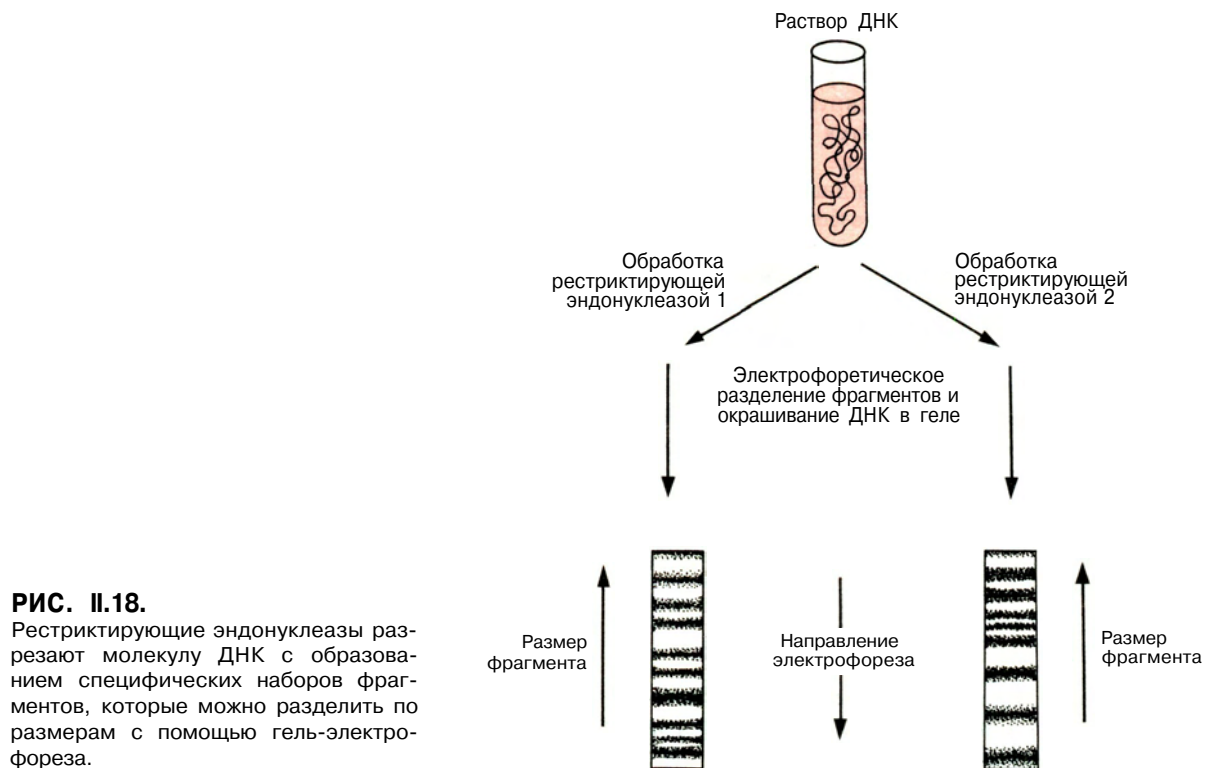
РИС. II.16.

Пример контролируемой хозяином рестрикции.



**РИС. II.17.**

Роль модифицирующей метилазы и рестриктирующей эндонуклеазы в рестрикции, контролируемой хозяином.

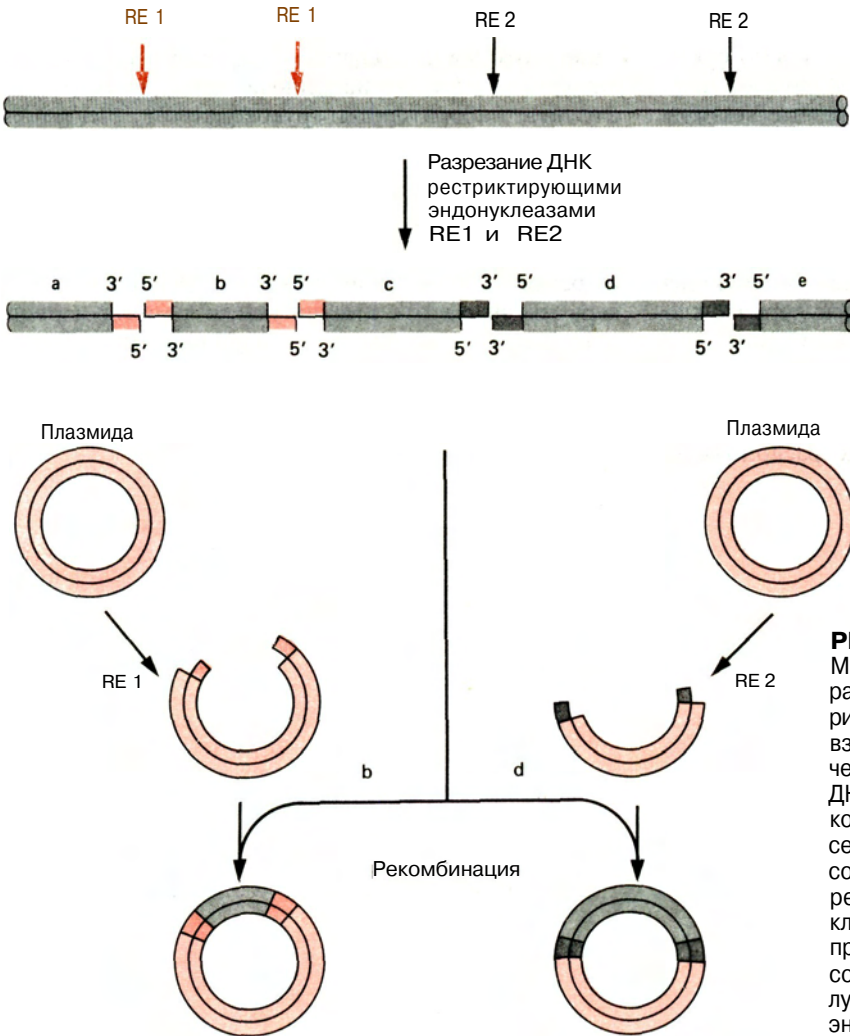
**РИС. II.18.**

Рестриктирующие эндонуклеазы разрезают молекулу ДНК с образованием специфических наборов фрагментов, которые можно разделить по размерам с помощью гель-электрофореза.

цируется в клетках хозяина, так что фаг может завершить репликацию при повторном инфицировании того же штамма. При этом его рост в неродственном штамме ограничивается, поскольку последний не содержит соответствующей системы модификации.

Было доказано, что модифицируемым фаговым компонентом является ДНК, а неспособность фага реплицироваться в неродственном штамме обусловлена деградацией инфицирующей фаговой ДНК (рис. II.17). Расщепление ДНК инициируется внесением нескольких разрывов в высокоспецифичных сайтах, после чего происходит полная неспецифическая деградация. Модификацией, защищающей некоторые инфицирующие фаговые ДНК, а также геном клетки от разрывов, является штамм-специфичное метилирование ДНК. С помощью генетических и биохимических исследований установлено, что метилирование ДНК происходит в строго

специфических участках. Рестрикция осуществляется эндонуклеазами, которые узнают аналогичные короткие специфические последовательности, лишённые метильных групп. Системы модификации и рестрикции всегда соответствуют друг другу, т.е. метилирование и разрезание происходят в одних и тех же последовательностях ДНК. В каждой системе в качестве мишеней для штамм-специфичных систем рестрикции-модификации используются свои короткие последовательности ДНК. Метилированная ДНК не разрезается по этой последовательности родственной рестриктирующей эндонуклеазой. Аналогично рестриктирующие эндонуклеазы разрезают только те ДНК, у которых не модифицированы соответствующие сайты рестрикции. Регуляция системы рестрикции и модификации осуществляется с помощью родственных наборов генов, и такие парные системы имеются у многих видов бактерий, бактериофагов и плазмид.



**РИС. II.19.**

Многие рестриктирующие эндонуклеазы разрезают молекулу ДНК несимметрично, в результате чего образуются взаимно комплементарные одноцепочечные концы. Любые два сегмента ДНК, имеющие такие концы, могут рекомбинировать *in vitro*. Если один из сегментов способен реплицироваться в соответствующей клетке-хозяине, то вся рекомбинантная молекула может быть клонирована и амплифицирована. В приведенном примере плазмидные ДНК соединены с двумя фрагментами, полученными с помощью рестриктирующих эндонуклеаз (верхняя часть рисунка).

При анализе генетической и физической организации сложных геномов особенно важными оказываются два свойства рестриктирующих эндонуклеаз. Первое связано с огромным диапазоном специфичностей, проявляемых в совокупности различными рестриктирующими эндонуклеазами, что позволяет разрезать практически любую ДНК на дискретные фрагменты самыми разными способами. Эти фрагменты можно разделить по размерам с помощью гель-электрофореза (рис. II.18). Распределение фрагментов по размерам, получающееся при расщеплении данной эндонуклеазной ДНК в специфических сайтах, является своего рода «отпечатком пальцев», характерным для этой ДНК. Второе свойство эндонуклеаз рестрикции связано со способностью многих из них осуществлять несимметричные разрезы двухцепочечной ДНК, в результате чего образуются фрагменты с комплементарными одноцепочечными концами (рис. II.19). Это позволяет проводить рекомбинацию ДНК *in vitro*. Любые два сегмента ДНК со взаимодополняющими концами могут объединиться, в результате чего происходит встраивание фрагментов в фаговую, плазмидную ДНК или другие потенциальные векторы. Такие специфические комбинации «вектор—вставка» можно получить в чистом виде путем молекулярного клонирования и амплификации в соответствующих хозяйских клетках (рис. II.12).

В последующих главах мы рассмотрим те уникальные свойства рестриктирующих эндонуклеаз, благодаря которым с помощью молекулярного клонирования удалось выделить специфические гены даже из наиболее сложных геномов. Вскоре после этого была получена детальная информация о структуре генов и их регуляторных элементов и построены подробные карты их расположения

в геноме. Препятствия, которые стояли на пути молекулярной генетики эукариотических геномов, были тем самым устранены.

Развитие методологии рекомбинантных ДНК и молекулярного клонирования само по себе привело к созданию новых областей биологических исследований. Однако полная реализация этих новых возможностей зависела от других проводимых одновременно разработок: развития методов фракционирования, которые позволяли бы разделять фрагменты ДНК, лишь незначительно различающиеся по длине (рис. II.18); создания относительно простых и быстрых процедур секвенирования ДНК длиной несколько тысяч нуклеотидов; получения специфически модифицированных генов путем внесения мутаций в их клонированные копии *in vitro*; генетической трансформации клеток, тканей и всего организма путем введения клонированных генов в соответствующие системы.

В четырех главах, составляющих часть II, описаны приемы и методы конструирования, клонирования, отбора и получения характеристик рекомбинантных ДНК. Приведенный материал не может служить лабораторным руководством. Основной акцент делается лишь на рассмотрении ключевых положений. Часто обращается внимание на ранние результаты, благодаря которым был разработан тот или иной метод. Ученые, получившие эти результаты, не могли предвидеть, как они скажутся на развитии технологии рекомбинантных ДНК. Теперь мы знаем, что именно успехи в молекулярной генетике прокариот и в изучении ферментов, осуществляющих синтез и деградацию нуклеиновых кислот, открыли путь к исследованию эукариотических геномов.

# ИНСТРУМЕНТАРИЙ: ФЕРМЕНТЫ

Возможность проведения различных манипуляций с ДНК *in vitro* всецело зависит от наличия очищенных ферментов, которые специфическим образом разрезают, модифицируют и соединяют молекулы. В настоящее время отсутствуют чисто химические методы, с помощью которых можно было бы осуществлять перестройку молекул ДНК с такими селективностью и разнообразием, которые характерны для ферментативных реакций. В то же время даже с помощью довольно небольшого числа ферментов можно получать рекомбинантные молекулы ДНК. Большинство этих ферментов были открыты при обстоятельствах, не связанных с их использованием при манипулировании с молекулами ДНК. На самом деле каждый фермент играет важную роль катализатора в том или ином химическом процессе, протекающем в организме, из которого он выделен. Использование ферментов в качестве инструмента при манипулировании с ДНК зависит, в частности, от их доступности и стабильности, а особенно от их чистоты, и прежде всего от того, свободны ли они от примесей, влияющих на ферментативную активность.

## 4.1. НУКЛЕАЗЫ

### а. Общие свойства

Нуклеазы позволяют специфическим образом модифицировать молекулы ДНК и РНК. Каждый фермент может быть отнесен к тому или иному классу в соответствии с его специфичностью или типом реакции, которую он катализирует. Так, ряд ферментов, подобно рестриктирующим эндонуклеазам, действует только на ДНК. Другие, подобно панкреатической РНКазе, гидролизуют только РНК (табл. 4.1). Есть ферменты, которые используют в качестве субстратов как ДНК, так и РНК. Одни нуклеазы предпочитают действовать либо на двухцепочечные, либо на одноцепочечные полинуклеотидные субстраты, другие не проявляют такой выраженной предпочтительности.

Нуклеазы можно также разделить на две следующие категории: **эксонуклеазы** и **эндонуклеазы**. Эксонуклеазы расщепляют полинуклеотидные субстра-

ты, имеющие свободные концы, при этом расщепление идет либо непосредственно с концов цепей, либо вблизи них. Разные экзонуклеазы расщепляют цепь предпочтительно либо с 5'-, либо с 3'-конца, но иногда они не проявляют такой специфичности. Эндонуклеазам свободные концы не требуются, поэтому данные ферменты могут гидролизовать кольцевые молекулы ДНК. Разрезание осуществляется по внутренним фосфодиэфирным связям, при этом образуются фрагменты разной длины. С помощью экзонуклеаз тоже могут образовываться короткие полинуклеотидные фрагменты, однако конечным продуктом во многих случаях являются нуклеозидмонофосфаты, поскольку экзонуклеазы осуществляют гидролиз, отщепляя последовательно один остаток за другим.

Наконец, нуклеазы отличаются одна от другой по тому, с какой стороны от межнуклеотидного фосфодиэфирного мостика осуществляется гидролиз (рис. 4.1). Одни ферменты делают разрез между фосфатом и 3'-гидроксильной группой с образованием 5'-фосфомоноэфирных продуктов, другие – между фосфатом и 5'-гидроксильной группой, в ре-



**РИС. 4.1.**

Нуклеазы могут расщеплять полинуклеотидную цепь с одной или с другой стороны от фосфодиэфирного мостика. Каждый фермент проявляет определенную специфичность в этом отношении.

Таблица 4.1. Типичные нуклеазы

Название <sup>1)</sup>	Источник	Специфичность	Предпочитаемая структура <sup>2)</sup>	Эндо или экзо	Продукт <sup>3)</sup>
*Bal 31	<i>Alteromonas espejiana</i>	ДНК ДНК	о.п. д.п.	Эндо Экзо $\left\{ \begin{array}{l} 5' \rightarrow 3' \\ 3' \rightarrow 5' \end{array} \right\}$	5'-p(Np) <sub>n</sub> H 5'-NMP
Exo I	<i>E. coli</i>	ДНК	о.п.	Экзо (3'→5')	5'-NMP
Exo III	<i>E. coli</i>	ДНК	д.п.	Экзо (3'→5')	5'-NMP
*Exo VII	<i>E. coli</i>	ДНК	о.п.	Экзо $\left\{ \begin{array}{l} 5' \rightarrow 3' \\ 3' \rightarrow 5' \end{array} \right\}$	5'-p(Np) <sub>n</sub> N
λ-Эксонуклеаза	<i>E. coli</i> (λ-фаг)	ДНК	д.п.	Экзо (5'→3')	5'-NMP
*Нуклеаза	Проростки	ДНК, РНК	о.п.	Эндо	5'-p(Np) <sub>n</sub> N
<i>Phaseolus aureus</i>	<i>Phaseolus aureus</i>				
*Нуклеаза	<i>Neurospora crassa</i>	ДНК, РНК	о.п.	Эндо	5'-p(Np) <sub>n</sub> N
<i>Neurospora crassa</i>					
Панкреатическая ДНКаза I	Крупный рогатый скот	ДНК	о.п., д.п.	Эндо	5'-p(Np) <sub>n</sub> N
Панкреатическая РНКаза	Крупный рогатый скот	РНК	о.п.	Эндо	(Np) <sub>n</sub> <sup>c</sup> <sub>u</sub> -p-3'
*РНКаза Н	Клетки	РНК•ДНК	д.п.	Эндо	5'-p(Np) <sub>n</sub> N
*РНКаза Н	Ретровирус	РНК•ДНК	д.п.	Экзо (3'→5')	5'-p(Np) <sub>n</sub> N
*Нуклеаза S1	<i>Aspergillus oryzae</i>	ДНК, РНК	о.п.	Эндо	5'-p(Np) <sub>n</sub> N
Фосфодиэстераза змеиного яда	<i>Crotalus adamanteus</i>	ДНК, РНК	о.п.	Экзо (3'→5')	5'-NMP
Фосфодиэстераза селезенки	Крупный рогатый скот	ДНК, РНК	о.п.	Экзо (5'→3')	3'-NMP
Стафилококковая нуклеаза	<i>Staphylococcus aureus</i>	РНК, ДНК	о.п.	Эндо	(Np) <sub>n</sub> N-p-3'

<sup>1)</sup> Ферменты, отмеченные звездочкой, описаны в тексте.

<sup>2)</sup> О.п. - одноцепочечная ДНК, д.п. - двухцепочечная.

<sup>3)</sup> N - любой нуклеотид; (Np)<sub>n</sub> - полинуклеотидная цепь, содержащая n остатков.

зультате чего образуются продукты с 3'-фосфомоноэфирными концами.

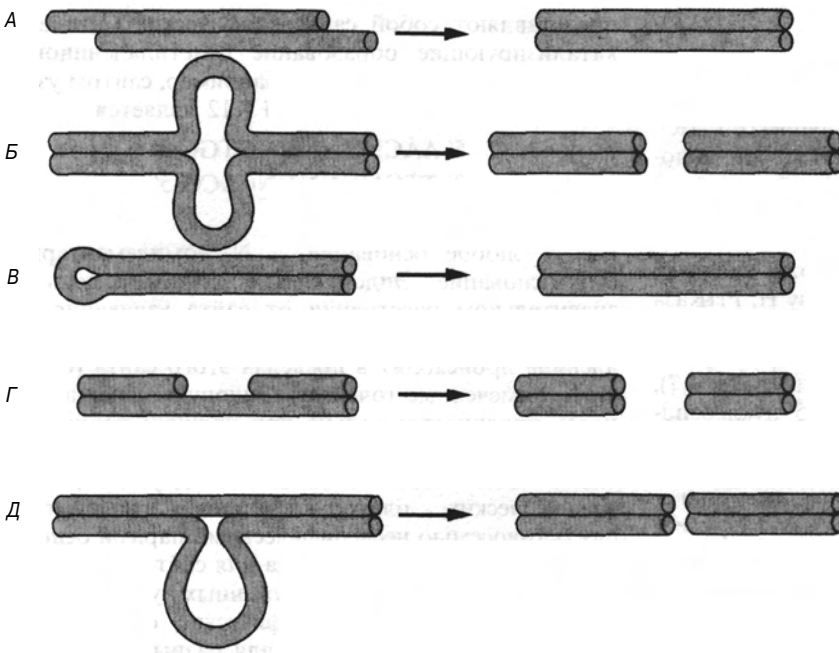
Большинство ферментов, перечисленных в табл. 4.1, упоминаются в тексте этой книги. Нуклеазы, отмеченные звездочкой, используются как рутинные реактивы в экспериментах по получению рекомбинантных молекул ДНК и поэтому рассматриваются более детально.

## б. Нуклеазы, специфичные в отношении одноцепочечной ДНК

**Эндонуклеазы.** Некоторые эндонуклеазы гидролизуют одноцепочечные молекулы ДНК примерно в тысячу раз быстрее, чем двухцепочечные. Такая специфичность используется во многих экспериментах — при конструировании рекомбинантных ДНК, при гетеродуплексном анализе и даже при анализе экспрессии генов. Некоторые реакции, катализируемые подобными эндонуклеазами, представлены на рис. 4.2.

Несколько разных эндонуклеаз такого типа были достаточно хорошо очищены и охарактеризованы, чтобы их можно было использовать в качестве реактивов. Одна из них, нуклеаза S1, получена из высушенных препаратов плесневого гриба *Aspergillus oryzae*; источниками двух других широко используемых ферментов являются *Neurospora* и *Mung beans*. Каждый из этих трех ферментов проявляет максимальную активность и точность распознавания одноцепочечных и двухцепочечных молекул при определенных условиях. Все три фермента гидролизуют как ДНК, так и РНК. При разрыве фосфоэфирных связей с помощью этих ферментов образуются 5'-монофосфатные и 3'-гидроксильные концы.

**Эксонуклеазы.** Эксонуклеаза *E. coli* exo VII специфична в отношении одноцепочечных ДНК. Она обладает необычной экзонуклеазной активностью в том смысле, что инициирует отщепление как с 5', так и с 3'-концов цепи, в то время как наиболее известные экзонуклеазы специфичны к какому-то одному концу (см., однако, описание экзонуклеазы

**РИС. 4.2.**

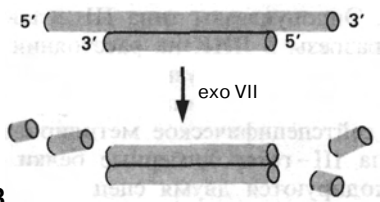
Некоторые реакции, катализируемые эндонуклеазами, специфичными в отношении одноцепочечных участков. А. Отщепление одноцепочечных хвостов у дуплексной ДНК. Б. Разрезание дуплексной ДНК с локально расплетенным участком на два фрагмента. В. Разрезание дуплексной ДНК со шпилькой на конце с образованием одного двухцепочечного дуплекса. Г. Разрезание дуплекса с пробелом с образованием двухцепочечных спиральных сегментов. Д. Разрезание дуплексной ДНК с одноцепочечной петлей в середине на двухцепочечные фрагменты.

*Bal 31* в разд. 4.1.в). Рис. 4.2, А и 4.3 иллюстрируют действие *exo VII* на две разные дуплексные ДНК. Хотя *exo VII* распознает концы цепей, она не отщепляет последовательно по одному нуклеотиду. Продуктами гидролиза отдельных цепей экзонуклеазой *exo VII* являются олигонуклеотиды длиной примерно 25 мономерных единиц, которые содержат 5'-фосфомоноэфирные концевые группы.

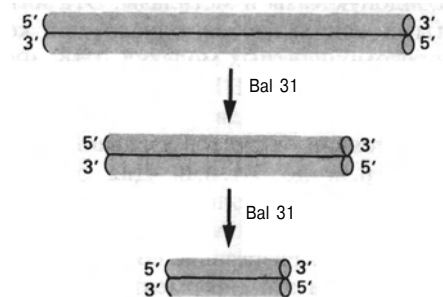
### в. Нуклеаза *Bal 31*

Псевдомонада *Alteromonas espejiana* секретирует единственную дезоксирибонуклеазу, получившую название *Bal 31*. В отношении одноцепочечной ДНК, в том числе одноцепочечных участков двухцепочечной ДНК, *Bal 31* ведет себя как эндонуклеаза, действуя аналогично другим эндонуклеазам, специфичным к одноцепочечным ДНК. Однако в отношении интактной двухцепочечной ДНК этот фермент проявляет экзонуклеазную активность, по-ви-

димо, благодаря тому, что он способен распознавать локальные одноцепочечные участки (или места нестабильности в дуплексах). *Bal 31* разрезает обе цепи на обоих концах дуплекса, т.е. осуществляет деградацию одновременно в направлениях 3'→5' и 5'→3'. В результате двухцепочечная молекула постепенно укорачивается (рис. 4.4.). Если эмпирически оценить скорость этого процесса, то с помощью фермента *Bal 31* можно получать фрагменты ДНК нужной длины. Хотя укорочение разных молекул ДНК происходит несинхронно, получается набор фрагментов, длина которых близка к заданной. В результате истощающегося гидролиза с помощью *Bal 31* промежуточные олигонуклеотидные продукты расщепляются до 5'-моонуклеотидов.

**РИС. 4.3.**

Реакция, катализируемая ферментом *exo VII*. Экзонуклеаза отщепляет одноцепочечные 5'- и 3'-концы с освобождением олигонуклеотидов длиной около 25 остатков.

**РИС. 4.4.**

Разрезание ферментом *Bal 31* двухцепочечной линейной ДНК.

## г. РНКазы Н

Существует группа ферментов, получивших название РНКазы Н потому, что они специфически расщепляют цепь РНК в гибридном дуплексе РНК–ДНК. РНКазы Н *E. coli* представляет собой эндонуклеазу, продуктами действия которой являются олигорибонуклеотиды с 5'-фосфомоноэфирными концами. Этот фермент широко используется как реактив (разд. 4.7). Клетки эукариот тоже содержат подобную эндонуклеолитическую РНКазу Н. РНКазы Н-экзонуклеолитическая активность присуща экзонуклеазе III *E. coli* (табл. 4.1) и обратным транскриптазам, кодируемым ретровирусами (разд. 4.7). Экзонуклеаза III расщепляет РНК до 5'-нуклеозидмонофосфатов в направлении 3'→5', а продуктами гидролиза цепи РНК с помощью обратной транскриптазы являются 5'-фосфорилированные олигорибонуклеотиды длиной от двух до десяти нуклеотидов.

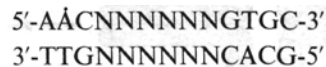
## 4.2. ЭНДОНУКЛЕАЗЫ РЕСТРИКЦИИ

Эволюция наделила различные виды бактерий уникальными эндонуклеазами, позволяющими им отличать их собственную ДНК от чужеродной. Тем самым природа снабдила ученых богатым набором высокоспецифичных реактивов для расщепления ДНК. При изучении ДНК большое значение имеют две важные особенности **рестриктирующих эндонуклеаз**. Первая связана с замечательной способностью фермента узнавать специфические короткие нуклеотидные последовательности в ДНК. Вторая состоит в том, что существует большое количество различных эндонуклеаз рестрикции, каждая из которых узнает специфическую последовательность.

### а. Три типа эндонуклеаз рестрикции

*Эндонуклеазы типов I и II.* Ферменты, относящиеся к группе эндонуклеаз типов I и II,— это сложные белки, обладающие активностями рестриктирующей эндонуклеазы и метилазы. Эти интересные ферменты не используются, однако, при конструировании рекомбинантных молекул ДНК. Ферменты типа I связываются с ДНК в специфических участках и затем производят двухцепочечные разрезы на разном расстоянии от сайтов узнавания, варьирующем от 400 п.н. до 7 т.п.н. Для осуществления ферментативного гидролиза ДНК необходимы  $Mg^{2+}$ , АТФ и S-аденозилметионин. Последний активирует фермент. Разрезание сопровождается гидролизом АТФ, при этом фермент утрачивает эндонуклеолитическую активность, но сохраняет АТФазную. Таким образом, эндонуклеазы типа I являются ДНК-зависимыми АТФазами. Кроме того, они

представляют собой сайтспецифические метилазы, катализирующие образование 6-метиладениновых остатков в сайте узнавания. Например, сайтом узнавания для фермента из *E. coli* K12 является



где N—любое основание, а  $\dot{N}$ —комплементарное ему основание. Эндонуклеаза разрезает цепь на значительном расстоянии от сайта узнавания, но при этом метилирование с образованием 6-метиладенина происходит в пределах этого сайта (остатки А, помеченные точками). Эндонуклеазная активность проявляется только при наличии полностью неметилированного сайта. Сайт узнавания для фермента *E. coli* K12 состоит из двух коротких специфических олигонуклеотидов, разделенных шестью-восемью неспецифическими парами оснований. Такая структура типична для сайтов узнавания ферментов типа I, обнаруженных у различных штаммов бактерий. Специфические олигонуклеотидные последовательности для разных ферментов различаются, но метилированные остатки А всегда находятся в одинаковых позициях.

Родственные системы рестрикции-модификации типа I кодируются аллельными локусами геномов различных кишечных бактерий (например, *E. coli* и *Salmonella typhimurium*). С каждой системой связаны три сцепленных гена: *hsdR*, *hsdM* и *hsdS*, расположенные в порядке, соответствующем порядку транскрипции. Полипептидные продукты *hsdM* и *hsdS* транслируются с одной двухцепочечной мРНК и вместе составляют метилазу. Первый из них обладает метилазной активностью, а второй осуществляет сайтспецифическое узнавание. Продукт гена *hsdR* обладает эндонуклеазной активностью. Все три полипептида содержатся в различных пропорциях в активных препаратах ферментов типа I.

Ферменты типа III, так же как и эндонуклеаза типа I, обладают нуклеазной и метилазной активностями. Однако несмотря на то, что ферменты типа III активируются S-аденозилметионином и для своей работы требуют АТФ, они не катализируют ее гидролиза. Эндонуклеазы типа III делают двухцепочечные разрезы в ДНК на расстоянии примерно 25 п.н. от своих сайтов узнавания. В присутствии АТФ и S-аденозилметионина те же ферменты катализируют сайтспецифическое метилирование. Ферменты типа III—гетеродимерные белки. Их субъединицы кодируются двумя сцепленными генами, локализованными у некоторых штаммов *E. coli* во внехромосомных геномах (фаговой ДНК или плазмиде). Каждый такой ген представляет собой отдельную транскрипционную единицу. Один генный

продукт узнает специфическую последовательность и обладает метилазной активностью, другой выполняет эндонуклеазную функцию.

Поскольку ферменты типа I не разрезают молекулы ДНК на репродуцирующиеся фрагменты и эндонуклеазные активности ферментов типа I и III конкурируют с метилированием, ни один из них не используется в качестве реагента при получении рекомбинантных ДНК.

**Эндонуклеазы типа II.** Рестриктирующие эндонуклеазы типа II являются основным инструментом при конструировании рекомбинантных молекул ДНК и при анализе структуры ДНК. Эти ферменты способны узнавать специфические короткие нуклеотидные последовательности и связываться с ними, но в отличие от эндонуклеаз типов I и III они производят двухцепочечные разрезы по специфическим фосфодиэфирным связям либо в пределах самого сайта узнавания, либо на вполне определенном небольшом расстоянии от него. Один и тот же фермент разрезает данную молекулу ДНК с образованием одинаковых наборов фрагментов. Длина фрагментов определяется расстоянием между специфическими последовательностями, узнаваемыми этим ферментом. Ферменты типа II гидролизуют фосфодиэфирные связи между 3'-гидроксильной группой и фосфатом, в результате чего образуется 5'-фосфомоноэфирная группа с одной стороны от разрыва и 3'-гидроксильная группа — с другой. Для функционирования ферментов необходим двухва-

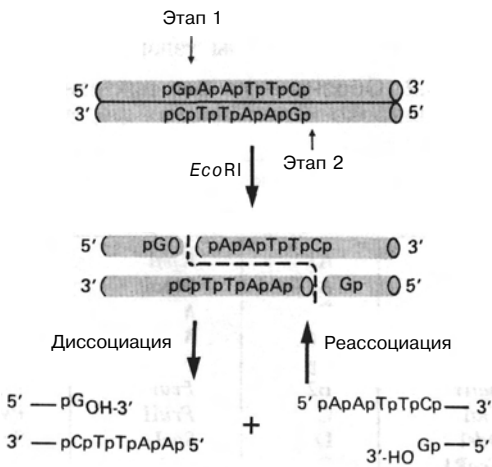
лентный катион, обычно  $Mg^{2+}$ , но ни АТФ, ни S-аденозилметионин не требуются.

## б. Типичная рестриктирующая эндонуклеаза типа II

**Эндонуклеаза *Eco* RI.** Широко используемая рестриктирующая эндонуклеаза *Eco* RI разрезает ДНК в тех местах, где встречается последовательность 5'-GAATTC-3' (рис. 4.5). Как и в случае других рестриктирующих эндонуклеаз типа II, эта последовательность является **палиндромом**, т.е. в обеих цепях точно напротив друг друга находятся одинаковые последовательности, читаемые в направлении 5'→3'. *Eco* RI гидролизует фосфодиэфирные связи между остатками G и A в каждой цепи. Поскольку эндонуклеаза *Eco* RI разрезает обе цепи в том месте, где встречается палиндромная последовательность 5'-GAATTC-3', молекула ДНК разрезается на характерный для нее набор фрагментов («фингерпринт» — «отпечатки пальцев», рис. 4.6).

**Липкие концы.** Эндонуклеаза *Eco* RI производит ступенчатые двухцепочечные разрезы, при этом у образующихся фрагментов ДНК на концах образуются короткие комплементарные одноцепочечные хвосты из четырех оснований — 5'-AATT-3' (рис. 4.5). В зависимости от ионной силы раствора и температуры эти комплементарные хвосты либо остаются спаренными, либо денатурируют, и тогда образуются отдельные фрагменты с короткими одноцепочечными выступами на 5'-концах. При соответствующих условиях комплементарные хвосты воссоединяются (рис. 4.5). Одноцепочечные концы, образующиеся при расщеплении ДНК эндонуклеазой *Eco* RI, получили название **липких**, поскольку они способны спариваться (как бы слипаться) друг с другом. Важным следствием образования ступенчатых разрывов является то, что фрагменты, получающиеся в результате обработки эндонуклеазой *Eco* RI двух разных ДНК (например, ДНК *E. coli* и дрожжей), могут соединяться с помощью липких концов. При этом различия, затрагивающие двухцепочечные спиральные сегменты указанных ДНК, на процесс соединения не влияют.

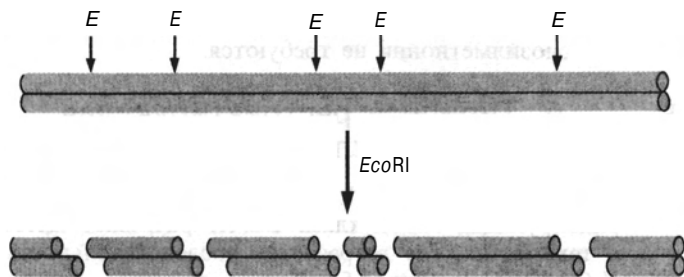
**Фермент.** В отличие от рестриктирующих эндонуклеаз типов I и III ферменты типа II, подобные *Eco* RI, не катализируют метилирования. Метилазную активность проявляет другой белок. Активная эндонуклеаза *Eco* RI представляет собой димер, состоящий из двух идентичных полипептидных цепей с мол. массой 31000. Исходя из гомодимерной структуры можно было бы ожидать, что фермент будет действовать симметрично на идентичные последовательности двух частей сайта узнавания и разрезать обе цепи одновременно. Однако на самом деле гидролиз происходит последовательно.



**РИС. 4.5.**

Сайт узнавания и разрезания для рестриктирующей эндонуклеазы *Eco* RI. Разрезание происходит в два этапа: сначала гидролизует одна цепь, потом — другая. При обычных условиях продуктами реакции являются два дуплексных фрагмента с комплементарными одноцепочечными концами. В условиях, способствующих стабилизации водородных связей между четырьмя нуклеотидами из разных цепей (например, при понижении температуры), фрагменты могут реассоциировать.



**РИС. 4.6.**

Схематическое изображение разрезания эндонуклеазой *Eco* RI протяженной двухцепочечной ДНК, содержащей несколько сайтов узнавания для этого фермента (E).

Белок взаимодействует с участком около десяти пар оснований; шесть из них находятся в пределах сайта узнавания, остальные являются фланкирующими. Именно эти фланкирующие последовательности, по-видимому, определяют выбор цепи, которая будет разрезана первой. Кроме того, фланкирующие пары оснований влияют на общую скорость разрезания в данном сайте узнавания, так что в разных сайтах одной молекулы ДНК гидролиз осуществляется с разной скоростью. Были получены кристаллы комплекса *Eco* RI и олигонуклеотида 5'-TCGCGAATTCGCG-3' и проведен их рентгеноструктурный анализ. Это позволило в деталях выяснить способ взаимодействия данного фермента и ДНК. Было установлено, что между двумя остатками аргинина и одним глутаматом (на каждой субъединице), с одной стороны, и парами оснований в сайте узнавания—с другой, образуются водородные связи, в результате чего происходит частичное раскручивание В-ДНК. О свойствах других рестриктирующих эндонуклеаз типа II известно значительно меньше, но, по-видимому, они аналогичны свойствам *Eco* RI.

### в. Различные группы рестриктирующих эндонуклеаз типа II

**Большой каталог реактивов.** На рис. 4.7–4.9 представлены некоторые из примерно 400 уже описанных рестриктирующих эндонуклеаз. У всех этих ферментов обнаружено около 90 разных сайтов узнавания. Каждый из ферментов получен из определенного прокариотического организма, что и отражено в его названии. Так, *Eco* RI ведет свое происхождение из штамма *Escherichia coli* K12; *Hae* II и *Hae* III—из *Haemophilus aegyptius*; *Bam* HI—из *Bacillus amyloliquefaciens* штамма H; *Mbo* I и *Mbo* II—из *Moraxella bovis* и т.д. Два разных фермента, имеющие одинаковые сайты узнавания, называются **изошизомерами**. Однако изошизомеры не обязательно делают разрез в одном и том же месте; сравните, например, *Xma* I и *Sma* I на рис. 4.7 и 4.8 соответственно. Чтобы упростить описание рестриктазных сайтов широко применяемых ферментов,

мы используем набор обозначений, представленных в табл. 4.2.

**Палиндромные сайты узнавания и разрезания.** Эндонуклеаза *Eco* RI—одна из многих рестриктирующих эндонуклеаз типа II (рис. 4.7), которые узнают различные палиндромные нуклеотидные последовательности и делают разрезы, в результате чего образуются фрагменты с комплементарными одноцепочечными хвостами. В пределах этой группы можно выделить несколько разновидностей ферментов. Например, *Eco* RI, *Hind* III и некоторые другие эндонуклеазы узнают последовательности из шести пар оснований; эндонуклеаза *Hinf* I узнает пять пар оснований, из которых центральным остатком (на рис. 4.7 он обозначен как N) может быть любой из четырех нуклеотидов. Эндонуклеаза *Bgl* I узнает шесть специфических нуклеотидных остатков, но они должны быть разделены в середине любыми другими пятью парами оснований. Эндонуклеаза *Not* I имеет сайт узнавания протяженностью восемь пар оснований; такие ферменты, как *Taq* I и *Mbo* I, имеют сайты узнавания из четырех

Таблица 4.2. Обозначения рестриктазных сайтов наиболее широко используемых ферментов

Рестриктирующая эндонуклеаза	Обозначение	Рестриктирующая эндонуклеаза	Обозначение
<i>Alu</i> I	A1	<i>Kpn</i> I	K
<i>Ava</i> I	A	<i>Mbo</i> I	M1
<i>Bam</i> HI	B	<i>Mbc</i> II	M2
<i>Bcl</i> I	Bc	<i>Msp</i> I	M
<i>Bgl</i> I	Bg	<i>Pst</i> I	P
<i>Bgl</i> II	B2	<i>Pvu</i> I	Pu
<i>Cla</i> I	C	<i>Pvu</i> II	Pv
<i>Dde</i> I	D	<i>Sac</i> I	Sc
<i>Eco</i> RI	E	<i>Sal</i> I	S
<i>Eco</i> RII	E2	<i>Sau</i> 3A	S3
<i>Eco</i> RV	E5	<i>Sma</i> I	Sm
<i>Hae</i> II	He	<i>Sph</i> I	Sp
<i>Hae</i> III	Ha	<i>Sst</i> I	Ss
<i>Hinc</i> II	Hc	<i>Taq</i> I	T
<i>Hind</i> III	H	<i>Xba</i> I	X
<i>Hinf</i> I	Hf	<i>Xho</i> I	Xh
<i>Hpa</i> I	HI	<i>Xma</i> III	Xm
<i>Hph</i> I	Hp		

нуклеотидов. Другая известная особенность касается свойств одноцепочечных концев, образующихся при действии различных рестриктирующих эндонуклеаз (рис. 4.7). Эндонуклеазы *Eco* RI, *Hind* III, *Mbo* I и *Cla* I образуют одноцепочечные хвосты, состоящие

Фермент	Сайт узнавания и разрезания
<i>Bam</i> HI	5' - GGATCC - - CCTAGG - 5'
<i>Bcl</i> I	5' - TGATCA - - ACTAGT - 5'
<i>Bgl</i> I	5' - GCCNNNN NGGC - - CGGN NNN NCCG - 5'
<i>Cla</i> I	5' - ATCGAT - - TAGCTA - 5'
<i>Eco</i> RI	5' - GAATTC - - CTTAAG - 5'
<i>Haell</i>	5' - RGC G CY - - Y C G C GR - 5'
<i>Hind</i> III	5' - AAGCTT - - TTCGAA - 5'
<i>Hinf</i> I	5' - GANTC - - CTNAG - 5'
<i>Hpa</i> II	5' - CCGG - - GGCC - 5'
<i>Mbo</i> I & <i>Sau</i> 3A	5' - GATC - - CTAG - 5'
<i>Msp</i> I	5' - CCGG - - GGCC - 5'
<i>Not</i> I	5' - GCGGCCGC - - CGCCGGCG - 5'
<i>Pst</i> I	5' - CTGCAG - - GACGTC - 5'
<i>Taq</i> I	5' - TCGA - - AGCT - 5'
<i>Xma</i> I	5' - CCGGG - - GGGCC - 5'

РИС. 4.7.

Группа рестриктирующих эндонуклеаз типа II, разрезающих палиндромные последовательности с образованием липких концов. Стрелками указаны точки разрезания. Фосфодиэфирный мостик, обозначаемый обычно буквой «р», на этой и последующих диаграммах не указан. Очевидно, что при разрезании ДНК рестриктирующими эндонуклеазами всегда образуются 5'-концевой фосфомоноэфир и 3'-концевой гидроксил (рис. 4.1). N - любое основание.  $\bar{N}$  - комплементарное ему основание. Буквами R и Y обозначены один из пуринов и один из пиримидинов соответственно.

Фермент	Сайт узнавания и разрезания
<i>Alu</i> I	5'-AGCT- -TCTGA-5'
<i>Hae</i> III	5'-GGCC- -CCGG-5'
<i>Hpa</i> I	5'-GTTAAC- -CAATTG-5'
<i>Sma</i> I	5'-CCCGGG- -GGGCCC-5'

РИС. 4.8.

Группа рестриктирующих эндонуклеаз типа II, разрезающих палиндромные последовательности с образованием тупых концов.

из 5'-концевых остатков каждой цепи, а эндонуклеазы *Pst* I и *Hae* II — из 3'-остатков.

Эндонуклеазы, имеющие различные сайты узнавания и разрезания, часто образуют идентичные липкие концы. Например, под действием эндонуклеаз *Mbo* I, *Bcl* I и *Bam* HI образуются 5'-концевые 5'-GATC-3'. В результате фрагменты, получающиеся при расщеплении данной ДНК любым из этих ферментов, при смешивании и отжиге будут соединяться друг с другом. Однако фрагменты, получающиеся из различных ДНК при разрезании эндонуклеазой *Bgl* I, вряд ли будут соединяться при отжиге, поскольку одноцепочечные хвосты могут содержать любую последовательность из трех нуклеотидов.

Другие рестриктирующие эндонуклеазы типа II (рис. 4.8) также узнают специфические палиндромные нуклеотидные последовательности, но разрезают фосфодиэфирный мостик в середине узнаваемой последовательности, в результате чего образуются фрагменты ДНК с **тупыми** двухцепочечными концами.

*Непалиндромные сайты узнавания с разрезанием на некотором расстоянии от них.* Ферменты типа II, но другой разновидности (рис. 4.9) тоже узнают специфические нуклеотидные последовательности, но гидролизуют фосфодиэфирные мостики вне этих последовательностей. При этом узнаваемая последовательность не является палиндромом. В этих случаях рестриктирующие эндонуклеазы, по-видимому, «отсчитывают» точное число пар оснований от узнаваемой последовательности и затем разрезают цепи. Механизм отсчета таков, что места гидролиза разных цепей смещены одно относительно

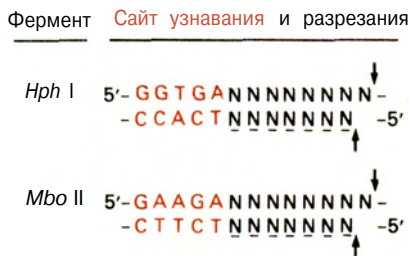


РИС. 4.9.

Группа рестриктирующих эндонуклеаз типа II, разрезающих цепи ДНК на определенном расстоянии от узнаваемой последовательности.

но другого на один нуклеотид. В результате образуются фрагменты ДНК с одноцепочечными выступами длиной только в один нуклеотидный остаток.

### г. Картирование сегментов ДНК с помощью рестриктирующих эндонуклеаз типа II

С помощью рестриктирующих эндонуклеаз можно разрезать сложные геномы или длинные сегменты ДНК на воспроизводимые наборы более мелких единиц. В свою очередь такие единицы можно разделить по размерам с помощью нескольких методов. Чаще всего для этого используют электрофорез фрагментов ДНК в полужидком геле, приготовленном на основе агарозы или полиакриламида (рис.4.10). Обычно подвижность двухцепочечного фрагмента в электрическом поле обратно пропорциональна логарифму его размера. Используя в качестве маркеров фрагменты известной длины, легко определить размер интересующего нас фрагмента с помощью линейки и простого графика. Для проведения такого анализа обычно достаточно менее 1 мкг ДНК, поскольку фрагменты легко выявляются при окрашивании соответствующим красителем, например бромистым этидием.

Набор фрагментов, получающихся при расщеплении ДНК определенной рестриктирующей эндонуклеазой, является своеобразным «отпечатком пальцев» этой ДНК. Определяя, какие фрагменты образуются при расщеплении ДНК несколькими рестриктазами по отдельности и в разных комбинациях, часто удается установить порядок расположения сегментов в исходной молекуле, т.е. построить физическую карту ДНК, где указано положение каждого сегмента (рис.4.11). Сложные геномы или большие молекулы ДНК дают при обработке рестриктирующими эндонуклеазами сложные наборы фрагментов, и для их анализа используют специальные компьютерные программы.

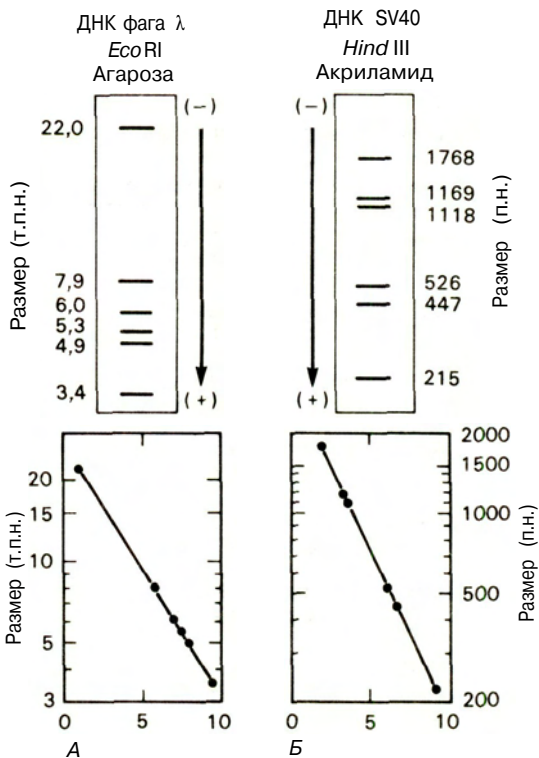
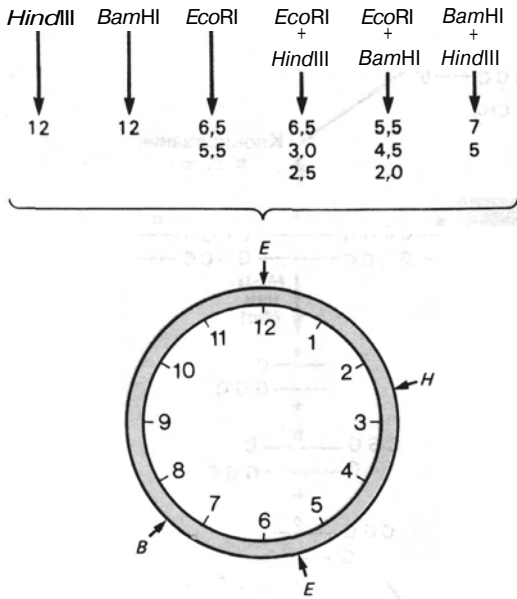


РИС. 4.10.

Электрофоретическое разделение фрагментов ДНК, получающихся при гидролизе последней рестриктирующими эндонуклеазами. А. Электрофорез в агарозном геле смеси фрагментов, образующихся при гидролизе ДНК бактериофага  $\lambda$  ферментом *Eco* RI. Б. Электрофорез в полиакриламидном геле смеси фрагментов, образующихся в результате гидролиза ДНК SV40 эндонуклеазой *Hind* III. В агарозном геле разделяют относительно крупные фрагменты (23-0,3 т.п.н.), а в полиакриламидном - более мелкие (6 т.п.н. - 2 п.н.). Варьируя процентное содержание агарозы или акриламида в геле, можно оптимизировать разделение фрагментов в том или ином диапазоне длин. Для визуализации полос, содержащих фрагменты разной длины, гель окрашивают бромистым этидием. В ультрафиолетовом свете различимы полосы, содержащие всего 50 нг ДНК. Под каждой из диаграмм приведены графики зависимости относительной подвижности фрагментов от логарифма их размеров.

### д. Защита ДНК посредством метилирования

**Родственные метилазы.** Геномы бактерий, кодирующие рестриктирующие эндонуклеазы, защищены от самодеградации с помощью специфических систем модификации. Метилазы, осуществляющие специфическую модификацию, узнают те же нуклео-

**РИС. 4.11.**

Построение рестрикционной карты кольцевой молекулы ДНК длиной 12 т.п.н. ДНК гидролизовали указанными эндонуклеазами по отдельности или в разных комбинациях и определяли размер фрагментов (в т.п.н.) с помощью электрофореза в агарозном геле. Зная размер фрагментов, можно установить положение сайтов узнавания для трех представленных ферментов.

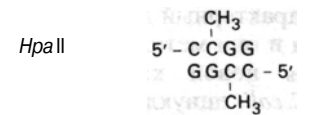
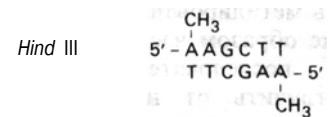
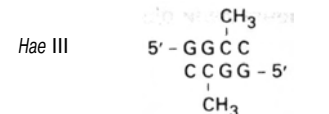
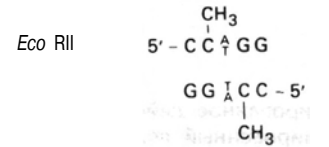
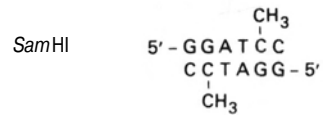
тидные последовательности, что и родственные им эндонуклеазы, и переносят метильную группу от S-аденозилметионина к специфическому адениновому (с образованием N<sup>6</sup>-метиладенина) или цитозиновому (с образованием 5-метилцитозина) остаткам в пределах сайта узнавания (рис. 1.4). В результате эндонуклеазное разрезание блокируется. На рис. П.17 представлены основные типы соответствующих реакций, а на рис. 4.12—характерные сайты метилирования для нескольких систем типа II. Некоторые метилазы достаточно хорошо очищены и могут использоваться при конструировании рекомбинантных молекул ДНК (разд. 6.2.б). Метилирование цепей, как и разрезание эндонуклеазами, происходит симметрично. Однако метилазы, которые были изучены, включая фермент *Eco*RI,—это мономерные белки, и метилирование разных цепей осуществляется в два отдельных этапа. Таким образом, несмотря на то, что эндонуклеазы и метилазы типа II узнают одну и ту же короткую последовательность ДНК и, по-видимому, эволюционировали параллельно, это разные белки.

**Идентификация метилированных сайтов в ДНК.** Метилирование—это весьма распространенное явление, совсем не обязательно связанное с рестрицией-модификацией. В результате некоторые сайты

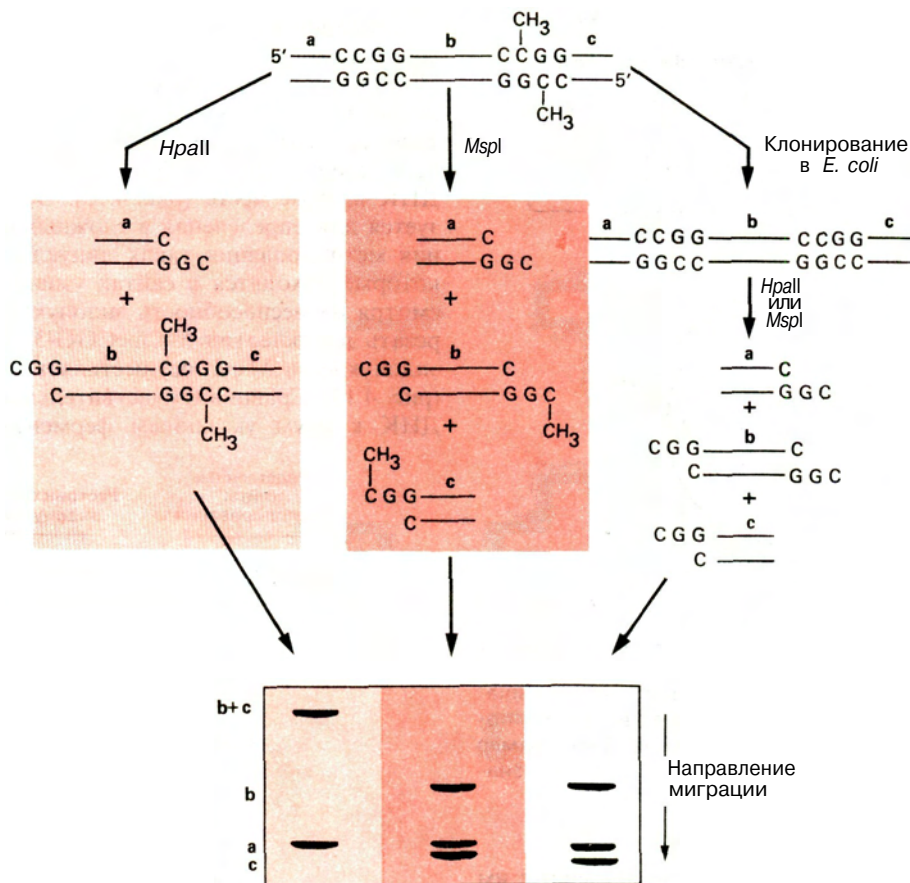
рестрикции в ДНК проявляют устойчивость к действию определенных эндонуклеаз. Например, в ДНК позвоночных часто встречаются динуклеотиды 5'-meCG-3' (разд. 1.1.a). Если сайт рестрикции эндонуклеазы *Hpa* II содержит такой метилированный динуклеотид, то фермент не может разрезать ДНК в этом месте (рис. 4.13). Этот факт используется для определения в сложных геномах состояния метилирования таких динуклеотидов 5'-CG-3', которые находятся в сайтах узнавания *Hpa* II. Несмотря на неспособность эндонуклеазы *Hpa* II разрезать динуклеотид 5'-СmeCGG-3', ее изоизомер, эндонуклеаза *Msp* I, осуществляет такое разрезание (рис. 4.13). Сравнивая чувствительность фрагментов ДНК к двум указанным ферментам, можно ло-

Родственные сайты метилирования

Рестриктирующая эндонуклеаза

**РИС. 4.12.**

Родственные сайты метилирования некоторых систем рестрикции-модификации.

**РИС. 4.13.**

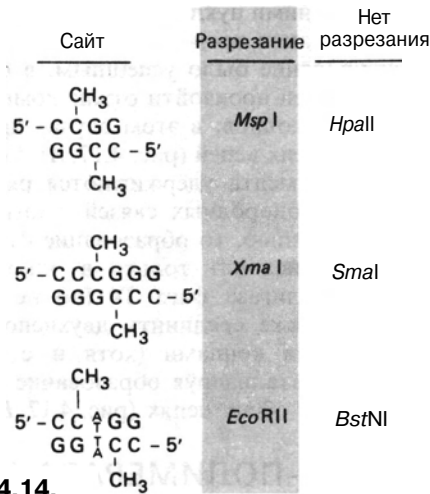
Дифференцированное действие *Msp I* и *Hpa II* на частично метилированный сегмент ДНК. После того как данный сегмент был клонирован и реплицирован в *E. coli*, оба фермента дали одинаковый набор фрагментов,

поскольку 5-метилцитозин был замещен цитозином. Внизу схематически представлены результаты гелеэлектрофореза смеси фрагментов ДНК, полученных после гидролиза в каждом из трех случаев.

каллизовать метилированный цитозин на карте. Подобным же образом часто встречающуюся в ДНК растений последовательность 5'-СмеС<sub>T</sub>GG-3' можно отличить от неметилированной формы, 5'-СС<sub>T</sub>GG-3', если сравнить продукты, получающиеся при расщеплении ДНК с помощью изошизомеров *Bst NI* и *Eco RII* (рис. 4.14).

**Независимые метилазы *E. coli*.** При включении фрагментов эукариотической ДНК в бактериальные векторы и их репликации в клетках *E. coli* или других прокариот характерный для них тип метилирования утрачивается и они метилируются способом, специфичным для новой клетки-хозяина. Например, в клетках *E. coli* динуклеотиды 5'-CG-3' не метилируются, и в результате сайты, ранее защищенные от действия эндонуклеазы *Hpa II*, становятся

чувствительными к ней (рис. 4.13). Однако два обычных фермента *E. coli*: *dam* ДНК-метилтрансфераза, которая метилирует остаток А в последовательности 5'-GATC-3', и *dcm* ДНК-метилтрансфераза, метилирующая остаток С в последовательности 5'-CC<sub>T</sub>GG-3', образуют новые устойчивые сайты; ни одна из этих метилаз не является частью системы рестрикции-модификации. Кроме того, эукариотическая ДНК, реплицированная в *E. coli*, становится устойчивой к эндонуклеазам *Mbo I* и *Bcl I* благодаря тому, что в пределах их сайтов узнавания находится 6-метиладенин. Заметим, что тот же самый 6-метиладенин в сайте 5'-GATC-3' не мешает работе эндонуклеаз *Bam HI*, *Bgl I* или *Sau 3A*; однако разрезание с помощью *Sau 3A* блокируется, если метилирован остаток С.



**РИС. 4.14.**

Пары рестриктирующих эндонуклеаз, с помощью которых можно определить локализацию остатков 5-метилцитозина.

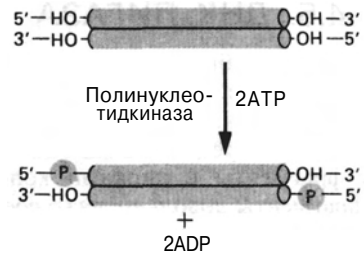
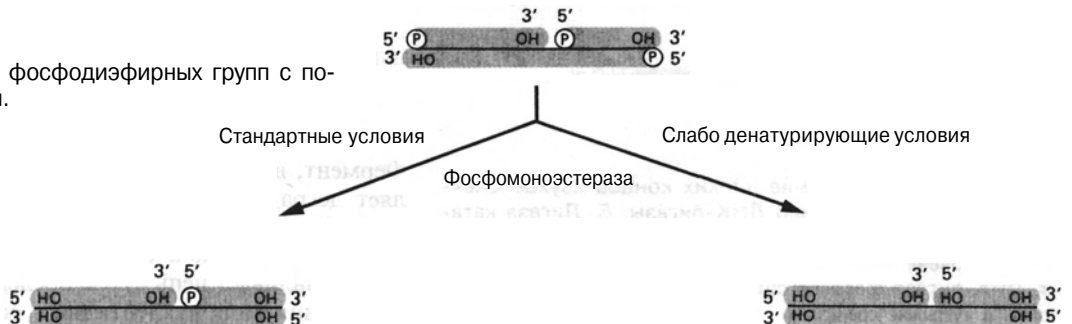
### 4.3. ФОСФОМОНОЭСТЕРАЗЫ

Часто при проведении экспериментов с рекомбинантными ДНК или при анализе структуры ДНК приходится отщеплять концевые фосфомоноэфирные группы (рис. 4.15). Известно множество неспецифических фосфомоноэстераз (фосфатаз). Эти ферменты, выделенные из клеток *E. coli* и кишечника теленка, были получены в высокоочищенном виде. Фосфатазы гидролизуют как 5'-, так и 3'-концевые фосфомоноэфиры в ДНК и РНК.

Фосфомоноэфирные группы, находящиеся на конце одноцепочечных разрывов в дуплексной ДНК или экранированные нависающей над ними второй цепью, подобные тем, которые образуются при разрезании эндонуклеазой *Pst* I (рис. 4.7), удаляются с помощью фосфатазы только при слабоденатурирующих условиях (например, при повышенной температуре).

**РИС. 4.15.**

Гидролиз концевых фосфодиэфирных групп с помощью фосфатазы.



**РИС. 4.16.**

Присоединение 5'-концевых фосфомоноэфирных групп с помощью полинуклеотидкиназы. В этих опытах используется <sup>32</sup>P-меченная АТФ, в результате чего образуется радиоактивно меченная ДНК.

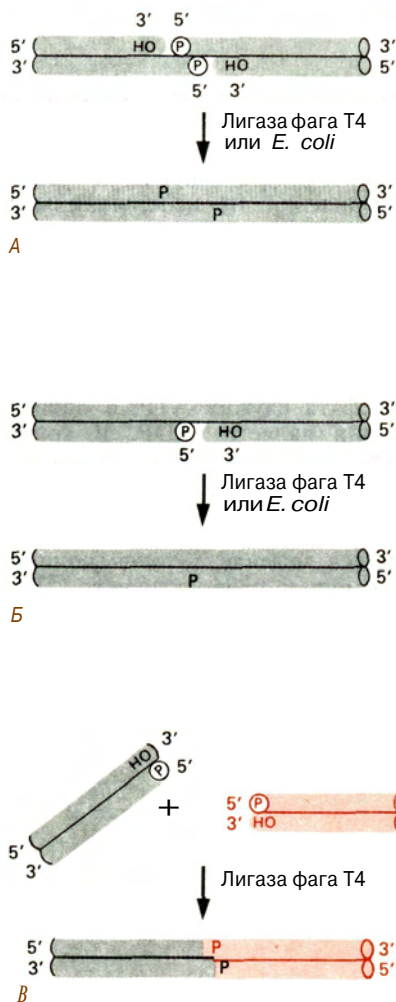
### 4.4. ПОЛИНУКЛЕОТИДКИНАЗА

Киназы составляют большой класс ферментов, катализирующих фосфорилирование многих биохимических соединений—от небольших молекул до очень крупных макромолекул, включая полипептиды и полинуклеотиды.

Полинуклеотидкиназа (рис. 4.16), широко используемая в качестве реактива в экспериментах с рекомбинантными ДНК, была выделена и очищена из клеток *E. coli*, инфицированных бактериофагом Т4. Фермент кодируется геномом бактериофага. Донором фосфата служит АТФ, а одним из продуктов реакции является АДФ. В ходе реакции специфически фосфорилируются 5'-концевые гидроксильные группы молекул ДНК и РНК; 3'-концевые гидроксильные группы не фосфорилируются. Одной из важных сфер применения полинуклеотидкиназы является мечение 5'-конца полинуклеотидной цепи с помощью радиоактивного <sup>32</sup>P с использованием АТФ, меченной по γ-фосфату. Последовательное действие фосфатазы (разд. 4.3) и полинуклеотидкиназы приводит к замещению немеченого 5'-концевого фосфомоноэфира на радиоактивный без каких-либо других изменений в цепи.

## 4.5. ДНК-ЛИГАЗА

Получение рекомбинантных молекул ДНК включает объединение *in vitro* сегментов ДНК из различных источников. Благодаря тому что при расщеплении ДНК определенными эндонуклеазами образуются фрагменты с липкими концами, при отжиге эти фрагменты довольно легко соединяются, однако водородные связи, удерживающие их вместе, в обычных условиях оказываются относительно слабыми и легко разрушаются. Для ковалентного сшивания фрагментов используют ДНК-лигазу, которая катализирует образование фосфодиэфирных



**РИС. 4.17.**

А. Ковалентное сшивание липких концов двухцепочечных молекул с помощью ДНК-лигазы. Б. Лигаза катализирует образование фосфодиэфирных связей в месте одноцепочечного разрыва в двухцепочечной молекуле. В. ДНК-лигаза фага Т4 сшивает двухцепочечные молекулы с тупыми концами.

связей между соседними нуклеотидами (разд. 2.1.д). Механизм лигирования представлен на рис. 2.20.

Чтобы лигирование было успешным, в объединяемых цепях должен произойти отжиг комплементарных («липких») концов; в этом случае происходит лигирование обеих цепей (рис. 4.17, А). Если два одноцепочечных сегмента удерживаются рядом за счет образования водородных связей с интактной комплементарной цепью, то образование фосфодиэфирной связи происходит только в одной цепи (рис. 4.17, Б). ДНК-лигаза фага Т4 (но не лигаза *E. coli*) способна также соединить двухцепочечные молекулы с тупыми концами (хотя и с низкой эффективностью), катализируя образование фосфодиэфирных связей в обеих цепях (рис. 4.17, В).

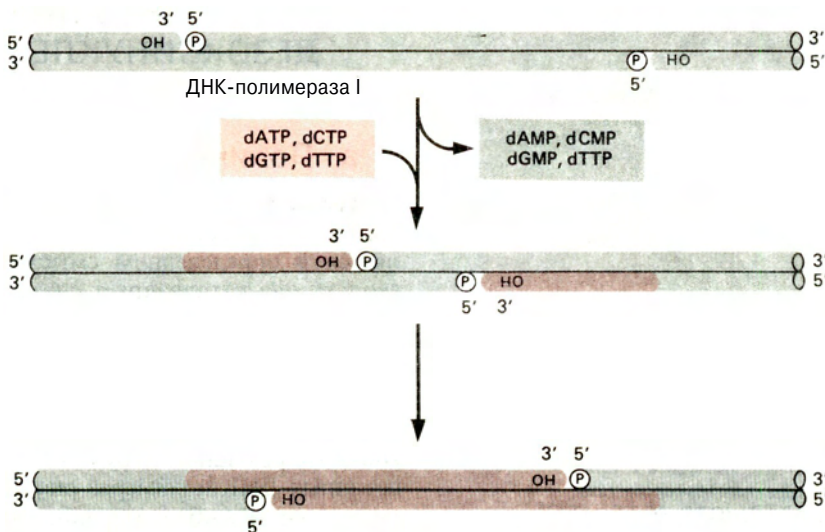
## 4.6. ДНК-ПОЛИМЕРАЗА I

### а. Полифункциональный фермент

Механизм, с помощью которого ДНК-полимераза катализирует полуконсервативный синтез новых цепей ДНК, был описан в разд. 2.1.д. Цепи удлиняются путем последовательного присоединения нуклеотидов к праймеру со свободной 3'-гидроксильной группой; выбор нуклеотидного остатка на каждом этапе определяется ДНК-матрицей. ДНК-полимераза I *E. coli* катализирует и некоторые другие важные реакции. Две из них имеют особое значение для экспериментов с рекомбинантными ДНК: это 3'→5'- и 5'→3'-экзонуклеазные реакции. В обоих случаях образуются продукты с 5'-фосфономоэфирными концами. Две указанные экзонуклеазные активности присущи разным участкам молекулы ДНК-полимеразы I, которые можно разделить, обработав белок протеолитическими ферментами. После разделения обнаруживается, что 3'→5'-экзонуклеазная и полимеразная активности связаны с большим по размеру полипептидом, содержащим карбоксильную группу на конце, а 5'→3'-экзонуклеазная активность — со вторым, более мелким фрагментом.

### б. Ник-трансляция

Реакции, катализируемые ДНК-полимеразой I, нашли широкое применение. Например, часто используется способность ДНК-полимеразы I катализировать одновременно как полимеризацию, так и 5'→3'-экзонуклеазное расщепление (рис. 4.18). Фермент, выступая в роли экзонуклеазы, осуществляет деградацию цепи в 5'→3'-направлении, начиная с 5'-конца одноцепочечной бреши (ника) в двухцепочечной ДНК, а выступая в роли полимеразы, восстанавливает цепь путем последовательного присоединения мононуклеотидных остатков к свобод-

**РИС. 4.18.**

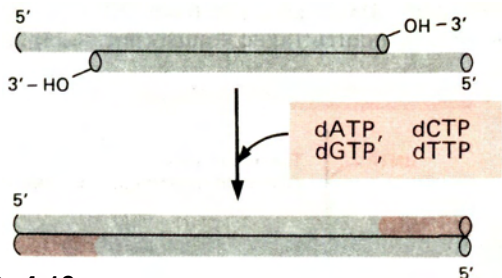
Ник-трансляция. В результате совместного действия 5'→3'-экзонуклеазной и полимеразной активностей ДНК-полимеразы I *E. coli* в присутствии α-<sup>32</sup>P-меченных дезоксирибонуклеотидтрифосфатов (на рисунке они заклю-

чены в окрашенный прямоугольник) одноцепочечный разрыв (ник) в двухцепочечной ДНК перемещается вдоль молекулы ДНК, при этом синтезируется <sup>32</sup>P-меченная ДНК.

ной 3'-гидроксильной группе на другом конце бреши. Собственно синтеза ДНК не происходит: брешь лишь перемещается вдоль цепи, чем и объясняется название этого процесса — **ник-трансляция**. Проводя ник-трансляцию в присутствии α-<sup>32</sup>P-меченных дезоксирибонуклеотидтрифосфатов, в качестве субстратов получают меченые фрагменты ДНК с высокой удельной активностью.

### в. Заполнение брешей

ДНК-полимераза способна превращать фрагменты ДНК с липкими концами в фрагменты с тупыми концами при наличии выступающего 5'-конца (рис. 4.19). 3'-гидроксильный конец служит праймером, 5'-выступ — матрицей; пробел на 3'-конце заполняется путем последовательного присоединения дезоксирибонуклеотидных остатков. Для заполнения лучше использовать большой карбокси-конце-

**РИС. 4.19.**

Образование тупых концов путем синтеза недостающего участка одной из цепей с помощью большого фрагмента ДНК-полимеразы I.

вой фрагмент ДНК-полимеразы I, получаемый в результате протеолиза; это позволяет избежать 5'→3'-экзонуклеазного расщепления матрицы (разд. 2.1.д). Липкие концы можно превратить в тупые и с помощью специфичной к одноцепочечной ДНК нуклеазы, однако в этом случае утрачивается несколько нуклеотидных остатков (рис. 4.2, А).

## 4.7. РНК-ЗАВИСИМЫЕ ДНК-ПОЛИМЕРАЗЫ (обратные транскриптазы)

Эти ферменты были выделены из РНК-содержащих онкогенных вирусов. Они позволяют синтезировать ДНК на РНК-матрице *in vitro* (разд. 2.2). Реакция, катализируемая обратными транскриптазами, аналогична стандартным реакциям с участием ДНК-полимераз и, как в случае с другими ДНК-полимеразами, нуждается в затравке (праймере). В качестве матрицы обычно используется комплементарная цепь ДНК. В результате образуется гибридная молекула РНК-ДНК. Часто удобным праймером служит короткая цепочка полидезоксириботимидиловой кислоты [poly(dT)], поскольку она способна спариваться с полирибоадениловой кислотой, обычно находящейся на концах эукариотических мРНК (рис. 4.20) (разд. 3.8.а); в результате такого спаривания инициируется синтез ДНК-копии этой РНК (**кДНК**).



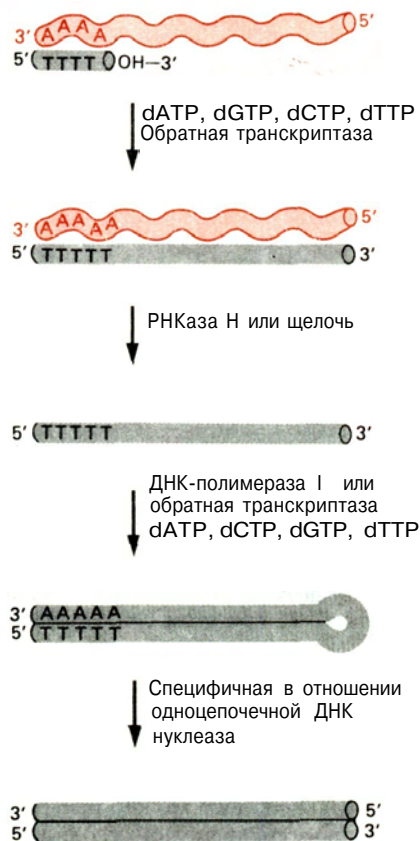


РИС. 4.20.

Схематическое изображение реакции, катализируемой обратной транскриптазой, и синтеза двухцепочечной кДНК.

кДНК-копия молекулы мРНК может быть превращена в двухцепочечную ДНК. Для этого прежде всего с помощью РНКазы Н или щелочного гидролиза удаляют исходную матричную РНК. Оставшаяся одноцепочечная кДНК служит собственным праймером (а также и матрицей) при синтезе второй комплементарной цепи ДНК. Как это происходит, не совсем ясно. Последняя реакция может катализироваться либо ДНК-полимеразой I, либо обратной транскриптазой. По-видимому, праймером служит короткая шпильчатая структура, образующаяся вблизи 3'-гидроксильного конца первой цепи. Конечный дуплекс все еще содержит шпильку на одном из концов, которая разрезается специфической в отношении одноцепочечной ДНК нуклеазой (см. рис. 4.2, В) с образованием двух полностью комплементарных цепей ДНК (дуплексная кДНК).

## 4.8. ТЕРМИНАЛЬНАЯ ДЕЗОКСИНУКЛЕОТИДИЛ-ТРАНСФЕРАЗА

### а. Полимеризация без матрицы

**Терминальная дезоксинуклеотидилтрансфераза** (часто называемая **терминальной трансферазой**) является в определенном смысле полимеразой, поскольку она катализирует синтез полидезоксирибонуклеотидов из дезоксирибонуклеозидтрифосфатов с высвобождением неорганического пирофосфата. Подобно ДНК-полимеразам, она не способна инициировать синтез новой полимерной цепи и поэтому требует присутствия праймера со свободной концевой 3'-гидроксильной группой. Однако в отличие от истинных ДНК-полимераз она не нуждается в матрице и не способна что-либо копировать вообще. Продукт полимеризации соответствует дезоксинуклеозидтрифосфатам, используемым в качестве субстрата (рис. 4.21). Если в реакции участвует dATP, то продуктом является полидезоксиадениловая кислота [poly(dA)], находящаяся на 3'-конце праймера; если используется dGTP, то к праймеру оказывается присоединенной poly(dG). Поскольку терминальная дезоксинуклеотидилтрансфераза не копирует матрицу, продуктом синтеза является одноцепочечный полимер. При использовании в качестве праймера двухцепочечной молекулы на каждом конце дуплекса образуются одинаковые 3'-хвосты. На самом деле терминальная дезоксинуклеотидилтрансфераза предпочитает одноцепочечные, а не двухцепочечные праймеры. Если имеется полностью двухцепочечный фрагмент ДНК, к которому нужно присоединить одноцепочечные хвосты, можно использовать различные методы получения одноцепочечных концов, способных служить праймерами (рис. 4.22).

### б. Синтез липких концов

С помощью дезоксинуклеотидилтрансферазы к молекулам ДНК, не имеющим липких концов, можно присоединить такие концы (рис. 4.23). Тупые

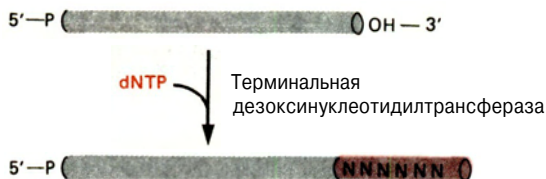
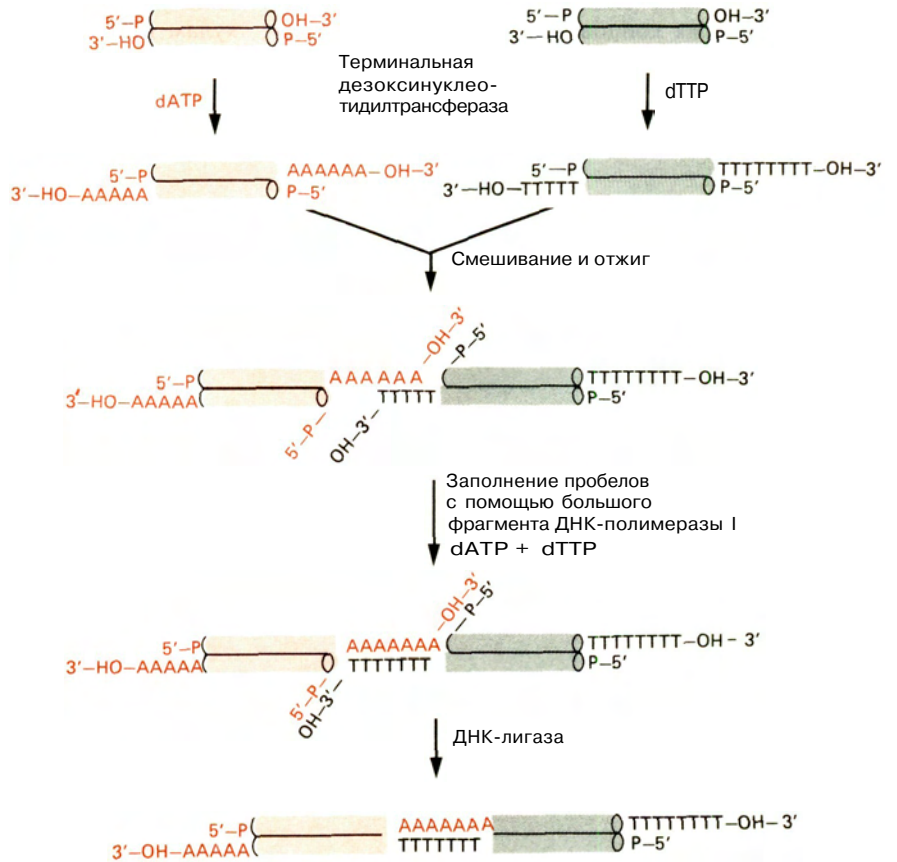
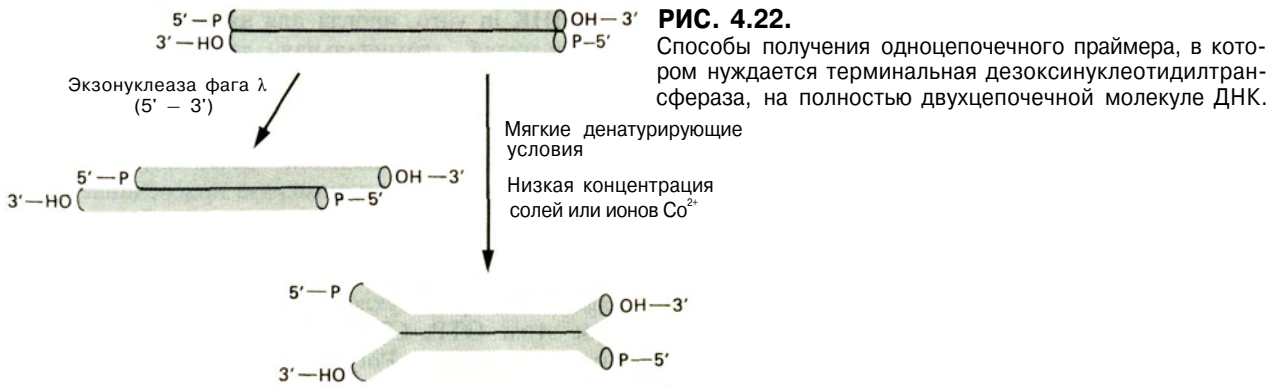


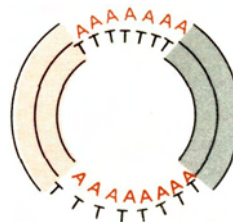
РИС. 4.21.

Реакция, катализируемая терминальной дезоксинуклеотидилтрансферазой. N - любой из четырех дезоксинуклеотидов, А, Т, G или С.



**РИС. 4.23.**

Образование липких концов с помощью терминальной дезоксинуклеотидилтрансферазы. Конечный продукт - линейная молекула - легко может замкнуться в кольцо.



концы образуются при механическом разрыве больших молекул ДНК (например, при быстром перемешивании раствора или продавливании его через узкое отверстие), при обработке ДНК с помощью рестриктирующих эндонуклеаз, дающих тупые концы, или при действии неспецифических ДНКаз. Например, если к одному набору фрагментов ДНК присоединить poly(dT), а к другому – poly(dA) и смешать эти фрагменты, то произойдет их соединение друг с другом. Пробелы, образующиеся из-за неравной длины poly(dT)- и poly(dA)-хвостов, могут быть заполнены с помощью ДНК-полимеразы I, а концы соединены ДНК-лигазой. Такой подход использовался при конструировании первой рекомбинантной молекулы ДНК. И хотя открытие рестриктирующих эндонуклеаз сильно упростило процедуру получения липких концов при рекомбинации моле-

кул ДНК *in vitro*, иногда для их создания все еще используется терминальная дезоксирибонуклеотидилтрансфераза.

#### 4.9. POLY(A)-ПОЛИМЕРАЗА

Подобно терминальной дезоксирибонуклеотидилтрансферазе, poly(A)-полимераза присоединяет нуклеотидные остатки к 3'-концу цепи без участия матрицы. Однако poly(A)-полимераза проявляет специфичность в отношении РНК. Праймером является полирибонуклеотид, а субстратом – АТР. При этом GTP, CTP, UTP или dATP не могут заменить АТР. Poly(A)-полимераза выполняет высокоспециализированную и важную функцию в экспериментах с рекомбинантными ДНК: ее используют для присоединения poly(A)-концов к молекулам РНК при синтезе кДНК (разд. 4.7).

## Глава 5

# ИНСТРУМЕНТАРИЙ: СИСТЕМЫ ХОЗЯИН–ВЕКТОР

В основе молекулярного клонирования лежит встраивание нужного фрагмента ДНК (вставки) в другую молекулу ДНК (вектор), которая способна реплицироваться в соответствующей клетке-хозяине (см. рис. П.1.2). Такое встраивание осуществляется *in vitro*, а затем образовавшиеся рекомбинантные молекулы ДНК вводятся в клетки. Векторная молекула должна содержать точку начала репликации (*ori*). Кроме того, для репликации нужны специфические ферменты и другие белки; их поставляет клетка-хозяин или они кодируются самим вектором. Вектором может быть любой небольшой внехромосомный элемент (например, плаزمид, ДНК фага или вируса). Каждый из этих элементов встречается в природе в клетках определенных видов, и большинство из них реплицируется только в природном хозяине или клетках близкородственных видов. В большинстве случаев эволюция механизма репликации протекала в направлении создания оптимальных условий для существования в клетках природного хозяина внехромосомных генетических элементов, при этом использовались метаболиты, ферменты и другие белки клетки-хозяина, а также ее аппарат белкового синтеза. Поэтому основным инструментом молекулярного клонирования всегда является двухкомпонентная система—совместимая комбинация хозяина и вектора.

При рекомбинации *in vitro* обычно образуется популяция молекул ДНК, в которой лишь некоторые имеют нужную структуру. Клонирование уникальных рекомбинантных молекул должно происходить при соблюдении следующих требований:

1. Условия, при которых популяция рекомбинантных ДНК смешивается с популяцией реципиентных клеток, должны быть такими, чтобы в каждую клетку попало только по одной рекомбинантной молекуле. Это позволяет изолировать рекомбинантные молекулы друг от друга.

2. Каждая реципиентная клетка в популяции должна быть отделена от всех остальных, чтобы можно было выделить клон клеток или вирусов, содержащих уникальную рекомбинантную молекулу.

3. Клетки или вирусы, получившие рекомбинантные ДНК, должны отличаться от клеток, их не получивших, так чтобы их можно было отобрать либо идентифицировать при скрининге.

4. Клетки, получившие нужную рекомбинантную молекулу, должны отличаться от клеток, содержащих другие рекомбинантные молекулы ДНК, чтобы можно было отобрать их или идентифицировать при скрининге.

В этой главе мы рассмотрим разнообразные типичные системы «хозяин-вектор» и первые три требования, которые имеют к этой теме прямое отношение, поскольку методы, используемые для скрининга и селекции, в основном зависят от определенных свойств используемой комбинации «хозяин-вектор». Методы выделения желаемого рекомбинантного клона (требование 4) описаны в гл. 6. Отобранные нами примеры иллюстрируют основные принципы, которые применяются в настоящее время при конструировании систем, используемых в сложных экспериментальных ситуациях.

Наиболее широко применяются такие комбинации, когда в роли хозяина выступает штамм *E. coli* K12, а в роли вектора—плазмиды и фаги *E. coli*. Предпочтение, отдаваемое этому штамму, обусловлено тем, что еще задолго до развития технологии рекомбинантных ДНК он уже широко использовался при изучении генетики микробов. Его генетические и физиологические свойства были детально исследованы и в коллекциях имелись многие хорошо охарактеризованные мутанты этого штамма. Кроме того, в клетках штамма *E. coli* K12 могут реплицироваться многие бактериофаги и плазмиды—потенциально полезные векторы. Знание свойств этого бактериального штамма и соответствующих векторов и обусловило его преимущественное использование в экспериментах с рекомбинантными ДНК. Были разработаны также другие системы хозяин-вектор, которые сначала применяли лишь в особых случаях. Например, если ставилась задача получить в большом количестве белок, кодируемый клонированным геном, имело смысл использовать системы с *Bacillus subtilis*. С распространением генетических манипуляций на клетки эукариот и в особенности с началом исследования экспрессии генов в клетках дрожжей, растений и животных появилась необходимость в разработке подходящих эукариотических систем хозяин-вектор.

Вскоре после того как начались работы с реком-

бинантными ДНК, ряд ученых выразили беспокойство по поводу того, что клетки или вирусы, содержащие вставки «чужеродной» ДНК, могут проявлять неожиданные и даже опасные свойства. Это стимулировало поиск ослабленных систем хозяин-вектор, что уменьшило бы вероятность инфицирования людей, работающих с этими системами, и других живых организмов. Некоторые природные или лабораторные варианты штамма *E. coli* K12 не способны выживать или широко распространяться в окружающей среде, поскольку для их развития необходимы весьма специфические экспериментальные создаваемые условия, обычно не встречающиеся в природе. С помощью классических генетических методов и технологии рекомбинантных ДНК были получены варианты штамма *E. coli* K12 и других штаммов, что расширило круг применяемых ослабленных хозяев. В одних случаях соображения о необходимости ограничения работ с рекомбинантными ДНК и требования, предъявляемые к экспериментам с этими ДНК, находились в едином русле. В других же применение ослабленных штаммов давало неожиданный импульс развитию техники экспериментирования. В процессе работы накапливалось все больше данных о необоснованности многих опасений (по крайней мере для большинства проводимых экспериментов), и требования к применению ослабленных систем хозяин-вектор становились все менее строгими.

## 5.1. *E. COLI*-СИСТЕМЫ: КЛЕТКИ-ХОЗЯЕВА

### а. Разносторонность хозяина

Несмотря на то что штамм *E. coli* K12—это уникальный, отличный от других штамм *E. coli*, он не однороден по своим свойствам и на самом деле представляет собой семейство родственных бактериальных штаммов, происходящих от одного исходного изолята. Известные в настоящее время члены этого семейства отличаются друг от друга мутациями в одном или нескольких генах. Некоторые штаммы *E. coli* K12 получены с помощью направленного отбора специфических по фенотипу клонов, но большинство из них имеют неидентифицированные аллельные особенности. Многие из этих особенностей несущественны для свойств субклона, используемого в качестве хозяина для рекомбинантных молекул ДНК, однако некоторые имеют важное значение. Эти последние свойства клеток-хозяев, которые проявляются в зависимости от используемого вектора, можно разбить на несколько классов: одни из них имеют отношение к точной репликации вектора, другие—к успешному

введению рекомбинантных векторов, третьи—к удобному отбору необходимого типа рекомбинантов.

### б. «Гостеприимство» хозяина

Клетки хозяина должны обеспечивать все условия для репликации векторной ДНК и не содержать элементов, подавляющих репликацию вектора или мешающих отбору. Так, нельзя использовать клетки, содержащие активную систему рестрикции, поскольку это угрожает целостности рекомбинантных ДНК, содержащих чувствительные сайты рестрикции. Также неприемлемы клетки, содержащие нормальные *dam*- и *dcm*-метилтрансферазы *E. coli*, поскольку в этих случаях образуются реплицированные рекомбинантные ДНК, устойчивые к используемым в экспериментах в качестве инструмента рестриктирующим эндонуклеазам (разд. 4.2.e). Применяя клетки, дефектные в отношении нормальных рекомбинационных функций, можно предотвратить возникновение нежелательных изменений во вставке.

Проявление других важных свойств клетки-хозяина зависит от того, какой используется вектор—плазмидный или фаговый. Например, если вектором служит ДНК бактериофага  $\lambda$ , то хозяйские клетки не должны быть лизогенными по этому фагу, поскольку такие клетки устойчивы к повторному заражению, что обусловлено присутствием  $\lambda$ -репрессора—белка cI, который выключает экспрессию всех генов, необходимых для лизиса (разд. 3.11.e). Плазмидные векторы обычно содержат маркерные гены, благодаря которым клетки-хозяева приобретают удобный для отбора фенотип, свидетельствующий о присутствии в них данного вектора. Поэтому исходные клетки-хозяева не должны иметь похожего фенотипа. Например, мутантные клетки *E. coli*, нуждающиеся для своего роста в определенной аминокислоте, являются подходящими хозяевами в комбинации с векторами, несущими недостающий ген. Если высеять культуру клеток, из которых только часть содержит указанный вектор, на среду, лишенную данной аминокислоты, то колонии будут образовывать лишь те клетки, которые несут этот вектор. Точно так же клетки, чувствительные к определенному антибиотику или токсину, можно использовать в комбинации с векторами, несущими гены устойчивости к этим агентам. Как правило, в системах с плазмидными векторами не используются клетки-хозяева, наследующие посторонние плазмиды, поскольку при этом затрудняется очистка рекомбинантных ДНК, которая зависит от физических свойств плазмидной ДНК.

## в. Доступность хозяина

Проблема введения изолированных молекул ДНК в живые клетки *E. coli*—процесс, называемый **трансфекцией**,—была решена эмпирически вскоре после начала работы с рекомбинантными молекулами ДНК. Было показано, что в присутствии  $\text{CaCl}_2$  устойчивые клетки становятся проницаемыми для ДНК и их трансфекция с помощью фаговой или плазмидной ДНК осуществляется всего за несколько минут. Природа этого высокоэффективного процесса остается малопонятной. Еще более эффективная передача фаговых геномов достигается в том случае, если они сначала упаковываются в фаговые частицы *in vitro*, а затем вводятся в клетку с помощью стандартной процедуры инфицирования (разд. 5.3). Очень важное требование в этом случае — способность клеток адсорбировать и реплицировать бактериофаг, поскольку не все штаммы *E. coli* являются перmissive хозяевами для всех бактериофагов *E. coli*.

## г. Некоторые примеры

Особые свойства различных вариантов штамма *E. coli* K12 лучше всего иллюстрирует рассмотрение некоторых широко используемых производных этого штамма (табл. 5.1). *E. coli* K12 С600 является одним из стандартных производных штамма K12, хотя различные клоны С600 имеют неодинаковые генотипы. Субклон RR1 не способен усваивать галактозиды из среды (обозначение *lac Y* указывает на отсутствие функциональной  $\beta$ -галактозидпермеазы). Плазмидный вектор, поставляющий  $\beta$ -галактозидпермеазный ген, сообщает реципиентным клеткам способность усваивать  $\beta$ -галактозиды, которые затем гидролизуются  $\beta$ -галактозидазой. Это очень удобный маркер, поскольку успешно трансфицированные клетки образуют голубые колонии на агаре, содержащем 5-бром-4-хлор-3-индолил- $\beta$ -D-галакто-

зид (названный Xgal) и их легко идентифицировать. Голубой хромофор образуется в результате гидролиза Xgal  $\beta$ -галактозидазой. Штамм RR1 несет мутацию *hsdR*, которая инактивирует эндогенную рестриктирующую эндонуклеазу *E. coli* K12; это сводит к минимуму деградацию поступившей в клетку плазмидной рекомбинантной ДНК. Дополнительная защита от деградации обуславливается тем, что в клетках этого штамма отсутствует основная дезоксирибонуклеаза *E. coli*—*endA*. Субклон HB101, близкий родственник RR1, имеет еще одно преимущество: он не способен к общей рекомбинации (*recA*).

Штамм  $\chi$ 1776 был сконструирован специально для работы с рекомбинантными ДНК в ответ на выраженное некоторыми учеными беспокойство о возможной опасности этих экспериментов. С помощью стандартных генетических методов в ДНК была внесена мутация (*dap*), в результате которой клетки могли синтезировать клеточную стенку только в тех случаях, когда в среду добавляли существенный ее компонент—редкую диаминопимелиновую аминокислоту. Эти и другие мутации, приводящие к повреждению белков, необходимых для образования клеточной стенки, делают клетки  $\chi$ 1776 очень чувствительными к некоторым антибиотикам и к лизису солями желчных кислот (в кишечнике животных) и ионными детергентами. Эта же чувствительность приводит к необходимости использовать специальные процедуры для проведения успешной трансфекции.

Удобные векторы были сконструированы на основе фага M13 (разд. 5.3), представителя группы так называемых Ff- (т.е. F-специфичных нитчатых) фагов. Эти фаги инфицируют только те клетки *E. coli*, которые несут половой фактор, называемый F-фактором (см. введение к ч. II). Поэтому хозяином для фага M13 должен быть штамм  $F^+$ , каким является штамм 71-18. У *E. coli* 71-18 *lac*-оперон делетирован ( $\Delta[lac-pro]$ ) из геномной ДНК, но он содержится в F-плазмиде ( $F' lac$ ). Для облегчения работы с этим штаммом  $F' lac$ -плазмиду маркируют делецией ( $\Delta M15$ ), соответствующей N-концу р-галактозидазы. Голубые бляшки, указывающие на присутствие активной  $\beta$ -галактозидазы, обнаруживаются на агаре с Xgal только в том случае, если вектор M13 содержит область (*lac Z $\alpha$* ) *lac*-оперона, отсутствующую в  $F' lac$ . Каждый из двух геномов ( $F' lac$   $\Delta M15$ ) и M13 *lac Z $\alpha$* ) детерминирует свою часть молекулы  $\beta$ -галактозидазы, и в результате их взаимодействия внутри клетки образуется активная  $\beta$ -галактозидаза.

Таблица 5.1. Некоторые варианты штамма *E. coli* K12, широко используемые в качестве клеток-хозяев

Вариант	Вектор	Полезные гены <sup>1)</sup>
С600	Плазмиды, фаг $\lambda$	
RR1	Плазмиды	<i>lacY</i> , <i>hsdR</i> , <i>endA</i>
HB101	»	<i>lacY</i> , <i>hsdR</i> , <i>endA</i> , <i>recA</i>
$\chi$ 1776	»	<i>dap</i>
71-18	Фаг M13	$\Delta[lac-pro]$ ; $F' lac$ ( $\Delta M15$ )
DH1 <sup>2)</sup>	Плазмиды	<i>hsdR</i> , <i>endA</i> , <i>recA</i>

<sup>1)</sup> Эти гены *E. coli* описаны в тексте.

<sup>2)</sup> Этот штамм особенно эффективно трансфицируется плазмидными векторами.

Таблица 5.2. Модульная организация плазмид

Модули	Плазмиды (приблизительные размеры в т. п. н.)				
	F (93)	R100 (100)	ColV (140)	ColE1 (6)	Ap201 (14)
Модуль репликации со строгим контролем (1-2 копии)	+	+	+	—	—
Модуль репликации с ослабленным контролем (10-30 копий)	—	—	—	+	+
Модуль конъюгации	+	+	+	—	—
Модуль резистентности к тетрациклину	—	+	—	—	—
Col-модуль	—	—	+	+	—
Модуль резистентности к ампициллину	—	—	—	—	+
Модуль резистентности к хлорамфениколу	—	+	—	—	—

## 5.2. E. COLI-СИСТЕМЫ: ПЛАЗМИДНЫЕ ВЕКТОРЫ

### а. Модульная структура плазмид

Существует много разных способов классификации плазмид, но, по-видимому, целесообразнее всего рассмотреть их здесь с точки зрения наличия в них модульных сегментов ДНК.

В табл. 5.2 суммированы данные о модульных особенностях нескольких встречающихся в природе плазмид. Каждый модульный сегмент может содержать один или несколько генов, или *цис-действующих* элементов,— таких, например, как область инициации репликации. Функционально родственные модули в различных и независимо выделенных плаزمиде часто, но не всегда оказываются и структурно родственными, как если бы они произошли от одной ДНК. Каждая плаزمида должна иметь по крайней мере один из нескольких репликационных модулей, которые позволяют ей автономно реплицироваться. В одном из классов, типичным представителем которого являются F-плазмиды, репликация и сегрегация регулируются согласованно с репликацией бактериального генома, и в каждой клетке содержится одна или две копии плазмиды. О таком типе репликации говорят как о репликации со **строгим контролем**. Второй тип репликационного модуля свободен от такого контроля, что приводит к существованию в клетке многих копий плазмид. Подобный тип репликации получил название репликации с **ослабленным контролем**.

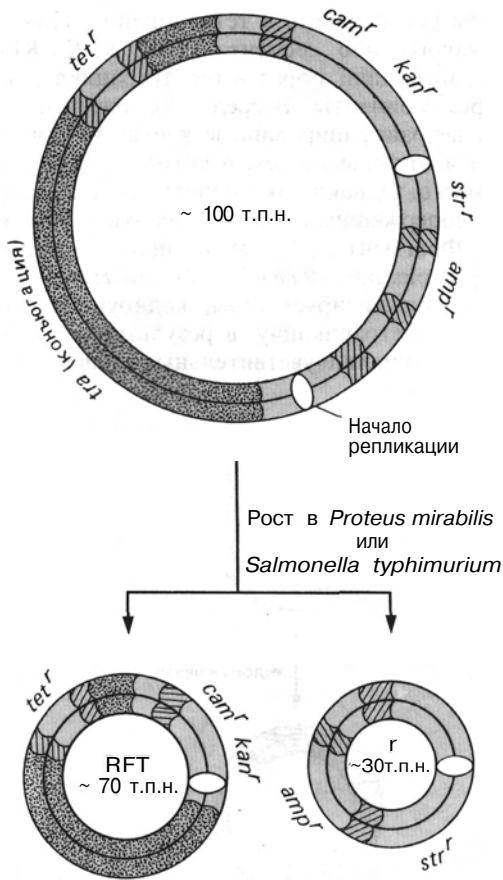
Наличие других модулей, не связанных с репли-

кацией, не является обязательным для каждой определенной плазмиды. Половые факторы, т.е. **конъюгативные** (или **самотрансмиссивные**) плазмиды, подобные F-фактору, имеют модули, содержащие гены и регуляторные области, необходимые для переноса плазмиды из одной клетки в другую (см. введение к ч. II). Половые факторы, несущие модуль ДНК, полученный от бактериальной хромосомы, отмечены апострофом, как в случае F'.

Модули другого типа содержат гены, белковые продукты которых инактивируют антибиотики. Плазмиды, несущие такие модули, часто называют **R-плазмидами** (от англ. resistance—устойчивость; табл. 5.3). Часто одна плазмида обеспечивает устойчивость к нескольким антибиотикам, при этом все или несколько генов резистентности разного типа могут быть сгруппированы в одном модуле. Не все R-плазмиды способны осуществлять функцию полового фактора, которая зависит от наличия в плазмиде модуля конъюгации. **Неконъюгативные** плазмиды (утратившие модуль конъюгации) не способны к самотрансмиссивности, но некоторые из них могут быть перенесены из одной клетки в другую, если они совмещены в донорной клетке с половым фактором. Половой фактор осуществляет перенос неконъюгативной плазмиды, которую в этом случае называют **мобилизуемой**.

Таблица 5.3. Гены резистентности к антибиотикам

Антибиотик (ген)	Кодируемые белки	Механизм резистентности
Ампициллин ( <i>amp</i> )	β-Лактамаза (пенициллиназа)	Гидролиз связи C—N в β-лактамном кольце
Хлорамфеникол ( <i>cam</i> )	Хлорамфеникол-ацетилтрансфераза	Ацетилирование САМ с помощью ацетил-СоА с образованием О-ацетил-САМ
Тетрациклин ( <i>tet</i> )	Охарактеризован недостаточно	Уменьшает способность клеток концентрировать тетрациклин
Стрептомицин ( <i>str</i> )	Стрептомицин-фосфотрансфераза (1) Стрептомицин-аденилатсинтетазы (2)	Фосфорилирование (1) или аденилирование (2) ОН-группы стрептомицина с помощью АТФ
Канамицин ( <i>kan</i> )	Канамицин-ацетилтрансфераза (1)	N-ацетилирование (1) или О-фосфорилирование (2) этих родственных антибиотиков
Неомицин ( <i>neo</i> )	Аминогликозид-фосфотрансфераза (2)	



**РИС. 5.1.**

Схематическое изображение (без соблюдения масштаба) генома типичной конъюгативной R-плазмиды и ее диссоциации на два независимых репликона: RTF- и r-плазмиды. Гены резистентности к антибиотикам и их продукты охарактеризованы в табл. 5.3. Светлыми овалами обозначены области начала репликации (*ori*). Точками выделены сегменты, необходимые для конъюгации, а штриховкой - мобильные элементы.

Имеется еще один тип модулей, встречающийся в составе как конъюгативных, так и неконъюгативных плазмид: *Col*-модули. *Col*-модули кодируют ген одного из нескольких белков - колицинов (антибактериальных агентов, продуцируемых бактериями). Колицины различаются по структуре и способу действия. Плазмиды, кодирующие определенный колицин, часто содержат гены, обеспечивающие иммунитет к этому колицину, страхуя таким образом клетку-продуцент от повреждений, вызываемых ее собственным средством защиты.

Некоторые плазмиды содержат также модули,

кодирующие системы рестрикции-модификации. Одним из таких примеров является система *Eco* RI.

**Типичная плазмида.** Модульное строение типичных R-плазмид схематически представлено на рис. 5.1. Около 50% последовательности из приблизительно  $10^5$  пар оснований этих плазмид гомологичны одному из участков F-плазмиды как по структуре, так и по функциям. Эта область включает гены, необходимые для конъюгации (*tra*) и репликации со строгим контролем. В F-подобной области находится также дополнительный модуль, содержащий гены, контролирующие резистентность к тетрациклину. Вторая половина R-плазмиды не родственна F и содержит модули, ответственные за резистентность к стрептомицину, сульфаниламидам, хлорамфениколу и канамицину.

**Модули как подвижные элементы.** Плазмиды, выделенные независимо и даже в различных частях света, часто содержат близкородственные модули. Например, многие плазмиды несут одни и те же гены резистентности к антибиотикам. На основании этих данных возникло представление о том, что между геномами происходит обмен некоторыми генетическими модулями в виде интактных сегментов ДНК. Сейчас стало ясно, что в бактериальных плаزمиде происходит множество перестроек и обменов. Например, R-плазмиды, присутствующие в клетках определенных бактерий, отличных от *E. coli*, особенно в *Salmonella typhimurium* или *Proteus mirabilis*, могут диссоциировать на две отдельные плазмиды, каждая из которых содержит одну из двух разных частей (рис. 5.1). F-подобные последовательности образуют независимо реплицирующуюся конъюгативную плазмиду, обозначаемую RTF, а остальная часть - другую независимо реплицирующуюся, но неконъюгативную плазмиду, обозначаемую просто r. Поскольку как RTF, так и r являются независимыми репликонами, каждая плазмида должна содержать модуль ДНК, обеспечивающий репликацию. Подобные перестройки не являются простым лабораторным курьезом: RTF- и r-плазмиды обнаружены в природе. Быстрое распространение генов резистентности к антибиотикам объясняется именно обменом модулями между плазмидами. Эти и другие данные, полученные при изучении генетики бактерий, привели к идентификации дискретных подвижных элементов, названных **инсерционными последовательностями (IS)** и **транспозонами** (гл. 10), которые способны перемещаться не только между плазмидами, но и между плазмидными и клеточными геномами, а также в пределах самого бактериального генома. Обычно многие модули в плазмиде являются подвижными элементами или фланкированы ими (рис. 5.1).

**Плазмиды как векторы.** Многие плазмиды используются в качестве векторов для молекулярного



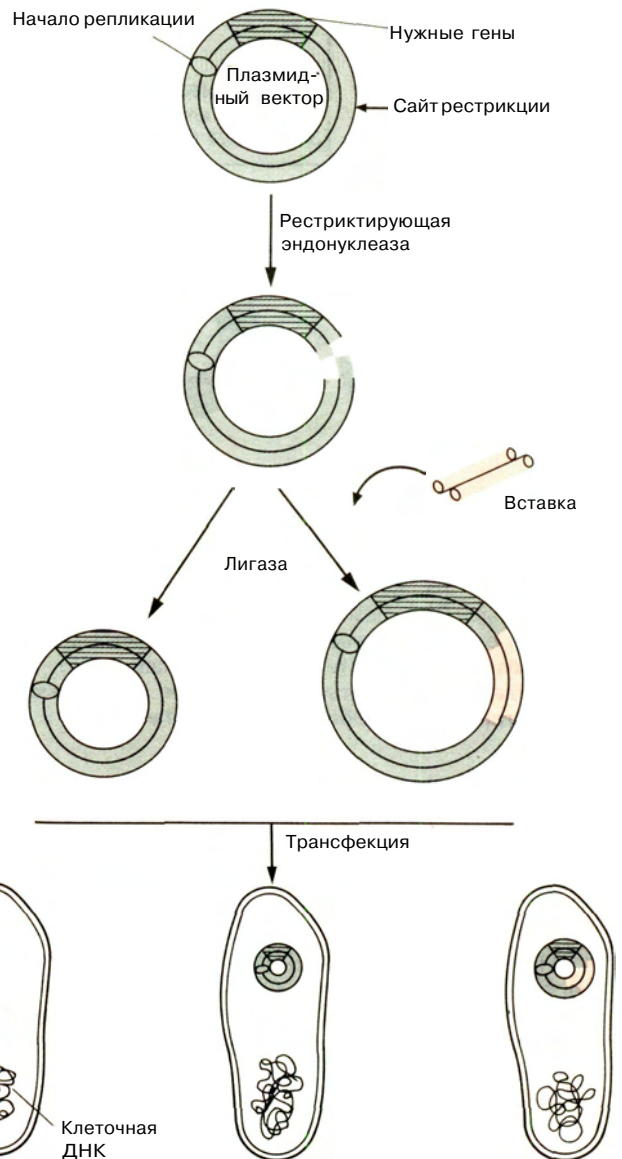
клонирования в *E. coli*. Плазмиды, исходно обнаруженные в природе, были впоследствии модифицированы, укорочены, реконструированы и подвергнуты рекомбинации как в клетках, так и в экспериментальных условиях. Целью всех этих манипуляций было получение универсальных плазмид или, напротив, плазмид, предназначенных для решения каких-то определенных экспериментальных задач. Лучшие из них содержали генетические маркеры, позволяющие упростить отбор рекомбинантных молекул.

## б. Конструирование векторов для отбора

Эффективность трансфекции *E. coli* K12 с помощью рекомбинантных плазмидных ДНК довольно низка. Какая-то часть клеток-хозяев вообще избегает трансфекции; другие клетки получают вектор без вставки, который либо не был разрезан рестриктирующей эндонуклеазой, либо подвергся лигированию, не включив вставки (рис. 5.2). Необходимо было разработать метод, позволяющий выявить или селективно вырастить те клетки, которые приобрели рекомбинантные векторы. Для этой цели в плазмидные векторы вводят гены, детерминирующие какие-то отличительные фенотипические особенности. Получив клетки, несущие рекомбинантные векторы, идентифицируют те, которые содержат нужную вставку, и клонируют их. Методы, используемые на этой конечной стадии молекулярного клонирования, описаны в гл. 6.

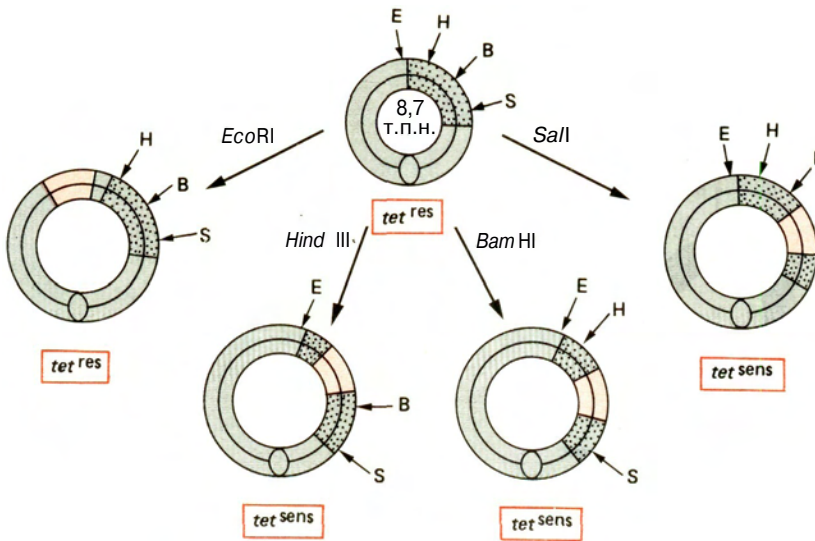
*Один селективный маркер.* В самых ранних экспериментах с рекомбинантными ДНК использовали плазмидный вектор pSC101. Он содержит всю информацию, необходимую для его репликации в клетках *E. coli*, а также гены устойчивости к тетрациклину. Кроме того, он несет по одному сайту узнавания для рестриктирующих эндонуклеаз *Eco* RI, *Hind* III, *Bam* HI и *Sal* I (рис. 5.3). Последние три сайта находятся в пределах области, ответ-

ственной за устойчивость к тетрациклину. При введении какого-либо сегмента ДНК в *Eco* RI-сайт после трансфекции образуются тетрациклиноустойчивые рекомбинанты. В среде, содержащей тетрациклин, нетрансфицированные клетки не образуют колоний в отличие от тех, в которых трансфекция имела место. Однако так происходит лишь с клетками, содержащими неизмененную плазмиду pSC101. Фрагмент ДНК, внедренный в любой из сайтов рестрикции, — *Bam* HI, *Sal* I или *Hind* III, — прерывает и инактивирует гены, кодирующие резистентность к тетрациклину, в результате чего появляются тетрациклинчувствительные трансфектанты.



**РИС. 5.2.**

При трансфекции *E. coli* с помощью рекомбинантных плазмидных ДНК одни клетки вообще не приобретают никакой плазмиды, другие получают исходную векторную молекулу без вставки, а третьи — рекомбинантную молекулу.

**РИС. 5.3.**

Структура плазмиды pSC101 и фенотипические свойства *E. coli* K12, трансфицированной этой плазмидой или различными ее рекомбинантными производными. В

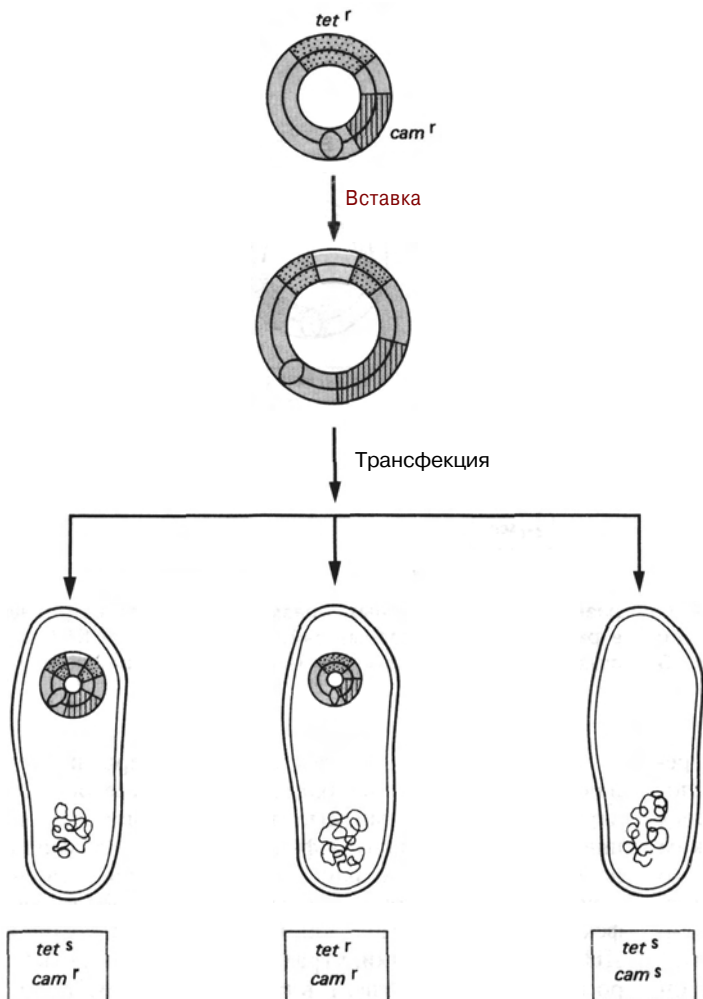
разных рекомбинантных плаزمидах вставка встроена в разные эндонуклеазные сайты. Клетки *E. coli* K12, содержащие плазмиды, чувствительны к тетрациклину.

Клетки, получившие при трансфекции такую рекомбинантную ДНК, отличаются от клеток, получивших pSC101, но не отличаются от исходных тетрациклинчувствительных клеток хозяина. Таким образом, присутствие в pSC101 только одного селективного маркера (резистентности к тетрациклину) не позволяет осуществить прямой отбор клеток, содержащих рекомбинантную плазмидную ДНК. Этот вектор имеет и другие недостатки: он содержит излишнюю ДНК, и, что более важно, регуляция его репликации осуществляется по строгому типу, так что в клетке-хозяине образуется лишь несколько копий плазмиды, т.е. выход рекомбинантной ДНК очень низок.

**Два селективных маркера.** При наличии двух селективных маркеров в плазмидном векторе можно провести отбор среди трех типов клеток (рис. 5.2): нетрансфицированных, получивших неизменную плазмиду или приобретших рекомбинантную плазмиду. Этот принцип отбора представлен на рис. 5.4, где используются два селективных маркера: гены устойчивости к антибиотикам тетрациклину и хлорамфениколу. Отбор основывается на индуцируемом вставкой нарушении, приводящем к инактивации только одного из маркированных генов. В приведенном случае нарушение произошло в гене резистентности к тетрациклину. Только те клетки, которые содержат рекомбинантную плазмиду, растут в присутствии хлорамфеникола и не способны расти в присутствии тетрациклина.

Помимо того что селективные маркеры позволяют выделять клоны бактериальных клеток, содержащие рекомбинантные плазмиды, они выполняют еще одну важную функцию. Если плазмиды не придает клетке-хозяину какие-то важные свойства, то клетки обычно размножаются более эффективно в отсутствие «лишней» внехромосомной ДНК. Поэтому клетки, утратившие плазмиду, быстро перерастают другие, и в результате рекомбинанты элиминируются. Другое дело, если клетки находятся в таких условиях, когда их жизнеспособность зависит от плазмиды. Например, клетки, содержащие рекомбинантный вектор, несущий ген резистентности к тетрациклину, успешно растут в среде, в которую добавлен тетрациклин, а клетки, утратившие вектор, элиминируются. Такое давление отбора на хозяйские клетки почти всегда оказывается необходимым для сохранения рекомбинантного плазмидного вектора.

**Свойства идеального плазмидного вектора.** Идеальный плазмидный вектор для молекулярного клонирования должен содержать минимальное количество ДНК, его репликация должна регулироваться по типу ослабленного, а не строгого контроля, чтобы обеспечить высокий выход ДНК; он должен содержать по крайней мере два селективных маркера и иметь только один сайт узнавания для одной рестриктирующей эндонуклеазы. Последнее свойство позволяет разрезать кольцевую ДНК в уникальном сайте, по которому затем и будет

**РИС. 5.4.**

Преимущества использования двух селективных маркеров. При индуцируемом вставкой нарушении одного из маркеров все три типа трансфицированных клеток имеют разные фенотипы.

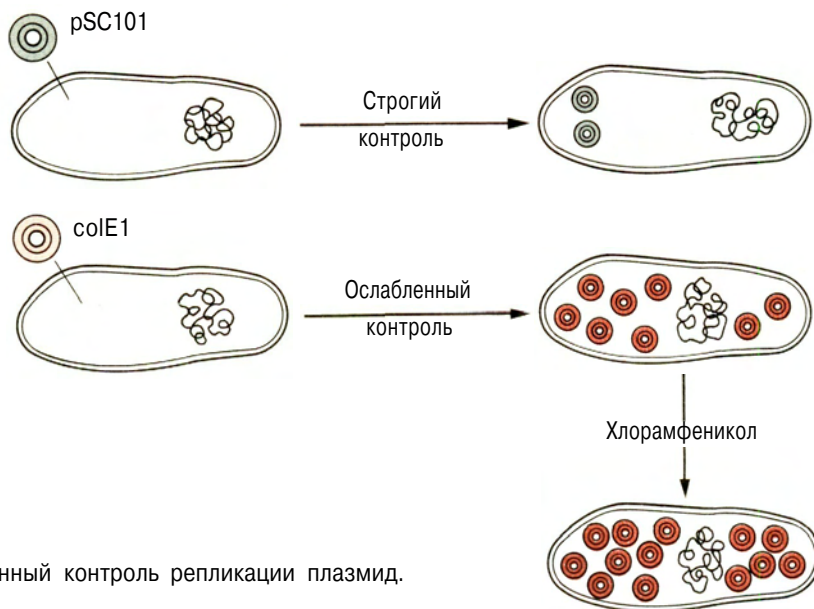
осуществлено лигирование со вставкой. Чтобы максимально облегчить отбор, уникальный сайт рестрикции необходимо включить в один из двух селективных маркеров. Вставка не должна прерывать последовательность, существенную для сохранения самой плазмиды. Векторы, которые приближаются к такому идеалу, были сконструированы из встречающихся в природе модулей ДНК с помощью классических генетических методов и методов, используемых при работе с рекомбинантными ДНК.

Модули, ответственные за репликацию с ослабленным контролем, встречаются в различных природных плаزمиды. Одна из таких плазмид, *colE1* (табл. 5.2), имеет еще одно преимущество: ее репликация осуществляется даже в условиях, когда синтез белков и хромосомной ДНК *E. coli* подавлен с помощью аминокислотного голодания или антибиотика хлорамфеникола (репликация *colE1*-типа описана в гл. 2). При таких условиях клетки не

делятся, однако в них накапливается несколько тысяч копий плазмидного генома (рис. 5.5). В результате из 1 л культуральной жидкости можно с легкостью получить около 1 мг плазмидной ДНК. Сама плазида *colE1* не часто используется в качестве вектора, но модуль *colE1*-репликации встроен в большое число других плазмид. Все такие плазмиды имеют преимущества, связанные с ослабленным контролем репликации, т.е. с накоплением плазмид при подавлении белкового синтеза. Некоторые из этих плазмид перечислены в табл. 5.4.

### в. Плазмидный вектор pBR322

Из всех векторов при работе с *E. coli* чаще всего используется плазида pBR322 (табл. 5.4). Этот вектор был сконструирован с помощью классических генетических методов (*in vivo*) и методов, применяемых при работе с рекомбинантными ДНК, и имеет многие свойства идеального плазмидного



**РИС. 5.5.**  
Строгий и ослабленный контроль репликации плазмид.

вектора для клонирования (рис. 5.6). Установлена полная нуклеотидная последовательность этого вектора длиной 4362 п.н. Поскольку pBR322 реплицируется по colE1-типу, удастся получить плазмидную ДНК с высоким выходом. Кроме того, плазида содержит два селективных маркера: гены устойчивости к ампициллину и тетрациклину. Каждый из 12 сайтов узнавания рестриктирующими эндонуклеазами встречается в молекуле лишь один раз, и это позволяет получать полноразмерные

линейные молекулы с разнообразными липкими концами. При трансфекции клеток *E. coli* K12 плазмидой pBR322 может образоваться более  $10^8$  успешно **трансформированных** клеток на 1 мкг плазмидной ДНК в зависимости от используемого штамма и условий. После внедрения сегментов ДНК в вектор pBR322 *in vitro* эффективность трансфекции падает на один или два порядка, что опять-таки определяется штаммом-хозяином и условиями трансфекции.

Таблица 5.4. Некоторые плазмидные векторы, используемые при работе с *E. coli*

Плазмидный вектор	Контроль репликации	Размер, п.н.	Селективные маркеры и некоторые уникальные сайты рестрикции <sup>1)</sup>
pSC101	Строгий	8700	Ген резистентности к тетрациклину ( <i>Hind</i> III, <i>Bam</i> HI, <i>Sal</i> I); <i>Eco</i> RI
pUC7	Ослабленный	2700	Ген резистентности к ампициллину <i>lacZα</i> <sup>2)</sup> ( <i>Eco</i> RI, <i>Bam</i> HI; <i>Hinc</i> II, <i>Pst</i> I)
pMB9	»	5500	Ген резистентности к тетрациклину ( <i>Hind</i> III, <i>Bam</i> HI, <i>Sal</i> I)
pBR322	»	4362	Ген резистентности к колицину <sup>3)</sup> ; <i>Eco</i> RI; <i>Hpa</i> I; <i>Sma</i> I Ген резистентности к тетрациклину ( <i>Hind</i> III, <i>Bam</i> HI, <i>Sal</i> I)
pDF41	Строгий	12800	Ген резистентности к ампициллину ( <i>Pst</i> I, <i>Pvu</i> I) <i>Eco</i> RI; <i>Ava</i> I; <i>Pvu</i> II; <i>Cla</i> I
pRK2501	»	11100	<i>trpE</i> -ген; <i>Bam</i> HI, <i>Eco</i> RI, <i>Hind</i> III, <i>Sal</i> I Ген резистентности к тетрациклину ( <i>Sal</i> I) Ген резистентности к канамицину ( <i>Hind</i> III, <i>Xho</i> I)

<sup>1)</sup> Сайты рестрикции, находящиеся внутри селективных маркеров, указаны в скобках после маркера.

<sup>2)</sup> *lacZα* содержит промотор и аминоконцевую часть β-галактозидазного гена (*lacZ*) *E. coli* (разд. 5.1.г и 5.2.г).

<sup>3)</sup> Колицин - это токсин, кодируемый некоторыми плазмидами; он активен в отношении *E. coli*.

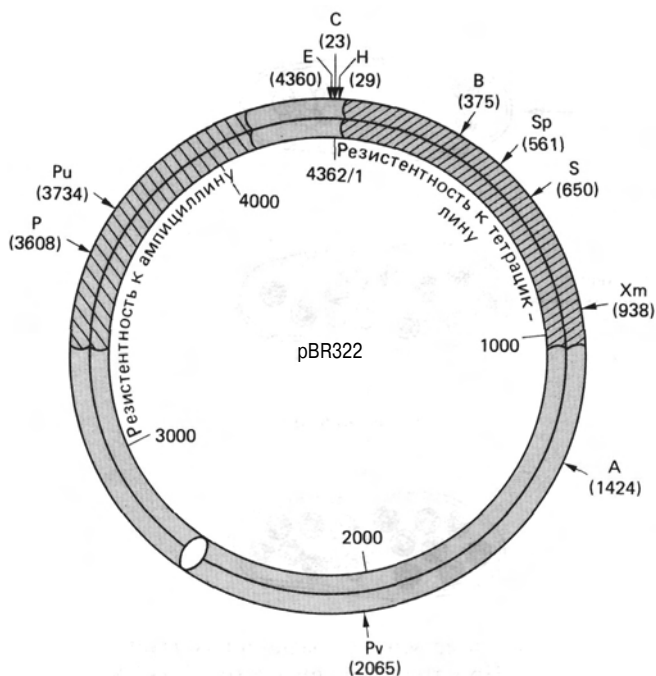


РИС. 5.6.

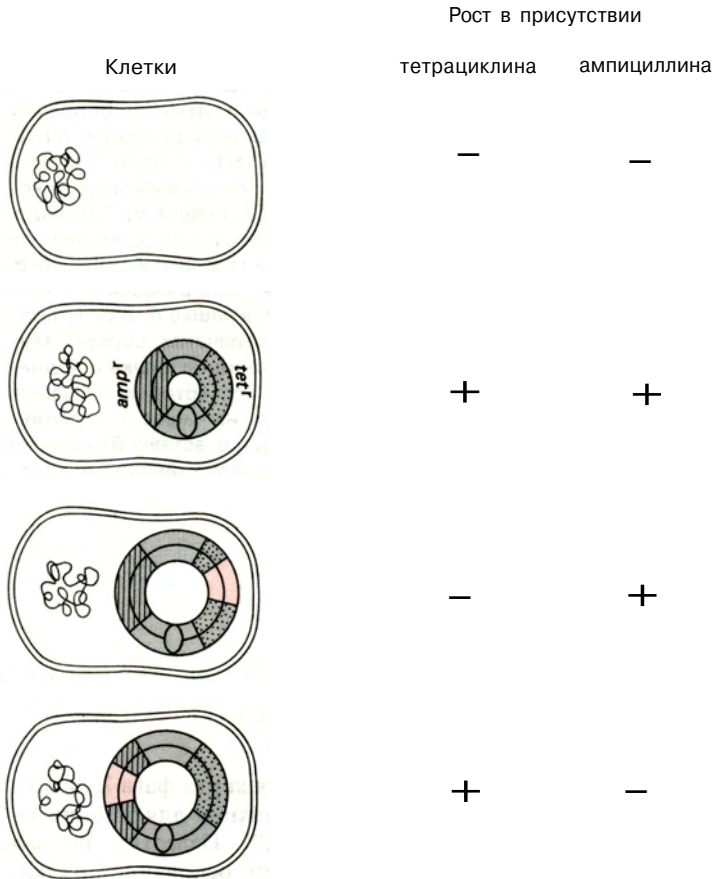
Плазмида pBR322. Указаны сайты узнавания для эндонуклеаз рестрикции, каждый из которых встречается в молекуле только один раз. Цифры около сайтов - число пар оснований.

Клетки *E. coli* K12, содержащие pBR322, выращивают в среде, содержащей ампициллин или тетрациклин либо оба этих антибиотика; при этом клетки, утратившие плазмиду, не растут в присутствии любого из двух антибиотиков (рис. 5.7). Если какой-либо сегмент ДНК внедряется в область pBR322, кодирующую устойчивость к тетрациклину (например, в сайт рестрикции для *Vam* H1 или *Sal*I), то трансформанты сохраняют способность расти на среде с ампициллином, но не растут на среде с тетрациклином. Если же вставка происходит в ампициллиновом гене (по сайтам рестрикции для *Pvu* I или *Pst* I), то наблюдается обратная ситуация. При любой локализации вставки происходит отбор только тех бактериальных колоний, которые содержат рекомбинантную плазмидную ДНК. Отбор осуществляют, посеяв трансфицированные клетки *E. coli* K12 на чашки с агаром, содержащим антибиотик, к которому клетки, несущие плазмиду, должны быть устойчивы. Все выросшие колонии будут содержать плазмиду. Клетки из таких резистентных к антибиотику клонов тестируют затем на способность расти в среде, содержащей второй антибиотик. Такая проверка позволяет отличить колонии, содержащие исходную плазмиду pBR322 (есть рост), от колоний, несущих рекомбинантную плазмиду (нет роста). Все эти операции можно выполнить за один или два дня с использованием метода отпечатков (рис. 5.8). Даже одну успешно трансформированную клетку из  $10^8$  высеянных можно обнаружить по появлению отдельной ко-

лонии. Преимущество использования вектора с двумя селективными маркерами становится очевидным, если сравнить получение клонов с рекомбинантными плазмидами в двух случаях: когда применяются векторы pBR322 или pSC101 (ср. данные, представленные на рис. 5.3 и 5.7)

#### г. Другие векторы, используемые для разных целей

Помимо pSC101 и pBR322 сконструировано множество других плазмидных векторов, применяемых для различных целей. В табл. 5.4 перечислены некоторые из них. Вектор pUC7 значительно меньше по размеру, чем pBR322, однако в нем сохранены необходимые компоненты, включая ген устойчивости к ампициллину. Вторым селективным маркером является часть гена *lacZ* (*lacZa*), которая вместе с остальной частью *lacZ*-гена клетки-хозяина ответственна за голубую окраску колоний, выращиваемых в среде с Xgal (разд. 5.1.г). Все уникальные сайты рестрикции, используемые для внедрения фрагмента ДНК, прерывают сегмент *lacZa* в векторе pUC7, в результате чего клетки с рекомбинантным вектором образуют на среде с Xgal «белые» (неокрашенные) колонии. Вектор pMB9 с ослабленным контролем репликации содержит два селективных маркера, один из которых кодирует устойчивость к колицину E1. Клетки, несущие pMB9, растут на среде, которая содержит очищенный колицин E1, а бесплазмидные клетки на этой

**РИС. 5.7.**

Фенотипические свойства клеток, содержащих плазмиду pBR322 со вставками и без них.

среде погибают. Модули репликации и устойчивости к тетрациклину у pMB9 такие же, как у плазмиды pBR322.

Вектор совершенно другого типа, pDF41, содержит модуль репликации, полученный из конъюгативной плазмиды F, и хромосомный ген (*trpE*) *E. coli*, кодирующий фермент, который участвует в биосинтезе триптофана. Это позволяет отбирать трансформантов, используя хозяйские клетки *trpE*<sup>-</sup> и среду, лишенную триптофана. Образование малого числа копий pDF41 оказывается полезным в экспериментах по изучению регуляции клонированных бактериальных генов в *E. coli*, поскольку при большом числе копий регуляторные эффекты могут маскироваться. Вектор pRK2501 ведет свое начало от плазмиды RK2, относящейся к плазмидам группы P. Отличительной чертой P-плазмид является их способность к конъюгационному переносу в грам-отрицательные бактерии многих родов (например, *Serratia marcescens*, *Pseudomonas aeruginosa*). Благодаря этому рекомбинантные плазмиды, сконструированные на основе pRK2501, могут использоваться в исследованиях, проводимых на многих видах бактерий.

### 5.3. *E. COLI*-СИСТЕМЫ: ФАГОВЫЕ ВЕКТОРЫ

#### а. Некоторые различия между плазмидными и фаговыми векторами

Когда используются фаговые векторы, жизнеспособным продуктом, содержащим сконструированную рекомбинантную ДНК, является популяция фаговых частиц. В отличие от ситуации с плазмидными векторами, когда производятся клонирование и отбор содержащих их клеток, рекомбинантный фаговый геном можно клонировать непосредственно. При таком клонировании образуются бляшки на газоне чувствительных клеток-хозяев (рис. 11.13). Фрагменты чужеродной ДНК, встроенные в популяцию фаговых векторных ДНК *in vitro*, и соответствующие рекомбинантные фаговые геномы можно затем ввести в клетки перmissive хозяина в виде изолированной ДНК (трансфекция) или реконструированных вирусных частиц (инфекция). В любом случае в клетке, инфицированной одной молекулой

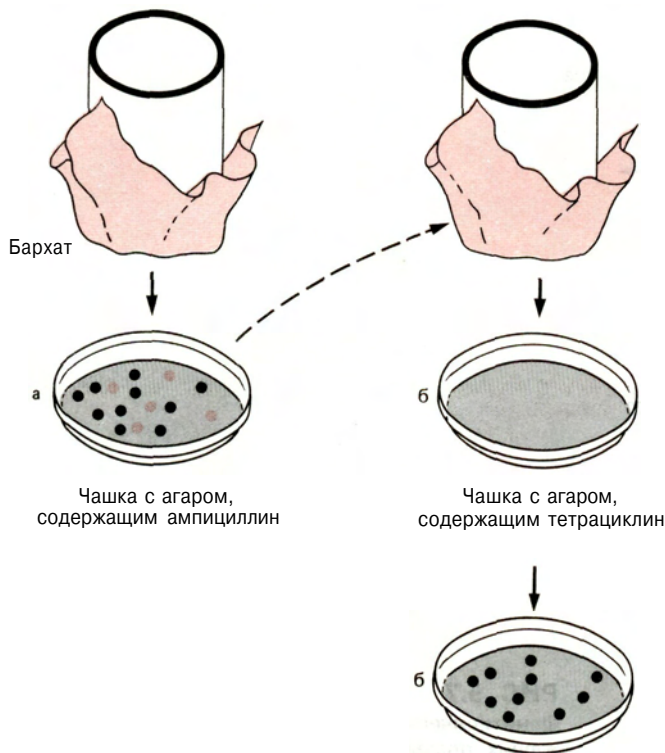


РИС. 5.8.

Выявления колоний, обладающих определенными фенотипическими свойствами, методом отпечатков. Клетки *E. coli* были трансфицированы плазмидой pBR322, содержащей вставку в тетрациклиновом гене (сайт *Tet* H1). Клетки, несущие такую плазмиду, образуют колонии на агаре, в который добавлен ампициллин (чашка а). Цилиндр, на торец которого натянут кусок стерильного бархата, прижимают сначала к агару в чашке а, а затем к свежему агару, содержащему тетрациклин, но не ампициллин, в чашке б. В результате на чашке б образуется отпечаток, соответствующий распределению колоний на чашке а. Однако некоторые из колоний (те, которые на чашке а окрашены) не могут расти на чашке б, поскольку они содержат плазмиду со вставкой, которая привела к инактивации генов устойчивости к тетрациклину.

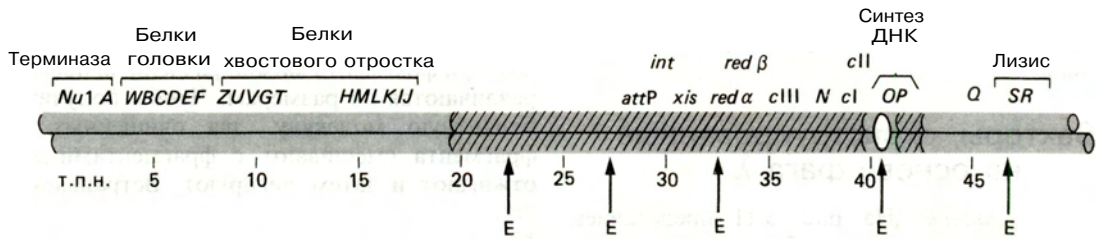
рекомбинантного вирусного вектора, образуется вирусное потомство, которое будет инфицировать близлежащие клетки и в конце концов лизировать их, приводя к формированию бляшек. Образование таких бляшек само по себе свидетельствует об успехе инфекции (или трансфекции). Однако для идентификации рекомбинантного фага, содержащего нужную вставку, часто бывают нужны дополнительные селективные условия или скрининг бляшек на содержание определенной последовательности ДНК (гл. 6).

Выбор типа вектора—плазмидного или фагового—зависит от цели эксперимента, размера внедряемого фрагмента и относительного содержания нужного фрагмента в применяемой смеси фрагментов ДНК. В целом фаговые векторы более эффективны, чем плазмидные, при клонировании крупных вставок, и скрининг большого числа негативных колоний на содержание специфической вставки провести проще, чем скрининг большого количества бактериальных колоний (гл. 6). Самыми распространенными векторами для *E. coli* являются два бактериофага —  $\lambda$  и M13.

## б. Фаг $\lambda$

*Геном фага  $\lambda$ .* Наиболее часто применяются фаговые векторы, полученные на основе умеренного

фага  $\lambda$  (гл. 2 и 3). Генетическая карта фага  $\lambda$  дикого типа с указанием сайтов рестрикции для эндонуклеаз представлена на рис. 5.9. Известна полная последовательность 48502 пар оснований, из которых состоит геном. Молекула ДНК фага  $\lambda$  содержит область начала репликации (*ori*), а также гены структурных вирусных белков (головки и хвостового отростка) и ферментов, участвующих в репликации ДНК, лизисе инфицированных клеток и установлении лизогении (разд. 3.11д). Приблизительно 1/3 генома, представленная непрерывным внутренним сегментом, находящимся на карте между точками 20 и 38 т.п.н., необязательна для успешной литической инфекции клеток; этот сегмент содержит гены, необходимые для установления лизогенного состояния. Необязательной является и другая небольшая по размеру область вблизи точки 43 т.п.н. Однако для успешной упаковки ДНК фага  $\lambda$  в вирионе ее длина должна быть больше, чем 38 т.п.н., и меньше, чем 52 т.п.н. Поэтому существенно, чтобы у всех векторов, сконструированных на основе ДНК фага  $\lambda$ , вся центральная необязательная область или ее часть была замещена сегментом соответствующей длины. Обычно многие или даже все гены, требующиеся для установления лизогенного состояния, удаляются, поэтому инфицирование всегда сопровождается лизисом клеток хозяина. Для упрощения конструирования рекомбинантных молекул используют производные фага  $\lambda$  дикого типа,



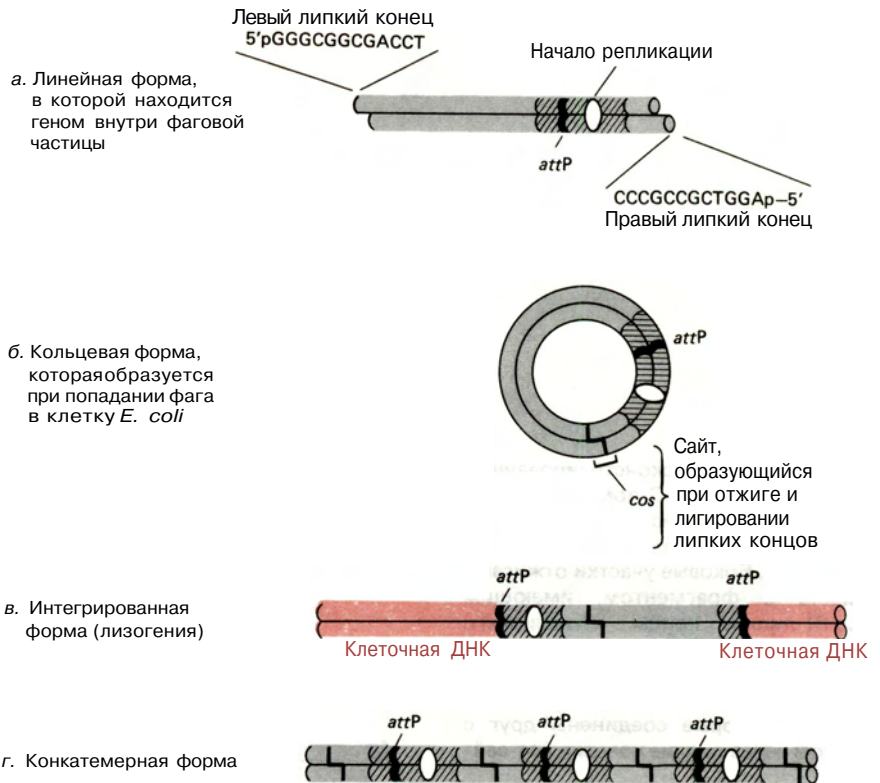
**РИС. 5.9.**

Линейный геном бактериофага  $\lambda$ . Заштрихованы гены, необязательные для литической инфекции. Светлый овал - область начала репликации. E - Eco RI-сайты.

утратившие некоторые из обычных сайтов рестрикции.  $\lambda$ -Векторы, в которые можно встраивать сегменты чужеродной ДНК путем замещения необязательной части генома без нарушения существенных областей, были получены из генома дикого типа с помощью мутаций и рекомбинации *in vivo*, а также методов, применяемых при работе с рекомбинантными ДНК.

**Множественные формы генома.** Во время своего жизненного цикла  $\lambda$ -геном может находиться в четырех разных формах. Внутри фаговой частицы ДНК находится в линейной форме (рис. 5.9 и 5.10.а). Линейная ДНК имеет одноцепочечные концы дли-

ной 12 нуклеотидов, способные к взаимному спариванию. При попадании в клетки *E. coli* геном фага  $\lambda$  замыкается в кольцо и сшивается с помощью ДНК-лигазы (рис. 5.10.б). Места спаривания липких концов получили название *cos*-сайтов. Кольцевая форма обязательна для осуществления репликации вирусной ДНК; она также служит предшественником при лизогенной интеграции ДНК фага в геном хозяина (рис. 1.15, В и разд. 2.4.г). После интеграции геном фага  $\lambda$  снова становится линейным (рис. 5.10.в). Однако благодаря удаленности сайта интеграции на фаговой ДНК (*attP*) от *cos*-сайта интегрированная линейная ДНК профага оказывается перемутированной относительно ДНК, находящейся в фаговой частице. Наконец, во время репликации ДНК фага  $\lambda$  по типу «катящегося кольца» образуются длинные конкатемеры, состоящие из многих



**РИС. 5.10.**

Четыре формы генома бактериофага  $\lambda$ . Все обозначения такие же, как на рис. 5.9.

г. Конкатемерная форма



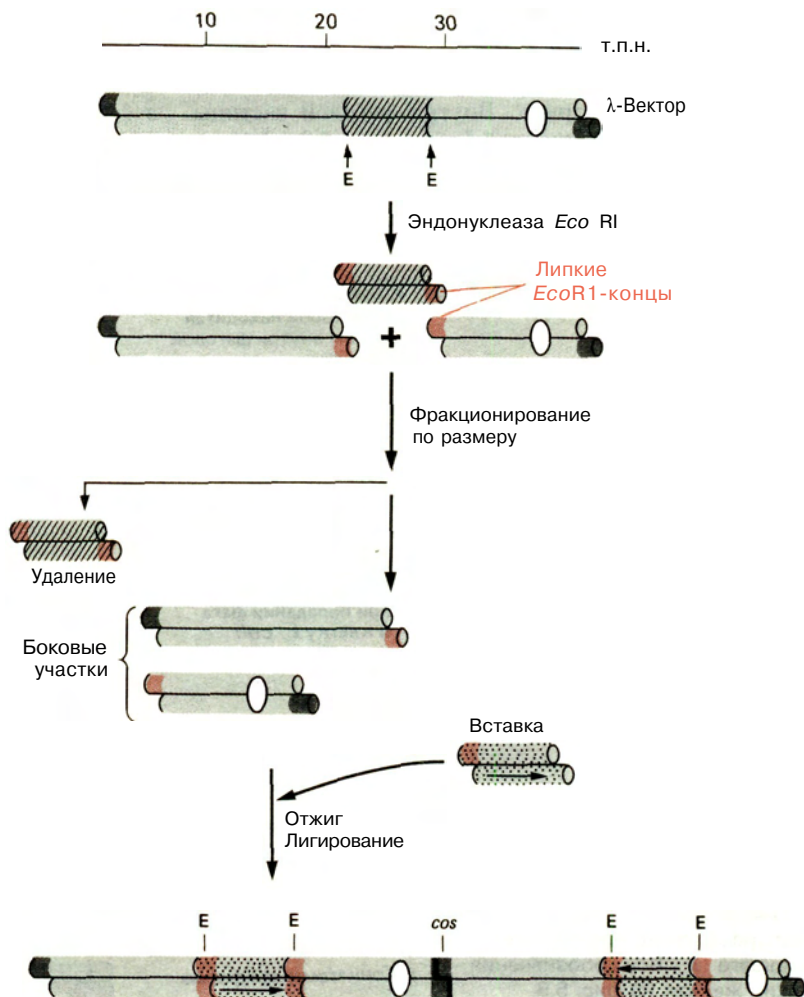
копий генома  $\lambda$  (рис. 2.32), которые разрезаются в *cos*-сайтах при образовании зрелой фаговой ДНК (рис. 5.10.г).

### в. Векторы, сконструированные на основе фага $\lambda$

**Типичные векторы.** На рис. 5.11 представлен один из обычных  $\lambda$ -векторов. Он короче генома фага  $\lambda$  дикого типа более чем на 10 т.п.н. Небольшая необязательная область вблизи 43 т.п.н. и часть центральной необязательной области в нем делегированы. Все гены, необходимые для литической инфекции, остаются интактными, и молекула имеет достаточно большой размер, чтобы упаковываться в частицы бактериофага. Сохранены только два из пяти исходных сайтов для *Eco* RI; они фланкируют необязательную часть ДНК протяженностью 7 т.п.н. После разрезания с помощью *Eco* RI образуются три фрагмента ДНК, один из которых соот-

ветствует левому плечу генома, второй—правому, а третий—необязательной центральной области. Эти три фрагмента можно разделить, поскольку они различаются по размерам. Чтобы получить рекомбинантную молекулу, два очищенных плечевых фрагмента смешивают с фрагментами-вставками, отжигают и затем лигируют. Встраиваемая ДНК должна иметь *Eco* RI-липкие концы и размер от 5 до 17 т.п.н., чтобы образовалась молекула, способная к упаковке. Следует заметить, что во время отжига и лигирования фрагменты могут соединиться двумя разными способами. Встраиваемая ДНК соединяется с двумя плечевыми фрагментами в любой из двух ориентаций с помощью липких *Eco* RI-концов, а плечевые фрагменты соединяются друг с другом с помощью своих липких  $\lambda$ -концов. Продуктом объединения является **конкатемер**—предшественник для упаковки вирусной ДНК в зрелые фаговые частицы.

Для разных целей было сконструировано мно-



**РИС. 5.11.**

Типичный вектор, сконструированный на основе фага  $\lambda$ . С помощью эндонуклеазы *Eco* RI из генома фага  $\lambda$  вырезают необязательную центральную часть. Боковые участки отжигают вместе с фрагментом, имеющим *Eco* RI-липкие концы, и лигируют. В результате образуются конкатемерные структуры, состоящие из боковых участков генома фага  $\lambda$  и вставки, которые соединены друг с другом через липкие концы и *cos*-сайты.

жество разных  $\lambda$ -векторов. Все они характеризуются двумя основными свойствами: 1) сами векторные молекулы могут размножаться, подобно фагу, в процессе литической инфекции, что позволяет легко получить векторную ДНК; 2) строго фиксированное положение сайтов рестрикции позволяет осуществить разрезание или удаление центральной несущественной области, а также получить сайты для лигирования вставки. Во многих векторах, сконструированных на основе ДНК фага  $\lambda$ , используются не *Eco* RI-, а другие сайты рестрикции, позволяющие получить нужные плечевые фрагменты и провести лигирование с разнообразными вставками. Различные  $\lambda$ -векторы включают вставки разных размеров. Те векторные молекулы, в которых делегирована вся несущественная область, способны включать вставки максимального размера—немногим более 24 т.п.н.

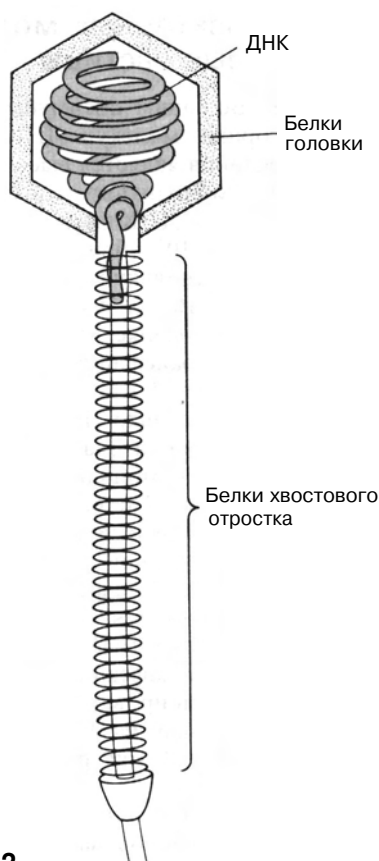
*Отбор рекомбинантного фага.* Как и плазмидные векторы, нужные рекомбинантные фаги отбираются из большого разнообразия побочных продуктов. Первый этап отбора основан на получении негативных колоний на газоне чувствительных клеток, поскольку не все возможные продукты лигирования способны осуществлять инфицирование. Например, в случае вектора, изображенного на рис. 5.11, побочные продукты объединения включают конкатемеры, содержащие только два плеча фага  $\lambda$ . Поскольку расстояние между *cos*-сайтами в этих молекулах меньше 38 т.п.н., они не могут упаковываться в фаговые частицы. Если  $\lambda$ -векторы содержат более протяженные плечевые фрагменты или если до отжига и лигирования неполностью удален несущественный сегмент ДНК фага  $\lambda$ , необходимо применять дополнительные методы отбора, поскольку и в том и в другом случае фаговая частица, не несущая вставку, может образовать бляшку. Чтобы отличить рекомбинантный фаг от нерекомбинантного, в некоторые  $\lambda$ -векторы включают сегмент ДНК, содержащий промотор, оператор и ген  $\beta$ -галактозидазы (*lac Z*) из *lac*-оперона *E. coli* (разд. 3.11.в). Большая часть гена *lacZ* удаляется из исходной ДНК фага  $\lambda$  при вырезании несущественной центральной области с помощью рестриктирующей эндонуклеазы. Если векторная ДНК соответствующим образом восстановлена и содержит ген  $\beta$ -галактозидазы, то на чашках с агаром, содержащим Xgal, появляются голубые бляшки; если же рекомбинантная векторная ДНК не содержит  $\beta$ -галактозидазного гена, то на той же среде образуются бесцветные бляшки. Таким образом, фаг, содержащий вставку, образует на индикаторных чашках бляшки другой окраски (разд. 5.1.г).

## г. Упаковка $\lambda$ -векторных молекул в фаговые частицы

*Упаковка in vitro.* Большим преимуществом использования  $\lambda$ -векторных молекул ДНК для молекулярного клонирования является высокая эффективность переноса рекомбинантных геномов в клетки *E. coli* K12. Теоретически ДНК фага  $\lambda$  дикого типа может дать примерно  $2 \cdot 10^{10}$  бляшкообразующих единиц на 1 мкг ДНК, поскольку мол. масса генома составляет примерно  $3 \cdot 10^7$  и каждая молекула способна дать начало бляшке. Однако клетки, обработанные  $\text{CaCl}_2$  и трансфицированные с помощью ДНК фага  $\lambda$  (разд. 5.1.в), способны образовать на 1 мкг ДНК только  $10^4$ – $10^6$  бляшек. Эффективность же бляшкообразования при использовании лигированных рекомбинантных ДНК на несколько порядков ниже. Такой низкий выход заставляет использовать относительно большие количества вставок и векторной ДНК при лигировании и трансфицировать большое число клеток, чтобы создать возможность для нахождения нужного сегмента ДНК. Однако ДНК фага  $\lambda$  можно упаковать с образованием инфекционных фаговых частиц *in vitro*, а при инфицировании этими частицами образуются  $10^8$  бляшек на 1 мкг ДНК фага  $\lambda$  дикого типа или  $10^6$  на 1 мкг рекомбинантных молекул ДНК фага  $\lambda$ .

*Конструирование фаговых головок и хвостовых отростков.* ДНК фага  $\lambda$  упакована в гексагональную **головку**, образованную несколькими белками (рис. 5.12). К головке прикреплен хвостовой отросток, построенный из других белков. Белки головки и хвостового отростка кодируются геномом фага  $\lambda$  (рис. 5.9) и синтезируются в клетках бактерий, инфицированных фагом. При литической инфекции происходит объединение белков головки с образованием пустых головок (рис. 5.13). Эти головки взаимодействуют с комплексом, образуемым конкатемерной  $\lambda$ -ДНК и фаговым белком **терминазой**, которая специфически связывается с ДНК вблизи *cos*-сайтов. Затем ДНК упаковывается в головку, начиная с левого конца молекулы, при этом полноразмерный линейный геном фага  $\lambda$  образуется в результате эндонуклеазного разрезания в *cos*-сайтах, катализируемого терминазой. В результате образуются липкие концы, типичные для зрелой линейной ДНК фага  $\lambda$ . На последнем этапе происходит прикрепление к головке уже сформированных хвостовых отростков, и процесс образования фаговой частицы завершается.

ДНК фага  $\lambda$  дикого типа или рекомбинантная ДНК упаковывается *in vitro*, если смешать при определенных условиях конкатемерную ДНК и предварительно сформированные пустые головки, хвостовые отростки и терминазу. Поскольку при

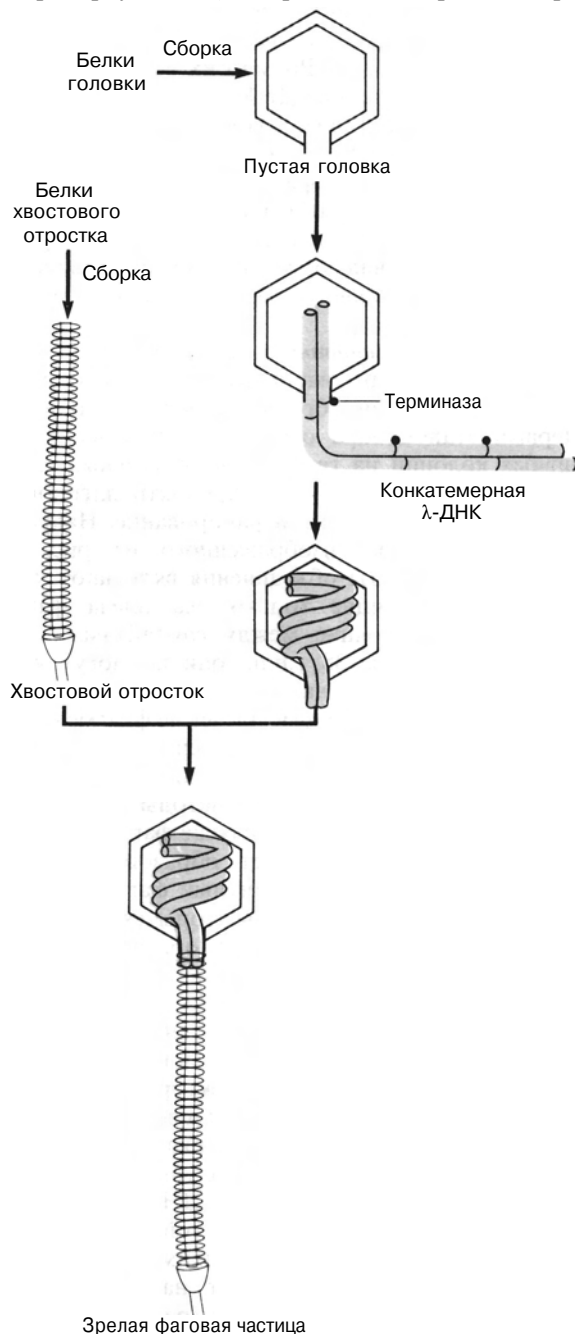


**РИС. 5.12.**  
Строение бактериофага  $\lambda$ .

инфицировании фагом  $\lambda$  дикого типа все эти белки быстро расходуются на образование новых частиц фага  $\lambda$ , для получения достаточных количеств головок, хвостовых отростков и терминазы, не связанной с ДНК, требуются специальные условия. Для этого выращивают культуры *E. coli*, инфицированные двумя разными штаммами фага  $\lambda$ , каждый из которых содержит мутацию в определенном гене (рис. 5.14): один из мутантов не способен синтезировать белок А (один из полипептидов терминазы), другой – белок Е головки. Обе культуры содержат большое количество других белков головки и хвостового отростка, но ни в одной из них не могут образовываться зрелые фаговые частицы. При смешивании экстрактов клеток, инфицированных указанными штаммами, образуются интактные пустые головки. Если в такой экстракт добавить фаговую ДНК (дикого типа или рекомбинантную), то произойдет ее упаковка и прикрепление к головке хвостового отростка, т.е. все те события, которые наблюдаются при инфицировании клеток бактериофагом дикого типа (рис. 5.13).

#### д. Фаг М13

Фаг М13 входит в группу нитчатых бактериофагов *E. coli*. Его геном представлен одноцепочечной кольцевой молекулой ДНК длиной примерно 6400 нуклеотидов, который упакован в напоминающий пробирку капсид, построенный по крайней мере



**РИС. 5.13.**  
Сборка фага  $\lambda$ .

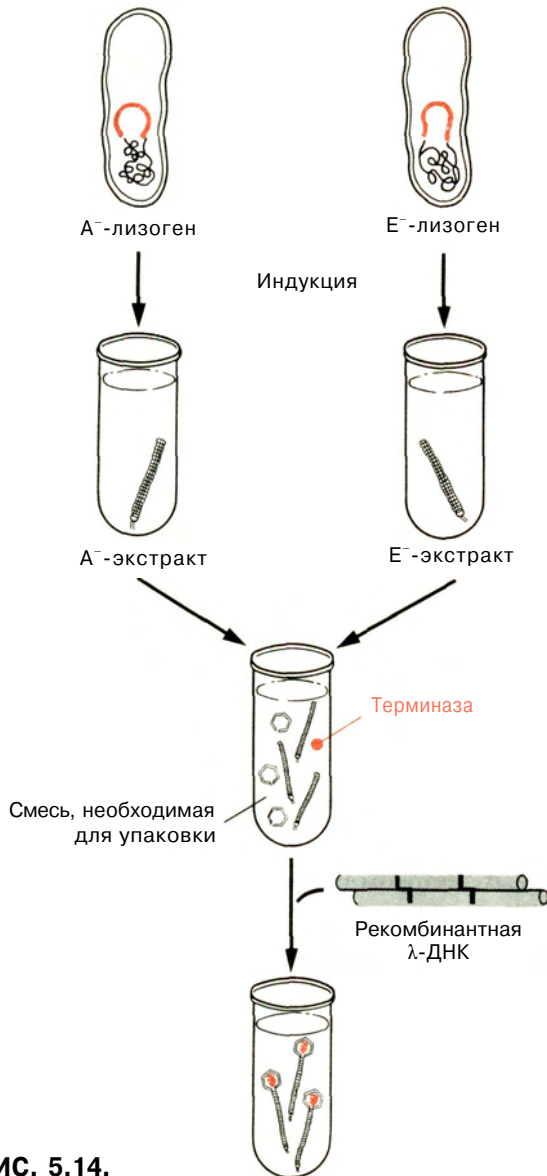


РИС. 5.14.

Приготовление экстрактов для осуществления упаковки *in vitro* ДНК фага  $\lambda$ . Используют два штамма *E. coli*, каждый из которых лизогенен в отношении определенного мутантного штамма фага  $\lambda$ . Один из мутантов не способен синтезировать белок А, другой - белок Е. Оба этих белка необходимы для упаковки ДНК фага  $\lambda$ . А<sup>-</sup> и Е<sup>-</sup>-экстракты смешивают и добавляют конкатемерную ДНК фага  $\lambda$ , которая связывается с терминазой прежде, чем происходит разрезание в *cos*-сайтах, и упаковывается в фаговые головки.

из трех разных кодируемых фагом белков. Фаг М13 способен инфицировать клетки *E. coli* только в том случае, если последние имеют тонкие, напоминающие волоски структуры, называемые **F-пилями**.

F-пили необходимы для прикрепления фага к клетке (рис. 5.15). Поскольку гены, ответственные за синтез F-пилей, кодируются половым фактором F, только F<sup>+</sup>-клетки проявляют чувствительность к фагу М13. Первым событием при репликации фага М13 в клетках *E. coli* является превращение одноцепочечной ДНК в двухцепочечную кольцевую форму (рис. 2.3; разд. 2.1.ж), т.е. синтез цепи, комплементарной инфицирующей цепи ДНК. Двухцепочечная ДНК детерминирует синтез примерно 300 своих копий, использующихся затем в качестве матриц при образовании одноцепочечных фаговых ДНК, упаковывающихся в фаговые частицы. Параллельно транскрибируются и транслируются фаговые гены, а на внутренней мембране клетки концентрируется основной капсидный белок. Зрелые одноцепочечные молекулы фаговой ДНК проходят через мембрану, захватывая при этом капсидный белок, в результате чего образуется большое количество зрелых фаговых частиц. Хотя инфицированные клетки не погибают, их рост замедляется, и на бактериальном газоне образуются мутные бляшки. Двухцепочечную репликативную форму ДНК можно легко выделить из инфицированных клеток, с тем чтобы использовать ее затем в качестве вектора, а одноцепочечная ДНК содержится в большом количестве в фаговых частицах в культуральной среде. Каждая клетка в каждом поколении высвобождает несколько сотен фаговых частиц. Одноцепочечные рекомбинантные молекулы являются особенно удобным объектом при изучении нуклеотидных последовательностей (гл. 7).

### е. Векторы, сконструированные на основе фага М13

Как и в случае фага  $\lambda$ , путем различных модификаций генома фага М13 дикого типа (рис. 5.16) были сконструированы разные удобные векторы. Для получения рекомбинантов использовались двухцепочечные репликативные формы, о которых говорилось выше. Одноцепочечные молекулы ДНК не применяют в качестве векторов, поскольку обычно их нельзя разрезать с помощью эндонуклеаз рестрикции типа II. Большая часть генома, состоящего примерно из 6400 нуклеотидных остатков, содержит гены, кодирующие белки оболочки, а также белки, участвующие в репликации и сборке. Область начала репликации состоит примерно из 150 нуклеотидов, находящихся внутри некодирующего сегмента длиной 507 нуклеотидов. В этот сегмент может быть помещена вставка без нарушения способности М13 к репликации.

*Типичный вектор.* Вектор, представленный на рис. 5.17, сконструирован на основе ДНК фага М13. Для того чтобы было удобно проводить отбор

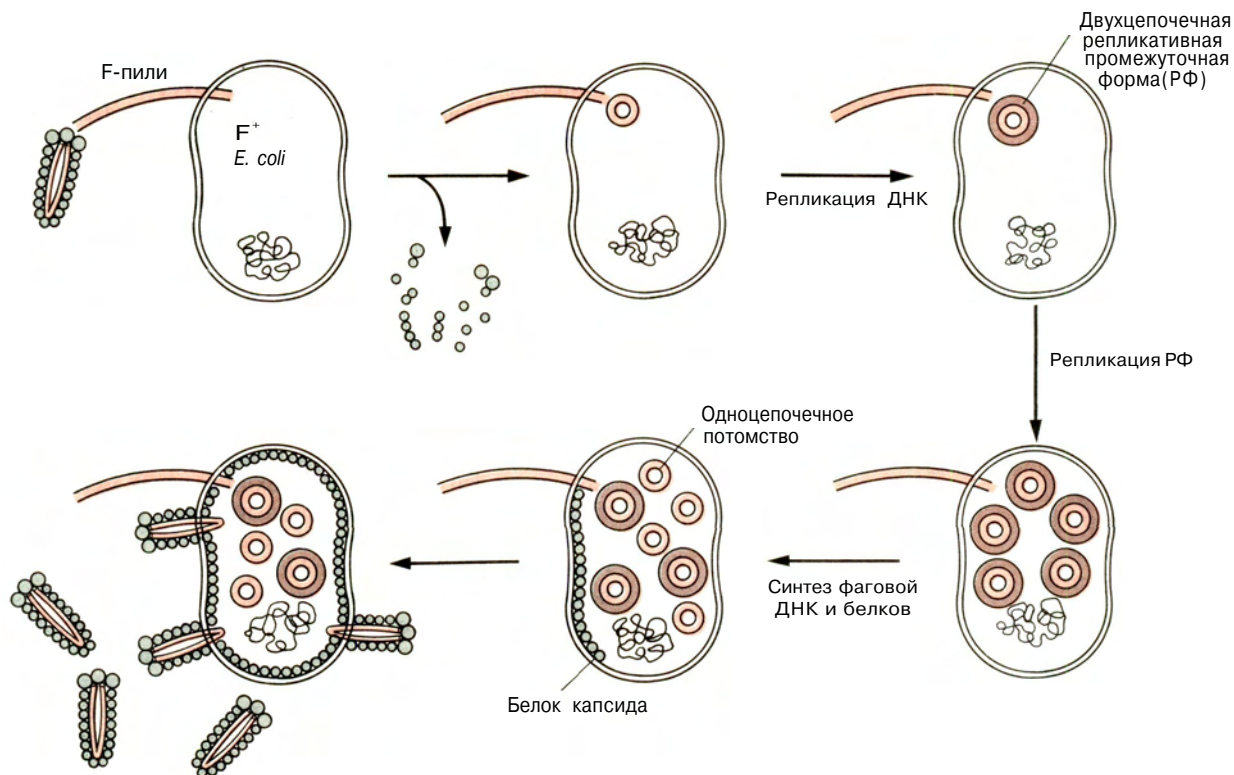


РИС. 5.15.

Жизненный цикл бактериофага M13. Обратите внимание, что инфицированная клетка и ее потомство непрерывно высвобождают новые фаговые частицы.

рекомбинантов, фаговый геном дикого типа модифицировали, встроив в некодирующую область длиной 507 п.н. часть *lac*-оперона *E. coli*, включающую оператор, промотор и 5'-кодирующую область β-галактозидазного гена. В эту короткую область в свою очередь был встроен сегмент из 42 п.н., который содержал несколько сайтов узнавания для уникальных рестриктирующих эндонуклеаз. Каждый из этих сайтов можно было использовать для включения фрагментов ДНК с соответствующими липкими концами. Такой сегмент получил название **полилинкера**. Его введение не нарушает ни кодирующую рамку, ни активность N-концевой части β-галактозидазы. В отсутствие вставки в полилинкере векторный геном детерминирует синтез N-концевой части β-галактозидазы, которая вместе с C-концевой частью β-галактозидазы, продуцируемой специальными штаммами *E. coli* (71-18 в табл. 5.1), образует активную β-галактозидазу (разд. 5.2.г). Бляшки, образуемые самим вектором при выращивании на агаре, содержащем Xgal, окрашиваются в голубой цвет. Однако после внедрения вставки в любой из рестрикционных сайтов по-

лилинкера синтез N-концевой части β-галактозидазы нарушается и образуются бесцветные бляшки. Таким образом, среди большого числа бляшек, образуемых вектором без вставки, легко обнаружить бляшки, содержащие рекомбинантный вирус. Были получены производные фага M13, содержащие другие сайты для рестриктирующих эндонуклеаз и позволяющие использовать другие приемы для отбора рекомбинантов. В некоторых случаях отбор

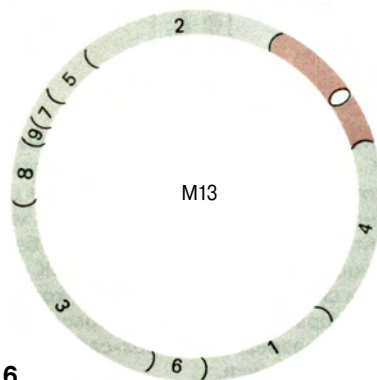
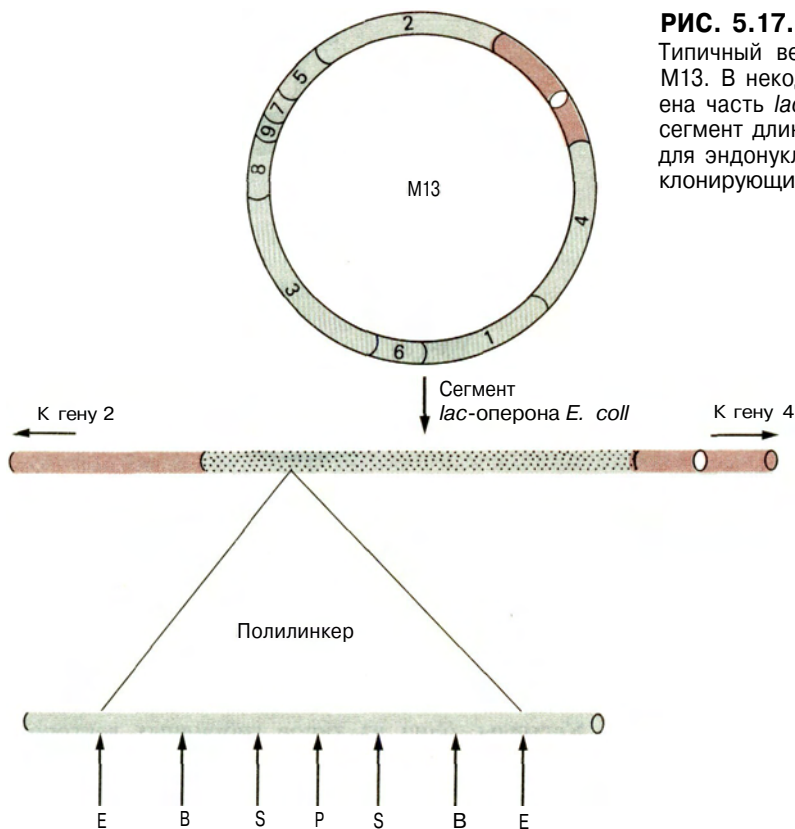


РИС. 5.16.

Карта генома фага M13. Цифрами отмечены разные гены (гены 2 и 5 ответственны за репликацию, остальные детерминируют образование капсида и сборку). Светлый овал - область начала репликации; некодирующий участок выделен цветом.

**РИС. 5.17.**

Типичный вектор, сконструированный на основе фага M13. В некодирующую область фагового генома встроена часть *lac*-оперона *E. coli*. Затем в ген *lacZ* встроено сегмент длиной 42 п.н., содержащий несколько сайтов для эндонуклеаз рестрикции (полилинкер, или мультиклонировующий сайт).

основывался на резистентности к антибиотикам, в других использовались гены, способные удовлетворять потребности соответствующих хозяев в определенных питательных веществах.

**Трансфекция.** Как и в случае с плазмидными векторами, обработка клеток  $\text{CaCl}_2$  (разд. 5.1.в) позволяет осуществлять трансфекцию чувствительных  $F^+$ -клеток рекомбинантными ДНК фага M13. После идентификации искомой бляшки и очистки получают клетки, из которых в процессе их роста и деления непрерывно высвобождаются фаговые частицы, содержащие рекомбинантную ДНК. Сконструированная рекомбинантная ДНК всегда является двухцепочечной молекулой, а ДНК фага M13 — одноцепочечной.

**Одноцепочечные рекомбинантные ДНК.** Системы, в которых в качестве вектора используется фаг M13, имеют несколько ценных свойств, характерных также для систем с плазмидными векторами (например, сохранение жизнеспособности клеток) и систем, где вектором служит фаг  $\lambda$  (рекомбинантные ДНК содержатся внутри фаговых частиц). Однако они обладают и некоторыми уникальными особенностями. Одна из них — это способность включать очень большие вставки, поскольку процесс упаковки ДНК не зависит от размера фагового

генома. Вторая особенность связана со способностью двухцепочечных рекомбинантных ДНК, сконструированных на основе репликативного интермедиата, превращаться во множество одноцепочечных молекул ДНК в фаговых частицах. Кроме того, система позволяет включать каждую из цепей ДНК встроеного фрагмента в разные рекомбинантные клоны. Векторные молекулы фага M13, точно так же как плазмидные и  $\lambda$ -векторы, включают вставки в любой из двух возможных ориентации (рис. 5.18). При трансфекции эти рекомбинанты попадают в разные клетки и образуют отдельные бляшки. Однако, поскольку только одна цепь двухцепочечной репликативной формы M13 используется в качестве матрицы для синтеза одноцепочечной фаговой ДНК, лишь одна из двух цепей репликативного интермедиата будет представлена во всех фаговых частицах данной бляшки (см. рис. 2.3). Таким образом, клонирование любой вставки всегда дает начало двум разным видам фага, каждый из которых содержит одну из цепей вставки. Это свойство очень удобно для проведения детального анализа нуклеотидных последовательностей и для синтеза специфических радиоактивно меченых ДНК-зондов (гл. 7).

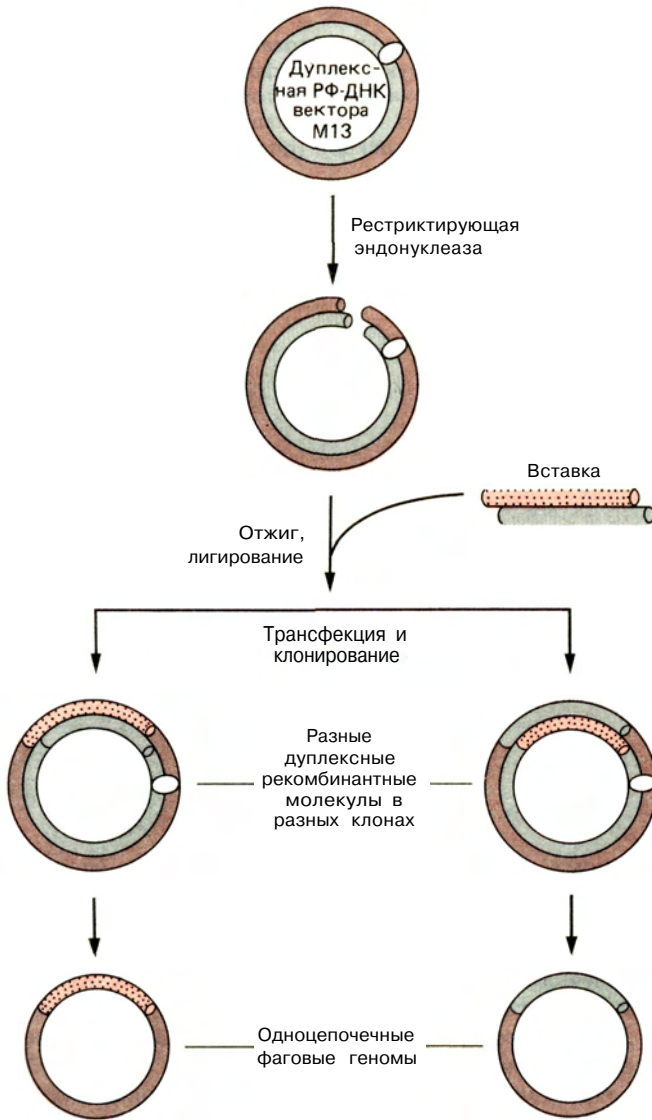


РИС. 5.18.

Векторы, сконструированные на основе M13, позволяют включать каждую из цепей встроенного фрагмента ДНК в разные клоны. Та цепь двухцепочечного вектора, которая идентична фаговому геному, выделена цветом. (Процесс репликации описан в разд. 2.1.ж.)

#### 5.4. E. COLI-СИСТЕМЫ: ПЛАЗМИДНО-ФАГОВЫЕ ВЕКТОРЫ

Поскольку плазмидные и фаговые векторы имеют свои преимущества, важные для молекулярного клонирования, были сконструированы молекулы ДНК, объединяющие свойства обоих этих векторов.

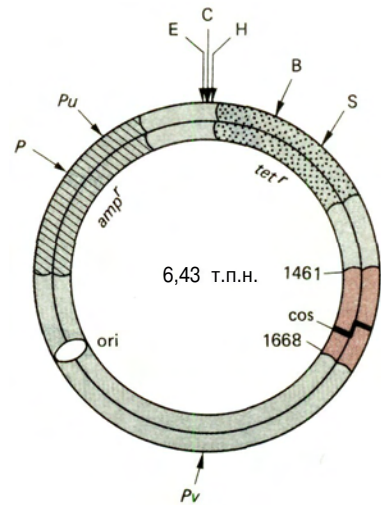


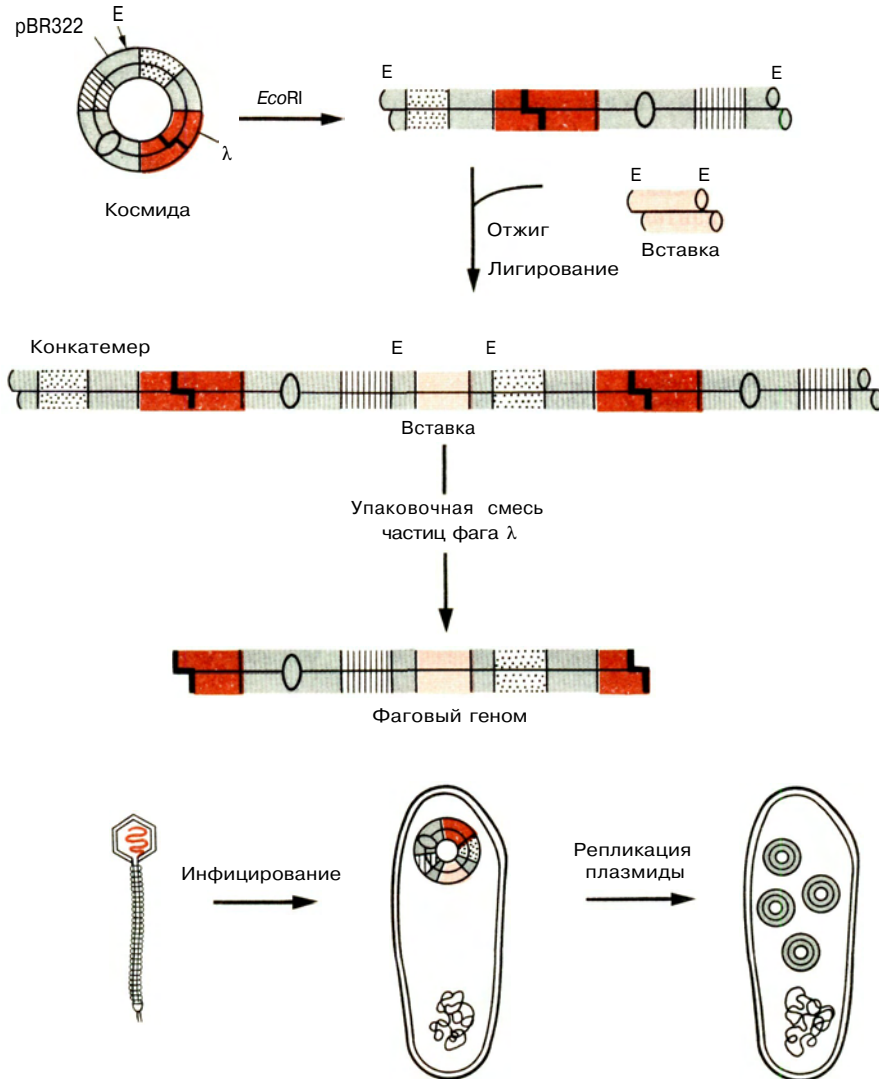
РИС. 5.19.

Типичная космида. pBR322-сегменты выделены серым цветом. Вместо фрагмента с 1461-го нуклеотида по 1688 (см.рис. 5.6) встроен участок генома фага  $\lambda$ .

#### а. Космиды

**Космиды**—один из типов гибридных векторов, которые реплицируются, используя плазмидный тип репликации, и обладают способностью упаковываться *in vitro* в оболочки частиц фага  $\lambda$ . Типичная космида (рис. 5.19) обладает способностью к репликации, содержит уникальные эндонуклеазные сайты и селективные маркеры, принадлежащие плазмидной ДНК, которая объединена с сегментом ДНК фага  $\lambda$ , содержащим соединенные липкие концы (*cos*-сайт). Космиды конструируют с помощью методов, применяемых при работе с рекомбинантными ДНК. При расщеплении конкатемерной ДНК фага  $\lambda$  с помощью соответствующих эндонуклеаз (рис. 5.10) образуются *cos*-сегменты, которые могут быть включены в стандартный плазмидный вектор. Для того чтобы могло произойти соединение по *cos*-сайту, достаточно фрагмента ДНК фага  $\lambda$  длиной всего лишь 250 п.н., включающего последовательности, необходимые для связывания терминазы и последующего разрезания. Важной чертой большинства космидных векторов для клонирования является их способность включать вставки ДНК размером до 45 т.п.н.

Если кольцевую космидную ДНК разрезать по какому-то уникальному сайту, смешать с фрагментами ДНК, содержащими соответствующие липкие концы, и провести отжиг, то образуются длинные конкатемеры (рис. 5.20). При смешивании этих конкатемеров с белками, осуществляющими упаковку ДНК фага  $\lambda$  (разд. 5.3.г), они разрезаются по *cos*-сайтам и ДНК упаковывается обычным путем.

**РИС. 5.20.**

Молекулярное клонирование с помощью космидного вектора (см. рис. 5.13 и 5.14, на которых показан процесс сборки фага  $\lambda$ ).

Этот процесс позволяет отобрать вставки большой протяженности, поскольку для того, чтобы ДНК могла упаковаться в головки фага  $\lambda$ , расстояние между *cos*-сайтами должно составлять от 38 до 52 т.п.н. Как и в случае с  $\lambda$ -векторами, смесь конкатемеров является сложной и часто содержит векторные фрагменты без всяких вставок или со множеством повторяющихся вставок.

Реципиентные клетки-хозяева приобретают упакованные космиды в результате инфицирования «фальшивыми» фаговыми частицами, причем этот процесс значительно более эффективен, чем транс-

фекция плазмидной ДНК. На 1 мкг ДНК со вставками можно получить от  $1 \cdot 10^4$  до  $5 \cdot 10^5$  колоний трансформированных клеток. Попав в клетку-хозяина, рекомбинантная ДНК амплифицируется и сохраняется в виде плазмиды (рис. 5.20).

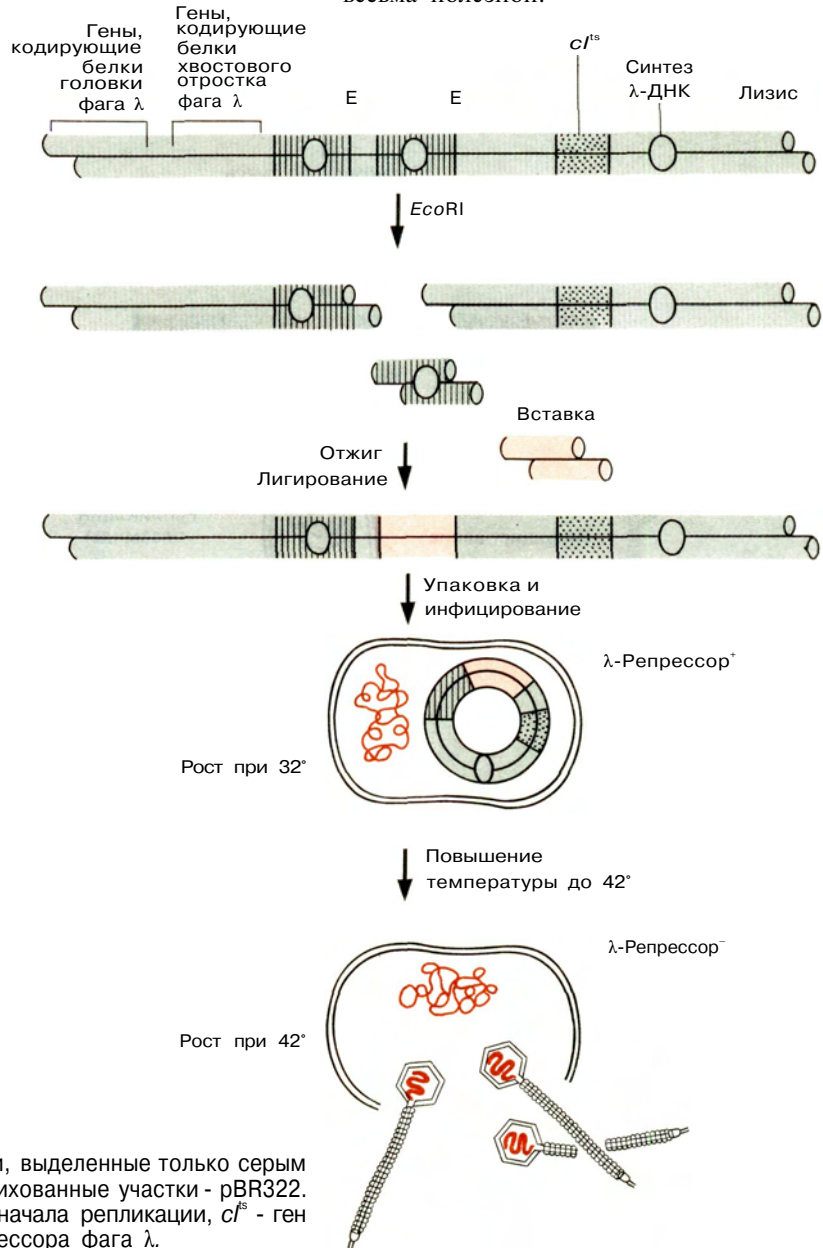
## б. Фазмиды

Космиды по существу являются плазмидными векторами, «присвоившими» *cos*-сайт фага  $\lambda$  для осуществления эффективной упаковки *in vitro*. Истинными гибридами между плазмидой и фагом являются **фазмиды** — линейные дуплексные молекулы ДНК, концы которых представляют собой сегменты ДНК фага  $\lambda$ , содержащие все гены, необходимые для литической инфекции, а средняя часть — линеаризованную плазмиду. Функции реп-



ликация как фага, так и плазмиды в фазмидах полностью сохранены. Обычно вектор содержит несколько тандемных повторов плазмидного сегмента, обеспечивающих необходимую для успешной упаковки ДНК протяженность генома. Один или несколько повторов в процессе конструирования рекомбинантных молекул заменяются вставкой (рис. 5.21). Подобно рекомбинантам фага  $\lambda$  и космид, фазмидные ДНК упаковываются *in vitro* перед инфекцией. Попадая в клетку *E. coli*, фазмида репли-

цируется как фаговая ДНК, в результате чего на газоне чувствительных бактерий образуются бляшки. Однако если вектор содержит ген, кодирующий репрессор фага  $\lambda$ , то фазмида реплицируется как плазмида, а не как фаг. Кроме того, если ген репрессора кодирует мутантный белок  $cI^{ts}$ , инактивирующийся при повышенной температуре (**температурочувствительный репрессор**), то фазмида реплицируется, как плазмида при низкой температуре и как фаг—при повышении температуры на несколько градусов. Такая маневренность может оказаться весьма полезной.



**РИС. 5.21.**

Типичная плазмида. Участки, выделенные только серым цветом, - ДНК фага  $\lambda$ , заштрихованные участки - pBR322. Овалом отмечены области начала репликации,  $cI^{ts}$  - ген термочувствительного репрессора фага  $\lambda$ .

## 5.5. ДРУГИЕ ПРОКАРИОТИЧЕСКИЕ СИСТЕМЫ ХОЗЯИН-ВЕКТОР

Системы с другими хозяином и вектором едва ли когда-нибудь смогут конкурировать с *E. coli*-системами в том, что касается выделения и очистки сегментов ДНК. Тем не менее возможность эффективного клонирования в других видах бактерий очень важна для изучения генетических свойств и экспрессии генов у организмов, представляющих научный, медицинский или промышленный интерес. Поэтому, несмотря на малое число хорошо охарактеризованных плазмид и бактериофагов и экспериментальные трудности при введении молекул ДНК в некоторые клетки, были разработаны полезные системы, способные служить векторами. Если отсутствовали подходящие для данного вида природные плазмиды или фаги, то иногда использовали плазмиды и фаги, полученные от других бактерий. Многие плазмиды и бактериофаги весьма требовательны и реплицируются только в одном или нескольких родственных видах клеток, но некоторые из них на удивление неприхотливы и репродуцируются в широком круге хозяев, принадлежащих к разным видам.

### а. Грамотрицательные организмы

R-плазида, изображенная на рис. 5.1, представляет собой пример плазмиды, способной размножаться во многих грамотрицательных бактериях. На основе производных таких плазмид можно сконструировать рекомбинантные ДНК, клонировать и амплифицировать их в клетках *E. coli*, а затем при необходимости перенести в другие бактерии.

Одна из таких систем клонирования была разработана на основе *Hemophilus influenzae* и встречающейся в природе неконъюгативной плазмиды *Hemophilus*. Плазида имеет протяженность около 6 т.п.н., содержит ген резистентности к ампициллину и образует множество копий в клетках хозяина. Поскольку *H. influenzae* природно компетентна в отношении трансформации, никаких специальных ухищрений при трансфекции не требуется. Полезные векторы получены также для азотфиксирующих бактерий *Klebsiella pneumoniae*. Они сконструированы на основе мелкого бактериофага P4, который способен размножаться как в клетках *E. coli*, так и в *K. pneumoniae* и в зависимости от условий реплицироваться или как фаг, или как плазида.

### б. Грамположительные организмы

*Bacillus subtilis*. Векторы, с помощью которых можно производить генетические манипуляции с

*Bacillus subtilis*, имеют особенно большое значение, поскольку этот организм широко используется в различных процессах ферментации, в частности при промышленном получении антибиотиков и ферментов. Кроме того, *B. subtilis* секретирует синтезируемые белки в среду, поэтому его можно использовать для выделения продуктов генов, кодируемых рекомбинантными векторами. Были выделены бактериофаги и плазмиды, несущие удобные селективные маркеры, например гены резистентности к антибиотикам, и на их основе сконструированы векторы. Одна из используемых плазмид получена из другого грамположительного организма, *Staphylococcus aureus*. Ее отличает то, что она способна реплицироваться и сохраняться в клетках как *Bacillus subtilis*, так и других видов *Bacillus* (рис. 5.22). Вектор, созданный на ее основе, имеет размер около 5 т.п.н. и содержит два селективных маркера. Один из них придает клеткам резистентность к канамицину, другой — ген *B. subtilis* — необходим для биосинтеза триптофана. При репликации векторной ДНК в клетках *B. subtilis* образуется до 50 ее копий на клетку. В составе вектора имеется несколько сайтов для разных эндонуклеаз рестрикции, которые используют для получения рекомбинантов. Встраивание фрагмента ДНК в единственный имеющийся *Bgl II*-сайт приводит к инактивации гена резистентности к канамицину, а в уникальный сайт *Hind III* — к инактивации триптофанового гена, в результате чего проявляется второй фенотипический маркер хозяйских клеток, которые сами утратили соответствующий активный ген. Как и в случае систем *E. coli*, плазмидные векторы с двумя маркерами позволяют легко идентифицировать и выделять рекомбинантные молекулы.

В отличие от *E. coli* виды *Bacillus* трансформируются даже без применения специальных методических приемов. Однако эффективность трансфекции заметно увеличивается, если вместо интактных клеток используют **протопласты**, которые получают в результате обработки клеток лизоцимом; при этом последние утрачивают клеточную стенку. После инкубации протопластов с трансформирующей ДНК в специальной питательной среде клеточная стенка регенерируется. Эффективность трансфекции варьирует от 100 до 10<sup>6</sup> трансформантов на 1 мкг векторной ДНК в зависимости от системы хозяин-вектор и метода трансфекции.

*Streptomyces*. Представители грамположительных организмов рода *Streptomyces* кодируют большинство известных применяемых в медицине антибиотиков. Поэтому получение удобных систем хозяин-вектор для *Streptomyces* очень важно как для изучения биосинтеза антибиотиков, так и для создания новых антибиотиков и повышения эффективности продуцентов. Наиболее разработанные в на-

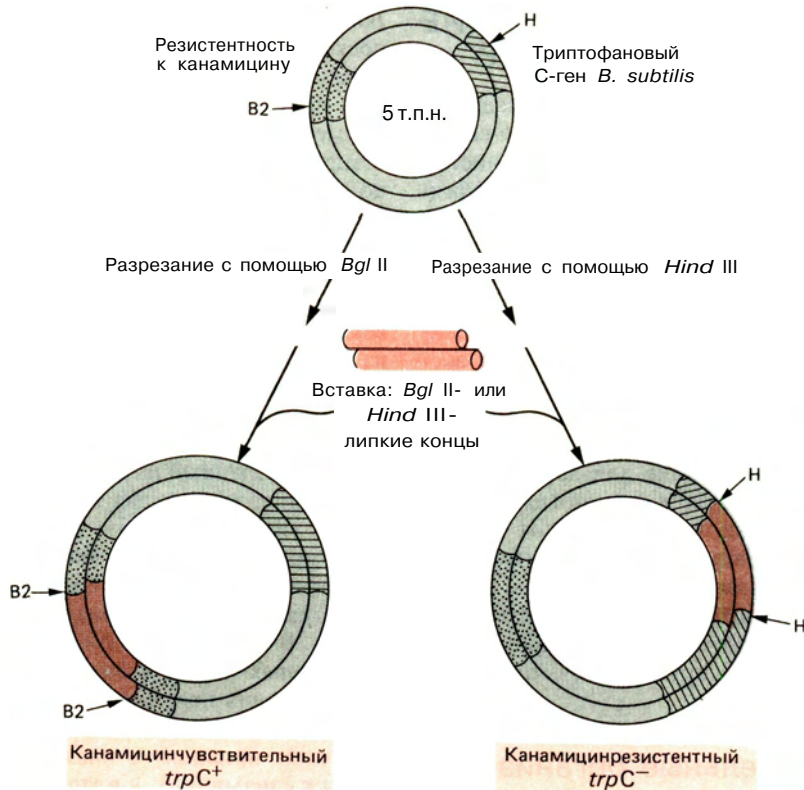


РИС. 5.22.

Система «хозяин-вектор» для *Bacillus*.

стоящее время векторные системы—это плазмиды, среди которых имеются как плазмиды с низким числом копий на клетку, так и плазмиды с высоким числом копий. При трансфекции и отборе необходимых рекомбинантов используют особые свойства стрептомицетов. Некоторые из плазмидных векторов *Streptomyces* являются конъюгативными и могут передаваться от одного вида к другому при конъюгации.

Одним из типичных векторов является плазида, представленная на рис. 5.23. Сегмент длиной примерно 11 т.п.н. получен от плазмиды, выделенной из *S. lividans*. Другая часть плазмиды содержит ген устойчивости к антибиотику метиленомицину А, полученный из хромосомы штамма *S. coelicolor*. Второй селективный маркер связан со способностью клеток, содержащих плазмиду, подавлять рост соседних клеток *Streptomyces*, утративших плазмиду. При посеве трансформированных клеток на газоне бесплазмидных клеток трансформанты выделяют из центра областей с подавленным ростом. Эта и подобные ей векторные молекулы можно трансфицировать в протопласты нескольких видов *Streptomyces*. После переноса протопластов в стандартную

среду происходит регенерация нормальных клеток. На 1 мкг плазмидной ДНК можно получить около  $10^5$  трансформантов.

## в. Челночные векторы

Плазмиды и бактериофаги эволюционируют синхронно с их природными хозяевами, поэтому

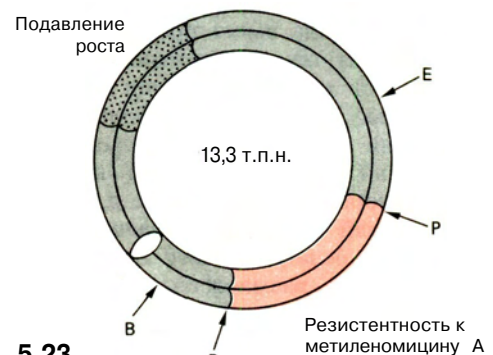


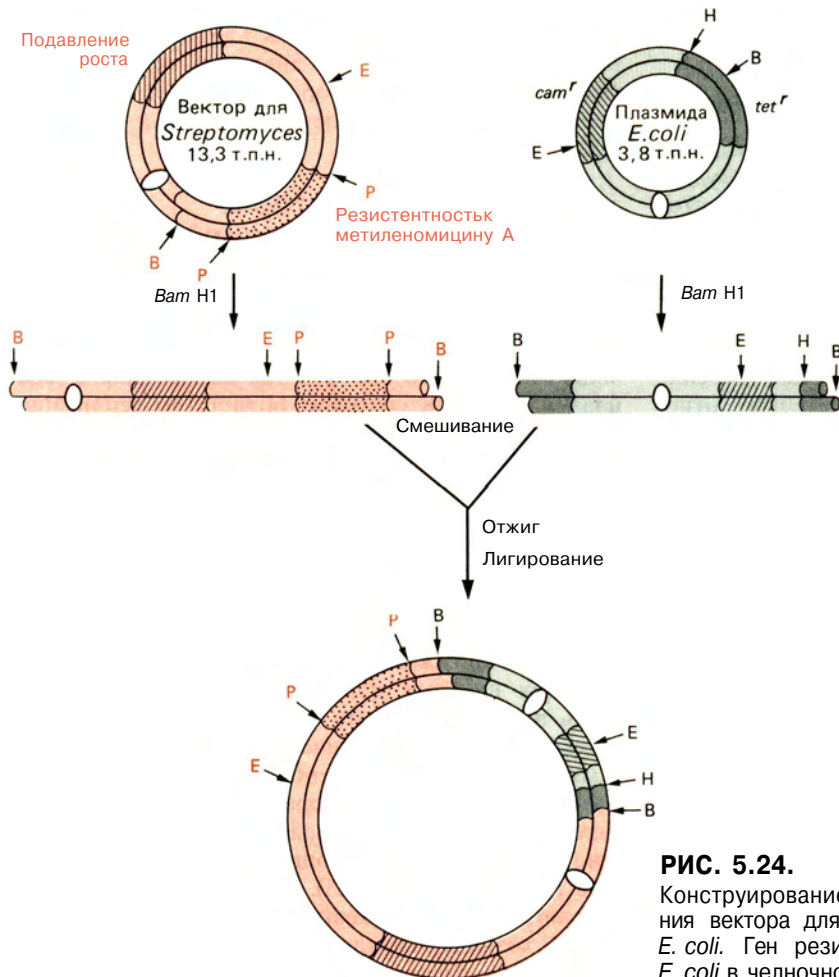
РИС. 5.23.

Вектор для *Streptomyces*. Большую часть векторной ДНК составляет плазида, выделенная из *S. lividans*; сюда входит область, ответственная за подавление роста соседних клеток *Streptomyces*, утративших плазмиду. Остальная часть векторной ДНК получена из *S. coelicolor* и содержит ген резистентности к метиленомицину А.

они реплицируются обычно только в клетках одного вида бактерий или небольшой группы видов. Векторы, способные реплицироваться в разных хозяевах, имеют определенные преимущества. Например, клонирование и выделение сегментов ДНК для структурного анализа наиболее удобно проводить, используя системы, где в качестве хозяина выступает *E. coli*. Однако, как правило, функциональный анализ клонированной вставки следует проводить в клетках того вида, откуда получена вставка. Нужный сегмент ДНК всегда можно клонировать в клетках *E. coli*, выделить его и затем переклонировать в векторе, совместимом с интересующими нас клетками-хозяевами. Однако значительно проще сразу провести клонирование с использованием вектора, который способен без дополнительных манипуляций существовать как в *E. coli*, так и альтернативном хозяине и реплицироваться в них. Для таких целей предназначены **челночные векторы**, которые содержат две области *ori*, соответствующие каждому из двух видов хозяев, а также гены, необ-

ходимые для репликации и не поставляемые хозяйскими клетками. Такие векторы конструируют с помощью методов, применяемых при получении рекомбинантных ДНК, и число их уже довольно велико. Одни из них способны существовать попеременно в клетках двух разных видов прокариот, другие — в двух прокариотических (обычно в *E. coli*) и эукариотических клетках (в том числе дрожжевых, растительных и животных). Большинство эукариотических векторов, описанных в последующих разделах, являются челночными.

Конструирование челночного вектора связано с решением специфических проблем, поскольку одна его часть всегда является «чужой», а другая — «родной» в альтернативных хозяйских клетках. «Чужая» часть не должна мешать репликации, направляемой «родной» частью вектора. Не следует конструировать векторы с таким сочетанием последовательностей, при котором он становится нестабильным в одном из двух хозяев, как это происходит при наличии последовательностей, мешающих реплика-



**РИС. 5.24.**

Конструирование челночного вектора путем объединения вектора для *Streptomyces* (рис. 5.23) и плазмиды *E. coli*. Ген резистентности к тетрациклину плазмиды *E. coli* в челночном векторе инактивирован.

ции или транскрипции в альтернативных хозяевах. На эффективность челночного вектора может влиять и присутствие активной системы рестрикции-модификации в одной или другой (или в обоих) из альтернативных хозяйских клеток. Эту проблему можно решить, используя мутантные клетки, утратившие активную эндонуклеазу рестрикции.

Один из челночных векторов, способных реплицироваться в клетках *Streptomyces* и *E. coli*, представлен на рис. 5.24. От плазмиды *Streptomyces* (рис. 5.23) вектор получил сигнальные последовательности, необходимые для репликации в клетках *Streptomyces*, и гены резистентности к метиленоминцину А, а от плазмиды *E. coli*—гены, ответственные за репликацию, и маркер резистентности к антибиотикам для сохранения в клетках *E. coli*. С помощью такой плазмиды можно клонировать сегмент ДНК *Streptomyces* в *E. coli*, а затем, размножив его, проводить функциональное тестирование в штаммах *Streptomyces*.

## 5.6. ЭУКАРИОТИЧЕСКИЕ СИСТЕМЫ ХОЗЯИН-ВЕКТОР: ДРОЖЖИ

### а. Универсальность и удобство

Лабораторные штаммы обычных пекарских дрожжей, *Saccharomyces cerevisiae*, являются удобными хозяевами для проведения экспериментов с рекомбинантными ДНК. Хотя дрожжи способны к половому размножению (см. гл. 10), обычно они размножаются путем почкования. В суспензии дрожжи растут в виде отдельных клеток, а на твердом субстрате образуют колонии подобно *E. coli*. В обычной питательной среде время удвоения клеток составляет 1,5–2,5 ч. Следовательно, относительно быстро и без особых затрат можно получить большое количество этих непатогенных организмов.

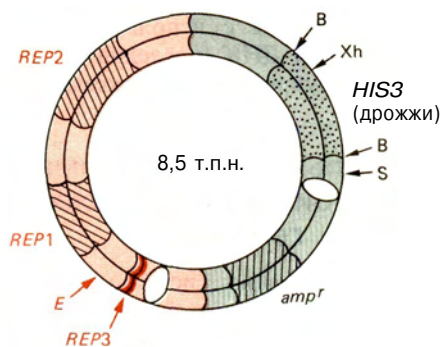
Гаплоидный геном *Saccharomyces cerevisiae* содержит  $1,4 \cdot 10^7$  п.н., что примерно в три раза больше, чем геном *E. coli*. Все эти пары оснований распределены по 17 (гаплоидным) хромосомам. Создана большая коллекция мутантов, дефектных в отношении метаболизма, биосинтеза или клеточного цикла, и картированы мутации, ответственные за эти дефекты. Вирусы, которые инфицировали бы клетки дрожжей, не обнаружены, но на основе единственной известной дрожжевой плазмиды сконструировано несколько векторов, используемых в настоящее время для введения рекомбинантных молекул ДНК в дрожжевые клетки. Имеются также векторы, которые содержат сегменты дрожжевого генома, обеспечивающие независимую репликацию векторной ДНК, а при наличии клонированной центромерной ДНК векторная ДНК даже ведет себя

подобно хромосомам при мейозе и митозе. Кроме того, клонированные фрагменты ДНК способны рекомбинировать с гомологичными сегментами дрожжевого генома, что приводит к стабильной **сайт-специфической трансформации** клеточного генома независимо от сохранения вектора. Эти методы вместе с классическими методами генетики дрожжей позволяют изучить данный эукариотический организм с такой же полнотой, какая ранее была доступна только для прокариот типа *E. coli*.

Жесткая полисахаридная клеточная стенка, окружающая мембраны дрожжевых клеток, создает барьер для молекул ДНК. Для того чтобы изолированная ДНК могла проникнуть в дрожжевую клетку, эту стенку необходимо удалить, но так, чтобы клетка осталась жизнеспособной. Для этого используют ферменты, разрушающие полисахариды. После обработки клеток соответствующими ферментами образуются **сферопласты**, способные поглощать ДНК после обработки их  $\text{CaCl}_2$ . Если затем эти сферопласты перенести в специальную питательную среду, то клеточная стенка восстановится. Проницаемые для ДНК клетки можно также получить, обработав их ацетатом лития. С помощью некоторых векторов получают около  $10^4$  трансформированных клеток *S. cerevisiae* на 1 мкг ДНК.

### б. Векторы, способные реплицироваться в клетках дрожжей

**Плазмидные векторы.** Некоторые штаммы *S. cerevisiae* содержат кольцевую двухцепочечную плазмиду длиной 6318 п.н. Эта так называемая **2 мкм-плазмидная ДНК** содержит область *ori* репликации ДНК, другую *цис*-действующую область (*REP3*) и два гена (*REP1* и *REP2*), которые вместе обеспечивают стабильное сохранение большого числа копий плазмиды—около 50 на дрожжевую клетку. Примерно 50% 2 мкм-плазмидной ДНК необходимы для ее репликации и сохранения в клетках, остальные 50% несут существенны. 2 мкм-ДНК не содержит маркерных генов, которые позволяли бы отбирать дрожжевые клетки, несущие плазмиду. Поэтому ни сама плазида, ни ее существенная часть не могут служить удобными векторами для дрожжевых клеток. Однако существенная часть 2 мкм-плазмиды была использована для конструирования челночного вектора путем объединения ее с плазмидой pBR322, которая содержала дрожжевой ген (*HIS3*), кодирующий один из ферментов биосинтеза гистидина (рис. 5.25). Используя такой вектор в сочетании с дрожжевым штаммом *his3<sup>-</sup>* в качестве хозяина, можно проводить отбор клеток, растущих в отсутствие гистидина. Оказалось, что



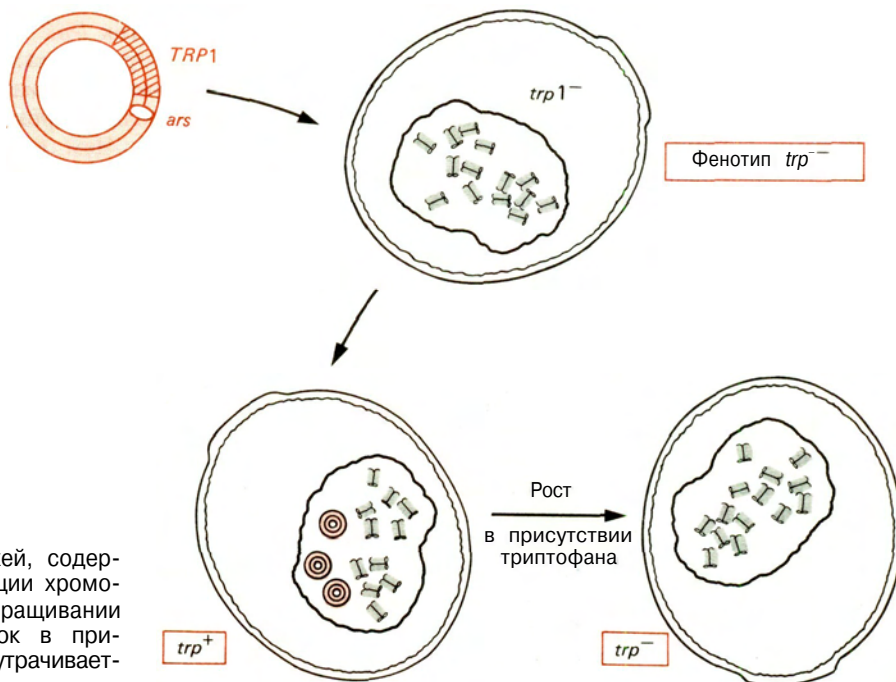
**РИС. 5.25.**

Схематическое изображение челночного вектора для системы дрожжи-*E. coli*, сконструированного на основе дрожжевой 2мкм-плазмиды (выделена цветом). Часть вектора, произошедшая от *E. coli*, представляет собой плазмиду рВВ322. Области начала репликации обозначены светлым овалом; область *HIS3*, полученная из ДНК хромосом дрожжевых клеток, выделена точками.

дрожжевой ген *HIS3* функционален также в клетках *E. coli* благодаря тому, что он содержит область, гомологичную промотору *E. coli*. Таким образом, дрожжевой ген *HIS3* может служить селективным маркером в бактериальных клетках, утративших аналогичный ген. Вектор содержит несколько уникальных рестриктазных сайтов, по которым можно встраивать определенные сегменты ДНК.

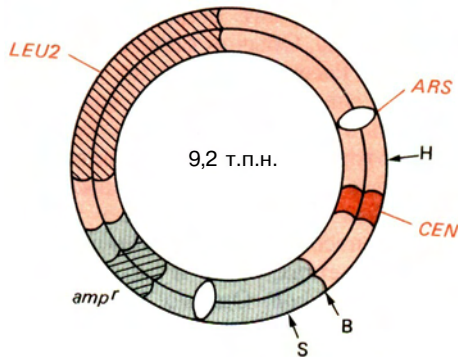
**Векторы ARS.** Дрожжевые векторы второго типа представляют собой кольцевую двухцепочечную

ДНК, которая содержит область *ori* репликации дрожжевого генома. Обычно такие векторы являются челночными. Рассмотрим вектор, изображенный на рис. 5.25. Если 2 мкм-сегмент удалить, то оставшаяся часть будет представлять собой плазмиду рВВ322, у которой в области, ответственной за резистентность к тетрациклину, содержится вставка в виде сегмента *HIS3* дрожжевого генома. Эта структура не способна к автономной репликации в дрожжевых клетках, хотя и может реплицироваться в клетках *E. coli*. Однако при клонировании в рВВ322 некоторых других генов дрожжевой хромосомы (не *HIS3*) образующиеся рекомбинантные молекулы могут реплицироваться автономно в клетках как *E. coli*, так и дрожжей! Среди генов, обеспечивающих способность к репликации, были *TRP1* и *ARG4*, которые кодируют ферменты, необходимые для биосинтеза триптофана и аргинина соответственно. Оказалось, что область *ori* хромосомы расположена достаточно близко от каждого из этих генов, чтобы включиться в клонируемые дрожжевые сегменты (рис. 5.26). Размер областей начала репликации *ori* не превышает 100 п.н. Такие области получили название **автономно реплицирующихся последовательностей** (или **ARS**). Все **ARS**-последовательности содержат одинаковый олигонуклеотид  $5' \text{---} \text{A}_7 \text{---} \text{TTTAT}_6 \text{---} \text{TTT}_7 \text{---} 3'$ , необходимый для функционирования. Этот олигонуклеотид длиной 11 п.н. окружен другими важными в функциональном отношении последовательностями, но они варьируют от одного **ARS**-элемента к другому. Любая



**РИС. 5.26.**

Плазмидный вектор для дрожжей, содержащий область начала репликации хромосомного происхождения. При выращивании *trp*<sup>+</sup>-трансформированных клеток в присутствии триптофана плазида утрачивается.



**РИС. 5.27.**

Мини-хромосомный челночный вектор для дрожжей. Сегменты, произошедшие из генома дрожжевых клеток, выделены цветом, pBR322-сегменты - серые.

кольцевая двухцепочечная ДНК, которая содержит *ARS*-последовательность, может служить дрожжевым вектором и использоваться для непосредственного клонирования в дрожжевых клетках. Такие молекулы способны стабильно существовать в клетках *S. cerevisiae* до тех пор, пока они являются единственным носителем какого-то необходимого гена (например, гена *TRP1* в клетках *trp1<sup>-</sup>*, выращиваемых в отсутствие триптофана), однако при снятии селективного давления они быстро утрачиваются.

**Мини-хромосомные векторы.** Дрожжевые векторы третьего типа ведут себя аналогично функциональной хромосоме, хотя и имеют меньшие размеры. Такие мини-хромосомные векторы реплицируются только один раз в каждом поколении клеток, и две образующиеся копии сегрегируют, как истинные хромосомы, распределяясь по дочерним клеткам во время мейоза и митоза. Необходимыми элементами функциональной мини-хромосомы являются три небольших сегмента хромосомной ДНК дрожжей (рис. 5.27). Один из них, *ARS*, содержит область *ori* репликации ДНК, второй создает условия для отбора. На рис. 5.27 представлена плаزمид, содержащая ген *LEU2*, которую можно применять в сочетании с хозяйскими клетками *leu2<sup>-</sup>*. Третий сегмент, *CEN*, регулирует число копий внутри клеток и отвечает за способность к сегрегации в течение многих поколений. Сегменты ДНК, которые придают мини-хромосоме такое замечательное свойство, — это центромеры, и происходят они из центромерных областей дрожжевых хромосом (отсюда обозначение *CEN*). Их размеры не должны превышать 500 п.н. Связь между структурой и функцией сегментов *CEN* описана в гл. 9; там же рассмотрены и молекулярные свойства дрожжевых теломер. В гл. 9 описано также, как с помощью клонирования этих двух важных хромосомных доменов удалось реконструировать функциональные

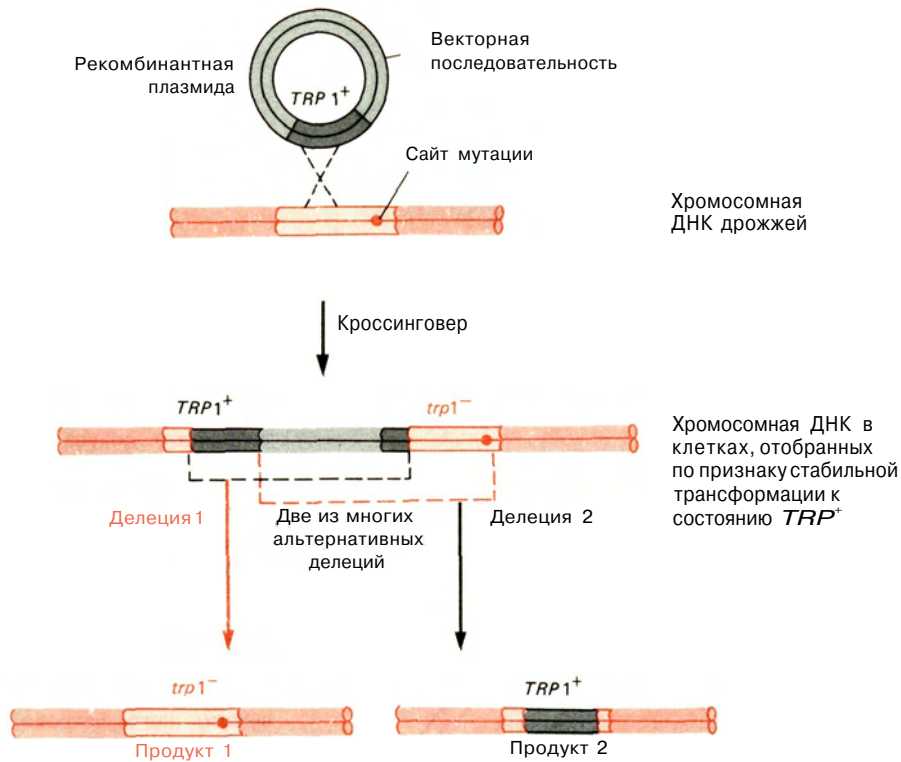
дрожжевые хромосомы с использованием линейных, а не кольцевых ДНК (последние применялись прежде при конструировании дрожжевых векторов).

## в. Стабильная трансформация при рекомбинации с дрожжевым геномом

Эта глава посвящена системам, с помощью которых рекомбинантные векторные молекулы могут независимо реплицироваться в соответствующих клетках. Однако большое значение имеет также способность рекомбинантных молекул осуществлять **стабильную трансформацию** клеточных геномов путем встраивания их в хромосомную ДНК. В этом случае новые генотип и фенотип очень стабильны, не зависят от давления отбора или сохранения внехромосомной рекомбинантной ДНК в будущих поколениях. Встроенная ДНК реплицируется одновременно с клеточным геномом. Такая стабильная трансформация происходит относительно редко и зависит от природы как донорной ДНК, так и реципиентных клеток.

**Стабильная трансформация рекомбинантными плазмидами.** Дрожжевые клетки можно стабильно трансформировать, хотя и с низкой частотой, с помощью рекомбинантных плазмид, которые не способны реплицироваться в дрожжах из-за отсутствия *ori* репликации. 1 мкг производной плазмиды pBR322, несущей дрожжевой ген *LEU2*, может трансформировать примерно 1 из  $10^7$  дрожжевых клеток *leu2<sup>-</sup>* в *LEU2<sup>+</sup>*. Если рекомбинантный вектор содержит, кроме того, элемент *ARS* и может реплицироваться автономно, то стабильная трансформация происходит с еще большей эффективностью. Чем выше число копий трансформирующей ДНК, образующихся в клетке, тем больше вероятность их встраивания в дрожжевой геном.

Встраивание происходит в результате рекомбинации между донорной рекомбинантной молекулой и реципиентным дрожжевым геномом. Рекомбинация может быть неспецифической, когда донорная ДНК встраивается в случайные хромосомные локусы, однако такие события довольно редки. Чаще встраивание происходит в специфическом хромосомном локусе, гомологичном последовательностям дрожжевой ДНК, включенной в рекомбинантную донорную молекулу. Например, плазида, несущая дрожжевой ген *TRP1* дикого типа и реплицирующаяся в дрожжевых клетках *trp1<sup>-</sup>*, как правило, рекомбинирует с гомологичным мутантным локусом *trp1* в дрожжевом геноме. В результате происходит стабильная трансформация клетки к независимому от триптофана состоянию. Встречаются два разных типа гомологичной (и соответственно



**РИС. 5.28.**

Реципрокная гомологичная рекомбинация между реплицирующимся плазмидным вектором и хромосомой дрожжей приводит к стабильной трансформации дрожжевых клеток. В приведенном примере вектор несет ген *TRP1* дикого типа (а также гены, необходимые для репликации). Гомологичная рекомбинация в мутантном локусе *trp1* дрожжевой хромосомы ведет к появлению в хромосомной ДНК двух копий триптофанового гена, функционирующей и нефункционирующей. Образующие-

ся клетки стабильно сохраняют фенотип *TRP1*<sup>+</sup> даже в отсутствие давления отбора (т.е. в присутствии триптофана). В последующих поколениях в результате гомологичной рекомбинации между двумя копиями (это довольно редкое событие) может произойти делеция векторных последовательностей вместе с одной из копий *trp*-гена, и тогда образуются либо *TRP1*<sup>-</sup>, либо *trp1*<sup>-</sup>-клетки.

сайт-специфической) рекомбинации. При кроссинговере между кольцевыми донорной и реципиентной молекулами происходит встраивание всего вектора в соответствующий хромосомный локус (рис. 5.28). Гомологичная рекомбинация между двумя копиями *trp*-гена может привести в последующих поколениях к восстановлению нормальной структуры ДНК и появлению клеток либо дикого, либо мутантного типа. При рекомбинации другого типа происходит конверсия гена — нереципрокная рекомбинация, при которой ген дикого типа в векторе используется для репарации мутантного сегмента в геноме (разд. 2.4.6).

*Сайт-специфическая трансформация с помощью линейных донорных молекул.* Кольцевые плазмиды, не способные реплицироваться в дрожжевых клетках, являются очень неэффективными донорами для

сайт-специфической трансформации. Однако трансформирующую активность таких молекул можно значительно повысить, если перед трансфекцией лиinearизовать их и если эти линейные двухцепочечные молекулы будут содержать на обоих концах последовательности, гомологичные последовательностям генома мишени (рис. 5.29). Линейную трансформирующую ДНК обычно получают путем эндонуклеазного расщепления рекомбинантной плазмиды, полученной после репликации в *E. coli*. Между двумя гомологичными концами может находиться по существу любая последовательность. Так, в известный геномный локус можно ввести специфические гены или нужные гены можно изменить или делетировать. Таким образом, в настоящее время генетику дрожжей можно изучать независимо от того, удастся или нет получить и отобрать специфических мутантов.



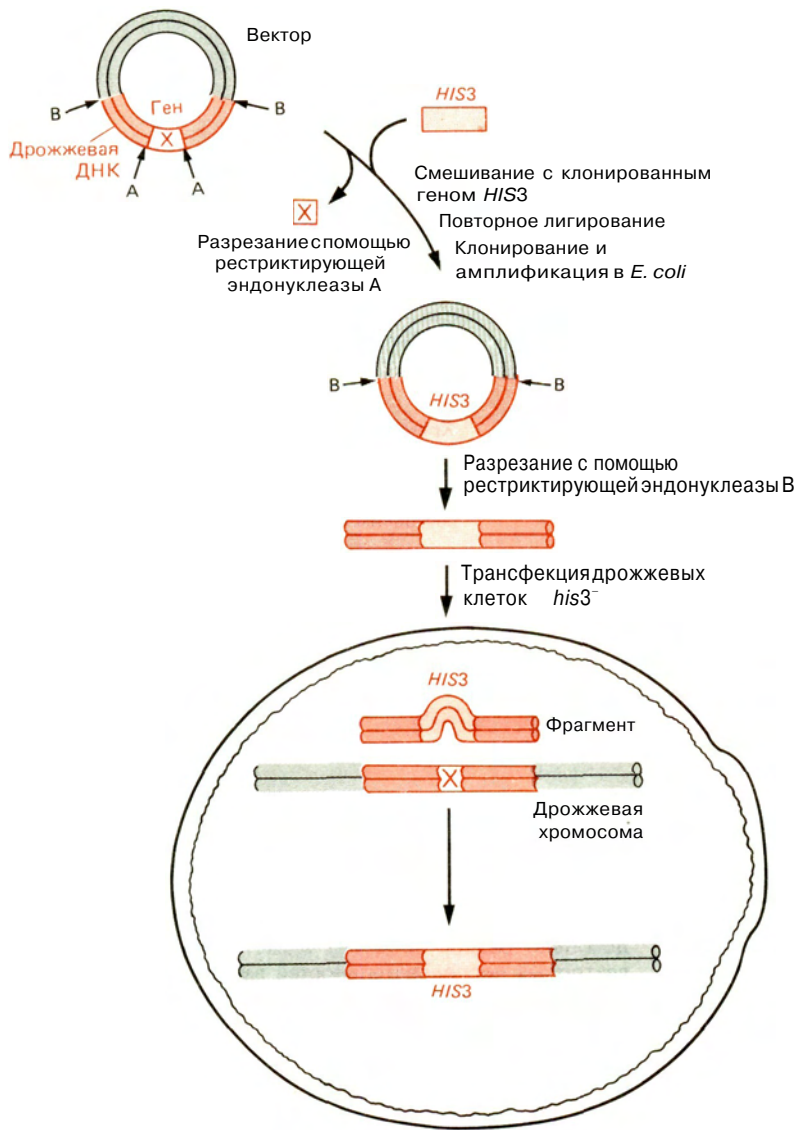


РИС. 5.29.

Сайт-специфическая трансформация геномного локуса дрожжей с помощью линейного фрагмента ДНК. Сегмент дрожжевой геномной ДНК, происходящий из того участка, который хотят трансформировать, клонируют в плазмидном векторе *E. coli*. Затем этот сегмент модифицируют с помощью инсерции, делеции или какой-либо мутации и клонируют и амплифицируют в клетках *E. coli*. В данном случае делетирован ген *X* и вместо него встроено селективный маркер *HIS3*. Плазмиду разрезают с помощью рестриктирующей эндонуклеазы в сайтах В и получают фрагмент дрожжевой ДНК в линейной форме, содержащий участки, гомологичные последовательностям генома-мишени. Трансфицируют дрожжевые клетки *his3<sup>-</sup>* и отбирают колонии *HIS3<sup>+</sup>*, утратившие ген *X*. Трансформированы могут быть как гаплоидные, так и диплоидные клетки. В диплоидных клетках встраивание происходит только в одну из двух гомологичных хромосом, которые сегрегируют при образовании гаплоидных спор.

## 5.7. ЭУКАРИОТИЧЕСКИЕ СИСТЕМЫ ХОЗЯИН-ВЕКТОР: ЖИВОТНЫЕ

### а. Трансформация клеток животных

*Нестабильная и стабильная трансформация.* Клонирование сегментов ДНК из клеток животных чаще всего проводят в клетках *E. coli*, и рекомбинантные ДНК, выделенные из этих бактерий, используют для структурного анализа. Однако для проведения функциональных исследований необходимо трансформировать клетки животных. Как и в случае дрожжевых клеток, возможна трансформация нескольких типов. Клетки животных можно

трансформировать путем инфекции или трансфекции с помощью вектора, реплицирующегося независимо как внехромосомный элемент. Некоторые из этих векторов представляют собой вирусные геномы, и в результате образуются вирионы с рекомбинантными ДНК. В других случаях векторы ближе по своим свойствам к плазмидам. Часто такие векторы обладают свойствами челночных векторов, использующих *E. coli* в качестве альтернативного хозяина. В отличие от дрожжевых и бактериальных клеток в клетках животных реплицирующиеся внехромосомные плазмидоподобные векторы часто оказываются нестабильными при размножении клеток после трансфекции, причем эта

нестабильность проявляется даже в случае применения селективного давления. Популяции таких трансформированных клеток часто могут сохраняться только временно (в течение нескольких дней), но этого достаточно для изучения экспрессии генов в векторной ДНК. На самом деле **временную экспрессию** можно зарегистрировать даже в тех случаях, когда вектор не способен реплицироваться в клетках животных. Тем не менее предпочтительнее использовать реплицирующиеся векторы, поскольку они содержатся в клетке в относительно большом числе копий, что в свою очередь приводит к увеличению количества генных продуктов и облегчает изучение экспрессии.

Некоторые векторы животных все-таки существуют в клетках животных в виде стабильно реплицирующихся внехромосомных ДНК. К ним относятся векторы, полученные на основе вируса папилломы крупного рогатого скота (разд. 5.7.в), и другие векторы, содержащие область *ori* репликации одного из герпесвирусов – вируса Эпштейна-Барр. В большинстве случаев стабильная трансформация связана с рекомбинационным внедрением трансформирующей ДНК в геном клетки. Как реплицирующиеся векторы, так и нереплицирующаяся ДНК способны вызывать **инсерционную трансформацию**. Интеграция в этих случаях осуществляется с помощью негомологичной рекомбинации в случайные участки генома. Именно этим данная ситуация отличается от ситуации в дрожжевых клетках, где трансформация чаще всего происходит за счет гомологичной рекомбинации между донорной ДНК и сайтом-мишенью дрожжевого генома.

**Клетки-хозяева.** Для трансформации с помощью рекомбинантных ДНК и для изучения экспрессии генов чаще всего используют животные клетки, выращенные в **культуре клеток** (или, менее строго, культуре тканей). Подходящую среду для эффективной экспрессии многих клонированных генов, в том числе и генов млекопитающих, обеспечивают также ооциты *Xenopus* (гл. 7). В последние годы были разработаны методы введения клонированных последовательностей ДНК в эмбрионы животных на ранних стадиях развития и даже в зиготы млекопитающих. ДНК способна встраиваться в геном таких эмбриональных клеток, вызывая их стабильную трансформацию. Получено жизнеспособное потомство, причем в тех случаях, когда в геноме первичных половых клеток введенная ДНК стабильно сохраняется, новые генотип и фенотип наследуются в последующих поколениях по законам Менделя. В случае плацентарных млекопитающих развитие нового потомства зависит от эффективности имплантации трансформированных ранних эмбрионов в матку матери. Системы для такой модификации первичных половых клеток от

животных (**трансгенные животные**) описаны во введении к ч. IV. Здесь мы рассмотрим случаи использования клеток животных, выращенных в культуре.

Многие типы клеток животных растут и делятся в подходящих питательных средах. При этом некоторые из них прочно прикрепляются к твердым поверхностям, например к дну пластиковой чашки, и продолжают делиться до тех пор, пока поверхность чашки целиком не покроется одиночным слоем клеток (**монослоем**). Другие типы клеток не прекращают деление после вступления в контакт с соседними клетками и образуют на чашке множество слоев; часто этими свойствами отличаются клетки ракового происхождения. Некоторые клетки животных не прикрепляются к поверхности и растут в суспензии подобно клеткам бактерий и дрожжей.

**Первичные** клеточные культуры, полученные из свежих тканей, обычно имеют ограниченное время жизни. Они делятся всего несколько раз, а затем рост их прекращается. Однако из таких культур можно получить **перевиваемые клеточные линии**, клетки которых по существу бессмертны. Такие линии широко используются в экспериментах с рекомбинантными ДНК. Они были получены из клеток различных организмов и органов, например из эмбрионов грызунов, почек обезьян, из *Drosophila* и раковых опухолей грызунов и человека.

При определенных условиях изолированные клетки животных растут и делятся, образуя дискретные колонии или клоны, подобно клеткам бактерий и дрожжей. Получить идеальный клон, ведущий начало от одной клетки, т.е. достичь популяционного единообразия в таких культурах, удается далеко не всегда. Часто линии клеток одинакового происхождения, размножавшиеся в разных лабораториях, проявляют неодинаковые свойства. Объясняется это скорее всего тем, что даже при весьма незначительном изменении условий роста может начаться быстрое размножение клеток, которые приобрели какую-либо благоприятную мутацию. Получены линии клеток, несущих специфические мутации, которые позволяют провести отбор колоний трансформированных клеток с новым генетическим маркером.

**Введение ДНК в клетки животных.** В качестве векторов часто используют геномы вирусов животных. Соответствующие рекомбинантные молекулы, содержащие вставку чужеродной ДНК, обычно упаковываются в вирионы и проникают в клетки хозяина путем инфицирования. В отличие от фага  $\lambda$  упаковку ДНК вирусов животных *in vitro* осуществить пока не удастся и она всегда происходит *in vivo*. Введение рекомбинантных геномов в клетки животных путем инфицирования представляет собой наиболее эффективный способ (по существу

инфицированной оказывается каждая клетка), однако особенности структуры некоторых рекомбинантных молекул препятствуют их упаковке. В таких случаях «голые» молекулы рекомбинантных ДНК могут быть введены в клетки путем трансфекции. Трансфекция оказывается необходимой и в том случае, когда используемые фрагменты ДНК не способны реплицироваться (они напоминают фрагменты, которые применяются для получения стабильной инсерционной трансформации). Для введения голых молекул ДНК в клетки животных и их ядра аналогично тому, как это делается в случае клеток прокариот и дрожжей, требуются специальные методические приемы. При необходимости можно осуществить прямую инъекцию ДНК в клетки и ядра, однако при этом удастся трансфицировать лишь относительно небольшое число клеток. Эту проблему можно решить, проинкубируя клетки с частицами, образовавшимися при совместном осаждении трансфицирующей ДНК и фосфата кальция. Механизм попадания таких частиц в клетки мало изучен, но интактная ДНК достигает ядра. Обычно таким способом удается трансфицировать не более 10% клеток, а для определенных типов клеток этот метод вообще не работает. В клетки многих типов ДНК можно ввести, если подвергнуть смесь клеток и ДНК действию высоковольтного электрического разряда (2–4 тыс. вольт). В основе этого метода, получившего название **электропорации**, по-видимому, лежит образование в мембранах клеток репарируемых отверстий, через которые и проникает ДНК.

**Фенотипический отбор трансформированных клеток.** Обычно после введения чужеродных последовательностей ДНК стабильно трансформируется менее 0,01% трансфицированных клеток животных. Выявление столь редких трансформантов среди миллионов неизмененных клеток и их выделение возможны лишь при наличии высокоэффективных методов отбора. В отличие от ситуации с *E. coli* для отбора животных клеток разработано всего не-

сколько методов. Широко используются два основных подхода: в одном из них применяют мутантные реципиентные клетки, в другом — нормальные. Как и в случае бактерий, при наличии у реципиентных клеток мутации в гене, дикий тип которого находится в составе клонированной трансформирующей ДНК, трансформанты можно отобрать, если выращивать клетки в условиях, позволяющих различать мутантный и дикий фенотипы. Например, клетки  $TK^{-1}$  грызунов содержат мутацию, в результате которой выводится из строя тимидинкиназа (фермент, превращающий тимидин в тимидин-5'-монофосфат). Такие клетки не способны расти в присутствии тимидина, если путь биосинтеза dTMP de novo блокирован аминоптеринем, подавляющим регенерацию тетрагидрофлата (рис. 5.30). Однако клетки, которые трансформированы с помощью ДНК, содержащей функциональный ген тимидинкиназы, способны расти и образовывать колонии в присутствии тимидина и аминоптерина (рис. 5.31). При таких условиях колонии могут образовываться даже в тех случаях, когда частота трансформации составляет менее  $1 \cdot 10^{-6}$ , т.е. это крайне чувствительный способ определения и выделения редких трансформантов  $TK^{-}$ . Ген тимидинкиназы, выделенный из ДНК вируса герпеса с помощью молекулярного клонирования, широко используется при этом способе отбора в сочетании с клетками  $TK^{-}$  грызунов. Подобный подход может найти применение и при других типах мутаций. К сожалению, из-за ограниченности числа линий мутантных клеток этот метод не удастся широко использовать. Чтобы преодолеть эту трудность, часто прибегают к трансфицированию клеток  $TK^{-}$  векторной ДНК, содержащей наряду с геном тимидинкиназы неселективный интересующий нас ген. Трансфицированные клетки, отобранные по фенотипу  $TK^{-}$ , почти всегда содержат ту часть векторной молекулы, которая несет интересующий исследователя ген. Фактически такая **котрансформация** не зависит от того, связаны ковалентно или не

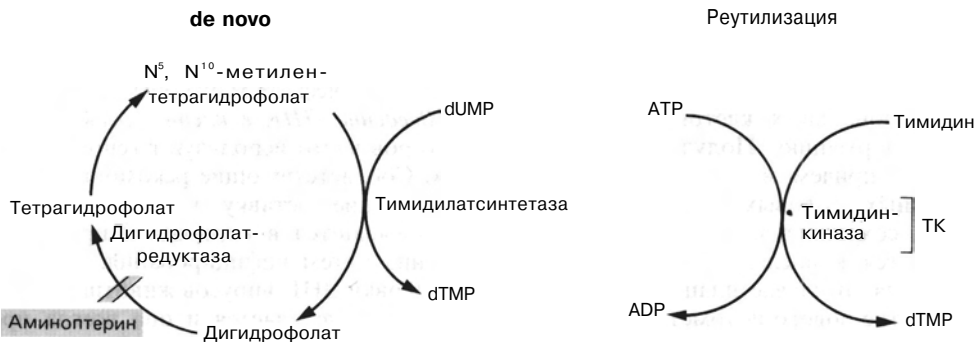
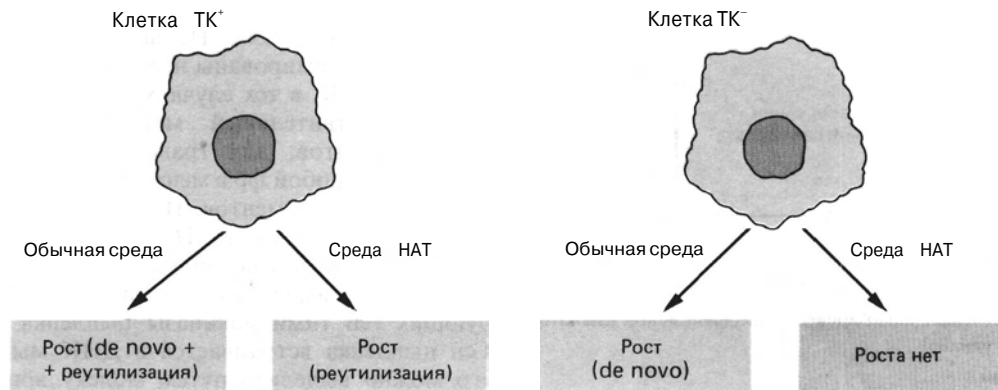


РИС. 5.30.

Синтез dTMP de novo и путем реутилизации. Синтез de novo блокируется аминоптеринем, в результате чего

клетки становятся зависимыми от тимидина и тимидинкиназы.

**РИС. 5.31.**

Фенотипы  $TK^+$ - и  $TK^-$ -клеток при росте на различных средах. Среда НАТ содержит гипоксантин, аминоптерин и тимидин. Гипоксантин добавляют для того, чтобы

связаны ген тимидинкиназы и упомянутый неселективный ген. Котрансформанты постоянно образуются при трансфицировании клеток смесью линейных фрагментов ДНК, содержащих два этих гена в разных молекулах. Более 50% клеток, отобранных как  $TK^+$ , содержат также и второй ген. Обычно оба гена интегрируются поблизости друг от друга, поскольку молекулы имеют тенденцию объединяться после проникновения в клетки до того, как произойдет интеграция.

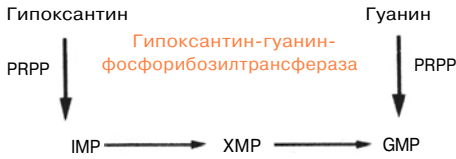
Котрансформация увеличивает число генов, которые с высокой эффективностью могут быть введены в клетки животных. Однако круг разных типов клеток ограничивается доступностью соответствующих мутантных линий. Более общий подход основан на превращении нормальных клеток в клетки, зависящие от трансфицирующего гена. Например, клетки многих видов не способны синтезировать гуаниловую кислоту (GMP) в среде, содержащей микрофеноловую кислоту—соединение, специфически подавляющее инозинат(ИМР)дегидрогеназу (рис. 5.32). Ингибирование снимается добавлением ксантина и введением функционального гена *E. coli gpt*, кодирующего ксантин-гуанин-фосфорибозилтрансферазу. Клетки животных обычно не способны использовать ксантин в качестве предшественника гуаниновых нуклеотидов, так как ксантин является очень плохим субстратом для аналогичного фермента животных—гипоксантин-гуанин-фосфорибозилтрансферазы. Поэтому клетки, трансформированные геном *gpt E. coli* и синтезирующие функциональный фермент ксантин-гуанин-фосфорибозилтрансферазу, можно селективно подрастить в среде, содержащей ксантин и микрофеноловую кислоту (в отсутствие гуанина). Подобным же образом ген *E. coli neo*, кодирующий аминокликозидфотрансферазу, можно использовать в качестве до-

обойти блок в пути синтеза пурина de novo, для которого требуется также тетрагидрофолат.

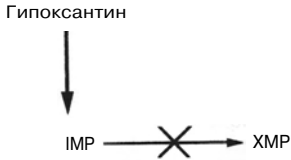
минантного маркера в сочетании с аминокликозидом, получившим название G-418. Это соединение является аналогом антибиотиков неомицина и канамицина, которые блокируют синтез белков в эукариотических клетках и инактивируются фосфотрансферазой (табл. 5.3). Гены *E. coli* обычно вводятся в клетку с помощью челночных векторов, которые клонируются и реплицируются в клетках *E. coli*. При конструировании рекомбинантных молекул нормальные промоторы *E. coli* замещают на транскрипционные регуляторные сигнальные последовательности, проявляющие активность в клетках животных. Гены *E. coli gpt* и *neo* можно использовать для котрансформации нормальных клеток геном, не имеющим селективного маркера. Доминантными селективными маркерами могут служить также бактериальный ген резистентности к гиромцину (*hph*) и ген млекопитающих, кодирующий устойчивую к метотрексату форму дигидрофолат-редуктазы.

**Векторы.** В качестве основы для конструирования векторных молекул широко использовался геном одного из паповавирусов приматов—вируса обезьян 40 (SV40). Те же причины, которые привели к выбору *E. coli* при разработке бактериальных систем клонирования (этот организм был достаточно хорошо изучен, и с ним было легко манипулировать), обусловили выбор SV40 в качестве вектора при работе с животными клетками. За годы, прошедшие с начала первых разработок векторов на основе вируса SV40, было получено большое количество данных о структурной организации и генетических свойствах геномов других вирусов эукариот. В настоящее время на основе некоторых из этих вирусов созданы интересные системы хозяин-вектор. Для млекопитающих это папилломавирусы, аденовирусы, вирус Эпштейна-Барр, вирус оспы

## А. Нормальные клетки животных



## Б. Клетки, рост которых ингибируется микрофеноловой кислотой (в отсутствие гуанина)



## В. Снятие ингибирования микрофеноловой кислотой (в отсутствие гуанина)

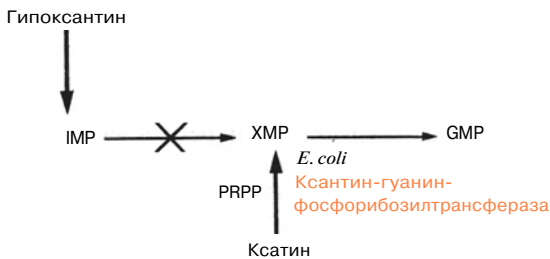


РИС. 5.32.

Синтез GMP при различных условиях. PRPP - фосфорибозилпирофосфат.

крупного рогатого скота и ретровирусы, а для насекомых — бакуловирусы.

Как мы уже отмечали, при работе с животными клетками используют два типа векторов. В некоторых случаях они аналогичны бактериальным плазидам или фаговым векторам. После трансфекции или инфицирования эти векторы реплицируются и сохраняются в виде независимых внехромосомных геномов или упаковываются в инфекционные вирусные частицы. В других случаях векторы не способны реплицироваться и используются для введения специфических сегментов ДНК в клетки с целью изучения временной экспрессии клонированного гена или получения стабильной трансформации при встраивании вектора в геном хозяина.

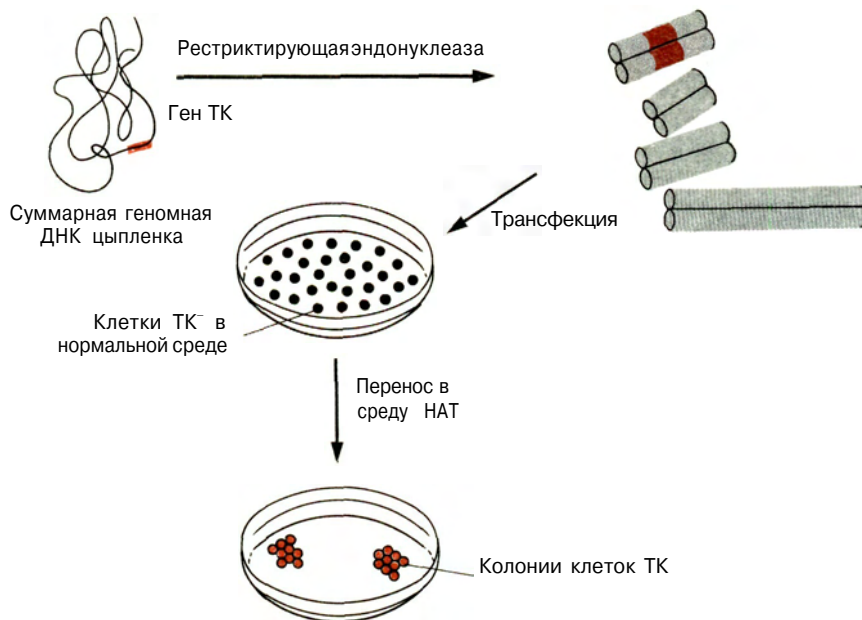
**Трансформация без векторов.** В этой главе мы рассматривали векторы, которые представляют собой молекулы ДНК со специфическими клони-

рованными вставками. Но поскольку клетки могут быть трансформированы нереплицирующимися сегментами ДНК, в тех случаях, когда имеется достаточно чувствительный метод отбора искомым трансформантов, для трансформации можно использовать любой фрагмент ДНК или смесь неколонированных фрагментов. Например, при трансформации мышинных клеток ТК суммарной ДНК цыпленка, подвергнутой эндонуклеазному расщеплению, можно выделить колонии клеток, экспрессирующих ген тимидинкиназы цыпленка (рис. 5.33). Ген цыпленка встраивается в ДНК мыши, откуда его можно выделить путем молекулярного клонирования с использованием специфических методов отбора. Этот подход оказался очень ценным при клонировании многих важных генов, в том числе онкогенов, содержащихся в раковых клетках.

## б. Векторы на основе SV40

**Вирус и его геном.** SV40 — это сферический вирус; его геном представлен одной замкнутой кольцевой двухцепочечной ДНК (рис. 5.34). Вирусная частица содержит три структурных кодируемых вирусом белка. VP1 — главный белок вириона, он является основной составной частью капсида; VP2 и VP3 представлены в меньших количествах. ДНК в вирусной частице находится в форме хроматина и связана с четырьмя гистонами (гл. 2). Природным хозяином SV40 являются макаки-резусы (*Macaca malatta*), но в лаборатории в качестве хозяев используют в основном клетки почек африканской зеленой мартышки (*Cercopithecus aethiops*). Расшифрована нуклеотидная последовательность всего генома SV40, состоящая из 5243 п.н., и установлены участки, кодирующие пять вирусных белков: малый Т-антиген, большой Т-антиген, VP1, VP2 и VP3. На рис. 5.34 приведена карта вирусного генома. Весь геном разделен на 10 частей; цифры у стрелок указывают число пар нуклеотидов. Первая пара находится в точке начала репликации ДНК (*ori*).

Вирион SV40, проникнув в клетку, направляется к ядру. Здесь происходит его «раздевание», а затем в строгой временной последовательности — экспрессия генов. Новые вирусные частицы появляются примерно через 30 ч после инфицирования монослоя клеток почки обезьяны, а через 4 сут все клетки разрушаются. Первыми транскрибируются гены двух ранних белков, малого и большого Т-антигенов, в направлении против движения часовой стрелки. Сайт полиаденилирования находится на карте вблизи нуклеотида 2550. Ни малый, ни большой Т-антигены не входят в состав зрелых вирусных частиц, однако большой Т-антиген играет важную роль в репликации вирусной ДНК, начинающейся вскоре после образования ранних белков.

**РИС. 5.33.**

Трансформация ТК<sup>-</sup>-клеток мыши в клетки ТК<sup>+</sup> с помощью ДНК цыпленка. Для трансформирования клеток ТК<sup>-</sup>, растущих в культуре, используют суммарную ДНК цыпленка, подвергнутую эндонуклеазному расщеплению. После трансфицирования клетки переносят в среду НАТ. Клетки, не получившие ДНК цыпленка или получившие фрагменты, не содержащие гена тимидинкиназы,

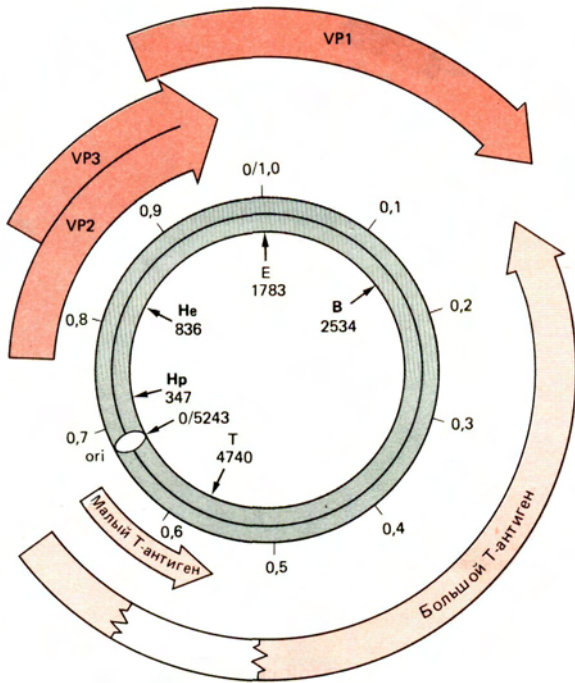
не способны расти в этой среде, а редкие трансформированные клетки, в которых произошла интеграция фрагмента ДНК с нужным геном в геном мыши, растут и образуют колонии. Следует заметить, что в тех случаях, когда рестриктирующая эндонуклеаза разрезает ген тимидинкиназы, трансформированные клетки не образуются.

Репликация начинается в точке *ori*, идет в обоих направлениях и заканчивается в месте встречи репликативных вилок. Затем происходит сшивание дочерних двухцепочечных молекул ДНК с образованием замкнутых кольцевых структур. В каждой инфицированной клетке образуется более  $10^5$  копий вирусного генома. Вскоре после инициации репликации ДНК замедляется синтез ранних белков и активизируется транскрипция «поздних» генов, которая начинается вблизи *ori* и протекает в направлении движения часовой стрелки, в результате чего синтезируются поздние белки VP1, VP2 и VP3. Накопление дочерних молекул ДНК и вирионных белков завершается сборкой новых вирусных частиц, их высвобождением и гибелью клетки. Инфекционный процесс в полной мере может быть иницирован в результате трансфекции с помощью очищенной ДНК вируса SV40.

**Несколько типов SV40-векторов.** Общие принципы конструирования векторов на основе SV40 аналогичны принципам, используемым на основе фага  $\lambda$ . Применяя соответствующие рестриктирующие эндонуклеазы, удаляют определенные области вирусного генома и замещают их другими сегментами ДНК. В зависимости от того, какой именно

сегмент вирусной ДНК удаляется, образуется тот или иной вектор. Различают три типа векторов, сконструированных на основе вируса SV40. Каждый из них имеет свои преимущества и недостатки при использовании в различных экспериментальных системах. Во-первых, это **трансдуцирующие SV40-векторы**, реплицирующиеся в соответствующих клетках-хозяевах (клетки почек обезьяны) и упаковывающиеся в вирионы. Они аналогичны специализированному трансдуцирующему фагам и векторам, сконструированным на основе фага  $\lambda$ . Во-вторых, это SV40-векторы, способные реплицироваться, но не упаковываться. Они называются **плазмидными SV40-векторами**. И наконец, существуют **пассивные трансформирующие векторы**, которые не способны ни реплицироваться, ни упаковываться. Они представляют собой молекулы, в которые включены небольшие сегменты генома SV40, способствующие экспрессии генов.

**Трансдуцирующие SV40-векторы.** Чтобы SV40-вектор мог упаковываться с образованием вирусных частиц, необходимо выполнение трех важных требований. Во-первых, вектор должен содержать сегмент размером 300 п.н., включающий точку *ori*. Это необходимо для обеспечения синтеза вирусной



**РИС. 5.34.**

Геном SV40. С наружной стороны карты приведены условные генетические единицы, с внутренней - число пар оснований. Указаны уникальные рестриктазные сайты. Участки, кодирующие пять вирусных белков, вынесены за пределы карты. Заметьте, что большой Т-антиген кодируется двумя отдельными группами генов (см. гл. 8).

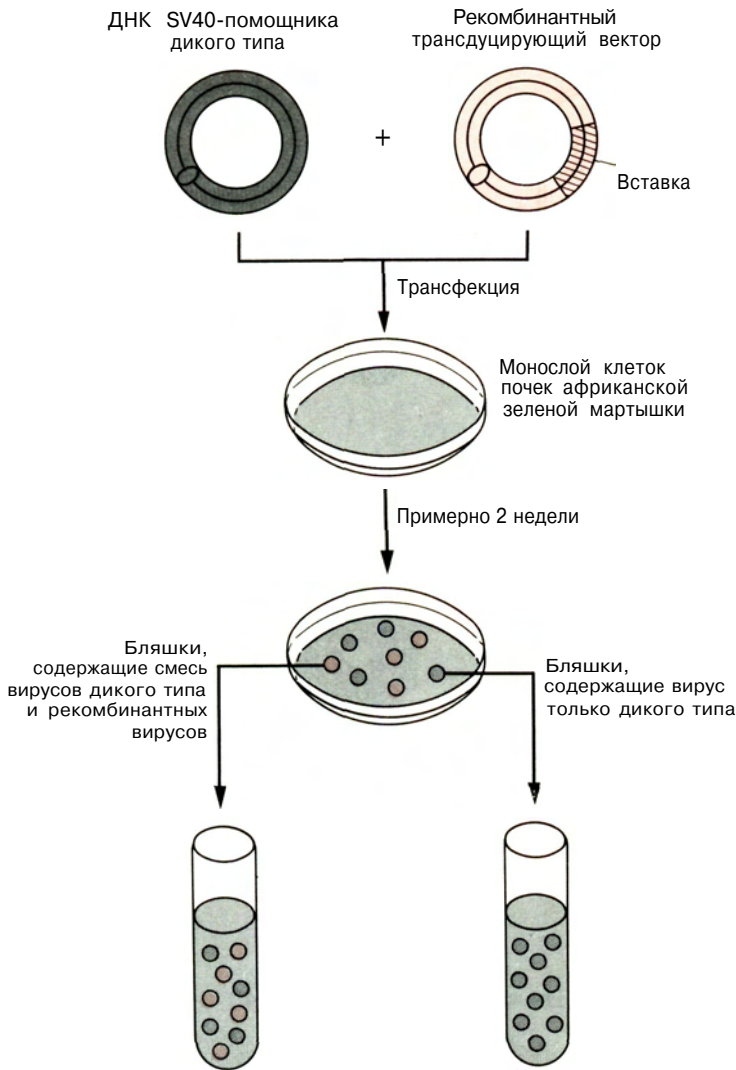
ДНК и регуляторных последовательностей, отвечающих за синтез мРНК. Во-вторых, суммарная длина рекомбинантной ДНК не должна превышать 5300 п.н. и быть меньше 3900 п.н., т.е. вектор должен иметь определенный размер, необходимый для упаковки ДНК в вирусные частицы. В-третьих, должны быть в наличии белки большой Т-антигена, VP1, VP2 и VP3, хотя их кодирование самим рекомбинантным геномом необязательно. К интактному геному SV40 уже практически невозможно добавить какой-то фрагмент ДНК, чтобы не было нарушено условие, при котором может происходить его упаковка. Поэтому общая стратегия состоит в делетировании кодирующей области вектора, замещении ее нужным фрагментом ДНК и обеспечении синтеза продукта утраченного гена с помощью аналогичного гена, входящего в состав генома либо вируса-помощника, либо хозяина (т.е. возмещении гена в *транс*-положении). Например, вирус-помощник дикого типа способен синтезировать белки Т или VP1, VP2 и VP3 в количестве, достаточном для удовлетворения как собственных нужд, так и нужд рекомбинантного вируса. На рис.

5.35 представлен эксперимент с трансфицированием клеточного монослоя смесью ДНК вируса SV40 дикого типа и ДНК дефектного вируса, несущей вставку. Клетки, получившие либо ДНК дикого типа, либо обе ДНК, подвергаются литической инфекции, и дочерние ДНК упаковываются в вирусные частицы.

Инфицированные клетки в монослое можно обнаружить по образующимся бляшкам. Отметим, что отличить бляшки, содержащие смешанное потомство, от тех, которые содержат только SV40 дикого типа, довольно трудно. Для этого нужно провести дополнительный анализ вирионов, полученных из отдельных бляшек, и идентифицировать те из них, которые содержат рекомбинантную ДНК. Однако эту проблему легко решить, если использовать вектор, сохраняющий по крайней мере один функциональный вирусный ген, и вирус-помощник, дефектный в отношении гена, поставляемого векторной молекулой. В этом случае негативные колонии будут образовываться только теми клетками, которые трансфицированы обоими ДНК. Например, если встроенный сегмент ДНК замещает участок, кодирующий поздние белки VP1, VP2 и VP3, то соответствующий вирус-помощник должен иметь мутацию в гене большого Т-антигена. Рекомбиант обеспечивает синтез большого Т-антигена, а вирус-помощник отвечает за синтез белков VP1, VP2 и VP3 (рис. 5.36). И наоборот, если встроенный сегмент ДНК замещает раннюю область SV40, то вирус-помощник должен обеспечивать синтез функционального большого Т-антигена, но не капсидных белков. В отличие от трансфекции с использованием в качестве помощника вируса дикого типа каждая бляшка, образующаяся после трансфекции векторной и дефектной вирусными ДНК, содержит смешанное потомство.

Утраченную функцию может возмещать и ген хозяина (рис. 5.37). Например, COS-клетки—это клетки почек африканской зеленой марышки, у которых в геном интегрирован ген большого Т-антигена вируса SV40. Они продуцируют достаточное количество большого Т-антигена, чтобы могла происходить репликация молекул, содержащих область *ori* репликации вируса SV40 размером примерно 80 пар оснований. Если такой трансдуцирующий вектор кодирует также белки VP1, VP2 и VP3, то вирусные частицы будут образовываться в клетках COS даже в отсутствие вируса-помощника. Поскольку никакой вирус-помощник не требуется, образующиеся вирионы содержат только геном трансдуцирующего вектора.

*SV40-плазмидные векторы.* SV40-векторы второго типа способны реплицироваться в клетках обезьян, но не могут упаковываться с образованием вирусных частиц (рис. 5.38). Для них существенно

**РИС. 5.35.**

Комплементация трансдуцирующего SV40-вектора с помощью вируса SV40 дикого типа. Монослой клеток почек африканской зеленой мартышки трансфицируют смесью ДНК SV40 дикого типа и трансдуцирующего SV40-вектора, в котором ранняя область (кодирующая большой Т-антиген) вирусной ДНК замещена вставкой чужеродной ДНК. Отметим, что область начала репликации вирусной ДНК должна оставаться интактной. Вставка чужеродной ДНК может замещать и позднюю область, кодирующую белки VP1, VP2 и VP3. В клетках, получивших ДНК обоих типов, вирусные функции, утраченные рекомбинантами, восполняет геном дикого типа. Обе ДНК реплицируются, и дочерние молекулы упаковываются в вирионы, которые совместно инфицируют соседние клетки. Вокруг первой инфицированной клетки на монослое появляется прозрачная зона - бляшка. Материал, полученный из некоторых бляшек, представляет собой смесь вирусов дикого типа и рекомбинантных вирусов. В других бляшках содержатся только частицы вируса дикого типа.

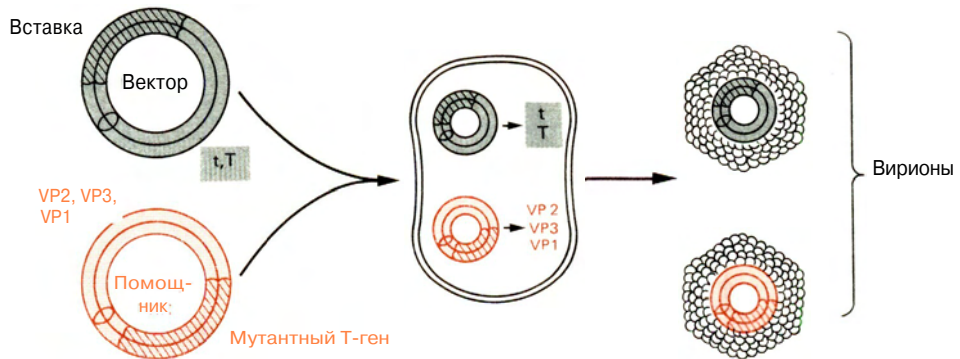
сохранение как области *ori* репликации, так и гена большого Т-антигена, в то время как гены белков VP1, VP2 и VP3 отсутствуют. С другой стороны, синтез большого Т-антигена может обеспечиваться и клетками COS. Плазмидные SV40-векторы могут быть челночными и существовать в клетках *E. coli* и животных, поскольку никаких ограничений на их размер, необходимый для упаковки, не налагается. Рекомбинантные молекулы получают в клетках *E. coli*, а затем трансфицируют ими клетки обезьяны. Такие клетки могут использоваться для изучения временной экспрессии генов, встроенных в векторные молекулы. Поскольку при репликации вектора образуется большое число копий, доза любого гена, содержащегося в векторе, увеличивается, что облегчает изучение соответствующих генных продуктов. Было бы очень ценно, если бы

такие SV40-плазмидные векторы вели себя в клетках животных аналогично тому, как ведут себя плазмиды в клетках бактерий. В этом случае рекомбинантный геном передавался бы дочерним клеткам при делении и мы имели бы стабильную систему, позволяющую изучать экспрессию генов. Однако плазмидные векторы оказываются нестабильными в большинстве клеток животных и не могут сохраняться в них бесконечно долго. Относительной стабильности можно достичь, если использовать клетки COS, поставляющие большой Т-антиген для репликации, вектор с селективным маркером, аналогичным гену *neo E. coli*, и среду G-418.

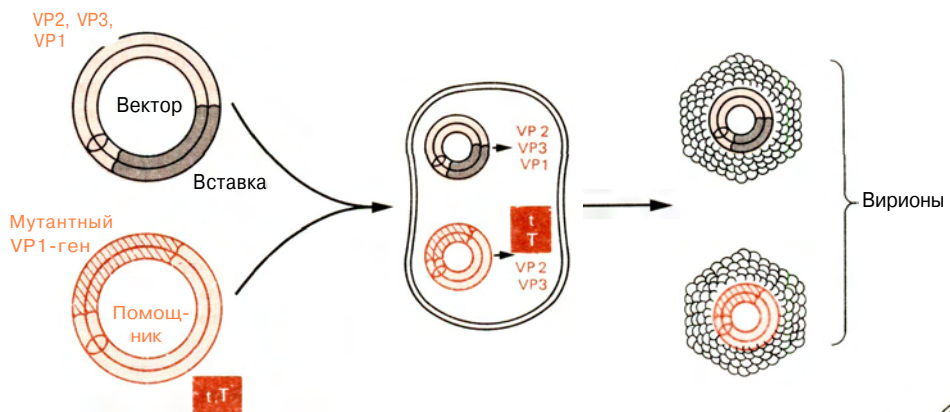
Стабильной трансформации удается достигнуть в тех случаях, когда SV40-плазмидные векторы интегрируются с геном клетки-хозяина. Частота такого события составляет  $10^{-5}$ – $10^{-3}$  в зависимости



## А. Замещение области поздних генов SV40



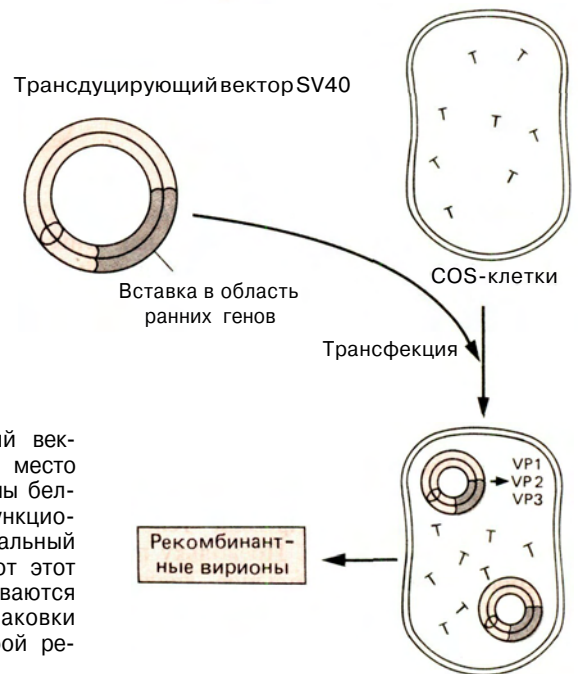
## Б. Замещение области ранних генов SV40

**РИС. 5.36.**

Комплементация трансдуцирующих векторов SV40 с помощью дефектного вируса-помощника. А. Вставка чужеродной ДНК замещает область поздних генов SV40. Последовательности, кодирующие белки VP1, VP2 и VP3, восполняются геномом помощника, который содержит мутантный ген большого Т-антигена. Б. Вставка чужеродной ДНК замещает область, кодирующую большой Т-антиген. Соответствующие гены предоставляет вирус-помощник, содержащий мутацию в гене белка VP1.

**РИС. 5.37.**

COS-клетки комплементируют трансдуцирующий вектор SV40, утративший ген большого Т-антигена, место которого занимает вставка чужеродной ДНК. Гены белков VP1, VP2 и VP3 не утрачены и нормально функционируют. Геном COS-клеток содержит функциональный ген большого Т-антигена, и клетки синтезируют этот белок (Т). После трансфекции в клетках оказываются все белки, необходимые для репликации и упаковки SV40. Все дочерние вирионы представляют собой рекомбинантные трансдуцирующие векторы.



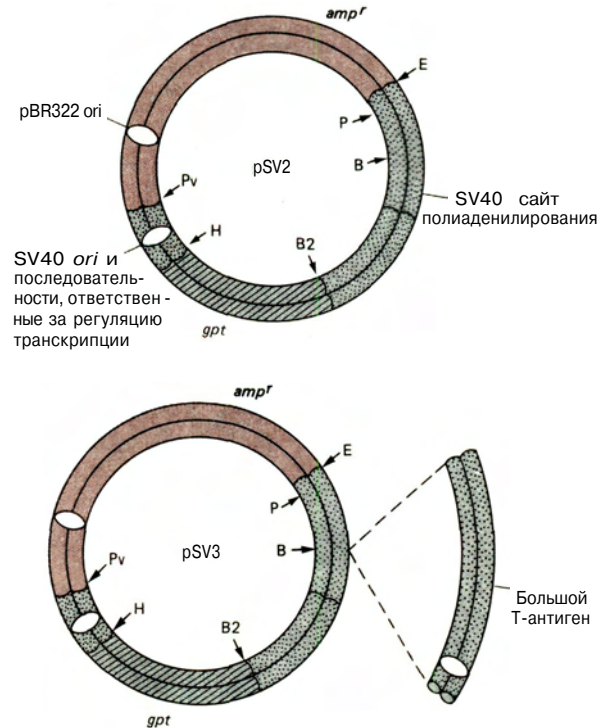
от используемой системы. Если вектор содержит селективный маркер, то колонии трансформированных клеток обезьяны можно выделить.

**Пассивные трансформирующие векторы.** Векторы третьего типа не способны к самостоятельной репликации. Они лишь доставляют в клетку клонированные сегменты ДНК, которые затем внедряются в клеточную ДНК. Образующиеся стабильно трансформированные клетки реплицируют приобретенную ДНК как часть своего собственного генома. Обычно такие векторы напоминают челночные, поскольку сначала для клонирования и выделения рекомбинантных ДНК используют клетки *E. coli*, а затем трансфицируют этой ДНК различные клетки млекопитающих (например, pSV2 в случае, представленном на рис. 5.38). Используются не только клетки почек обезьяны, поскольку для таких векторов не требуются специфические условия для репликации. Если вектор содержит подходящий маркер, то можно отобрать колонии трансформантов. Маркером может служить ген тимидинкиназы вируса герпеса или гены *gpt* и *neo E. coli*. Сегменты ДНК вируса SV40, включаемые в подобные векторы, обычно содержат области регуляции транскрипции и сайты полиаденилирования.

### в. Векторы на основе вируса папилломы крупного рогатого скота

**Вирус и его геном.** Папилломавирусы провоцируют образование бородавок — доброкачественных новообразований кожи у многих млекопитающих; они относятся к той же группе, что и SV40. Папилломавирусы характеризуются высокой видоспецифичностью. У них имеется оболочка — капсид; геном представляет собой замкнутую кольцевую двухцепочечную ДНК длиной около 8 т.п.н. ДНК вируса папилломы крупного рогатого скота, как и ДНК SV40, навита на нуклеосомные коры, образуемые четырьмя клеточными гистонами. Изучение биологических и генетических свойств папилломавирусов сильно осложнялось тем, что отсутствовали клеточные линии, способные поддерживать репликацию вируса. Кроме того, вирус приходилось выделять непосредственно из бородавок. Однако с развитием технологии рекомбинантных ДНК небольшие количества папилломавирусной ДНК, которые удавалось выделить из бородавок, были очищены и амплифицированы с помощью молекулярного клонирования в клетках *E. coli*. Генетический анализ показал, что ДНК папилломавирусов может служить весьма полезным уникальным вектором для клеток животных.

Полностью установлена нуклеотидная последовательность ДНК папилломавирусов крупного рогатого скота и человека. Геном папилломавируса



**РИС. 5.38.**

SV40-плазмидные векторы для клонирования в *E. coli*. Эти челночные векторы содержат сегмент длиной примерно 2,3 т.п.н, происходящий из плазмиды pBR322 и содержащий область *ori*, и ген ампициллиназы. Сегмент вектора, происходящий из генома SV40, находится на генетической карте вируса между нуклеотидами 5171 и 270 и содержит область *ori* и регуляторные последовательности для осуществления ранней транскрипции; он присоединен к 5'-концу кодирующей области гена *gpt E. coli*. К 3'-концу гена *gpt* тоже присоединены SV40-последовательности, которые на генетической карте вируса находятся в окрестности 2500 н.п. Они содержат сигналы полиаденилирования мРНК в клетках животных и сигналы, необходимые для созревания мРНК. Эти векторы были получены клонированием в *E. coli*. Вверху представлен вектор pSV2-плазмидный вектор, содержащийся в COS-клетках, не способный реплицироваться в нормальных клетках обезьяны из-за отсутствия большого Т-антигена. В случае нормальных клеток обезьяны или других клеток млекопитающих вектор pSV2 используют при изучении временной экспрессии гена *gpt* или какого-то другого гена, встроенного, например, в сайт *Bat H1* или *Eco RI*. Внизу представлен вектор pSV3, не нуждающийся в клетках COS для репликации и действующий как плазмидный вектор в нормальных клетках обезьяны, поскольку он содержит ген большого Т-антигена. Этот ген вместе с последовательностью, ответственной за регуляцию его транскрипции, встроен в сайт *Bat H1* плазмиды pSV2. Как pSV2, так и pSV3 могут включаться в геном клетки-хозяина и вызывать стабильную трансформацию. На  $10^4$ – $10^5$  трансформируемых клеток приходится одна клетка, способная образовать колонию в присутствии микофеноловой кислоты.

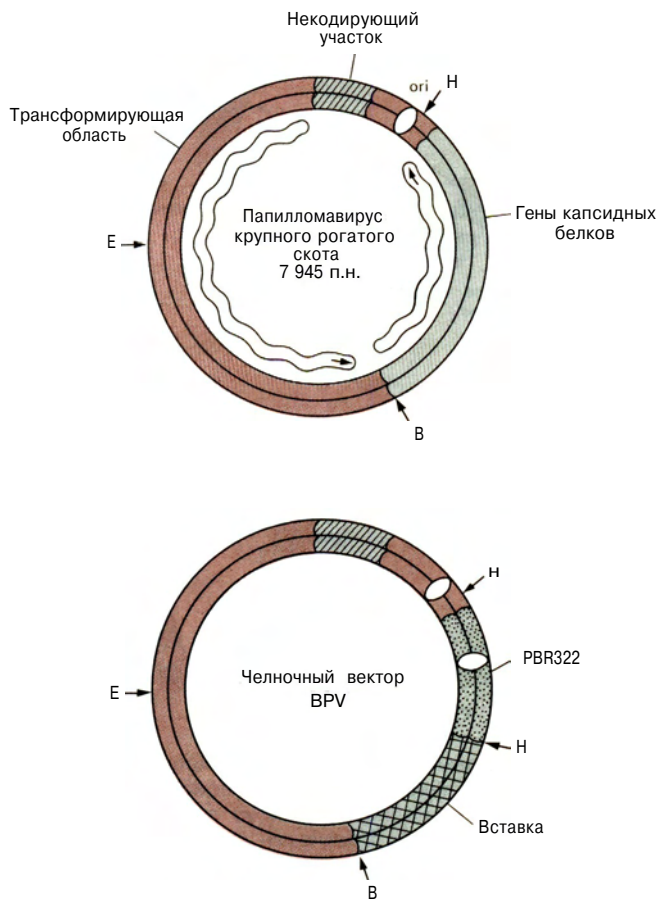


РИС. 5.39.

Схематическое представление генетических карт папилломавируса крупного рогатого скота (BPV) и челночного BPV-вектора. Для трансформации достаточно области, включающей некодирующий регуляторный сегмент (на рисунке она выделена цветом). Наиболее вероятное местоположение области *ori* отмечено овалом; вторая область *ori* способна функционировать при определенных условиях. С внутренней стороны кольца схематически представлена транскрипция генов, необходимых для трансформации, и генов капсидного белка. В вектор включены только трансформирующие вирусные последовательности, а также сегменты *pBR322*, необходимые для репликации и отбора в клетках *E. coli*.

крупного рогатого скота (рис. 5.39) используют в качестве основы для конструирования векторов животных. В отличие от вируса SV40 папилломавирус крупного рогатого скота (BPV) реплицируется в виде стабильной плазмиды в клетках грызунов и многих клетках быка. Вирионы при этом не образуются, поскольку экспрессия капсидного белка, насколько это известно, происходит только в терминально дифференцированных эпидермальных клетках. Инфицированные клетки не погибают, и геном BPV передается дочерним клеткам. Такие геномы трансформируют клетки, вызывая в них генетические изменения, в частности превращая нормальные клетки в раковые. Трансформанты в таких случаях можно выявить благодаря тому, что они образуют возвышающиеся над монослоем колонии, так называемые **фокусы**. Клоны трансформированных клеток можно получить в культуре аналогично тому, как получают колонии из отдельных клеток *E. coli*. Подобно многим трансформированным клеткам эукариот, эти клетки растут быстрее своих нормальных двойников. Для поддержания трансформированного состояния необходи-

мы только те гены, которые на рис. 5.39 обозначены как «трансформирующие»; на их долю приходится 69% длины всего генома. Стабильная трансформация с помощью ДНК вируса BPV не связана с интеграцией вирусной ДНК в геном клетки; достаточно, что ДНК находится в клетке в состоянии плазмиды. Каждая трансформированная клетка содержит от 10 до 100 копий генома вируса BPV.

**Конструирование векторов.** Векторы на основе ДНК папилломавируса можно сконструировать лигированием чужеродных вставок с «трансформирующим» сегментом вирусного генома длиной примерно 5500 п.н. (рис. 5.39). Обычно челночные векторы получают путем клонирования эукариотического компонента в плазмидном векторе *E. coli* (например, в *pBR322*). Перед трансфекцией эукариотических клеток векторные молекулы можно разрезать с помощью рестриктирующей эндонуклеазы, чтобы отделить *pBR322* от ДНК BPV (и вставки), поскольку некоторые плазмидные последовательности подавляют репликацию ковалентно связанной с ними ДНК вируса папилломы. Этими линейными сегментами ДНК трансфицируют животные клетки, при этом сегмент, содержащий вставку и трансформирующую область вируса BPV, замыкается в кольцо и реплицируется как независимая плаزمид. На основе BPV-вектора сконструировано множество разных вариантов. Например, при введении в вектор гена *neo E. coli*, функционирующего в клетках животных, трансформированные клетки можно отбирать на среде с аминогликозидом G-418. Благодаря этому отпадает

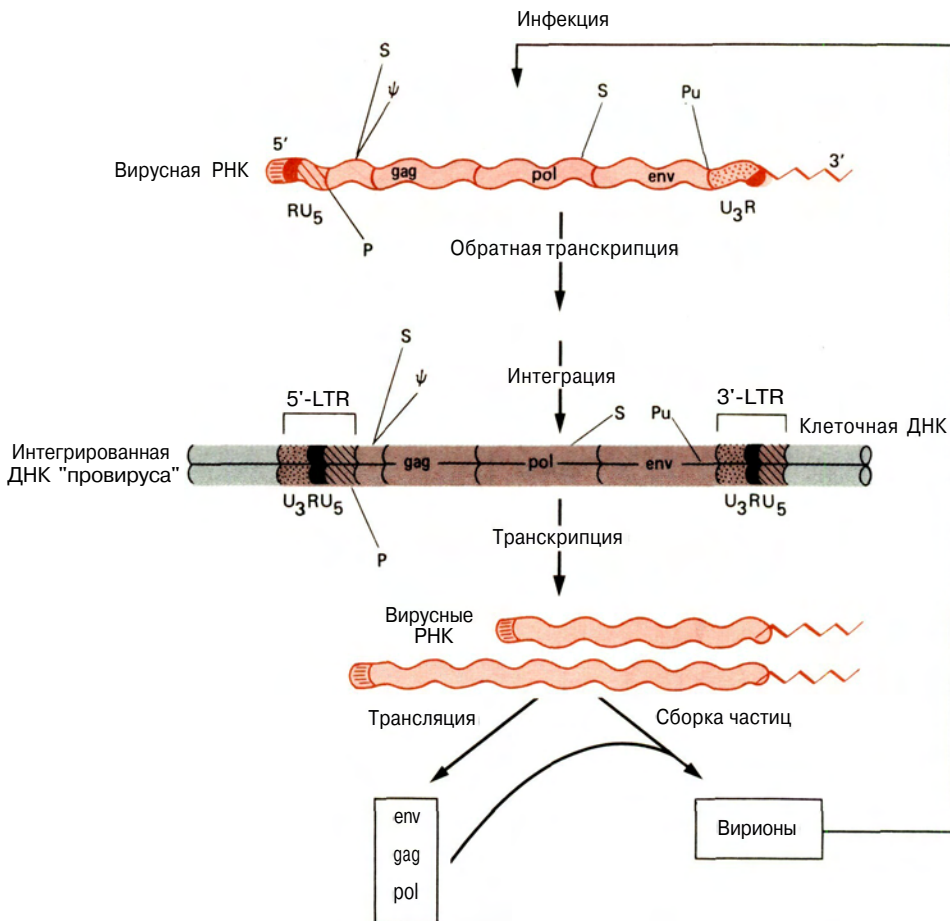
необходимость в идентификации фокуса и тем самым расширяется круг чувствительных клеток, в который теперь могут входить клетки, не проявляющие такого морфологического признака.

### г. Векторы на основе ретровирусов

*Жизненный цикл ретровирусов* (разд. 2.2.а). Репликация одноцепочечного РНК-генома ретровирусов включает следующие стадии: образование двухцепочечной ДНК, являющейся копией РНК, с помощью обратной транскриптазы; встраивание вирусной ДНК в геном хозяина; транскрипцию встроенной ДНК с образованием мРНК и дочерних вирусных геномов; синтез вирусных структурных белков и обратной транскриптазы; сборку новых

вирусных частиц и высвобождение их из клетки (рис. 5.40). Наряду со структурными белками вирусные частицы содержат обратную транскриптазу, которая необходима для инициирования жизненного цикла вируса. Интегрированный вирусный геном получил название **провируса**. Его включение в геном клетки вызывает стабильную трансформацию последней. Трансформированные клетки часто идентифицируют по появлению плотных фокусов из быстрорастущих клеток, напоминающих опухолевые. Многие ретровирусы являются онкогенными.

Необходимым условием продуктивности жизненного цикла ретровируса является наличие в его геноме определенных нуклеотидных последовательностей, способных функционировать только в *цис*-положении. В то же время продукты вирусных



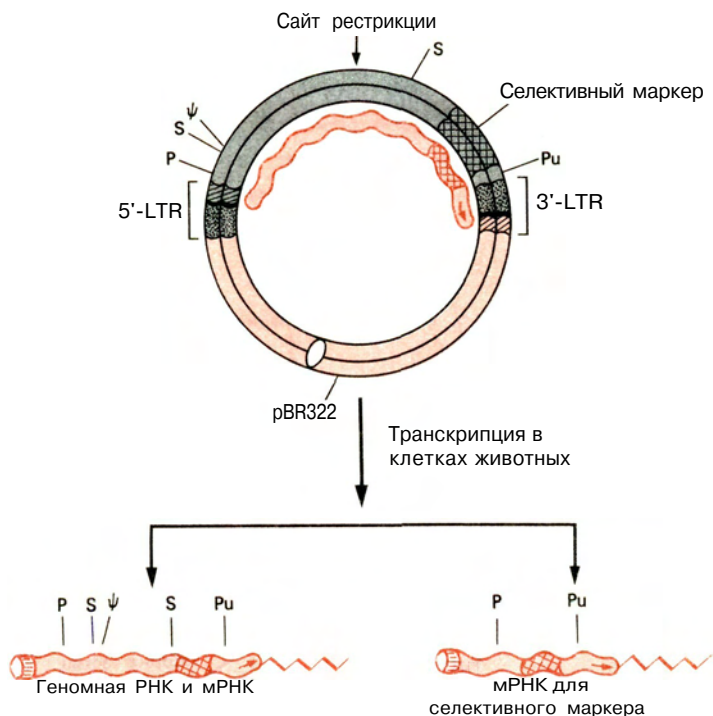
**РИС. 5.40.**

Жизненный цикл ретровируса (более подробно см. гл. 2). После инфицирования клеток РНК-геном ретровируса превращается в дуплексную ДНК с помощью обратной транскриптазы, присутствующей в вирионе. В результате интеграции дуплексной ДНК с клеточным геномом образуется провирус. ДНК провируса фланкирована длинными прямыми концевыми повторами (5'-LTR и 3'-LTR). На 5'-LTR находятся сигналы регуля-

ции транскрипции. В результате транскрипции провирусной ДНК образуются мРНК и новые вирусные геномы, а в результате трансляции мРНК - вирусные белки gag, pol и env. Далее происходит сборка вирусных частиц, которые отпочковываются от плазматической мембраны. Экспрессия невирусных генов, встроенных в центральную кодирующую область, также находится под контролем 5'-LTR.

генов – группоспецифичные антигены (**gag**) и белки оболочки (**env**), а также обратная транскриптаза (**pol**) – могут поставляться вирусом-помощником или провирусом, находящимися в *транс*-положении. Структурные элементы, которые должны находиться в *цис*-положении, включают: 1) элементы, участвующие в обратной транскрипции (они обозначаются как R, U5, U3, P и Pu); 2) короткую последовательность  $\psi$ , необходимую для упаковки геномов в вирионы; 3) две последовательности (S), необходимые для осуществления сплайсинга РНК с образованием функциональной мРНК, ответственной за синтез белка оболочки (рис. 5.40). Длинные прямые концевые повторы (LTR) на одном из концов дуплексной ДНК вирусного генома и провируса содержат элементы, ответственные за регуляцию транскрипции, которые способствуют транскрипции вирусных генов.

*Конструирование векторов на основе ретровирусов.* В основе конструирования ретровирусных векторов лежат три важных принципа. Во-первых, векторы должны сохранять последовательности вирусного генома, необходимые для репликации, экспрессии генов и упаковки. Синтез вирусных белков может определяться генами, находящимися в *транс*-положении. Во-вторых, чужеродные сегменты ДНК должны встраиваться в кодирующие области генома или отчасти замещать их. В-третьих, ретровирусные векторы обычно предназначены для использования в качестве трансдуцирующих векторов с упаковкой рекомбинантного генома в вирион. Сами ретровирусы представляют собой природные трансдуцирующие вирусы: они рекомбинируют с клеточными геномами и трансдуцируют клеточные гены во вновь инфицируемые клетки как часть провируса.



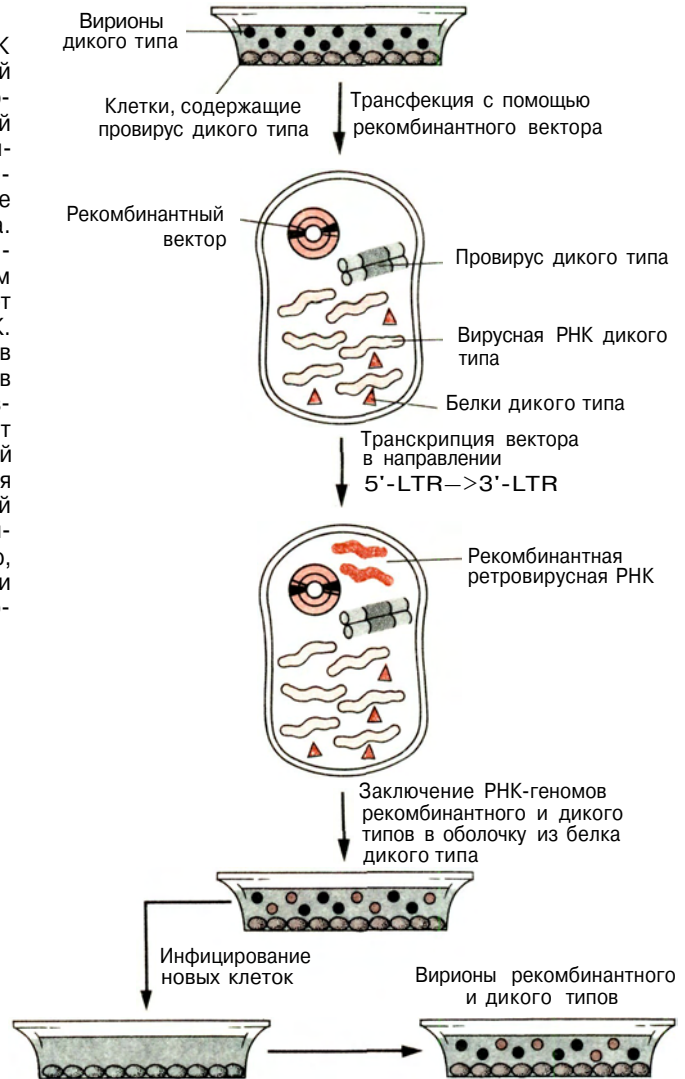
**РИС. 5.41.**

Челночный вектор, сконструированный на основе ретровируса. pBR322-последовательности позволяют осуществлять отбор, репликацию и клонирование в *E. coli*. К этим плазмидным последовательностям присоединен модифицированный ретровирусный геном, фланкированный 5'- и 3'-LTR. Вектор содержит и другие последовательности (P,  $\psi$ , S и Pu), необходимые для репликации, образования мРНК и сборки ретровируса, а также селективный маркер (например, гены *leo* и *gpt* *E. coli* или ген тимидинкиназы вируса герпеса). В уникальный сайт рестрикции или вместо каких-то участков ДНК ретровируса могут быть включены другие сегменты ДНК. После трансфекции клеток животных на 5'-LTR инициируется транскрипция, которая заканчивается в

точке терминирования транскрипции, находящейся на 3'-LTR. Одни транскрипты являются вирусными геномами, которые упаковываются в белковую оболочку. Белки оболочки кодируются либо вектором, либо генами, находящимися в *транс*-положении в составе вируса-помощника или генома клеток-хозяев. Другие транскрипты используются как мРНК. В приведенном примере селективный маркер находится в положении, обычно занимаемом геном *env*. Транскрипты, которые должны играть роль мРНК для генов этой области, утрачивают все последовательности, находящиеся между S-сайтами, в результате сплайсинга промежуточной РНК и воссоединения двух концов по S-сайтам.

**РИС. 5.42.**

Превращение дуплексной рекомбинантной ДНК ретровирусного вектора в инкапсидированный РНК-геном в клетках хозяина, содержащих провирус дикого типа. Клетки в чашке, изображенной в верхней части рисунка, содержат интегрированный провирус дикого типа. В этих клетках синтезируются РНК-геномы дикого типа и вирусные белки и происходит сборка вирионов дикого типа. Вскоре после трансфекции клеток двухцепочечным рекомбинантным ретровирусным вектором (рис. 5.41) в результате транскрипции, идущей от 5'-LTR к 3'-LTR, образуется рекомбинантная РНК. РНК-геномы рекомбинантного и дикого типов упаковываются в оболочку, состоящую из белков дикого типа, и в среде накапливается смесь разных вирионов. Такую жидкую среду используют в дальнейшем как базовый инокулят, содержащий рекомбинантный вирус и вирус-помощник, для инфицирования новых клеток. После обратной транскрипции в некоторых вновь инфицированных клетках происходит интеграция как дикого, так и рекомбинантного провирусов. Такие клетки используют впоследствии как продуценты вирионов разных типов.



ДНК-форма геномов многих ретровирусов была клонирована в *E. coli*-системах и поэтому доступна для дальнейших манипуляций. Эти геномы встраивают в рекомбинантные векторы *in vitro* и полученные формы клонируют и реплицируют в *E. coli*. Типичный вектор содержит: 1) фрагменты плазмиды pBR322, необходимые для репликации в клетках *E. coli* и отбора; 2) LTR-последовательности (длинные прямые концевые повторы) вирусного генома, включая сегменты R, U5 и U3; 3) последовательности P, Pu, ψ и S; 4) селективный маркер, например ген *neo E. coli*; 5) по крайней мере один уникальный сайт для рестриктирующей эндонуклеазы, по которому можно встраивать дополнительные сегменты ДНК без какого-либо нарушения существенной части вирусного генома (рис. 5.41).

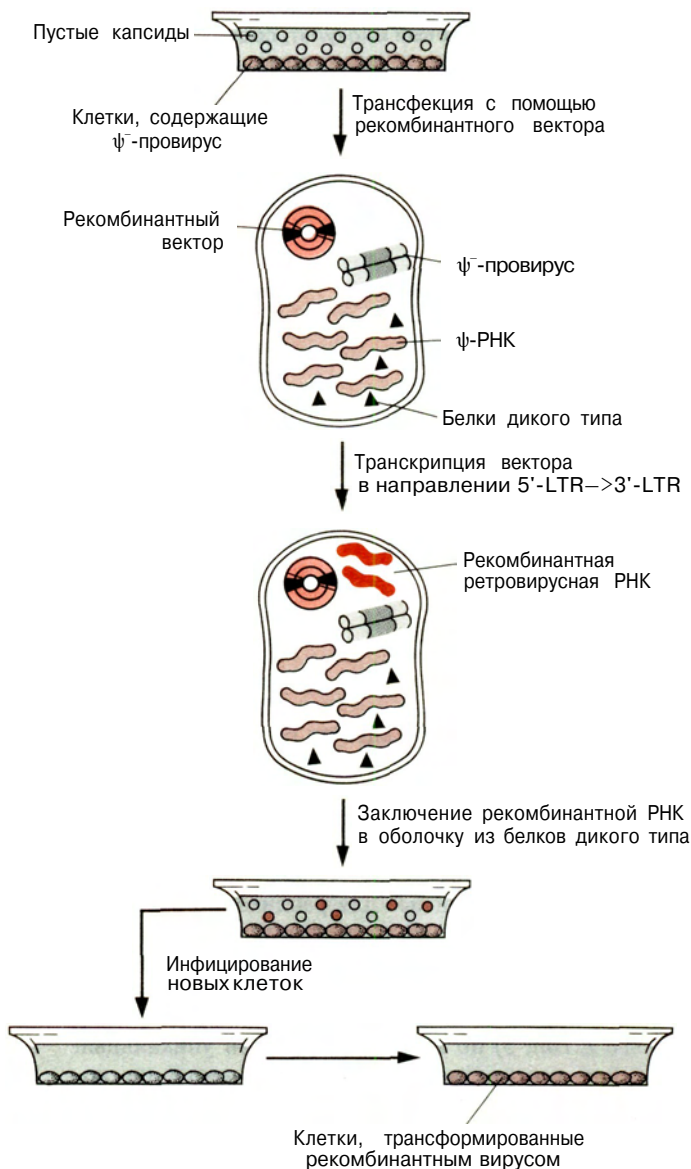
Векторы, сконструированные путем клонирования в клетках *E. coli*, трансфицируются непосредственно в соответствующие клетки животных. На 5'-конце LTR находится элемент, ответственный за регуляцию транскрипции с помощью РНК-полимеразы II, на 3'-конце — сигнал терминации транскрипции. Соответственно синтезируемая РНК по существу представляет собой вирусный геном (рис. 5.41), за исключением одного или нескольких невирусных сегментов, находящихся между 5'- и 3'-концами. Следует отметить, что *E. coli*-последовательности вектора не транскрибируются и не присутствуют в рекомбинантной РНК ретровирусного генома.

Рекомбинантная ретровирусная РНК может быть упакована в вирионы, если в среде присутствуют необходимые вирусные белки. Такие белки

образуются, если клетки хозяина содержат полностью функциональный провирус, который обеспечивает синтез вирусных белков, находясь в *транс*-положении (рис. 5.42). В этом случае вирусное потомство представляет собой смешанную популяцию рекомбинантных вирионов и вирионов вируса-помощника дикого типа. Такой препарат можно использовать для повторного инфицирования новых клеток, причем необходимый вирус-помощник поставляется одновременно с рекомбинантным вирусом. Можно получить и чистые препараты рекомбинантных вирионов, если провирус в исходных клетках дефектен в отношении упаковки. Такие хозяйские клетки получают при инфицировании их

делеционным мутантом, утратившим  $\psi$ -последовательности (рис. 5.43). Несмотря на неспособность вирусной РНК, образуемой провирусом  $\psi^-$ , упаковываться, рекомбинантные геномы сохраняют способность к упаковке, инфицированию новых клеток и установлению провирусного состояния, поскольку их  $\psi$ -последовательность остается интактной. Вирионы содержат белки, необходимые для осуществления обратной транскрипции и включения в клеточный геном в процессе следующего инфекционного цикла.

*Использование ретровирусных векторов.* Трансдуцирующие вирусные векторы имеют одно преимущество, связанное с тем, что инфицирование



**РИС. 5.43.**

Превращение дуплексной рекомбинантной ДНК ретровирусного вектора в инкапсидированный РНК-геном в клетках хозяина, несущих провирус  $\psi^-$ . Клетки в чашке, изображенной в верхней части рисунка, содержат интегрированный провирус с делегированной  $\psi$ -последовательностью. Несмотря на синтез вирусной РНК и белков, инкапсидации РНК не происходит. В результате в среде накапливаются только пустые неинфекционные капсиды. Вскоре после трансфекции с помощью двухцепочечного рекомбинантного ретровирусного вектора в результате транскрипции, идущей в направлении 5'-LTR → 3'-LTR, образуется рекомбинантная РНК. Эта РНК заключается в оболочку из вирусных белков дикого типа, и в среде накапливаются рекомбинантные вирионы. Такой инокулят содержит только пустые капсиды и рекомбинантные вирионы. При инфицировании этим инокулятом новых клеток происходит репликация рекомбинантного генома с помощью вирусной обратной транскриптазы. ДНК интегрирует как провирус и образовавшиеся трансформированные клетки экспрессируют гены, содержащиеся в рекомбинантном провирусе.

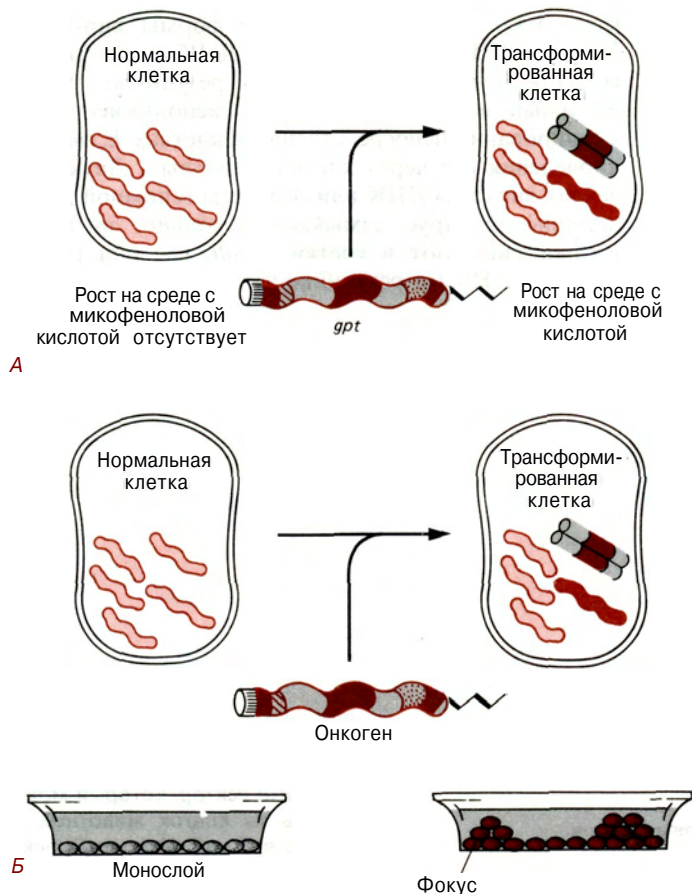


РИС. 5.44.

Отбор клеток, трансформированных ретровирусными векторами. **А.** Рекомбинантный геном содержит ген *gpt E.coli*. После проникновения вириона в клетку происходит образование дуплексной ДНК, катализируемое обратной транскриптазой, присутствующей в рекомбинантном вирионе. Эта ДНК встраивается в клеточный геном, образуя провирус. Трансформированные клетки можно отобрать благодаря экспрессии гена *gpt*, снимающей ингибирующий эффект микрофеноловой кислоты. Если в рекомбинантную молекулу включен какой-то неселективный ген, он тоже экспрессируется. **Б.** Ретровирусный геном содержит онкоген. Трансформированные клетки можно идентифицировать по их способности образовывать фокусы, тогда как нетрансформированные клетки образуют лишь монослой. В этом случае в трансформированных клетках также экспрессируется неселективный ген, включенный в трансформирующий геном.

значительно эффективнее, чем трансфекция. После того как получен препарат вновь сконструированного вектора путем трансфекции клеток, обладающих функциями помощника, все последующие операции можно осуществлять с помощью инфицирования. Кроме того, если целью эксперимента является трансформация клетки, то помощник не требуется. Рекомбинантный геном образует провирус, и клетки стабильно трансформируются. Трансформированные колонии можно отбирать несколькими способами (рис. 5.44). В тех случаях, когда вектор содержит функциональный маркер, подобный генам *neo* или *gpt E.coli*, трансформируемые клетки можно отобрать, добавить в среду G-418 или микрофеноловую кислоту соответственно. Отметим, что последовательности в 5'-LTR, ответственные за регуляцию транскрипции, способствуют синтезу функциональной мРНК, направляемому генами *E.coli*, в клетках животных. Если ген *E.coli* (или любой другой ген) встроен в участок гена *env* (как на рис. 5.41), то S-сайты должны оставаться интактными, чтобы было обеспечено образование функциональной мРНК. Фокусы, образуемые трансформированными клетками, можно идентифициро-

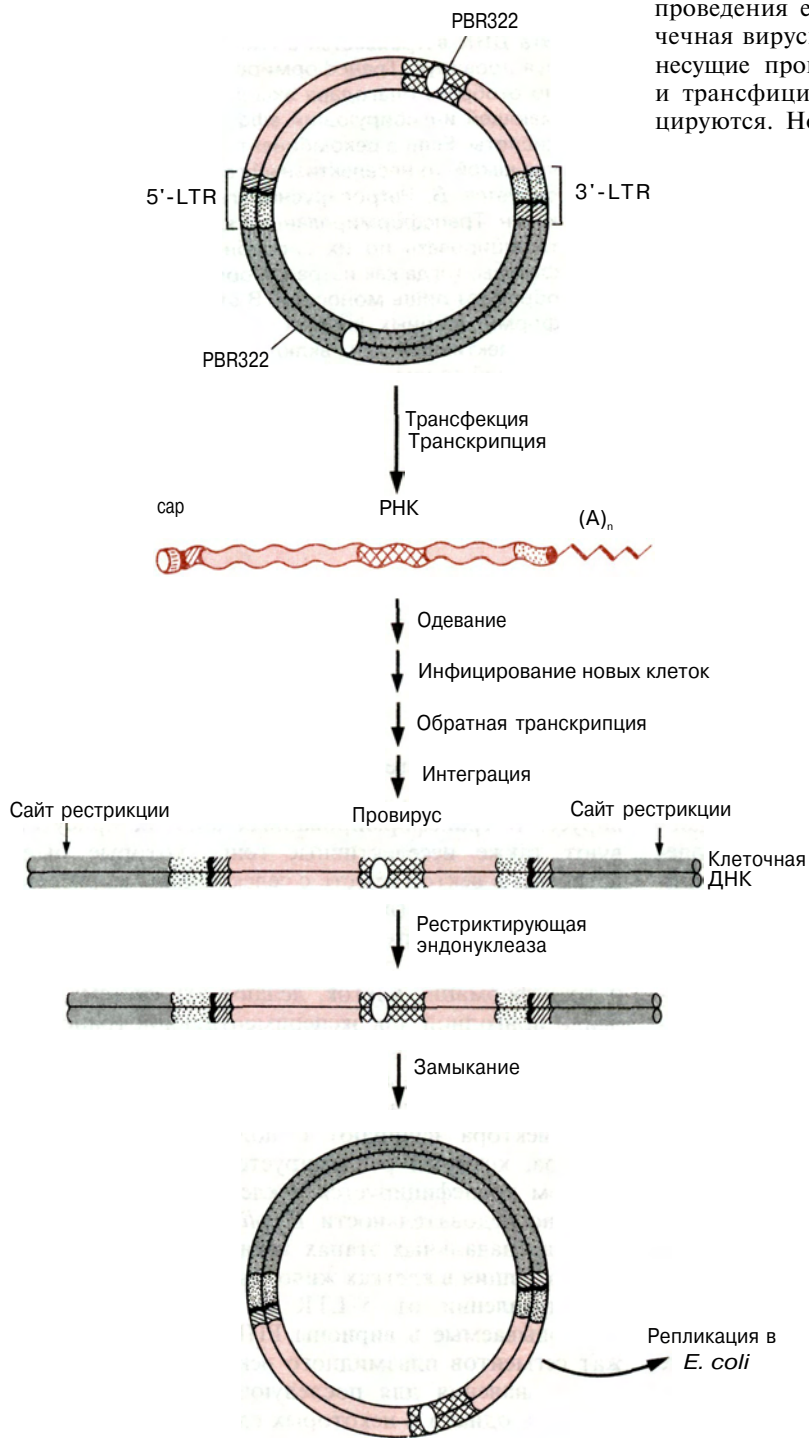
вать по их характерной морфологии, если вектор сохранил онкоген, присутствовавший в исходном вирусе. В трансформированных клетках присутствуют также неселективные гены, которые были встроены в вектор вместе с селективным маркером. Описанные свойства трансдуцирующих рекомбинантных ретровирусов, в особенности эффективность, с которой осуществляются инфицирование и трансформация клеток, делают эту систему особенно пригодной для экспериментальной трансформации клеток очень ранних зародышей и даже всего организма животных.

Конструирование рекомбинантного ретровирусного вектора начинают с получения челночного вектора, который реплицируется в клетках *E.coli*, а затем трансфицируется в клетки животных. Однако последовательности *E.coli*-плазмид утрачиваются на начальных этапах трансфекции, поскольку транскрипция в клетках животных протекает только в направлении от 5'-LTR к 3'-LTR (рис. 5.41). Упаковываемые в вирионы РНК-геномы не содержат сегментов плазмидного вектора. Часто это не имеет значения для последующего экспериментирования, однако в некоторых случаях возможность



введения рекомбинантных молекул снова в клетки *E. coli* оказывается существенной. Чтобы решить эту проблему, сегмент *E. coli*-плазмиды, например плазмиды pBR322, встраивают между последовательностями LTR при первоначальном конструировании

(рис. 5.45). В этом случае все формы вирусного генома—РНК, двухцепочечная ДНК и провирус — сохраняют сегмент pBR322. В результате ретровирусный вектор можно будет использовать для клонирования непосредственно в клетках *E. coli*, без проведения его через клетки животных. Двухцепочечная вирусная ДНК или сегменты геномной ДНК, несущие провирус, замыкают в кольцо, лигируют и трансфицируют в клетки *E. coli*, где они реплицируются. Новые рекомбинантные молекулы часто



**Рис. 5.45.**

Ретровирусный вектор, который можно выделить из клеток животных и клонировать в *E. coli*. Исходный вектор по существу аналогичен вектору, представленному на рис. 5.41, за исключением того, что между 5'-LTR и 3'-LTR встроены дополнительные pBR322-последовательности, ответственные за репликацию вектора в *E. coli*. После трансфекции происходит транскрипция вектора, идущая в направлении 5'-LTR→3'-LTR, с образованием вирусных РНК-геномов, включающих сегмент pBR322. Геномы упаковываются в оболочку из вирусных белков, синтез которых детерминируется генами, находящимися в *транс*-положении. В новых клетках, инфицированных этими вирусами, образуются дуплексная ДНК и провирус, еще сохранивший сегмент pBR322. Если ДНК, выделенную из трансформированных клеток, разрезать с помощью рестриктирующих эндонуклеаз, то образуются фрагменты, содержащие провирус, и множество других фрагментов. При замыкании этих фрагментов в кольцо и трансфицировании ими клеток *E. coli* к репликации оказываются способными лишь молекулы, содержащие провирус.

содержат сегменты геномной ДНК, которые фланкировали провирусную ДНК в геноме хозяина. Поскольку внедрение провируса в функциональные гены или вблизи них часто вызывает мутагенный или онкогенный эффект (либо и тот и другой), анализ таких фланкирующих последовательностей можно использовать для выделения гена, определяющего данный фенотип.

## 5.8. ЭУКАРИОТИЧЕСКИЕ СИСТЕМЫ ХОЗЯИН-ВЕКТОР:РАСТЕНИЯ

### а. Общие положения

Растительные клетки не содержат собственных плазмид. В этом случае в качестве основы для конструирования трансформирующих векторных систем в принципе могут использоваться независимо реплицирующиеся геномы различных растительных вирусов. Такие системы были созданы на основе генома вируса мозаики цветной капусты. Однако все наиболее совершенные системы векторов растений получены на основе плазмид из семейства необычных бактериальных плазмид, носящих название *pTi*. Эти плазмиды образуют природную систему трансформации, с помощью которой осуществляется перенос сегментов плазмидной ДНК в геномы разнообразных двудольных растений.

В отсутствие специфического вектора прямая трансформация по крайней мере некоторых растительных клеток осуществляется с помощью трансфекции фрагментами чужеродной ДНК, добавленными в культуральную среду. Как и клетки животных, клетки растений поглощают ДНК и она интегрируется с клеточным геномом, в результате чего образуются стабильно трансформированные клетки. Однако эффективность прямой трансформации весьма низка. Отобрать трансформанты, появляющиеся с частотой примерно 1 на  $10^6$  обработанных клеток, можно лишь с помощью высокочувствительных методов. Для повышения эффективности прямого введения ДНК в клетки растений можно использовать метод электропорации. В этом случае трансформированными становятся до 1% клеток, а кроме того, такой способ можно применять в случае как однодольных, так и двудольных растений. С помощью рекомбинантных ДНК были трансформированы различные растительные клетки, в том числе клетки табака, петунии, томатов и подсолнечника. Для трансформации часто используют протопласты, полученные путем разрушения жестких клеточных стенок с помощью целлюлазы, в результате чего клетки становятся проницаемыми для ДНК. При переносе трансфицированных протопластов в соответствующую среду клеточная стенка восстанавливается.

### б. Плазида *pTi*-А, индуцирующая образование опухолей

Семейство конъюгативных плазмид *pTi*-А обнаружено в грамотрицательных бактериях *Agrobacterium tumefaciens*. Замкнутые кольцевые двухцепочечные плазмиды *pTi* имеют размер от 150 до 250 т.п.н. и содержат разнообразные уникальные гены (рис. 5.46 и табл. 5.5). Бактерии *A. tumefaciens* попадают в пораженные ткани растений, а затем плазмиды проникают в растительные клетки и рекомбинируют с клеточной ДНК. В растительном геноме имеется несколько специфических сайтов интеграции. Трансформированные клетки образуют опухоли, получившие название **корончатых галлов** (рис. 5.47). Штаммы *A. tumefaciens*, утратившие плазмиду *pTi*, не вызывают образования корончатых галлов. Клетки, выделенные из корончатых галлов, в отличие от нормальных клеток растений быстро растут в лабораторной питательной среде даже в отсутствие гормонов роста растений. Если их привить на здоровое растение, то образуется новый галл. Иногда экспрессия опухолевого фенотипа подавляется и из клеток галла регенерируются нормальные ткани растений. В лабораторных условиях трансформацию осуществляют путем инокуляции *A. tumefaciens* в пораненные растительные ткани (в ткани стебля, корней или кусочков листьев) или путем совместного культивирования бактерий

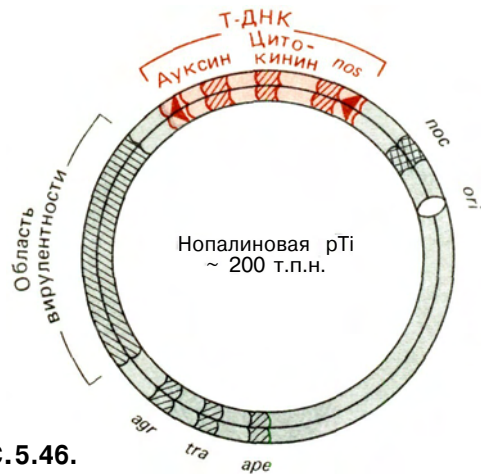


РИС. 5.46.

Схематическое изображение генетической карты типичной нопапиновой плазмиды *pTi*. Участок, выделенный цветом, - сегмент Т-ДНК, встроенный в ДНК клеток корончатого галла. Некоторые из представленных на схеме генов охарактеризованы в табл. 5.5 (ген *agr* ответствен за чувствительность к агроцину-84). Стрелками указаны последовательности из 25 п.н. на правой и левой границах Т-ДНК. Правосторонняя последовательность необходима для интеграции Т-ДНК в растительный геном. Показано положение области *ori* (для репликации в *A. tumefaciens*).

Таблица 5.5. Некоторые гены Ti-плазмид

Октопиновые плазмиды	Нопалиновые плазмиды
Шесть генов вирулентности ( <i>vir</i> )	Шесть генов вирулентности ( <i>vir</i> )
Ген синтеза октопина ( <i>ocs</i> )	Ген синтеза нопалина ( <i>nos</i> )
Ген катаболизма октопина ( <i>occ</i> )	Ген катаболизма нопалина ( <i>noc</i> )
Ген исключения фага AP1 ( <i>ape</i> )	Ген, необходимый для конъюгации ( <i>tra</i> )
Ген биосинтеза ауксина ( <i>iaa</i> )	Ген биосинтеза ауксина ( <i>iaa</i> )

с протопластами клеток растений. При совместном культивировании трансформируется более 10% протопластов табака.

**Геномы *pTi*.** Обычно в ДНК клеток корончатого галла бывает встроена лишь небольшая (от 13 до 25 т.п.н.) часть генома *pTi*, однако эта часть всегда включает определенный сегмент, получивший название **Т-ДНК** (рис. 5.46). В клетках корончатых галлов с интегрированной Т-ДНК синтезируются РНК и белки, которые ответственны за опухолевый фенотип этих клеток. Кроме того, образуются ферменты, катализирующие синтез аргининовых производных, известных под названием **опины**. В нетрансформированных клетках опины не синтезируются. *pTi*-плазмиды классифицируют в соответствии с теми опинами, синтез которых они индуцируют. Например, октопиновые *pTi*-плазмиды индуцируют синтез октопина в трансформированных клетках растений, а нопалиновые *pTi*-плазми-

ды—синтез нопалина (табл. 5.5). Однако каждая *pTi*-плазида имеет также гены, экспрессируемые в *A. tumefaciens*; они находятся вне Т-ДНК. Некоторые из них, например гены вирулентности (*vir*), нужны для переноса Т-ДНК в клетки растений и, следовательно, существенны для образования опухоли. Другие связаны с катаболизмом тех же опинов, чей синтез кодируется Т-ДНК. Экспрессия генов метаболизма опинов индуцируется в клетках бактерий самими опинами. Таким образом, *pTi*-плазмиды обеспечивают клетки бактерий источником метаболита (путем трансформации растительных клеток) и механизмом использования этого метаболита в качестве источника углерода, азота и энергии. Секрета опинов трансформированными клетками растений создает для *A. tumefaciens* очень благоприятные условия, поскольку другие распространенные почвенные бактерии не способны метаболизировать эти необычные соединения.

Система *pTi*—*A. tumefaciens* дает замечательный пример включения прокариотической плазмидной ДНК в эукариотический геном в природных условиях. В действительности сама плазида является, по-видимому, природной химерой. Она содержит два набора генов, один из которых активен в клетках бактерий, другой—в клетках растений. Регуляция работы генов, входящих в состав Т-ДНК, осуществляется с помощью элементов, оперирующих в клетках растений, тогда как гены в остальной части плазмиды находятся под контролем бактериальных промоторов. Следует отметить, что для образования опухоли требуются гены обеих групп.

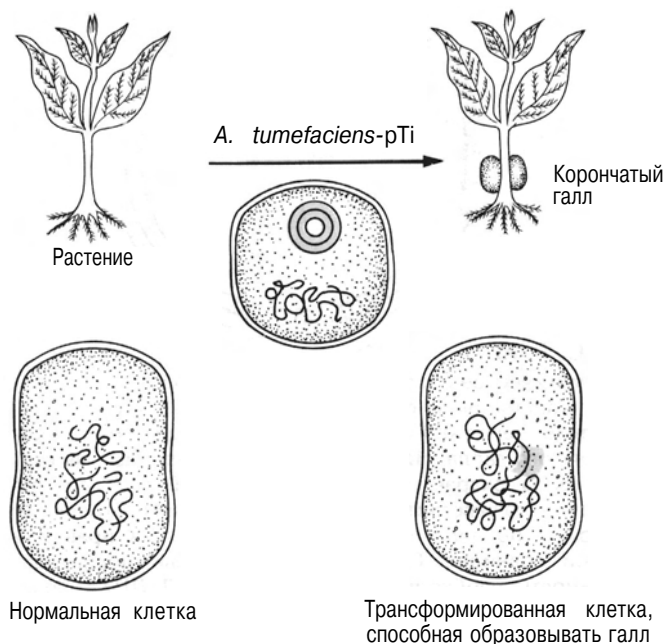


РИС. 5.47.

*pTi*-Индукцируемая трансформация и образование корончатого галла.

Гены *vir*, экспрессируемые в клетках бактерий, существенны для переноса Т-ДНК в растительный геном. Другие гены Т-ДНК экспрессируются в растениях и ответственны за опухолевый фенотип клеток корончатого галла. Эти онкогены кодируют белки, катализирующие образование гормонов роста растений—ауксина и цитокинина.

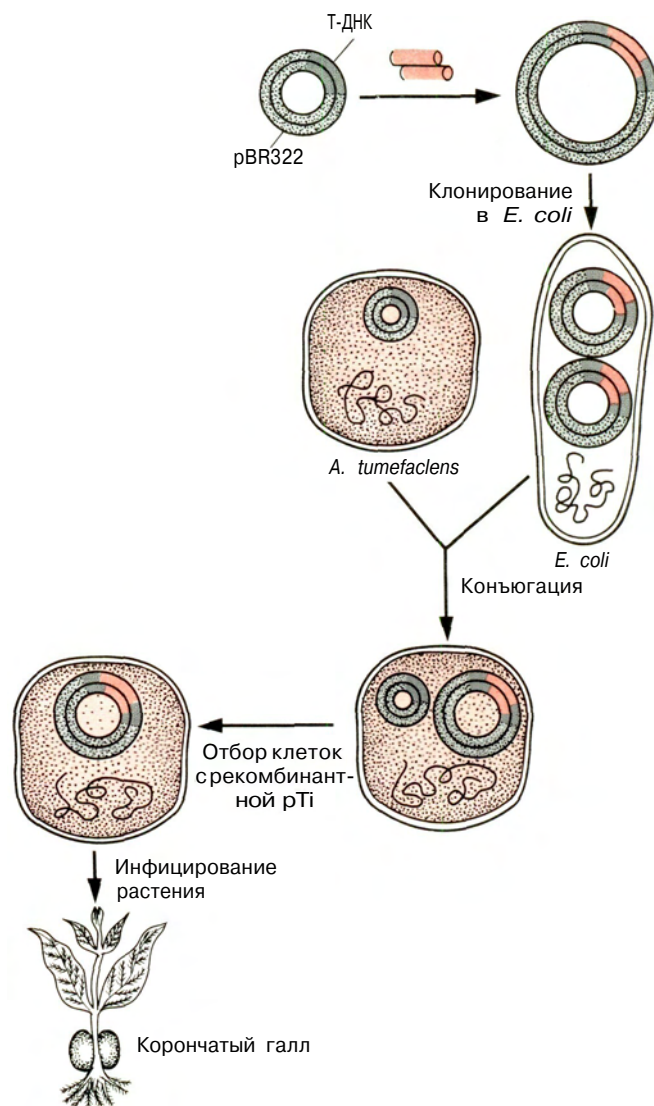
**Инсерционная трансформация с участием рТi.** Механизм инсерционной трансформации плазмидой рТi отличается от механизма трансформации другими эукариотическими системами, описанных в данной главе, но имеет некоторое сходство с бактериальной конъюгацией. В хромосоме *A. tumefaciens* закодирована информация о функциях, необходимых для прикрепления бактерий к клеткам растений. Плазмида рТi кодирует *цис*- и *транс*-функции, нужные для интеграции. Для осуществления интеграции Т-ДНК на правом ее конце должен присутствовать сегмент из 25 п.н.; переносятся только те последовательности, которые расположены слева от этой области. Подобная же последовательность встречается на левом конце Т-ДНК. По-видимому, она не требуется для интеграции, но помогает обозначить конец интегрируемой Т-ДНК. *Транс*-функции обеспечиваются продуктами *vir*-генов (рис. 5.46). Среди них имеется сайт-специфическая эндонуклеаза, которая разрезает нижнюю цепь Т-ДНК в пределах обеих пограничных последовательностей. 3'-конец ДНК рТi, ближайший к правостороннему разрезу, служит праймером для синтеза ДНК, замещающей нижнюю цепь Т-ДНК. Свободная цепь Т-ДНК переносится в растительные клетки, начиная с 5'-конца правой пограничной последовательности к 3'-концу. Механизм ее включения в случайные сайты ДНК клеток растений остается неясным.

### в. Конструирование векторов рекомбинантных ДНК с использованием рТi

Использование плазмид рТi в качестве вектора рекомбинантных ДНК зависит от того, можно ли встроить нужный фрагмент в область Т-ДНК, провести отбор и поддерживать рекомбинантные молекулы в клетках *A. tumefaciens*, с тем чтобы ввести их затем в клетки растений. Из-за отсутствия уникальных сайтов рестрикции в большой молекуле рТi прямое включение в нее чужеродной ДНК невозможно, но существует обычный способ встраивания генов в Т-ДНК. Сама Т-ДНК была вначале клонирована в *E. coli* (с использованием плазмиды рBR322). Нужные сегменты ДНК встраивают в подходящие сайты рестрикции в Т-ДНК выделенной плазмиды и полученные рекомбинанты повторно клонируют в *E. coli* (рис. 5.48). Если такие плазмиды ввести путем конъюгации в клетки *A. tumefaciens*,

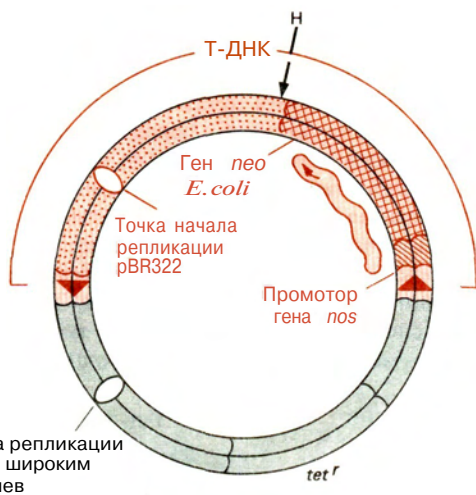
которые уже несут плазмиду рТi дикого типа, то Т-ДНК с встроенным фрагментом перейдет в результате гомологичной рекомбинации в область Т-ДНК интактной плазмиды рТi. Отбор клеток *A. tumefaciens*, несущих рекомбинантные плазмиды рТi, можно осуществить, если в чужеродную ДНК включить маркирующий ген типа гена *neo E. coli*. Репликация рекомбинантной плазмиды рТi в клетках *A. tumefaciens* и последующее инфицирование растений ведут к образованию корончатых галлов, несущих сегмент рекомбинантной Т-ДНК.

Используя те преимущества, которые дает различие между *цис*- и *транс*-функциями плазмиды,



**РИС. 5.48.**

Конструирование и использование промежуточного рТ-вектора.



**РИС. 5.49.**

Челночный вектор «Т-ДНК-*E. coli*». Участок, выделенный серым цветом, происходит из R-плазмиды *E. coli* (разд. 5.2), способной реплицироваться в широком круге грамотрицательных бактерий, включая *A. tumefaciens*. Он содержит точку *ori*, гены, необходимые для репликации и конъюгационного переноса путем мобилизации, и ген, ответственный за резистентность к тетрациклину (*tet<sup>r</sup>*) и позволяющий проводить отбор. Стрелками указаны пограничные последовательности Т-ДНК, которые функционируют как *цис*-последовательности при интеграции с ДНК растительной клетки. Т-ДНК содержит точку *ori* репликации рВR322 и ген *neo E. coli*, который способен экспрессироваться в растительных клетках, поскольку он сцеплен с элементом регуляции транскрипции, обычно инициирующим транскрипцию гена *nos* плазмиды рТi (направление транскрипции показано стрелкой). Встраивание дополнительных сегментов в Т-ДНК может происходить по единственному сайту *Hind* III.

сконструировали относительно простую векторную рТi-систему. Основной вектор—челночная плазида, способная реплицироваться как в *E. coli*, так и в широком круге других бактерий (включая *A. tumefaciens*), содержащая оба пограничных сегмента Т-ДНК из 25 п.н. (рис. 5.49). Путем клонирования в *E. coli* уже получены производные векторы, содержащие различные вставки слева от правостороннего пограничного сегмента из 25 п.н. Если перенести такие молекулы конъюгативным путем в клетки *A. tumefaciens*, несущие плазмиду рТi дикого типа, то они реплицируются, а продукты генов, необходимые для встраивания рекомбинантной плазмиды в ДНК растений, будут поставлены генами, находящимися в *транс*-положении. Растительные клетки, инфицируемые такими клетками *A. tumefaciens*, интегрируют либо Т-ДНК дикого типа, либо рекомбинантную Т-ДНК, либо и ту и другую. Функциональные маркерные гены, включенные в рекомбинантную Т-ДНК, экспрессируются в интегрированном состоянии в ДНК растений, если они оказываются под контролем элементов генома эукариотической клетки, контролирующих синтез РНК. Для этого, например, можно подключить маркерный ген к 3'-концу регуляторного участка гена *nos* в рекомбинантной Т-ДНК. Если вместо помощника дикого типа используется плазида-помощник рТi, из которой делетированы либо *цис*-активная последовательность длиной 25 п.н., либо весь сегмент Т-ДНК, то способностью к переносу обладает только рекомбинантная Т-ДНК.

Многие клетки растений **тотипотентны** (т.е. способны размножаться, дифференцироваться и образовывать целые плодоносящие растения). Векторные системы, сконструированные на основе плазмид рТi, в принципе можно использовать для получения генотипически измененных растений из трансформированных клеток, и такие опыты уже поставлены. Кроме того, выполненные исследования уже позволили получить данные об экспрессии генов и процессах дифференцировки в нескольких видах растений.

# СРЕДСТВА: КОНСТРУИРОВАНИЕ, КЛОНИРОВАНИЕ И ОТБОР РЕКОМБИНАНТНЫХ МОЛЕКУЛ ДНК

В этой главе мы рассмотрим процесс получения клонированных ДНК. На первом этапе готовят фрагменты ДНК (вставки), пригодные для последующего лигирования с векторной молекулой. Следующий этап состоит в самом лигировании. Эти процессы осуществляются *in vitro*. Далее рекомбинантные молекулы вводят в клетки, где они амплифицируются путем репликации. Затем проводят клонирование, отбор и дальнейшую амплификацию.

## 6.1. ВСТАВКИ

### а. Общие положения

При конструировании рекомбинантных молекул обычно используют сложные смеси потенциальных вставок, и в результате образуется целый набор клонов, из которого путем скрининга и отбора получают нужные рекомбинантные молекулы. Скрининг и отбор можно значительно упростить, если обогатить исходную смесь нужным сегментом; в этом случае для выявления искомого рекомбинанта приходится проверять значительно меньше рекомбинантных клонов. С одной стороны, для конструирования рекомбинантных ДНК можно использовать очищенные фрагменты ДНК, а с другой – само молекулярное клонирование является наиболее простым и эффективным способом очистки фрагментов. Клонирование представляет собой также наиболее эффективный способ получения большинства фрагментов геномной ДНК в значительных количествах.

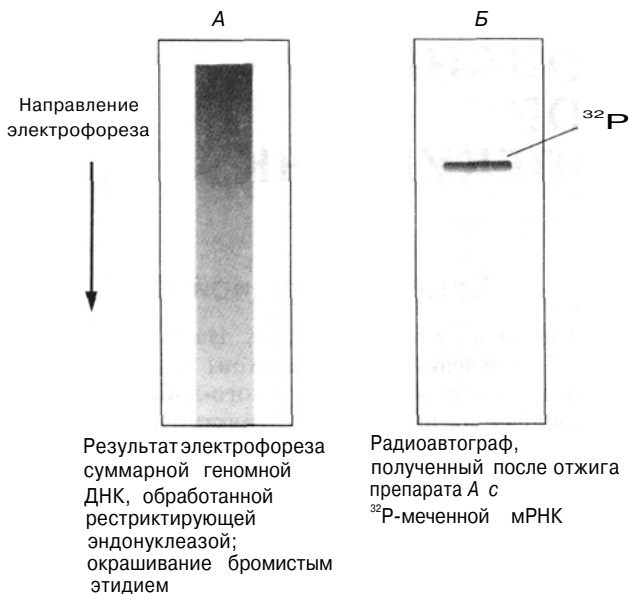
Существует три источника вставок для клонирования: 1) геномная ДНК, фрагментированная либо с помощью рестриктирующих эндонуклеаз, либо физическими методами, например с помощью ультразвука; 2) синтетические фрагменты ДНК, полученные химическим или ферментативным методом либо путем объединения этих двух методов; 3) сегменты ДНК (кДНК), полученные с помощью ферментативного копирования РНК-матрицы *in vitro*. При расщеплении ДНК эндонуклеазами часто образуются фрагменты, способные к непосредственному лигированию с подходящими липкими или тупыми концами вектора. В других случаях соответствующие концы присоединяют к фрагментам перед лигированием.

### б. Вставки геномной ДНК

*Эндонуклеазное расщепление.* Наиболее прямой способ получения вставок состоит в расщеплении суммарной геномной ДНК какого-либо организма или вируса с помощью рестриктирующей эндонуклеазы. Метод отличается высокой воспроизводимостью: специфический фермент разрезает данную ДНК с образованием уникального набора фрагментов. Если при эндонуклеазном расщеплении образуются липкие концы, соответствующие концам векторной молекулы, то клонирование осуществляется сразу или после обогащения популяции фрагментами для получения нужной вставки.

*Обогащение смеси нужными фрагментами.* Для эффективного разделения сложных смесей фрагментов используют два метода: **электрофорез** в полужидкой среде (разд. 4.2.г; рис. 4.10) и **жидкостную хроматографию высокого разрешения с обращенной фазой**. Однако ни один из этих методов не позволяет получить фрагменты в чистом виде. Обычно препараты содержат примеси других фрагментов, имеющих близкую длину или обладающих такой же способностью к элюции. Тем не менее можно получить значительное обогащение смеси в отношении нужного фрагмента, что облегчает его выявление в большой коллекции клонов.

При электрофорезе и хроматографии разделение фрагментов ДНК основывается на их различии по размеру и нуклеотидному составу. Если размер генома невелик, то образуется относительно небольшое число хорошо разделяющихся фрагментов (см. рис. 4.10); их легко выделить, собрав нужные фракции элюата с хроматографической колонки или вырезав кусочки агарозного геля. Однако если расщепляется крупный геном, то хорошо отделяются лишь некоторые фрагменты. Чаще получается непрерывный набор фрагментов всевозможных размеров (рис. 6.1, А). Чтобы понять, в чем тут дело, проведем простой расчет. Типичный гаплоидный геном млекопитающих содержит примерно  $3 \cdot 10^9$  п.н. Грубую оценку числа фрагментов, образующихся при исчерпывающем расщеплении эндонуклеазой, можно получить, разделив размер генома на предполагаемое среднее расстояние (в парах оснований) между двумя соседними сайтами для данной рестриктирующей эндонуклеазы (в

**РИС. 6.1.**

А. Электрофоретическое разделение фрагментов, полученных в результате расщепления эукариотической геномной ДНК с помощью рестриктирующей эндонуклеазы. Для визуализации фрагментов проводят окрашивание бромистым этидием. Слой геля толщиной от 0,5 мм до нескольких миллиметров наносят на твердую (например, стеклянную) подложку. Обычно гели готовят из полиакриламида или агарозы в забуференном солевом растворе. Б. Препарат ДНК переносят с геля на нитроцеллюлозный фильтр по методу, описанному в подписи к рис. 6.2, и гибридизуют с очищенной  $^{32}\text{P}$ -меченой мРНК. На рисунке представлен радиоавтограф, полученный при наложении рентгеновской пленки на нитроцеллюлозный фильтр.

предположении случайного распределения четырех оснований). Для фермента, узнающего участок из шести пар оснований (например, *Eco* RI или *Hind* III), среднее расстояние между сайтами будет равно  $4^6$  (4096) п.н. Если размер сайта узнавания равен четырем парам нуклеотидов, то он будет встречаться примерно один раз на каждые  $4^4$  (256) п.н. При расщеплении ДНК млекопитающих ферментами, которые разрезают молекулу по сайтам узнавания длиной шесть пар нуклеотидов, образуется примерно  $7 \cdot 10^5$  уникальных фрагментов, а ферментом, сайт узнавания которого составляет четыре пары нуклеотидов, —  $12 \cdot 10^6$  фрагментов. В результате получается непрерывное распределение фрагментов по длинам. При этом некоторые фрагменты вообще невозможно выявить, поскольку они представлены в очень малом количестве.

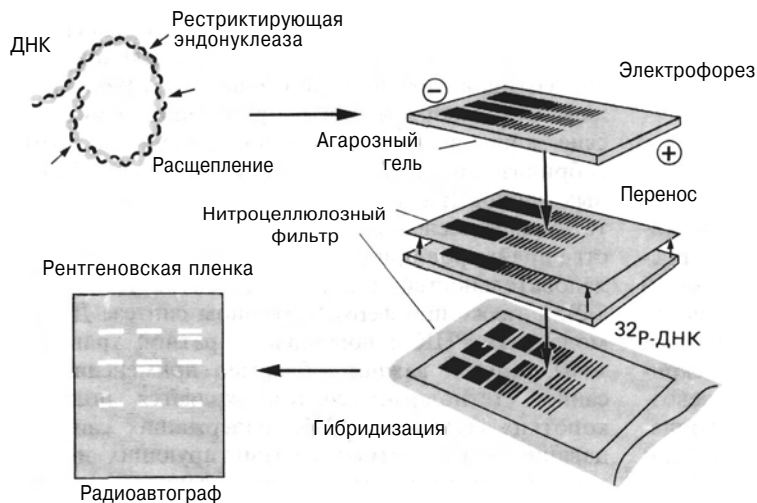
*Идентификация специфических фрагментов.* Проблемы, возникающие при идентификации фрагментов, иллюстрирует рис. 6.1, А. При расщеплении

примерно 50 мкг геномной ДНК и последующем электрофорезе препарата масса сегмента ДНК длиной 1000 п.н., встречающегося в геноме лишь один раз, составляет всего  $17 \cdot 10^{-6}$  мкг, или 17 пг. Как идентифицировать этот специфический фрагмент? Если имеется гомологичная ДНК или препарат комплементарной РНК, которые можно использовать в качестве зонда при гибридизации, то интересующий нас фрагмент можно обнаружить при отжиге зонда с фрагментами после их денатурации. Так, в качестве зонда можно использовать очищенную мРНК, соответствующую гену, который мы хотим клонировать. Иногда в роли зонда выступает гомологичный ген, клонированный в другом организме и имеющий достаточно близкую структуру. Чтобы выявить гибрид, нужно использовать меченый зонд. Чаще всего его метят радиоактивным изотопом.

На рис. 6.1, Б представлен результат отжига в геле  $^{32}\text{P}$ -меченой мРНК, используемой в качестве зонда, с денатурированными фрагментами (см. рис. 6.1, А). Геномный фрагмент, комплементарный данной мРНК, образует с ней дуплекс, и его выявляют по локализации в геле  $^{32}\text{P}$ -РНК. Определив примерный размер интересующего нас фрагмента, можно элюировать материал соответствующей области другого геля, не прошедшего гибридизацию, и использовать этот материал для клонирования. Аналогичным образом можно тестировать фракции после хроматографического разделения фрагментов на колонке.

*Блоттинг.* Почти во всех случаях, подобных описанным выше, ДНК перед гибридизацией переносят с геля на более подходящую подложку. Этот метод (**блоттинг**) широко используется, например, для идентификации специфических РНК и белков после их электрофоретического разделения (см. разд. 7.7 и 6.4 соответственно). Блоттинг ДНК назван по имени его изобретателя блоттингом по Саузерну (или саузерн-блоттингом); соответствующие методики для РНК и белков получили название нозерн-блоттинга и вестерн-блоттинга.

Гибридизация меченого зонда с фрагментами ДНК, находящимися в геле, происходит очень неэффективно. Поэтому перед отжигом фрагменты переносят из геля на более подходящую твердую подложку, обычно на нитроцеллюлозный фильтр или нейлон. Сначала гель обрабатывают щелочью, чтобы произошла денатурация ДНК. Затем на него накладывают твердую подложку и пропускают через гель буферный раствор (рис. 6.2). В результате фрагменты ДНК переходят на подложку с сохранением их взаимного расположения. Далее инкубируют подложку в растворе, содержащем  $^{32}\text{P}$ -зонд, при такой ионной силе и температуре, которые обеспечивают образование водородных связей меж-

**РИС. 6.2.**

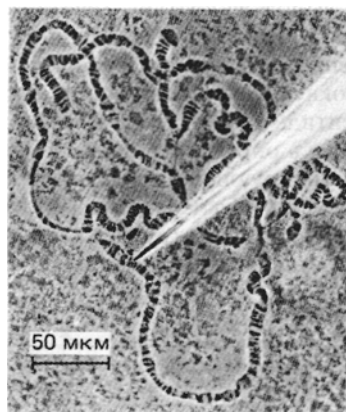
Блоттинг ДНК. На пластину геля накладывают нитроцеллюлозный фильтр того же размера, что и гель, и пропускают буферный раствор. Фрагменты ДНК переносятся на нитроцеллюлозный фильтр и фиксируются на нем. Аналогичные процедуры используются и при переносе РНК и белка с геля на твердую подложку.

ду зондом и комплементарным фрагментом ДНК. Чувствительность метода зависит от удельной радиоактивности зонда и позволяет регистрировать фрагменты, когда их количество измеряется пикограммами. Например,  $^{32}\text{P}$ -меченные зонды, полученные с помощью ник-трансляции (разд. 4.6.6), имеют удельную радиоактивность более 100 имп./мин на 1 пг, что достаточно для визуализации радиоавтографов.

**Случайная фрагментация геномной ДНК.** Длинные молекулы ДНК легко ломаются даже при незначительных гидродинамических напряжениях, и их можно подвергнуть случайной фрагментации с помощью ультразвука, путем быстрого перемешивания раствора или пропуска его через небольшое отверстие. Для получения случайного набора перекрывающихся фрагментов можно использовать и рестриктирующие эндонуклеазы, если проводить опыт в таких условиях, когда расщепление происходит лишь в небольшом числе имеющихся сайтов. Средний размер образующихся фрагментов зависит от величины прикладываемого усилия и от концентрации эндонуклеазы. Обычно ДНК выделяют из большой популяции клеток, поэтому исходные ее препараты всегда представлены большим числом идентичных геномов. Каждый сегмент ДНК присутствует в конечном наборе во фрагментах разных размеров, что не позволяет очистить его перед клонированием. Тем не менее случайные наборы фрагментов имеют преимущество перед наборами, получаемыми в результате полного гидролиза рестриктирующей эндонуклеазой, поскольку вероятность нахождения хотя бы одной копии нужного сегмента целиком в одном фрагменте у них значительно выше.

**Вырезание определенных сегментов хромосом.** Если положение данного гена в хромосоме известно

и хромосома достаточно велика, чтобы проводить с ней различные манипуляции, можно выщипать ее часть, которая содержит нужный ген. Этим требованиям удовлетворяют политенные хромосомы слюнных желез *Drosophila*. Каждая такая хромосома содержит более 1000 копий ДНК, расположенных параллельно друг другу, так что небольшой сегмент может включать 1000 копий определенного гена. Генетические и цитогенетические исследования позволили более или менее точно установить положение многих генов *Drosophila*. Используя фазово-контрастную микроскопию, можно локализовать область хромосомы, содержащую определенный ген, и вырезать ее тонкой иглой (рис. 6.3). Таким методом получают сегменты генома длиной около 200 т.п.н., которые далее разрезают рестриктирующими эндонуклеазами и включают в векторные

**РИС. 6.3.**

Вырезание определенной области политенной X-хромосомы *Drosophila* тонкой стеклянной иглой. (С любезного разрешения U. Pirotta.)



молекулы. Достижимое при этом обогащение весьма значительно, поскольку на долю 200 т.п.н. приходится лишь около 0,1% от  $1,8 \cdot 10^8$  п.н., составляющих весь геном.

### В. Синтетические вставки

Успехи, достигнутые в создании методов химического синтеза, позволили синтезировать из простых нуклеозидов молекулы ДНК длиной до 50 остатков. Такие молекулы, которые трудно получить другим путем, можно затем клонировать и выделить в больших количествах. Например, скрининг определенного гена может оказаться невозможным в отсутствие соответствующего зонда. Однако такой ген можно синтезировать *in vitro*, если исходя из аминокислотной последовательности соответствующего полипептида установить нуклеотидную последовательность. Именно таким способом были впервые клонированы в клетках *E. coli* сегменты, кодирующие гормоны соматостатин и инсулин. Эти сегменты получали, синтезируя отдельные олигонуклеотидные блоки и объединяя их затем при помощи ДНК-лигазы. На рис. 6.4 показан процесс синтеза гена соматостатина. Было синтезировано восемь коротких одноцепочечных фрагментов ДНК. Их отжигали и получали двухцепочечные молекулы, стабилизированные водородными связями, образующимися между комплементарными концевыми последовательностями. Поскольку все синтетические продукты имели на 5'-концах гидроксильные группы, перед лигированием концы фосфорилировали с помощью полинуклеотидкиназы и АТР. На рис. 6.4 соответствующие фосфатные группы окрашены. Впоследствии из 66 коротких синтетических фрагментов был воссоздан ген инсулина длиной 514 п.н. Следует отметить, что в связи с вырожденностью генетического кода установление истинной нуклеотидной последовательности аминокислотных остатков в соответствующем полипептиде оказывается невозможным. Тем не менее удалось определить и синтезировать функциональную кодирующую последовательность.

Синтезировать нуклеотидную последовательность целого гена очень трудно. Однако большую ценность представляет синтез даже коротких участ-

ков кодирующей последовательности, поскольку в отсутствие других зондов участки длиной 15–20 нуклеотидов можно использовать в качестве специфических зондов при гибридизации. Синтетические зонды используются при саузерн- и нозерн-гибридизации, а также для скрининга рекомбинантных клонов (разд. 6.4). Кроме того, короткие синтетические последовательности применяются в качестве праймеров при определении нуклеотидных последовательностей длинных сегментов ДНК (разд. 7.2), а также при ферментативном синтезе ДНК на молекулах РНК с помощью обратной транскриптазы. Другой важной областью применения описанной технологии синтеза является получение коротких сегментов ДНК, содержащих сайты узнавания для известных рестриктирующих эндонуклеаз. Такие сегменты-линкеры лигируют с тупыми концами фрагментов ДНК, встраиваемых в векторные молекулы (разд. 6.2).

**Химический синтез полидезоксинуклеотидов.** Для химического синтеза полидезоксинуклеотидов широко применяют два метода. В обоих случаях исходными строительными блоками являются дезоксирибонуклеозиды, и оба метода включают этап связывания мононуклеотидов и олигонуклеотидов. Методы полностью автоматизированы; при этом используется установка «Синтезатор ДНК».

Независимо от метода первым условием является защита тех реакционноспособных групп дезоксинуклеозидов, которые не участвуют в связывании. Так, аминогруппы дезоксиаденозина и дезоксицитозина обычно бензоилируют, а к аминогруппе гуанозина присоединяют остаток изомасляной кислоты (рис. 6.5). Тимидин, у которого аминогруппа отсутствует, используется немодифицированным. 5'-гидроксильные группы обычно защищают путем образования эфирной связи с 4,4'-диметокситрифенилметильным соединением, которое называют диметокситритилом, или сокращенно  $(MeO)_2Tr$  (рис. 6.6). По завершении синтеза все присоединенные группы должны быть удалены. Свободные аминогруппы восстанавливаются при мягком щелочном гидролизе аминокислотных единиц, а диметокситритильные группы удаляются с помощью мягкого кислотного гидролиза.

**Фосфаттриэфирный (или фосфотриэфирный) метод.** Нуклеозид с присоединенной диметокситри-

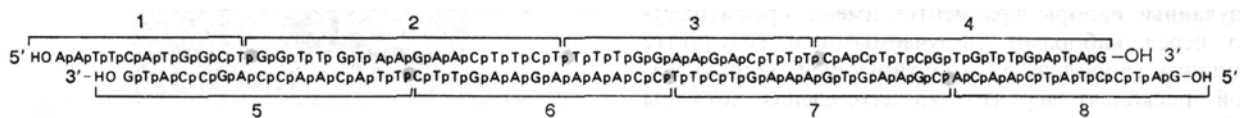
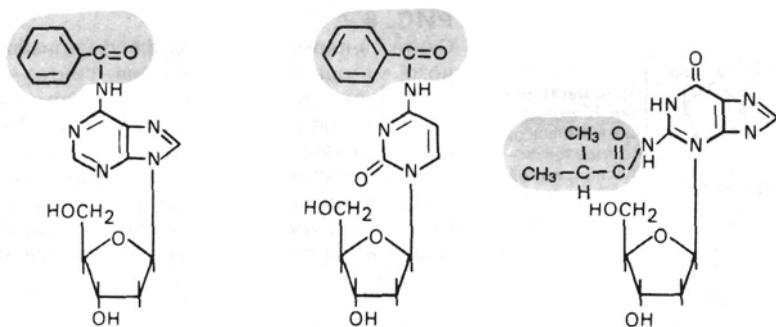


РИС. 6.4.

Синтез гена соматостатина путем лигирования коротких одноцепочечных олигодезоксирибонуклеотидов, синтезированных химическим методом. Цветные кружки -

фосфомоноэфирные группы (р), присоединенные с помощью полинуклеотидкиназы в присутствии АТР.



N-бензоилдезоксаденозин

N-бензоилдезоксцитидин

N-изобутирилдезоксигуанозин

**РИС. 6.5.**

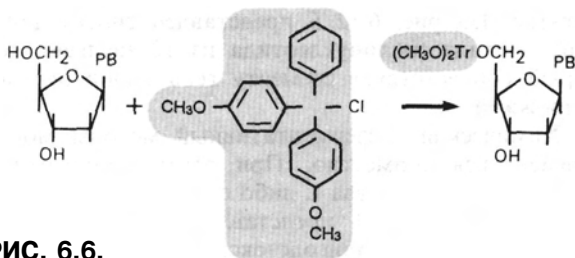
N-ацилзащищенные дезоксирибонуклеозиды - исходные продукты для химического синтеза полидезоксирибонуклеотидов.

тильной группой можно без каких-либо дополнительных модификаций использовать для синтеза фосодиэфира путем присоединения к 3'-ОН-группе, например, *n*-хлорфенилфосордихлоридата (рис. 6.7). Этот диэфир может служить прямым предшественником 5'-конца новой цепи. Далее такой диэфир превращается в триэфир с помощью, например, β-цианоэтанола. Затем 5'-диметокситритильная группа удаляется путем мягкого кислотного гидролиза, в результате чего образуется реакционноспособная 5'-гидроксильная группа (рис. 6.7). Диэфир и триэфир смешивают в присутствии реагентов, стимулирующих их связывание (рис. 6.8), обычно арилсульфоновых соединений, например триизопропилбензолсульфонилхлорида. Точный механизм связывания неизвестен. Продуктом реакции (рис. 6.8, А) является полностью защищенный динуклеотид. При удалении всех защищающих групп образуется простой динуклеозидмонофосфат (рис. 6.8, Б). Существенно, что полностью защищенный динуклеотид является хорошим исходным продуктом для создания более крупных молекул. При обработке кислотой (например, бензолсульфоновой) или  $ZnBr_2$  диметокситритильная группа удаляется и динуклеотид может играть роль 3'-конца растущей цепи (рис. 6.8, В). При обработке же триэтиламином преимущественно гидролизуется цианоэтиловый эфир, в результате чего динуклеотид приобретает реакционноспособный 3'-конец, который может слу-

жить 5'-концом удлиняющейся цепи (рис. 6.8, Г). Связывание двух динуклеотидов дает тетра-нуклеотид (рис. 6.8, Д). Аналогично соответствующим образом защищенные мононуклеотиды и олигонуклеотиды могут соединяться в длинные блоки. В зависимости от выбора исходного материала синтез цепей ДНК идет в направлении 3'→5' или 5'→3'.

Фосфаттриэфирный метод можно упростить и процесс синтеза ускорить, если фиксировать один из концевых нуклеотидов с помощью гидроксильной группы на твердом носителе. В принципе фиксирована может быть любая гидроксильная группа, но обычно реакция идет лучше, если закреплен 3'-гидроксил. Обычно фиксация осуществляется путем образования эфирной связи между 3'-ОН-группой и карбоксильной группой на носителе, например на пористом стеклянном шарике (рис. 6.9). В этом случае первый нуклеотид оказывается фиксированным, а 3'-ОН-группа — защищенной. Далее последовательно добавляются нуклеотиды или олигонуклеотиды с 5'-ОН-группами, защищенными диметокситритилом. Перед каждым следующим этапом присоединения диметокситритил удаляется (рис. 6.7), в результате чего образуется свободная 5'-ОН-группа, по которой может происходить дальнейшее наращивание цепи. По окончании синтеза полидезоксинуклеотид отсоединяют от носителя с помощью щелочного гидролиза (рис. 6.9). При такой процедуре синтеза не нужно проводить дополнительную очистку после каждого акта наращивания цепи, как при других способах. Метод с использованием фиксации на носителе был автоматизирован, при этом синтез полимера длиной 10–20 нуклеотид занимал несколько суток.

**Фосфиттриэфирный (или фосфорамидитный) метод.** Основным исходным продуктом в этом случае тоже являются дезоксинуклеозиды с защищенными аминогруппами и 5'-диметокситритильной группой (рис. 6.5 и 6.6). На схеме, представленной на рис. 6.10, 3'-гидроксильная группа концевой нуклеозиды защищена путем фиксации ее на твердом носителе с помощью эфирной связи. Пред-

**РИС. 6.6.**

Защита 5'-ОН-группы с помощью ди-*n*-метокситритилхлорида. На этом и следующих рисунках РВ - защищенное пуриновое или пиримидиновое основание.

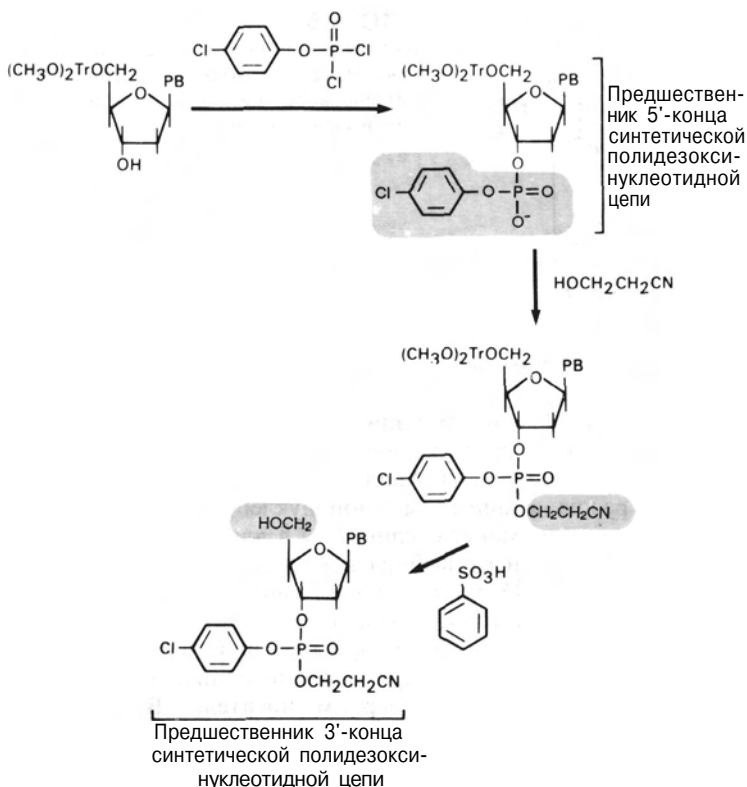


РИС. 6.7.

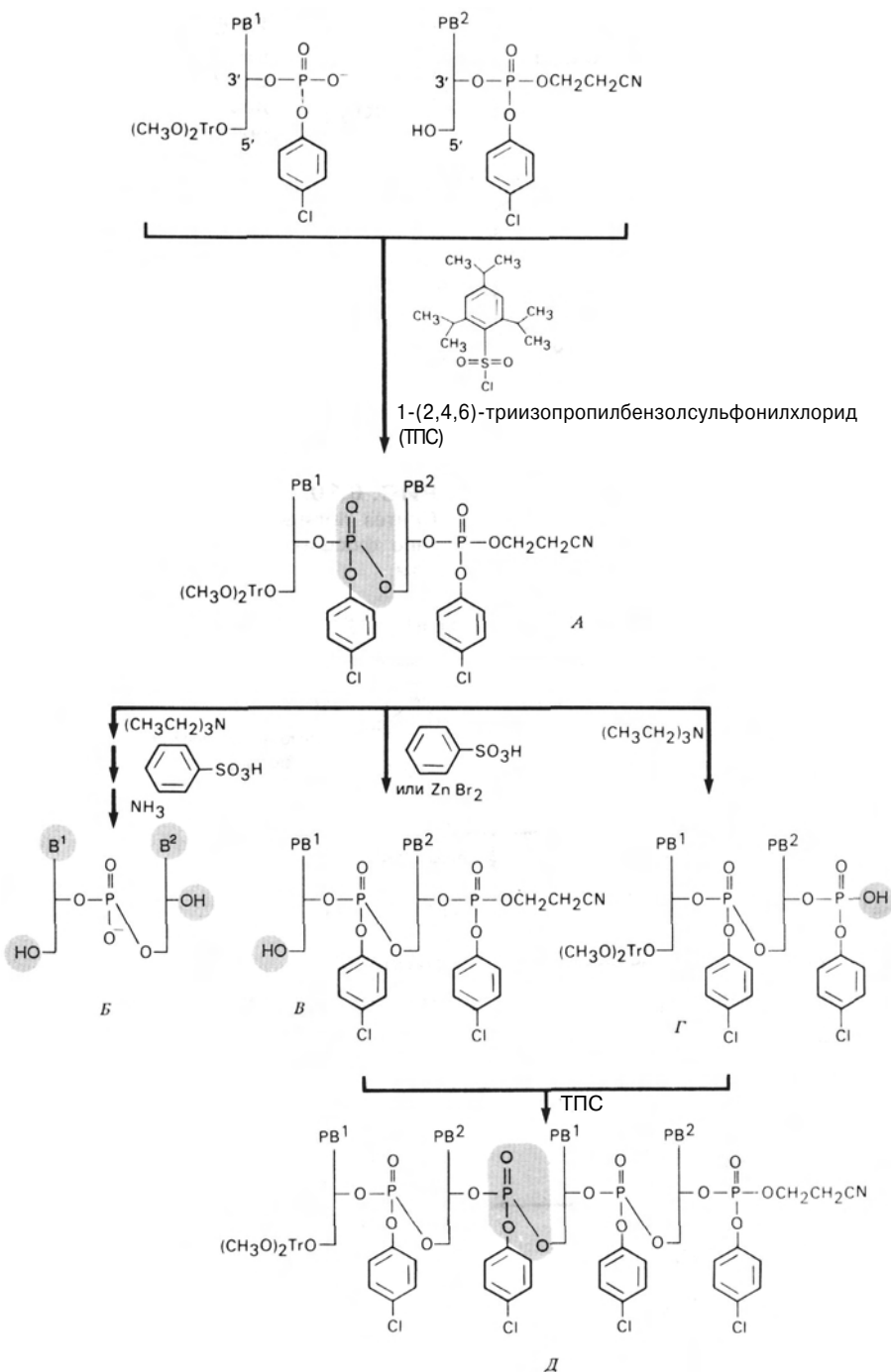
Синтез интермедиатов для фосфаттриэфирного метода. Первая реакция - синтез защищенного фосфодиэфира с использованием *p*-хлорфенилфосфордихлорида. Этот фосфодиэфир может быть превращен в фосфотриэфир с помощью  $\beta$ -цианоэтанола, а 5'-ди-*p*-метокситригидрильная группа может быть удалена при мягком кислотном гидролизе с помощью бензолсульфоновой кислоты.

Предшественником следующего остатка является защищенный дезоксирибонуклеозид-3'-фосфорамидит, который активируется с помощью тетразола. Скорость процесса и его эффективность зависят от фосфорамидитов, стабильных и эффективно связывающихся, легко синтезируемых агентов (рис. 6.11). Промежуточным продуктом реакции связывания является фосфит. Он окисляется до фосфата (триэфира) с помощью иода. Как и в предыдущем случае, триэфир защищает новую межнуклеотидную связь от гидролиза во время последующих этапов синтеза. По завершении синтеза цепи удаляют различные защищающие группы и освобождают полидезоксинуклеотид от твердого носителя с помощью щелочного гидролиза. Фосфитный метод с твердыми носителями, полученными на основе силикагелей (или с использованием пористых стеклянных шариков), применяют при автоматизированном последовательном синтезе олигодезоксирибонуклеотидов. Присоединение каждого нуклеотидного остатка занимает менее 15 мин, при этом могут синтезироваться цепи длиной до 50 остатков с хорошим выходом.

**Ферментативные методы.** Ферментативный синтез коротких полидезоксирибонуклеотидов с заданной последовательностью представляется менее сложным, чем химический синтез, и соответствующи-

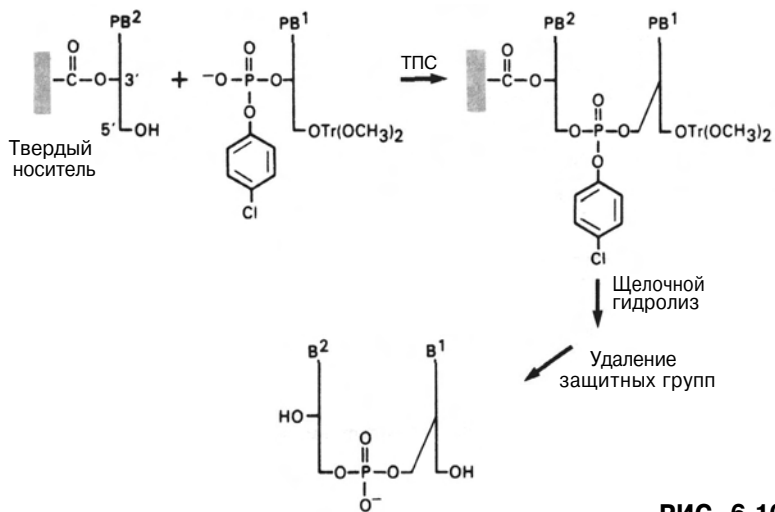
е процедуры более просты в исполнении (если не применяется автоматический ДНК-синтезатор). Для ферментативного синтеза не нужно использовать защитные соединения. Однако реакции имеют свои ограничения и с трудом поддаются контролю. При синтезе используется бактериальный фермент полинуклеотидфосфорилаза, специфичный в отношении рибонуклеотидов, но способный катализировать и полимеризацию дезоксирибонуклеотидов (хотя и с небольшой скоростью), если вместо  $Mg^{2+}$  использовать  $Mn^{2+}$  (рис. 6.12, А). Субстратами всегда служат дезоксирибонуклеозидфосфаты, при этом никакой матрицы не требуется. К праймерам длиной не менее трех остатков присоединяются за один раз один или два дезоксирибонуклеотидных остатка. На рис. 6.12, Б представлен синтез додекануклеотида (полинуклеотида из 12 звеньев), который соответствует сегменту гена изоцитохрома с дрожжей.

Химический и ферментативный методы могут применяться совместно. При этом используется либо ДНК-полимераза I, либо обратная транскриптаза. На рис. 6.13 представлены два частично комплементарных олигодезоксирибонуклеотида, синтезированных химическими методами. Дуплексную молекулу легко получить, используя одноцепочечные участки в качестве матрицы, на которой син-

**РИС. 6.8.**

Химический синтез дидезоксирибонуклеотида фосфат-триэфирным методом. К полностью защищенной динуклеотиду (рис. А) после удаления защищающих групп могут присоединяться другие звенья. Показана структу-

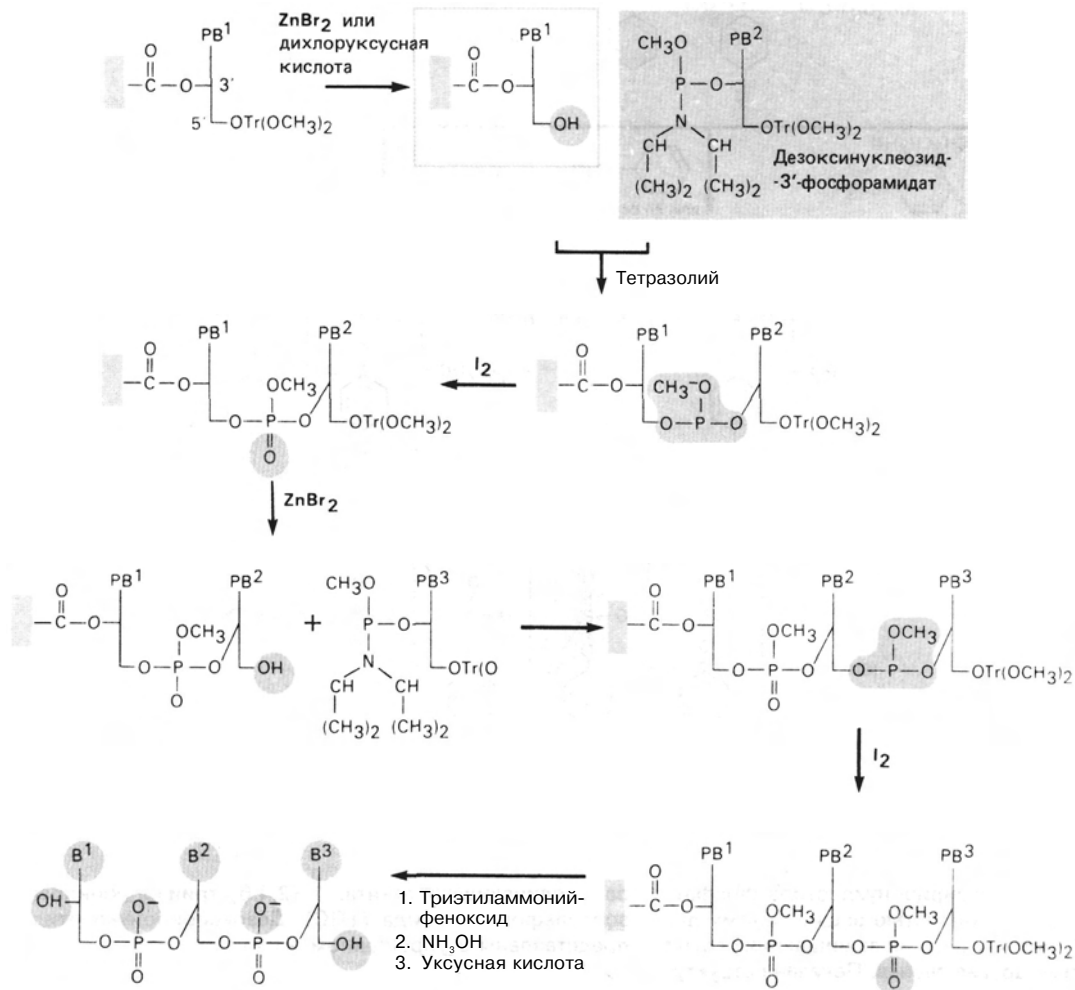
ра соединяющего агента, 1-(2,4,6)-триизопропилбензолсульфонилхлорида (ТПС). Дезоксирибонуклеотиды представлены в упрощенном виде.

**РИС. 6.9.**

Синтез олигодезоксирибонуклеотида фосфаттриэфирным методом на твердом носителе [ТПС - 1-(2,4,6)-триизопропилбензолсульфонилхлорид].

**РИС. 6.10.**

Синтез олигодезоксирибонуклеотида на твердом носителе фосфиттриэфирным (фосфорамидитным) методом.



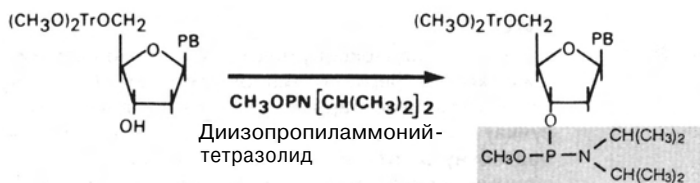


РИС. 6.11.

Синтез дезоксирибонуклеозид-3'-фосфорамидита - интермедиата, используемого в фосфиттриэфирном (фосфорамидитном) методе.

тезируется вторая цепь с помощью ДНК-полимеразы I или обратной транскриптазы (см. рис. 4.19).

### г. Копирование РНК с образованием ДНК

Методы, используемые для анализа РНК, менее совершенны, чем аналогичные методы исследования ДНК. Чтобы определить первичную структуру молекулы РНК, проще всего получить с нее ДНК-копию, клонировать эту копию и определить ее нуклеотидную последовательность. Молекулы ДНК, синтезируемые путем ферментативного копирования РНК-матрицы, получили название **кДНК** (от «комплементарной ДНК» или «ДНК-копии»). В качестве матриц могут использоваться очищенные РНК или смесь молекул РНК, например мРНК из какой-либо ткани или популяции клеток.

Хорошим источником РНК являются высокодифференцированные ткани и клетки, продуцирующие в большом количестве определенные мРНК. Например, в первых успешных работах по исследованию семейства овальбуминовых генов (гл. 7), в которых была применена техника работы с рекомбинантными молекулами ДНК, использовали мРНК, выделенную из специализированных клеток, синтезирующих в большом количестве определенные белки. Клонированные кДНК использовались в качестве зондов для выделения соответствующих генов.

**Копирование эукариотической мРНК.** Общие принципы синтеза двухцепочечной кДНК на мРНК описаны в разд. 4.7 (см. рис. 4.20). На первом этапе последовательность мРНК копируют с помощью обратной транскриптазы и получают комплементарную одноцепочечную ДНК. РНК служит матрицей, а в качестве праймера используется короткая последовательность *oligo*(dT), комплементарно спаренная с *oligo*(rA) на конце РНК. На втором этапе осуществляют деградацию РНК-матрицы либо с помощью рибонуклеазы, либо путем щелочного гидролиза. Затем используют одноцепочечную кДНК в качестве матрицы для синтеза комплементарной цепи ДНК с помощью ДНК-полимеразы I или обратной транскриптазы. На этом этапе праймер не нужен, поскольку на 3'-конце первой структуры временно образуется шпилька варьирующей длины, которая и служит праймером. На последнем этапе эту шпильку разрезают с помощью нуклеазы, специфичной в отношении одноцепочечной ДНК, и получают дуплексную кДНК.

**Копирование других РНК.** Для получения кДНК на основе геномов некоторых РНК-содержащих вирусов или других молекул РНК, не имеющих на конце *oligo*(rA), необходимы некоторые ухищрения. В частности, к 3'-ОН-концу РНК можно присоединить участок *poly*(A), используя *poly*(A)-полимеразу (разд. 4.9) и АТР. Далее всю последовательность реакций можно проводить так же, как и раньше. Альтернативный подход состоит в отжиге с данной

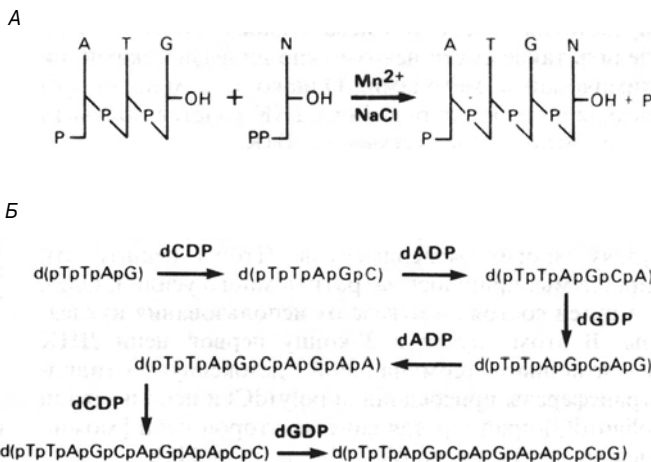
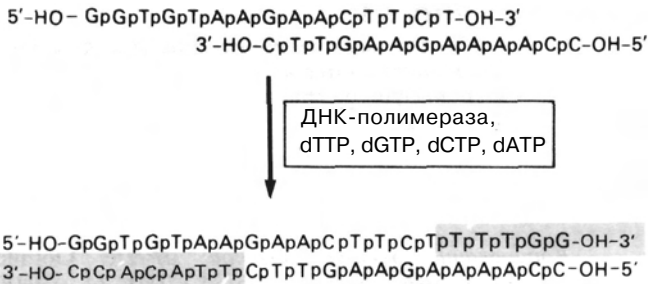


РИС. 6.12.

А. Присоединение одного дезоксирибонуклеотидного остатка к олигодезоксирибонуклеотиду, катализируемое полинуклеотидфосфорилазой. Б. Ферментативный синтез сегмента гена изоцитохрома с дрожжей. На каждом этапе присоединяются один или два новых дезоксирибонуклеотидных остатка.

**РИС. 6.13.**

Синтез полидезоксинуклеотида с помощью химического и ферментативного методов. Две частично комплементарные цепи (верхняя часть рисунка) синтезированы химическим методом и подвергнуты отжигу. Образовавшийся несовершенный дуплекс используется в качестве матрицы и праймера для синтеза полностью дуплексной молекулы с помощью ДНК-полимеразы I *E. coli* или обратной транскриптазы.

РНК олигодезоксирибонуклеотида, комплементарного какому-либо участку молекулы, для образования праймера.

**Осложнения.** На рис. 4.20 схематически представлена некая идеализированная схема синтеза кДНК. На практике же на каждом этапе, начиная с приготовления очищенного препарата мРНК, возникают сложности, в результате чего конечный продукт обычно представляет собой смесь молекул кДНК, большинство из которых оказываются более короткими, чем полноразмерная РНК. Перечислим некоторые из возникающих проблем. 1. мРНК может частично деградировать при выделении под действием рибонуклеаз. 2. мРНК может копироваться неполностью, в результате чего на 5'-конце кДНК будет нехватать каких-то последовательностей. Чем длиннее молекула РНК, тем больше вероятность, что эта проблема возникнет. 3. Синтез второй цепи ДНК может остановиться прежде, чем кончится матрица, так что 3'-конец РНК будет пропущен. Оставшийся конец первой цепи будет удален с помощью нуклеазы, специфичной в отношении одноцепочечных молекул. 4. Нуклеаза может отщепить не только шпильку, но и концы дуплекса.

Все это, а также тот факт, что даже высокоочищенные препараты РНК никогда не бывают абсолютно гомогенными, приводит к тому, что препараты двухцепочечных молекул кДНК всегда представляют собой смесь разных молекул. Разделить такие смеси невозможно ни физическими, ни химическими методами. Однако с помощью молекулярного клонирования кДНК удается получить абсолютно чистые сегменты ДНК.

Серьезные проблемы, возникшие при синтезе полноразмерных копий кДНК на РНК-матрицах с помощью стандартных методов, определили неудачу многих экспериментов. Чтобы решить эти проблемы, пришлось затратить много усилий. Один из путей состоял в отказе от использования нуклеазы. В этом случае к 3'-концу первой цепи ДНК с помощью терминальной дезоксирибонуклеотидилтрансферазы присоединяли poly(dC) и использовали oligo(dG)-праймер для синтеза второй цепи [можно использовать также oligo(dA) и oligo(dT)].

На рис. 6.14 представлен другой метод, в котором устранен главный недостаток, свойственный стандартному методу, — синтез неполных копий РНК. Во-первых, oligo(dT)-праймер включается в состав плазмидного вектора *E. coli*, так что кДНК может синтезироваться непосредственно на векторе. В результате исключаются этапы выделения и лигирования. Во-вторых, не применяется специфичная к одноцепочечной ДНК нуклеаза, благодаря чему последовательности, соответствующие 3'-концу мРНК, не утрачиваются. В-третьих, молекулы ДНК, синтез которых не доходит до конца РНК-матрицы, содержат заглубленный 3'-конец, который является неэффективным субстратом для терминирующей трансферазы (разд. 4.8). При таком подходе получают популяцию рекомбинантных ДНК, обогащенную полноразмерными копиями РНК в виде кДНК, способными осуществлять трансфекцию клеток-хозяев.

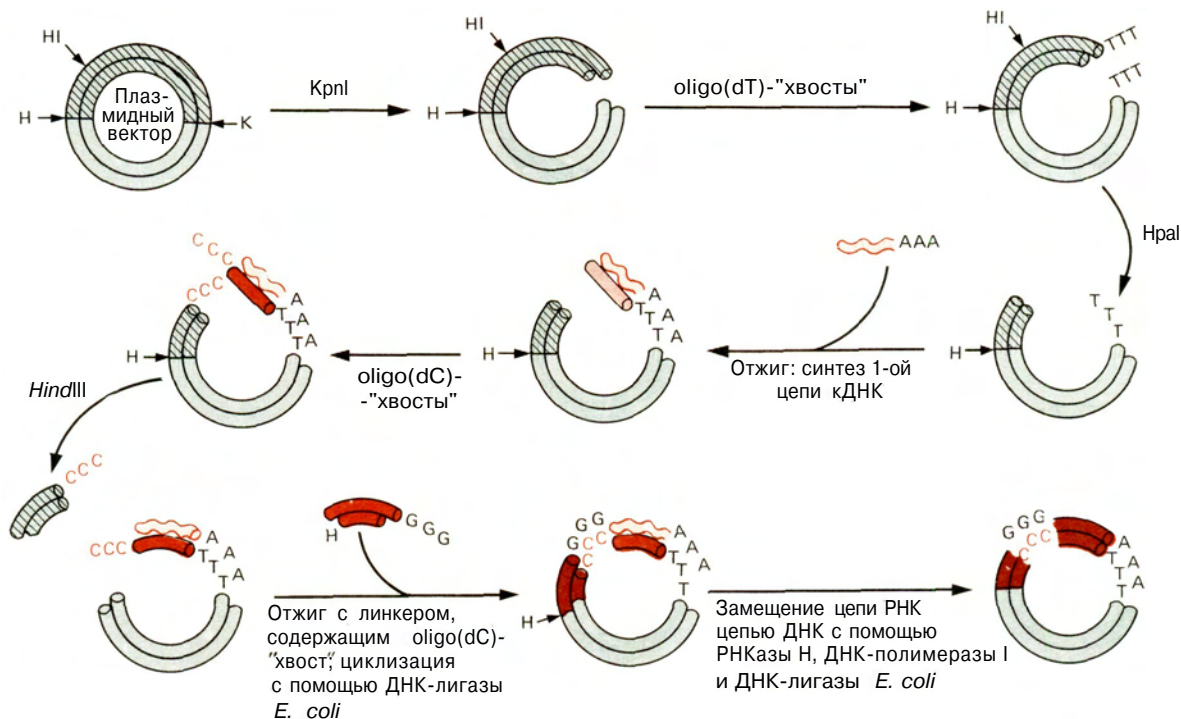
Немного модифицировав этот метод, можно получить плазмиды, содержащие молекулы кДНК в форме, способной экспрессироваться в клетках эукариот. Модификация состоит во включении эукариотического промотора и сигналов процессинга РНК в линкерный сегмент, показанный на рис. 6.14.

## 6.2. ЛИГИРОВАНИЕ ВЕКТОРА СОВСТАВКОЙ

После того как получены вектор и вставка, можно приступать к конструированию рекомбинантных ДНК. Обычно нужно осуществить только два двухцепочечных соединения независимо от того, представляет ли собой вектор линейную плазмиду или вирусный геном или два плеча ДНК фага  $\lambda$  объединены в одну линейную молекулу по *cos*-сайтам. Если вставки имеют тупые концы, то перед лигированием для его облегчения к ним часто присоединяют липкие концы.

### а. Соединение концов

Рассмотрим четыре возможные ситуации: вектор и вставка имеют полностью совпадающие липкие концы (рис. 6.15, А); две разные пары совпадающих

**РИС. 6.14.**

Клонирование полноразмерных молекул кДНК. Кольцевой специально сконструированный плазмидный вектор *E. coli*, полученный из rBR322, превращают в линейную форму с помощью рестриктирующей эндонуклеазы (*KpnI*), которая разрезает молекулу в одном из сайтов в несущественной области. С помощью терминальной дезоксирибонуклеотидилтрансферазы к 3'-концам молекулы присоединяют остатки дТ. Затем один из oligo(dT)-участков удаляют, используя соответствующую рестриктирующую эндонуклеазу (*HpaI*), в результате чего образуется тупой конец. При этом способе oligo(dT), используемый в качестве праймера в первой реакции обратной транскрипции, является составной частью плазмидного вектора. Вектор содержит также уникальный сайт для рестриктирующей эндонуклеазы (в данном случае *HindIII*), расположенный вблизи конца, противоположного тому, который содержит oligo(dT), в несущественной области. После отжига с РНК и синтеза первой цепи кДНК к соответствующим 3'-концам при-

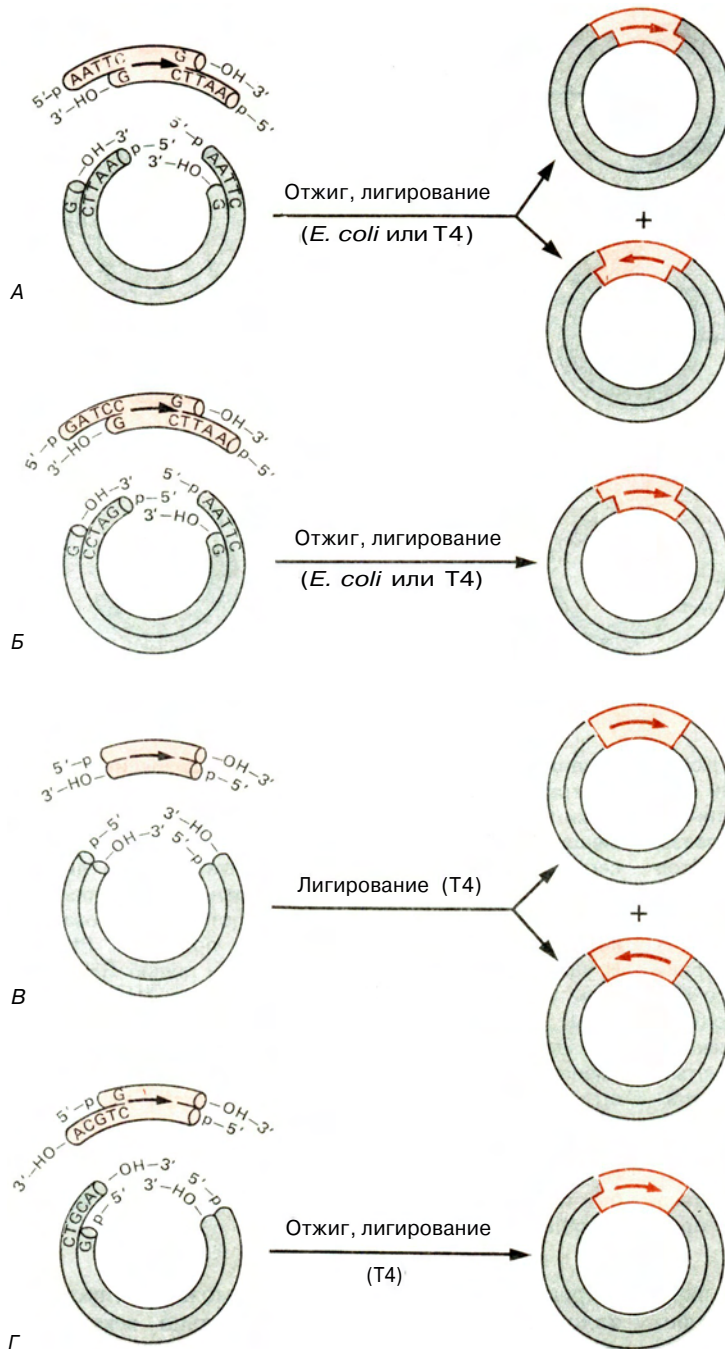
соединяют с помощью терминальной дезоксирибонуклеотидилтрансферазы «хвосты» oligo(C). Затем лишние остатки С убирают, используя эндонуклеазу *HindIII*, и замещают путем лигирования с заранее подготовленным сегментом с G-«хвостом» на одном конце и липким *HindIII*-сайтом - на другом. С помощью этого фрагмента два конца конструкции объединяются с образованием кольцевой структуры. Наконец, с помощью РНКазы *H* удаляют цепь РНК и замещают ее цепью ДНК (при этом G-«хвост» является праймером для ДНК-полимеразы), а кольцевую структуру сшивают ДНК-лигазой. При всех манипуляциях необходимо поддерживать условия, способствующие сохранению гетеродуплексной структуры, поскольку РНК удерживается на векторе благодаря комплементарному спариванию. Заметим также, что гибридная РНК-ДНК устойчива к действию *HindIII*, в противном случае расщепление происходило бы также по любому из находящихся во вставке *Hind III*-сайту.

концов (рис. 6.15, Б); тупые концы (рис. 6.15, В); пару совпадающих липких и пару тупых концов (рис. 6.15, Д). Структура вектор-вставка, в которой каждый из компонентов содержит один липкий и один тупой концы (рис. 6.15, Д), может образоваться только одним способом. То же самое относится к случаю, когда вектор и вставка имеют по два разных совпадающих липких конца (рис. 6.15, Б). В тех же случаях, когда каждая молекула имеет два идентичных липких конца (рис. 6.15, А) или два тупых конца (рис. 6.15, В), вставка может быть

встроена в любой из двух возможных (противоположных) ориентации относительно вектора, в результате чего образуются два разных продукта. Липкие концы вектора и вставки соединяют с помощью ДНК-лигазы в условиях, способствующих образованию водородных связей между комплементарными участками. Для объединения тупых концов ДНК-лигаза Т4 и фрагменты должны присутствовать в высоких концентрациях, поскольку лигаза имеет низкое сродство к тупым концам.

Помимо лигирования вектора и вставки может

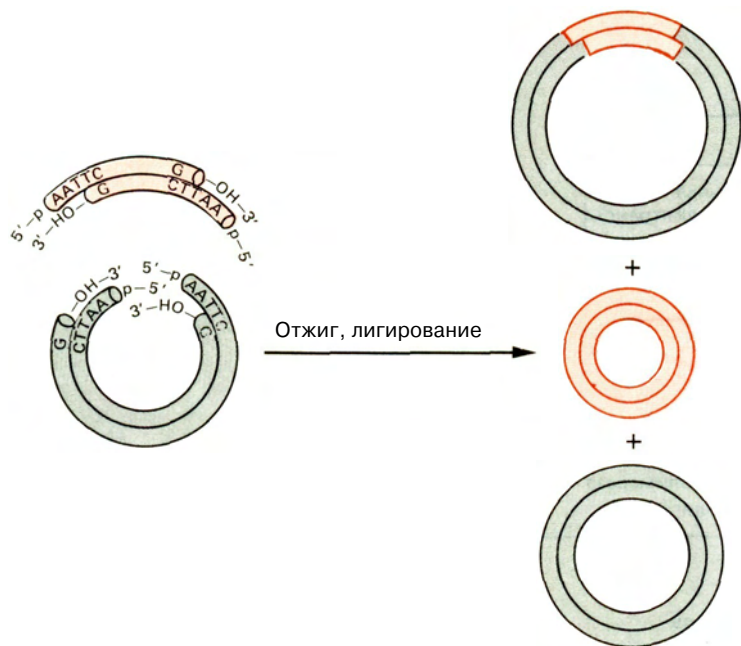


**РИС. 6.15.**

Лигирование вставки и вектора. Напомним, что ДНК-лигаза *E. coli* не способна катализировать воссоединение фрагментов с тупыми концами (разд. 4.5).

происходить внутримолекулярное лигирование этих двух компонентов, в результате чего выход рекомбинантных молекул снижается (рис. 6.16). Вероятность таких побочных реакций можно уменьшить, проводя перед лигированием дефосфорилирование либо вектора, либо вставки (рис. 6.17). В отсутствие 5'-фосфоэфирных групп внутри-

молекулярное лигирование осуществляться не может. Обычно дефосфорилируют векторную молекулу, чтобы свести к минимуму ее способность к амплификации после трансфекции. Вставка, замкнувшаяся в кольцо, не способна к репликации, поэтому с ней возникает меньше проблем. Как показано на рис. 6.17, рекомбинантная молекула,

**РИС. 6.16.**

Побочные продукты лигирования: внутри-молекулярное спаривание.

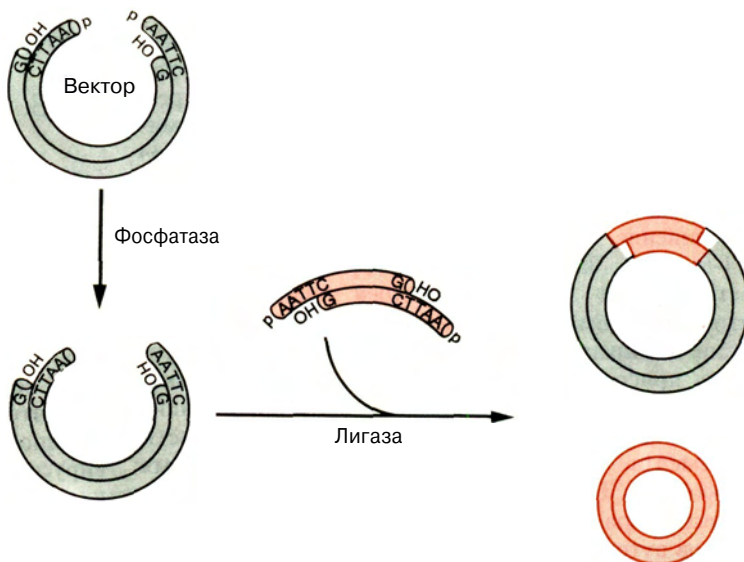
образовавшаяся из дефосфорилированной векторной молекулы, содержит по одному пробелу в каждой из цепей. Такие молекулы легко репарируются в клетках соответствующих хозяев.

### б. Присоединение липких концов

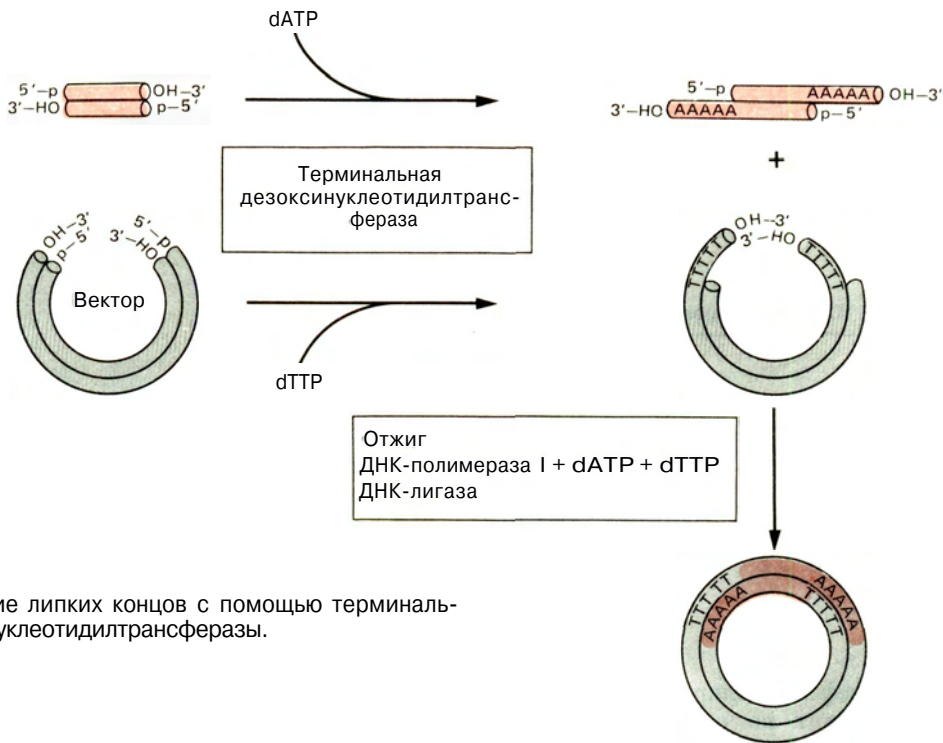
Встраивание сегментов ДНК в векторные молекулы облегчается, если эти сегменты и вектор имеют одинаковые липкие концы. Существуют два удобных способа образования липких концов у сегментов ДНК с тупыми концами (таких, как

кДНК): образование одноцепочечных «хвостов» и введение сайтов рестрикции.

*Образование одноцепочечных «хвостов».* Если с помощью терминальной дезоксирибонуклеотидилтрансферазы нарастить каждый из 3'-гидроксильных концов двухцепочечной вставки с тупыми концами и присоединить комплементарные нуклеотиды к концам линейной векторной ДНК, то при отжиге образуется кольцевая рекомбинантная молекула (рис. 6.18). Если размер «хвостов» одинаков, то концы легко сшиваются ДНК-лигазой. Однако такая ситуация встречается редко, поэтому перед

**РИС. 6.17.**

Дефосфорилирование, препятствующее внутримолекулярному лигированию векторной молекулы.

**РИС. 6.18.**

Присоединение липких концов с помощью терминальной дезоксирибонуклеотидилтрансферазы.

лигированием бреши заполняют с помощью ДНК-полимеразы I. Напомним также, что для терминальной дезоксирибонуклеотидилтрансферазы требуются одноцепочечные 3'-гидроксильные концы в качестве праймеров и что для эффективного наращивания дуплексных молекул ДНК с тупыми концами или молекул, у которых 3'-гидроксильные концы заглублены (разд. 4.8), необходимы особые условия.

**Введение сайтов рестрикции.** Короткие фрагменты ДНК с тупыми концами, содержащие последовательности с сайтами для специфических рестриктирующих эндонуклеаз, получают путем химического синтеза. Структуры двух таких «линкеров» и способ получения с их помощью липких концов у вставки с тупыми концами показаны на рис. 6.19. Соединение линкера со вставкой катализируется T4-лигазой. Затем эту ДНК обрабатывают соответствующей рестриктирующей эндонуклеазой, чтобы получить липкие концы. На этом последнем этапе возникают проблемы. Если фрагмент, к которому присоединены линкеры, содержит сайт узнавания для используемого фермента, то образование липких концов приводит к нарушению целостности фрагмента. Эту проблему можно решить двумя способами. Первый состоит в использовании другого линкера. Второй основан на защите внутренних сайтов путем метилирования с помощью метилазы,

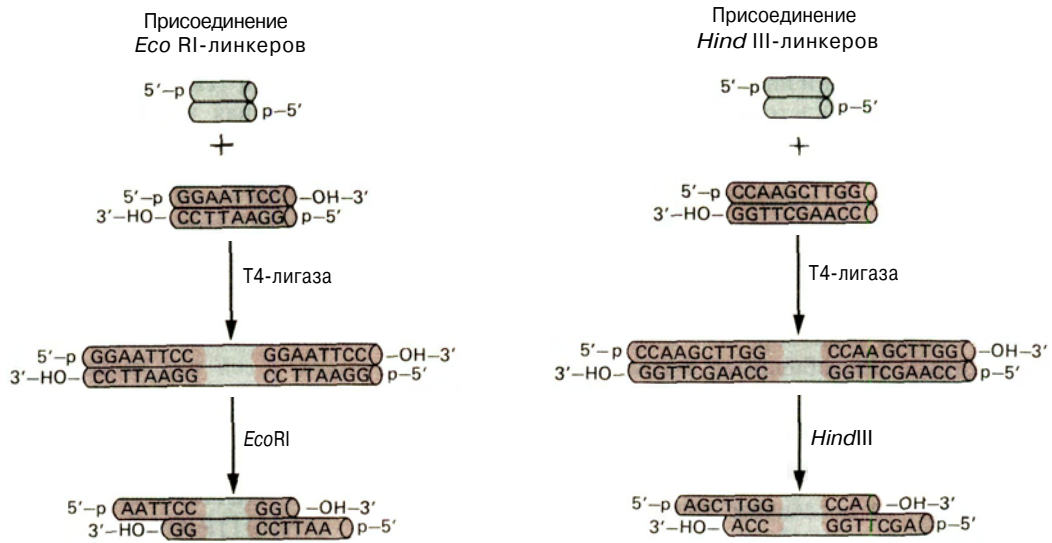
специфичной в отношении данного сайта (рис. 6.20 и разд. 4.2.д).

Линкеры используют также для образования одинаковых липких концов у молекул вставки и вектора, имеющих разные липкие концы. Последние сначала превращают в тупые с помощью специфичной к одноцепочечным участкам нуклеазы или путем достройки заглубленных концов с помощью ДНК-полимеразы I или обратной транскриптазы, а затем присоединяют линкеры.

### 6.3. ИНФЕКЦИЯ, ТРАНСФЕКЦИЯ И КЛОНИРОВАНИЕ

#### а. Перенос рекомбинантных молекул из пробирки в клетку

Сконструированные рекомбинантные молекулы ДНК вводят в клетки или вирусные частицы для клонирования и амплификации с помощью методов, описанных в гл. 5. Для разных систем хозяин-вектор применяются разные методы. Напомним, что рекомбинантные ДНК, сконструированные на основе бактериальных и дрожжевых плазмид или вирусов эукариот, трансфицируются в хозяйские клетки только после того, как клеточные мембраны (или стенки) становятся проницаемыми. Рекомбинанты, сконструированные с использованием  $\lambda$ -век-



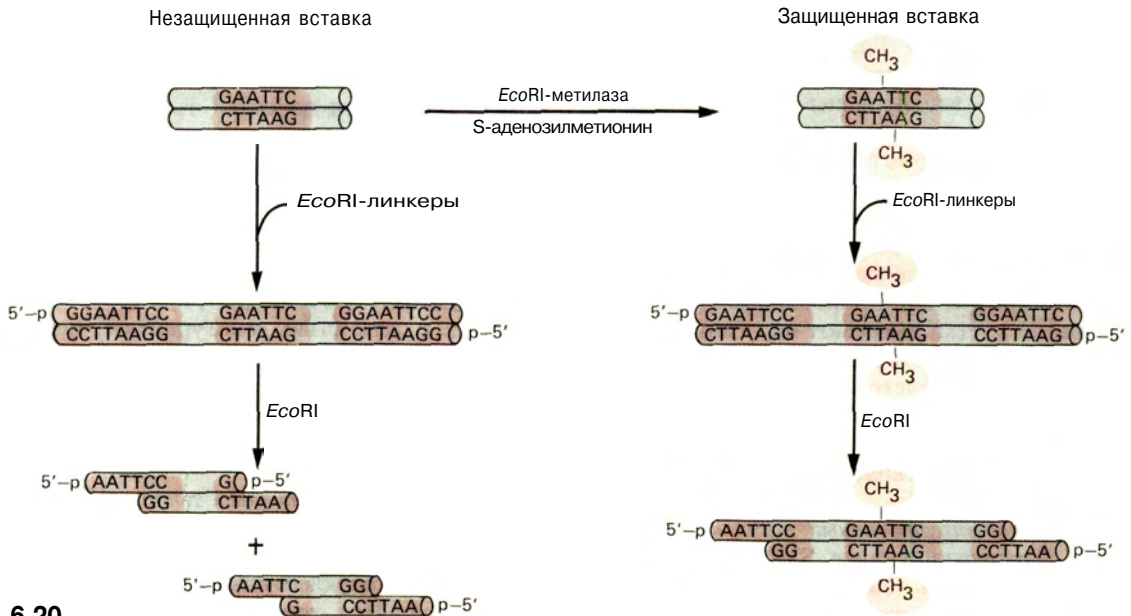
**РИС. 6.19.** Образование липких концов с помощью линкеров.

торов или космид, упаковываются в фаговые частицы, которые затем используются для инфицирования перmissive клеток *E. coli* K12.

Обычно трансфекция происходит не очень эффективно. Рекомбинантный геном включается лишь в часть обработанных клеток. Выращивая клетки в условиях, при которых проявляется зависимость

от векторных генов, можно идентифицировать нужные клетки и отобрать их.

*Одна молекула ДНК на клетку.* При планировании эксперимента по молекулярному клонированию и его проведении необходимо соблюдать два основных принципа. Во-первых, после конструирования *in vitro* отдельные молекулы рекомбинантных ДНК должны быть введены в разные структуры. Ни одна из клеток не должна получить более одной плазмидной молекулы или вирусной частицы. Во-



**РИС. 6.20.** Сайты для рестриктирующей эндонуклеазы, находящиеся во встроенном фрагменте, защищены путем мети-

лирования до присоединения линкеров.

вторых, эти структуры должны быть способны к репликации.

В результате трансфекции и инфекции образуются популяции самых разных клеток либо популяции вирусов или бактериофагов. Одни члены популяции содержат нужные рекомбинантные молекулы, другие несут рекомбинанты, содержащие нежелательные вставки, или векторные молекулы вообще без вставки, третьи (если используются плазмиды и некоторые эукариотические вирусные векторы) представляют собой неизмененные клетки. Кроме того, могут встречаться разнообразные непредсказуемые и аберрантные рекомбинанты, в том числе векторы, в которые включены две или более вставки, и рекомбинанты, у которых в результате рекомбинации уже в клетке-хозяине произошло изменение вставки. Таким образом, трансфекция и инфекция не являются последним этапом эксперимента по получению рекомбинантной ДНК, а представляют лишь один из его этапов. Далее нужно провести клонирование и идентификацию нужного рекомбинанта.

## б. Клонирование

Необходимым условием успешного клонирования является возможность разделения всех трансфицированных или инфицированных клеток. Только в этом случае каждый клон будет представлен отдельной колонией бактериальной или животной клетки или негативной фаговой либо вирусной колонией и будет содержать определенную рекомбинантную ДНК.

*Плазмидные векторы.* Трансфицированные клетки высевают на агар таким образом, чтобы они были полностью изолированы одна от другой и каждая могла дать начало отдельной колонии. Обычно векторы содержат по крайней мере один селективный маркер, по которому проводится отбор трансфицированных клеток. При этом нетрансфицированные клетки не могут образовывать колонии в используемой среде.

*Фаговые векторы.* Инфицированные (или трансфицированные) клетки высевают на газоне чувствительных клеток, на котором в результате инфицирования соседних клеток образуются фаговые бляшки. Некоторые векторные системы имеют встроенные маркеры, позволяющие легко отбирать рекомбинанты среди реконструированных интактных векторов (разд. 5.3).

*Векторы, сконструированные на основе вирусов животных.* Трансдуцирующие SV40-векторы, способные продуцировать вирусные частицы (разд. 5.7.б), клонируются так же, как и фаговые векторы. Однако, поскольку такие SV40-векторы нуждаются в вирусе-помощнике, выполняющем утраченные

ими функции, вирусные бляшки (зоны лизиса на монослое клеток) могут содержать как рекомбинантные геномы, так и геномы помощника. Плазмиды, сконструированные на основе вируса SV40, и пассивно трансформирующие векторы, которые не продуцируют вирусных частиц, сначала клонируют как челночные векторы в клетках *E. coli*. Обычно они содержат селективный маркер, определяющий фенотип животной клетки, что позволяет отобрать клоны трансформированных клеток. Аналогично рекомбинантные векторы на основе папилломавирусов крупного рогатого скота и ретровирусов клонируют в клетках *E. coli* как челночные векторы, а затем трансфицируют ими клетки животных.

## 6.4. СКРИНИНГ КЛОНИРОВАННЫХ ПОПУЛЯЦИЙ РЕКОМБИНАНТНЫХ МОЛЕКУЛ

### а. Обнаружение нужного клона

После разделения рекомбинантных молекул и получения отдельных клонов встает трудно разрешимая задача — обнаружение нужного клона или клонов. Идентификация клона основывается на том, что вставка в рекомбинантной ДНК детерминирует какое-то уникальное свойство содержащей ее клетки. Это свойство может определяться структурой самой вставки (т.е. нуклеотидной последовательностью) либо быть связанным с ее функцией (т.е. продуктом экспрессии клонированного гена). На нем основывается скрининг популяции клонов с целью идентификации одного нужного клона. Методы скрининга должны быть очень чувствительными, поскольку иногда приходится идентифицировать один клон из сотен тысяч или даже миллионов клонов.

### б. Отжиг с комплементарным полинуклеотидом

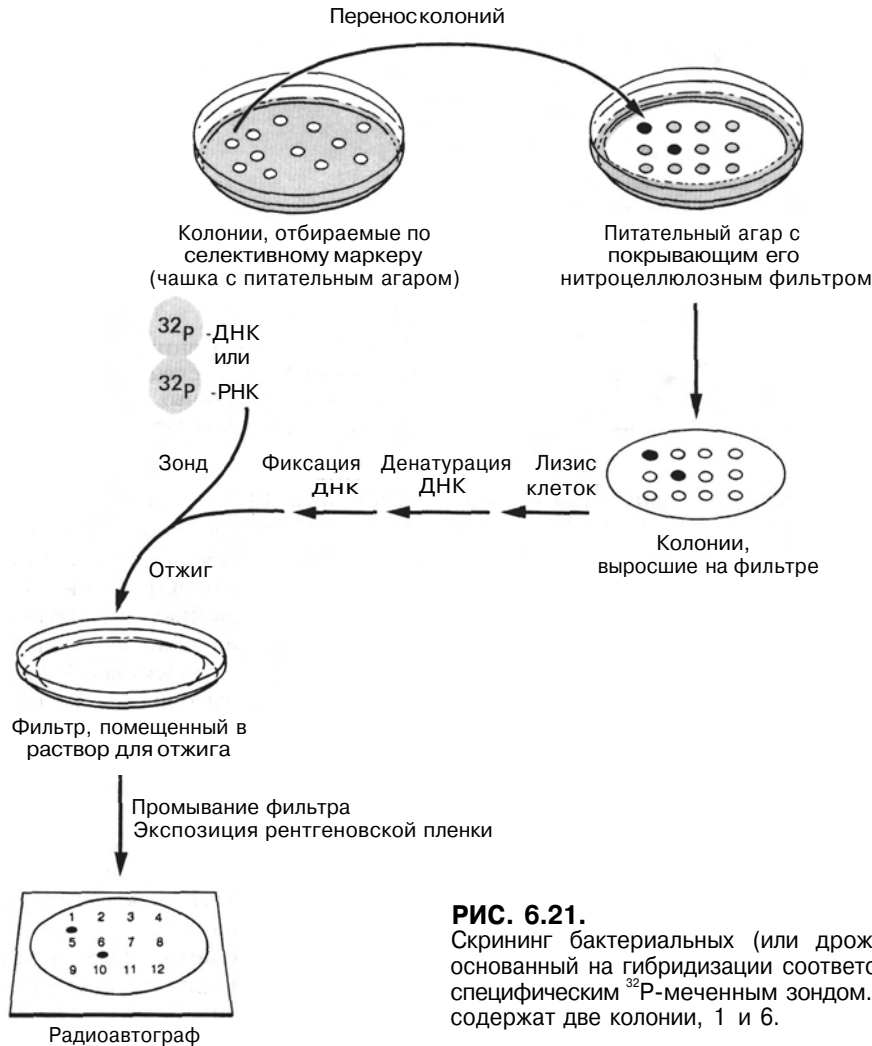
Комплементарные одноцепочечные полинуклеотиды, будь то РНК или ДНК, при соответствующих условиях с легкостью ренатурируют, несмотря на присутствие большого избытка неродственных цепей. Благодаря этому исследователь получает мощный инструмент для обнаружения специфической вставки, который можно использовать в разных ситуациях.

*Отжиг с радиоактивным зондом.* Предположим, что мы имеем большую популяцию клонированных рекомбинантных молекул и проводим отжиг с определенным полинуклеотидным зондом. Тогда мечеными становятся только те рекомбинантные молекулы, которые содержат комплементарности, комплементарные зонду. Успех эксперимента зави-

сит прежде всего от наличия нужного зонда.

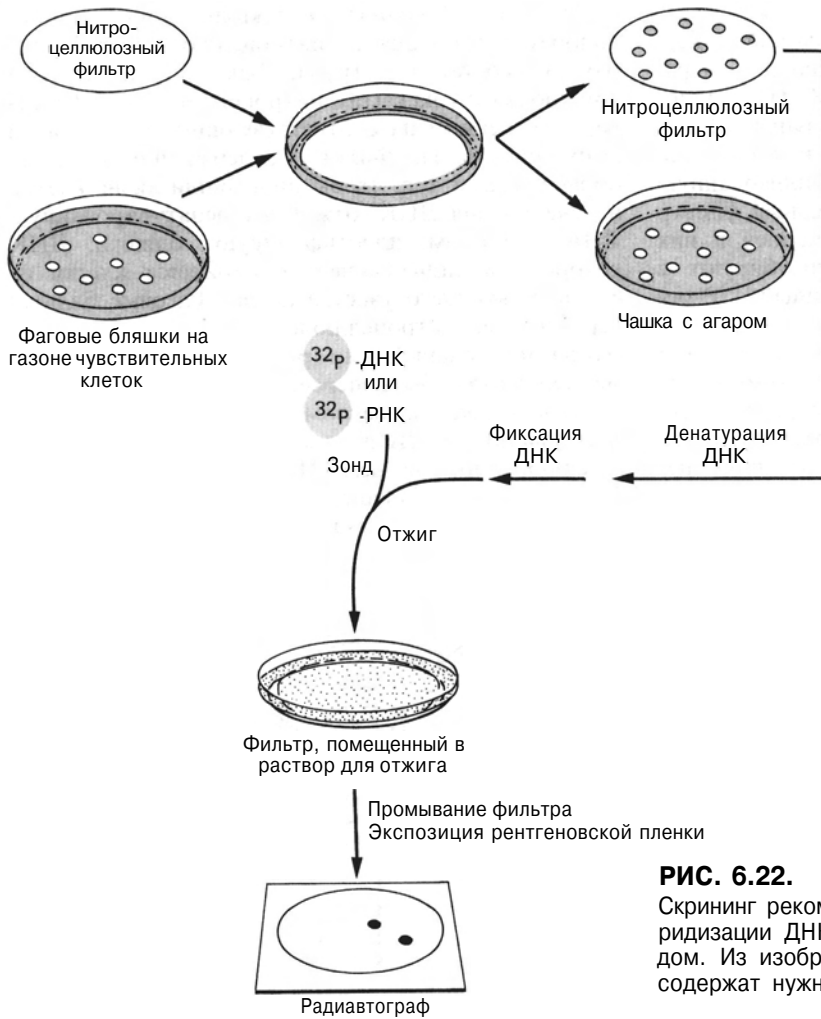
Для скрининга большого числа колоний, содержащих плазмиды, а также фаговых бляшек разработаны удобные и быстрые методы. На тех же принципах основан и скрининг SV40-бляшек. Метод состоит в том, что бактериальные колонии или бляшки переносят с агара на иммобилизованную подложку, например на нитроцеллюлозный фильтр (разд. 6.1.6), и подвергают содержащуюся в них ДНК денатурации. Затем фильтр инкубируют в растворе, содержащем денатурированный  $^{32}\text{P}$ -меченный зонд, в условиях, способствующих ренатурации комплементарных цепей. Если колония или бляшка содержит ДНК, комплементарную зонду, то на радиоавтограмме в соответствующем месте будет наблюдаться потемнение. Несмотря на то что каждая бляшка или колония содержит лишь небольшое количество ДНК, ее удается выявить благодаря высокой удельной радиоактивности зондов.

Получить большие коллекции клонированных колоний или бляшек на нитроцеллюлозных фильтрах не составляет труда. Бактерии и дрожжи хорошо растут на самом нитроцеллюлозном фильтре, помещенном на соответствующий питательный агар (рис. 6.21). Фильтр, содержащий колонии, обрабатывают так, чтобы произошли лизис клеток и денатурация ДНК, отжигают денатурированную ДНК с зондом, идентифицируют бляшки, ДНК которых гибридизовалась, и отбирают клетки из соответствующего участка агара. Фаговые бляшки переносят на нитроцеллюлозный фильтр, наложив его на питательный агар (рис. 6.22). За один раз на фильтр может быть перенесено до 10000 бляшек с одной чашки диаметром 10 см. Если осторожно снять фильтр, то ДНК из каждой бляшки останется фиксированной на нем. После отжига и идентификации нужных бляшек можно отобрать фаговые частицы из бляшек, оставшихся на чашке с агаром.



**РИС. 6.21.**

Скрининг бактериальных (или дрожжевых) колоний, основанный на гибридизации соответствующей ДНК со специфическим  $^{32}\text{P}$ -меченным зондом. Нужную вставку содержат две колонии, 1 и 6.

**РИС. 6.22.**

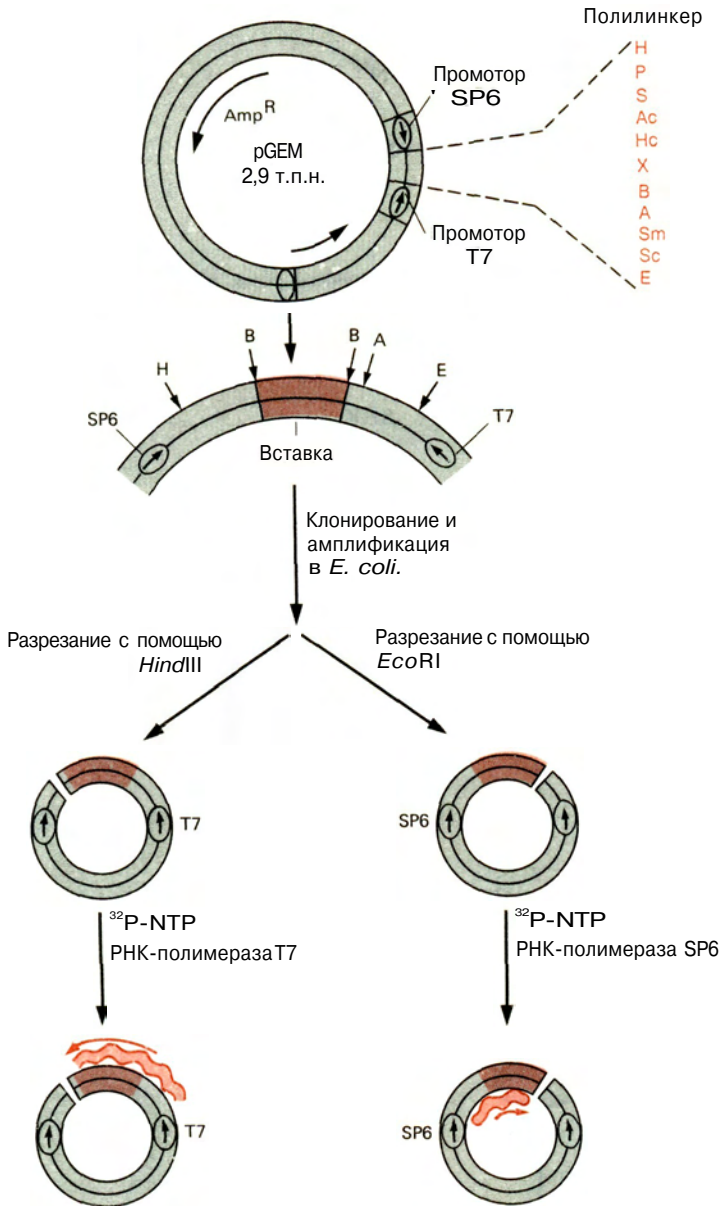
Скрининг рекомбинантных бляшек, основанный на гибридизации ДНК со специфическим  $^{32}\text{P}$ -меченным зондом. Из изображенных на рисунке десяти бляшек две содержат нужную вставку.

Аналогичным образом можно выявить упакованные рекомбинантные молекулы вируса SV40. Питательный агар, покрывающий инфицированный монослой клеток обезьяны, удаляют, весь монослой переносят на нитроцеллюлозный фильтр и идентифицируют нужные бляшки путем отжига с соответствующим радиоактивно меченным зондом. Вирионы, содержащие рекомбинантные ДНК, отбирают из агарового слоя.

В качестве зондов можно использовать высокоочищенные мРНК и денатурированные гомологичные (обычно клонированные) ДНК или кДНК. Особенно удобны одноцепочечные зонды, поскольку их не нужно денатурировать и тестируемые клонированные ДНК не конкурируют с другими гибридизующимися цепями ДНК во время отжига. мРНК-зонды уже являются одноцепочечными, а одноцепочечные кДНК-зонды легко получить, скопировав только одну цепь (см. рис. 4.20). Одно-

цепочечные РНК-зонды можно также приготовить из рекомбинантной плазмиды, содержащей последовательность зонда, встроенную в специальный вектор, например рGEM (рис. 6.23). Этот вектор содержит два разных высокоспецифичных бактериальных промотора, фланкирующих вставку в рекомбинантных молекулах. Используя очищенную соответствующим образом линейаризованную плазмиду в качестве матрицы *in vitro* и ту или иную из РНК-полимераз прокариот, синтезируют РНК, которая представляет собой копию одной из двух цепей вставки.

Если известна аминокислотная последовательность полипептида, кодируемого данным геном, то можно восстановить нуклеотидную последовательность этого гена и затем синтезировать его химическими методами. Даже в тех случаях, когда в качестве зонда используется участок ДНК длиной всего 15 нуклеотидов, можно получить относитель-

**РИС. 6.23.**

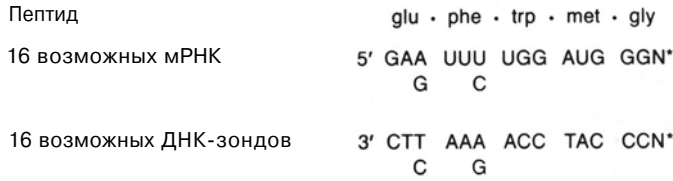
Рекombинантная плазмидная ДНК, на которой синтезируется радиоактивно меченная одноцепочечная РНК. Плазмида рGEM содержит последовательности, происходящие из рBR322, и два прокариотических промотора. Промотор SP6 получен из бактериофага SP6 *Salmonella typhimurium* и способен функционировать только со специфичной РНК-полимеразой SP6. Промотор T7 получен из бактериофага T7 *E. coli* и функционирует только со специфичной T7-РНК-полимеразой. Между этими двумя промоторами находится полилинкер, содержащий множество сайтов для эндонуклеаз рестрикции. В данном случае вставка включена в сайт *Bam* H1, но в принципе она может включиться в любой из этих сайтов. Прежде чем осуществлять транскрипцию *in vitro*, очищенную рекombинантную плазмидную ДНК разрезают с помощью рестриктирующей эндонуклеазы и получают линейные молекулы. В зависимости от сайта рестрикции, структура которого в свою очередь зависит от того, какую РНК-полимеразу используют - SP6 или T7, получают РНК-копию одной или другой цепи вставки.

но достоверный результат. Из-за вырожденности генетического кода невозможно точно восстановить уникальную последовательность кодирующего участка. Чтобы решить эту проблему, синтезируют смесь олигонуклеотидов, различающихся одним или несколькими нуклеотидами (рис. 6.24). Это не так трудно, как кажется. Вместо того чтобы использовать один защищенный мононуклеотид на определенном этапе химического синтеза, используют смесь защищенных мононуклеотидов.

*Тестирование на синтез специфического полипептида in vitro.* Иногда не удается провести прямой скрининг клонов потому, что просто отсутствуют

в достаточной степени очищенный подходящий зонд. В таком случае можно использовать непрямые методы, хотя обычно они неудобны при массовом скрининге клонированных популяций. Большинство обычно применяемых непрямых методов основаны на двух принципах. Во-первых, можно транслировать соответствующую мРНК *in vitro* и получить идентифицируемый полипептид. Во-вторых, можно использовать тот факт, что трансляция подавляется, если одноцепочечная мРНК ренатурирует с комплементарной клонированной ДНК. РНК, входящая в дуплекс РНК-ДНК, не способна функционировать как мРНК.



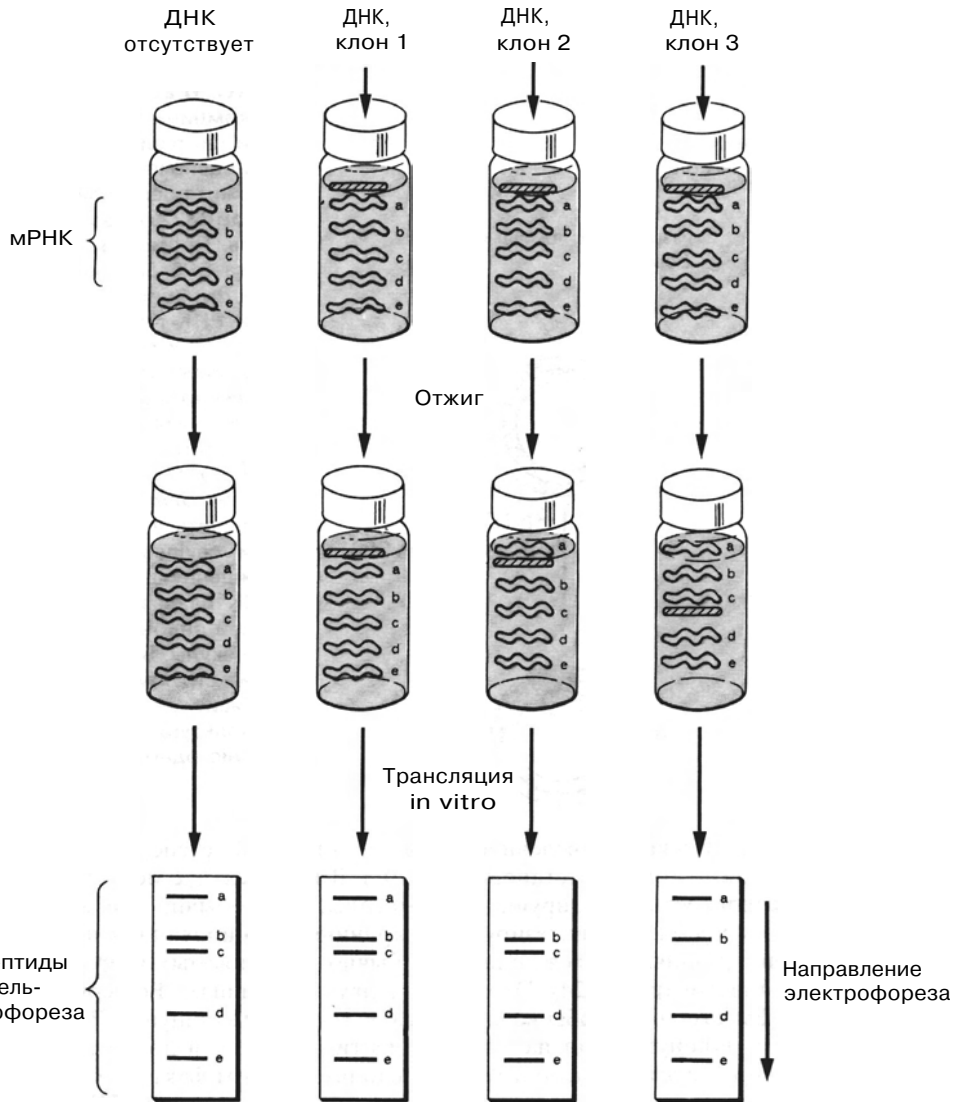


\* N – любой из четырех нуклеотидов

**РИС. 6.24.**

Данному пентапептиду соответствуют несколько нуклеотидных последовательностей в соответствующей области мРНК. Синтезировав химическим путем смесь олигодезоксирибонуклеотидов, комплементарных этим

мРНК, мы можем получить зонд для идентификации мРНК и клонов, содержащих сегменты, кодирующие данный пептид.



**РИС. 6.25.**

Метод HART (hybrid-arrested translation). Клоны 2 и 3 содержат молекулы ДНК, отвечающие молекулам

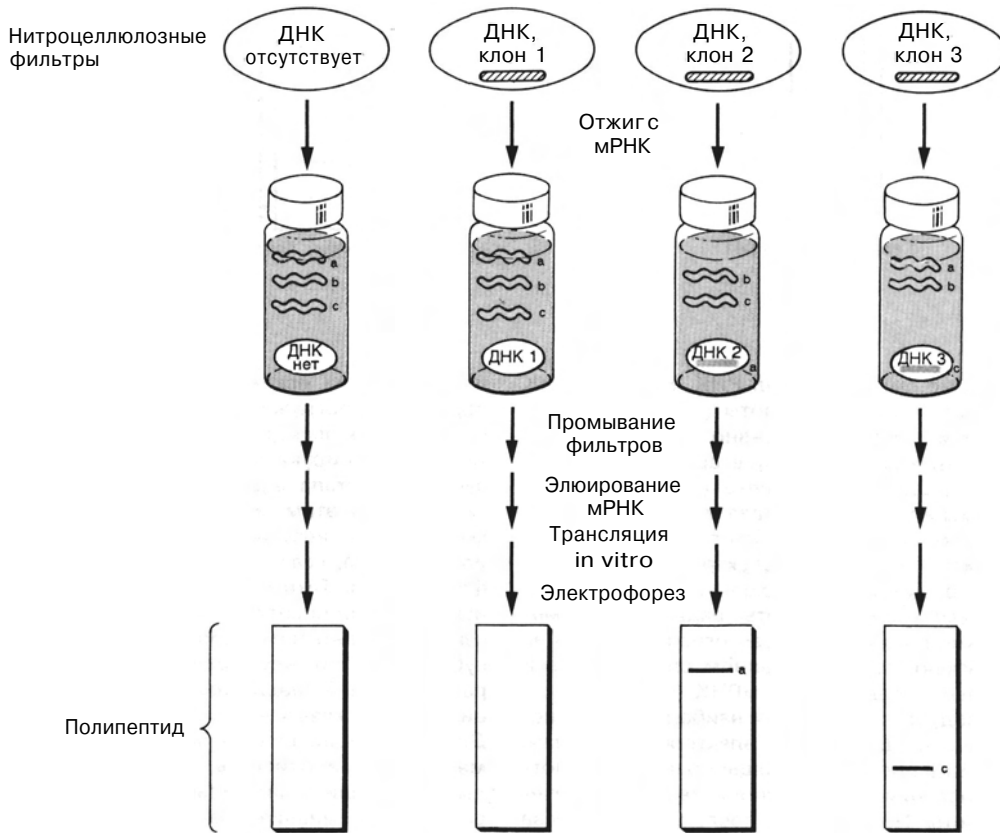
мРНК а и с соответственно (дополнительные данные о трансляции in vitro представлены на рис. 6.27).

Если проинкубировать смесь мРНК *in vitro* в присутствии необходимых компонентов, то на них синтезируются соответствующие полипептиды. Смесь этих полипептидов можно разделить по размерам с помощью электрофореза. Эффективная трансляция *in vitro* осуществляется только с одноцепочечных мРНК. Если же до начала трансляции какая-то мРНК ренатурировалась с комплементарной ДНК и образовался гибридный дуплекс ДНК-РНК, то информация с этой мРНК не будет считываться. С помощью процедуры, получившей название НАТ (от англ. hybrid-arrested translation; рис. 6.25), клонированные рекомбинантные ДНК очищают, денатурируют и отжигают (в растворе) с препаратами мРНК до начала трансляции *in vitro*. Молекулы ДНК, гомологичные определенному гену, гибридизуются с мРНК, и среди продуктов трансляции не обнаруживается соответствующий полипептид.

Второй метод идентификации рекомбинантных молекул с помощью трансляции *in vitro* получил название **гибридизационной селекции** (рис. 6.26). Из популяции рекомбинантов выделяют и очищают

рекомбинантные ДНК. Полученные препараты ДНК подвергают денатурации и каждый из них фиксируют на твердой подложке, например на нитроцеллюлозном фильтре. Каждый из препаратов инкубируют со смесью мРНК (в растворе). Гибридизация и связывание РНК на фильтре происходят только в тех случаях, когда РНК оказывается комплементарной иммобилизованной клонированной ДНК. Все остальные РНК удаляются с фильтра при промывании. Связанную РНК отделяют от ДНК и тестируют на синтез специфического полипептида в системе трансляции *in vitro*. Это позволяет идентифицировать соответствующий рекомбинант.

Для облегчения идентификации полипептидов, образующихся при трансляции смеси мРНК *in vitro*, используют разные методы. Если препарат мРНК обогатить специфической мРНК, то полипептид, кодируемый такой мРНК, иногда удается идентифицировать среди продуктов трансляции по его размеру и количеству (рис. 6.27, А). Когда используется сложная смесь мРНК и полипептидов, интересующий нас полипептид можно идентифици-



**РИС. 6.26.**

Гибридизационная селекция. Денатурированная ДНК из клонов 2 и 3 «отбирает» мРНК для полипептидов а и с

соответственно (дополнительные данные о трансляции *in vitro* представлены на рис. 6.27).

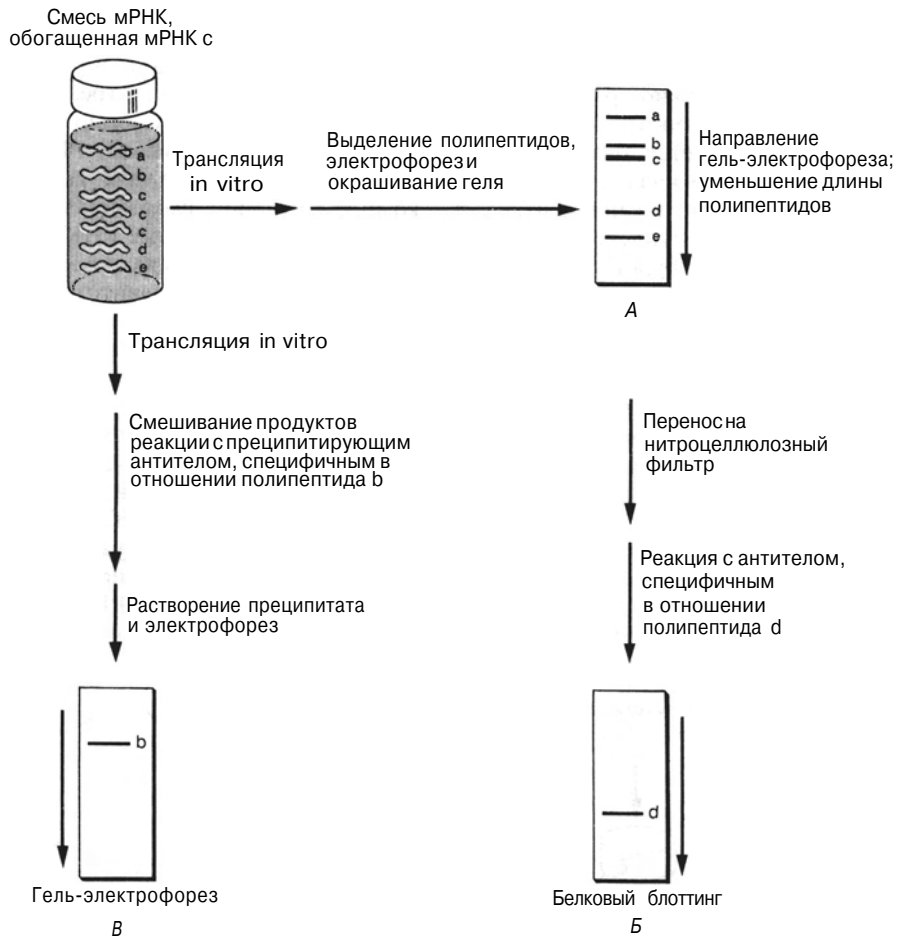


РИС. 6.27.

Трансляция *in vitro* смеси молекул РНК. Наряду с РНК в смеси присутствуют аминокислоты (иногда по крайней мере одна из них радиоактивно меченная), АТФ и клеточные экстракты, содержащие другие компоненты, необходимые для синтеза белка (например, ферменты, рибосомы и тРНК). А. Каждая мРНК дает начало соответствующему полипептиду. Эти полипептиды разделяют по размерам с помощью электрофореза и идентифицируют окрашиванием подходящим красителем. Если используемые аминокислоты радиоактивно меченные, то можно получить радиоавтограф, наложив на гель рентгеновскую пленку. В данном случае в смеси особенно обильно представлена мРНК с и соответственно среди продуктов трансляции наиболее многочислен полипептид с. Б. Анализ электрофореграммы, представленной на рис. А, с помощью белкового блоттинга. Для трансляции *in vitro* использовались меченные аминокислоты. Полипептиды после электрофореза были перенесены с геля на иммобилизованную подложку, где они фиксировались в результате адсорбции. Обычно для переноса используют электрооэлюцию, а не просто блоттинг. Процесс аналогичен рассмотренному

на рис. 6.2, за исключением того, что гель вместе с подложкой погружают в буферный солевой раствор между двумя электродами. Далее подложку с полипептидами инкубируют в растворе, содержащем соответствующие антитела, в данном случае антитела к полипептиду d. При этом можно использовать радиоактивно меченные или несущие флуоресцентную метку антитела или антитела, связанные с легко идентифицируемыми ферментами. Таким способом определяют положение искомого полипептида в исходном геле. Положение связанного антитела можно определить также, проинкубировав подложку с радиоактивно меченым препаратом белка *A Staphylococcus aureus*. Этот белок специфически связывается с иммуноглобулином IgG. В. Для трансляции *in vitro* использовали радиоактивно меченные аминокислоты. Образовавшиеся полипептиды смешивали с антителами к полипептиду b. Осадок растворяли и анализировали с помощью гель-электрофореза. Полученный радиоавтограф содержал только одну полосу, отвечающую радиоактивно меченному полипептиду b.

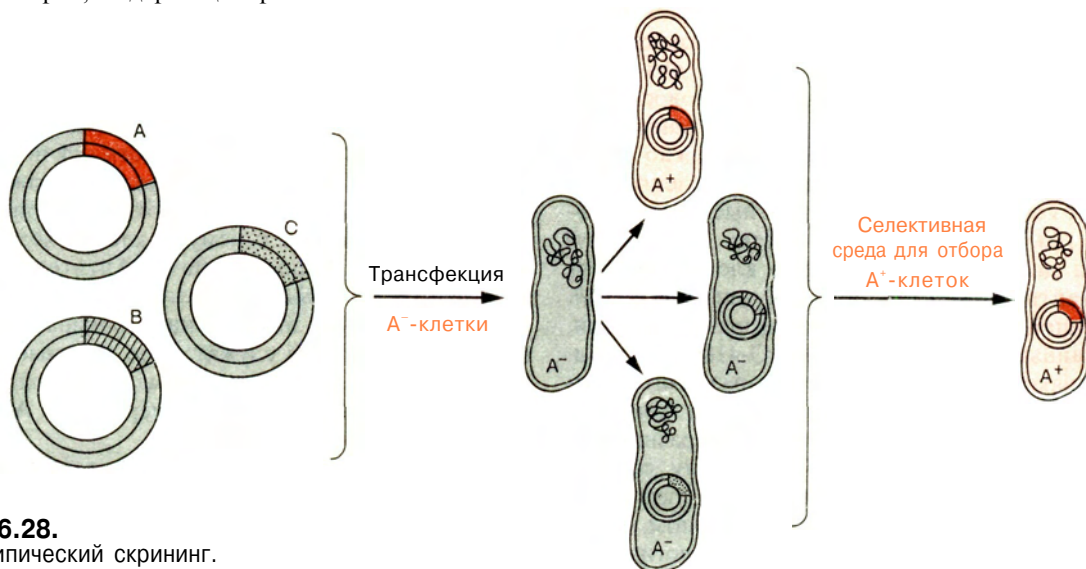
ровать по его способности взаимодействовать со специфическими антителами. В таких случаях применяют метод белкового блоттинга (вестерн-блоттинг), аналогичный блоттингу ДНК (рис. 6.27, Б). При другом подходе нужный полипептид осаждают в присутствии специфических антител из смеси продуктов трансляции *in vitro* (рис. 6.27, В), осадок растворяют и подвергают электрофорезу. В принципе в геле должны присутствовать только два белка—интересующий нас полипептид и иммуноглобулин.

#### в. Определение экспрессии гена в клетках

Экспрессирующийся ген, входящий в состав рекомбинантной ДНК, при определенных условиях придает клетке-хозяину специфические фенотипические свойства. Метод отбора, основанный на выявлении этого фенотипа, позволяет идентифицировать и выделять соответствующий клон точно так же, как стандартные селективные маркеры в векторных молекулах позволяют отбирать трансформированные клетки. К стандартным селективным маркерным системам, приведенным в гл. 5, относятся клетки *E. coli*, утратившие  $\beta$ -лактамазу (и поэтому чувствительные к ампициллину), которые используются для отбора трансформантов, содержащих векторы с геном *amp*, а также клетки млекопитающих, имеющие фенотип ТК<sup>-</sup>, которые позволяют отобрать трансформанты, содержащие векторы с геном тимидинкиназы. С помощью **фенотипического отбора** можно отбирать специфические клоны непосредственно из популяций бактерий, дрожжевых или животных клеток, трансформированных смесью рекомбинантных векторов, содержащих различные гены бакте-

рий, дрожжей или животных соответственно. Возможность проведения такого отбора зависит от наличия хозяйских клеток с определенным фенотипом, чаще всего клеток, дефектных в отношении продукта гена, который должен быть клонирован. Другие варианты фенотипического отбора, используемые при работе с клетками эукариот, описаны ниже и в разд. 5.7. Если отсутствуют нужные хозяйские клетки, то вместо фенотипического отбора можно использовать иммунологический скрининг. В таких случаях клоны отбирают с помощью антител, специфичных в отношении продукта данного гена.

**Фенотипический отбор.** Фенотипический отбор из популяции трансформированных клеток—это самый прямой путь выделения нужного клона (рис. 6.28). На практике этот метод, вообще говоря, ограничивается клонированием прокариотических генов и некоторых генов дрожжей и других грибов в клетках прокариот и эукариотических генов в клетках эукариот. В общем виде он был описан в разд. 5.6.б, когда мы рассматривали способ конструирования дрожжевого плазмидного вектора. Чтобы клонировать ген А (или кДНК-А), смесь фрагментов ДНК (т.е. фрагментов, полученных из всего клеточного генома) встраивали в векторные молекулы и трансфицировали ими клетки, мутантные по гену А ( $A^-$ ). Клетки выращивали в условиях, требующих присутствия продукта гена А, так что колонии могли образовывать лишь те клетки, в которых синтезируется продукт гена А, присутствующего в векторе. Наиболее подходящими для проведения такого отбора являются клетки-хозяева, в которых делетирован либо весь ген А, либо его часть. В противном случае приходится выявлять и разграничивать ревертанты и трансфор-



**РИС. 6.28.**  
Фенотипический скрининг.

манты. Кроме того, делеционные мутанты исключают вероятность проявления фенотипа  $A^+$  в результате супрессии фенотипа  $A^-$  какими-то другими генами.

*Экспрессия эукариотических генов в бактериальных системах хозяин-вектор.* Ни фенотипический отбор, ни иммунологический скрининг обычно неприменимы при выделении клеток *E. coli*, содержащих определенный рекомбинантный эукариотический ген. Исключение составляют некоторые гены дрожжей и других грибов, способные экспрессироваться в клетках *E. coli*. Однако большинство генов эукариот не могут должным образом экспрессироваться в клетках *E. coli*, поскольку они утратили функциональные промоторы и содержат некодирующие участки в составе кодирующих последовательностей (гл. 8). Другое дело, если рекомбинанты содержат кДНК-копии эукариотических мРНК, поскольку некодирующие участки выщепляются во время созревания мРНК в эукариотических клетках. Клетки *E. coli* не способны осуществлять такой сплайсинг. В результате кДНК транслируются в правильные полипептиды в *E. coli* (рис. 6.29). Для того чтобы клетки *E. coli* смогли обеспечить транскрипцию и трансляцию клонированной эукариотической кДНК, в соответствующие сайты вектора обычно включают *E. coli*-промотор и сигналы трансляции. Подобным же образом, для того чтобы кДНК эукариот могла экспрессироваться в эукариотических клетках, вектор необходимо снабдить соответствующими сигналами регуляции транскрипции. Поскольку в норме геномные регуляторные сигналы обычно находятся вне транскрибируемой части гена, в 5'- или 3'-фланкирующей области, они отсутствуют в мРНК или кДНК.

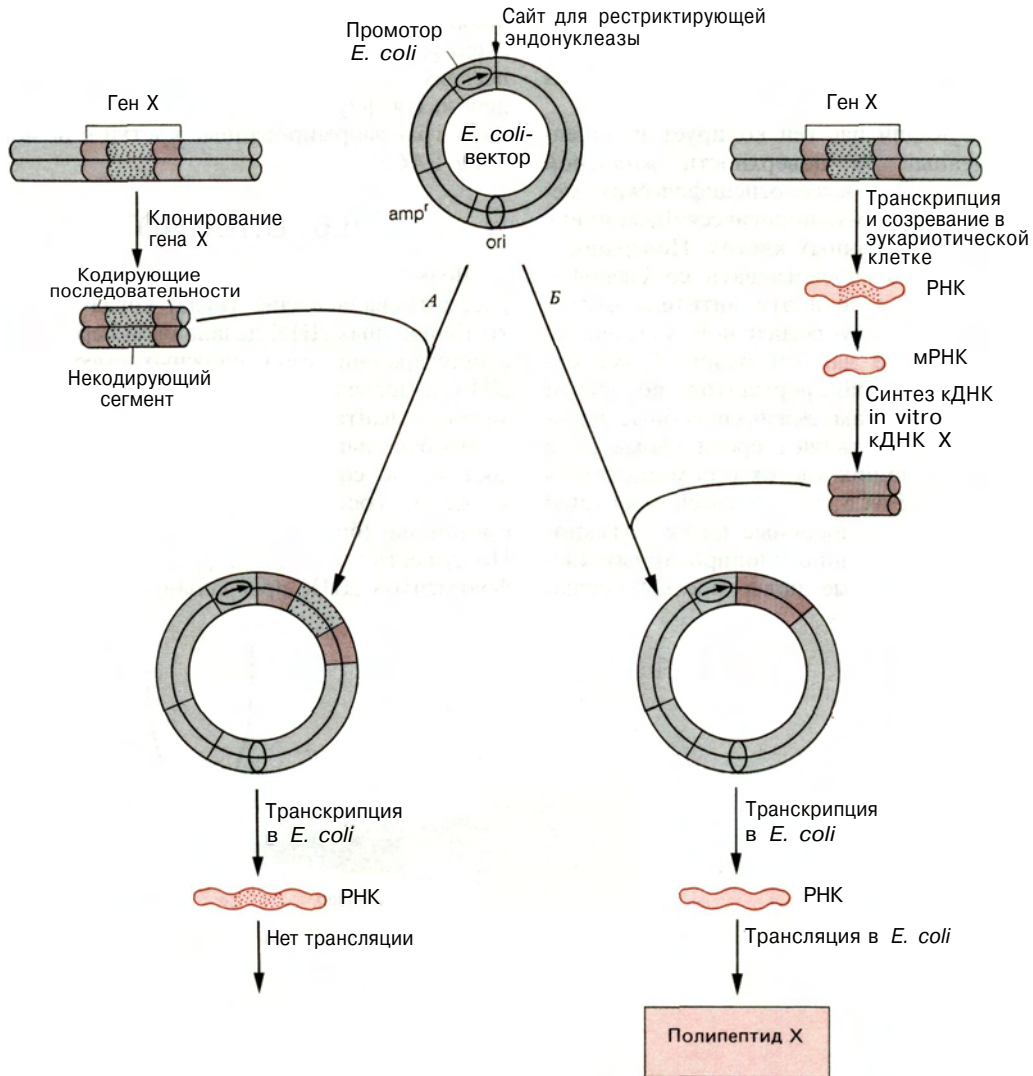
*Другие методы фенотипического скрининга в эукариотических клетках.* Фенотипический скрининг редко применяется при работе с диплоидными эукариотическими клетками, поскольку для этого необходима линия клеток, у которых повреждены оба аллеля интересующего нас гена. Идентифицировать трансформированные клетки можно с помощью котрансформации, применяя селективный маркер и соответствующие мутантные клетки (например, ген тимидинкиназы и клетки ТК<sup>-</sup> грызунов) или доминантный маркер и клетки, ингибируемые антибиотиком (например, ген *neo* *E. coli*, кодирующий аминокликозидтрансферазу, и клетки, обработанные аминокликозидом G-418; см. разд. 5.17). Однако при этом невозможно различить нужный котрансформант и сопутствующие контрансформированные колонии, содержащие разные рекомбинантные молекулы ДНК. Поэтому обычно скрининг основывается на использовании специфических свойств, проявляемых нужным клоном. В качестве примера можно привести морфологические

изменения, наблюдаемые при превращении нормальных клеток в опухолевые, который описан в разделах, посвященных векторам на основе папилломавирусов крупного рогатого скота и ретровирусов (разд. 5.7). Клетки, трансформированные к онкогенному фенотипу, образуют характерную колонию, или **фокус**, в отличие от обычных клеток, образующих монослой. Этот метод использовали, например, при отборе рекомбинантов, содержащих клеточные онкогены из геномной ДНК животных.

Другой метод скрининга применяют при клонировании клеток животных, в которых экспрессируется рекомбинантный ген (или кДНК), кодирующий секретлируемый полипептид. При этом используют клетки хозяина, в норме не образующие полипептид, и тестируют среду, в которой находятся трансформированные клетки, на присутствие в ней данного полипептида. Если продуктом является гормон, то его присутствие в среде можно обнаружить с помощью какого-либо удобного высокочувствительного биологического метода. Таким же способом были клонированы гены, кодирующие факторы роста, которые стимулируют пролиферацию специфических клеток-мишеней. Если активность полипептида как такового трудно измерить, то используют специфичные к нему антитела. Этот метод непригоден для одновременного скрининга большой популяции клонированных клеток, содержащих разные рекомбинантные молекулы, поскольку для этого пришлось бы проводить слишком много (возможно, миллион) отдельных тестов. Вместо этого смесь трансформированных клеток подразделяют на удобное число отдельных групп и тестируют эти группы. Группу, дающую положительный ответ, снова подразделяют на подгруппы, вновь выделяют позитивную подгруппу, подразделяют ее, тестируют и так далее до тех пор, пока не будет идентифицирован один позитивный клон.

*Иммунологический скрининг.* Для скрининга рекомбинантной популяции в отношении экспрессии данного гена используют более общий метод, в основе которого лежит иммунологический подход. Синтезируемый белок в этом случае не должен быть функционально активным; необходимы лишь специфические антитела к нему и клетки хозяина, не синтезирующие данный антиген. Антитела взаимодействуют с белковыми молекулами одного клона. Их выявляют по аккумуляции радиоактивности, источником которой служат либо меченые антитела, либо специфические меченые реагенты, узнающие иммуноглобулины (например, белок *A Staphylococcus aureus*).

Было разработано несколько модификаций техники скрининга с участием антител применительно к *E. coli*-системам хозяин-вектор. Все они сходны между собой, хотя применяемые приемы зависят,

**РИС. 6.29.**

Экспрессия эукариотических генов в *E. coli*-системах хозяин-вектор. Сегмент эукариотической ДНК встроен вблизи промотора *E. coli*, чтобы обеспечить его транскрипцию в клетках бактерий. Эукариотические гены (такие, как показанный здесь ген X) обычно являются прерывистыми из-за наличия в них некодирующих сегментов (на рисунке они выделены точками), поэтому РНК-транскрипт гена X, образующийся в клетках *E. coli*, не способен транслироваться в правильный полипептид

(А). При синтезе функциональной мРНК в эукариотических клетках некодирующие участки подвергаются сплайсингу (гл. 8), в результате чего кДНК, синтезированная на мРНК соответствующего гена X, утрачивает некодирующую последовательность. При встраивании такой кДНК в бактериальный вектор и введении получившейся рекомбинантной молекулы в клетки *E. coli* образуется РНК, способная транслироваться в правильный продукт гена X (Б).

например, от того, какого рода клоны—бактериальные или фаговые—будут тестироваться. В одной из модификаций этого метода антитела равномерно распределяют по поверхности твердого носителя, например пластикового или бумажного диска, и затем прижимают его к агаровому слою, содержащему лизированные колонии или фаговые бляшки. Если специфический антиген синтезируется каким-

либо клоном, он будет связываться с антителом на диске. Антиген можно локализовать, обработав снятый с агара диск другими радиоактивно меченными (обычно с помощью  $^{125}\text{I}$ ) антителами, промыв диск и получив радиоавтограф. Эта процедура зависит от способности двух антител (одного—покрывающего диск и второго— $^{125}\text{I}$ -меченного) связываться с разными антигенными детерминан-

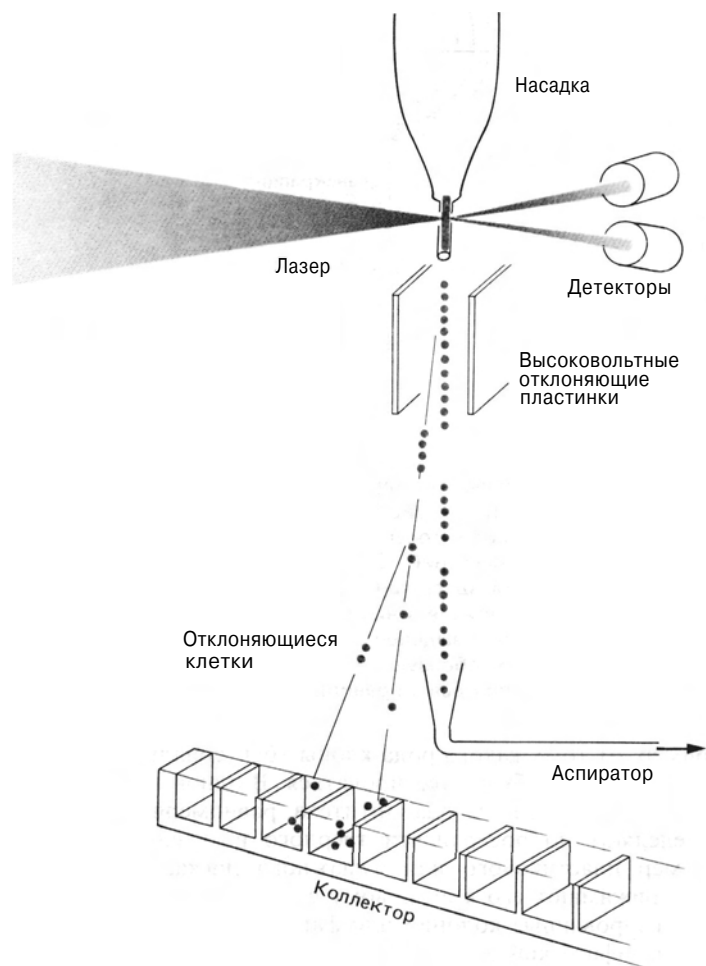
тами на одном и том же полипептиде. Как и при скрининге путем отжига с ДНК- или РНК-зондами, позитивный клон дает четкое темное пятно на рентгеновской пленке.

Если интересующий нас ген кодирует полипептид, расположенный на поверхности животной клетки, то с помощью высокоспецифических методов можно провести иммунологический скрининг и отбор трансформированных клеток. Поверхностные белки могут взаимодействовать со специфическими антителами, а если эти антитела связываются с какой-либо флуоресцентной меткой, то клетки, синтезирующие данный белок, будут отличаться от других трансформантов по своим флуоресцентным свойствам. Жизнеспособные флуоресцирующие клетки выявляют среди громадного числа нефлуоресцирующих клеток с помощью специального прибора (FACS, от fluorescence-activated cell sorter). Далее жизнеспособные клетки размножают и получают популяцию клонированных клеток (рис. 6.30). Некоторые поверхностные белки

являются рецепторами, с которыми связываются специфические лиганды (например, инсулин). Если исследуемый ген кодирует такой рецептор, то, используя флуоресцирующий лиганд, можно отделить трансформированные клетки с помощью прибора FACS.

## 6.5. БИБЛИОТЕКИ

Возможность проведения шотган-экспериментов рассматривалась еще тогда, когда технология рекомбинантных ДНК делала свои первые шаги. Идея амплификации очень сложных смесей фрагментов ДНК и последующего обнаружения одного из них казалась фантастичной, а сейчас она представляется почти банальной. Основной принцип этого подхода состоит в создании коллекции рекомбинантных молекул, составляющих полный геном данного организма (**библиотеки рекомбинантных молекул**). По существу, полный нефракционированный набор фрагментов ДНК превращают в соответствующий



**РИС. 6.30.**

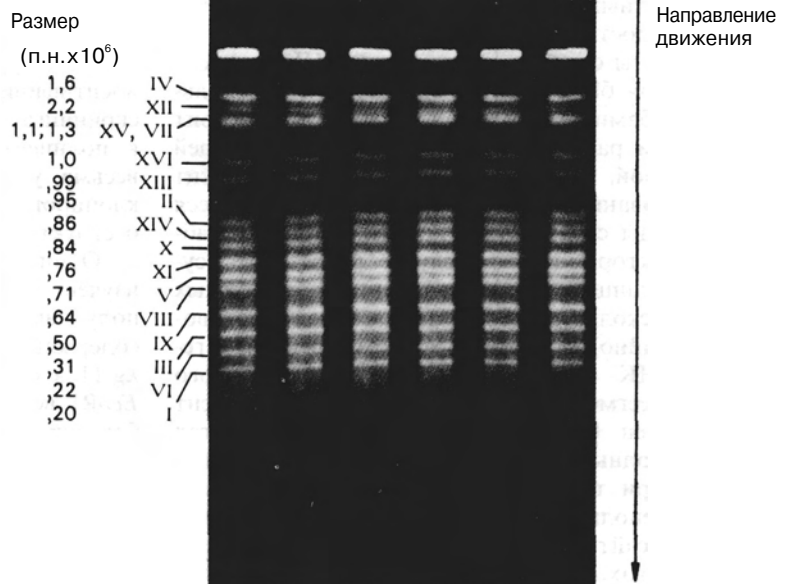
Схематическое изображение устройства для отбора флуоресцирующих клеток. Популяцию клеток обрабатывают антителами, специфичными к определенному поверхностному белку. Антитела связаны с флуоресцирующим красителем, так что клетки, несущие антиген, становятся способными к флуоресценции. Клетки одну за другой пропускают через узкую насадку, помещенную между лазером и детектором флуоресценции. Клетки, интенсивность флуоресценции которых превышает некий фиксированный уровень, отклоняются и попадают в ячейки для культивирования, где они размножаются и могут быть клонированы. Нефлуоресцирующие клетки собираются с помощью аспиратора. Таким методом можно осуществить скрининг  $\sim 10^5$  клеток за 1 с. [D.R. Parks et al., Proc. Natl. Acad. Sci. U.S.A., **76** (1979), p. 1962.]

набор стабильных рекомбинантов, который можно сохранять и многократно использовать для клонирования различных вставок. При этом применяются *E. coli*-системы хозяин-вектор, где в качестве вектора используются плазмиды или бактериофаги, поскольку они более пригодны для получения большого числа рекомбинантных молекул и удобны для хранения.

Наиболее важны два типа библиотек. Одна из них, создаваемая из суммарной геномной ДНК, в принципе должна содержать все гены данного организма. Однако такая цель обычно оказывается недостижимой, поскольку некоторые последовательности ДНК не удается клонировать. Один из вариантов геномной библиотеки представляет собой всю ДНК какой-то одной хромосомы. Для создания такой библиотеки необходимо выделить ДНК из отдельных хромосом. Хромосомы человека фракционируют с помощью методов, аналогичных используемым при сортировке флуоресцирующих и нефлуоресцирующих клеток. Метод основан на разной эффективности связывания флуоресцирующих красителей с хромосомами, которая зависит от размера хромосом и GC-содержания ДНК. Раствор окрашенных хромосом пропускают с высокой скоростью через узкое отверстие, на которое сфокусирован пучок света, испускаемого лазером. Хромосомы поочередно пересекают этот пучок, который индуцирует их специфическую флуоресценцию. Детектор определяет интенсивность флуоресценции и изменяет направление движения тех хромосом, интенсивность флуоресценции которых превышает заданную величину, таким образом, что они попадают в специальный коллектор. Изменяя этот

заранее установленный уровень интенсивности флуоресценции, можно отбирать разные хромосомы.

Отдельные хромосомы дрожжей и некоторых простейших невозможно идентифицировать цитогенетическими методами (т.е. с помощью окрашивания и микроскопирования), однако ДНК из них все же удается выделить. Для выделения применяют мягкие методы, чтобы избежать разрыва молекул. Затем ДНК подвергают электрофорезу в агарозном геле в условиях, позволяющих разделить молекулы длиной до 2000 т.п.н. (рис. 6.31). Традиционные методы электрофореза не позволяют разделять дуплексные молекулы, размер которых значительно превышает 20 т.п.н. Скорость миграции столь больших молекул уже не зависит от их размера. Однако если вместо постоянного однонаправленного электрического поля, приложенного к обычному гелю, использовать поле, ориентация которого многократно меняется, то даже очень большие молекулы можно будет разделить по размерам. По-видимому, этот феномен объясняется (по крайней мере частично) механизмом прохождения молекул ДНК через поры агарозного геля. Предполагается, что молекулы вытягиваются в направлении поля, а затем при изменении этого направления переориентируются. Время переориентации зависит от длины цепи и угла между направлениями поля; они и определяют конечное расстояние, на которое перемещается молекула. В простейшем варианте **импульсного электрофореза** электрический ток подается импульсами, при этом направления поля примерно перпендикулярны друг другу, а само поле неоднородно. При более сложных распределениях



**РИС. 6.31.**

Импульсный электрофорез. Приведена фотография геля, окрашенного бромистым этидием. Хромосомную ДНК дрожжей (*S. cerevisiae*) подвергли электрофорезу в течение 24 ч в условиях однородного поля (CHEF hexagonal array apparatus). Направление поля изменяли на 120° каждые 80 с. Дрожжевые хромосомы (они обозначены римскими цифрами) идентифицировали путем гибридизации с предварительно картированными зондами. Слева указаны их размеры в миллионах пар нуклеотидов. [G. Chu, D. Vollrath, R.W. Davis, Science, **234** (1986) p. 1582.]



используют однородные поля и оптимизируют угол между двумя направлениями, с тем чтобы повысить разрешение. Импульс обычно длится примерно минуту.

Библиотеки второго типа включают последовательности, составляющие всю мРНК, обнаруживаемую в определенных клетках. В этом случае популяцию мРНК превращают в популяцию молекул кДНК, которые затем клонируют. **Геномные библиотеки** представляют собой собрание генов и последовательностей ДНК; в **библиотеках кДНК** представлены продукты экспрессии этих генов в форме мРНК.

### а. Геномные библиотеки

Геномные библиотеки обычно создают с помощью векторов, сконструированных на основе бактериофага  $\lambda$  или космиды (рис. 6.32). Эти векторы содержат большие вставки, благодаря чему минимизируется число рекомбинантов, необходимых для составления библиотеки. Например, если создается  $\lambda$ -библиотека генома млекопитающего, содержащего  $3 \cdot 10^9$  п.н. (на гаплоид) при средней длине вставки 17 т.п.н., то весь геном будет представлен  $3 \cdot 10^9 / 1,7 \cdot 10^4 = 1,8 \cdot 10^5$  рекомбинантами. На самом деле для создания библиотеки, вероятность обнаружения в которой определенного сегмента генома превышает 99%, нужно получить  $\sim 10^6$  отдельных рекомбинантных молекул, поскольку лигирование отдельных фрагментов происходит случайно. Так, некоторые фрагменты могут быть включены более чем в одну векторную молекулу, а другие могут вообще не участвовать в лигировании и упаковке. Полная космидная библиотека содержит меньше рекомбинантных молекул, поскольку размер вставки может достигать 45 т.п.н.

Фрагменты суммарной геномной ДНК для конструирования библиотек можно получать несколькими способами. Наиболее удобный из них состоит в частичном расщеплении ДНК рестриктирующей эндонуклеазой, сайт узнавания которой содержит шесть оснований (или более), а образующиеся липкие концы соответствуют липким концам выбранного вектора. В этом случае разрезание осуществляют лишь в ограниченном числе возможных сайтов. Поскольку выбор таких сайтов производится случайно и в используемом препарате геномной ДНК содержится много копий любого геномного сегмента, практически каждый сегмент ДНК должен быть представлен во фрагментах ДНК, пригодных по своему размеру для клонирования. При таком подходе, однако, может получиться неполная библиотека. Те части генома, в которых сайты узнавания рестриктирующих эндонуклеаз находятся слишком далеко друг от

друга, не смогут включиться в жизнеспособный рекомбинантный фаг. С другой стороны, те области генома, в которых сайты узнавания тесно сгруппированы, будут разделены на очень короткие фрагменты, и не все из них будут представлены в библиотеке.

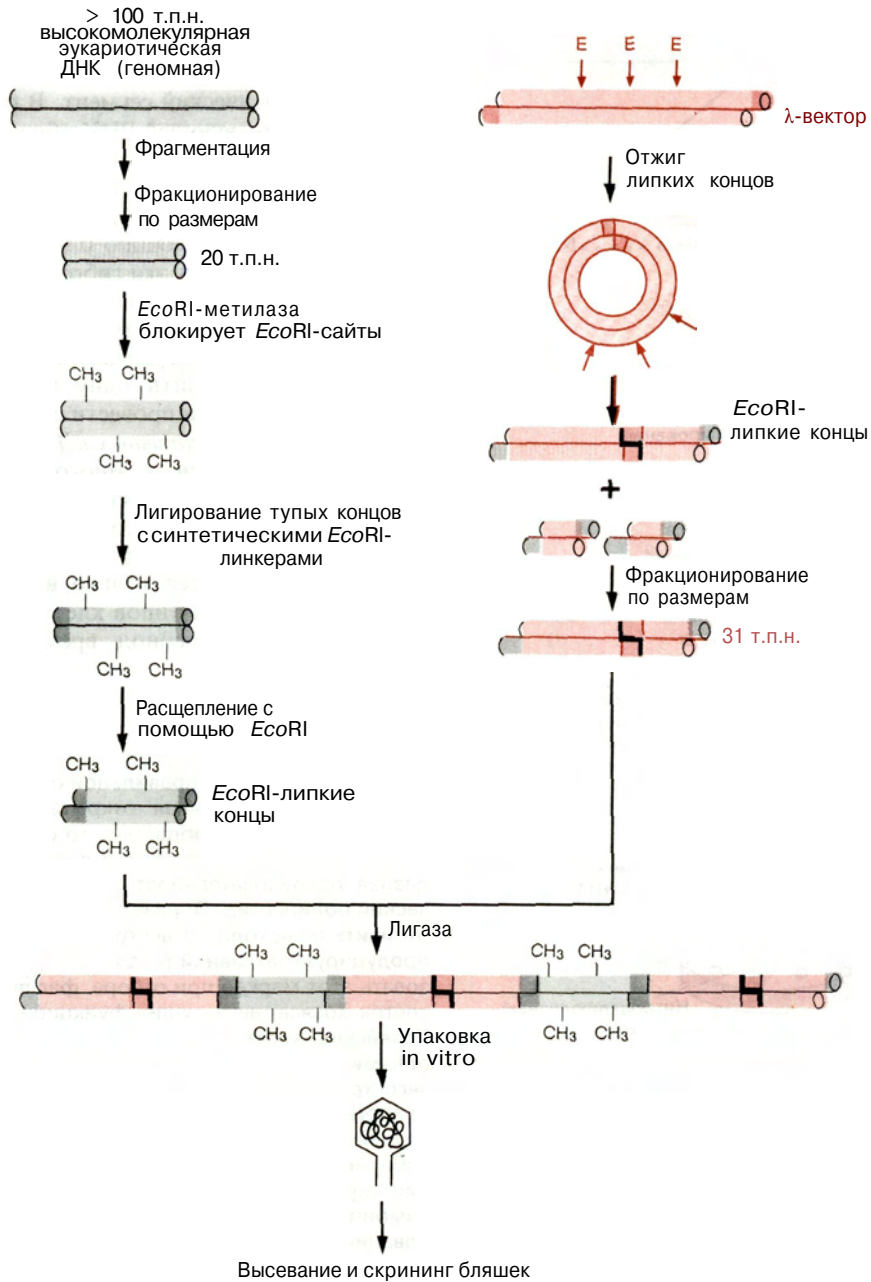
Вообще говоря, более репрезентативную библиотеку можно получить при частичной фрагментации геномной ДНК, осуществляемой более случайно, чем с помощью фермента с сайтом узнавания из шести пар оснований. Для этого можно использовать гидродинамические методы деградации ДНК или очень ограниченную обработку эндонуклеазами с сайтами узнавания из четырех пар оснований, например *Alu I* и *Hae III*.

### б. Библиотеки кДНК

Для создания библиотек кДНК обычно используют плазмидные или фаговые векторы, сконструированные на основе бактериофага  $\lambda$ . Принципы конструирования не отличаются от описанных в разд. 6.1.г, за исключением того, что в этих случаях чаще используют смесь разных мРНК, а не очищенную мРНК. Полученные с помощью специально сконструированных векторов или определенным образом отобранные библиотеки кДНК могут использоваться для клонирования последовательностей мРНК даже в тех случаях, когда аминокислотная последовательность белка и кодирующая нуклеотидная последовательность неизвестны и даже отсутствует гомологичный зонд. Соответствующие примеры мы рассмотрим ниже.

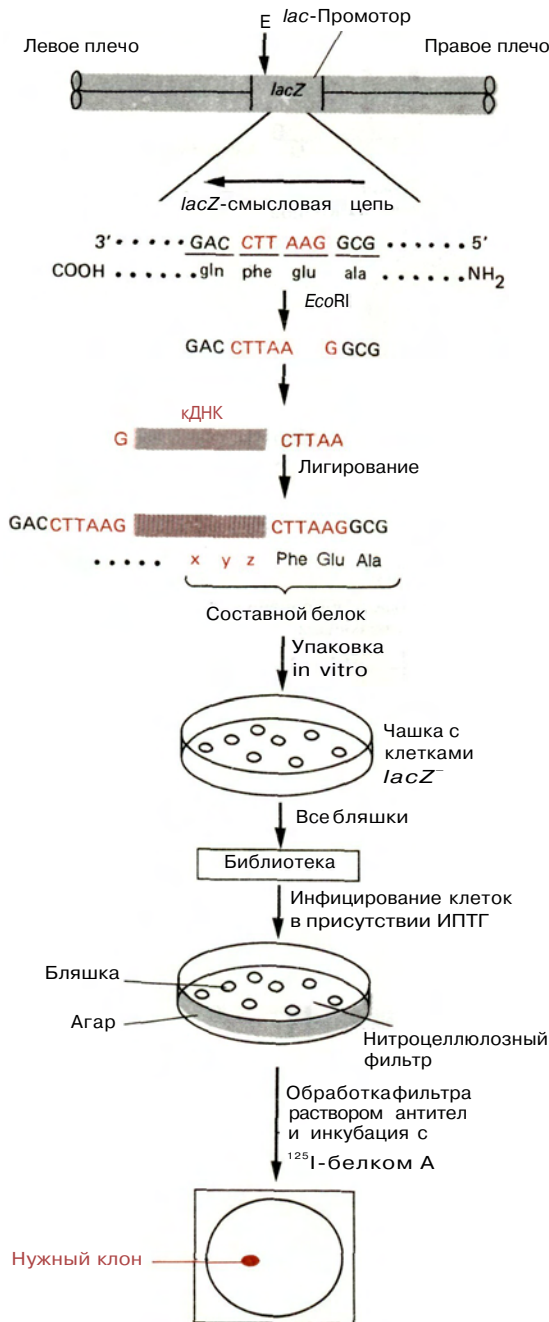
*Библиотека экспрессирующихся кДНК.* Очищенную кДНК или смесь разных кДНК можно лигировать с векторами, специально сконструированными для осуществления транскрипции и трансляции кодирующей области кДНК. Нужный клон идентифицируют с помощью иммунологического скрининга с применением антител, специфичных к полипептиду, кодируемому данной кДНК. Эта весьма удобная методика позволяет осуществить клонирование даже тогда, когда ничего не известно о структуре нужного нам гена или белка.

Одним из векторов, широко применяемых для изучения экспрессии, является производное фага  $\lambda$ , получившее название  $\lambda$ gt11 (рис. 6.33). Когда кДНК, содержащие *EcoRI*-линкеры, включаются в вектор  $\lambda$ gt11 в его единственном сайте для эндонуклеазы *EcoRI*, вставки оказываются в области, кодирующей бактериальный ген  $\beta$ -галактозидазы (*lacZ*). Примерно одна из шести кДНК-вставок включается в подходящей для транскрипции ориентации и в фазе с рамкой считывания  $\beta$ -галактозидазы. (Копирование мРНК с образованием дуплексной кДНК редко происходит безошибочно, и лишь одна примерно из

**РИС. 6.32.**

Создание геномной библиотеки. Высокомолекулярную геномную ДНК фрагментируют путем частичного гидролиза рестриктирующей эндонуклеазой или с помощью физических методов. Все другие этапы описаны в гл. 5 и 6. Отметим, что способ получения «плечевых» последовательностей λ-вектора отличается от представленного на рис. 5.11. В данном случае векторную ДНК

сначала переводят в кольцевую форму путем воссоединения *cos*-сайтов. Затем несущественную центральную часть отделяют от «плечей» с помощью рестриктирующей эндонуклеазы и проводят фракционирование. Далее плечевые последовательности соединяют со вставками с помощью липких *EcoRI*-концов. [T. Maniatis et al., *Cell*, **15** (1978), p. 687.]



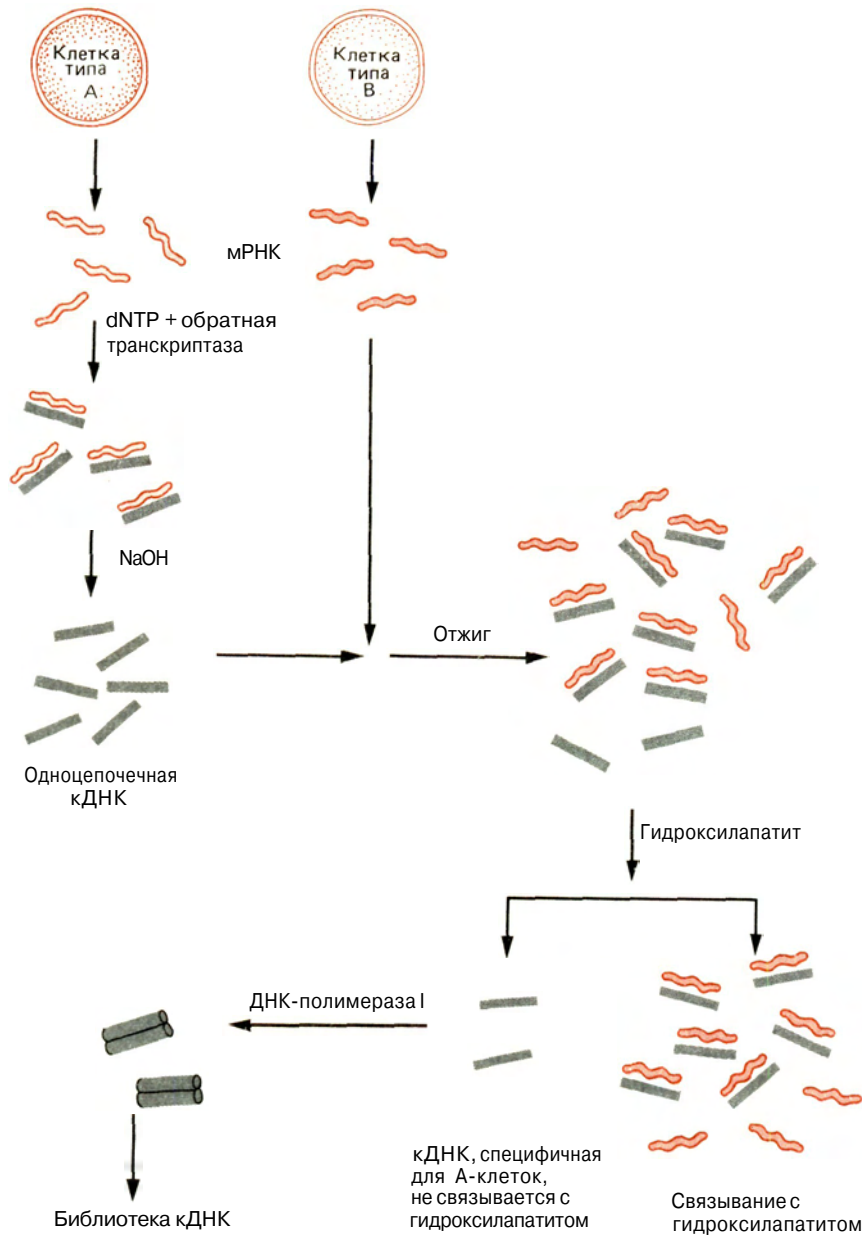
**РИС. 6.33.**

Клонирование кДНК в векторе  $\lambda$ gal11 - производном фага  $\lambda$ . Длина векторной ДНК составляет 44 т.п.н. ДНК содержит копию гена  $\beta$ -галактозидазы (*lacZ*) *E. coli*, вблизи 3'-конца которого находится сайт для эндонуклеазы *Eco*RI; кДНК с *Eco*RI-линкерами встраиваются в этот сайт, и рекомбинантные молекулы упаковываются в  $\lambda$ -частицы *in vitro*. При встраивании нарушается целостность кодирующей области *lacZ*, в результате чего активная  $\beta$ -галактозидаза не образуется. Однако если

трех молекул кДНК начинает синтезироваться в правильном сайте.) При индукции  $\beta$ -галактозидазного гена с помощью  $\beta$ -галактозида в *lac*-промоторе начинается транскрипция, которая распространяется на эукариотический сегмент. В результате трансляции соответствующей РНК образуется гибридный белок, у которого на N-конце находится  $\beta$ -галактозидазный полипептид, а на С-конце — эукариотический. Библиотека кДНК, созданная с помощью  $\lambda$ gt11, образует популяцию бляшек, отдельные члены которой содержат эти гибридные белки. Бляшку, содержащую интересный нас белок (и, следовательно, кДНК), идентифицируют по ее способности связывать соответствующее антитело. На одной чашке можно выявить одну позитивную бляшку среди  $10^4$  и без труда провести скрининг  $10^6$  бляшек.

**Библиотеки селективных кДНК.** Развитие и дифференцировка сложных многоклеточных организмов из одной оплодотворенной яйцеклетки осуществляются с помощью высокоорганизованной системы дифференцированной экспрессии генов. Одни гены экспрессируются только в одном или строго ограниченном числе типов клеток, другие — в какой-то определенный период времени. В результате

кДНК встраивается в правильной ориентации и при этом образуется непрерывная открытая рамка считывания с кодирующей *lacZ*-областью, то синтезируется составной белок, на N-конце которого находится  $\beta$ -галактозидазная последовательность, а на С-конце — эукариотический полипептид. Образующиеся рекомбинанты легко отличить от исходного вектора  $\lambda$ gt11, поскольку они не продуцируют активной  $\beta$ -галактозидазы. Чтобы использовать этот маркер при отборе, фаг высевают на газоне клеток-хозяев, не несущих функционального гена *lacZ*. Хозяйские клетки содержат также множество копий плазмиды с геном *lac*-репрессора (*lacI*), в результате чего транскрипция гена *lacZ* в  $\lambda$ gt11 подавляется. Это предотвращает потерю фаговых частиц из-за вредного воздействия составного белка на клетки или фаг. Далее собирают фаговые частицы из бляшек и создают библиотеку. Для выделения и клонирования нужного рекомбинанта часть библиотеки используют для инфицирования свежих клеток. Бляшки получают на нитроцеллюлозном фильтре, помещенном поверх агара, в присутствии изопропилтио- $\beta$ -D-галактопиранозида (ИПТГ) - индуктора транскрипции гена *lacZ*, в результате чего синтезируется составной белок. Затем нитроцеллюлозный фильтр обрабатывают раствором, содержащим антитела к интересующему нас полипептиду, и после промывания идентифицируют бляшки, содержащие антитела, по способности антител к IgG связываться с <sup>125</sup>I-меченным белком А из клеток *Staphylococcus aureus*. Помеченные таким образом бляшки визуализируют с помощью радиоавтографии и отбирают из агара.

**РИС. 6.34.**

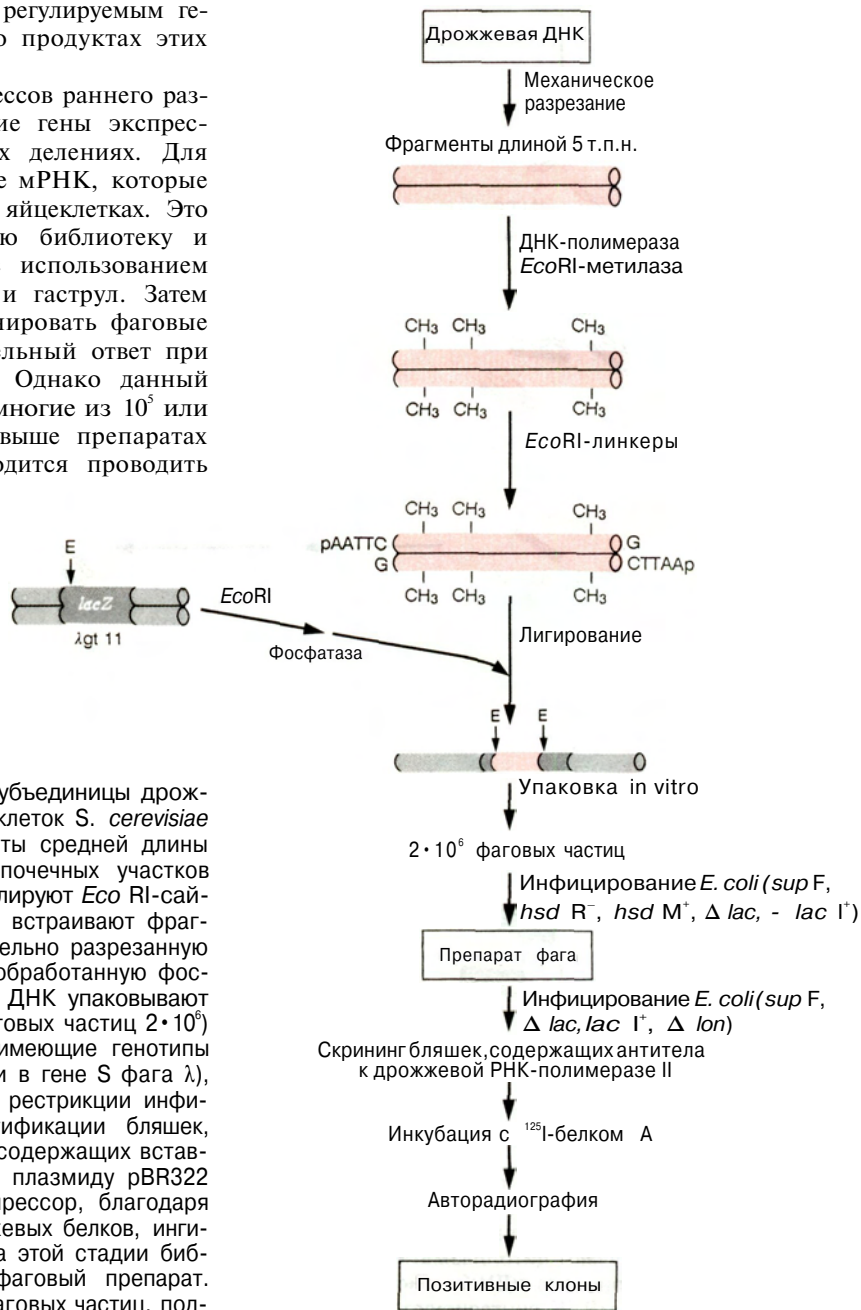
Создание библиотеки кДНК, представляющих собой копии мРНК, специфичных для определенного типа клеток или тканей. Из клеток или тканей двух родственных типов, А и В, получают цитоплазматическую мРНК. Специфическую библиотеку кДНК конструируют для клеток типа А. На А-мРНК синтезируют одноцепочечную кДНК, которую затем отжигают с избыточным количеством В-мРНК, в результате чего образуются дуплексные гибриды комплементарных А-кДНК и В-мРНК. А-кДНК, специфичные только для клеток А, остаются одноцепочечными и могут быть отделены от дуплексных гибридных молекул и оставшейся РНК с

помощью гидроксилапатита, который при соответствующих условиях связывает гибриды и свободные РНК, но не одноцепочечную РНК (разд. 7.5.б). Затем одноцепочечную А-кДНК используют в качестве матрицы для синтеза второй цепи кДНК и получают библиотеку дуплексных кДНК. Эта библиотека представлена почти исключительно молекулами кДНК-копиями мРНК, специфичными для клеток типа А. В качестве альтернативы обогащенную кДНК используют как зонд для гибридизации с полной библиотекой кДНК, полученной на А-мРНК, для определения клонов кДНК, соответствующих специфичной мРНК клеток типа А.

популяции цитоплазматических мРНК в различных тканях и клетках оказываются представленными разными молекулами. Такая дифференцированная экспрессия генов может использоваться для клонирования кДНК, соответствующих регулируемым генам, даже в тех случаях, когда о продуктах этих генов ничего не известно.

Например, при изучении процессов раннего развития *Xenopus* важно знать, какие гены экспрессируются при первых клеточных делениях. Для этого нужно идентифицировать те мРНК, которые присутствуют в гастрале, но не в яйцеклетках. Это можно сделать, создав геномную библиотеку и проведя раздельный скрининг с использованием препаратов РНК из яйцеклеток и гастрал. Затем можно идентифицировать и клонировать фаговые бляшки, которые дают положительный ответ при отжиге только с РНК гастрал. Однако данный метод имеет тот недостаток, что многие из  $10^5$  или более мРНК в двух указанных выше препаратах являются одинаковыми и приходится проводить

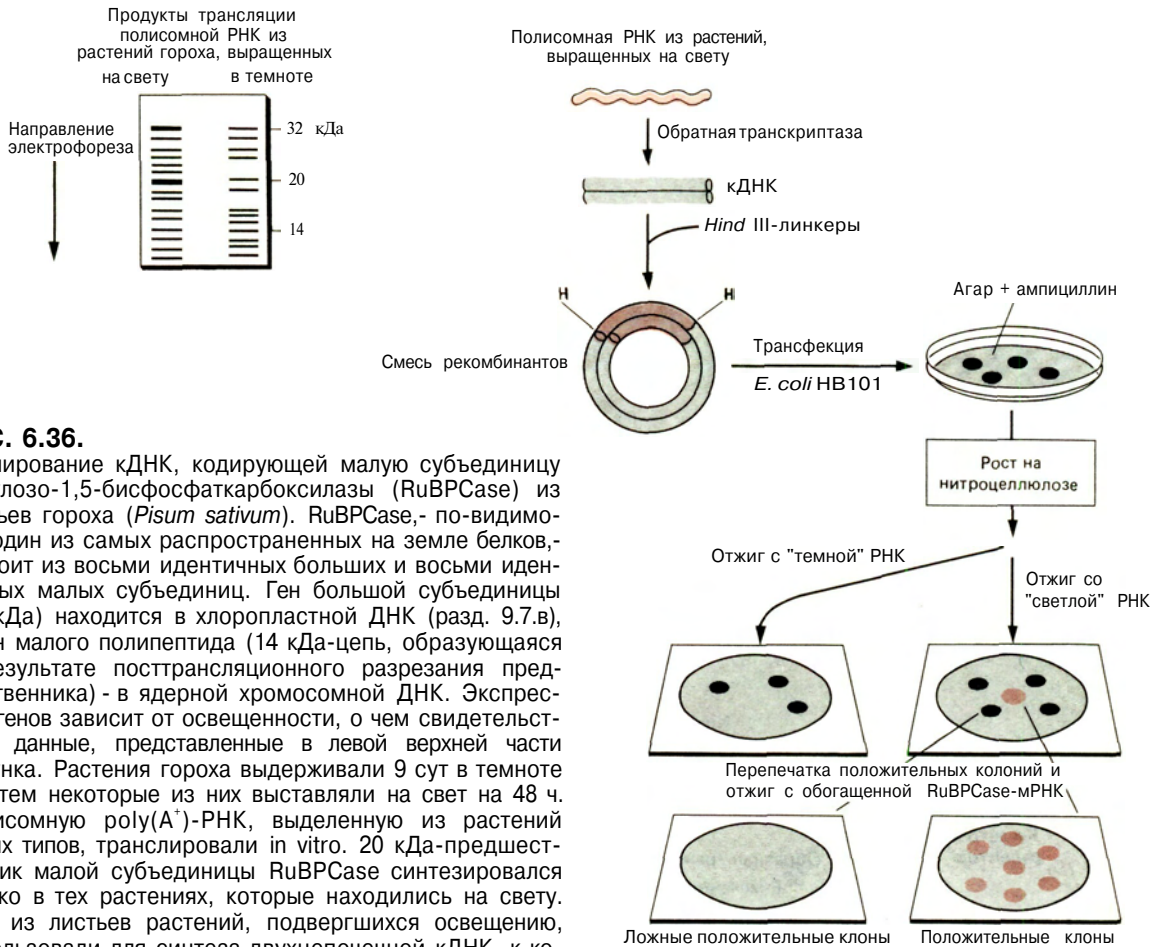
скрининг громадного числа бляшек, чтобы обнаружить те немногие, которые проявляют специфичность. Более эффективным является создание библиотеки кДНК на основе гастральной РНК, кото-



**РИС. 6.35.**

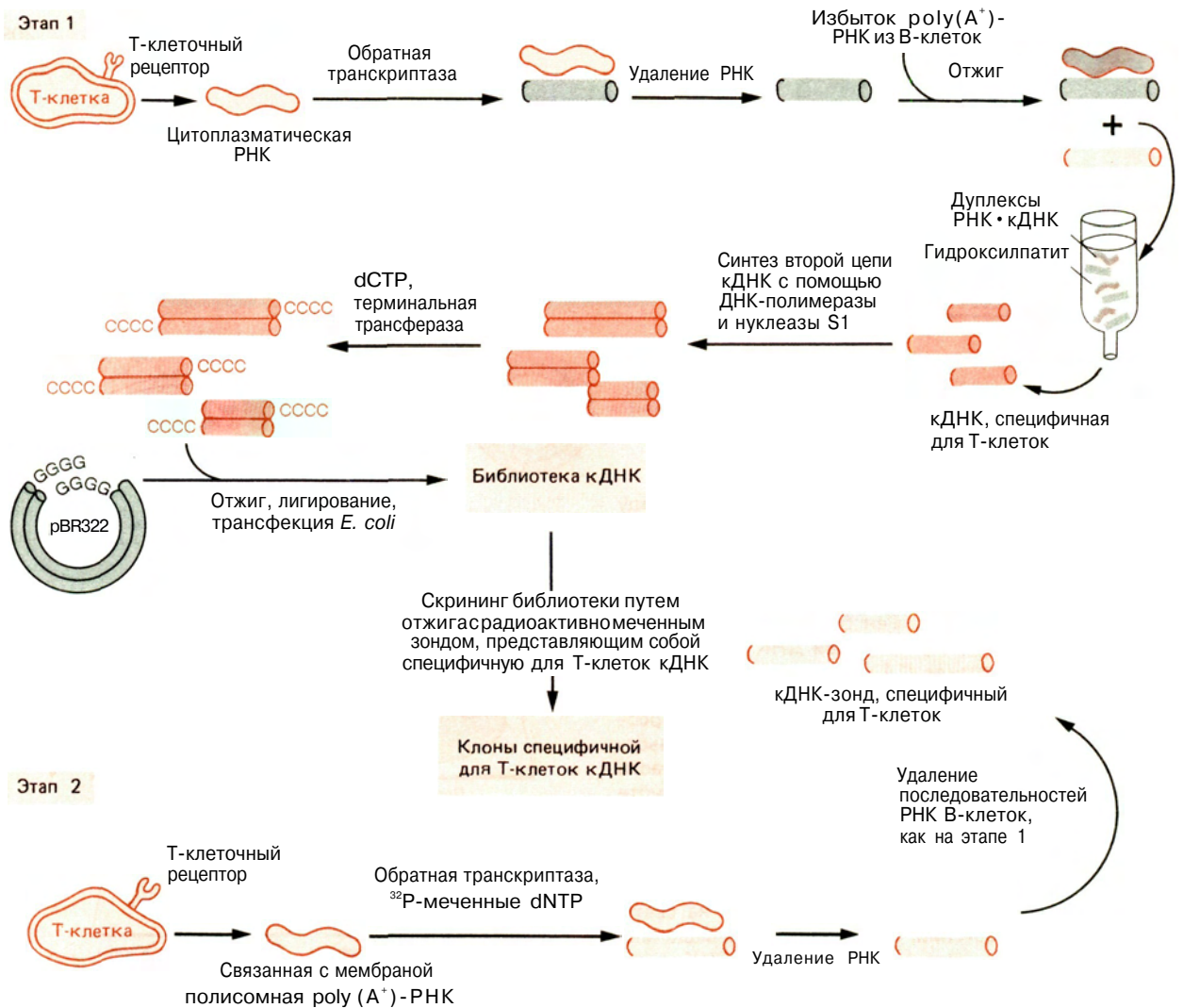
Клонирование генов, кодирующих субъединицы дрожжевой РНК-полимеразы II. ДНК из клеток *S. cerevisiae* механически разрезают на фрагменты средней длины 5 т.п.н. После заполнения одноцепочечных участков с помощью ДНК-полимеразы метилируют Eco RI-сайты, присоединяют Eco RI-линкеры и встраивают фрагменты в ДНК фага  $\lambda$ gt11, предварительно разрезанную с помощью эндонуклеазы Eco RI и обработанную фосфомоноэстеразой. Реконбинантные ДНК упаковывают в вирионы фага  $\lambda$  in vitro (выход фаговых частиц  $2 \cdot 10^6$ ) и инфицируют ими клетки *E. coli*, имеющие генотипы *sup F* (для супрессии амбер-мутации в гене S фага  $\lambda$ ), *hsd R*<sup>-</sup>, *hsd M*<sup>+</sup> (для предотвращения рестрикции инфицирующей ДНК),  $\Delta lac$  (для идентификации бляшек, образуемых фагом  $\lambda$ gt11, и бляшек, содержащих вставки дрожжевой ДНК) и содержащие плазмиду pBR322 с геном *lac I* (кодирующим *lac*-репрессор, благодаря чему подавляется экспрессия дрожжевых белков, ингибирующих рост клеток или фага). На этой стадии библиотеку размножают и получают фаговый препарат. Часть препарата, содержащую  $10^6$  фаговых частиц, подвергают скринингу для выявления фаговых частиц, образующих белки, которые связываются с антителами к дрожжевой РНК-полимеразе II (см. подпись к рис. 6.33). Для скрининга используют штамм *E. coli*, в котором делегирован важный протеазный ген,  $\Delta lon$ , чтобы уменьшить деградацию дрожжевых белков. Повторный

скрининг 60 положительных бляшек позволяет отобрать множество клонов, содержащих кодирующие последовательности для нескольких субъединиц РНК-полимеразы II. [R.A. Young, R.W. Davis, *Science*, **222** (1983), p. 778.]

**РИС. 6.36.**

Клонирование кДНК, кодирующей малую субъединицу рибулозо-1,5-бисфосфаткарбоксилазы (RuBPCase) из листьев гороха (*Pisum sativum*). RuBPCase, по-видимому, один из самых распространенных на земле белков, состоит из восьми идентичных больших и восьми идентичных малых субъединиц. Ген большой субъединицы (55 кДа) находится в хлоропластной ДНК (разд. 9.7.в), а ген малого полипептида (14 кДа-цепь, образующаяся в результате посттрансляционного разрезания предшественника) - в ядерной хромосомной ДНК. Экспрессия генов зависит от освещенности, о чем свидетельствуют данные, представленные в левой верхней части рисунка. Растения гороха выдерживали 9 сут в темноте и затем некоторые из них выставляли на свет на 48 ч. Полисомную poly(A<sup>+</sup>)-РНК, выделенную из растений обоих типов, транслировали *in vitro*. 20 кДа-предшественник малой субъединицы RuBPCase синтезировался только в тех растениях, которые находились на свету. РНК из листьев растений, подвергшихся освещению, использовали для синтеза двухцепочечной кДНК, к которой были присоединены Hind III-линкеры. ДНК лигировали с разрезанной в Hind III-сайте плазмидой pBR322 и полученными рекомбинантными плазмидами трансформировали штамм HB101 *E. coli*. Трансформированные бактериальные колонии, устойчивые к ампициллину, выращивали на нитроцеллюлозных фильтрах, помещенных на поверхность питательного агара в чашках. После фиксации ДНК на фильтрах клоны, содержащие кДНК RuBPCase, идентифицировали по их способности гибридизоваться с радиоактивно меченной poly(A<sup>+</sup>)-РНК из подвергавшихся освещению листьев и по отсутствию гибридизации с РНК из листьев растений, выращенных в темноте. Ложные положительные клоны элиминировали, проводя гибридизацию исходно положительных клонов с радиоактивно меченной РНК, обогащенной по мРНК RuBPCase. Обогащение проводили фракционированием РНК по размерам с помощью электрофореза в полиакриламидном геле. Искомая мРНК представляла собой молекулу длиной 900 п.н., присутствующую в подвергшихся освещению (но не в выращенных в темноте) листьях, при трансляции которой *in vitro* образовывалась малая субъединица RuBPCase. Наконец, гибридные клоны были идентифицированы также при гибридизационной селекции (разд. 6.4.б). [J.R. Bedbrook, S.M. Smith, R.J. Ellis, *Nature*, **287** (1980), p. 692.]

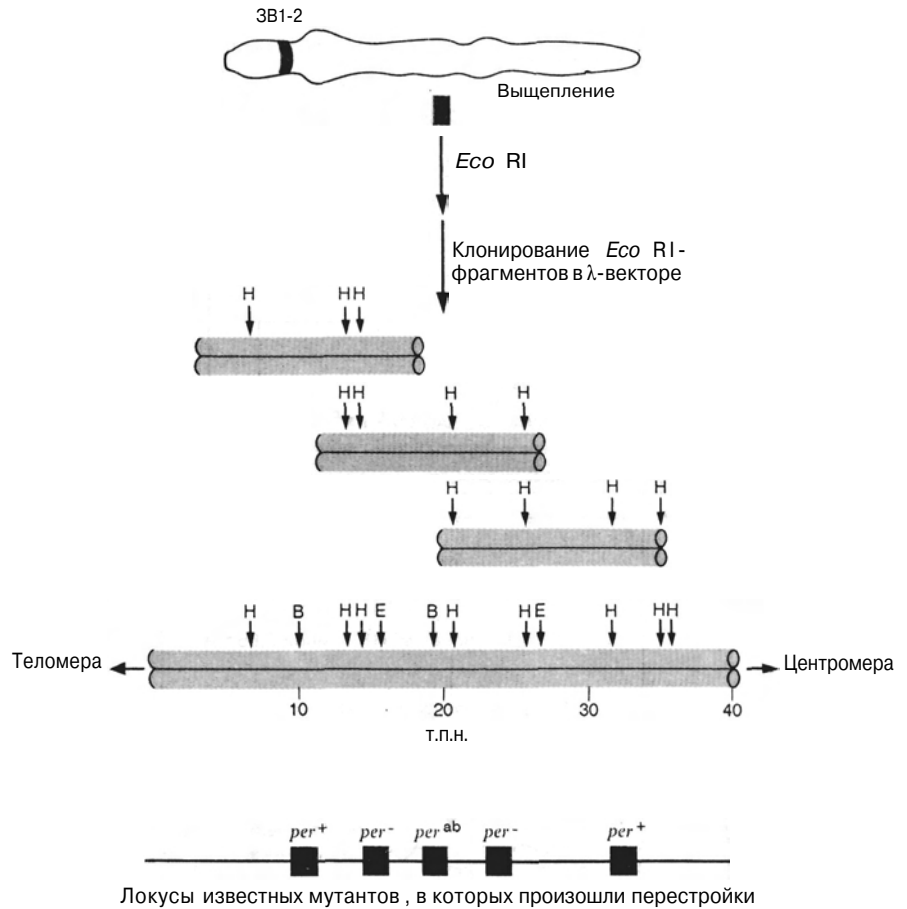
рая была предварительно обогащена последовательностями мРНК, специфичными для стадии гастролы (рис. 6.34). Такое обогащение можно получить следующим образом. С помощью обратной транскриптазы на мРНК из гастрол синтезируют одноцепочечную кДНК. Затем все кДНК, гомологичные мРНК, присутствующим как в яйцеклетках, так и в гастралах, удаляют путем образования гибридов кДНК•мРНК между кДНК гастрол и мРНК яйцеклеток; эти гибриды адсорбируются на гидроксипатите, а в растворе остаются только одноцепочечные кДНК, специфичные для гастрол (см. разд. 7.5.б). Из этих кДНК получают обычным способом дуплексные молекулы кДНК и создают библиотеку кДНК. Такая библиотека обычно бывает сильно обогащена молекулами кДНК, соответствующими гастролоспецифичной мРНК. Аналогичный подход можно использовать в случае мРНК, выделенной из любых двух родственных тканей или клеток. Вариации на эту тему включают использование в качестве зондов молекул кДНК, обогащенных описанным выше методом. В этом случае



**РИС. 6.37.**

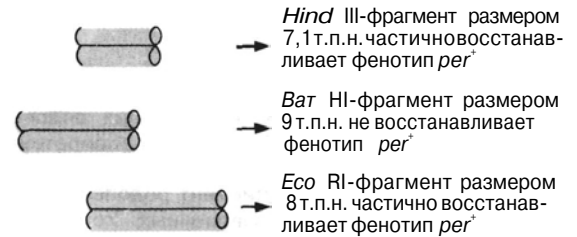
Клонирование кДНК крысы, кодирующей β-цепь Т-клеточного рецептора. Т-лимфоциты играют ключевую роль в клеточном и гуморальном иммунитете. Их функционирование опосредуется характерным гетеродимерным поверхностным рецептором (разд. 10.6.в). Первый этап клонирования гена β-цепи состоит в создании рВВ322-библиотеки кДНК, обогащенной мРНК, специфичной для Т-клеток. В качестве матрицы для синтеза первой цепи кДНК используется цитоплазматическая Т-клеточная poly(A<sup>+</sup>)-РНК. Молекулы кДНК отжигали с цитоплазматической poly(A<sup>+</sup>)-РНК из лимфоцитов другого типа, В-клеток, и дуплексы РНК • кДНК удаляли из смеси путем связывания их с гидроксилпатитом (разд. 7.5.б). При этом удалялось более 95% молекул кДНК, соответствующих молекулам мРНК, общим для В- и Т-клеток. Оставшуюся (5%) специфичную для Т-клеток кДНК использовали в качестве матрицы для синтеза второй цепи кДНК. После обработки нуклеазой S1 присоединяли С-«хвосты» и лигировали молекулы с плазмидой рВВ322, предварительно разрезанной в *Pst*I-сайте и содержащей G-«хвосты». Рекомбинантные молекулы использовали для получения трансформиро-

ванных колоний *E. coli*. Второй этап состоит в получении радиоактивно меченного зонда для скрининга библиотеки кДНК. В качестве зонда использовали первую цепь кДНК, синтезированную на цитоплазматической мембраносвязанной полисомной poly(A<sup>+</sup>)-РНК из Т-клеток. (Напомним, что, поскольку Т-клеточный рецептор находится на плазматической мембране клетки, он транслируется на рибосомах, связанных с мембранами эндоплазматического ретикулума.) Для обогащения кДНК в отношении специфичных для Т-клеток транскриптов ее гибридизовали с В-клеточной poly(A<sup>+</sup>)-РНК и дуплексы РНК • кДНК удаляли с помощью гидроксилпатита. Оставшуюся в растворе радиоактивно меченную специфичную для Т-клеток кДНК использовали для скрининга библиотеки. Из 5000 проверенных клонов библиотеки 35 были положительными. Пять из них гибридизовались с В-клеточной РНК на РНК-блотах. Один из оставшихся 30 содержал последовательности, кодирующие β-цепь Т-клеточного рецептора. До успешного клонирования кДНК Т-клеточный рецептор не удавалось выделить и о его структуре не было ничего известно. [S.M. Hedrick et al., *Nature*, **308** (1984), p. 149.]

**РИС. 6.38.**

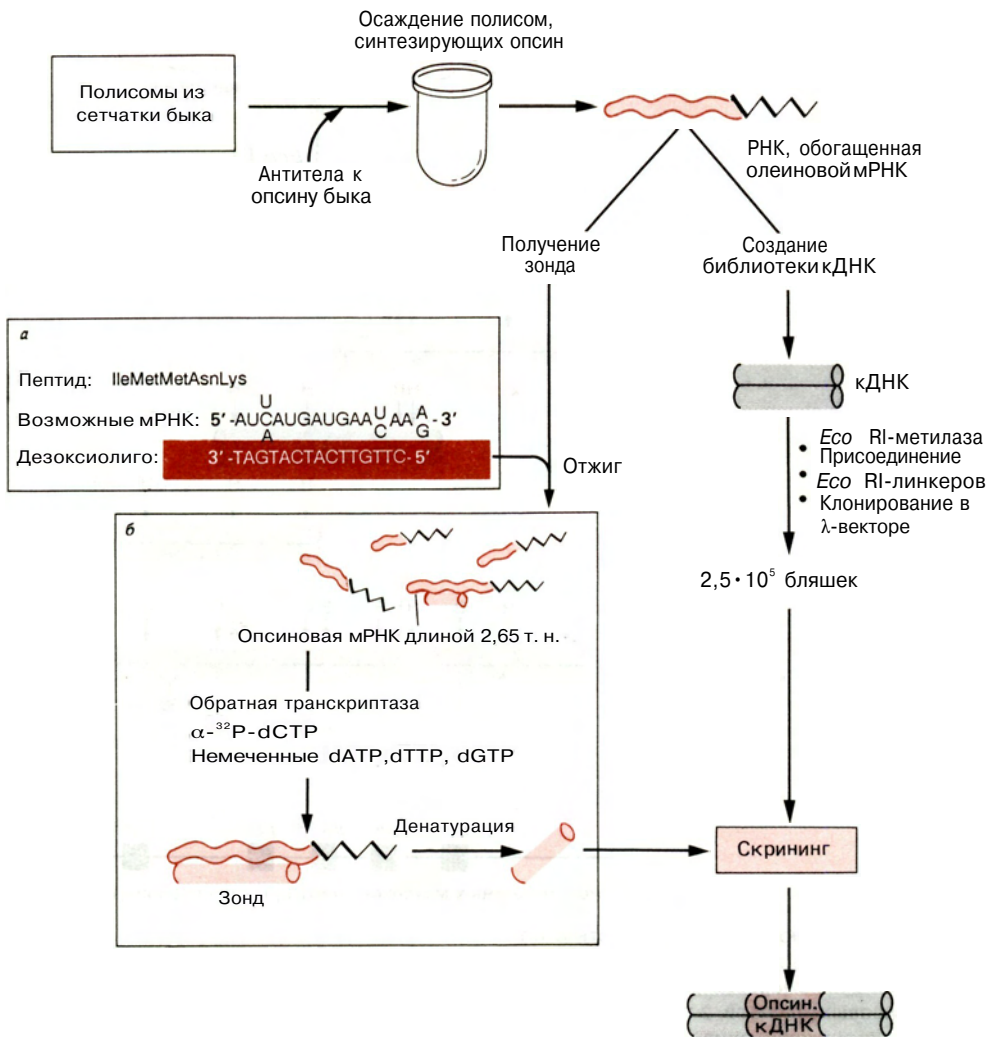
Клонирование гена, регулирующего ритмичность в поведении *Drosophila*. Мутации в локусе *per* области 3B1-2 X-хромосомы *Drosophila melanogaster* вызывают аномалии циркадных (24-часовых) ритмов, а также изменение частоты повторения песни ухаживания у мужской особи *Drosophila*. Такие мутации часто бывают связаны с хромосомными перестройками. Из этой части нормальной X-хромосомы были выщеплены сегменты ДНК, подвергнуты частичному гидролизу рестриктирующей эндонуклеазой и клонированы в λ-векторе. Клоны с перекрывающимися последовательностями идентифицировали с помощью перекрестной гибридизации и обычных карт эндонуклеазных сайтов (здесь указаны не все сайты). В результате была построена карта длинного сегмента ДНК, включающего локус *per*. Субклонированные области этого сегмента использовались в качестве зондов для построения карт X-хромосом, имеющих перестройки вблизи локуса *per* и известных как *per*<sup>+</sup>, *per*<sup>-</sup> или *per*<sup>ab</sup> (аномальное ритмическое поведение). Таким способом была установлена корреляция между утратой нормального фенотипа и сегментом ДНК длиной 15 т.п.н., находящимся вблизи центральной части карты. Некоторые субклоны при трансдукции с помощью Р-элемента частично восстанавливали циркадные ритмы и периодичность песен ухаживания у мух с делециями *per* (см. разд. 10.2.6). Таким образом

Трансдуцирующие субклонированные участки



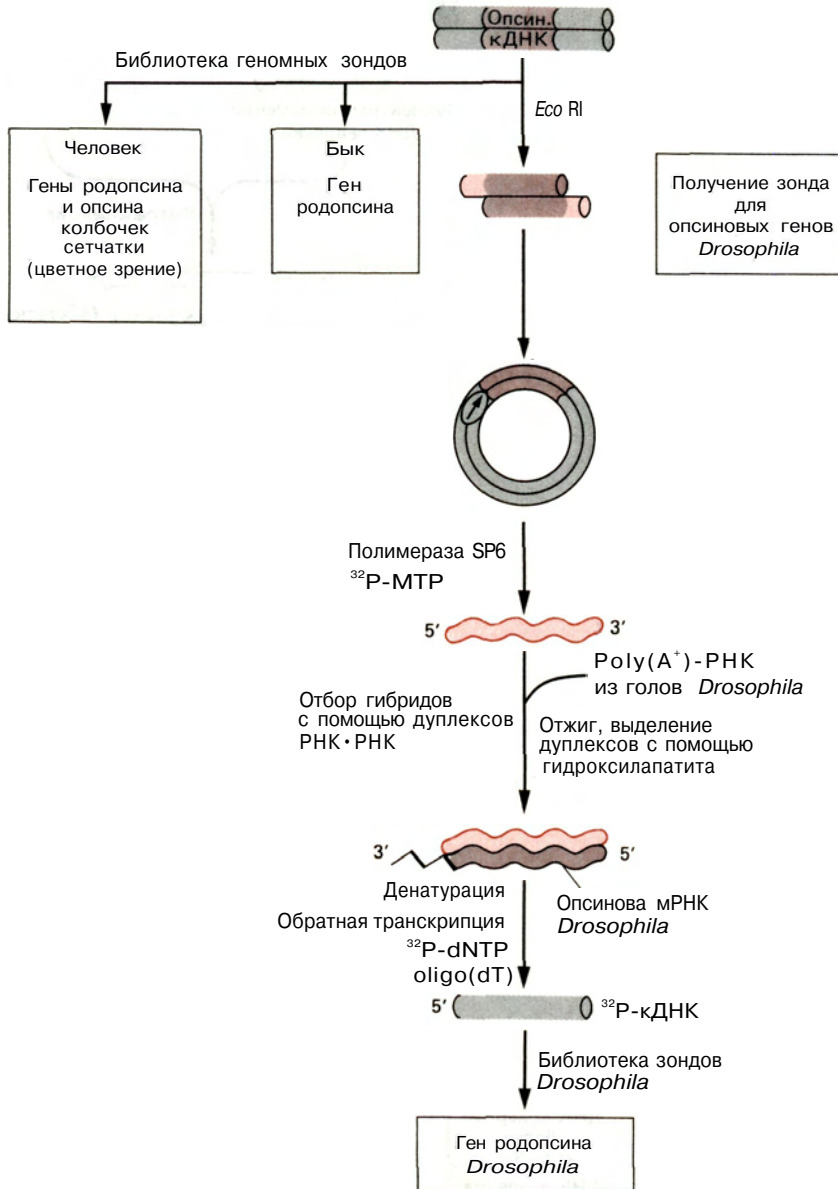
удалось выделить функциональный ген, регулирующий поведение *Drosophila*, не зная точно положение этого гена и не располагая какими-либо данными о его продукте [T.A. Bargiello, M.W. Young, Proc. Natl. Acad. Sci. USA., **81** (1984), p. 2142; W. Zehring et al., Cell, **39** (1984), p. 369; Y. Citri et al., Nature, **326** (1987), p. 42.]



**РИС. 6.39.**

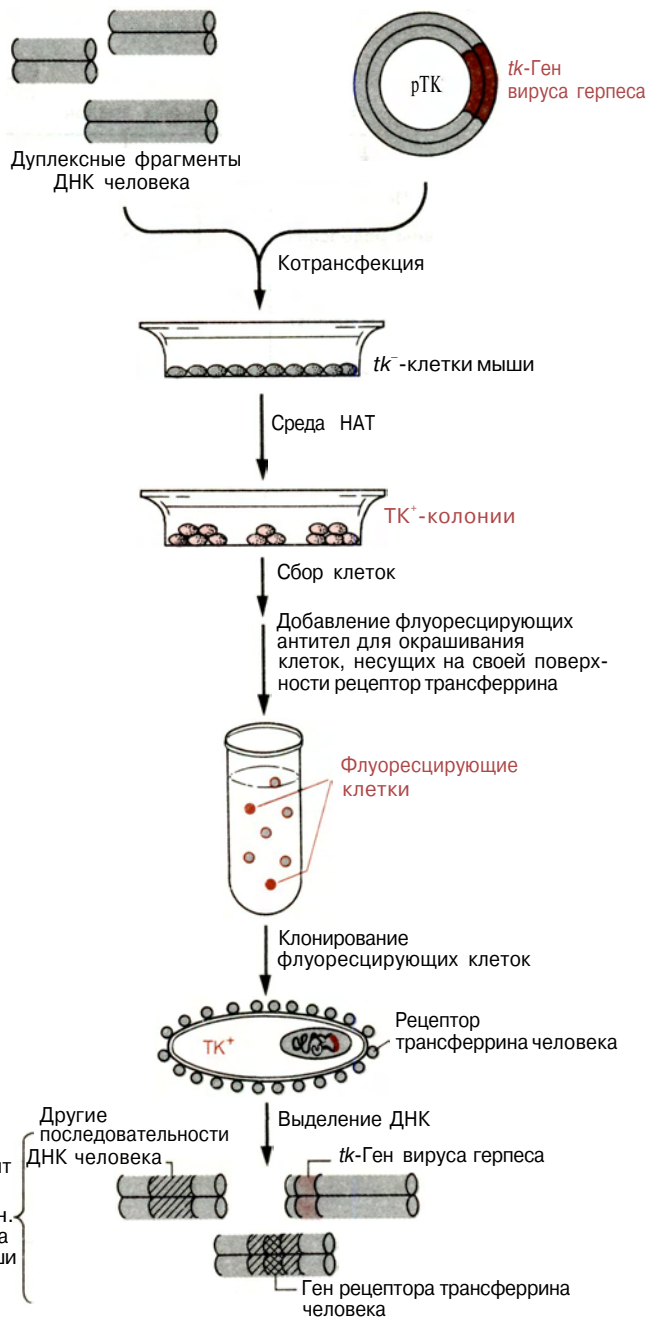
Клонирование генов и кДНК, кодирующих родопсин быка, человека и *Drosophila*. Опсины - это апопротеины, входящие в состав фоторецепторов сетчатки (зрительных пигментов). Зрительный пигмент родопсин содержится в палочках сетчатки позвоночных. Установлена его аминокислотная последовательность. Выделены соответствующие кДНК и ген из клеток быка. Из полисом сетчатки быка была получена poly(A<sup>+</sup>)-РНК, обогащенная мРНК опсина. С помощью  $\lambda$ -вектора стандартными методами была создана библиотека кДНК. Ее скрининг проводили с использованием специфического зонда, полученного в два этапа (а и б): а - химический синтез дезоксиолигонуклеотида длиной 15 остатков, комплементарного мРНК, последовательность которой была воссоздана исходя из данных о последовательности пяти аминокислот, находящихся вблизи СН<sub>3</sub>-конца молекулы опсина; б - использование этого дезоксиолигонуклеотида из 15 звеньев в качестве специфического

праймера для обратной транскрипции мРНК опсина, присутствующей в сетчатке быка. Зонд был радиоактивно помечен при обратной транскрипции с помощью  $\alpha$ - $^{32}$ P-dCTP. Олигонуклеотид длиной 15 звеньев позволяет идентифицировать олеиновые клоны, однако специфичность теста увеличивается с увеличением длины зонда. Из  $2,5 \cdot 10^5$  клонов библиотеки специфическую реакцию с зондом дали 0,2%. Анализ нуклеотидных последовательностей восьмью из них показал, что они кодируют части молекулы опсина. Клоны кДНК использовали в качестве зондов для скрининга геномных библиотек *Drosophila*, быка и человека. В библиотеке *Drosophila* никаких положительных рекомбинантов обнаружено не было, однако с помощью специальных методов была выявлена последовательность, отдаленно родственная последовательности *Drosophila*. кДНК быка вырезали из ДНК  $\lambda$ -вектора с помощью рестриктирующей эндонуклеазы и повторно клонировали ее в век-



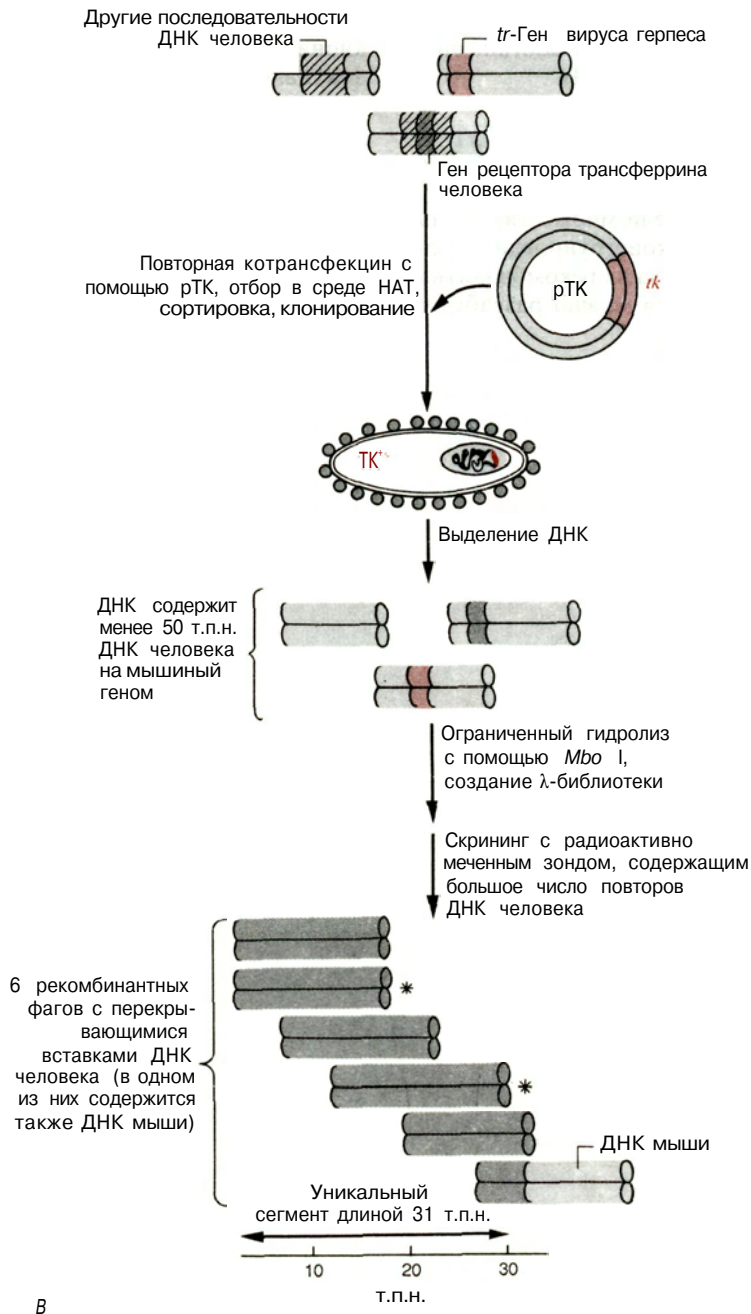
торе типа rGEM. Радиоактивно меченную РНК, соответствующую несмысловой цепи, синтезировали, используя РНК-полимеразу SP6, и отжигали с poly(A<sup>+</sup>)-РНК, выделенной из голов *Drosophila*. Дуплексы РНК•РНК, содержащие мРНК опсина *Drosophila*, осаждали на гидроксилпатите (разд. 7.5.6). После денатурации дуплексов РНК *Drosophila* использовали в качестве матрицы для обратной транскрипции в присутствии oligo(dT)-праймера и радиоактивно меченных дезоксирибонуклеотидов, а образовавшуюся радиоактивно меченную кДНК - для отбора клонов из геномной библиотеки

*Drosophila*. Успех процедуры зависит от характеристик дуплексов РНК•РНК. Неправильно спаренные дуплексы РНК•РНК отличаются большей стабильностью, чем соответствующие неправильно спаренные дуплексы ДНК•ДНК. [J. Nathans, D.S. Hogness, *Cell*, **34** (1983), p. 807; C.S. Zuker, A.F. Cowman, G.M. Rubin, *Cell*, **40** (1985), p. 851.]

**РИС. 6.40.**

Клонирование гена человека, кодирующего специфический трансферриновый рецептор, находящийся на клеточной поверхности. Трансферрин - это  $\text{Fe}^{3+}$ -связывающий сывороточный белок, участвующий в транспорте железа внутрь клетки. Поглощение клеткой  $\text{Fe}^{3+}$ -трансферринового комплекса опосредуется специфическим рецептором, представляющим собой гомодимерный гликопротеин клеточной мембраны. А. Для клонирования гена рецептора мышиные  $tk^-$ -клетки котрансфицировали высокомолекулярными фрагментами суммарной ДНК человека и рекомбинантной плазмидой, содержащей ген тимидинкиназы ( $tk$ ) вируса герпеса. После отбора в среде НАТ  $TK^+$ -клетки смешивали с антителами к рецептору трансферрина человека. Клетки, связавшие антитела (и, следовательно, несущие рецептор), обрабатывали флуоресцентной меткой, отбирали с помощью сортера и клонировали. ДНК, выделенная из клонированных клеток, содержала значительно

большее количество ДНК человека (более  $3 \cdot 10^3$  т.п.н. на мышиный геном), чем требовалось для кодирования одного гена. (Количество ДНК человека оценивали в присутствии большого избытка мышиной ДНК путем отжига с зондом, содержащим последовательности ДНК, повторяющиеся почти в  $10^6$  случайных мест, разбросанных по всему геному человека, и не встречающиеся в ДНК мыши [разд. 9.5.в.] Б. Чтобы уменьшить количество нежелательной ДНК человека, еще раз про-



водили котрансфекцию с участием ДНК из этих клеток, отбор в среде НАТ, сортировку и клонирование. ДНК из полученных вторичных трансформантов, содержащую менее 50 т.п.н. ДНК человека на мышиный геном, использовали для создания геномной библиотеки. Было отобрано шесть фаговых клонов, которые гибридизовались с зондом, содержащим повторяющиеся последовательности ДНК человека, разбросанные по всему геному. Все шесть клонов содержали перекрывающиеся

сегменты ДНК. Две вставки из шести (отмечены звездочками) встраивали в фаговые векторы для воссоздания уникального сегмента ДНК размером 31 т.п.н., ответственного за синтез поверхностного рецептора трансферрина человека, после трансфекции в мышиные клетки. [L.C. Kuhn, A. McClelland, F.H. Ruddle, *Cell*, **37** (1984), p. 95.]

библиотеку кДНК подвергают скринингу, выявляя те клоны, которые гибридизуются с обогащенным мРНК (или кДНК)-зондом (рис. 6.34).

## 6.6. НЕКОТОРЫЕ СТРАТЕГИИ КЛОНИРОВАНИЯ ГЕНОВ И кДНК

В этой главе мы рассмотрели множество разных методов, используемых при конструировании, клонировании и отборе специфических рекомбинантных генов. На основе этих методов можно разработать

множество стратегий для получения определенных клонированных генов или кДНК. При выборе эффективной и действенной стратегии важно учитывать природу гена и его продукта, а также происхождение вида. Чтобы представить себе все многообразие методов работы с рекомбинантными молекулами ДНК, лучше всего рассмотреть конкретные примеры клонирования различных генов и кДНК. На рис. 6.35–6.40 представлены эксперименты по клонированию, иллюстрирующие разные стратегии. Другие примеры описаны в последующих главах.

# ПРОДУКТЫ РЕКОМБИНАЦИИ: ХАРАКТЕРИСТИКА И МАНИПУЛИРОВАНИЕ

После отбора клонированных рекомбинантных молекул нужно охарактеризовать содержащуюся в них вставку. Для этого необходимо ответить на несколько вопросов. Прежде всего — действительно ли рекомбинантная молекула содержит нужный сегмент ДНК? Далее — включился ли сегмент целиком или только частично? Соответствует ли он во всех деталях геномной ДНК, из которой происходит? Является ли нужная последовательность функциональной? В этой главе мы рассмотрим разные подходы к характеристике клонированных фрагментов ДНК. Прежде всего нужно очистить рекомбинантную ДНК от ДНК клеток хозяина, РНК и белков. Эта процедура довольно проста и включает экстракцию, физическое разделение компонентов и ферментативное расщепление РНК и белков. Так, плазмидную ДНК можно отделить от геномной благодаря ее малому размеру и кольцевой структуре (крупный кольцевой геном, подобный геному *E. coli*, обычно при выделении разрывается на линейные фрагменты). Фаговую ДНК можно получить в свободном от клеточной ДНК виде путем очистки фаговых частиц и последующего выделения из них ДНК.

## 7.1. МАКРОСТРУКТУРА КЛОНИРОВАННОЙ ВСТАВКИ

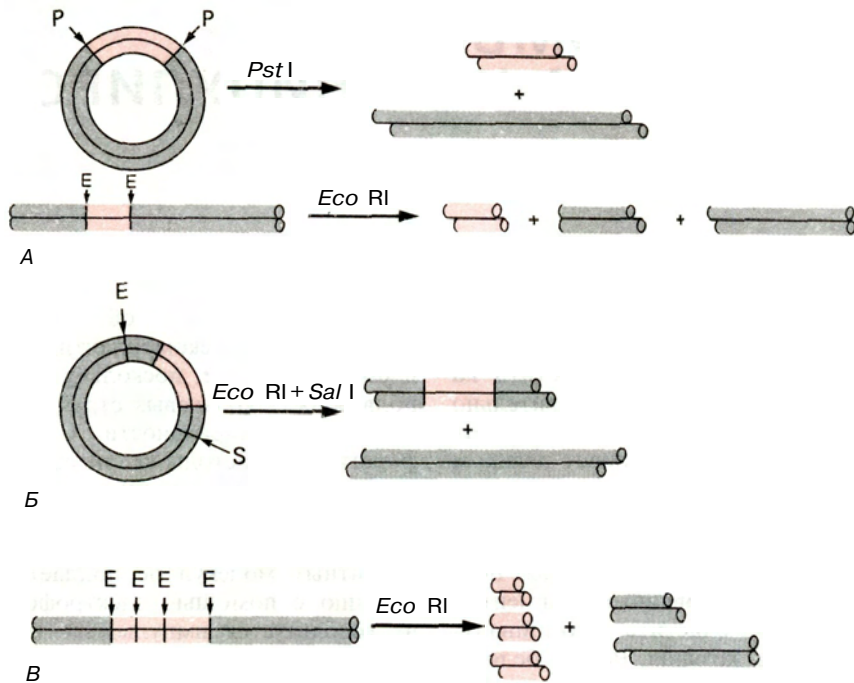
Для определения макроструктуры вставки сначала устанавливают ее размер и положение сайтов для рестриктирующих эндонуклеаз. Поскольку клонированные фрагменты геномной ДНК часто бывают длиннее или короче, чем интересующая нас область, необходимо также определить положение и длину нужного сегмента внутри вставки.

### а. Размер вставки

Размер векторной молекулы обычно бывает известен, поэтому размер вставки можно оценить, зная полную длину рекомбинантной молекулы и предположив, что никакого deletирования во время клонирования не произошло. Размер многих рекомбинантных молекул можно определить с помощью

гель-электрофореза. Чтобы получить точные данные, кольцевые рекомбинантные молекулы превращают в линейные, поскольку электрофоретическая подвижность кольцевых структур зависит от степени их сверхспиральности. Стандартные электрофоретические методы непригодны для тестирования молекул, длина которых превышает 15 т.п.н, поскольку крупные молекулы мигрируют в геле слишком медленно. Поэтому размер многих рекомбинантных молекул не поддается прямому определению с помощью электрофореза, если только не используется импульсный электрофорез (разд. 6.5.а). Альтернативный способ оценки размера рекомбинантных молекул состоит в измерении их длины с помощью электронной микроскопии. Здесь тоже очень важны стандартные размеры, поскольку число пар оснований в сегменте определенной длины на электронной микрофотографии может зависеть от конфигурации молекулы и способа приготовления препарата.

Другой подход к определению размера вставки состоит в обращении тех реакций, которые использовались при конструировании рекомбинантных молекул, т.е. в вырезании вставок с помощью соответствующих эндонуклеаз рестрикции (рис. 7.1, А). Размер линейной вставки определяют с помощью гель-электрофореза. Поскольку карта сайтов для эндонуклеаз рестрикции в векторной молекуле обычно бывает известна, идентифицировать встроенный фрагмент довольно легко. Если при лигировании произошла элиминация эндонуклеазных сайтов на концах вставки (т.е. лигировались тупые концы), можно использовать альтернативную процедуру: разрезать векторную молекулу по сайтам, находящимся в участках, которые фланкируют вставку; в этом случае вставка будет вырезана вместе с сегментами вектора известных размеров (рис. 7.1, Б). Иногда ситуация осложняется тем, что сама вставка содержит сайты узнавания для используемой эндонуклеазы (рис. 7.1, В). В таком случае сумма размеров всех фрагментов, образующихся из вставки, даст ее общую длину. Если вставка слишком велика, то ее разрезание и суммирование могут оказаться даже необходимыми для более точного определения длины.

**РИС. 7.1.**

Определение размера вставки в рекомбинантной молекуле ДНК. А. Обращение процессов, которые использовались для лигирования вектора со вставкой при

клонировании. Б. Разрезание по сайтам, фланкирующим вставку. В. Суммирование размеров рестрикционных фрагментов, образовавшихся из вставки.

## б. Картирование сайтов для рестриктирующих эндонуклеаз

Принципы, лежащие в основе построения карты сайтов для рестриктирующих эндонуклеаз в каком-либо сегменте ДНК, описаны в разд. 4.2.г. В тех случаях, когда вектор имеет небольшие размеры и не слишком сложен, положение сайтов иногда можно определить непосредственно в самой рекомбинантной молекуле. Однако часто оказывается необходимым сначала вырезать вставку и очистить ее от векторных сегментов.

## в. Субклонирование

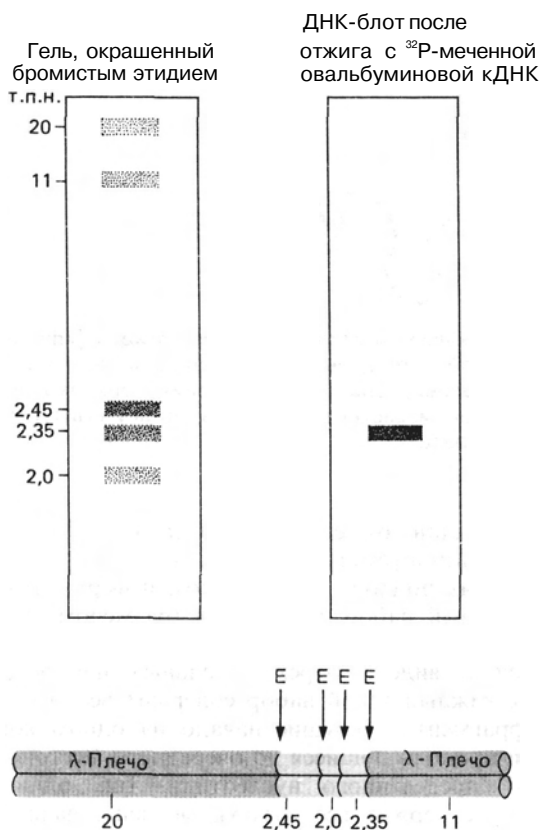
Длинные клонированные вставки, входящие в состав  $\lambda$ - или космидного вектора, часто бывают весьма неудобными для манипулирования. Поэтому после построения частичной рестрикционной карты вставку можно разделить на более мелкие фрагменты путем **субклонирования**. В принципе субклонирование не отличается от других методов создания рекомбинантных ДНК, просто в этих случаях в качестве исходного генома используют рекомбинантный клон. Для субклонирования обычно применяют рестриктирующие эндонуклеазы, сайты

узнавания которых состоят из четырех пар оснований, поскольку они способны разрезать ДНК во многих местах.

## г. Определение положения интересующего нас сегмента во вставке

Для определения положения сегментов в небольших рекомбинантных геномах используются те же подходы, что и при нахождении специфических сегментов в больших геномах до клонирования (разд. 6.1.б). После построения рестрикционной карты вставки установить точную локализацию нужного сегмента не составляет труда. Для этого нужен только подходящий меченый зонд, аналогичный ДНК- или РНК-зонду, используемому для отбора клона в начале клонирования.

На рис. 7.2 представлены данные, касающиеся гена овальбумина курицы. Овальбумин—это белок, состоящий из одной полипептидной цепи (43500 Да); он синтезируется в яйцевом курицы после стимуляции стероидным гормоном эстрогеном (или прогестероном) и накапливается в яичном белке. У кур-несушек стимуляция гормоном происходит в естественных условиях, а у цыплят

**РИС. 7.2.**

Инкубация овальбуминовой мРНК со смесью фрагментов, образовавшихся при расщеплении ДНК фага λ *Eco*RI-эндонуклеазой и содержащих последовательности овальбуминового гена цыпленка. Гидролизат подвергнут электрофорезу и блот-гибридации. Внизу представлена рестрикционная карта вставки. Цифры - длины фрагментов в т.п.н.

образование яйцевода и синтез овальбумина индуцируются при введении эстрогена. Гормон индуцирует транскрипцию гена овальбумина, в результате чего содержание овальбуминовой мРНК в клетке увеличивается практически от нуля до 50000 молекул через две-три недели и падает до 10 молекул на клетку после прекращения воздействия эстрогена. Яйцеводы кур-несушек представляют собой хороший источник овальбуминовой мРНК, поскольку она составляет в них около 50% суммарной мРНК. Очищенную мРНК используют затем для приготовления радиоактивного зонда, применяемого для идентификации клонированного гена овальбумина. Зондом может служить радиоактивно меченная одноцепочечная кДНК, синтезированная непосредственно на мРНК с помощью обратной транскриптазы, или клонированная дуплексная кДНК. С по-

мощью таких зондов из популяции λ-векторов, содержащих вставки геномной ДНК курицы, полученные при ограниченном гидролизе эндонуклеазой *Eco*RI, были отобраны клонированные последовательности овальбуминового гена. Вставка, представленная на рис. 7.2, имела длину 6,8 т.п.н. Из нее образовались три *Eco*RI-фрагмента, при этом только фрагмент длиной 2,35 т.п.н. отжигался с овальбуминовым кДНК-зондом.

Положение искомой последовательности в клонированном сегменте можно определить также с помощью электронной микроскопии. Для этого зонд и рекомбинантную ДНК смешивают, денатурируют и отжигают, а затем готовят препарат для микроскопирования. Комплементарные участки зонда и исследуемой молекулы образуют дуплексы, которые на электронных микрофотографиях выглядят как относительно толстые нити, а некомплементарные участки остаются одноцепочечными и имеют вид более тонких нитей. Гетеродуплекс (РНК • ДНК), образующийся между овальбуминовым рекомбинантом, описанным в разд. 7.2, и овальбуминовой мРНК, показан на рис. 7.3. Такие гетеродуплексы формируются при условиях, благоприятных для гибридизации типа РНК • ДНК, а не ДНК • ДНК. В сегменте ДНК размером 2,35 т.п.н., который гибридизовался с мРНК (рис. 7.2), в образовании двухцепочечной структуры участвовали только около 550 п.н. Таким образом, гетеродуплексный анализ гораздо более информативен, чем определение положения кодирующих последовательностей в клоне. Он показывает, что, хотя мРНК содержит около 1800 п.н., в клонированном сегменте ДНК представлено менее трети длины кодирующих последовательностей гена овальбумина. Кроме того, 550 п.н. в клонированной ДНК, комплементарные мРНК, не образуют одну непрерывную последовательность, а распределены по четырем участкам, разделенным одноцепочечными петлями, т.е. кодирующие последовательности геномной ДНК чередуются с некодирующими. Как описано в разд. 8.5, такие **промежуточные последовательности**, или **интроны**, типичны для эукариотических генов.

## 7.2. ТОНКАЯ СТРУКТУРА КЛОНИРОВАННОЙ ВСТАВКИ: НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ

Чтобы до конца выяснить структуру, функцию и происхождение клонированного сегмента ДНК, необходимо установить его первичную структуру — нуклеотидную последовательность. Быстрые и точные методы определения последовательностей ДНК



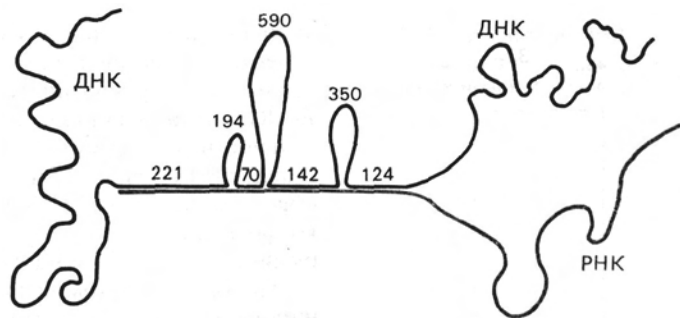
**РИС. 7.3.**

Схема гетеродуплекса, построенная по электронно-микроскопической фотографии. Этот гетеродуплекс образуется между овальбуминовым рекомбинантом, описанным в разд. 7.2 (черный цвет), и овальбуминовой мРНК (выделена цветом). Указаны длины различных

участков молекулы в числе пар нуклеотидов (или числе нуклеотидов, если речь идет об одноцепочечных участках). Они определены путем сравнения с длинами полинуклеотидов известных размеров, присутствующих в том же образце.

были созданы вскоре после разработки методов, используемых в работе с рекомбинантными молекулами. В принципе сейчас можно провести секвенирование молекулы ДНК любой длины. Определение последовательности сегментов ДНК протяженностью в сотни и даже тысячи нуклеотидов представляет собой рутинную процедуру, а примерно до 1975 г. это была очень трудная задача. Для этого с помощью РНК-полимеразы сначала получали РНК-копии ДНК, а затем определяли нуклеотидную последовательность кРНК. Точность и полнота процесса копирования часто были недостаточны, а секвенирование РНК занимало много времени. Сейчас секвенируют саму ДНК, а для секвенирования РНК обычно вначале получают кДНК-копии.

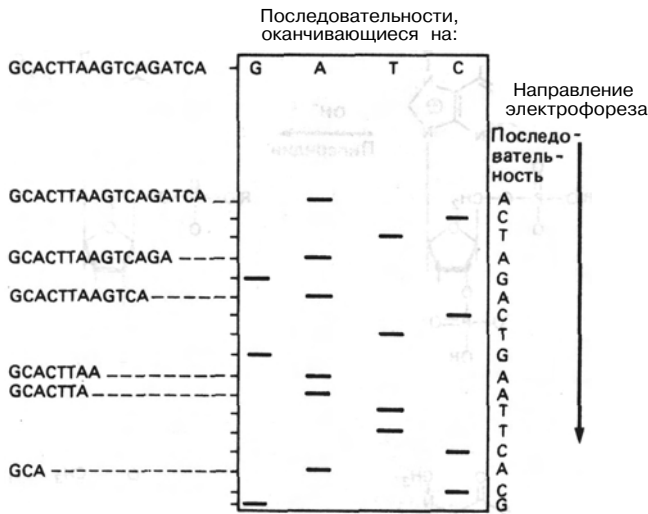
### а. Общие принципы

Методы секвенирования ДНК можно разбить на две категории. В основе одних лежат химические реакции, в которых используются непосредственно фрагменты очищенной ДНК. Во втором случае используют ДНК-копии очищенных сегментов, полученные ферментативным путем. Эти подходы имеют и некоторое сходство. Прежде всего фрагменты ДНК обычно очищают, что легко осуществить с помощью клонирования. Далее, и в том, и в другом случае за один раз секвенируется только одна цепь ДНК. Для повышения точности лучше провести секвенирование каждой цепи дуплексной молекулы и сравнить результаты. За один раз используют только одну из цепей, помеченную радиоактивным изотопом. Определяют нуклеотидную последовательность только этой цепи, и именно о ней мы будем говорить в последующих разделах. Третья общая особенность указанных методов состоит в том, что в обоих случаях образуется набор радиоактивно меченных одиночных цепей всех воз-

можных длин—от единицы до  $n$ , где  $n$ —полная длина секвенируемой молекулы.

В обоих подходах, химическом и ферментативном, полный набор фрагментов (размером от нескольких единиц до  $n$ ) на самом деле бывает представлен в виде четырех отдельных наборов. В идеале каждый такой набор содержит все возможные фрагменты, берущие начало на одном конце цепи и продолжающиеся до очередного местоположения определенного нуклеотида. Так, один из наборов содержит все возможные цепи, начинающиеся в одном сайте и оканчивающиеся в тех местах, где встречается дезоксиаденозин (рис. 7.4). Второй набор содержит все возможные цепи, начинающиеся в том же сайте, а заканчивающиеся на всех встречающихся поочередно дезоксицитидиновых остатках. Третий и четвертый наборы представлены фрагментами, оканчивающимися на остатках тимидина или дезоксигуанозина.

Химический и ферментативный методы отличаются друг от друга способом получения четырех наборов фрагментов. В первом случае этот набор получают путем разрезания предварительно радиоактивно меченной цепи четырьмя разными способами, а во втором четыре радиоактивно меченных набора фрагментов получают путем копирования немеченой цепи ДНК. Последним этапом в обоих методах является разделение фрагментов, составляющих каждый из наборов, по длинам с помощью электрофореза в полиакриламидном геле в денатурирующих условиях. При этом удается разделить полинуклеотидные цепи, отличающиеся друг от друга только одним нуклеотидным остатком. Поскольку фрагменты несут радиоактивную метку, их легко выявить с помощью радиоавтографии, при этом достаточно лишь небольшого количества ДНК—порядка пикомолей или даже меньше. Все четыре набора, полученные из одного фрагмента



ДНК, подвергают одновременному электрофорезу на параллельных дорожках одной пластины геля. Сканируя каждую дорожку, расшифровывают всю последовательность. Используя специальные электрофоретические методы, можно разделить цепи размером от 1 до 300 или более нуклеотидных остатков. Молекулы, длина которых превышает несколько сотен нуклеотидов, фрагментируют (обычно с помощью рестриктирующих эндонуклеаз или путем субклонирования) и затем определяют нуклеотидную последовательность каждого фрагмента.

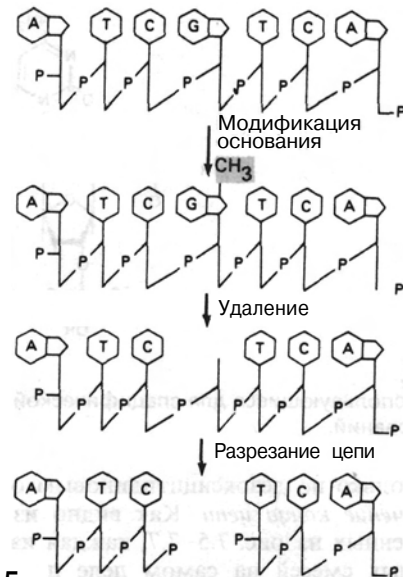
## б. Химическое секвенирование

Получение наборов фрагментов, имеющих общие концы. Химическое секвенирование основано на специфической модификации различных пуриновых и пиримидиновых оснований. Эти модифицированные основания выщепляют затем из полимерной цепи с сохранением сахарофосфатного остова. Далее гидролизуют относительно нестабильные фосфодиэфирные связи, соседствующие с сайтом, где находилось удаленное модифицированное основание, в результате чего цепь разрывается. Все эти реакции схематически представлены на рис. 7.5, где в качестве примера рассмотрена модификация гуаниновых остатков. Реакции, специфичные в отношении разных оснований, приведены на рис. 7.6. Представлены две реакции, одна из которых специфична в отношении только гуанина, а другая — только цитозина, и две реакции, специфичные в отношении либо пуринов, либо пиримидинов. Каждую из четырех реакций проводят с использованием отдельного препарата ДНК, в результате чего получают четыре набора фрагментов, необходимых для точного определения последовательности. Глубина

## РИС. 7.4.

Результат одновременного электрофореза четырех наборов олигодезоксирибонуклеотидов, синтезированных на матрице, изображенной в левом верхнем углу. Каждый набор содержит все возможные цепи, начинающиеся от 5'-конца и оканчивающиеся в тех местах, где находится тот или иной из дезоксирибонуклеозидов — дезоксигуанозин, дезоксиаденозин, дезоксицитидин или тимидин. В качестве примера слева приведен набор цепей, оканчивающихся на дезоксиаденозине.

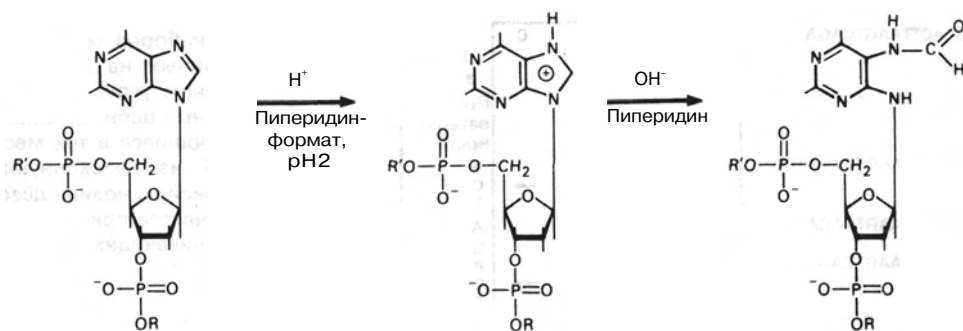
каждой реакции строго ограничена, так что в каждой молекуле ДНК данного препарата в реакцию вступает в среднем только одно из чувствительных оснований. В результате популяция идентичных молекул ДНК превращается в ходе реакций, в которых участвуют случайные реакционноспособные основания, в набор цепей, начинающихся в данном сайте на одном конце цепи и заканчивающихся на одном из реакционноспособных оснований. В примере, приведенном на рис. 7.7, разрезание цепи про-



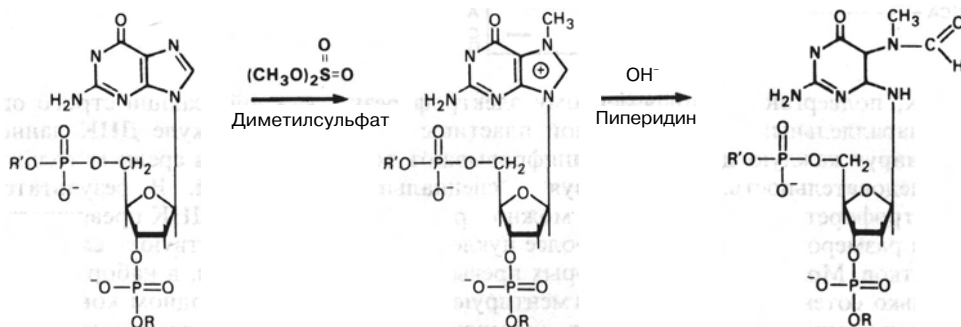
## РИС. 7.5.

Три этапа химического секвенирования: модификация основания, удаление модифицированного основания, разрыв цепи в месте расположения свободной дезоксирибозы. В данном примере модифицируемым основанием является гуанин.

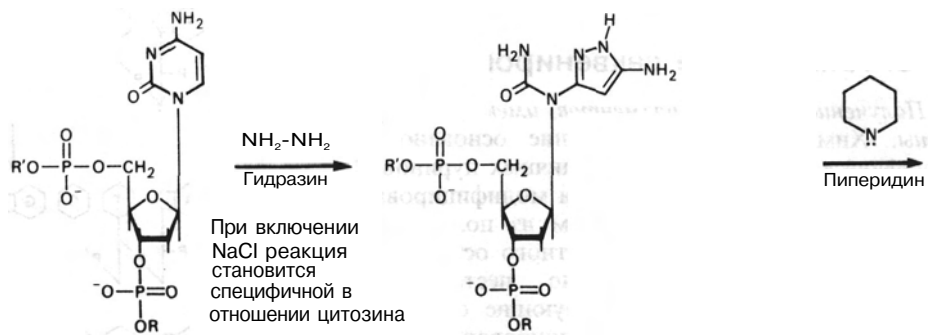
Разрезание по пуриновым остаткам



Разрезание по гуаниновым остаткам



Разрезание по пиримидиновым остаткам

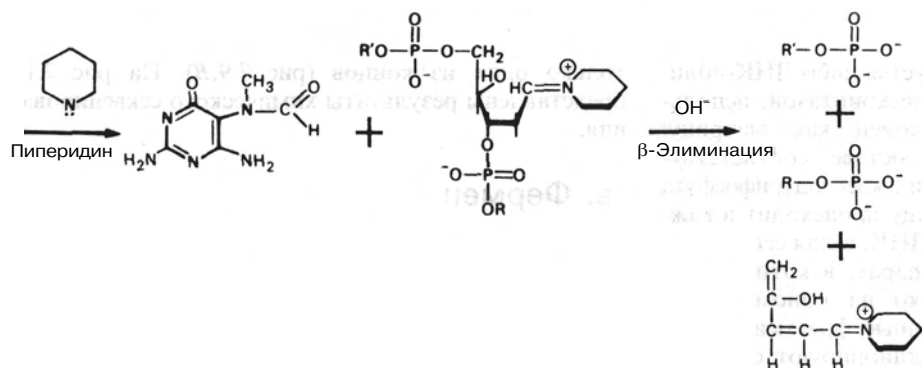
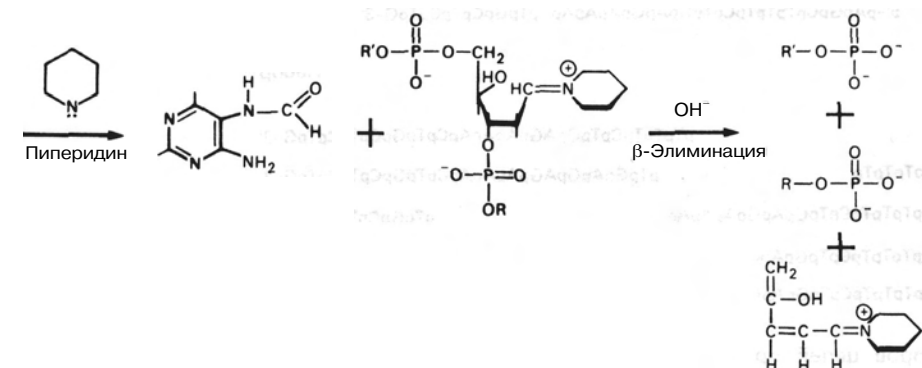
**РИС. 7.6.**

Реакции, используемые для специфической модификации оснований.

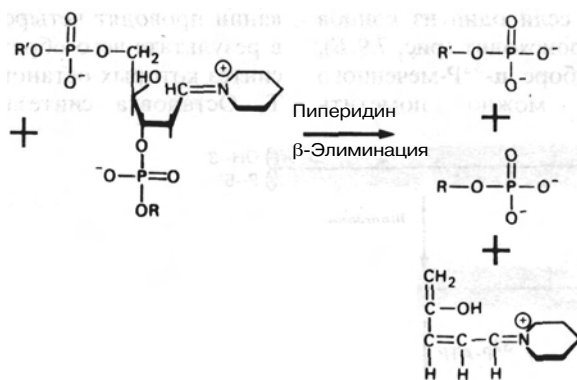
исходит только по дезоксицитидиновым остаткам.

<sup>32</sup>*P*-мечение конца цепи. Как видно из данных, представленных на рис. 7.5–7.7, каждая из четырех реакционных смесей на самом деле представлена двумя наборами цепей, получающихся в результате гидролиза препарата ДНК. Один включает все фрагменты, начинающиеся на 5'-конце исходной цепи, другой—все фрагменты, начинающиеся на 3'-конце. Нужную нам информацию о последова-

тельности можно получить, используя любой из наборов, и выбор одного из них зависит от того, как был помечен исходный фрагмент—с 3'- или 5'-конца—перед проведением химических реакций, специфичных в отношении определенного основания. Немеченый материал нельзя визуализировать с помощью радиоавтографии. Если исходным материалом является дуплексная ДНК, то помечена должна быть только одна из двух цепей, в противном случае конечные продукты будут представлены двумя разными наборами меченых цепей и данные будет трудно интерпретировать.



Смесь  
фрагментов



Введение метки в 5'-конец осуществляют с помощью полинуклеотидкиназы и  $\gamma$ - $^{32}\text{P}$  [АТР] (рис. 7.8). Если присутствуют 5'-концевые фосфоноэфирные группы, то их необходимо предварительно удалить с помощью фосфоноэстеразы. Одноцепочечные молекулы секвенируют сразу после введения метки. Дуплексную ДНК после осуществления киназной реакции разделяют на две комплементарные цепи с помощью денатурации и электрофореза (разд. 7.2.в). Меченый двухцепочечный фрагмент можно

также разрезать на две части по определенному внутреннему эндонуклеазному сайту и образовавшиеся фрагменты разделить с помощью электрофореза. В этом случае секвенируемые молекулы на самом деле являются двухцепочечными, но одна из цепей не содержит метки и поэтому остается «невидимой».

Введение метки в 3'-концы дуплексных цепей целесообразно проводить в том случае, когда на 5'-концах имеются выступы. Удлинение более

$$5'\text{-pArGpCpTpTpTpCpTpGpArGpArArApCpTpGpCpTpCpTpG-3'}$$

5'-Набор	3'-Набор
5'-pArGp	pTpTpTpCpTpGpAGpArArApCpTpGpCpTpCpTpG-3'
5'-pArGpCpTpTpTp	pTpGpArGpAGpArArApCpTpGpCpTpCpTpG-3'
5'-pArGpCpTpTpTpCpTpGpArGpArArAp	pTpGpCpTpCpTpG-3'
5'-pArGpCpTpTpTpCpTpGpArGpArArApCpTpGp	pTpCpTpG-3'
5'-pArGpCpTpTpTpCpTpGpArGpArArApCpTpGpCpTp	pTpG-3'

**РИС. 7.7.**

Образование родственных наборов цепей при химическом секвенировании. В данном примере модифицируются и элиминируются остатки дезоксицитидина. Для

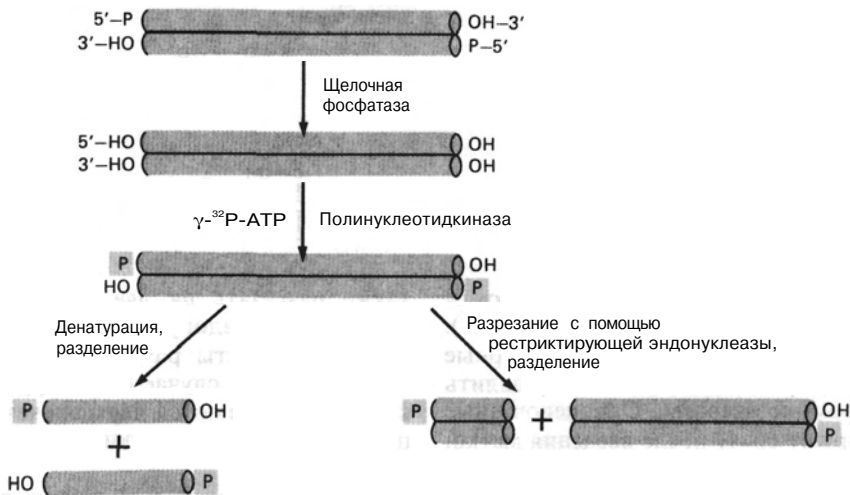
секвенирования могут использоваться или 5'-, или 3'-наборы в зависимости от того, какой из концов (5' или 3') исходной ДНК был радиоактивно помечен.

коротких 3'-концов катализируется либо ДНК-полимеразой I, либо обратной транскриптазой, использующими выступающий 5'-конец как матрицу;  $^{32}\text{P}$  включается в цепь в составе соответствующего  $\alpha\text{-}^{32}\text{P}$ -меченного дезоксинуклеозидтрифосфата (рис. 7.9, А). Мечение по 3'-концу происходит в каждой из цепей двухцепочечной ДНК, а для секвенирования необходимо иметь препарат, в котором все молекулы несут метку только на одном конце. Поэтому меченые двухцепочечные фрагменты расщепляют рестриктазой и фракционируют субфрагменты с помощью гель-электрофореза или разделяют цепи ДНК. Однако, если концы дуплексного фрагмента не идентичны, часто оказывается возможным прямое получение одноцепочечного  $^{32}\text{P}$ -меченного конца. Например, если один из концов тупой, то его удлинения не происходит (рис. 7.9, Б), и при соответствующем подборе  $\alpha\text{-}^{32}\text{P}$ -меченного дезоксинуклеозидтрифосфата можно пометить

только один из концов (рис. 7.9, В). На рис. 7.10 представлены результаты химического секвенирования.

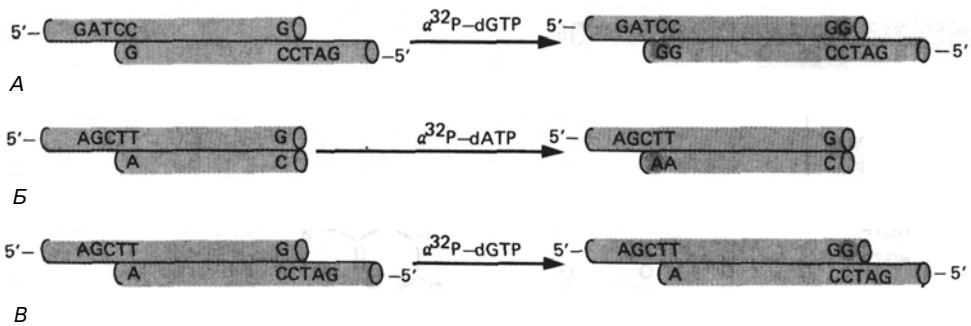
## в. Ферментативное секвенирование

*Получение наборов фрагментов, имеющих общие концы.* Для секвенирования ДНК с помощью ферментативного копирования необходимо иметь одноцепочечную ДНК, а также короткий полидезоксинуклеотидный праймер, комплементарный небольшому участку ДНК. Одиночная цепь служит матрицей, а новая цепь наращивается, начиная с 3'-гидроксильной группы праймера путем присоединения дезоксирибонуклеотидтрифосфатов. При секвенировании проводят четыре типа реакций копирования, в результате чего образуются четыре набора цепей, синтез которых остановлен на остатках А, С, G или Т. Остановка синтеза происходит в результате

**РИС. 7.8.**

Введение 5'-концевой метки в двухцепочечную ДНК, используемую для химического секвенирования. Пред-

ставлены два способа получения фрагментов, содержащих радиоактивную метку только в одной цепи.

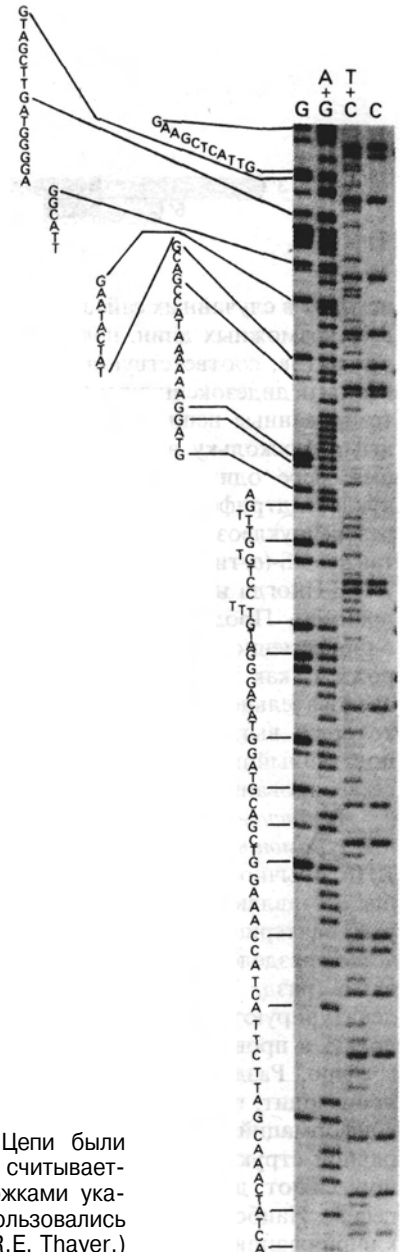
**РИС. 7.9.**

Стратегия мечения 3'-гидроксильного конца фрагментов ДНК путем заполнения пробелов  $\alpha$ - $^{32}\text{P}$ -мечеными дезоксирибонуклеозидтрифосфатами с помощью ДНК-полимеразы I или обратной транскриптазы.

присоединения дидезоксирибонуклеотида, у которого отсутствует 3'-гидроксильная группа, что препятствует дальнейшему удлинению цепи. Реакции проводят таким образом, чтобы вероятность остановки синтеза на любом другом остатке, кроме заданного, была очень мала.

ДНК-полимераза I (а также обратная транскриптаза) может утилизировать 2',3'-дидезоксирибонуклеозидтрифосфатные аналоги нормальных дезоксирибонуклеозидтрифосфатных субстратов. Как и при обычной полимеризации, соответствующий дидезоксирибонуклеотид присоединяется к 3'-гидроксильному концу праймерной цепи (рис. 7.11), однако образующийся при этом продукт не может служить праймером для дальнейшего роста цепи, и реакция останавливается.

Для проведения четырех разных реакций копирования комплекс «матрица-праймер» инкубируют в присутствии всех четырех обычных дезоксирибонуклеозидтрифосфатов и какого-то определенного дидезоксирибонуклеозидтрифосфата. Например, в одной реакции используют ddATP, dATP, dGTP, dCTP, dTTP, в другой – dATP, ddGTP, dGTP, dCTP, dTTP и т.д. Удлинение цепи с помощью матричного копирования происходит до тех пор, пока вместо очередного дезоксирибонуклеотида не присоединится дидезоксирибонуклеотид; в этот момент рост цепи прекращается. Поскольку терминация синтеза про-

**РИС. 7.10.**

Результаты химического секвенирования. Цепи были помечены на 5'-концах. Последовательность считывается в направлении 5'→3'. Надписи над дорожками указывают, какие специфические реакции использовались в каждом случае. (С любезного разрешения R.E. Thayer.)

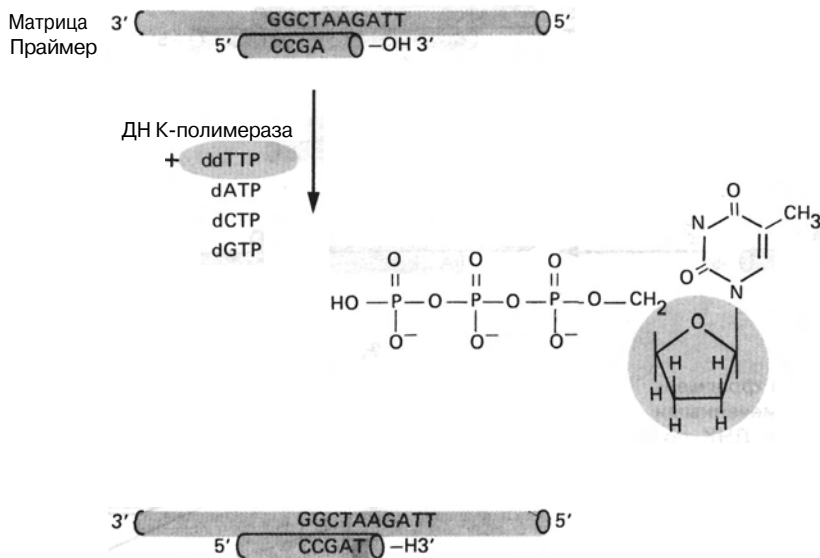


РИС. 7.11.

Присоединение дидезоксинуклеотида к растущей цепи останавливает синтез ввиду отсутствия 3'-гидроксильной группы, необходимой для продолжения процесса.

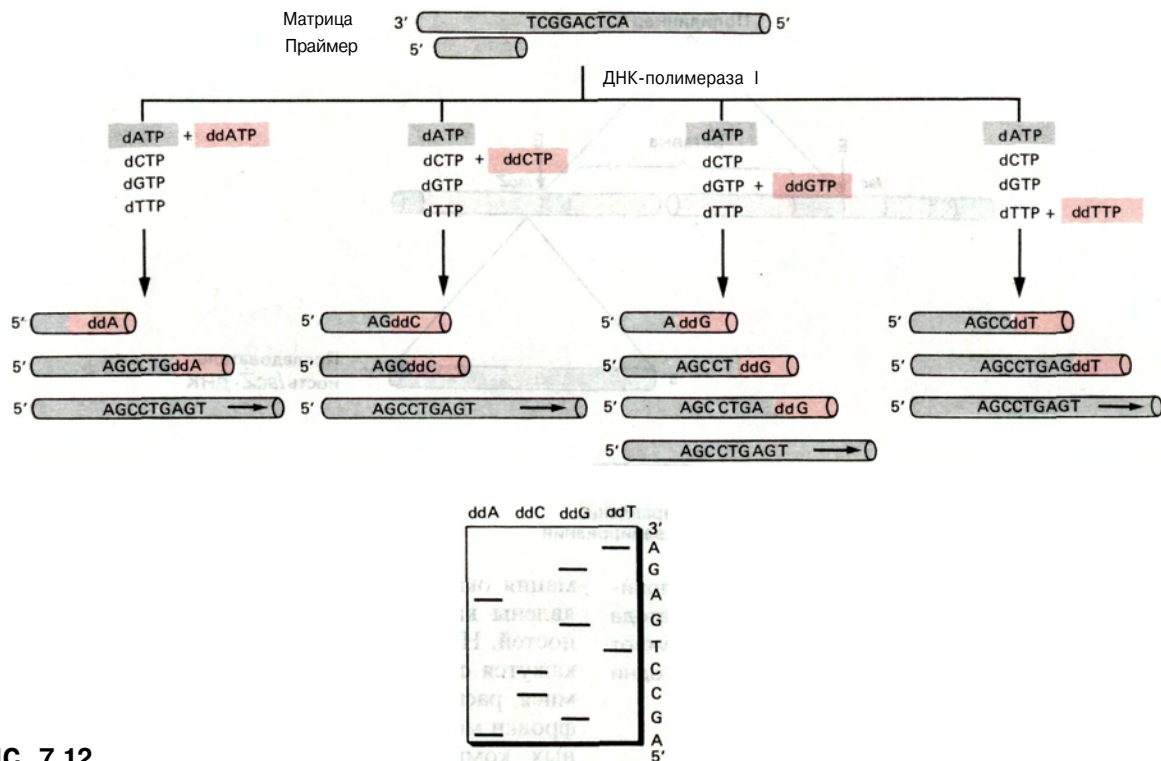
исходит в случайных сайтах, образуется набор цепей всех возможных длин, начиная с 5'-конца праймера до сайтов, соответствующих положению присоединенного дидезоксинуклеотида (рис. 7.12). Новосинтезированные цепи являются радиоактивно мечеными, поскольку в реакции используется по меньшей мере один радиоактивно меченный дезоксинуклеозидтрифосфат. Часто им является  $\alpha$ - $^{32}$ P-дезоксинуклеозидтрифосфат, но это может быть также  $^{35}$ S-( $\alpha$ -тио)-дезоксинуклеозидтрифосфат (рис. 7.13). Иногда используется радиоактивно меченный праймер. Продукты четырех реакций подвергают одновременному электрофорезу на отдельных дорожках, как при химическом секвенировании. Последовательность считывают с радиоавтографа, который выглядит так же, как и радиоавтограф, полученный при использовании химического метода секвенирования (рис. 7.10).

*Получение одноцепочечной ДНК для секвенирования с помощью дидезокси-терминаторов.* Молекулы ДНК обычно являются двухцепочечными; исключения составляют лишь одноцепочечные ДНК некоторых бактериофагов. Как мы уже говорили, физическое разделение цепей дуплекса не всегда осуществимо (разд. 7.2.a). Для разделения цепей дуплекс денатурируют нагреванием или обработкой щелочью и проводят гель-электрофорез или хроматографию. Разделение двух комплементарных цепей происходит, по-видимому, благодаря различиям их конформаций, обусловленным различиями во вторичной структуре. Если разделения не произошло, применяют другие методы получения одиночных цепей. Наиболее эффективный из них состоит в клонировании фрагмента ДНК в векторе, скон-

струированном на основе бактериофага M13 (разд. 5.3.e). Любая клонируемая вставка всегда будет фланкирована одними и теми же последовательностями векторной ДНК (рис. 7.14), что позволяет использовать стандартные праймеры. Их синтезируют химическими методами, и всегда можно выбрать удобный праймер. Кроме того, не составляет труда получить отдельные клоны, каждый из которых содержит одну из двух комплементарных цепей ДНК (разд. 5.3.e), благодаря чему легко осуществить секвенирование обеих цепей.

Комбинируя клонирование в M13 и секвенирование с помощью дидезокси-терминатора, мы можем получить эффективный способ определения последовательностей молекул ДНК, длина которых достигает нескольких тысяч пар нуклеотидов. Для секвенирования длинных дуплексных молекул их разрезают механически на фрагменты и затем проводят клонирование. Каждая образующаяся бляшка содержит определенные фрагменты или цепи. Случайно отбирая клоны, определяют последовательность вставок. Поскольку вставки образуются в результате разрыва исходной молекулы ДНК в случайных местах, некоторые фрагменты содержат перекрывающиеся последовательности. С помощью компьютера сравнивают последовательность каждого фрагмента с последовательностями других фрагментов (разд. 7.3), выявляют перекрывающиеся области и располагают отдельные фрагменты в определенном порядке. Таким способом были определены последовательности митохондриальной ДНК человека (16569 пар нуклеотидов) и ДНК  $\lambda$ -фага (48502 пары нуклеотидов).

*Другое применение методов секвенирования.* Сек-

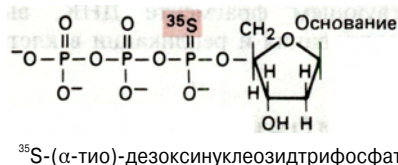
**РИС. 7.12.**

Дидезокси-метод секвенирования. Приведено схематическое изображение получающегося радиоавтографа. Использовался радиоактивно меченный  $\alpha$ - $^{32}\text{P}$ -дезокси-АТФ.

венирование с помощью дидезокси-терминаторов может применяться для прямого определения последовательностей РНК или ДНК, причем даже в том случае, если в смеси помимо секвенируемой цепи присутствуют неродственные молекулы. Важным элементом при этом является праймер, комплементарный небольшому интересующему нас участку и лишь с малой вероятностью ассоциирующий с другими молекулами, присутствующими в смеси. Подобной специфичностью обычно обладает праймер длиной 20 оснований с уникальной последовательностью. При таких условиях с помощью обратной транскриптазы синтезируют кДНК. Этот метод применим только в том случае, если мы уже достаточно много знаем об интересующих нас РНК или ДНК, но круг таких систем все время расширяется.

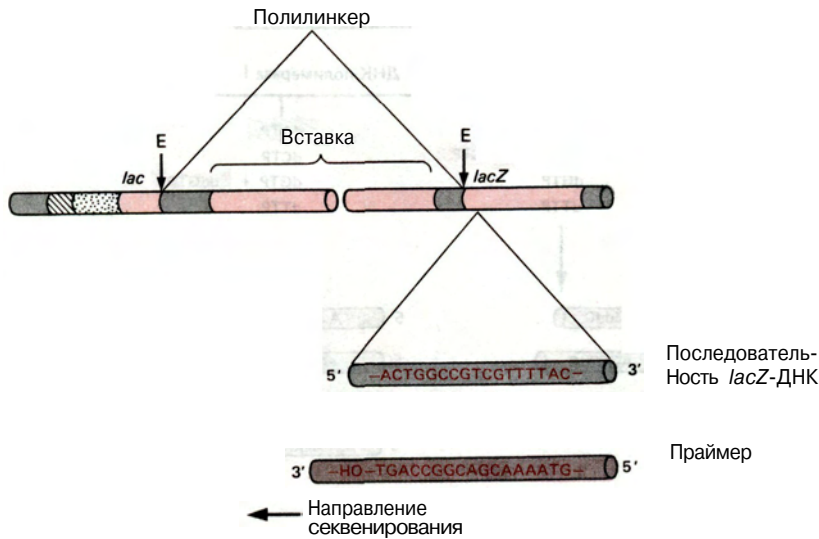
Метод химического секвенирования может применяться для анализа специфических последовательностей ДНК, присутствующих в смеси наряду с большим числом разных молекул ДНК, в том числе даже если речь идет об одном гене в суммарной геномной ДНК млекопитающих. Прежде всего ДНК разрезают с помощью рестриктирующей эндонуклеазы, так что интересующий нас сегмент

оказывается во фрагменте длиной в несколько сотен пар оснований. Затем проводят химические реакции, специфичные к определенным основаниям, в результате чего образуются четыре набора фрагментов. Такие наборы включают все возможные по длине фрагменты с интересующей нас областью, начинающиеся в сайте рестрикции, по которому была разрезана ДНК. Здесь же присутствует множество других фрагментов из остальной части генома. Фрагменты разделяют путем электрофореза в секвенирующем геле, затем переносят на нейлоновый фильтр и отжигают с  $^{32}\text{P}$ -меченным зондом, который комплементарен последовательности-мишени, находящейся вблизи одного из сайтов раз-

**РИС. 7.13.**

Для идентификации синтезированных цепей в дидезокси-методе используется один или несколько радиоактивно меченных дезоксирибонуклеозидтрифосфатов. Обычно это либо  $\alpha$ - $^{32}\text{P}$ -трифосфат, либо  $^{35}\text{S}$ -( $\alpha$ -тио)-трифосфат (представленный здесь). Отметим, что замещение атома кислорода атомом серы при  $\alpha$ -Р лишь незначительно изменяет способность молекулы служить субстратом для ДНК-полимеразы I.



**РИС. 7.14.**

Секвенирование вставок одноцепочечной ДНК, клонированных в векторе M13, с помощью дидезокси-метода (см. разд. 5.3.е, где описан этот вектор). Независимо от вставки при секвенировании могут использоваться одни и те же одноцепочечные праймеры.

резания. Зонд ассоциирует со всеми фрагментами, образовавшимися начиная с одного конца последовательности-мишени, в результате чего на радиоавтограмме образуется типичная «секвенирующая лестница». Таким методом можно определить, например, нуклеотидную последовательность мутантных форм ранее клонированного гена. Он может использоваться также для определения характера метилирования специфической, уже клонированной области генома, поскольку гидразин реагирует с 5'-метилцитозиновыми остатками менее эффективно, чем с цитозиновыми и тиминовыми. О наличии 5'-метилцитозина в геле можно судить по отсутствию цитозинового остатка, представленного в соответствующем фрагменте ДНК, выделенном после клонирования и репликации в клетках *E. coli* (см. рис. 4.13).

### 7.3. КОМПЬЮТЕРНЫЙ АНАЛИЗ НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК

Простота разделения сегментов ДНК путем клонирования и их последующий анализ позволили определить нуклеотидную последовательность многих ДНК. Но хотя стремление биологов к получению точных в химическом отношении данных продолжало стимулировать эти исследования, инфор-

мация оказывалась бесполезной, если не были выявлены какие-то особенности этих последовательностей. На первый взгляд длинные ряды оснований кажутся совершенно загадочными и не поддающимися расшифровке. К счастью, проблему расшифровки можно решить с помощью высокоэффективных компьютерных программ и для рутинного использования имеются программы на нескольких стандартных компьютерных языках. Возможности различных программ в значительной степени перекрываются. Их можно использовать при работе как на больших компьютерах, так и на мини- и микрокомпьютерах. Вообще говоря, эти программы пригодны для поиска как структурных, так и биологических особенностей последовательностей.

#### а. Хранение информации о первичной структуре

Для ввода и хранения данных о последовательностях и последующего поиска, проводимого исследователем вручную или с помощью специальной аналитической программы, используются стандартные процедуры редактирования. Большинство этих процедур предусматривают также возможность коррекции хранящейся информации. Ввод данных часто осуществляют путем типирования последовательности в компьютерном терминале после прочтения и регистрации данных секвенирующего геля. Однако имеются также программы для прямого ввода исходных данных. В принципе такие программы позволяют свести к минимуму ошибки, допускаемые при неправильном считывании гелей или неточной регистрации данных. Существует два подхода к такой автоматизации. В первом случае детектор автоматически сканирует радиоавтографы

Нуклеотидная последовательность:

```

5109      5119      5129      5139      5149      5159
TCCTTTCAAG ACCTAGAAGG TCCATTAGCT GCAAAGATTC CTCTCTGTTT AAAACTTTAT

5169      5179      5189      5199      5209      5219
CCATCTTTGC AAAGCTTTTT GCAAAAAGCCT AGGCCTCCAA AAAAGCTCC CCACTACTTC

5229      5239      1      6      16      26      36
TGGAATAGCT CAGAGGCCGA GCGCGCCTCG GCCTCTGCAT AAATAAAAAA AATTAGTCAG

46      56      66      76      86      96
CCATGGGGCG GAGAATGGGC GGAAGTGGGC GGAGTTAGGG GCGGGATGGG CGGAGTTAGG

106      116      126      136      146      156
GGCGGGACTA TGGTTGCTGA CTAATTGAGA TGCATGCTTT GCATACTTCT GCCTGCTGGG

166      176      186      196      206      216
GAGCCTGGGG ACTTCCACA CCTGGTTGCT GACTAATTGA GATGCATGCT TTGCATACTT

226      236      246      256
CTGCCTGCTG GGGAGCCTGG GGACTTTCCT CACCCTAACT GACA

```

Комплементарная цепь:

```

251      241      231      221      211      201
TGTCAGTTAG GGTGTGGAAG GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG

191      181      171      161      151      141
CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT

131      121      111      101      91      81
ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCATAGTCC CGCCCCTAAC TCGCCCATC

71      61      51      41      31      21
CCGCCCTAA CTCCGCCAG TTCGCCCAT TCTCCGCCC ATGGCTGACT AATTTTTTTT

11      1      5234      5224      5214      5204
ATTTATGCAG AGGCCGAGGC CGCCTCGGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC

5194      5184      5174      5164      5154      5144
TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTTTGCAAAG ATGGATAAAG TTTTAAACAG

5134      5124      5114      5104
AGAGGAATCT TTGCAGCTAA TGGACCTTCT AGGTCTTGAA AGGA

```

Последовательность  
содержит 404 нуклеотида

95 (23,5)	Аденин
96 (23,8)	Цитозин
110 (27,2)	Гуанин
103 (25,5)	Тимин
0	Другие

Распределение динуклеотидов:

	Число	Проценты	Ожидаемые проценты
AA	32	7,9	5,5
AC	18	4,5	5,6
AG	24	6,0	6,4
AT	20	5,0	6,0
CA	21	5,2	5,6
CC	23	5,7	5,6
CG	9	2,2	6,6
CT	43	10,7	6,1
GA	24	6,0	6,4
GC	35	8,7	6,5
GG	44	10,9	7,4
GT	7	1,7	6,9
TA	18	4,5	6,0
TC	20	5,0	6,1
TG	33	8,2	6,9
TT	32	7,9	6,5

### РИС. 7.15.

Секвенирование области генома SV40, включающей точку начала репликации. Числа обозначают то же, что и на рис. 5.34. Остаток 1 находится непосредственно в этой точке и соседствует с остатком 5243 в кольцевом геноме. Вся последовательность читается в направлении 5'→3', слева направо, сверху вниз. Это типичные дан-

ные на выходе компьютера; сюда входят: последовательность, введенная в компьютер, комплементарная последовательность, нуклеотидный состав, частота встречаемости нуклеотидов и частоты, ожидаемые для последовательности со случайным распределением тех же нуклеотидов.

и передает данные непосредственно в компьютер. Во втором данные автоматически регистрируются от сенсора, управляемого вручную. При втором подходе исследователь является арбитром при любых затруднениях.

## б. Структурный анализ

*Первичная структура.* Если имеются данные о нуклеотидной последовательности одной цепи, то большинство программ позволяют построить комплементарную цепь, вычислить нуклеотидный состав, выявить участки, богатые пуринами, пиримидинами или определенными сочетаниями оснований, и определить частоту встречаемости различных динуклеотидов (рис. 7.15). Могут быть выявлены специфические субпоследовательности в пределах определенного сегмента, что часто используется для нахождения сайтов для рестриктирующих эндонуклеаз (рис. 7.16). В программу введена информация о сайтах узнавания для известных ферментов, и по одной команде выдаются сведения о положении этих сайтов для каждого из ферментов, числе ожидаемых фрагментов, образующихся при расщеплении ими ДНК, размере каждого фрагмента в парах оснований и его процентном отношении к общей длине сегмента, а также о нуклеотидных остатках, соответствующих концам каждого фрагмента. Существуют также программы, которые могут предсказать, какие продукты будут получены при совместном действии двух или нескольких эндонуклеаз. При этом предсказания делаются как для линейных, так и для кольцевых молекул. Полученная информация может использоваться для подтверждения данных секвенирования путем сравнительного анализа ожидаемого и реального результата действия эндонуклеаз.

Имеющиеся программы осуществляют также поиск характерных особенностей последовательностей, включая прямые (рис. 7.17) и обратные (рис. 7.18) повторы. Выявляются не только полностью совпадающие сегменты, но и сегменты с той или иной степенью несовпадения; для этого допускаются определенные отклонения параметров, заложенных в программу, от фиксированного значения. В примере, представленном на рис. 7.17, перечислены все повторы длиной не менее шести пар оснований и гомологичные не менее чем на 75%. Предполагается, что максимальный размер образующихся петель равен двум основаниям.

Можно получить данные о гомологии или частичной гомологии различных последовательностей. Эти данные могут касаться структуры одного и того же гена у двух разных организмов или родственных генов одного организма. Такой сравнительный анализ широко используется при

изучении эволюции на молекулярном уровне. Для сравнения двух последовательностей между собой или одной последовательности с группой других последовательностей разработаны специальные алгоритмы. Может использоваться описанный ранее вероятностный анализ. По результатам сравнения информации о последовательностях, полученной в разных лабораториях, был создан центральный банк данных, из которого всегда можно затребовать информацию о тех или иных последовательностях. К 1989 г. в банках имелась информация о последовательностях более 20 миллионов пар оснований, представляющая данные о многих генах и организмах.

Помимо этого общего применения, обнаружение сходства различных последовательностей является

	Сайты		Фрагменты		
	#	п.н.	Длина	Концы	
ALU 1, OXA 1, (AGCT)	3		п.н. %		
		5126	277 (68,6)	5226	260
		5172	54 (13,4)	5172	5226
		5226	46 (11,4)	5126	5172
		27 ( 6,7)	5100	5126	
AVA 3 (ATGCAT)	2	126	270 (66,8)	5100	126
		198	72 (17,8)	126	198
			62 (15,3)	198	260
BGL 1 (GCCNNNNNGGC)	1	5235	268 (66,3)	5235	260
			136 (33,7)	5100	5235
DDE 1 (CTNAG)	1	5228	275 (68,1)	5228	260
			129 (31,9)	5100	5228
HAE 3 (GGCC)	4	5191	254 (62,9)	6	260
		5234	92 (22,8)	5100	5191
		5243	43 (10,6)	5191	5234
		6	9 ( 2,2)	5234	5243
			6 ( 1,5)	5243	6
HIND 3 (AAGCTT)	1	5171	332 (82,2)	5171	260
			72 (17,8)	5100	5171
HINF 1 (GANTC)	1	5135	368 (91,1)	5135	260
			36 ( 8,9)	5100	5135

**РИС. 7.16.**

Компьютерные данные о положении некоторых сайтов для рестриктирующих эндонуклеаз в сегменте ДНК SV40, представленном на рис. 7.15. Указаны число сайтов (#) для каждого фермента, их положение (п.н.), а также длина каждого фрагмента, образующегося при расщеплении исходного фрагмента, как в числе пар оснований, так и в процентном отношении от длины всего сегмента (в скобках). Указаны номера остатков на концах фрагментов.

```

49 GAATGGGCGGA 59
81 GA TGGGCGGA 90
0.909 a
3E-05 b
1E+00 c

52 TGGGCGGA 59
62 TGGGCGGA 69
1.000
4E-05
2E+00

53 GGGCGGAACT 62
96 GGGCGGGACT 105
0.900
9E-05
4E+00

57 GGAACTGGGCGGAGTTAGGGGCG 79
80 GGA TGGGCGGAGTTAGGGGCG 100
0.913
0E+00
0E+00

62 TGGGCGGAGTTAGGGGCGGGA 82
83 TGGGCGGAGTTAGGGGCGGGA 103
1.000
0E+00
0E+00

.107 TGGTTGCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGTGGGGAGCCTGGGGACTTTCACACC 178
179 TGGTTGCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGTGGGGAGCCTGGGGACTTTCACACC 250
1.000
0E+00
0E+00

127 TGCATGCTT TGCATACT 143
208 TGCATACTTCTGCCTGCT 225
0.778
1E-05
5E-01

136 TGCATACTTCTGCCTGCT 153
199 TGCATGCTT TGCATACT 215
0.778
1E-05
5E-01

```

THE NUMBER OF MATCHES IS 22

```

MAXPROB= 9.99E-05 EXPECT= 4 MINRATIO=7.50E-01;
LOOPLNGTH= 2 DISTANCE= 404;

```

составной частью секвенирования очень длинных последовательностей с помощью дидезокси-метода, объединенного с клонированием в фаге M13 (разд. 7.2.в). Используя компьютер для обнаружения перекрывающихся последовательностей, мы не только делаем работу менее утомительной, но и можем

THE DYAD SYMMETRIES ARE:

```

5177 TTTGCAAA 5184
5172 AAACGTTT 5165
1.000
4E-05
2E-01

1 GCCTCGGCCTCTG 13
5242 CGGAGCCGGAGAC 5230
1.000
0E+00
0E+00

```

THE NUMBER OF MATCHES IS 2

```

MAXPROB= 9.99E-05 EXPECT= 0
LOOPLNGTH= 2

```

```

MINRATIO= 7.50E-01;
LOOPDIST= 20;

```

### РИС. 7.17.

Компьютерные данные о некоторых повторах, встречающихся в сегменте ДНК SV40, представленном на рис. 7.15. Программа не учитывала повторы длиной менее 7 п.н., повторы, гомологичные менее чем на 75% (MINRATIO), а также повторы, вероятность обнаружения которых (MAXPROBE) при такой же степени совпадений в случайной последовательности превышает  $1 \cdot 10^{-4}$  (это условие записывается в виде 9.99E-05). Реальная гомологичность данного повтора (в %) обозначена на рисунке буквой *a*, а вероятность его появления - буквой *b*. Ожидаемое число повторов с такой же гомологией в последовательности со случайным распределением оснований и такими же нуклеотидным составом и длиной обозначено буквой *c*. Максимально допустимый размер петли равен двум нуклеотидам. При  $b < 1 \cdot 10^{-5}$  (1E-05) этот параметр принимается равным 0, и с тоже автоматически становится равным нулю. В анализируемой последовательности выявлено 22 повтора (на рисунке приведено 8 из них), в то время как при случайном распределении оснований их ожидаемое число равно 4.

быстро решить, следует ли нам пытаться получить дополнительные данные.

*Вторичная структура.* Разработаны программы, позволяющие предсказать стабильную внутри-молекулярную вторичную структуру одноцепочечных РНК или ДНК. Они позволяют, например, построить модель укладки цепи тРНК и рРНК, и многие из таких моделей получили экспериментальное подтверждение в опытах с использованием нуклеаз, специфичных к одноцепочечным участкам. Расчеты основаны на предположениях о вероятности образования определенных пар оснований, на термодинамических свойствах разных пар оснований и данных о стабильности спирали.

## в. Биологическое значение

С помощью компьютерных программ можно перевести информацию с языка нуклеотидов на язык аминокислот в соответствии с правилами генетического кода. При этом указываются все три возможные рамки считывания и стоп-кодоны трансляции. В тех случаях, когда аминокислотная последовательность кодируемого полипептида известна, идентифицировать правильную рамку считывания не составляет труда. В противном случае правиль-

### РИС. 7.18.

Компьютерные данные об обратных повторах в сегменте ДНК SV40, изображенном на рис. 7.15. Обозначения такие же, как на рис. 7.17.

Трансляция начинается с позиции 1

```

5' GGC CTC GGC CTC TGA GCT ATT CCA GAA 5214 5184
   ARG LEU GLY LEU END ALA ILE PRO GLU VAL VAL ARG ARG LEU PHE TRP ARG PRO ARG LEU

   TTG CAA AAA GCT TTG CAA AGA TGG ATA AAG TTT TAA ACA GAG AGG AAT CTT TGC AGC TAA 5124
   LEU GLN LYS ALA LEU GLN ARG TRP ILE LYS PHE END THR GLU ARG ASN LEU CYS SER ASN

   TGG ACC TTC TAG GTC TTG AAA GGA GTG CCT GGG GGA ATA TTC CTC TGA TGA GAA AGG CAT 5094 5064
   TRP THR PHE END VAL LEU LYS GLY VAL PRO GLY GLY ILE PHE LEU END END GLU ARG HIS
  
```

Трансляция начинается с позиции 2

```

5' GCC TCG GCC TCT GAG CTA TTC CAG AAG 5213 5183
   ALA SER ALA SER GLU LEU PHE GLN LYS END END GLY GLY PHE PHE GLY GLY LEU GLY PHE

   TGC AAA AAG CTT TGC AAA GAT GGA TAA AAG TTT AAA CAG AGA GGA ATC TTT GCA GCT AAT 5123
   CYS LYS LYS LEU CYS LYS ASP GLY END SER PHE LYS GLN ARG GLY ILE PHE ALA ALA ASN

   GGA CCT TCT AGG TCT TGA AAG GAG TGC CTG GGG GAA TAT TCC TCT GAT GAG AAA GGC ATA 5093 5063
   GLY PRO SER ARG SER END LYS GLU CYS LEU GLY GLU TYR SER SER ASP GLU LYS GLY ILE
  
```

Трансляция начинается с позиции 3

```

5' CCT CGG CCT CTG AGC TAT TCC AGA AGT AGT GAG GAG GCT TTT TTG GAG GCC TAG GCT TTT 5212 5182
   PRO ARG PRO LEU SER TYR SER ARG SER SER GLU GLU ALA PHE LEU GLU ALA END ALA PHE

   GCA AAA AGC TTT GCA AAG ATG GAT AAA 5122 5122
   ALA LYS SER PHE ALA LYS MET ASP LYS VAL LEU ASN ARG GLU GLU SER LEU GLN LEU MET

   GAC CTT CTA GGT CTT GAA AGG AGT GCC TGG 5092 5062
   ASP LEU LEU GLY LEU GLU ARG SER ALA TRP GLY ASN ILE PRO LEU MET ARG LYS ALA TYR
  
```

**РИС. 7.19.**

Компьютерные данные о трансляции части генома SV40 (остатки 5243-5062, рис. 5.34) при трех возможных рамках считывания. Представленная цепь ДНК имеет такую же последовательность, как и мРНК для ранних

белков (смысловая цепь). Правильной является третья рамка считывания. Иницирующий кодон ATG (AUG) выделен цветом. Стоп-сигнал трансляции обозначен как «end».

ную рамку можно выбрать, исходя из ее длины. Рамки, в которых часто встречаются стоп-кодоны, вряд ли могут считаться правильными. На рис. 7.19 показаны часть генома SV40, кодирующая участок вблизи NH<sub>2</sub>-конца малого и большого Т-антигенов (см. рис. 5.34), и полученные с помощью компьютера аминокислотные последовательности для всех трех рамок считывания. Правильной является третья рамка. Можно также рассчитать частоту, с которой встречается каждый кодон, и таким образом выявить предпочтительные кодоны для определенных аминокислот. Обращение этой процедуры позволяет воссоздать возможные нуклеотидные последовательности, исходя из известной аминокислотной последовательности полипептида, хотя эта задача не имеет однозначного решения вследствие вырожденности генетического кода. Центральный банк данных содержит информацию об аминокислотных последовательностях всех проанализированных полипептидов, что позволяет сравнивать

последовательности изучаемого и известных белков.

С помощью компьютеров можно вести поиск специфических нуклеотидных последовательностей, например промоторов или участков связывания с рибосомами. Они позволяют также обнаружить последовательности со специфическими свойствами. Например, наличие сходных последовательностей в разных неродственных во всех других отношениях генах или вблизи них наводит на мысль об их возможной регуляторной роли. Точно так же присутствие гомологичных последовательностей в кодирующих областях различных генов может свидетельствовать об общности эволюционного происхождения этих генов. Данный вопрос детально рассмотрен в ч. III книги. Однако важно осознавать, что статистический анализ еще не является достаточным для того, чтобы делать вывод о биологической роли той или иной последовательности. Об этом можно говорить только после проведения независимого биологического исследования.

## 7.4. ОПРЕДЕЛЕНИЕ ПОЛОЖЕНИЯ КЛОНИРОВАННЫХ СЕГМЕНТОВ В ГЕНОМАХ

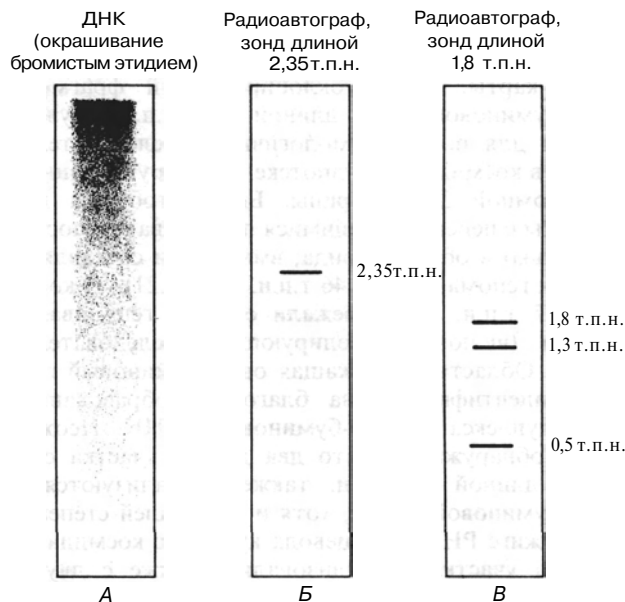
В идеале установление положения клонированного сегмента ДНК означает определение фланкирующих его последовательностей и того хромосомного сайта, по которому этот сегмент был включен в хромосому. Конечная цель такого картирования состоит в полном описании структуры хромосомы.

### а. Молекулярная локализация

#### *Характеристика клонированного сегмента.*

Прежде всего необходимо доказать, что клонированный сегмент действительно является репликой определенного геномного сегмента. Существует ли в геноме гомологичный сегмент с точно так же расположенными сайтами для рестриктирующих эндонуклеаз? Этот вопрос имеет важное значение, поскольку во время репликации в системе хозяин-вектор иногда происходит клонирование случайных последовательностей, образовавшихся в результате рекомбинаций, делеций или мутаций. Основной подход состоит в использовании клонированной ДНК в качестве зонда для анализа продуктов эндонуклеазного расщепления суммарной геномной ДНК, из которой произошел данный клонированный сегмент. Такой подход аналогичен описанному в разд. 6.1.6 с тем отличием, что зондом в данном случае является клонированная вставка, меченная  $^{32}\text{P}$ . В эксперименте, результаты которого иллюстрирует рис. 7.20, *Б*, клонированный зонд представлял собой субклонированный сегмент (2,35 т.п.н.) овальбуминового гена, показанный на рис. 7.2; материал, подвергнутый электрофорезу, — это *EcoRI*-гидролизат ДНК курицы (рис. 7.20, *А*). С зондом гибридизовался материал только одной полосы, представляющий собой фрагменты длиной 2,35 т.п.н. Этот эксперимент показывает, что 1) клонированная вставка в самом деле является фрагментом ДНК курицы; 2) в геноме гибридизующийся сегмент расположен между двумя *EcoRI*-сайтами, находящимися друг от друга на расстоянии 2,35 т.п.н.; 3) ни один из фрагментов любого другого размера не гибридизуется с зондом. Таким образом, клонированный фрагмент длиной 2,35 т.п.н., по-видимому, представляет собой истинную геномную последовательность и два аллеля в диплоиде являются идентичными, поскольку локализация этих *EcoRI*-сайтов совпадает.

Еще одна последовательность овальбуминового гена обнаружена в *EcoRI*-фрагменте длиной 1,8 т.п.н. После того как он был клонирован и использован в качестве зонда для анализа *EcoRI*-гидро-



**РИС. 7.20.**

Отжиг  $^{32}\text{P}$ -меченных клонированных сегментов овальбуминового гена длиной 2,35 т.п.н. (*Б*) и 1,8 т.п.н. (*В*) с *EcoRI*-фрагментами суммарной ДНК цыпленка (*А*).

лизата ДНК курицы, были обнаружены три гибридизующихся фрагмента (рис. 7.20, *В*) длиной 1,8; 1,3; 0,5 т.п.н. соответственно. Эти данные свидетельствуют о неидентичности двух аллельных овальбуминовых генов. Один из них содержит дополнительный *EcoRI*-сайт, разделяющий фрагмент длиной 1,8 т.п.н. на две части. Интересно, что столь простая процедура лежит в основе генетического анализа. Наследование двух или более аллелей легко прослеживается с помощью эндонуклеазного расщепления, электрофореза, ДНК-блоттинга и гибридизации между соответствующим зондом и ДНК, выделенной из разных источников. Единственным требованием является наличие полиморфных эндонуклеазных сайтов в аллелях любого гена или сегмента ДНК. Для проведения таких экспериментов достаточное количество ДНК можно получить из очень небольшого кусочка ткани или из нескольких миллилитров крови, поэтому метод может использоваться для анализа ДНК большинства видов, в том числе и человека.

*Построение карты молекулы.* Имея сравнительно короткие сегменты молекулы ДНК, такие, например, как ранее описанный субклонированный фрагмент длиной 2,35 т.п.н., можно построить детальную карту молекулы вплоть до установления ее нуклеотидной последовательности. Более длинные клонированные последовательности, подобные тем, которые присутствуют в  $\lambda$ -векторах или космидах,

часто содержат несколько генов, а также другие сегменты ДНК, и могут использоваться для расширения карты. Так, субклонированный фрагмент овалбуминового гена длиной 2,35 т.п.н. служил зондом для поиска гомологичных последовательностей в космидной библиотеке, сконструированной из геномной ДНК курицы. Были отобраны две космиды с перекрывающимися последовательностями только в области зонда; вместе они составляли участок генома длиной 46 т.п.н. (рис. 7.21), из которых 7,7 т.п.н. принадлежали самому гену овалбумина (включая некодирующие последовательности). Область, содержащая овалбуминовый ген, была идентифицирована благодаря образованию гетеродуплекса с овалбуминовой кДНК. Неожиданно обнаружилось, что два других участка сегмента длиной 46 т.п.н. также гибридизуются с овалбуминовой кДНК, хотя и в меньшей степени. При отжиге РНК из яйцевода курицы с космидами эти два участка гибридизовались также с двумя другими мРНК, отличными от овалбуминовой. Было высказано предположение, что в сегменте длиной 46 т.п.н. локализованы еще два гена, последовательность которых сходна с последовательностью гена овалбумина. Данное предположение получило подтверждение после проведения рестрикционного и гетеродуплексного анализа, а также после определения нуклеотидной последовательности. Эти два гена, обозначенные как X и Y (их функции и продукты остаются неизвестными), транскрибируются, как и овалбуминовый ген, в присутствии стероидных гормонов.

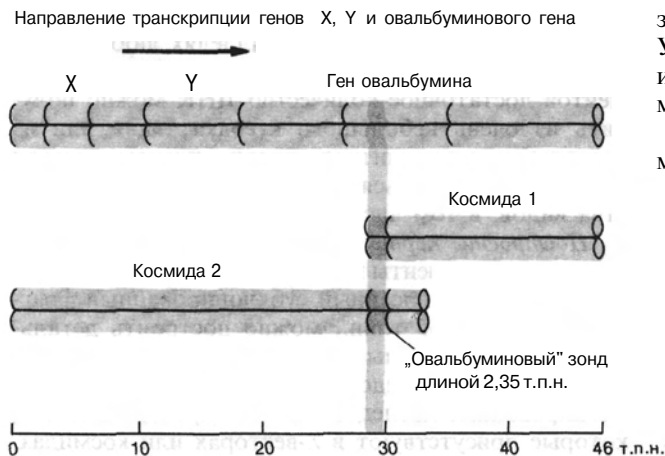
Методы, использующиеся при картировании овалбуминового гена и генов X и Y, лежат в основе широко распространенного подхода к построению детальных генетических карт, получившего название **прогулки по хромосоме**. Уникальный сегмент ДНК, примыкающий к одному из концов клонированного

фрагмента, очищают (путем субклонирования) и используют для зондирования библиотеки геномной ДНК. Одни гибридизующиеся клоны полностью перекрываются с исходным клонированным сегментом, другие же включают новые сегменты, в результате чего карта расширяется. При повторении этой процедуры совершается «прогулка» на большее расстояние. Данный метод не очень удобен для картирования тех геномов, которые содержат большое число рассеянных повторов, поскольку в этом случае трудно очистить уникальные последовательности, используемые в качестве зондов.

## б. Хромосомная локализация

Классическое генетическое картирование основывается на получении определенных мутаций и анализе частот рекомбинаций (см. введение к ч. I). У *Drosophila* генетические карты удалось расширить и уточнить путем установления корреляций между генетическими данными и хромосомными aberrациями типа делеций, инверсий и транслокаций, которые визуально проявляются как изменения в характере исчерченности политенных хромосом. Однако подобные методы непригодны для анализа хромосом большинства растений и животных. Генетический анализ немногочисленных популяций с большим периодом генерации весьма затруднителен, поскольку политенность встречается редко, а хромосомы довольно многочисленны, имеют небольшие размеры и с трудом поддаются идентификации. Например, у млекопитающих установление корреляций между фенотипическими изменениями и делециями, транслокациями и инверсиями позволяет локализовать лишь ограниченное число генов в специфических хромосомах или отдельных их областях. С развитием методов получения клонированных сегментов ДНК были разработаны универсальные процедуры картирования, которые не зависят от фенотипического проявления мутаций. Удалось локализовать многие гены, в том числе и гены человека, в специфических областях хромосом.

Мы рассмотрим два наиболее распространенных метода. С помощью одного из них положение



**РИС. 7.21.**

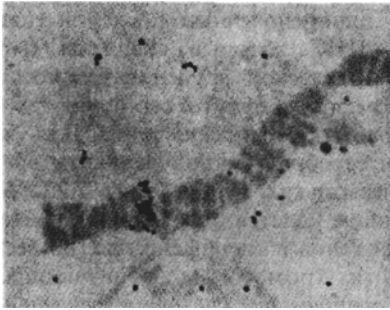
Используя две космиды с перекрывающимися последовательностями, можно построить детальную генетическую карту области овалбуминового гена цыпленка. Сегмент размером 46 т.п.н. помимо гена овалбумина содержит два родственных гена, X и Y, с неизвестными функциями. [A. Royal et al., *Nature*, **279** (1979), p. 125.]

клонированного сегмента ДНК устанавливают по его способности гибридизоваться со специфическим участком в практически интактных хромосомах. Второй позволяет идентифицировать хромосому, которая несет определенную последовательность, путем гибридизации подходящего меченого зонда с гибридными клетками, содержащими различные наборы лишь из нескольких хромосом данного вида. В третьем методе, описанном во введении к ч. IV, используются клонированные фрагменты ДНК для выявления полиморфизма рестрикционных сайтов (для примера см. рис. 7.20). Такой полиморфный маркер применяют затем в качестве генетического маркера для построения карт сцепления эукариотических хромосом—аналогично тому, как в классическом генетическом анализе используют фенотипические маркеры.

*Картирование хромосом с помощью гибридизации.* РНК-зонды или денатурированные ДНК-зонды

могут гибридизоваться с соответствующими последовательностями денатурированной ДНК, входящими в состав хромосом с характерными морфологическими признаками. Положение радиоактивно меченого зонда определяют, нанеся на препарат, подготовленный для микроскопического исследования, чувствительную эмульсию; в том месте эмульсии, которое контактирует с гибридным участком, появляются темные зерна серебра. Проведя микроскопическое исследование всего препарата, можно локализовать зонд. Довольно точные данные удается получить в случае больших политенных хромосом *Drosophila* в основном благодаря тому, что положение темных областей можно соотнести с характером исчерченности хромосом (рис. 7.22, А).

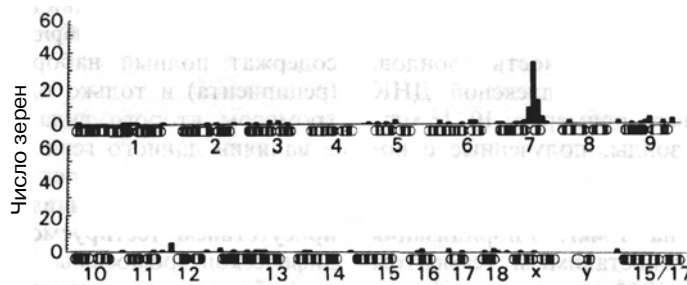
Однако определение положения единичных генов в небольших и весьма многочисленных хромосомах растений и позвоночных путем гибридизации *in situ* затруднено по двум причинам. Во-первых, не всегда легко идентифицировать отдельные хромосомы.



А



Б



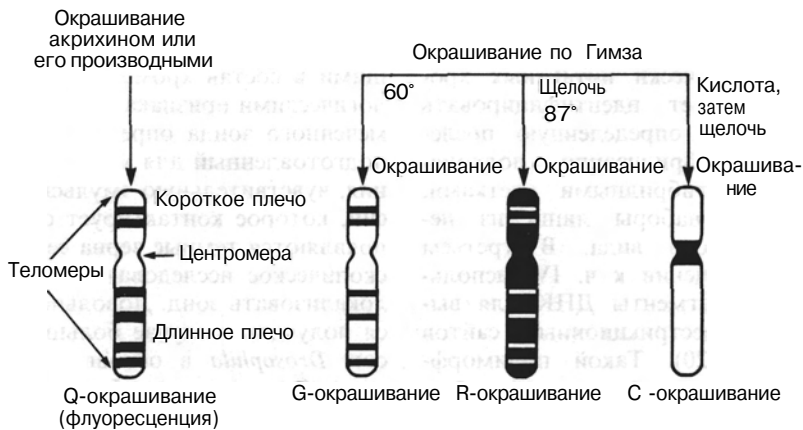
В

РИС. 7.22.

А. Локализация клонированного гена легкой цепи 2 миозина на политенной хромосоме 3 *Drosophila* с помощью гибридизации *in situ*. Темные зерна серебра группируются на участке, гомологичном зонду. [С любезного разрешения J. Toffenetti, D. Mischke, M.L. Pardue. *J. Cell. Biol.*, **104** (1987), p. 19.] Б. Локализация единичной последовательности, присутствующей в хромосоме 7 свиньи, с помощью гибридизации *in situ*. Зондом служил клонированный участок гена, кодирующего главный антиген гистосовместимости. Стрелкой указаны зерна серебра на хромосоме 7. Анализ

распределения зерен по одной хромосоме не позволяет однозначно локализовать искомую последовательность, поскольку отдельные зерна встречаются и в других областях хромосомы. Для этого нужно просмотреть не менее 100 хромосом. На диаграмме рис. В представлено распределение числа зерен по хромосомным сайтам, построенное по результатам анализа более 100 хромосом 7. У всех остальных хромосом встречаются лишь отдельные зерна, случайно разбросанные по разным локусам. [M. Rabin et al., *Cytogenet. Cell. Genet.*, **39** (1985), p. 206.]



**РИС. 7.23.**

Разные методы окрашивания хромосом млекопитающих. Положение полос при Q- и G-окрашивании одинаково, а при R-окрашивании - обратное. Кариотип человека, представленный на рис. 1.9, окрашен по Гимза (G-ок-

раска). При кислотно-щелочном окрашивании по Гимза окрашивается гетерохроматин центромеры (C-окраска) (см. также рис. 1.7).

Это зависит от их размера, положения центромеры и характерного рисунка из темных и светлых полос, образующегося после окрашивания фиксированных метафазных хромосом определенными красителями (рис. 7.23). Кроме того, сами полосы захватывают слишком большие отрезки ДНК. Например,  $3 \cdot 10^9$  пар нулеотидов гаплоидного генома человека содержатся менее чем в 1000 различных полосах (кариотип человека представлен на рис. 1.9), а у растений это число даже меньше. В результате с помощью чувствительной эмульсии можно установить лишь, в какой области хромосомы находится интересующий нас сегмент. Более того, размер зерен серебра и его разброс уменьшают разрешающую способность примерно до  $10^7$  пар оснований.

Вторая проблема — чувствительность зондов. Масса одной копии фрагмента дуплексной ДНК длиной 1 т.п.н. составляет примерно  $10^{-12}$  мкг. Радиоактивно меченные зонды, полученные с помощью ник-трансляции (разд. 4.6.б), редко имеют удельную радиоактивность, значительно превышающую  $10^8$  расп./мин на 1 мкг. Гибридизация такого зонда с одним сопоставимым сегментом в хромосоме дает только  $10^{-4}$  расп./мин на 1 мкг, или примерно один распад в неделю. Чтобы достичь необходимой чувствительности, следует одновременно использовать несколько специальных приемов. Например, чтобы получить зонды с удельной радиоактивностью примерно  $10^9$  расп./мин на 1 мкг, можно провести ник-трансляцию с использованием в качестве субстрата  $^{125}\text{I}$ -5-иод-dCTP. Гибридизацию следует проводить в присутствии декстрансульфата, что увеличивает скорость процесса примерно в 100 раз, в результате чего усиливается сигнал. Число зерен также может возрасти,

если одноцепочечные последовательности векторной ДНК, которые сцеплены с гибридизуемой эукариотической последовательностью в клонированном зонде, будут в свою очередь гибридизоваться с избытком молекул зонда. Для увеличения числа зерен серебра можно увеличить время экспозиции до нескольких недель (рис. 7.22, Б и В). Таким способом удается определить положение единичных копий генов в специфических, но весьма обширных областях хромосом млекопитающих.

*Картирование хромосом, основанное на получении гибридных соматических клеток.* Для локализации клонированного сегмента в определенной хромосоме используют межвидовые соматические гибриды. Обычно такие **гибридные соматические клетки** содержат полный набор хромосом одного вида (реципиента) и только одну или небольшое число хромосом второго вида (донора). Чтобы судить о наличии данного гена в определенной донорной хромосоме, анализируют группу разных линий гибридных клеток и устанавливают корреляцию между присутствием тестируемого гена и наличием специфической хромосомы.

Гибридные соматические клетки образуются при 1) слиянии клеток двух разных видов; 2) слиянии клеток одного вида с мини-клетками другого, содержащими одну или несколько донорных хромосом; 3) трансфекции реципиентных клеток препаратом очищенных хромосом донора (рис. 7.24). Чтобы элиминировать неслившиеся донорные и реципиентные клетки, применяют соответствующие селективные условия. Например, донорные клетки могут проявлять чувствительность к определенным лекарственным веществам, а реципиентные клетки могут быть мутантами, растущими лишь в особых

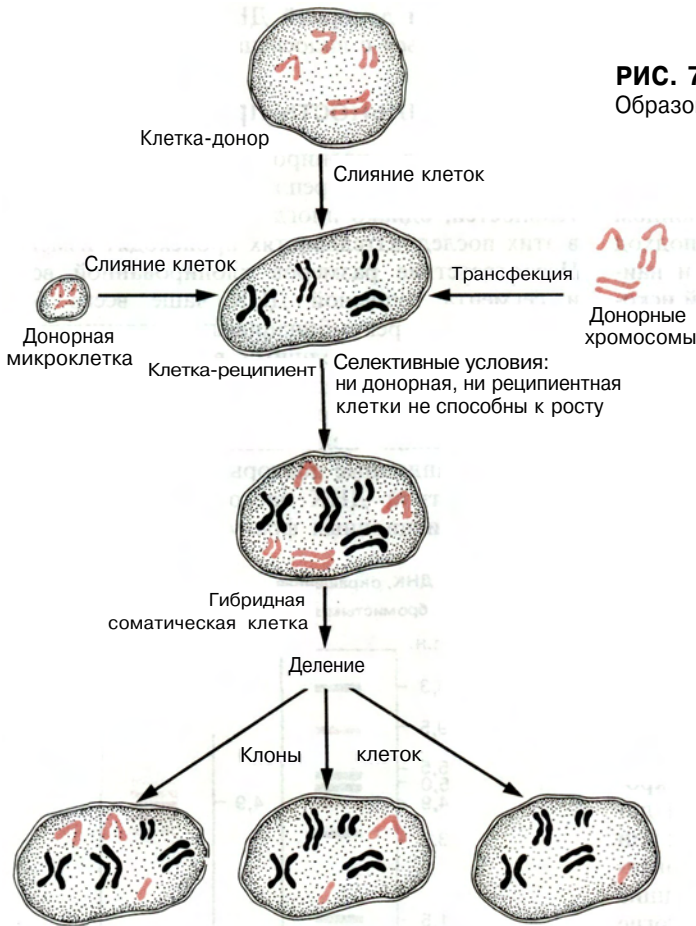


РИС. 7.24.

Образование гибридов соматических клеток.

условиях (например, клетки ТК<sup>-</sup> не способны расти в среде НАТ; см. рис. 5.31). В присутствии таких препаратов и в среде НАТ растут только гибридные клетки, несущие в донорной хромосоме функциональный ген тимидинказы.

Если отобранные гибридные клетки выращивать в неселективных условиях, то донорные хромосомы в конце концов утратятся в силу более или менее случайных причин. Но клоны гибридных клеток можно получить в любой момент и идентифицировать резидентные донорные хромосомы по характеру исчерченности, выявлению ранее картированных последовательностей ДНК или ферментативных маркеров либо (последнее более предпочтительно) с использованием всех трех методов. Таким способом можно получить банки различных линий гибридных клеток, каждая из которых содержит определенный набор донорных хромосом.

Альтернативный подход состоит в сохранении селективных условий. Поддерживая условия, благоприятные для данного гена, в течение многих клеточных поколений (например, используя известные

мутантные клетки реципиента и соответствующую среду), можно получить линии клеток только с одной стабильной донорной хромосомой. Другой способ получения гибридных клеток с одной донорной хромосомой состоит в трансфекции препаратом, обогащенным этой хромосомой.

При конструировании гибридов возникает вопрос, какой из видов будет служить донором, а какой — реципиентом. Ответ на него обычно можно найти лишь экспериментальным путем. Как правило, гибридные клетки «человек-мышь» утрачивают хромосомы человека.

После идентификации чужеродных хромосом в соматических клетках гибридов используют три подхода, для того чтобы установить, присутствует ли тестируемый ген (или сегмент ДНК) в определенной хромосоме. Первый основан на выявлении корреляции между наличием этой хромосомы и экспрессией тестируемого гена с образованием определенного продукта. Для этого измеряют ферментативную активность или выявляют взаимо-

действие со специфическими антителами; при этом необходимо, чтобы используемые методы позволяли различать продукты генов донорных и реципиентных клеток. Например, два белка—продукта генов могут иметь разную электрофоретическую подвижность, как, например, в случае галактокиназ грызунов и приматов. При втором подходе используют гибридизацию между клонированным зондом и набором хромосом гибрида *in situ*; этот подход имеет уже описанные ограничения. Третий и наиболее общий подход связан с идентификацией искомого гена (или сегмента ДНК) путем отжига соответствующего ДНК- или РНК-зонда с ДНК, выделенной из гибридных клеток и перенесенной на нитроцеллюлозный фильтр. Отжиг с эндонуклеазными фрагментами часто позволяет выявить донорный ген во фрагментах, имеющих размеры, типичные для донорного генома, даже в тех случаях, если гены донора и реципиента являются перекрестно гибридизующимися. Поскольку трудно получить гибриды с одной донорной хромосомой, обычно исследуют целый набор гибридных клеток, которые содержат разные донорные хромосомы. Если этот набор достаточно обширен, то присутствие последовательности зонда обычно можно связать с присутствием определенной хромосомы.

Используют также гибридные соматические клетки, содержащие лишь часть определенной хромосомы донора. Например, некоторые линии донорных клеток содержат хромосому с делецией или хромосому, несущую небольшой транслоцированный участок из другой хромосомы, и за прошедшие годы были выделены и охарактеризованы многие линии гибридных клеток, содержащие делетированные или составные донорные хромосомы. Например, с помощью таких гибридных клеток гены тяжелой цепи иммуноглобулина человека были локализованы в положении 32 длинного плеча хромосомы 14 (14q32). Были проанализированы две линии гибридных клеток мыши и человека, каждая из которых содержала одну из пары реципрокных транслокаций между хромосомой 14 и X-хромосомой. Зонд, представляющий собой клонированный ген иммуноглобулина, гибридизовался только с ДНК из гибрида, содержащего часть хромосомы 14, соответствующую q32.

В результате анализа клонов из библиотек, содержащих ДНК одной хромосомы, были получены детальные карты участков индивидуальных хромосом (разд. 6.5.а). Наиболее прямой путь получения таких библиотек состоит в использовании отдельных очищенных хромосом. Другой подход основан на использовании ДНК из гибридных соматических клеток, содержащих только одну донорную хромосому. Из такой библиотеки могут быть отобраны рекомбинанты, содержащие после-

довательности донорной ДНК, по их способности гибридизоваться с видоспецифичными зондами.

## В. Нестабильность при клонировании

Как правило, клонированные вставки представляют собой точные реплики геномных последовательностей, однако иногда во время клонирования в этих последовательностях происходят изменения. Несоответствие размеров клонированной вставки и сегмента геномной ДНК чаще всего бывает обусловлено рекомбинациями, делециями или вставками, произошедшими во время клонирования. Нередко наблюдаются вариации в размерах вставки или в ее рестрикционной карте при повторном выделении. Если клонированный фрагмент содержит тандемные повторы, то при расщеплении рекомбинантной ДНК часто получается сложная смесь рестрикционных фрагментов (рис. 7.25). На-

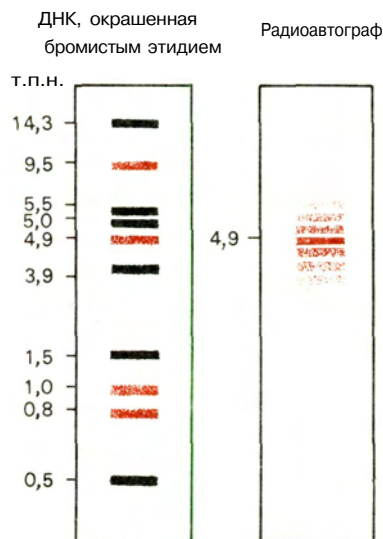


РИС. 7.25.

Нестабильность рекомбинанта, полученного на основе  $\lambda$ -вектора, во время репликации. Сегмент, содержащий тандемный повтор, клонировали в  $\lambda$ -векторе и после размножения фага из одной очищенной бляшки выделяли ДНК. Далее ДНК гидролизовали с помощью рестриктирующей эндонуклеазы, подвергли электрофорезу и гель окрашивали бромистым этидием (слева). Затем осуществляли перенос ДНК по Саузерну и отжигали ее с клонированным зондом, содержащим такую же повторяющуюся последовательность. Вставка длиной 4,9 т.п.н. содержала большую часть повторяющегося сегмента. Однако с зондом ассоциировали и другие фрагменты ДНК—как более крупные, так и более мелкие (справа). Такие фрагменты содержались в субмолярных количествах и не обнаруживались при окрашивании геля бромистым этидием. Встроенные фрагменты изображены в виде цветных полосок,  $\lambda$ -фрагменты—в виде черных.

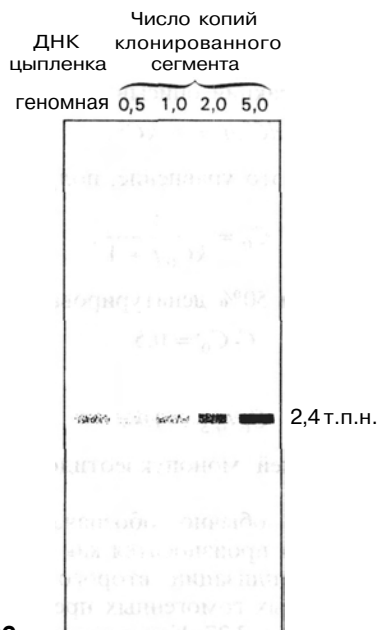
пример, некоторые фрагменты могут быть представлены в количестве меньшем, чем моль-эквивалент (т.е. в субмолярном количестве). Это тоже свидетельствует о происшедших во время репликации рекомбинантной молекулы перестройках, что приводит к образованию смешанной популяции рекомбинантов. Перестройка состоит в гомологичной рекомбинации, сопровождающейся делециями и амплификациями клонированных tandemно повторяющихся единиц. Подобная нестабильность клонированных повторов в *E. coli*-системах хозяин-вектор уменьшается, но не устраняется полностью в клетках, мутантных по функциям рекомбинации (например, в клетках *recA*<sup>-</sup>, *recBC*<sup>-</sup>).

### 7.5. ОПРЕДЕЛЕНИЕ ЧИСЛА КОПИЙ ДАННОЙ ПОСЛЕДОВАТЕЛЬНОСТИ В ГЕНОМЕ

Ген овальбумина присутствует в гаплоидном геноме курицы в единственном экземпляре, однако многие гены и некоторые другие сегменты ДНК встречаются в геноме множество раз (гл. 9). Типичными примерами геномов с многократно повторяющимися последовательностями ДНК являются эукариотические геномы. Иногда несколько копий в таком геноме бывают абсолютно идентичны, и тогда при отжиге смеси рестрикционных фрагментов геномной ДНК с меченым гомологичным зондом на электрофореграмме наблюдается одна полоска, четко видимая даже в том случае, когда количество расщепляемой ДНК не превышает 1 мкг. Обычно для обнаружения сегмента, представленного в геноме млекопитающих лишь однократно, требуется не менее 10 мкг ДНК, поэтому появление полосы при количестве ДНК 1 мкг и менее свидетельствует о том, что данный сегмент представлен в геноме многократно. Если разные копии последовательности различаются по одной или нескольким парам оснований, то у них могут быть и неодинаковые сайты для рестриктирующих эндонуклеаз. В результате на радиоавтограмме будет присутствовать множество полос или она будет представлять собой одно размытое темное пятно; это зависит от числа копий повторяющейся последовательности.

#### а. Оценка числа копий с помощью гибридизации ДНК–ДНК

Число копий можно оценить, проведя денситометрический анализ радиоавтограммы. При этом для повышения точности в отдельные лунки на этой же пластине геля вносят известное количество немеченой клонированной ДНК как образец геномной ДНК (рис. 7.26). Плотность радиоавтографа линей-



**РИС. 7.26.**

Определение числа копий. Одна полоска содержит *Eco* RI-гидролизат 30 мкг ДНК цыпленка, другие - клонированный сегмент гена овальбумина длиной 2,35 т.п.н. в количествах, эквивалентных 0,5; 1; 2 и 5 копиям (на гаплоидный геном) овальбуминового гена в общем количестве, представленном в 30 мкг ДНК цыпленка. Электрофорез проводили в агарозном геле. На рисунке дано схематическое изображение радиоавтографа, полученного после отжига ДНК, перенесенной на фильтр, с <sup>32</sup>P-меченым фрагментом размером 2,35 т.п.н. Плотность полоски с ДНК цыпленка соответствует одной гаплоидной копии.

но зависит от количества ДНК в геле, поскольку зонд присутствует в избытке. Сравнивая между собой эти показатели для геномной ДНК и ДНК стандарта, мы можем оценить число копий.

#### б. Оценка числа копий по кинетике реассоциации ДНК

Более точную информацию о числе копий можно получить, исследуя кинетику реассоциации ДНК. Кинетические методы стали применяться в этой области по крайней мере на десять лет раньше методов, описанных в разд. 7.5.а, и именно с их помощью были получены первые данные о том, что эукариотические геномы содержат много повторов. Для того чтобы понять суть кинетического метода, необходимо вывести несколько уравнений.

Скорость реассоциации отдельных комплементарных цепей в растворе зависит от концентрации ДНК и подчиняется кинетике второго порядка. Если

$C_0$  и  $C$ —это суммарные концентрации денатурированной ДНК в момент времени  $t_0$  и в момент времени  $t$  после начала гибридизации соответственно, то скорость реакции описывается уравнением

$$dC/dt = -kC^2. \quad (1)$$

Проинтегрировав это уравнение, получим

$$C/C_0 = \frac{1}{kC_0 t + 1}. \quad (2)$$

При гибридизации 50% денатурированной ДНК

$$C/C_0 = 0,5 \quad (3)$$

и

$$C_0 t_{0,5} = 1/k, \quad (4)$$

где  $C$ —число молей мононуклеотидов на 1 л,  $t$ —время в секундах.

Величина  $C_0 t$  обычно обозначается словом, которое пишется и произносится как «Cot» («кот»).

Кинетика гибридизации второго порядка нескольких очищенных гомогенных препаратов ДНК представлена на рис. 7.27. Как и во многих экспериментах по гибридизации, ДНК перед денатурацией и реассоциацией была разрезана на фрагменты длиной около 400 пар оснований. Обратите внимание на то, что шкала Cot логарифмическая; это облегчает сравнительный анализ данных. Соответствие кинетических кривых простой реакции второго порядка следует из того, что ренатурация на 90% укладывается в интервал значений Cot, охватывающий два порядка величины.

Конкретные значения  $Cot_{0,5}$  зависят от общего размера генома, поскольку  $C$ —это суммарная кон-

центрация ДНК, хотя, чем больше размер генома, тем меньше истинная концентрация каждого гибридизующегося фрагмента. Другими словами, в более сложном геноме каждый гибридизующийся фрагмент составляет меньшую часть суммарной ДНК и поэтому гибридизуется при более высоких значениях Cot. Таким образом, кинетические кривые позволяют оценить размер генома, если у нас имеется препарат геномной ДНК известного размера, который может служить стандартом, и если кривые соответствуют кинетике второго порядка. Если  $N_a$  и  $N_b$ —размеры геномов  $a$  и  $b$ , то

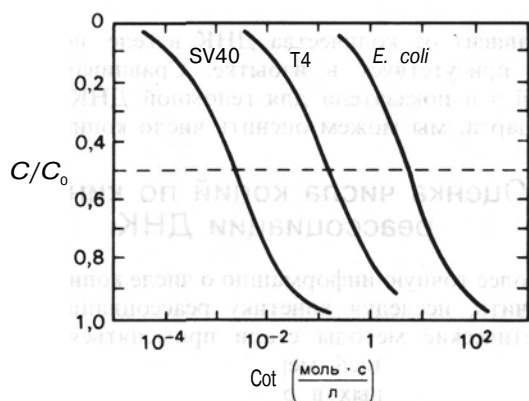
$$\frac{[Cot_{0,5}]_a}{[Cot_{0,5}]_b} = \frac{N_a}{N_b}. \quad (5)$$

Характер кинетических кривых второго порядка для ДНК SV40, T4 и *E. coli* позволяет сделать вывод о том, что каждый геномный сегмент представлен в геноме только один раз. Аналогичные кривые для эукариотической геномной ДНК имеют более сложную форму. В качестве примера на рис. 7.28 приведены данные для ДНК *Drosophila*. Гибридизация протекает в широком диапазоне значений Cot, поскольку многие сегменты ДНК повторяются, причем одни из них многократно, другие менее часто, а третьи—лишь несколько раз. Те сегменты, которые встречаются в геноме только один раз, имеют низкую концентрацию и гибридизуются при самых высоких значениях Cot. Большинство высокоповторяющихся последовательностей, концентрация которых наиболее высока, гибридизуются очень быстро и характеризуются самыми низкими значениями Cot. Предположив, что гибридизация сегментов ДНК, представленных в единственном числе, следует кинетике второго порядка, можно из сложных Cot-кривых с помощью метода наименьших квадратов выделить отдельную кривую для однокопийной последовательности и оценить ее Cot.

Для таких последовательностей *Drosophila* значение  $Cot_{0,5}$  равно 28,6 моль · с/л при экспериментальных условиях, которые использовались в случае, представленном на рис. 7.28. Это значение принято за стандартное; сравнивая с ним значение  $Cot_{0,5}$  для какого-то клонированного фрагмента ДНК *Drosophila*, можно рассчитать число копий этого фрагмента по формуле

$$\frac{28,6}{Cot_{0,5} \text{ (клон)}} = \text{Число копий}. \quad (6)$$

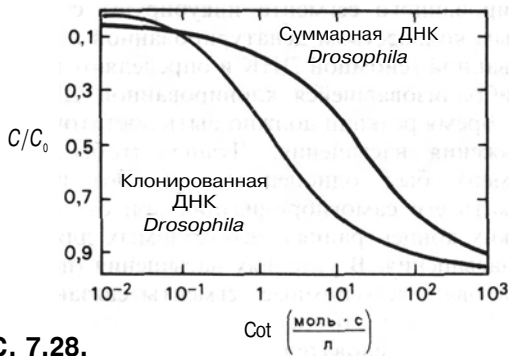
Вторая кривая на рис. 7.28 относится к гибридизации клонированного фрагмента в присутствии суммарной геномной ДНК *Drosophila*. Клонированный фрагмент помечен радиоизотопом, благодаря чему кинетику его гибридизации можно регистрировать отдельно. Заметим, что концентрация фрагмен-



Размер генома (п.н.)	SV40	T <sub>4</sub>	<i>E. coli</i>
	$5,2 \times 10^5$	$1,7 \times 10^6$	$4 \times 10^9$

**РИС. 7.27.**

Кинетика реассоциации денатурированной ДНК SV40, фага T4 и *E. coli* (смысл всех обозначений описан в разд. 7.5.6). [R.J. Britten, D.E. Kohne, Science, **161** (1968), p. 529, с изменениями.]

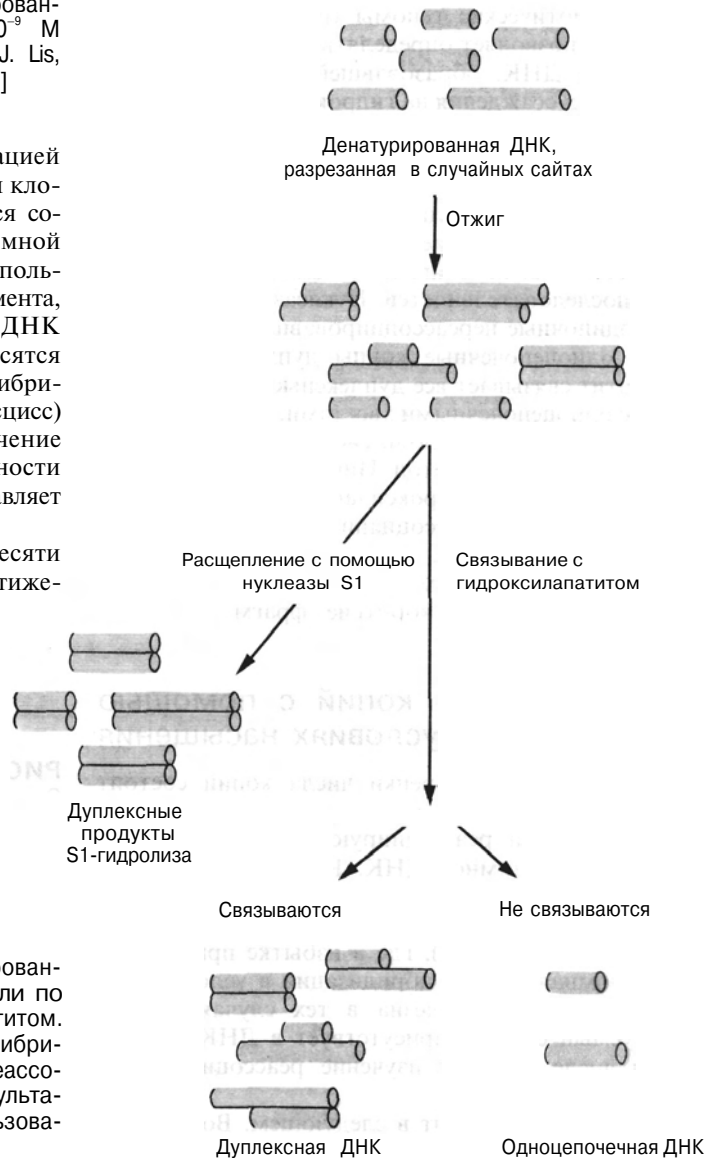


**РИС. 7.28.**

Кинетика реассоциации суммарной геномной ДНК *D. melanogaster* и клонированного фрагмента ДНК *Drosophila* при избытке ДНК *D. melanogaster*. Клонированный фрагмент представлен в количестве  $4,5 \cdot 10^{-9}$  М (см. разд. 7.5.б, где описан этот эксперимент). [J. Lis, L. Prestidge, D.S. Hogness, Cell, **14** (1978), p. 901.]

та достаточно мала, так что его самогибридизацией можно пренебречь. Эффективная концентрация клонированной последовательности обеспечивается соответствующими последовательностями геномной ДНК. Поэтому значение  $C_0t$  по оси абсцисс, используемое для получения значения  $Cot$  для фрагмента, рассчитывается из концентрации геномной ДНК *Drosophila*, а  $C$  и  $C_0$  по оси ординат относятся к фракции клонированной ДНК, которая гибридизовалась. Вот почему значения  $C_0t$  (ось абсцисс) действительно относятся к геномной ДНК. Значение  $Cot_{0,5}$  для клонированной последовательности (рис. 7.28) равно 1,06, а число копий составляет  $28,6/1,06 = 27$ .

Для определения числа копий от одной до десяти этот метод малопригоден, поскольку для достиже-



**РИС. 7.29.**

Исследования кинетики реассоциации денатурированной ДНК с помощью S1-нуклеазного метода или по связыванию дуплексных форм с гидроксипатитом. Через разные промежутки времени после начала гибридизации отбирают пробы и определяют долю реассоциировавшей ДНК. На рисунке сравниваются результаты, полученные этими двумя методами с использованием одного и того же препарата.

рая должна быть ниже температуры плавления дуплексной ДНК; см. рад. 1.1.e), ионной силы, вязкости раствора и размера гибридизующихся фрагментов. Все эти параметры должны быть строго фиксированными.

Процент гибридизовавшейся ДНК измеряют разными способами. Во-первых, можно использовать тот факт, что при определенных условиях с гидроксилпатитом связывается дуплексная, но не одноцепочечная ДНК. Во-вторых, можно определять в разные моменты времени процент ДНК, которая становится устойчивой к нуклеазе, специфичной в отношении одноцепочечной ДНК (S1). Результаты, получаемые с помощью этих двух методов, различаются, если анализируются сложные эукариотические геномы (рис. 7.29). Нуклеазный метод позволяет определять только количество дуплексной ДНК, образовавшейся при реассоциации, а метод осаждения на гидроксилпатите дает не только количество дуплексной ДНК, но и ДНК с неспаренными одноцепочечными хвостами. Такие дуплексы образуются, в частности, в результате случайных разрывов одиночных цепей еще до реассоциации с образованием фрагментов подходящих размеров, а также благодаря наличию повторяющихся последовательностей. Нуклеаза S1 расщепляет все одиночные нерассоциировавшие цепи, а также все одноцепочечные концы дуплексов. Гидроксилпатит связывает все дуплексные цепи, включая и цепи с одноцепочечными хвостами, и поэтому дает более высокую оценку доли гибридизовавшихся цепей, чем нуклеазный метод. Иными словами, в опытах с применением гидроксилпатита мы получаем большую степень реассоциации ДНК и большую скорость реассоциации, чем в опытах с S1-нуклеазой. Эти различия уменьшаются, если используются относительно короткие фрагменты ДНК (~ 400 п.н.).

## в. Оценка числа копий с помощью гибридизации в условиях насыщения

Еще один метод оценки числа копий состоит в определении общего количества ДНК клонированного сегмента, реассоциирующегося с известным количеством геномной ДНК. Клонированный сегмент должен присутствовать в избытке в отличие от ситуации, описанной в разд. 7.5.6 (метод кинетической реассоциации), где в избытке присутствовала геномная ДНК. Гибридизация в условиях насыщения особенно полезна в тех случаях, когда исследуемый сегмент присутствует в ДНК в очень малом числе копий и изучение реассоциации затруднено.

Эксперимент состоит в следующем. Возрастающее количество препарата радиоактивно меченного

клонированного сегмента инкубируют с фиксированным количеством денатурированной и фрагментированной геномной ДНК и определяют количество гибридизовавшейся клонированной ДНК (рис. 7.30). Время реакции должно быть достаточным для достижения насыщения. Лучше, чтобы меченый фрагмент был одноцепочечным; это позволяет избежать его самогибридизации при относительно высоких концентрациях, необходимых для получения насыщения. В условиях насыщения (плато) все клонированные геномные сегменты связаны с зондом. Высота плато дает суммарное число гомологичных последовательностей в геноме, которое можно рассчитать исходя из удельной радиоактивности меченого фрагмента. Для повышения точности метода проводят калибровочные эксперименты, в которых к препаратам геномной ДНК добавляют известные количества немеченого клонированного фрагмента. В эксперименте, результаты которого представлены на рис. 7.30, количество гибридизовавшегося радиоактивного фрагмента при насыщении эквивалентно двум копиям гена на

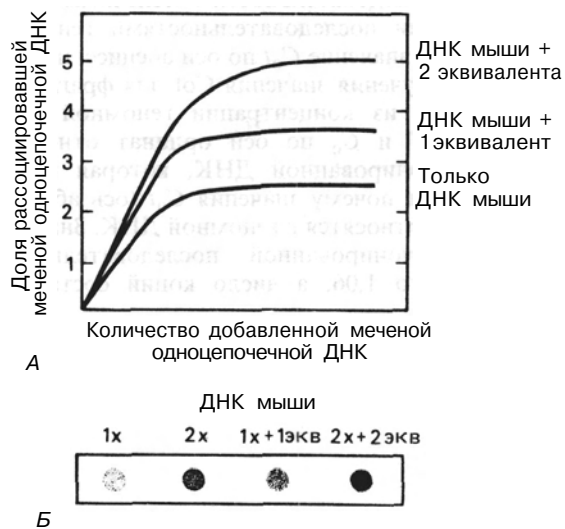


РИС. 7.30.

Оценка числа копий с помощью метода реассоциации в условиях насыщения. А. Отжиг разных количеств препарата  $^3\text{H}$ -меченой цепи клонированного сегмента ДНК с одинаковым количеством фрагментированной денатурированной геномной ДНК (в растворе). Две верхние кривые отражают ситуацию, когда в реакцию смесь добавляли клонированный нерадиоактивный сегмент в количестве, эквивалентном одной и двум копиям. Б. Фиксация препаратов денатурированной геномной ДНК в виде пятен на нитроцеллюлозном фильтре (рис. 7.39). После инкубации препаратов с  $^{32}\text{P}$ -меченой ДНК фильтр промывали и либо получали радиоавтограф (как в данном случае), либо вырезали пятна и подсчитывали число импульсов в каждом из них.

гаплоидный геном. Чтобы упростить эксперимент, препараты денатурированной геномной ДНК фиксируют в виде пятен на нитроцеллюлозном фильтре. В усовершенствованном варианте сначала определяют насыщающую концентрацию радиоактивного зонда, а затем инкубируют зонд с разным количеством геномной ДНК, фиксированной на нитроцеллюлозном фильтре.

## 7.6. ИЗМЕНЕНИЕ КЛОНИРОВАННЫХ СЕГМЕНТОВ: ПОЛУЧЕНИЕ МУТАНТОВ

### а. Общие положения

В основе классического генетического анализа лежит получение случайных мутаций, вызывающих наследуемые фенотипические изменения. Разработка новых молекулярно-генетических методов привела к созданию так называемой **обратной генетики**. В охарактеризованные клонированные гены вносят специфические мутации и затем устанавливают корреляцию между изменениями в определенных участках этих генов и изменением фенотипа или локализируют регуляторные элементы. Например, если в кодирующую часть гена внести мутации, то на нем будет синтезироваться модифицированный полипептид. Это позволяет изучать влияние аминокислотной последовательности белка на его конформацию или ферментативную активность. Аналогично нуклеотидная замена в предполагаемом регуляторном участке может привести к подавлению или стимуляции транскрипции, что подтвердит предположение о функциональной роли данного участка.

Существуют разные способы внесения мутаций в клонированные фрагменты или небольшие геномы, но мы рассмотрим лишь немногие из них. Во всех случаях успех зависит от двух условий. Во-первых, должна быть хорошо известна структура исходной ДНК. Как минимум, необходимо знать подробную карту сайтов рестрикции, но еще лучше, если известна вся последовательность изучаемого фрагмента. Во-вторых, поскольку все методы дают некую смесь продуктов, нужный элемент следует очистить с помощью клонирования, амплифицировать и охарактеризовать. При некоторых типах мутагенеза мутации образуются в случайных местах, при других – в определенных сайтах; о последних говорят как о сайт-специфических (или сайт-направленных) мутациях.

### б. Делеционные мутанты

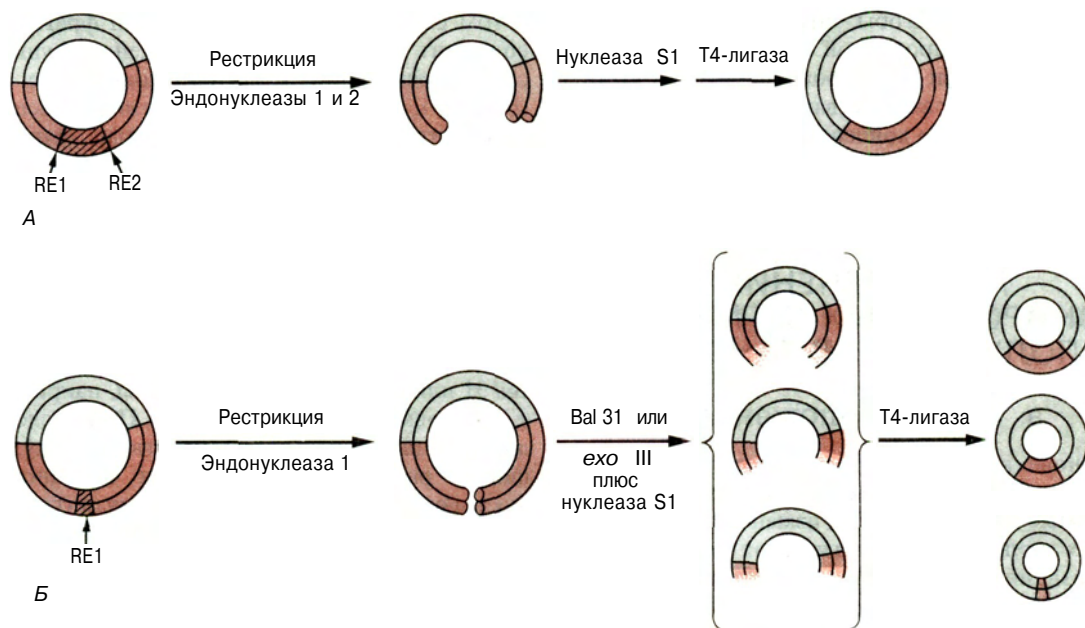
*Использование рестриктирующих эндонуклеаз.* Внести делецию в определенный участок можно

в том случае, если клонированный фрагмент ДНК содержит два близко расположенных уникальных сайта рестрикции в интересующей нас области (рис. 7.31, А). После эндонуклеазного расщепления фрагмента образующиеся «хвосты» отщепляют с помощью специфичной к одноцепочечным участкам нуклеазы и тупые концы вновь сшивают. Однако столь простой подход применим лишь в редких случаях. Чаще используют другой метод, при котором делецию создают в окрестности одного уникального сайта рестрикции (рис. 7.31, Б). После эндонуклеазного разрезания фрагмента соседние нуклеотиды удаляют с помощью эндонуклеазы *Bal 31* или какой-либо экзонуклеазы, например *exoIII E. coli*, совместно с нуклеазой *S1*. При последующем лигировании образуется семейство молекул с делециями, расположенными вокруг исходного эндонуклеазного сайта. Делеционные мутанты очищают с помощью молекулярного клонирования. В одном из вариантов этого метода перед образованием кольцевой молекулы к концам присоединяют рестриктазные линкеры. При этом в молекулу вводятся эндонуклеазные сайты, которые могут оказаться полезными для исследования ее свойств, секвенирования и последующего моделирования.

При помощи более сложных манипуляций может быть получен целый набор делеций разной длины, берущих начало от одного исходного сайта (рис. 7.32). Для этого рекомбинантную молекулу обрабатывают двумя способами. Из одной ее части (А) удаляют с помощью двух рестриктирующих эндонуклеаз (*RE1* и *RE2*) большой сегмент, включающий мишень для делеций. Вторую часть (В) линеаризуют с помощью одного из двух ферментов (*RE1*) и затем обрабатывают ферментом *Bal 31*. Далее в результате разрезания вторым ферментом (*RE2*) получают набор фрагментов уменьшающейся длины (*B'*, *B''*, *B'''* на рис. 7.32). Лигирование этих фрагментов с частью А и последующее клонирование дает набор молекул с делециями, начинающимися в сайте *RE1*.

*Использование неспецифических эндонуклеаз.* В сверхспиральные кольцевые молекулы могут быть внесены единичные разрывы с помощью неспецифических эндонуклеаз, например ДНКазы I. При этом образуется целый набор линейных молекул с разрывами в разных сайтах. После внесения разрыва и расширения deletированной области осуществляют лигирование с помощью методов, описанных ранее для устранения эндонуклеазных разрывов. При таком подходе используют некоторые интересные свойства ДНКазы I: в присутствии  $Mn^{2+}$  этот фермент расщепляет сразу обе цепи дуплексной ДНК, а в присутствии  $Mg^{2+}$  гидролизует за один раз только одну цепь, в результате чего в ДНК появляются одноцепочечные пробелы. Кроме того, фермент расщепляет сверхспиральную



**РИС. 7.31.**

Получение делеционных мутантов с помощью рестриктирующих эндонуклеаз. Серым цветом выделена ДНК плазмидного вектора, коричневым - вставка. А. Использо-

вание двух соседних уникальных рестрикционных сайтов - RE1 и RE2. Б. Использование одного уникального рестрикционного сайта.

ДНК быстрее, чем линейную дуплексную, вследствие локальной неупорядоченности ДНК в сверхспиральном состоянии. Поэтому после непродолжительной обработки сверхспиральной ДНК ДНКазой I в присутствии  $Mg^{2+}$  образуются в основном полноразмерные линейные дуплексные молекулы. Обработка их с помощью *Bal 31* или экзонуклеазы, а затем S1-нуклеазы расширяет образовавшийся ранее пробел. После лигирования отдельные продукты можно получить с помощью молекулярного клонирования.

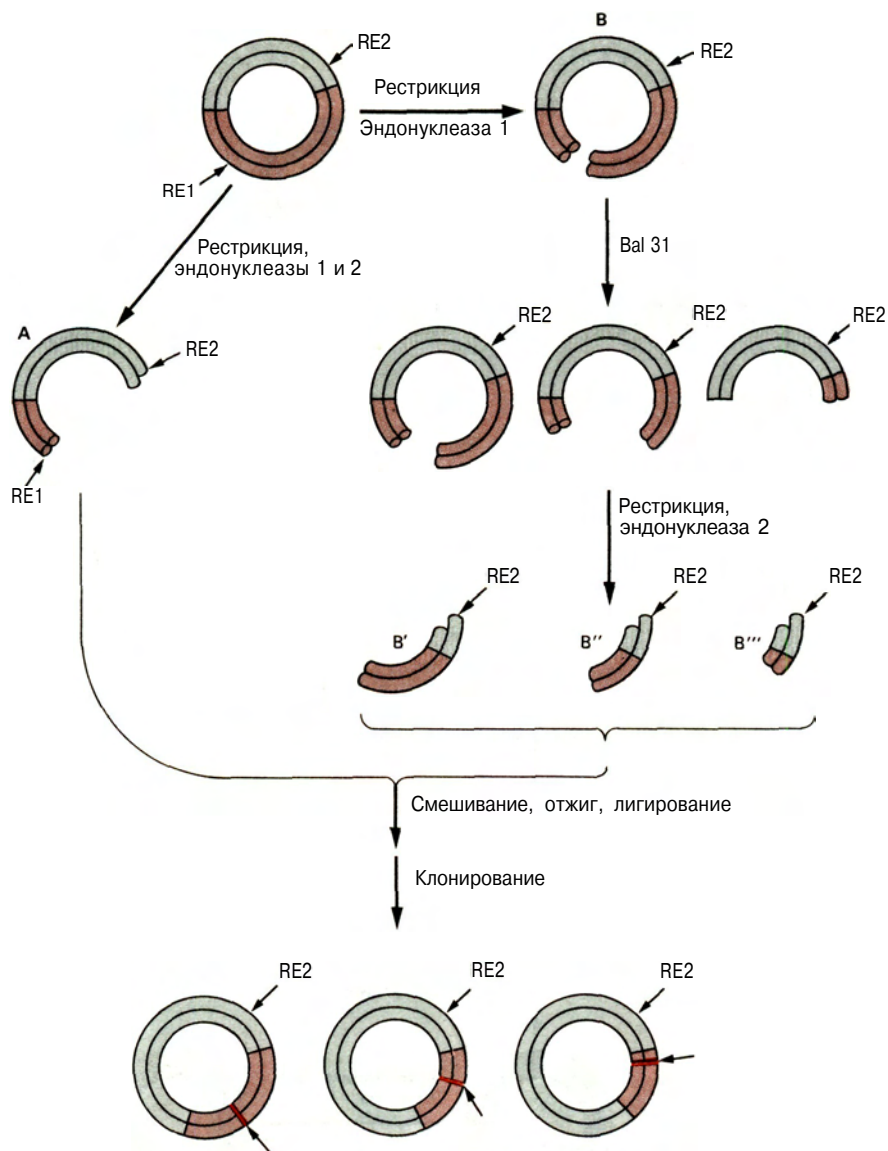
с помощью ДНКазы I в присутствии  $Mn^{2+}$ ), то появление новых сайтов рестрикции в наборе клонированных мутантов поможет построить физическую карту мутаций.

### в. Инсерционные мутанты

Принципы конструирования инсерционных мутантов сходны с описанными выше для делеционных мутантов. Клонированный сегмент ДНК расщепляют по одному из сайтов с помощью рестриктирующей эндонуклеазы или неспецифической эндонуклеазы. При необходимости заполняют пробелы на концах образовавшейся линейной молекулы или отщепляют одноцепочечные «хвосты» с помощью нуклеазы и осуществляют лигирование в присутствии сегмента, который хотят встроить в молекулу. В качестве вставки может использоваться синтетический фрагмент, содержащий множество сайтов для рестрикционных эндонуклеаз, — так называемый полилинкер. Если первое расщепление было неспецифичным (например, если оно было осуществлено

### г. Точечные мутации

*Химический мутагенез.* Для получения точечных мутаций в определенном участке молекулы чаще всего используют дуплексную кольцевую ДНК, содержащую короткий одноцепочечный участок. Один из способов создания таких сайт-специфических пробелов состоит в обработке сверхспиральной ДНК соответствующей рестриктирующей эндонуклеазой в присутствии бромистого этидия, который встраивается между плоскостями пар оснований и вносит нарушения в структуру дуплекса (рис. 7.33, А). При этих условиях многие рестриктирующие эндонуклеазы разрезают только одну из цепей в соответствующих сайтах. По-видимому, при встраивании бромистого этидия в обычную дуплексную молекулу ДНК разрезания вообще не происходит, а в сверхспиральной молекуле разрезается только одна цепь. Не все рестриктирующие эндонуклеазы ведут себя подобным образом, но все же число их достаточно велико. После разрезания молекулы с помощью экзонуклеазы в дуплексной молекуле создают небольшой одноцепочечный пробел в месте разреза. Альтернативный способ полу-

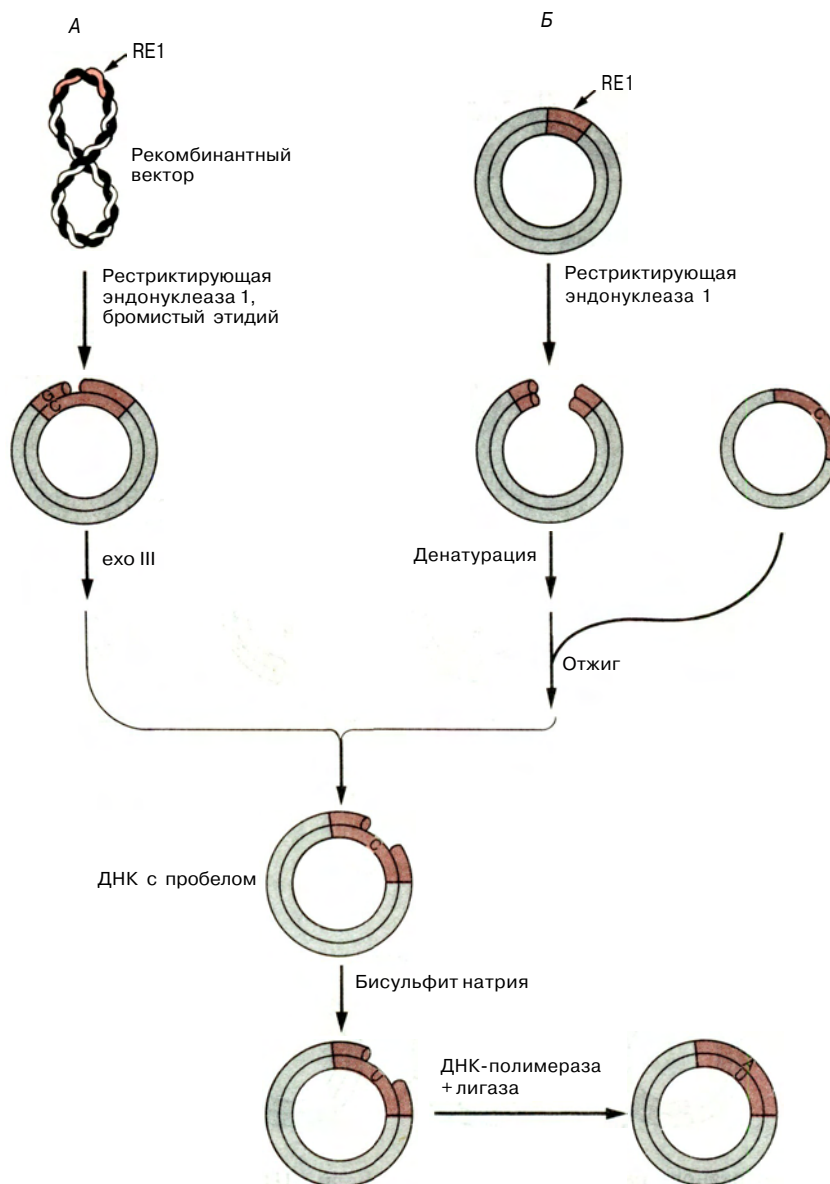
**РИС. 7.32.**

Получение набора делеций разного размера, берущих начало от одного исходного сайта, находящегося во вставке. Сайт RE1 после образования кольцевых молекул не восстанавливается. Его положение в этих структурах указано стрелками.

чения дуплексной молекулы с пробелом состоит в использовании векторной системы на основе фага M13 (рис. 7.33, Б). Для этого создают одноцепочечный M13-рекомбинант, содержащий нужный сегмент, а также другой, двухцепочечный, рекомбинант, содержащий такой же сегмент, но с делецией. Этот двухцепочечный рекомбинант денатурируют и реассоциируют с одноцепочечным, в результате чего образуется гетеродуплекс с пробелом.

Если ДНК, содержащую одноцепочечный пробел, обработать бисульфитом натрия, то в одноцепочечном участке произойдет дезаминирование остатков цитозина с образованием урацила, т.е. дезаминируются только определенные остатки цитозина. Если пробел невелик, а число дезаминированных остатков цитозина ограничено благодаря малому времени реакции и низкой концентрации бисульфита, то возникает очень небольшое число специфических мутаций. Далее пробел заполняют с помощью ДНК-полимеразы и осуществляют лигирование. Вместо исходных С•С-пар в молекуле ДНК теперь содержатся пары У•А, которые после репликации превращаются в Т•А-пары.

*Мутагенное копирование.* Специфические мута-

**РИС. 7.33.**

Индукция сайт-специфических точечных мутаций в одноцепочечных участках дуплексной ДНК с помощью бисульфита натрия. Одноцепочечный участок во вставке получали путем расщепления сверхспиральной дуплексной ДНК рестриктирующей эндонуклеазой в присут-

вии бромистого этидия и последующей обработки продуктом экзонуклеазы III (А) или с помощью образования гетеродуплекса между интактной вставкой и вставкой с делецией, клонированными в M13 (Б).

ствии другого типа можно получить, если при заполнении пробела вместо нормального дезоксирибонуклеозидтрифосфата использовать его мутагенный аналог. ДНК-полимераза I способна использовать в качестве субстратов различные аналоги обычных дезоксирибонуклеозидтрифосфатов. Некоторые из них являются мутагенами. Например, N<sup>6</sup>-гидрокси-десоксицитидин-5'-трифосфат (НО-dCTP) может

включаться в синтезируемую цепь в положение, соответствующее А или G в матричной цепи в зависимости от того, находится ли он в иминной или аминной форме соответственно (рис. 7.34). Если пробел в дуплексной ДНК заполняется с помощью НО-dCTP вместо dTTP, то напротив А будет находиться НО-dC (рис. 7.35). После трансфекции и репликации наряду с нормальными формами будут

**РИС. 7.34.**

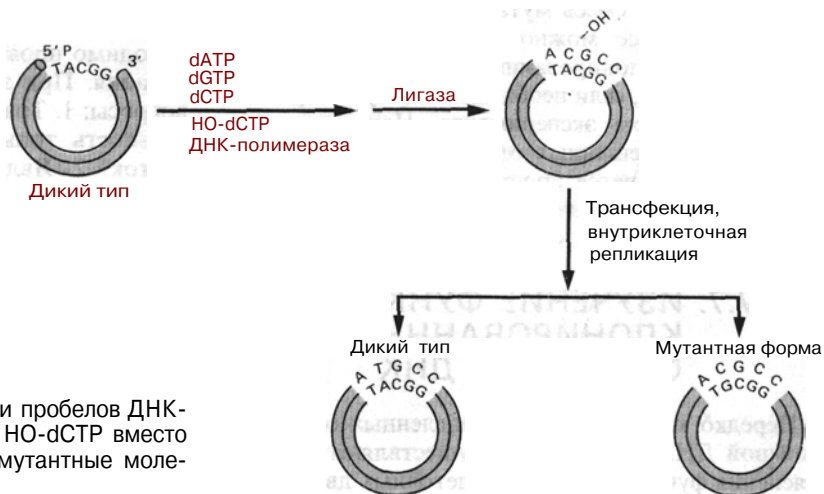
Таутомерные формы N<sup>6</sup>-гидроксицитидина.

обнаруживаться мутантные геномы с транзициями T • A → C • G. Проведя отбор, эти мутанты можно клонировать. Аналогично замещение dCTP на HO-dCTP приводит к транзициям C • G → T • A.

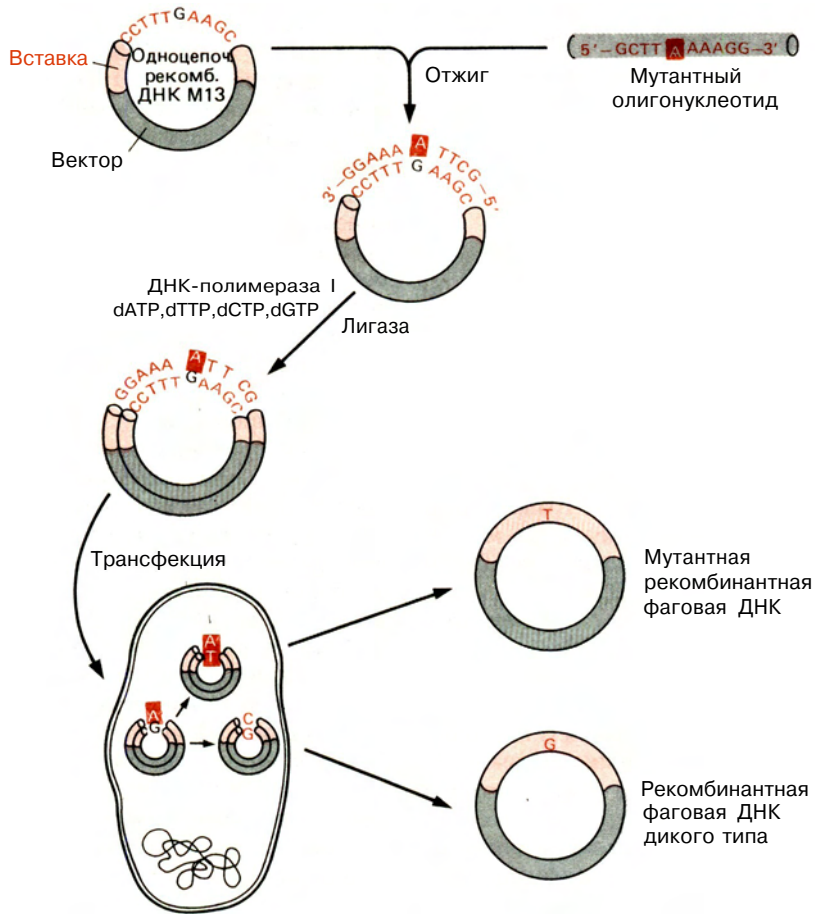
*Сайт-специфический мутагенез с применением синтетических олигодезоксирибонуклеотидов.* Олигодезоксирибонуклеотиды можно синтезировать в больших количествах, что позволяет разработать достаточно универсальные методы получения сайт-специфических точечных мутаций в клонированных сегментах ДНК. В основе одного из таких методов лежит образование гетеродуплекса между одноцепочечным синтетическим олигодезоксирибонуклеотидом, содержащим мутантную последовательность, и комплементарной одноцепочечной рекомбинантной векторной ДНК, несущей соответствующий сегмент дикого типа (рис. 7.36). Для этого ген, в котором мы хотим получить мутацию, клонируют, например, в фаге M13 и получают одноцепочечную кольцевую рекомбинантную вирусную ДНК. Затем с этой кольцевой молекулой отжигают

синтетический олигодезоксирибонуклеотид длиной от 8 до 20 нуклеотидов, содержащий мутантную последовательность. Этот олигодезоксирибонуклеотид выполняет роль праймера, а оставшийся одноцепочечный участок — роль матрицы при синтезе ДНК *in vitro* с помощью ДНК-полимеразы I. После копирования всей кольцевой молекулы начало и конец новой цепи соединяют лигазой. Образовавшаяся дуплексная молекула содержит неспаренные основания в мутантной последовательности. Интересно, что фаговое потомство, вышедшее из одной инфицированной клетки, сегрегирует как смешанная популяция фагов дикого типа и мутантных рекомбинантов, которые можно разделить при последующем клонировании.

При втором подходе, так называемом каскетном мутагенезе, участок клонированной ДНК дикого типа замещают синтетическим дуплексным олигодезоксирибонуклеотидом, содержащим мутантную последовательность (рис. 7.37). В приведенном примере клонированная вставка встроена в дуплексный вектор типа pBR322. В простейшем случае для вырезания участка, в который мы хотим внести мутацию, используют уникальные рестрикционные сайты, встречающиеся во вставке, но не в векторе. Если такие сайты отсутствуют, приходится прибегать к дополнительным ухищрениям. Дуплексный олигонуклеотид получают путем отжига двух синтетических комплементарных цепей, каждая из которых содержит соответствующую замену основания. Кроме того, эти цепи синтезируют таким образом, чтобы дуплекс, который они образуют, имел соответствующие липкие концы. В отличие от гетеродуплексного метода, при трансфекции и клонировании измененного рекомбинанта геномом дикого типа не образуется. Однако если используют смесь синтетических олигодезоксирибонуклеотидов, содержащих альтернативные основания в мутант-

**РИС. 7.35.**

Мутации, возникающие при заполнении пробелов ДНК-полимеразой в результате включения HO-dCTP вместо dTTP. После репликации появляются мутантные молекулы с транзициями T • A → C • G.

**РИС. 7.36.**

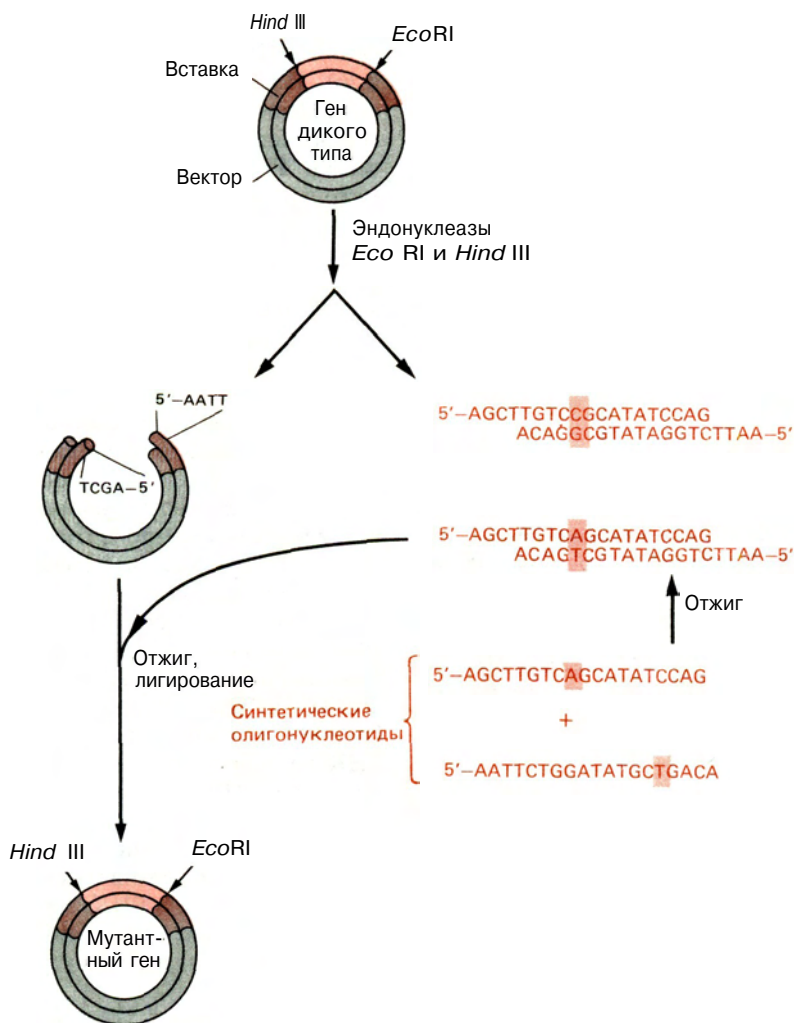
Сайт-специфический мутагенез с использованием синтетических олигонуклеотидов: гетеродуплексный метод. Синтетический олигонуклеотид служит праймером при синтезе мутантной вставки в рекомбинантном M13-векторе.

ных участках, то смесь мутантов сегрегирует после трансфекции и ее можно разделить с помощью последующего клонирования. Такой подход оказывается полезным, если необходимо получить разные мутации в одном эксперименте. Напомним, что синтез таких смешанных олигодезоксирибонуклеотидов осуществляется простым использованием на нужной стадии химического синтеза не одного мононуклеотида, а их смеси.

## 7.7. ИЗУЧЕНИЕ ФУНКЦИЙ КЛОНИРОВАННЫХ СЕГМЕНТОВ ДНК

Нередко клонирование определенных сегментов геномной ДНК, или кДНК, осуществляют с целью выяснения функций их внутриклеточных двойников.

Исходя из результатов структурного анализа клонированных последовательностей, указывающих на присутствие в них открытых рамок считывания, можно предположить, что они содержат гены. Часто удается выявить специфические промоторные последовательности и другие регуляторные элементы. Однако, чтобы подтвердить эти данные, необходимо провести прямые функциональные исследования. При этом нужно ответить на следующие вопросы: 1. Транскрибируется ли данная последовательность только в одном или нескольких типах клеток? 2. Являются ли транскриптами молекулы мРНК? 3. Влияют ли на транскрипцию изменения, происходящие в клетке? 4. Содержит ли клонированный сегмент промоторы, терминаторы или другие регуляторные сигналы, и если да, то как они работают? 5. Какова связь между структурой клонированного сегмента и структурой внутриклеточных транскриптов? 6. Могут ли транскрипты транслироваться в полипептид? Для ответа на все эти вопросы используют различные экспериментальные подходы в зависимости от того, какая именно система анализируется. Многие из них мы описали в последую-

**РИС. 7.37.**

«Кассетный» мутагенез. Участок, в который мы хотим внести мутацию, вырезают из клонированного рекомбинантного вектора по двум фланкирующим уникальным рестрикционным сайтам и вместо него включают синтетический дуплексный олигонуклеотид, содержа-

щий нуклеотидную замену и имеющий липкие концы, совпадающие с липкими концами в линейном векторе. Две цепи синтетического олигонуклеотидного дуплекса синтезируют отдельно и затем отжигают.

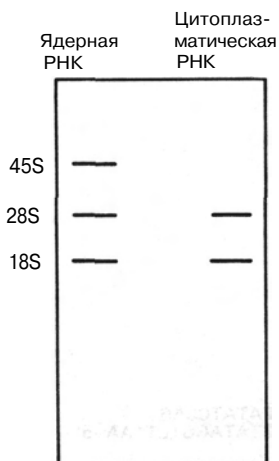
щих главах, но некоторые широко используемые приемы, лежащие в основе большинства методов, мы рассмотрим в этой главе.

### а. Характеристика внутриклеточных транскриптов, соответствующих клонированным сегментам ДНК

Говоря о любом клонированном сегменте генома, нам прежде всего необходимо ответить на вопросы, связанные с его транскрипцией *in vivo*. Очень важными являются также данные о структурном сходстве между клонированной кДНК и внут-

риклеточными родственными транскриптами. В основе соответствующих исследований лежат три метода: РНК-блоттинг, анализ с использованием специфичных к одноцепочечным ДНК нуклеаз и копирование РНК, выделенной из клеток, с помощью обратной транскриптазы. Все методы включают предварительное выделение и очистку РНК из целых клеток или специфических внутриклеточных органелл, таких, как ядро или цитоплазма. Результативность всех методов зависит от способности РНК образовывать гетеродуплексы с комплементарной клонированной ДНК.

**РНК-блоттинг.** Блоттинг РНК (нозерн-блоттинг) аналогичен блоттингу ДНК (разд. 6.1.6). Он

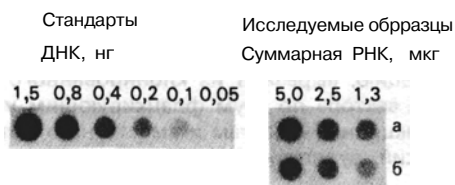
**РИС. 7.38.**

Схематическое представление РНК-блота. Показаны две дорожки в геле: одна отвечает ядерной, другая - цитоплазматической РНК. После электрофореза РНК переносят на нитроцеллюлозный фильтр, отжигают ее с  $^{32}\text{P}$ -рибосомной РНК и затем накладывают фильтр на рентгеновскую пленку. Три полосы на радиоавтографе ядерной РНК соответствуют 18S-рРНК, 28S-рРНК и их 45S-предшественнику (гл. 8). В препарате цитоплазматической РНК 45S-РНК отсутствует.

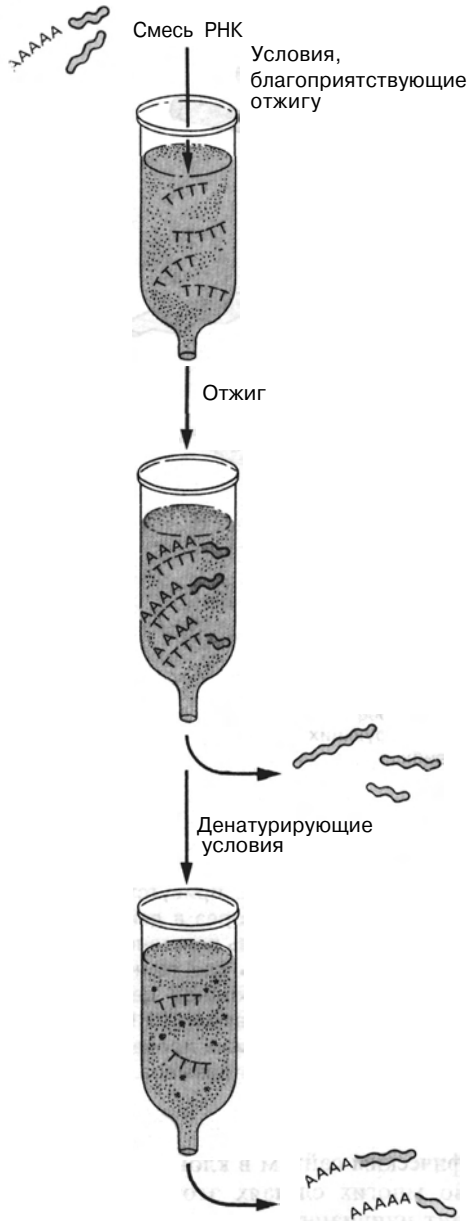
состоит в следующем. Выделенную РНК разделяют по размерам с помощью электрофореза в агарозном геле (рис. 7.38). Обычно электрофорез проводят в условиях, способствующих денатурации РНК, чтобы свести к минимуму влияние вторичной структуры молекулы на ее электрофоретическую подвижность. Щелочные условия для этой цели не подходят ввиду лабильности фосфодиэфирных связей в молекуле РНК в этих условиях. Поэтому используют такие агенты, как глиоксаль, формальдегид или мочевины. Затем РНК переносят на иммобилизованную подложку, стараясь сохранить распределение молекул РНК. Далее используют меченую ДНК (например, интересующий нас клон) в качестве зонда для выявления на фильтре соответствующих молекул РНК. Фильтр инкубируют с ДНК в условиях, благоприятствующих гибридизации. Промыв фильтр для удаления избыточной ДНК, с помощью радиоавтографии устанавливают положение зонда, а следовательно, и положение гомологичной РНК в том геле, в котором проводился электрофорез. Таким способом выявляют продукты транскрипции клонированного сегмента ДНК. Если на параллельной дорожке этого же геля одновременно провести разделение смеси РНК или молекул одноцепочечной ДНК известного размера, то можно оценить размер транскриптов. Кроме того, РНК-блоттинг позволяет оценить количество РНК, синтезированной в клетках, из которых она получена. Метод оценки

аналогичен используемому при определении числа копий ДНК на нитроцеллюлозных фильтрах (рис. 7.26). Плотность полосы на рентгеновской пленке пропорциональна количеству присутствующей гомологичной РНК. Для оценки количества специфической РНК в смеси можно также использовать методы, представленные на рис. 7.30 и в разд. 7.5.в (рис. 7.39). Как и ранее, желательно, чтобы меченый зонд был одноцепочечным, поскольку он не должен реассоциировать с комплементарной цепью ДНК вместо РНК.

С помощью всех этих довольно простых методов можно получить обширную информацию о функциональных свойствах клонированного сегмента ДНК. Сюда относятся не только данные о способности к транскрибированию, но и оценка числа транскриптов и ее зависимость от типа клеток или внеклеточной среды. Анализируя РНК из очищенных клеточных компонентов, можно установить, где локализируются разные транскрипты — в ядре, цитоплазме или полисомах. Часто очень важным является вопрос о полиаденилировании гомологичной РНК, поскольку полиаденилирование является характерным признаком большинства эукариотических мРНК. Разделить полиаденилированную [ $\text{poly}(\text{A}^+)$ ] и неполиаденилированную ( $\text{poly}(\text{A}^-)$ ] РНК не составляет труда, поскольку полиаденилированная РНК спаривается при соответствующих условиях с  $\text{poly}(\text{dT})$  или  $\text{poly}(\text{rU})$ . Сами полимеры обычно фиксируют на инертной твердой подложке, что упрощает отделение несвязанной  $\text{poly}(\text{A}^-)$ -РНК. Фракцию  $\text{poly}(\text{A}^-)$ -РНК элюируют при денатурирующих условиях (рис. 7.40). Затем РНК из каждой фракции подвергают электрофорезу, блоттингу и

**РИС. 7.39.**

Оценка количества специфической РНК в смеси молекул РНК с помощью «дот-блоттинга». Два препарата РНК (а и б) в трех вариантах каждый (количества 5,0; 2,5 и 1,3 мкг) фиксировали на нитроцеллюлозном фильтре и затем отжигали с избытком  $^{32}\text{P}$ -зонда, представляющего собой одноцепочечную ДНК (клонированную в векторе М13). Количество зонда, гибридизовавшегося в каждом из этих случаев, определяли с помощью радиоавтографии. Количество РНК в препарате, комплементарной зонду, оценивали путем сравнения плотности пятен, опытных и контрольных, полученных при нанесении известных количеств клонированного сегмента ДНК. Стандарты готовили из дуплексной денатурированной ДНК. (С любезного разрешения J. Skowronski.)

**РИС. 7.40.**

Фракционирование смеси  $\text{poly(A)}^+$ - и  $\text{poly(A)}^-$ -РНК. Смесь пропускают через колонку, содержащую  $\text{oligo(dT)}$ , связанную с целлюлозой, в условиях, благоприятствующих отжигу  $\text{oligo(T)}$  с  $\text{poly(A)}$ -«хвостами» РНК. РНК, у которой  $\text{poly(A)}$ -«хвосты» отсутствуют, проходит через колонку, не задерживаясь. Затем  $\text{poly(A)}^+$ -РНК элюируют с колонки при денатурирующих условиях.

тестированию на способность гибридизоваться с клонированной ДНК. Если РНК, идентифицированная с помощью клонированной ДНК, выделена из цитоплазмы и полиаденилирована, то скорее всего она представляет собой мРНК. Идентификация становится более надежной, если, кроме того, РНК связана с полисомами.

Еще одной характеристикой мРНК, гибридизовавшихся с клонированным сегментом ДНК, является их размер. Иногда РНК имеет больший размер, чем клонированный сегмент; это означает, что в клоне представлена лишь часть гена. В других случаях РНК оказывается короче клонированного сегмента, что свидетельствует о наличии в клонированной последовательности дополнительных последовательностей, не представленных в транскрипте. Это могут быть геномные последовательности, фланкирующие транскрибируемую область, или некодирующие последовательности, прерывающие кодирующую область и подвергающиеся сплайсингу во время созревания мРНК (разд. 7.1.г и гл. 8). Все эти взаимоотношения между клонированным сегментом ДНК и гомологичной клеточной РНК можно установить более точно, используя специфичные к одноцепочечным ДНК нуклеазы или осуществляя обратную транскрипцию.

*Исследование родства между клонированным сегментом ДНК и внутриклеточной РНК с помощью ДНК–РНК-гибридизации и специфичных к одноцепочечным ДНК нуклеаз.* Эксперимент включает три этапа (рис. 7.41). На первом клонированный фрагмент ДНК метят радиоактивным изотопом (например, с помощью ник-трансляции; см. рис. 4.18) и денатурируют. Можно также (и это более предпочтительно) использовать одноцепочечную ДНК. На втором этапе ДНК гибридизуют с выделенной клеточной РНК в растворе в условиях, благоприятных для образования ДНК–РНК-гибридов. Любые последовательности ДНК, присутствующие в клонированном фрагменте, но не представленные в РНК, остаются одноцепочечными. Их удаляют с помощью нуклеазы, специфичной к одноцепочечной ДНК. И наконец, гибридные молекулы денатурируют и подвергают гель-электрофорезу. Размер образующейся радиоактивной ДНК определяют с помощью радиавторграфии, сравнивая положение исследуемой ДНК с положением молекул ДНК известного размера.

Этот основной эксперимент можно проводить во многих разных вариантах, каждый из которых позволяет выявить свои особенности структуры РНК в зависимости от того, какая именно нуклеаза используется, является ли РНК суммарной или очищенной цитоплазматической  $\text{poly(A)}^+$ -мРНК. Один из таких вариантов представлен на рис. 7.41, где в ДНК имеется один протяженный участок, компле-



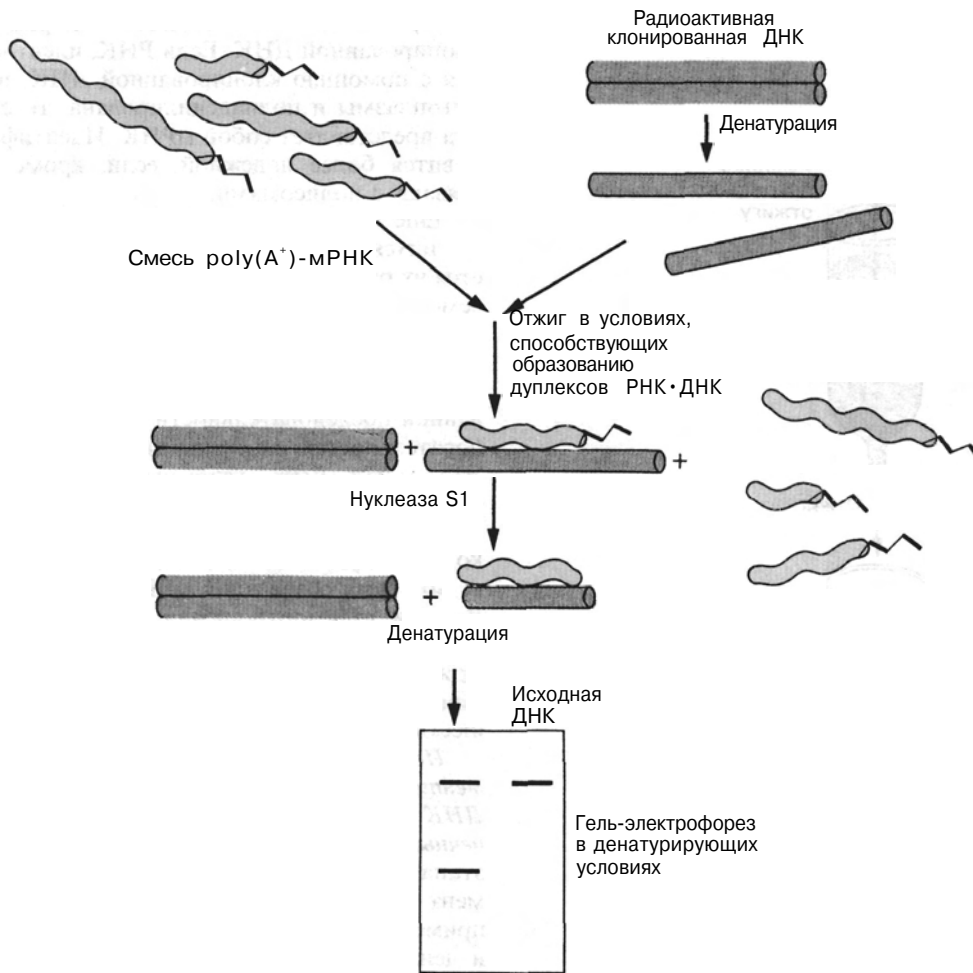


РИС. 7.41.

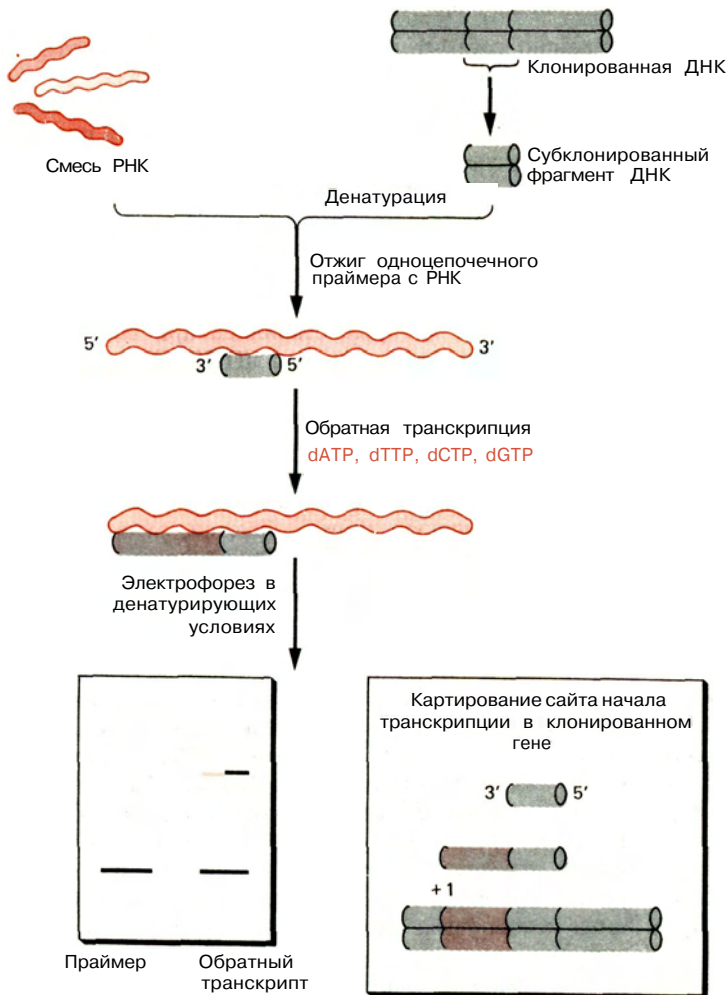
Установление структурного родства между клонированным сегментом ДНК и гомологичной клеточной РНК с помощью специфичной к одноцепочечной ДНК нуклеазы S1. В приведенном примере смесь цитоплазматических молекул poly(A<sup>+</sup>)-РНК отжигали с денатурированным препаратом радиоактивно меченой клонированной ДНК в условиях, благоприятствующих образованию дуплексов РНК·ДНК. С помощью нуклеазы S1 удаляли ту часть ДНК, которая не образует гетеродуп-

лекса с гомологичной РНК, присутствующей в смеси. Затем проводили электрофорез в денатурирующих условиях и накладывали на гель рентгеновскую пленку. На радиоавтографе появлялись две темные полосы. Одна из них соответствовала исходной цепи ДНК, другая — реассоциировавшей РНК. Размер РНК можно оценить, если на параллельной дорожке провести электрофорез маркерных ДНК известного размера.

ментарный РНК, и не спарены только концы молекулы. Метод позволяет также выявить любую область, где отсутствует комплементарность между РНК и ДНК, в том числе последовательности, аналогичные представленным на рис. 7.3. Другие варианты метода описаны в следующих главах. Во многих экспериментах такого рода используется специфичная к одноцепочечной ДНК эндонуклеаза S1 (разд. 4.1.6), поэтому метод называют **S1-картированием**.

S1-картирование часто используют для установления соответствия между началом молекулы РНК

и специфическим сайтом в клонированном сегменте ДНК. Во многих случаях этот сайт представляет собой сайт инициации транскрипции. Точность эксперимента значительно повышается, если используется небольшой ДНК-зонд, что позволяет измерить длину защищенного сегмента ДНК с точностью до одного нуклеотида. Для получения такого зонда обычно субклонировывают строго определенные участки длинного фрагмента ДНК. Размер защищенного фрагмента ДНК определяют с помощью электрофореза в таких же гелях, что и при определении нуклеотидной последовательности.



**Удлинение праймера:** исследование родства между внутриклеточной РНК и клонированным фрагментом ДНК с помощью обратной транскриптазы. Для проведения этого эксперимента необходим относительно короткий ДНК-зонд, который гибридизуется с участком РНК, желательнее находящимся на расстоянии не более 100 пар оснований от предполагаемого 5'-конца РНК (рис. 7.42). ДНК-зонд отжигают с РНК, и он играет роль праймера при обратной транскрипции, а матрицей служит РНК. Поскольку обратная транскриптаза прекращает копирование, когда достигает 5'-конца РНК, размер продукта позволяет установить, где начинается РНК. Обычно праймер бывает радиоактивно меченым, что позволяет легко определить положение продукта в геле. Поскольку ДНК-зонд специфичен только к конкретной РНК, исследуемый препарат может представлять собой смесь разных РНК.

**РИС. 7.42.**

Идентификация 5'-конца молекулы РНК с помощью удлинения праймера. Относительно короткий синтетический или субклонированный фрагмент ДНК, гомологичный участку интересующей нас РНК, находящемуся вблизи 5'-конца, отжигают с РНК. Затем с помощью обратной транскриптазы копируют РНК, используя фрагмент ДНК в качестве праймера. Копирование РНК-матрицы останавливается, когда обратная транскриптаза доходит до ее 5'-конца. Проводя гель-электрофорез в денатурирующих условиях, можно оценить размер удлиненной праймерной ДНК и найти сайт, где начинается РНК. Обычно используется радиоактивно меченная праймерная ДНК, что облегчает идентификацию продукта.

## б. Функциональное тестирование клонированной ДНК

Клонированные сегменты ДНК можно транскрибировать и транслировать в эукариотических системах либо *in vitro* с использованием клеточных экстрактов или очищенных ферментов, либо *in vivo* после введения клонированного сегмента в клетку.

**Системы *in vitro*.** Для приготовления бесклеточных экстрактов, содержащих РНК-полимеразы I, II и III, обычно используют разнообразные эукариотические клетки. В анализируемом сегменте должны присутствовать последовательности, ответственные за регуляцию транскрипции; необходимы также вспомогательные белки (или факторы), стимулирующие транскрипцию. При фракционировании таких экстрактов получают очищенные или частично очищенные полимеразы и факторы, с помощью которых можно детально изучить механизмы транс-

крипции (гл. 8). Для осуществления трансляции мРНК с образованием соответствующих полипептидов используют другие типы клеточных экстрактов (см. гл. 6).

*Системы in vivo.* Для изучения функций клонированного сегмента ДНК лучше всего использовать целые клетки, поскольку это позволяет следить как за процессом транскрипции, так и за процессом трансляции. Наиболее изученной биологической тест-системой являются ооциты лягушки (*Xenopus*). Ооциты—это крупные клетки объемом до 1 мкл, которые можно многократно получать от одной лягушки. ДНК инъецируют непосредственно в ядро (так называемый зародышевый пузырек), при этом обычно достаточно 5 нг препарата. ДНК в комплексе с гистонами ооцита упаковывается с образованием хроматина, далее осуществляются ее транскрипция и трансляция, и в течение нескольких часов можно идентифицировать продукты экспрессии генов—мРНК и полипептиды. Для проведения эксперимента часто оказывается достаточно того количества РНК и полипептидов, которое синтезируется одним ооцитом. Во многих отношениях экспрессия инъецированных клонированных генов в ооцитах протекает нормально. Исходные ядерные транскрипты генов РНК-полимераз II и III подвергаются процессингу, а зрелые мРНК и тРНК переходят в цитоплазму.

Рекомбинантные ДНК могут быть введены и в эукариотические клетки в культуре. Молекулы можно инъецировать непосредственно в отдельные клетки, однако проще вводить клонированные молекулы в большую популяцию клеток одновременно. При этом используется методика трансфекции, описанная в гл. 5 в связи с клонированием в эукариотических клетках. По крайней мере часть молекул, введенных таким способом, достигает ядер.

## 7.8. СИНТЕЗ ПОЛИПЕПТИДОВ, КОДИРУЕМЫХ КЛОНИРОВАННЫМИ СЕГМЕНТАМИ ЭУКАРИОТИЧЕСКОЙ ДНК

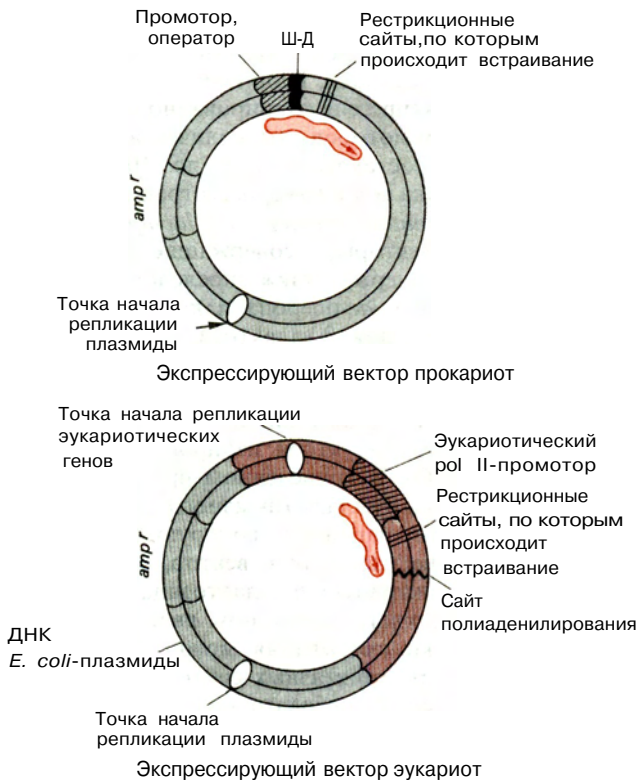
В предыдущих разделах мы уже не раз останавливались на синтезе *in vitro* и *in vivo* полипептидов, кодируемых вставками в рекомбинантных векторах. Скрининг по фенотипу зависит от экспрессии интересующего нас гена в клетках хозяина (гл. 5 и 6). Отбор нужных клонов такими методами, как НАРТ и гибридизационная селекция, осуществляется с помощью трансляции *in vitro* (разд. 6.4), а функциональный анализ клонированного сегмента после введения его в исходную клетку основывается на изучении как транскрипции, так и трансляции

(разд. 7.7.б). Цель многих экспериментов по клонированию состоит в синтезе нужного полипептида, а иногда—в получении его в количествах, достаточных для проведения фенотипического анализа. В ряде случаев цель клонирования состоит в получении антител. В этом разделе мы опишем некоторые системы хозяин-вектор, специально созданные для продукции нужного белка.

Синтез белков, предназначенных для проведения тех или иных исследований или для применения в медицине и промышленности,— это одна из самых первых задач, которые ставились при разработке технологии рекомбинантных ДНК. В настоящее время ряд поступающих в продажу ферментов, используемых для конструирования рекомбинантных молекул ДНК, в свою очередь синтезируются под контролем генов, клонированных в плазмидных векторах *E. coli*. Используемые при этом препараты имеют высокую степень очистки, относительно недороги и включают ДНК-полимеразу I, ДНК-лигазу *E. coli*, обратную транскриптазу. Другие белки и их мутантные формы, полученные после сайт-специфического мутагенеза соответствующих клонированных генов, используют для установления связи между структурой белка и его функцией. К ним относятся алкогольдегидрогеназа,  $\beta$ -лактамаза, цитохром *c*- и тирозил-тРНК-синтаза. С помощью методов клонирования были синтезированы важные в клиническом отношении белки, которые трудно было получить иным способом. Сюда относятся интерфероны, инсулин и гормон роста человека. Были синтезированы белки, кодируемые некоторыми вирусами животных (в том числе вирусами человека, такими, как вирус гепатита В) и простейшими (например, *Plasmodium falciparum*, возбудителем малярии человека). Они использовались в качестве антигенов в некоторых вакцинах.

### а. Выбор системы экспрессии

Для экспрессии уже клонированного гена выбор хозяина имеет такое же важное значение, как и при самом клонировании. Наиболее подходящими хозяевами для получения белка часто оказываются клетки *E. coli*, поскольку с ними просто манипулировать и их легко выращивать в больших объемах. После открытия интронов, прерывающих кодирующие участки многих эукариотических генов, для использования в экспрессирующих векторах чаще стали применять клонированные кДНК, а не сами гены. Однако при синтезе активных форм определенных эукариотических белков в клетках *E. coli* возникает ряд проблем. Так, первичные продукты трансляции некоторых эукариотических генов должны подвергнуться специфическим посттрансляционным модификациям, прежде чем образуются функ-

**РИС. 7.43.**

Общие свойства экспрессирующих векторов. Даже если вектор был сконструирован для экспрессии в клетках эукариот, он обычно является членистым и содержит последовательности, необходимые для репликации в *E. coli*. Это позволяет получить достаточное количество рекомбинантных молекул. Указаны точка инициации транскрипции и направление транскрипции от промотора, контролирующего экспрессию вставки. Ш-Д - последовательность Шайна-Дальгарно.

циональные молекулы. Обычно эти модификации состоят в протеолизе, фосфорилировании, ацетилировании и гликозилировании. Ферменты, катализирующие эти реакции в клетках *E. coli*, обычно отсутствуют. Отчасти для решения этой проблемы были разработаны эффективные эукариотические системы экспрессии. В животных системах хозяин-вектор в качестве кодирующих последовательностей не обязательно использовать кДНК, поскольку эти клетки способны удалять интроны.

Векторы, предназначенные для эффективной транскрипции и трансляции клонированного сегмента, содержат элементы, повышающие эффективность репликации вектора и экспрессии соответствующих генов (рис. 7.43). Репликация как таковая для экспрессии генов не требуется, но, обеспечивая образование большого числа матриц для транскрипции, она повышает содержание мРНК и полипептидов в клетке. Таким образом, экспрессирующие векторы обычно содержат точку начала репликации и кодируют некоторые другие функции, необходимые для эффективной внутриклеточной репликации и не обеспечиваемые хозяйскими клетками. Транскрипция зависит от наличия в векторе промотора, который соответствует присутствующей в клетках РНК-полимеразе, т.е. промотора *E. coli* для экспрессии в клетках бактерий и промотора,

используемого РНК-полимеразой II, для экспрессии в эукариотических клетках. Этот промотор должен быть расположен в векторе таким образом, чтобы клонированный ген мог встроиться в правильной ориентации и за промотором по ходу транскрипции. Эукариотические экспрессирующие векторы должны также иметь сайт для расщепления и полиаденилирования транскрипта-предшественника функциональной мРНК (гл. 8). Для точной трансляции в начале кодирующей области должен находиться кодон АТГ, а в конце — один из стоп-кодонов. Последний обычно содержится в клонированном сегменте, но стартовый кодон часто отсутствует во вставке кДНК, поэтому его функции должен выполнять вектор. Для трансляции в *E. coli* на 5'-конце на расстоянии примерно 10 нуклеотидов от иницирующего кодона должен находиться сайт связывания с рибосомой, последовательность Шайна-Дальгарно (разд. 3.6.a).

В экспрессирующие векторы часто бывают включены дополнительные элементы. Обычно они предназначаются для увеличения выхода полипептидов или для упрощения процедуры очистки нужного белка от других клеточных белков. Одним из основных факторов, ответственных за уменьшение выхода, является нестабильность многих белков в чужеродной клеточной среде.

Чтобы добиться экспрессии гена (или кДНК), обычно каждый раз приходится решать уникальную экспериментальную задачу. Поэтому многие экспрессирующие векторы, хотя и обладают некоторыми общими свойствами, имеют свои особенности. Ниже мы рассмотрим системы, в которых используются оригинальные экспрессирующие векторы, предназначенные для конкретных целей.

## б. Экспрессирующие векторы, используемые в *E. coli*

Векторы, сконструированные для осуществления эффективного белкового синтеза в *E. coli*, обычно происходят от рBR322 или другой высококопийной плазмиды; это необходимо для получения максимального числа матриц для транскрипции.

**Промоторы.** Большинство векторов, предназначенных для изучения экспрессии генов, содержат один из сильных промоторов, описанных в разд. 3.11: *lac*-промотор с соответствующим оператором, *trp*-промотор и оператор, а также P<sub>L</sub>-промотор бактериофага λ. Обычно используют мутантный *lac*-промотор, называемый UV5 (рис. 3.7), поскольку это позволяет осуществлять транскрипцию с большей скоростью, чем в случае промотора дикого типа, а кроме того, активность сохраняется даже в отсутствие положительного эффектора CAP-cAMP (разд. 3.11.в). Другой часто используемый промотор, *lac*, представляет собой синтетическую конструкцию: его -35-последовательность представляет собой -35-последовательность *trp*-промотора, а блок Прибнова (-10-последовательность) — соответствующую последовательность промотора UV5. Таким образом, *lac*-промотор идентичен оптимальному промотору *E. coli*, структура которого была определена исходя из данных, полученных при сравнении последовательностей многих природных промоторов (рис. 3.6). Как и ожидалось, он оказался весьма эффективным.

Помимо такого достоинства, как эффективность, эти четыре промотора обладают еще одним ценным свойством: они позволяют осуществлять временной контроль инициации транскрипции, поскольку связаны с регуляторными элементами. Это очень важно, так как накопление больших количеств чужеродного для *E. coli* белка может подавлять рост клеток и тем самым ограничивать конечный выход белка. Впрочем, если клетки удастся вырастить до высокой плотности, а затем индуцировать к оптимальной экспрессии клонированного гена, то часто можно достигнуть оптимального выхода белка. Если белок, который предполагается синтезировать, нестабилен в клетках *E. coli*, то молекулы, синтезированные в начале цикла роста, вероятно, деградируют ко времени достижения культурой высокой плот-

ности. Поэтому имеет смысл отсрочить экспрессию до того момента, когда плотность культуры повысится достаточно сильно. Если используются векторы, в которых экспрессия клонированного гена зависит от этих промоторов, то количество нужного белка обычно составляет от 1 до 10% уровня суммарного белка в клеточном экстракте.

**Экспрессирующий вектор с *lac*-промотором.** Плазмидные векторы, содержащие промотор *lac* UV5, часто содержат также последовательность Шайна-Дальгарно *lac*-оперона и кодирующие последовательности для β-галактозидазы (рис. 7.44). Поскольку экспрессия, иницируемая на участке с *lac*-оператором-промотором, находится под контролем *lac*-оперона, транскрипцию можно иницировать с помощью индуктора изопропил-β-D-тио-галактозида (IPTG). В приведенном примере вместо восьмого кодона β-галактозидазного гена встроены полилинкер, что нарушает кодирующую рамку. Клетки *E. coli*, несущие такой вектор, не способны синтезировать активную β-галактозидазу (в предположении, что клеточный геном является *lacZ*<sup>-</sup>). Две части рамки считывания можно совместить, если в один из рестриктазных сайтов полилинкера встроить фрагмент с определенной рамкой считывания. В этом случае трансляция мРНК может иницироваться в начале кодирующего участка β-галактозидазного гена, проходить через вставку и заканчиваться на обычном стоп-кодоне. В результате клетки, содержащие плазмиду с такой вставкой, будут продуцировать активную β-галактозидазу, поскольку первые восемь аминокислот молекулы несут незначительную нагрузку на фермент и чужеродные аминокислоты не будут влиять на нее. Гибридный полипептид можно очистить, при этом показателем степени очистки будет служить удельная активность β-галактозидазы. В некоторых случаях гибридный белок можно расщепить и удалить протяженную карбоксильную β-галактозидазную часть. Однако и сами гибридные белки находят применение: с их помощью получают антитела к чужеродному белку (а также к β-галактозидазе).

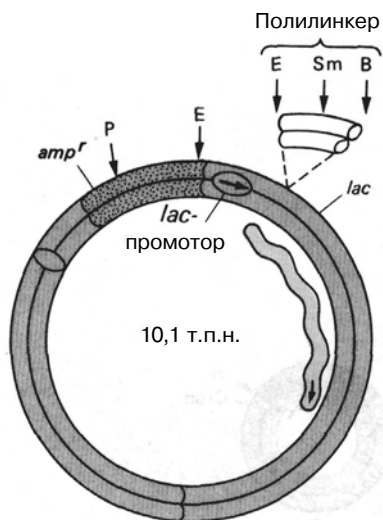
**Экспрессирующий вектор с *trp*-промотором.** В векторе, представленном на рис. 7.45, перед полилинкером находятся *trp*-промотор, оператор, последовательность Шайна-Дальгарно, аттенуатор, а также примерно 300 кодонов гена *trp* *E. coli*. Транскрипцию можно контролировать, поместив содержащие плазмиду клетки *E. coli*, выращенные до высокой плотности, в среду без триптофана (голодание) и добавив затем 3-β-индолилукриловую кислоту. Последняя конкурирует с оставшимся триптофаном за связывание с *trp*-репрессором, однако образует неактивный комплекс, в результате чего происходит дерепрессия промотора. Как и в случае с вектором,

содержащим *lac*-промотор, при встраивании чужеродного гена в рамку считывания по одному из сайтов рестрикции в полилинкере синтезируется гибридный белок. Трансляция останавливается, дойдя до случайного стоп-кодона вектора.

В случае, представленном на рис. 7.45, вставка кодирует обратную транскриптазу ретровируса. После трансфекции и дерепрессии в клетках *E. coli* синтезируется ферментативно активный гибридный белок. Для повышения стабильности обратной транскриптазы большую часть избыточных последовательностей гена *trpE* делетируют с помощью некоторых манипуляций и переклонирования. Несколько аминокислот *trpE*, остающихся на N-конце образовавшегося полипептида, слабо влияют на активность обратной транскриптазы.

*Экспрессирующий вектор, использующий P<sub>L</sub>-промотор бактериофага λ.* Вектор, представленный на рис. 7.46, позволяет синтезировать эукариотические белки, не сливающиеся с чужеродным полипептидом. Транскрипция с этого про-

мотора подавляется λ-репрессором, который образуется профагом, присутствующим в клетке-хозяине. Применение клеток, лизогенизированных фагом, который кодирует чувствительный к нагреванию репрессор (cI<sup>857</sup>), позволяет индуцировать экспрессию гена путем переноса клеток, выращенных при 30°C, в среду с температурой 42°C. Кроме промотора вектор содержит часть последовательности Шайна-Дальгарно гена cII фага λ. Между промотором и этой последовательностью находится группа регуляторных элементов фаговой ДНК, которые функционируют в *цис*-положении, ослабляя транскрипционную полярность. (Явление транскрипционной полярности состоит в уменьшении эффективности транскрипции, которое иногда наблюдается при увеличении расстояния от промотора.) Эти последовательности с антиполярным эффектом функционируют при наличии продукта гена *N* фага λ, поставляемого лизогенизированными хозяйскими клетками, которые обычно используются с вектором этого типа. В таком векторе сегмент, содержащий связывающийся с рибосомой участок Ш-Д последовательности, включает также кодон ATG cII-гена фага λ. Более того, ATG перекрывает *Bam*HI-сайт, унаследованный от плазмиды pBR322 (молекулы, на основе которой сконструирован вектор). Эукариотические кодирующие участки встраивают в *Bam*HI-сайт разными способами, зависящими от кодирующей последовательности. Очень важным моментом является встраивание кодирующей последовательности таким образом, чтобы она находилась в рамке с ATG. Таким образом, вектор предоставляет все регуляторные элементы, необходимые для транскрипции и трансляции, за исключением стоп-кодона, который обычно находится в пределах клонированных эукариотических кДНК.



Последовательности, примыкающие к полилинкеру

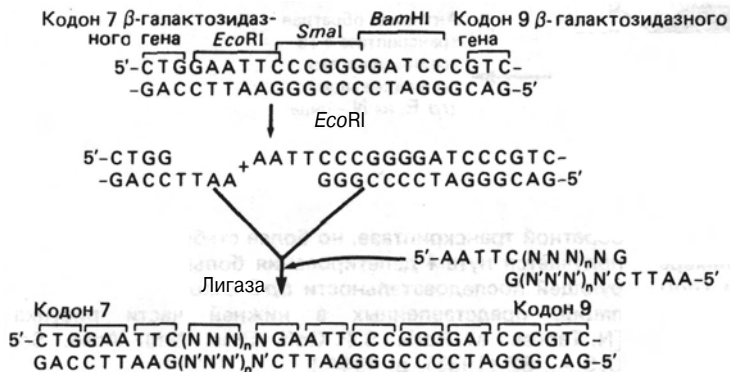
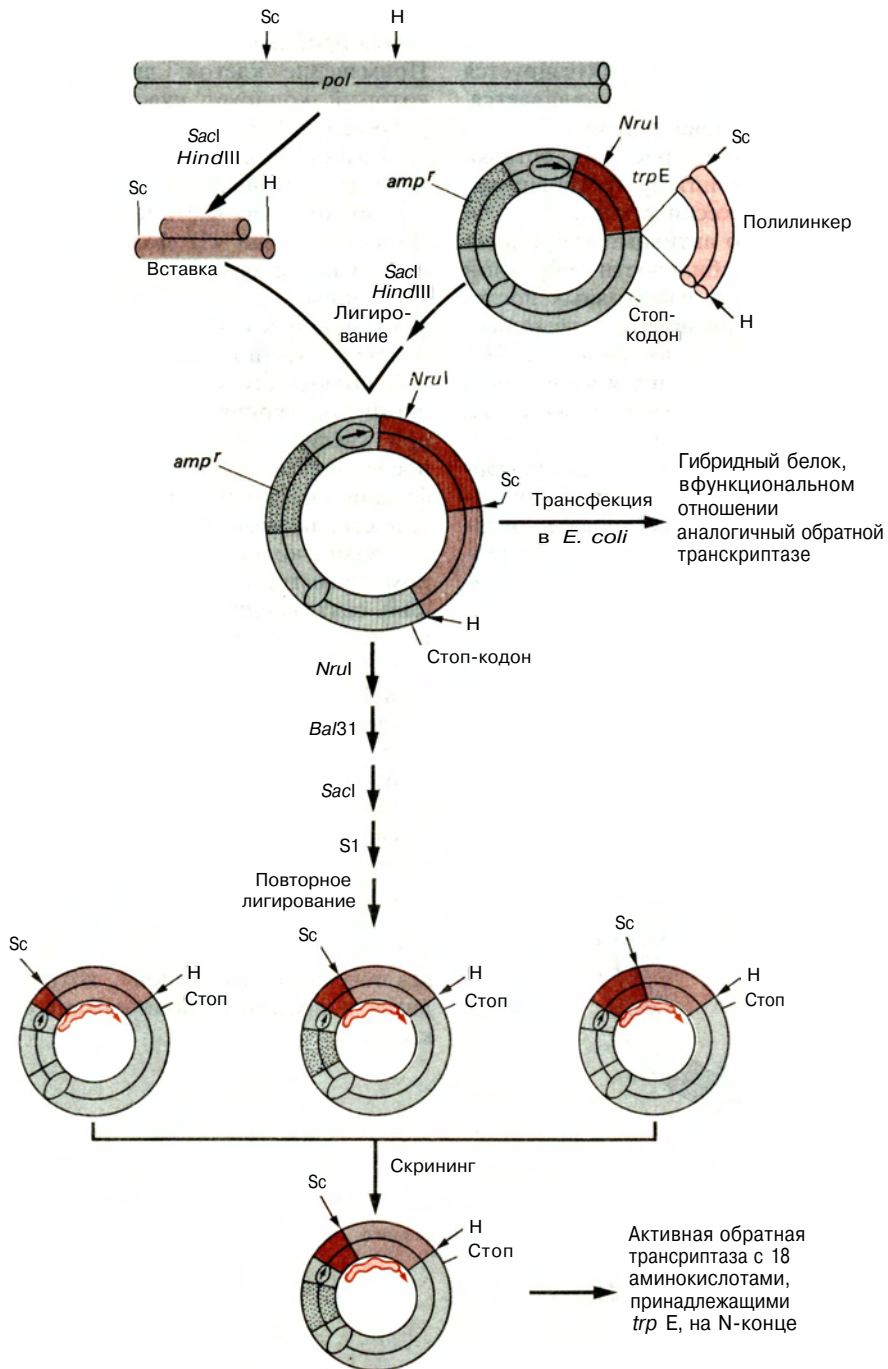


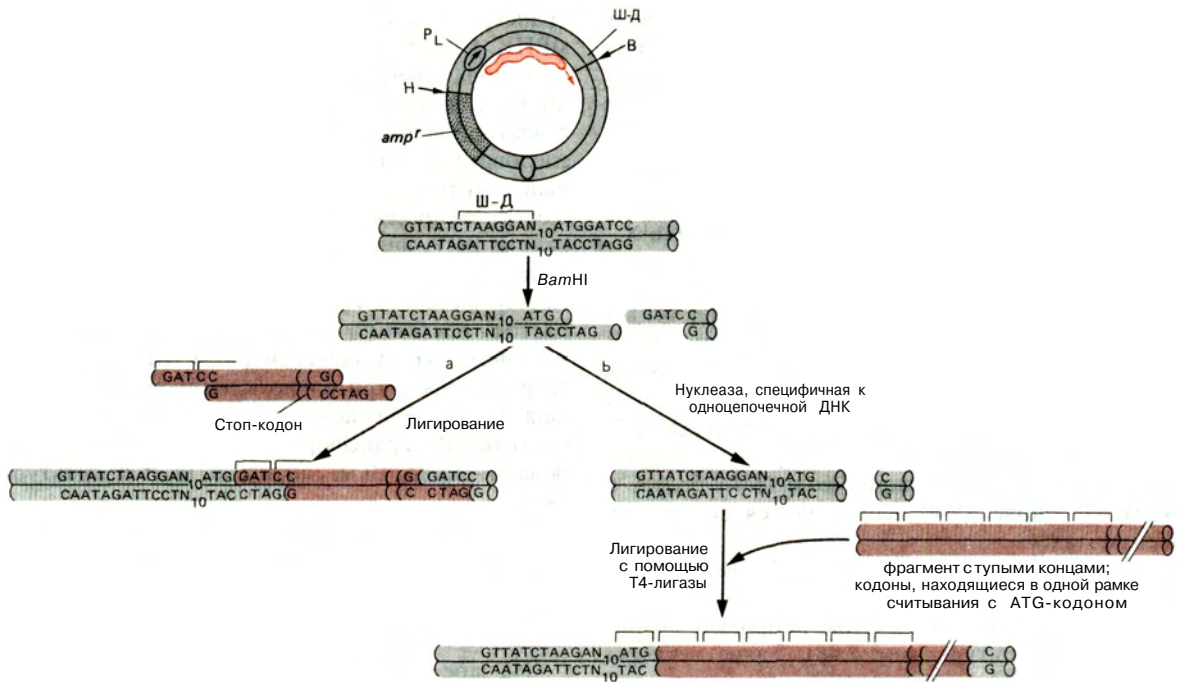
РИС. 7.44.

Вектор для экспрессии в *E. coli*, содержащий промотор *lac* UV5 и последовательность Шайна-Дальгарно, связывающуюся с рибосомами. Перед седьмым кодоном β-галактозидазного гена находится полилинкер, в который по любому из рестриктазных сайтов могут встраиваться вставки. Если в один из таких рестриктазных сайтов встроить фрагмент с нужной рамкой считывания, то две части β-галактозидазного гена можно соединить с образованием одной длинной непрерывной рамки. В результате в клетке будет синтезироваться гибридный полипептид. В приведенном примере вставка встроена в *Eco*RI-сайт; N и N' - спаренные основания.

**РИС. 7.45.**

Вектор для экспрессии в *E. coli*, содержащий *trp*-промотор и регуляторные сигналы. В приведенном примере кодирующая часть гена обратной транскриптазы (*pol*) ретровируса (разд. 5.7.г) встроена в полилинкер, прилегающий к кодирующему участку гена *trp E*. Образующийся гибридный белок функционально аналогичен

обратной транскриптазе, но более стабильный фермент получается путем делетирования большей части кодирующей последовательности *trp E* с помощью манипуляций, представленных в нижней части рисунка. [N. Tanese, M. Roth, S.P. Goff, Proc. Natl. Acad. Sci. U.S.A., **82** (1985), p. 4944.]

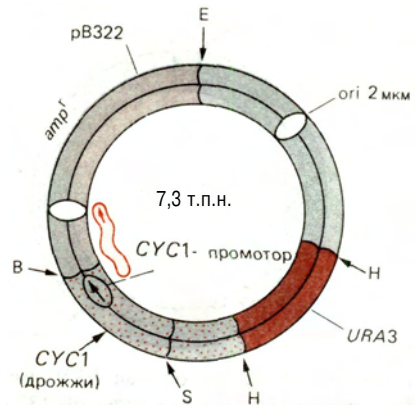
**РИС. 7.46.**

Вектор для экспрессии в *E. coli*, содержащий промотор  $P_L$  бактериофага  $\lambda$ . Регуляторная область этого вектора сконструирована путем лигирования различных элементов  $\lambda$ -генома с частью плазмиды pBR322. Сразу за ATG-кодоном генома фага  $\lambda$  находится *Bam* HI-сайт, в который могут быть встроены эукариотические кодирующие последовательности. Разработано множество

приемов, позволяющих присоединять кодирующую последовательность к вектору таким образом, чтобы новые кодоны находились в одной рамке считывания с ATG; два из них представлены на этом рисунке (варианты а и б). [F.M. Rosenberg, Y.-S. Ho, A. Shatzman, *Methods in Enzymology*, **101** (1983), p. 123.]

### в. Экспрессирующие векторы, используемые в клетках дрожжей

Наиболее эффективные в отношении экспрессии генов дрожжевые векторы сконструированы на основе 2 мкм-кольцевой плазмиды дрожжей (разд. 5.6.6). Эта плазмида обеспечивает образование в дрожжевых клетках большого числа копий рекомбинантной ДНК до тех пор, пока сегмент *REP3* находится в *цис*-положении, а функциональные гены *REP1* и *REP2* — либо в *цис*-, либо в *транс*-положении (например, в составе другой плазмиды). Применяются несколько разных регулируемых дрожжевых промоторов; в примере, представленном на рис. 7.47, это промотор гена *CYC1*, кодирующего изо-1-цитохром *c*. Этот промотор и связанные с ним регуляторные сигналы составляют область размером в несколько сотен пар оснований, что типично для подобных сложных эукариотических областей, в которых транскрипция регулируется с помощью РНК-полимеразы II (гл. 8). Сегменты ДНК, трансляцию которых мы хотим осуществить, встраивают в рамку считывания в *Bam* HI-сайте. В этом случае

**РИС. 7.47.**

Вектор для экспрессии в дрожжевых клетках. Эта конструкция сходна с дрожжевым плазмидным вектором, представленным на рис. 5.25. Кроме сегментов плазмиды pBR322 он содержит часть 2 мкм-плазмиды, обеспечивающую образование большого числа копий рекомбинантной ДНК в клетках дрожжей, селективный маркер (дрожжевой ген *URA3*), промотор и регуляторные последовательности дрожжевого гена *CYC1*, а также удобный для клонирования *Bam* HI-сайт во втором кодоне кодирующей последовательности *CYC1*.



они непосредственно прилегают к иницирующему кодону ATG гена *CYC1*. Сайт полиаденилирования в векторе отсутствует, но обычно кДНК содержат такой сайт, который расположен за стоп-кодоном трансляции.

### г. Экспрессирующие векторы, используемые в клетках животных

Для экспрессии клонированных генов в клетках животных были сконструированы два типа векторов; в основе одного из них лежит геном SV40, а другого—геном папилломавируса крупного рогатого скота (BPV). Основные свойства этих двух систем вирусных векторов описаны в разд. 5.7. Векторы на основе папилломавирусов особенно полезны для синтеза больших количеств белка, а векторы на основе SV40 используются во многих экспериментах.

Типичные векторы содержат либо весь геном BPV, либо его часть, необходимую для стабильной трансформации хозяйских клеток, например мышечных клеток в культуре (рис. 7.48). Такие векторы часто включают эукариотический ген (или часть гена) вместе с промотором, сигналом полиаденилирования и другими регуляторными областями. Между сайтом инициации транскрипции и стартовым кодоном трансляции находится удобный рестриктазный сайт. Когда в этот сайт встраивается определенная кодирующая последовательность, она

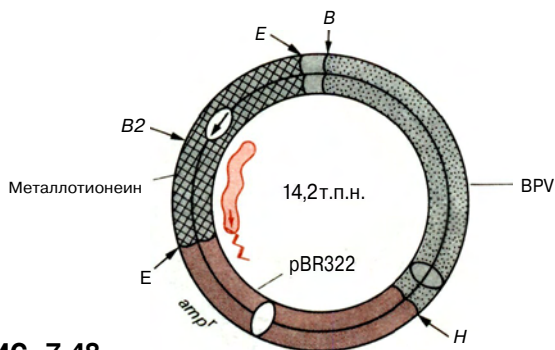


РИС. 7.48.

Экспрессирующий вектор на основе BPV. Вектор содержит последовательности плазмиды pBR322, позволяющие ему реплицироваться в *E. coli*, трансформирующую область генома BPV, ген мыши, кодирующий белок металлотиионеин. Между сайтом инициации транскрипции и первым кодоном гена металлотиионеина в рестриктазный *Bgl*III-сайт встроена кодирующая последовательность, включающая стартовый и стоп-кодны. Клетки мыши, трансформированные рекомбинантным вектором, синтезируют мРНК (стрелка с внутренней стороны кольца), которая включает вставку, и транслируют ее с образованием соответствующего полипептида.

транскрибируется с образованием соответствующей мРНК. Синтез этой мРНК начинается от стартового кодона транскрипции и заканчивается на сигнале полиаденилирования, находящемся во вставке или в векторной последовательности. Кодирующая последовательность должна содержать свой собственный стартовый и стоп-кодны. Одним из преимуществ экспрессирующих векторов, созданных на основе BPV, является то, что трансформированные ими клетки сохраняются в культуре в течение многих месяцев. В результате непрерывного деления клеток постоянно синтезируется нужный белок. В одном из экспериментов в BPV-вектор была встроена предварительно клонированная кодирующая часть гена поверхностного антигена вируса гепатита В. Трансформированные клетки секретировали около 10 мг антигена на 1 л культуры за 24 ч.

## 7.9. ФЕРМЕНТАТИВНАЯ АМПЛИФИКАЦИЯ СЕГМЕНТОВ ДНК И РНК

На заре развития методов молекулярного клонирования применение альтернативных способов амплификации специфических сегментов ДНК или РНК, присутствующих в больших популяциях молекул, казалось маловероятным. Однако в настоящее время такой метод разработан и широко используется. Он получил название **полимеразной цепной реакции (ПЦР)**. С его помощью можно получать микрограммы ДНК-копии сегментов ДНК или РНК, даже когда они присутствуют в препарате в виде единственной молекулы.

Вся полимеразная цепная реакция осуществляется *in vitro* с использованием ДНК-полимеразы и олигонуклеотидных праймеров, комплементарных двум 3'-концам участков, ограничивающих амплифицируемый дуплексный сегмент. Таким образом, разработка ПЦР-метода стала возможной лишь после того, как были созданы методы быстрого и недорогого химического синтеза олигонуклеотидов (разд. 6.1.в). Для осуществления реакции необходимо знать нуклеотидную последовательность того участка, который мы хотим амплифицировать, чтобы можно было синтезировать соответствующие олигонуклеотидные праймеры. При творческом применении ПЦР-метода он существенно дополняет методы молекулярной генетики, и если бы часть II книги не была фактически завершена до того, как этот метод был окончательно разработан, его основы были бы обязательно изложены в гл. 5–7.

**Полимеразная цепная реакция.** Для того чтобы амплифицировать какой-то сегмент ДНК, синтезируют два олигонуклеотидных праймера, каждый из которых комплементарен одному из двух 3'-концов

участков, ограничивающих амплифицируемый дуплексный сегмент. В ходе ПЦР-реакции необходимо копировать последовательность каждой из цепей, находящуюся между участками, с которыми спариваются олигонуклеотидные праймеры. После спаривания праймеров с денатурированной ДНК, содержащей амплифицируемый сегмент, осуществляют встречное удлинение праймеров с помощью ДНК-полимеразы в присутствии четырех дезоксирибонуклеозидтрифосфатов (рис. 7.49). Образующиеся дуплексные ДНК денатурируют, вновь спаривают с праймерами и проводят ДНК-полимеразную реакцию. Этот цикл (денатурация, спаривание, синтез) может быть повторен примерно 60 раз с удвоением в каждом цикле количества дуплексного сегмента ДНК. За  $n$  циклов образуется  $2^n$  копий дуплексного сегмента, ограниченного праймерами.

Используя термостабильную ДНК-полимеразу, выделенную из термофильных бактерий *Thermus aquaticus*, можно осуществлять множество ПЦР-циклов при однократном введении фермента. Сначала смешивают ДНК, избыточное количество праймерных молекул, дезоксирибонуклеозидтрифосфаты и полимеразу. Цикл запускают, нагрев смесь до температуры, обеспечивающей денатурацию ДНК; затем охлаждают раствор до температуры, необходимой для отжига праймеров. Далее устанавливают температуру, при которой может происходить синтез ДНК. Весь процесс, который длится несколько часов, можно автоматизировать, используя нагреватели, работа которых регулируется с помощью компьютера.

Использование ДНК-полимеразы *T. aquaticus* не только позволяет автоматизировать процесс, но дает и другое преимущество. Этот фермент наиболее активен в температурном интервале 70–75°C. При таких условиях спаривание олигонуклеотидных праймеров (их длина составляет примерно 20 остатков) и ДНК происходит более специфично, чем при 37°C – оптимальной температуре для ДНК-полимеразы *E. coli*. В результате ассоциация праймеров происходит более точно, что сводит к минимуму амплификацию нежелательных сегментов ДНК, особенно в присутствии всей геномной ДНК. Точность отжига повышается также при подборе оптимального температурного режима, ионной силы, длины праймера и его нуклеотидного состава. Специфичность может быть достаточно высокой, чтобы осуществлялась амплификация двух разных сегментов в одном и том же препарате ДНК в присутствии двух пар праймеров.

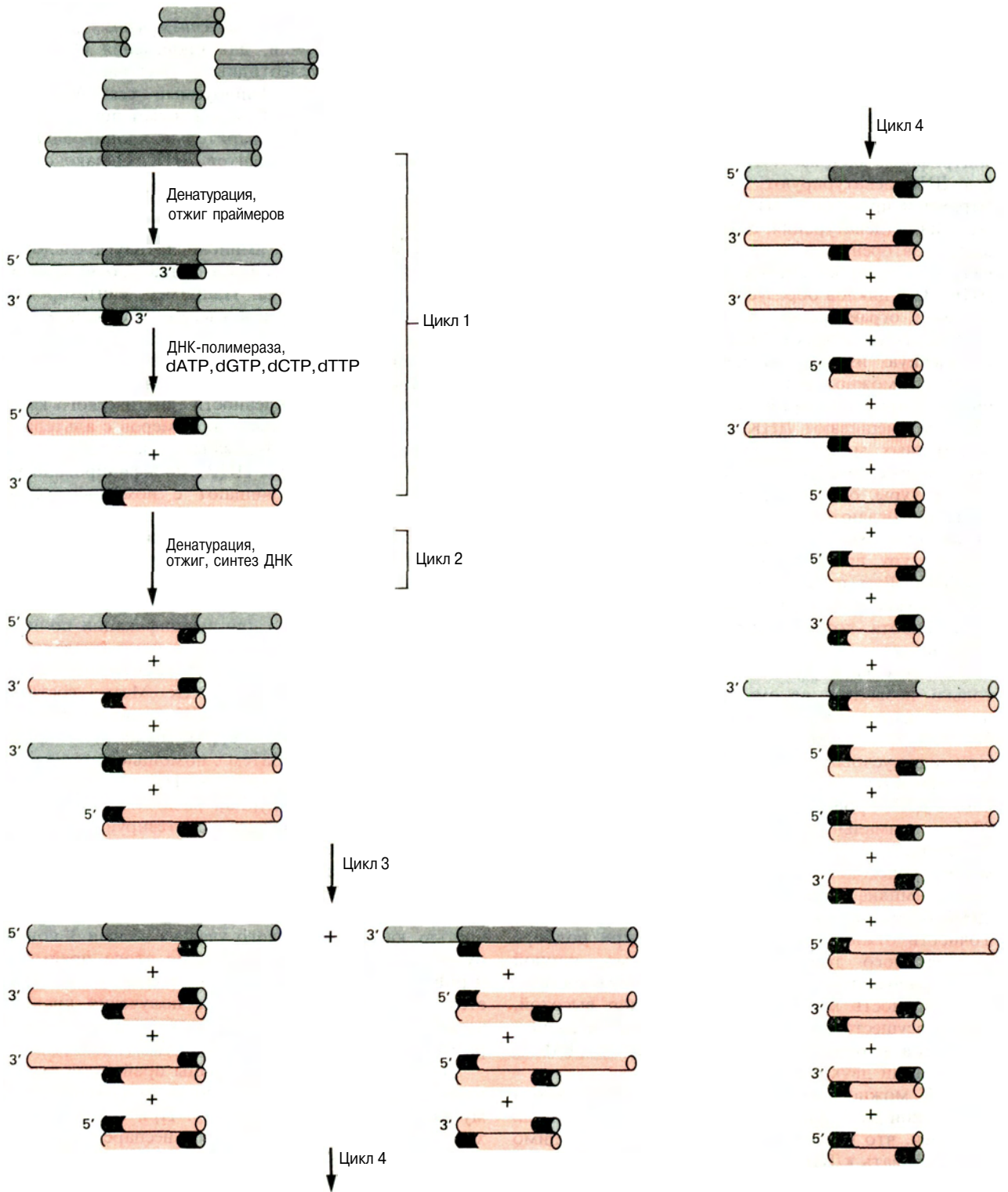
РНК можно амплифицировать с образованием дуплексной ДНК аналогичным образом с той лишь разницей, что в начале первого цикла необходимо синтезировать кДНК-копию РНК. Если РНК представлена молекулами мРНК, то роль одного из

праймеров в ПЦР, а также в реакциях с участием обратной транскриптазы могут выполнять короткие poly(A)-цепочки.

*Разные варианты полимеразных цепных реакций.* Как мы уже говорили, для проведения ПЦР необходимо знать нуклеотидные последовательности, фланкирующие амплифицируемый сегмент. Это подразумевает, что ПЦР-метод может применяться только при наличии предварительно клонированных и секвенированных сегментов ДНК. Однако с помощью относительно простых модификаций можно значительно расширить возможности метода ПЦР. В одном из вариантов можно выделить определенный ген, если известна аминокислотная последовательность лишь короткого участка соответствующего очищенного белка. Например, синтезировав праймеры длиной 20 пар нуклеотидов на основании данных о последовательности двух концов пептидного сегмента длиной в 20 аминокислот, можно амплифицировать геномный фрагмент длиной 60 п.н. Вследствие вырожденности генетического кода при этом используют смесь праймеров с альтернативными основаниями в нужных положениях (разд. 6.4.6). После проведения ПЦР амплифицированный сегмент из 60 п.н. очищают с помощью гель-электрофореза или клонирования и используют в качестве зонда для идентификации интересующей нас последовательности при анализе библиотек геномной ДНК или кДНК.

Еще один пример универсальности метода ПЦР дает использование его для амплификации полной кДНК-копии мРНК исходя из данных последовательности одного небольшого участка либо соответствующего пептида, либо мРНК. Метод состоит в следующем. Суммарную мРНК копируют с образованием одноцепочечной кДНК, используя короткий poly(dT)-праймер. Затем с помощью терминальной трансферазы к 3'-концу кДНК присоединяют «хвост» (например, poly(dG); разд. 4.8); используя короткий poly(dC)-праймер, синтезируют вторую цепь ДНК. До этого момента никакой специфичности в отношении определенной мРНК не проявляется. На следующем этапе используют олигонуклеотидный праймер с такой же последовательностью, как и у короткого участка вблизи 3'-конца нужной нам мРНК. С помощью этого праймера и poly(dC) проводят ПЦР; специфичность оказывается достаточно высокой для получения практически чистого продукта, представляющего полную мРНК.

В другом варианте в 5'-концы праймеров включают последовательности, соответствующие какому-либо рестриктазному сайту. При отжиге с исходной матрицей они остаются неспаренными, но амплифицируемый фрагмент приобретает концевые рестриктазные сайты. Они могут оказаться полез-



ными при последующем клонировании или включении в вектор для изучения экспрессии.

*Применение ПЦР.* Возможности метода ПЦР во всей полноте еще не раскрыты, но с его помощью уже удалось получить новые данные о структуре разных РНК, генов и геномов, а также о процессах, протекающих на геномном уровне. Одно из его очень важных применений состоит в определении характера мутаций. После того как ген клонирован, с помощью двух праймеров амплифицируют соответствующий геномный сегмент, полученный от разных особей данного вида. Секвенирование этого сегмента позволяет установить природу мутации, произошедшей в данном гене, или выявить полиморфизм, будь то замена одной пары оснований или делеция, инсерция или перестройка. Благодаря простоте метода с его помощью можно диагностировать генетические заболевания и проводить популяционно-генетические исследования. Используя праймеры, которые специфически ассоциируются с определенными мутантными формами гена, можно также классифицировать мутации.

С помощью зондов, комплементарных геномам таких инфекционных агентов, как вирусы и бактерии, могут быть идентифицированы геномы этих агентов на фоне значительного избытка ДНК клетки-хозяина. В частности, вполне реальной представляется диагностика СПИДа с помощью этого метода. Еще одно его применение состоит в анализе структуры продуктов реакций рекомбинации.

Одним из интересных и потенциально наиболее эффективных применений ПЦР является амплификация сегментов ДНК внутри отдельных клеток. Такие опыты проводились как на диплоидных клетках, так и на сперматозоидах человека. Использование сперматозоидов имеет особое значение для генетического анализа видов, включая человека, которые не могут подвергаться экспериментальному скрещиванию. Этот метод позволяет оценить частоту мейотической рекомбинации непосредственно в больших популяциях сперматозоидов, если использовать пару праймеров, специфичных в отношении полиморфных или мутантных форм сцепленных генов.

#### **РИС. 7.49.**

Полимеразная цепная реакция (ПЦР), позволяющая амплифицировать специфический сегмент ДНК. Два олигонуклеотидных праймера, по одному для каждой цепи, ограничивают амплифицируемый сегмент. Вначале ДНК денатурируют, затем производят отжиг праймеров, после чего ДНК-полимераза катализирует удлинение цепи в присутствии четырех нуклеозидтрифосфатов (цикл 1). Второй и последующие циклы запускают повышением температуры, вызывая денатурацию дуплексных молекул. В каждом цикле количество нужного сегмента удваивается. После нескольких циклов в смеси преобладают дуплексы, соответствующие нужному сегменту (короткие молекулы ДНК). Реакция идет эффективно, если размер амплифицируемого сегмента не превышает 2 т.п.н.

# Литература

## Введение

Приведены классические статьи и книги, в которых описываются эксперименты, заложившие теоретическую и техническую основу методов работы с рекомбинантными молекулами ДНК. Список дан в хронологическом порядке.

- F. D'Herelle.* 1926. *The Bacteriophage and Its Behavior.* Williams and Wilkins, Baltimore.
- O.T. Avery, C.M. Macleod, M. McCarty.* 1944. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. I. Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III. *J. Exp. Med.*, **79**, 137-158.
- M. Delbruck.* 1946. Experiments with Bacterial Viruses (Bacteriophages). *Harvey Lectures*, **41**, 161-187.
- J. Lederberg, E.L. Tatum.* 1946. Gene Recombination in *E.coli*. *Nature*, **158**, 558. (This paper describes conjugation.)
- N. Zinder, J. Lederberg.* 1952. Genetic Exchange in *Salmonella*. *J. Bacteriology*, **64**, 679-699. (В этой статье рассматривается трансдукция.)
- L.L. Cavalli-Sforza, J. Lederberg, E.M. Lederberg.* 1953. An Infective Factor Controlling Sex Compatibility in *Bacterium coli*. *J. Gen. Microbiol.*, **8**, 89-103.
- A. Lwoff.* 1954-1955. Control and Interrelations of Metabolic and Viral Diseases of Bacteria. In *Harvey Lectures*, series L, pp. 92-111. Academic Press, New York.
- F. Jacob, P. Schaeffer, E.L. Wollman.* 1960. Episomic Elements in Bacteria. *Symp. Soc. Gen. Microbiol.*, **10**, 67-91.
- A.D. Kaiser, D.S. Hogness.* 1960. The Transformation of *Escherichia coli* with Deoxiribonucleic Acid Isolated from Bacteriophage  $\lambda$ -dg. *J. Mol. Biol.*, **2**, 392-415.
- W. Arber, D. Dussoix.* 1962. Host specificity of DNA Produced by *Escherichia coli*. I. Host Controlled Modification of Bacteriophage  $\lambda$ . *J. mol. Biol.*, **5**, 18-36, 37-49.
- T. Watanabe.* 1963. Infective Heredity of Multiple Drug Resistance in Bacteria. *Bacteriological Reviews*, **27**, 87-115.
- W. Arber.* 1965. Host Controlled Modification of Bacteriophage. *Annu. Rev. Microbiol.*, **19**, 365-378.

- T.J. Kelly, Jr., H.O. Smith.* 1970. A Restriction Enzyme from *Hemophilus influenzae*. II. Base Sequence of the Recognition Site. *J. Mol. Biol.*, **51**, 393-409.
- H.O. Smith, K.W. Wilcox.* 1970. A Restriction Enzyme from *Hemophilus influenzae*. I. Purification and Chemical Properties. *J. Mol. Biol.*, **51**, 371-391.
- K.J. Danna, D. Nathans.* 1971. Sequence Specific Cleavage of Simian Virus 40 DNA by Restriction Endonucleases of *Hemophilus influenzae*. *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2913-2917.
- J. Hedgpeh, H.M. Goodman, H.W. Boyer.* 1972. DNA Nucleotide Sequence Restricted by the RI Endonuclease. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3448-3452.
- D.A. Jackson, R.H. Symons, P. Berg.* 1972. Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing  $\lambda$  Phage Genes and the Galactose Operon of *E.coli*. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2904-2909.
- J.E. Mertz, R.W. Davis.* 1972. Cleavage of DNA by RI Restriction Endonuclease Generates Cohesive Ends. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3370-3374.
- H. Temin, D. Baltimor.* 1972. RNA Directed DNA Synthesis and RNA Tumor Viruses. *Advances in Virus Research*, **17**, 129-186.
- S.N. Cohen, A.C.Y. Chang, H.W. Boyer, R.B. Hellings.* 1973. Construction of Biologically Functional Bacterial Plasmids in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3240-3244.
- P.E. Lobban, A.D. Kaiser.* 1973. Enzymatic End-to-End Joining of DNA Molecules. *J. Mol. Biol.*, **78**, 453-471.
- J.F. Morrow, S.N. Cohen, A.C.Y. Chang, H.W. Boyer, H.M. Goodman, R.B. Helling.* 1974. Replication and Transcription of Eucaryotic DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1743-1747.

В следующих работах описано историческое развитие методов работ с рекомбинантными ДНК.

- J. Cairns, G.S. Stent, J.D. Watson* (eds.). 1966. *Phage and the Origins of Molecular Biology.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

- S.N. Cohen*. 1975. The Manipulation of Genes. *Sci. American*, **233**, 25-33.
- M.F. Singer*. 1979. Introduction and Historical Background. In J.K. Setlow and A. Hollaender (eds.). *Genetic Engineering*, vol. 1, pp. 1-13. Plenum, New York.

## Общие работы для всех глав части II

Некоторые монографии и лабораторные руководства, описывающие основы и практические детали методов работы с рекомбинантными ДНК.

- P.D. Boyer* (ed.). 1981, 1982. *The Enzymes*, 3rd ed., vols. 14 and 15. Academic Press, New York.
- T. Maniatis, E.F. Fritsch, J. Sambrook*. 1982. *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. [Имеется перевод: Маниатис Т., Фрич Э., Сэмбрук Дж. Методы генетической инженерии. Молекулярное клонирование.- М.: Мир, 1984.]
- S.M. Linn, R.J. Roberts* (eds.). 1983. *Nucleases*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- J.D. Watson, J. Tooze, D.T. Kurtz*. 1983. *Recombinant DNA, a Short Course*. Scientific American Books, New York.
- D.M. Glover* (ed.). 1985, 1987. *DNA Cloning*, vols. 1, 2 and 3. IRL Press, Oxford, England, and Washington, D.C. [Имеется перевод: Клонирование ДНК. Методы. Под ред. Д. Гловера.- М.: Мир, 1988, 1989.]
- R.W. Old, S.B. Primrose*. 1985. *Principles of Gene Manipulation*, 3rd ed. Blackwell Scientific Publications, Oxford, England.
- P.H. Pouwels, B.E. Enger-Valk, W.J. Brammer*. 1985. *Cloning Vectors*. Elsevier, Amsterdam.
- F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.A. Smith, J.G. Seidman, K. Struhl* (eds.). 1987. *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley, New York.
- S.L. Berger, A.R. Kimmel* (eds.). 1987. *Guide to Molecular Cloning Techniques*. Academic Press, San Diego.
- 4.1**
- R.J. Crouch, M.L. Dirksen*. 1983. Ribonuclease H. In S.M. Linn and R.J. Roberts (eds.), *Nucleases*, pp. 211-241. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- 4.2**
- D. Nathans*. 1979. Restriction Endonucleases, Simian Virus 40 and the New Genetics. *Science*, **206**, 903-910.
- S.E. Halford*. 1983. How does EcoRI Cleave Its Recognition Site on DNA? *Trends Biochem. Sci.*, **8**, 455-460.
- C. Kessler, H.J. Holtke*. 1986. Specificity of Restriction Endonucleases and Methylases—A Review. *Gene*, **47**, 1-52.
- J.A. McClarin, C.A. Frederick, B.C. Wang, P. Greene, H.W. Byer, J. Grable, J.M. Rosenberg*. 1986. Structure of the DNA—EcoRI Endonuclease Recognition Complex at 3 Angstrom Resolution. *Science*, **234**, 1526-1541. [Миниобзор Т.Д. Ричмонд, в котором обсуждается эта работа, был опубликован в *Nature*, **326**, 18-19 (1987).]
- M. Nelson, M. McClelland*. 1987. The Effect of Site-Specific Methylation on Restriction-Modification Enzymes. *Nucleic Acids Res.* **15** (supplement) r219-r230.
- B.-Q. Qiang, I. Schildkraut*. 1987. NotI and SfiI: Restriction Endonucleases with Octanucleotide Recognition Sequences. *Methods in Enzymology*, **155**, 260-301.
- R.J. Roberts*. 1987. Restriction Enzymes and Their Isoschizomers. *Nucleic Acids Res.* **15** (supplement) r189-r217. (Это обширный список рестриктирующих эндонуклеаз, который ежегодно обновляется.)
- J.M. Rosenberg, J.A. McClarin, C.A. Frederick, B.-C. Wang, J. Grable, H.W. Boyer, P. Greene*. 1987. Structure and Recognition Mechanism of EcoRI endonuclease. *Trends Biochem. Sci.* **12**, 395-398.
- 4.3**
- G. Chaconas, J.H. van de Sande*. 1980. 5'-<sup>32</sup>P Labeling of RNA and DNA Restriction Fragments. *Methods in Enzymology*, **65**, 75-85.
- 4.5**
- I.R. Lehman*. 1974. DNA Ligase: Structure, Mechanism and Function. *Science*, **186**, 790-797.
- 4.6**
- P.W.J. Rigby, M. Dieckmann, C. Rhodes, P. Berg*. 1977. Labeling DNA to High Specific Activity in vitro by Nick Translation in vitro with DNA Polymerase I. *J. Mol. Biol.*, **113**, 237-251.
- 4.8**
- R. Roychoudhury, R. Wu*. 1980. Terminal Transferase-Catalyzed Addition of Nucleotides to the 3' Termini of DNA. *Methods in Enzymology*, **65**, 43-64.
- 5.1**
- M. Mandel, A. Higa*. 1970. Calcium-Dependent Bacteriophage DNA Infection. *J. Mol. Biol.*, **53**, 159-162.

## Литература к указанным разделам глав

- B.J. Bachmann.* 1983. Linkage Map of *Escherichia coli* K12, Edition 7. Microbiol. Rev., **47**, 180-230.
- R. Wu, L. Grossman* (eds.). 1987. Methods in Enzymology 153.
- ### 5.2
- S. Falkow.* 1975. Infectious Multiple Drug Resistance. Pion, London. (Книга посвящена открытию плазмид и их биологии.)
- F. Bolivar, K. Backman.* 1979. Plasmids of *E.coli* as Cloning Vectors. Methods in Enzymology, **68**, 245-267.
- R.P. Novick.* 1980. Plasmids. Sci. American, **243**, 102-127.
- ### 5.3
- J. Messing, J. Vieira.* 1982. A new Pair of M13 Vectors for Selecting Either DNA Strand of Double-Digest Restriction Fragments. Gene, **19**, 269-276.
- A. Campbell.* 1983. Bacteriophage  $\lambda$ . In J.A. Shapiro (ed.), Mobile Genetic Elements, pp. 65-103. Academic Press, New York.
- M. Feiss.* 1986. Terminase and the Recognition, Cutting, and Packaging of  $\lambda$  Chromosomes. Trends in Genetics, **2**, 100-104.
- ### 5.4
- J. Collins, B. Hohn.* 1978. Cosmids: A Type of Plasmid Gene-Cloning Vector That is Packageable in vitro in Bacteriophage  $\lambda$  Heads. Proc. Natl. Acad. Sci. U.S.A. **75**, 4242-4246.
- J. Collins.* 1979. *E.coli* Plastids Packageable in vitro in  $\lambda$  Bacteriophage Particles. Methods in Enzymology, **68**, 309-326.
- B. Hohn.* 1979. In vitro packaging of  $\lambda$  and Cosmid DNA. Methods in Enzymology, **68**, 299-309.
- S. Brenner, G. Cesareni, J. Karn.* 1982. Plasmids: Hybrids between ColE1 Plasmids and *E.coli* Bacteriophage  $\lambda$ . Gene, **17**, 27-44.
- ### 5.5
- P.S. Lovett, K.M. Keggins.* 1979. *Bacillus subtilis* as a Host for Molecular Cloning. Methods in Enzymology, **68**, 342-357.
- D.A. Hopwood, M.J. Bibb, K.F. Chater, T. Kieser.* 1987. Plasmid and Phage Vectors for Gene Cloning and Analysis in *Streptomyces*. Methods in Enzymology, **153**, 116-166.
- ### 5.6
- C. Ilgen, P.J. Farabaugh, A. Hinnen, J.M. Walsh, G.-R. Fink.* 1979. Transformation of Yeast. In J.K. Setlow and A. Hollander (eds.). Genetic Engineering, vol.1, pp. 117-132. Plenum, New York.
- D. T. Stinchcomb, K. Struhl, R.W. Davis.* 1979. Isolation and Characterization of a Yeast Chromosomal Replicator. Nature, **282**, 39-43.
- L. Clarke, J. Carbon.* 1980. Isolation of a Yeast Centromere and Construction of Functional Small Circular Chromosomes. Nature, **287**, 504-509.
- J.R. Broach.* 1982. The Yeast Plasmid 2 $\mu$  Circle. Cell, **28**, 203-204.
- R.J. Rothstein.* 1983. One-step Gene Disruption in Yeast. Methods in Enzymology, **101**, 202-211.
- K. Struhl.* 1983. The New Yeast Genetics. Nature, **305**, 391.
- R.K. Mortimer, D. Schild.* 1985. Genetic Map of *Saccharomyces cerevisiae*, Edition 9. Microbiol. Rev., **49**, 181-213.
- J. Campbell.* 1988. Eukaryotic DNA Replication: Yeast Bares its ARSs. Trends Biochem. Sci., **13**, 212-217.
- ### 5.7
- S.P. Goff, P. Berg.* 1976. Construction of Hybrid Viruses Containing SV40 and A Phage DNA Segments and Their Propagation in Cultured Monkey Cells. Cell, **9**, 695-705.
- C. Shih, B.-Z. Shilo, M.P. Goldfarb, A. Danenberg, R.A. Weinberg.* 1979. Passage of Phenotypes of Chemically Transformed Cells via Transfection of DNA and Chromatin. Proc. Natl. Acad. Sci. U.S.A. **76**, 5714-5718.
- D.H. Homer.* 1980. DNA Cloning in Mammalian Cells with SV40 Vectors. In J.K. Setlow, A. Hollander (eds.), Genetic Engineering, vol. 2, pp. 83-101. Plenum, New York.
- R.C. Mulligan, P. Berg.* 1980. Simian Virus 40 Mediated Expression of a Bacterial Gene in Mammalian Cells. Science, **209**, 1422-1427.
- A. Pellicer, D. Robins, B. Wold, R. Sweet, J. Jackson, I. Lowy, J.M. Roberts, G.K. Sim, S. Silverstein, R. Axel.* 1980. Altering Genotype and Phenotype by DNA-mediated Gene Transfer. Science, **209**, 1414-1422.
- R.C. Mulligan, P. Berg.* 1981. Selection for Animal Cells That Express the *E.coli* Gene Coding for Xanthine-Guanine-Phosphoribosyl Transferase. Proc. Natl. Acad. Sci. U.S.A. **78**, 2072-2076.
- B.H. Howard.* 1983. Vectors for Introducing Genes into Cells of Higher Eukaryotes. Trends Biochem. Sci., **8**, 209-212.
- C.L. Cepko, B.E. Roberts, R.C. Mulligan.* 1984. Construction and Applications of a Highly Transmissible Murine Retro virus Shuttle Vector. Cell, **37**, 1053-1062.
- J. Yates, N. Warren, D. Reisman, B. Sugden.* 1984. A cis-Acting Element from Epstein-Barr Viral Genome That Permits Stable Replication of Recombinant Plasmids in Latently Infected Cells. Proc. Natl. Acad. Sci. U.S.A. **81**, 3806-3810.
- H.M. Temin.* 1986. Retrovirus Vectors for Gene Transfer: Efficient Integration into and Expression of Exogenous DNA in Vertebrate Cell Genomes. In:

- R. Kucherlapati (ed.), *Gene Transfer*, pp. 149-158. Plenum, New York.
- D. DiMaio. 1987. Papillomavirus Cloning Vectors. In: N.P. Salzman and P.M. Howley (eds.). *The Papovaviridae*, vol. 2. The Papillomaviruses, pp. 293-320. Plenum, New York.
- C. Cepko. 1988. Retrovirus Vectors and Their Application in Neurobiology. *Neuron.*, **1**, 345-353.
- 5.8**
- M.W. Bevan, M.-D. Chilton. 1982. T-DNA of the *Agrobacterium* TI and RI Plasmids. *Annu. Rev. Genet.*, **16**, 357-384.
- M.-D. Chilton. 1983. A Vector for Introducing New Genes into Plants. *Sci. American*, **249**, 50-60.
- P. Zambryski, H.M. Goodman, M. Van Montagu, J. Schell. 1983. *Agrobacterium* Tumor Induction. In: J. A. Shapiro (ed.). *Mobile Genetic Elements*, pp. 505-535. Academic Press, New York.
- R.B. Horsch, J.E. Fry, N.L. Hoffman, D. Eichholtz, S.G. Rogers, R.T. Fraley. 1985. A Simple and General Method for Transferring Genes into Plants. *Science*, **227**, 1229-1231.
- R. Deblaere, A. Reynaerts, H. Hofte, J.-P. Hernalsteens, J. Leemans, M. Van Montagu. 1987. Vectors for Cloning in Plant Cells. *Methods in Enzymology*, **153**, 277-292.
- M. Fromm, J. Callis, L.P. Taylor, V. Walbott. 1987. Electroporation of DNA and RNA into Plant Protoplasts. *Methods in Enzymology*, **153**, 351-366.
- A. Weissbach, H. Weissbach (eds.). 1988. *Methods for Plant Molecular Biology*. Academic Press. Orlando.
- P. Zambryski, J. Tempe, J. Schell. 1989. Transfer and Function of T-DNA Genes from *Agrobacterium* Ti and Ri Plasmids in Plants. *Cell*, **56**, 193-201.
- 6.1**
- E.M. Southern. 1974. Detection of Specific Sequences Among DNA Fragments Separated by Gel Electrophoresis. *J. Mol. Biol.*, **98**, 503-517.
- H. Okayama, P. Berg. 1982. High Efficiency Cloning of Full Length cDNA. *Mol. Cell. Biol.*, **2**, 161-170.
- V. Pirotta. 1984. Chromosome Microdissection and Microcloning. *Trends Biochem. Sci.*, **9**, 220-221.
- M.H. Caruthers. 1985. Gene Synthesis Machines: DNA Chemistry and Its Uses. *Science*, **230**, 281-285.
- R. Wu, L. Grossman (eds.). 1987. Chemical Synthesis and Analysis of Oligodeoxynucleotides. *Methods in Enzymology*, **154**, 221-326.
- 6.4**
- M. Grunstein, D.S. Hogness. 1975. Colony Hybridization: A Method for the Isolation of Cloned DNAs That Contain a Specific Gene. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961-3965.
- R.A. Kramer, J.R. Cameron, R.W. Davis. 1976. Isolation of Bacteriophage  $\lambda$  Containing Yeast Ribosomal RNA Genes: Screening by in situ RNA Hybridization to Plaques. *Cell*, **8**, 227-232.
- W.D. Benton, R.W. Davis. 1977. Screening  $\lambda$ gt Recombinant Clones by Hybridization to Single Plaques in situ. *Science*, **196**, 180-182.
- B.M. Paterson, B.E. Roberts, E.L. Kuff. 1977. Structural Gene Identification and Mapping by DNA-mRNA Hybrid-Arrested Cell-Free Translation. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4370-4374.
- M. Grunstein, J. Wallis. 1979. Colony Hybridization. *Methods in Enzymology*, **68**, 379-389.
- R. Wu (ed.). 1979. Screening and Selection of Cloned cDNA. *Methods in Enzymology*, **68**, 389-454.
- R.A. Young, R.W. Davis. 1983. Efficient Isolation of Genes Using Antibody Probes. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1194-1198.
- 6.5**
- T. Maniatis, R.C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G.K. Sim, A. Efstratiadis. 1978. The Isolation of Structural Genes from Libraries of Eucaryotic DNA. *Cell*, **15**, 687-701.
- A. Efstratiadis, L. Villa-Komaroff. 1979. Cloning of Double-Stranded cDNA. In: J. K. Setlow and A. Hollaender (eds.). *Genetic Engineering*, vol. 1, pp. 15-36. Plenum, New York.
- T.D. Sargent, I.B. Dawid. 1983. Differential Gene Expression in the Gastrula of *Xenopus laevis*. *Science*, **222**, 135-139.
- G.F. Carle, M.V. Olson. 1984. Separation of Chromosomal DNA Molecules from Yeast by Orthogonal-Field-Alter nation Gel Electrophoresis. *Nucleic Acids Res.* **12**, 5647-5664.
- D.C. Schwartz, C.R. Cantor. 1984. Separation of Yeast Chromosome-Sized DNAs by Pulsed Field Gradient Gel Electrophoresis. *Cell*, **37**, 67-75.
- G. Chu, D. Vollrath, R.W. Davis. 1986. Separation of Large DNA Molecules by Contour-Clamped Homogeneous Electric Fields. *Science*, **234**, 1582-1585.
- R. Wu, L. Grossman (eds.). 1987. cDNA Cloning. *Methods in Enzymology*, **154**, 3-83.
- 6.6**
- E.M. Tobin, J. Silverthorne. 1985. Light Regulation of Gene Expression in Higher Plants. *Annu. Rev. Plant Physiol*, **36**, 596-593.
- R.J. Konopka. 1987. Genetics of Biological Rhythms in *Drosophila*. *Annu. Rev. Genet.*, **21**, 227-236.
- M. Heisenberg. 1988. Designing Heat-Shock Clocks. *Nature*, **333**, 19-20.
- 7.1**
- M. Thomas, R.L. White, R.W. Davis. 1976. Hybridization of RNA to Double-Stranded DNA: Formation



- of R Loops. Proc. Natl. Acad. Sci. U.S.A. **73**, 2294-2298.
- R.L. White, D.S. Hogness.* 1977. R Loop Mapping of the 18S and 28S Sequences in the Long and Short Repeating Units of *Drosophila melanogaster* rDNA. Cell, **10**, 177-192.
- L. Grossman, K. Moldave* (eds.). 1980. Determination of Fragment Ordering. Methods in Enzymology, **65**, 429-494.
- R.C. Parker.* 1980. Determination of DNA Fragment Sizes. Methods in Enzymology, **65**, 415-426.
- M.M. Gottesman* (ed.). 1987. Methods in Enzymology, 151.
- R. Wu* (ed.). 1987. Methods in Enzymology, 155.
- ## 7.2
- A.M. Maxam, W. Gilbert.* 1977. A New Method for Sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**, 560-564.
- F. Sanger, S. Nicklen, A.R. Coulsen.* 1977. DNA Sequencing with Chain-Terminating Inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5467.
- F. Sanger.* 1980. Determination of Nucleotide Sequences in DNA. Nobel Lecture, December 8. Bioscience Reports, **1**, 3-18.
- G. M. Church, W. Gilbert.* 1984. Genomic Sequencing. Proc. Natl. Acad. Sci. U.S.A. **81**, 1991-1995.
- K.B. Millis, F.A. Faloona.* 1987. Specific Synthesis of DNA in vitro via a Polymerase-Catalyzed Chain Reaction. Methods in Enzymology, **155**, 335-350.
- L.M. Smith, R.J. Kaiser, J.Z. Sanders, L.E. Hood.* 1987. The Synthesis and Use of Fluorescent Oligonucleotides in DNA Sequence Analysis. Methods in Enzymology, **155**, 260-301.
- D.R. Engelke, P.A. Hoener, F.S. Collins.* 1988. Direct Sequencing of Enzymatically Amplified Human Genomic DNA. Proc. Natl. Acad. Sci. U.S.A. **85**, 544-548.
- R.J. Saiki, D.H. Gelfand, S. Stoffel, S.J. Scharff, R. Higuchi, G.T. Horn, K.B. Mullis, H.A. Erlich.* 1988. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. Science, **239**, 487-491.
- ## 7.3
- T.R. Gingeras, R.J. Roberts.* 1980. Steps Toward Computer Analysis of Nucleotide Sequences. Science, **209**, 1322-1328.
- C.L. Queen, L.J. Korn.* 1980. Computer Analysis of Nucleic Acids and Proteins. Methods in Enzymology, **65**, 595-609.
- D. Soil, R.J. Roberts* (eds.). 1984. The Applications of Computers to Research. Nucleic Acids, part I and II. IRL Press, Oxford, England.
- M.J. Bishop, C.J. Rowlings.* 1987. Nucleic Acid and Protein Sequence Analysis: A Practical Approach. IRL Press, Oxford, England.
- R.F. Doolittle.* 1987. Of URFS and ORFS: A Primer on How to Analyze Derived Amino Acid Sequences. University Science Books, Mill Valley, California.
- G. von Heijne.* 1987. Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit? Academic Press, New York.
- C. DiLisi.* 1988. Computers in Molecular Biology: Current Applications and Emerging Trends. Science, **240**, 47-52.
- Nucleic Acids Research 16(5A) 1988, весь выпуск.
- ## 7.4
- M.E. Harper, G.F. Saunders.* 1981. Localization of Single Copy DNA Sequences on G-Banded Human Chromosomes by in situ Hybridization. Chromosoma, **83**, 431-439.
- F.H. Ruddle.* 1982. A New Era in Mammalian Gene Mapping: Somatic Cell Genetics and Recombinant DNA Methodologies. Nature, **294**, 115-119.
- P. D'Eustachio, F.H. Ruddle.* 1983. Somatic Cell Genetics and Gene Families. Science, **220**, 919-928.
- P.M. Rosenburg, Y.-S. Ho, A. Shatzman.* 1983. Methods in Enzymology, **101**, 123.
- ## 7.5
- J.G. Wetmur, N. Davidson.* 1968. Kinetics of Renaturation of DNA. J. Mol. Biol, **31**, 349-370.
- S.S. Longacre, B. Mach.* 1979. A Precise Quantitation of Gene Number by Saturation Hybridization Using Cloned DNA. Nucleic Acids Res., **6**, 1241-1258.
- ## 7.6
- E. Melgar, D.A. Goldthwait.* 1968. Deoxyribonucleic Acid Nucleases II. The Effects of Metals on the Mechanism of Action of Deoxyribonuclease I. J. Biol. Chem., **243**, 4409-4416.
- T.E. Shenk, J. Carbon, P. Berg.* 1976. Construction and Analysis of Viable Deletion Mutants of Simian Virus 40. J. Virology, **18**, 664-671.
- C. A. Hutchison III, S. Phillips, M.H. Edgell, S. Gillam, P. Jahnke, M. Smith.* 1978. Mutagenesis at a Specific Position in a DNA Sequence. J. Biol. Chem., **253**, 6551-6560.
- R.M. Myers, L.S. Lerman, T. Maniatis.* 1985. A General Method for Saturation Mutagenesis of Cloned DNA Fragments. Science, **229**, 242-247.
- D. Shortle, D. Botstein.* 1985. Strategies and Application of in vitro Mutagenesis. Science, **229**, 1193-1201.
- M. Smith.* 1985. In vitro Mutagenesis. Annu. Rev. Genet. **19**, 423-462.
- R. Wu, L. Grossman* (eds.). 1987. Site-Specific Mutagenesis and Protein Engineering. Methods in Enzymology, **154**, 329-533.

## 7.7

- A.J. Berk, P.A. Sharp.* 1978. Structure of the Adenovirus 2 Early mRNAs. *Cell*, **14**, 695-711.
- D.R. Parks, V.M. Bryan, V.T. Oi, L.A. Herzenberg.* 1979. Antigen-Specific Identification and Cloning of Hybridomas with a Fluorescence-Activated Cell Sorter. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1962-1966.
- J.B. Gurdon, M.P. Wickens.* 1983. The Use of *Xenopus* Oocytes for the Expression of Cloned Genes. *Methods in Enzymology*, **101**, 370-386.
- P. Kavathas, V.P. Sukhatme, L.A. Herzenberg, J.R. Partis.* 1984. Isolation of the Gene Encoding the Human T-Lymphocyte Differentiation Antigen Leu-2(T8) by Gene Transfer and cDNA Subtraction. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7688-7692.

## 7.8

- W. Gilbert, L. Villa-Komaroff.* 1980. Useful Proteins from Recombinant Bacteria. *Sci. American*, **243**, 74-95.
- R. Wu, L. Grossman* (eds.). 1987. Expression Vectors. *Methods in Enzymology*, **153**, 385-563.

## 7.9

- H.A. Erlich, D.H. Gelfand, R.K. Saiki,* 1988. Specific DNA Amplification. *Nature*, **331**, 461-462.
- M.A. Frohman, M.K. Dush, G.R. Martin.* 1988. Rapid Production of Full-Length cDNAs from Rare Transcripts: Amplification Using a Single Gene-Specific Oligonucleotide Primer. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8998-9002.
- H. Li., U.B. Gyllensten, X. Cui, R. Saiki, H.A. Erlich, N. Arnheim.* 1988. Amplification and Analysis of DNA Sequences in Single Human Sperm and Diploid Cells. *Nature*, **335**, 414-417.
- N. Gautam, M. Baetscher, R. Aebersold, M.I. Simon.* 1989. A G Protein Gamma Subunit Shares Homology with ras Proteins. *Science*, **244**, 971-974.
- E.Y. Loh, J.F. Elliott, S. Cwirla, L.L. Lanier, M.M. Davis.* 1989. Polymerase Chain Reaction with a Single-Sided Specificity: Analysis of T Cell Receptor  $\delta$  Chain. *Science*, **243**, 217-220.

# Оглавление

От редактора перевода . . . . .	5	д. Гибридные спирали ДНК–РНК . . . . .	54
Предисловие . . . . .	6	1.3. Структура белков . . . . .	56
Благодарности . . . . .	10	а. Компоненты белков и соединяющие их химические связи . . . . .	56
<hr/>		б. Размер и форма белков . . . . .	58
<b>Часть I. Молекулярные основы наследственности: обзор</b>		в. Чем определяется конформация белка . . . . .	59
Введение . . . . .	11	<hr/>	
Хромосомы . . . . .	16	<b>Глава 2. Репликация, сохранение и модификация генома . . . . .</b>	
Клеточный цикл . . . . .	16	<hr/>	
Мейоз и образование гамет . . . . .	18	2.1. Репликация ДНК . . . . .	67
Строение хромосом . . . . .	18	а. Матричная функция ДНК при репликации . . . . .	67
Наследование одиночных признаков	23	б. Репликация начинается в определенных точках . . . . .	68
Независимая сегрегация и независимое комбинирование . . . . .	25	в. Репликация ДНК полуконсервативна . . . . .	72
Связь между генами и хромосомами	25	г. Комплементарное копирование оснований, перенос дезоксирибонуклеотидов и лигирование ДНК при репликации . . . . .	74
Рекомбинация . . . . .	25	д. Ключевые ферменты, участвующие в синтезе ДНК . . . . .	78
Связь между генами и белками . . . . .	26	е. Для репликации необходимо раскручивание спирали . . . . .	82
Гены и ДНК . . . . .	27	ж. Инициация образования новых цепей ДНК и их рост в репликативных вилках . . . . .	86
Генетика - молекулярная наука . . . . .	30	з. Терминация репликации ДНК и расхождение дочерних спиралей . . . . .	92
Перенос генетической информации в клетке . . . . .	31	2.2. Репликация РНК с образованием ДНК . . . . .	93
Структура и сохранение геномной ДНК . . . . .	32	а. Репликация геномов ретровирусов	93
Экспрессия и регуляция генов . . . . .	33	б. Некоторые ДНК-содержащие вирусы используют для репликации обратную транскрипцию . . . . .	97
<hr/>		2.3. Репарация ДНК . . . . .	97
<b>Глава 1. Молекулы генетического аппарата</b>	38	а. Репарация путем прямого восстановления исходной структуры . . . . .	99
1.1. Структура и поведение ДНК . . . . .	39	б. Репарация путем замены модифицированных остатков . . . . .	100
а. Компоненты молекулы ДНК и соединяющие их химические связи	39	в. Значение репарации ДНК . . . . .	102
б. Спиральная структура ДНК . . . . .	41	2.4. Рекомбинация ДНК . . . . .	103
в. Альтернативные формы двойной спирали ДНК . . . . .	42	<hr/>	
г. Размер молекул ДНК . . . . .	44		
д. Разнообразие форм ДНК . . . . .	44		
е. Денатурация и ренатурация ДНК	47		
ж. Упаковка ДНК в хромосомах	49		
1.2. Структура и поведение РНК . . . . .	52		
а. Типы РНК и их распространенность	52		
б. Компоненты молекулы РНК и соединяющие их химические связи	52		
в. Структура РНК . . . . .	52		
г. Денатурация и ренатурация РНК	54		

а. Типы рекомбинации . . . . .	104		
б. Общая рекомбинация между гомологичными молекулами ДНК . . . . .	105		
в. Ферменты, участвующие в общей рекомбинации . . . . .	108		
г. Сайт-специфическая рекомбинация . . . . .	110		
2.5. Репликация . . . . .	111		
<hr/>			
Глава 3. Аппарат экспрессии генов и его логика . . . . .	115		
<hr/>			
3.1. Основные положения процесса экспрессии генов . . . . .	115		
а. Транскрипция ДНК в РНК . . . . .	115		
б. Соответствие между нуклеотидными триплетами и аминокислотами . . . . .	116		
в. Расшифровка кода с помощью тРНК . . . . .	116		
г. Правильная инициация трансляции . . . . .	116		
д. Трансляция кодонов и соединение аминокислот . . . . .	117		
е. Регуляция экспрессии генов на разных этапах образования РНК и белка . . . . .	117		
3.2. Транскрипция: передача информации о нуклеотидной последовательности ДНК на уровень РНК . . . . .	117		
а. Синтез РНК на ДНК-матрице . . . . .	118		
б. ДНК-зависимые РНК-полимеразы . . . . .	121		
в. Транскрипция инициируется в особых нуклеотидных последовательностях . . . . .	123		
г. Терминация транскрипции и отделение цепей РНК . . . . .	125		
3.3. Процессинг РНК у прокариот . . . . .	127		
а. Группы генов, кодирующих рРНК и тРНК . . . . .	127		
б. Разрезание рРНК-тРНК-котранскриптов . . . . .	128		
в. Образование зрелых тРНК из более крупных транскриптов . . . . .	128		
3.4. Генетический код . . . . .	130		
а. Аминокислотная последовательность белков соответствует нуклеотидной последовательности кодирующих их генов . . . . .	131		
б. Соответствие между аминокислотами и их кодонами . . . . .	132		
в. Расшифровка генетического кода . . . . .	132		
г. Избыточность генетического кода . . . . .	136		
д. Универсальность генетического кода . . . . .	137		
3.5. Аппарат трансляции . . . . .	138		
а. Присоединение аминокислот к «родственным» тРНК . . . . .	138		
		б. На рибосомах осуществляются спаривание аминоксил-тРНК с кодонами и сборка белковых цепей . . . . .	142
		3.6. Трансляция мРНК у прокариот . . . . .	145
		а. Условия инициации . . . . .	146
		б. Элонгация полипептидной цепи . . . . .	150
		в. Терминация элонгации полипептидной цепи . . . . .	151
		3.7. Некоторые общие особенности процесса трансляции . . . . .	151
		а. Одновременная трансляция молекулы мРНК более чем одной рибосомой . . . . .	152
		б. Трансляция бактериальных мРНК может осуществляться параллельно транскрипции . . . . .	152
		в. Рибосомы начинают новый раунд после трансляции кодирующей последовательности . . . . .	153
		г. Взаимодействие кодона и антикодона . . . . .	153
		3.8. Трансляция мРНК у эукариот . . . . .	157
		а. Особые модификации мРНК эукариот . . . . .	159
		б. Инициация трансляции на 5'-кэпированных концах малыми рибосомными субчастицами . . . . .	160
		в. Элонгация и терминация полипептидной цепи . . . . .	160
		3.9. Ингибиторы транскрипции и трансляции . . . . .	160
		а. Ингибирование РНК-полимеразы . . . . .	160
		б. Ингибирование трансляции . . . . .	161
		3.10. Судьба синтезированных белков . . . . .	163
		а. Посттрансляционная модификация полипептидных цепей . . . . .	163
		б. Доставка эукариотических белков к клеточным мембранам и проникновение через них . . . . .	164
		в. Транспорт белков в эукариотические клеточные органеллы . . . . .	170
		г. Транспорт белков в клетках прокариот . . . . .	172
		3.11. Регуляция генной экспрессии . . . . .	172
		а. Регуляция содержания РНК в процессе биосинтеза . . . . .	172
		б. Согласованная регуляция экспрессии прокариотических генов . . . . .	173
		в. Регуляция экспрессии лактозного оперона . . . . .	174
		г. Регуляция экспрессии триптофанового оперона . . . . .	178
		д. Временная регуляция генной экспрессии в жизненном цикле бактерии фагаλ . . . . .	181

е. Трансляционная регуляция экспрессии некоторых генных про- дуктов . . . . .	185
Литература . . . . .	189

## Часть II. Триумф рекомбинантных ДНК

Введение . . . . .	195
Введение новой генетической информации в клетки бактерий . . . . .	196
Трансформация бактерий . . . . .	196
Конъюгация . . . . .	197
Трансдукция . . . . .	198
Принципы клонирования . . . . .	202
Концепция рекомбинантной ДНК . . . . .	203
Важные открытия . . . . .	204
Бактериальные плазмиды . . . . .	204
Рестриктирующие эндонуклеазы . . . . .	207
Глава 4. Инструментарий: ферменты . . . . .	211
4.1. Нуклеазы . . . . .	211
а. Общие свойства . . . . .	211
б. Нуклеазы, специфичные в отношении одноцепочечной ДНК . . . . .	212
в. Нуклеаза <i>Bal 31</i> . . . . .	213
г. РНКазы <i>H</i> . . . . .	214
4.2. Эндонуклеазы рестрикции . . . . .	214
а. Три типа эндонуклеаз рестрикции . . . . .	214
б. Типичная рестриктирующая эндонуклеаза типа II . . . . .	215
в. Различные группы рестриктирующих эндонуклеаз типа II . . . . .	216
г. Картирование сегментов ДНК с помощью рестриктирующих эндонуклеаз типа II . . . . .	218
д. Защита ДНК посредством метилирования . . . . .	218
4.3. Фосфомоноэстеразы . . . . .	221
4.4. Полинуклеотидкиназа . . . . .	221
4.5. ДНК-лигаза . . . . .	222
4.6. ДНК-полимераза I . . . . .	222
а. Полифункциональный фермент . . . . .	222
б. Ник-трансляция . . . . .	222
в. Заполнение брешей . . . . .	223
4.7. РНК-зависимые ДНК-полимеразы (обратные транскриптазы) . . . . .	223
4.8. Терминальная дезоксирибонуклеотидилтрансфераза . . . . .	224
а. Полимеризация без матрицы . . . . .	224
б. Синтез липких концов . . . . .	224
4.9. poly(A)-полимераза . . . . .	226
Глава 5. Инструментарий: системы хозяин-вектор . . . . .	227
5.1. <i>E. coli</i> -системы: клетки-хозяева . . . . .	228

а. Разносторонность хозяина . . . . .	228
б. «Гостеприимство» хозяина . . . . .	228
в. Доступность хозяина . . . . .	228
г. Некоторые примеры . . . . .	229
5.2. <i>E. coli</i> -системы: плазмидные векторы . . . . .	229
а. Модульная структура плазмид . . . . .	229
б. Конструирование векторов для отбора . . . . .	232
в. Плазмидный вектор pBR322 . . . . .	234
г. Другие векторы, используемые для разных целей . . . . .	236
5.3. <i>E. coli</i> -системы: фаговые векторы . . . . .	237
а. Некоторые различия между плазмидами и фаговыми векторами . . . . .	237
б. Фагλ . . . . .	238
в. Векторы, сконструированные на основе фага λ . . . . .	240
г. Упаковка λ-векторных молекул в фаговые частицы . . . . .	241
д. Фаг M13 . . . . .	242
е. Векторы, сконструированные на основе фага M13 . . . . .	243
5.4. <i>E. coli</i> -системы: плазмидно-фаговые векторы . . . . .	246
а. Космиды . . . . .	246
б. Фазмиды . . . . .	247
5.5. Другие прокариотические системы хозяин-вектор . . . . .	249
а. Грамотрицательные организмы . . . . .	249
б. Грамположительные организмы . . . . .	249
в. Челночные векторы . . . . .	250
5.6. Эукариотические системы хозяин-вектор: дрожжи . . . . .	252
а. Универсальность и удобство . . . . .	252
б. Векторы, способные реплицироваться в клетках дрожжей . . . . .	252
в. Стабильная трансформация при рекомбинации с дрожжевым геномом . . . . .	254
5.7. Эукариотические системы хозяин-вектор: животные . . . . .	256
а. Трансформация клеток животных . . . . .	256
б. Векторы на основе SV40 . . . . .	260
в. Векторы на основе вируса папилломы крупного рогатого скота . . . . .	265
г. Векторы на основе ретровирусов . . . . .	267
5.8. Эукариотические системы хозяин-вектор: растения . . . . .	273
а. Общие положения . . . . .	273
б. Плазмида pTi-A, индуцирующая образование опухолей . . . . .	273
в. Конструирование векторов рекомбинантных ДНК с использованием pTi . . . . .	275

Глава 6. Средства: конструирование, клонирование и отбор рекомбинантных молекул ДНК. . . . .	277	7.3. Компьютерный анализ нуклеотидных последовательностей ДНК . . . . .	328
6.1. Вставки . . . . .	277	а. Хранение информации о первичной структуре. . . . .	328
а. Общие положения. . . . .	277	б. Структурный анализ . . . . .	330
б. Вставки геномной ДНК . . . . .	277	в. Биологическое значение . . . . .	331
в. Синтетические вставки . . . . .	280	7.4. Определение положения клонированных сегментов в геномах. . . . .	333
г. Копирование РНК с образованием ДНК. . . . .	285	а. Молекулярная локализация . . . . .	333
6.2. Лигирование вектора со вставкой . . . . .	286	б. Хромосомная локализация . . . . .	334
а. Соединение концов. . . . .	286	в. Нестабильность при клонировании . . . . .	338
б. Присоединение липких концов . . . . .	289	7.5. Определение числа копий данной последовательности в геноме . . . . .	339
6.3. Инфекция, трансфекция и клонирование . . . . .	290	а. Оценка числа копий с помощью гибридизации ДНК–ДНК . . . . .	339
а. Перенос рекомбинантных молекул из пробирки в клетку. . . . .	290	б. Оценка числа копий по кинетике реассоциации ДНК. . . . .	339
б. Клонирование . . . . .	292	в. Оценка числа копий с помощью гибридизации в условиях насыщения . . . . .	342
6.4. Скрининг клонированных популяций рекомбинантных молекул . . . . .	292	7.6. Изменение клонированных сегментов: получение мутантов. . . . .	343
а. Обнаружение нужного клона . . . . .	292	а. Общие положения . . . . .	343
б. Отжиг с комплементарным полинуклеотидом . . . . .	292	б. Делеционные мутанты. . . . .	343
в. Определение экспрессии гена в клетках . . . . .	299	в. Инсерционные мутанты . . . . .	344
6.5. Библиотеки . . . . .	302	г. Точечные мутации. . . . .	344
а. Геномные библиотеки. . . . .	304	7.7. Изучение функций клонированных сегментов ДНК. . . . .	348
б. Библиотеки кДНК. . . . .	304	а. Характеристика внутриклеточных транскриптов, соответствующих клонированным сегментам ДНК . . . . .	349
6.6. Некоторые стратегии клонирования генов и кДНК. . . . .	316	б. Функциональное тестирование клонированной ДНК. . . . .	353
Глава 7. Продукты рекомбинации: характеристика и манипулирование . . . . .	317	7.8. Синтез полипептидов, кодируемых клонированными сегментами эукариотической ДНК . . . . .	354
7.1. Макроструктура клонированной вставки . . . . .	317	а. Выбор системы экспрессии . . . . .	354
а. Размер вставки . . . . .	317	б. Экспрессирующие векторы, используемые в <i>E. coli</i> . . . . .	356
б. Картирование сайтов для рестриктирующих эндонуклеаз . . . . .	318	в. Экспрессирующие векторы, используемые в клетках дрожжей . . . . .	359
в. Субклонирование . . . . .	318	г. Экспрессирующие векторы, используемые в клетках животных . . . . .	360
г. Определение положения интересующего нас сегмента во вставке . . . . .	318	7.9. Ферментативная амплификация сегментов ДНК и РНК. . . . .	360
7.2. Тонкая структура клонированной вставки: нуклеотидная последовательность . . . . .	319	Литература. . . . .	364
а. Общие принципы . . . . .	320		
б. Химическое секвенирование . . . . .	321		
в. Ферментативное секвенирование . . . . .	324		

Максин Сингер, Пол Берг  
ГЕНЫ И ГЕНОМЫ

В 2-х томах  
Том 1

Зав. редакцией канд. биол. наук *М.Д. Гроздова*  
Ведущий редактор *Н.Н. Шафрановская*  
Редактор *В.И. Николаева*  
Художник *В.П. Медников*  
Технический редактор *Л.П. Бирюкова*  
Корректор *В.И. Николаева*

Лицензия ЛР № 010174 от 20.05.97 г.

Сдано в набор 11.04.95г. Подписано к печати 3.03.98. Формат 84 x 108 <sup>1</sup>/<sub>16</sub>.  
Бумага офсетная. Печать офсетная. Гарнитура таймс. Объем 11,75 бум. л. Усл.  
печ. л. 39,48. Усл. кр.-отт. 79,38. Уч. изд. л. 41,69. Изд. № 4/9020. Тираж 5000 экз.  
Зак. № 243. С 031

Издательство «Мир» Государственного комитета Российской Федерации  
по печати  
129820, ГСП, Москва, И-110, 1-й Рижский пер., 2

ОАО «Можайский полиграфкомбинат»  
143200, г. Можайск, ул. Мира, 93



9 785030 028491



**Максин Сингер** родилась в Нью-Йорке. Окончила Йельский университет. В 1957 г. получила степень доктора философии в области биохимии. С 1956 по 1975 г. работала в Национальном институте артрита и нарушений обмена веществ. В 1975 г. перешла в Национальный институт рака, где в 1979 г. возглавила лабораторию биохимии. В 1988 г. д-р Сингер стала Президентом института Карнеги, сохраняя за собой руководство лабораторией. Как член Национальной академии наук она возглавляла редколлегию журнала *Proceedings of the National Academy of Sciences*. В 1988 г. получила Президентскую премию за достижения в области науки, что является высочайшей наградой, присуждаемой гражданским лицам. В настоящее время д-р Сингер занимается изучением мобильных элементов генома человека.

**Пол Берг** родился в Нью-Йорке. Окончил Университет шт. Пенсильвания. В 1952 г. получил степень доктора философии в области биохимии. Продолжил свою деятельность в Институте клеточной физиологии в Копенгагене и в Университете Вашингтона в Сент-Луисе. Затем перешел в Станфордский университет, где и продолжает работать в качестве профессора биохимии и директора Бекмановского центра по молекулярно-генетическим исследованиям в медицине. В 1980 г. Берг получил Ласкеровскую премию в области медицины и Нобелевскую премию по химии за свои исследования нуклеиновых кислот, в частности рекомбинантных ДНК. Занимается анализом экспрессии генов эукариот.

