

# Структура биополимеров

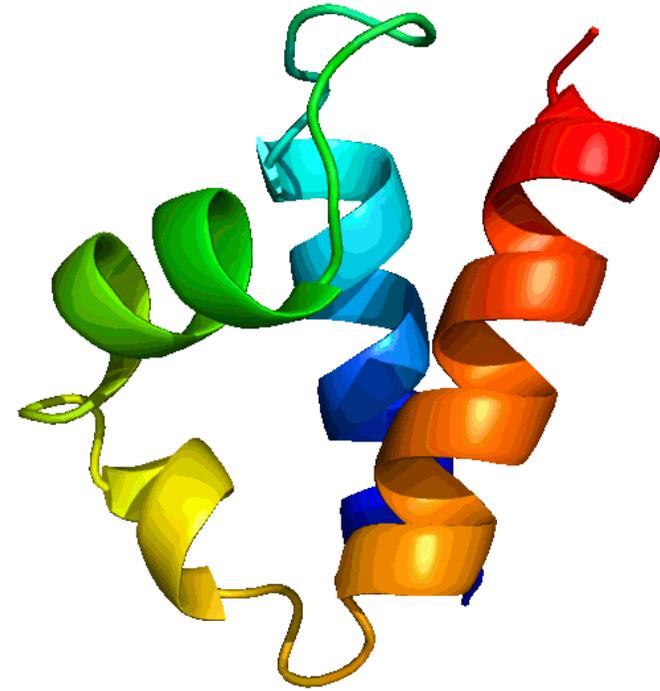
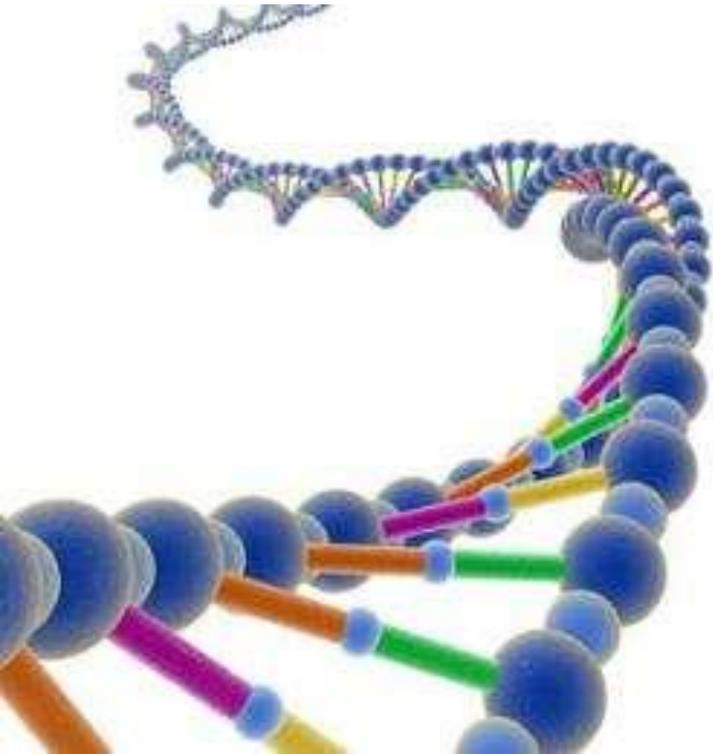
д.б.н., проф. Анна Карягина

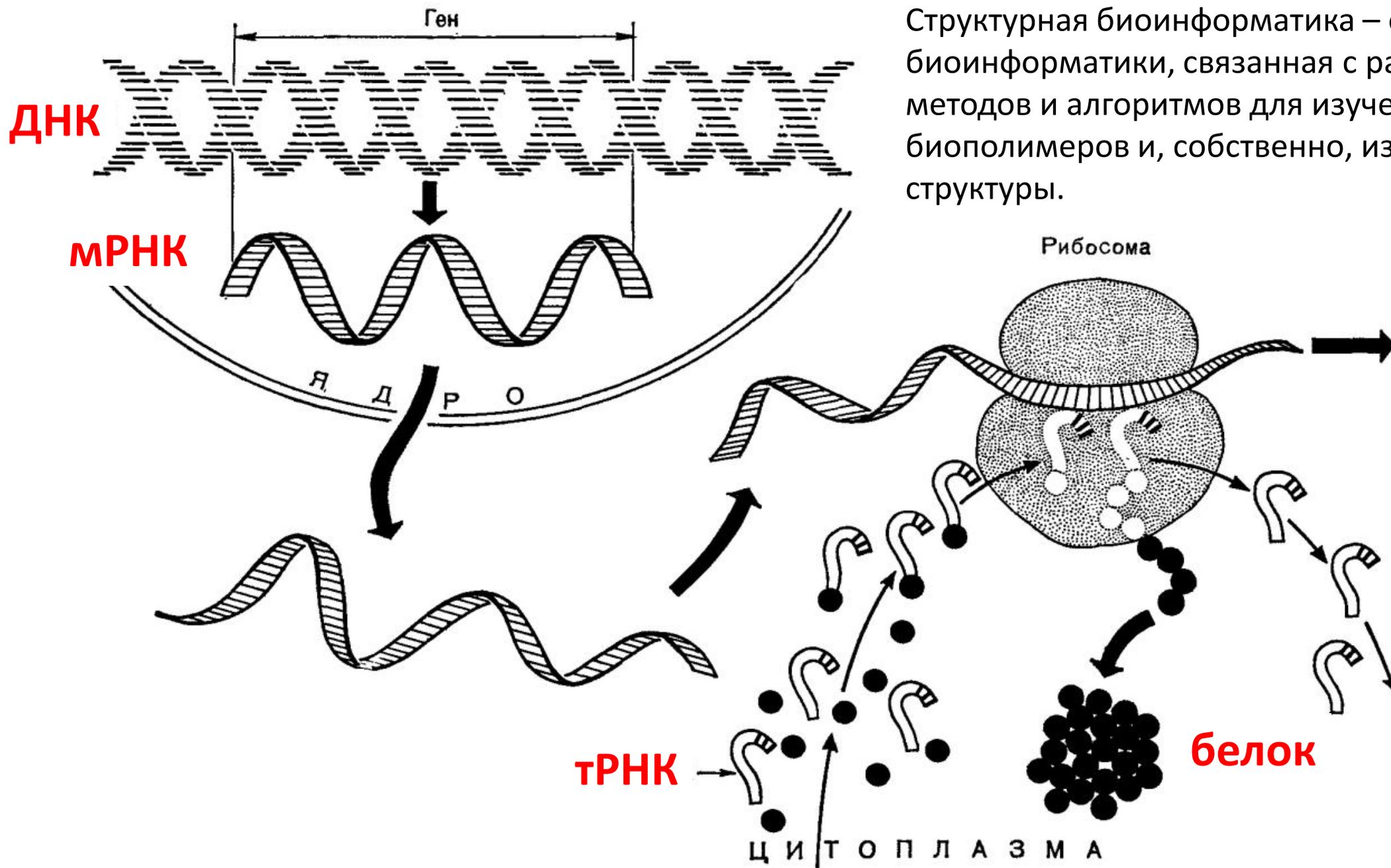
НИИ ФХБ им. А.Н. Белозерского МГУ им. М.В Ломоносова,

ФГБУ «НИЦ эпидемиологии и микробиологии им. Н.Ф. Гамалеи»

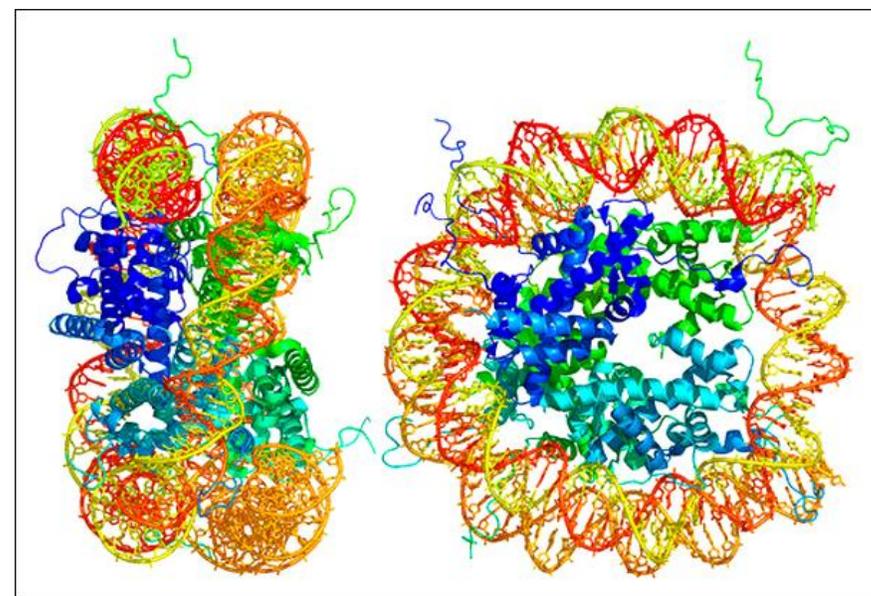
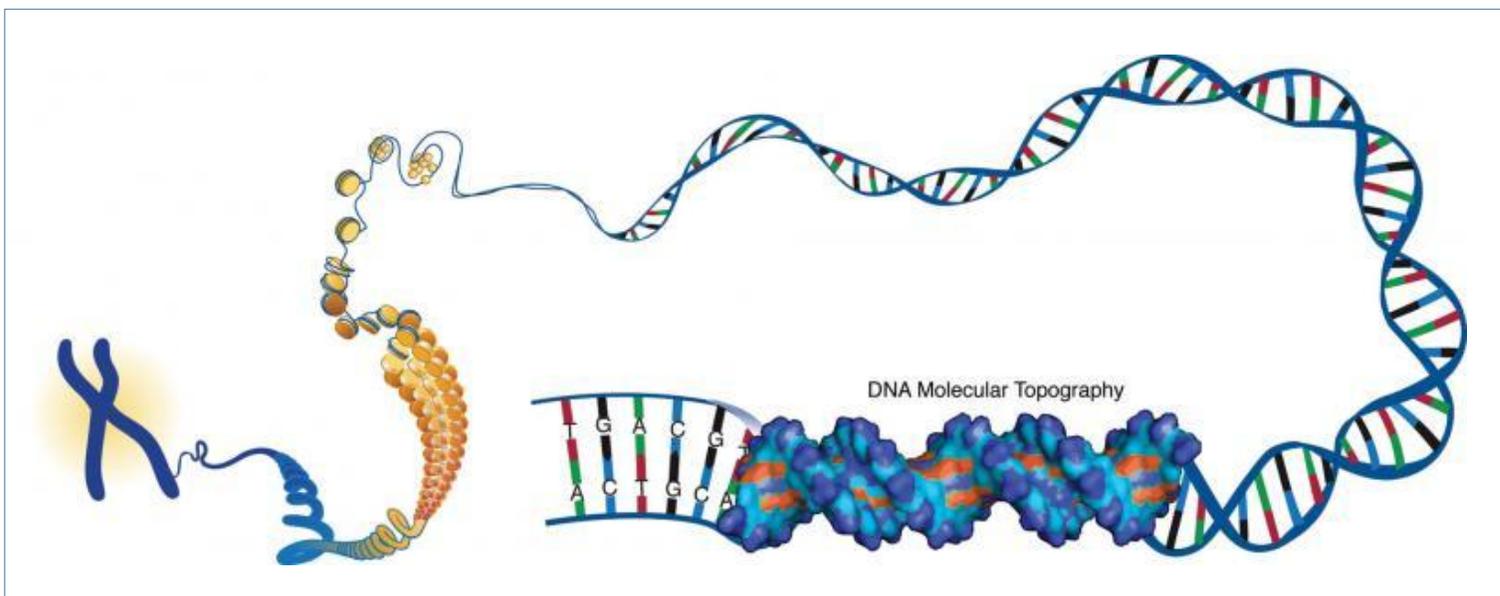
Минздрава РФ

e-mail akaryagina@gmail.com





Структурная биоинформатика – область биоинформатики, связанная с разработкой методов и алгоритмов для изучения структур биополимеров и, собственно, изучением их структуры.



У биополимеров есть последовательность составляющих их мономерных блоков. Цепочка биополимера имеет структуру, причем, уровней структурной организации много.

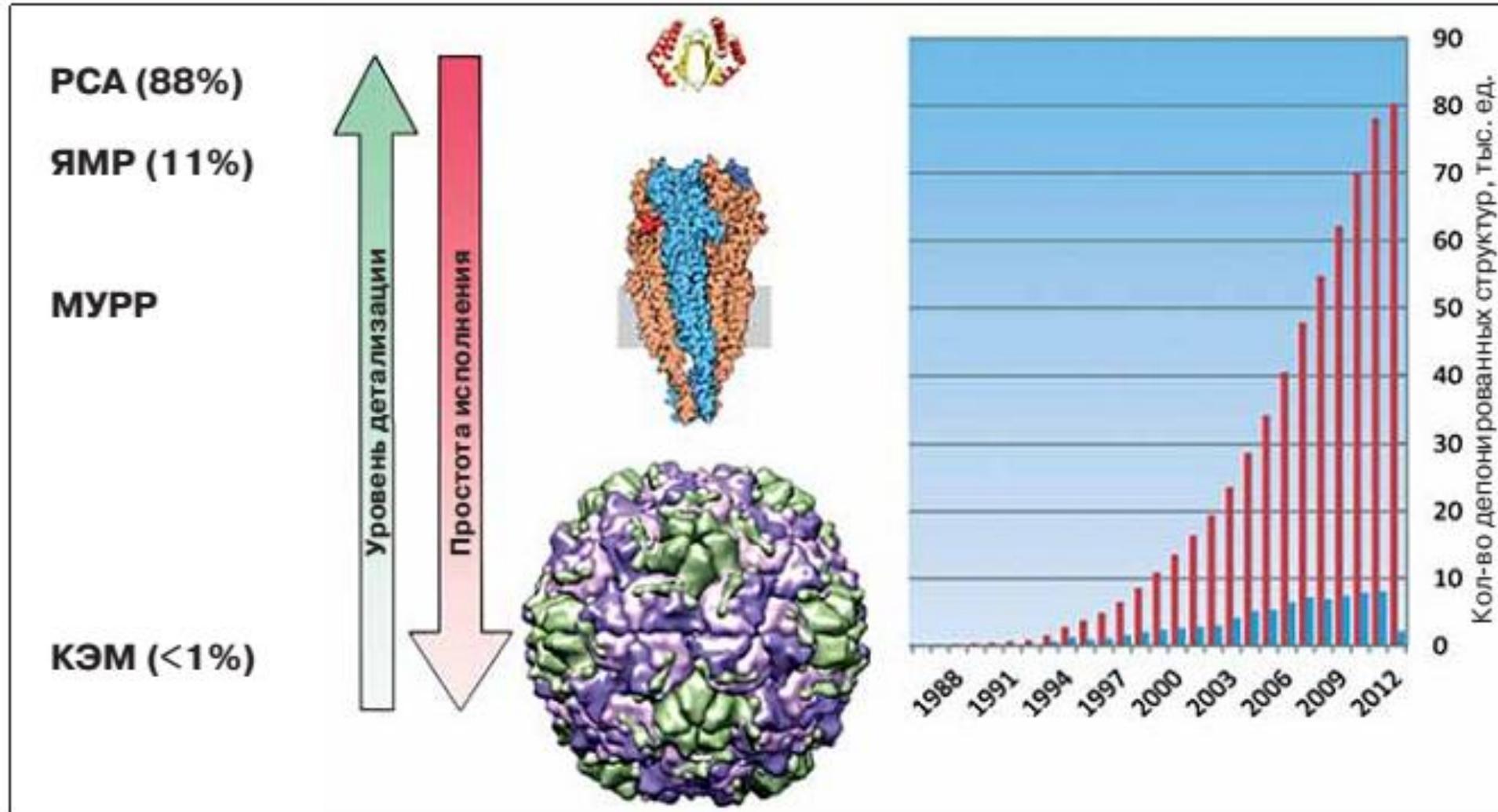
Расшифрованные последовательности белков:

- Reviewed (**569 213**) Swiss-Prot
- Unreviewed (**245 871 724**) TrEMBL

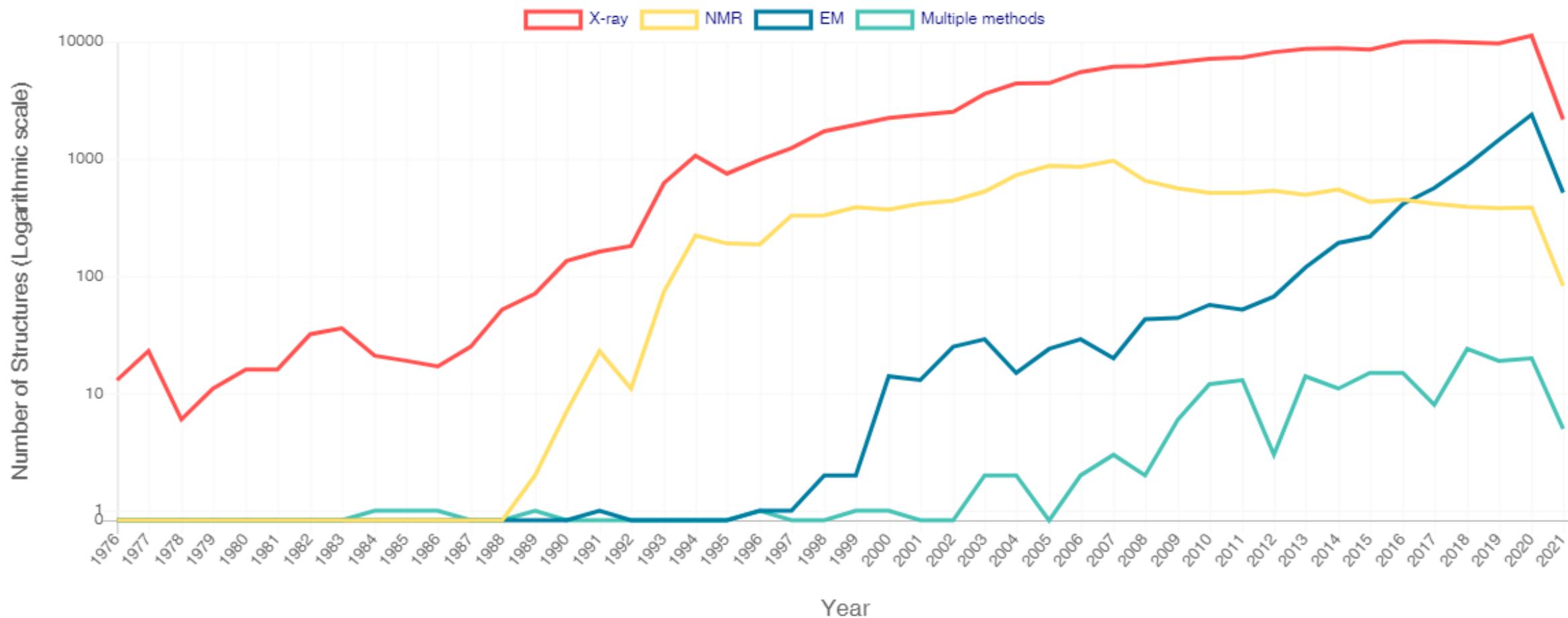
Расшифрованные структуры белков и их комплексов:

- 202292 структуры макромолекул
- 61535 структур человеческих белков
- 59395 разных белковых последовательностей
- 8101 разных человеческих белков
- 11369 структур ДНК, РНК и комплексов с ними

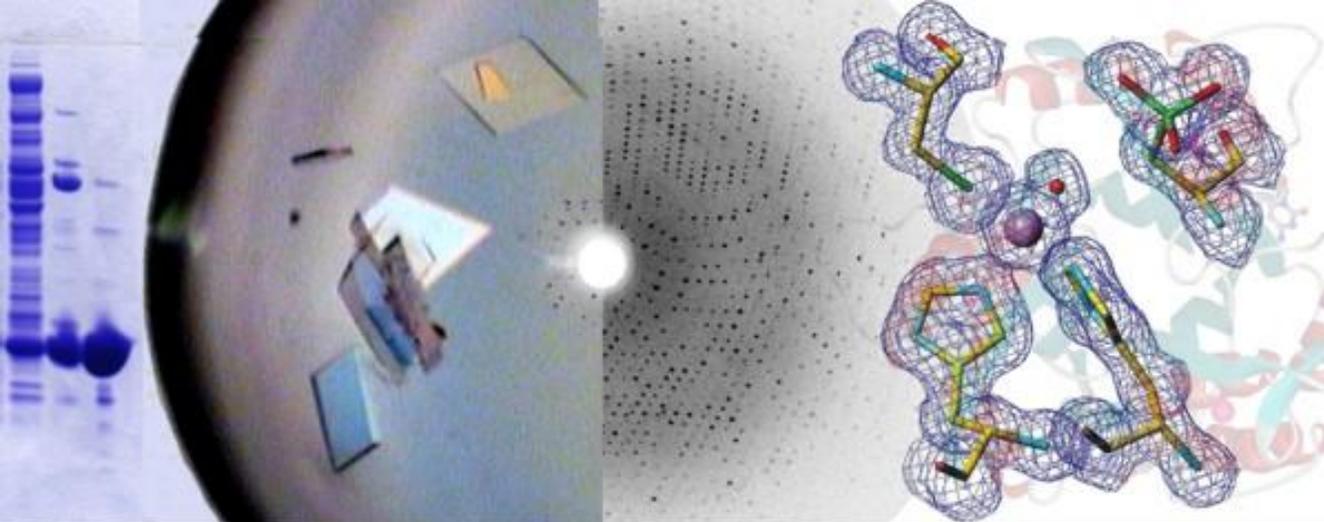
Данные о структурах биополимеров хранятся в RCSB PDB  
(Research Collaboratory for Structural Bioinformatics Protein Data Bank)



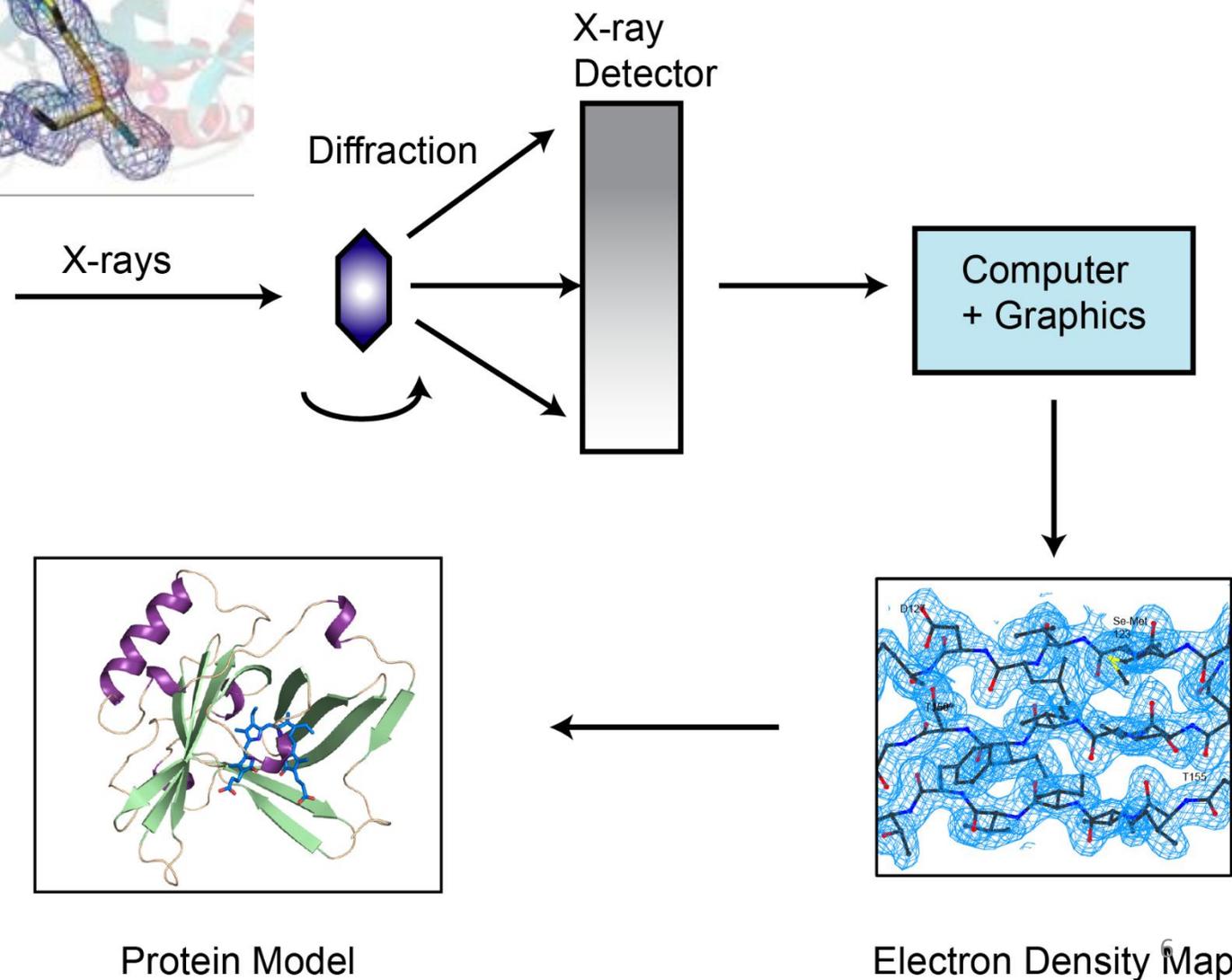
Основные методы исследования пространственных структур макромолекул  
и статистика роста числа структур в банке данных RCSB ([www.rcsb.org](http://www.rcsb.org)).



- X-ray: X-RAY DIFFRACTION, FIBER DIFFRACTION, or POWDER DIFFRACTION
- NMR: SOLUTION NMR or SOLID-STATE NMR
- EM: ELECTRON MICROSCOPY or ELECTRON CRYSTALLOGRAPHY or ELECTRON TOMOGRAPH.
- MULTIPLE METHODS: Multiple experimental methods. For example, if a structure is solved by X-RAY DIFFRACTION AND NEUTRON DIFFRACTION, it will be counted only in this category.



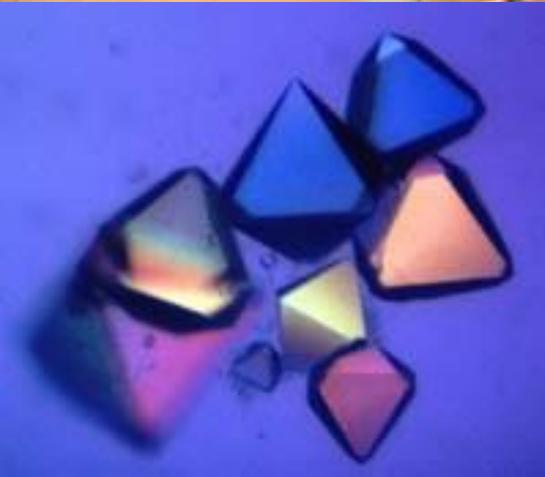
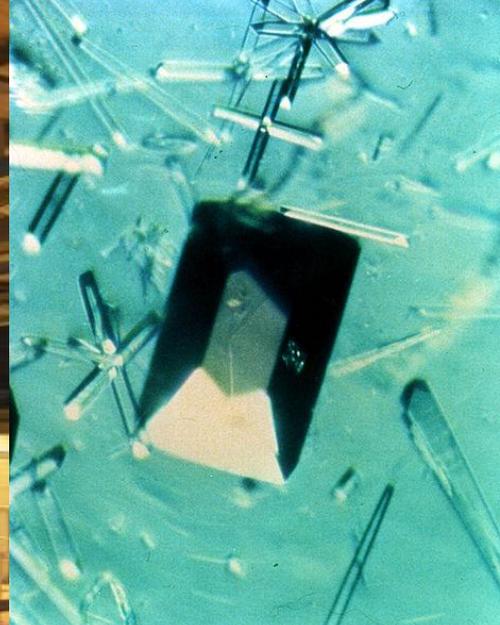
В основе **рентгеноструктурного анализа** лежит явление дифракции рентгеновских лучей на трёхмерной кристаллической решётке.



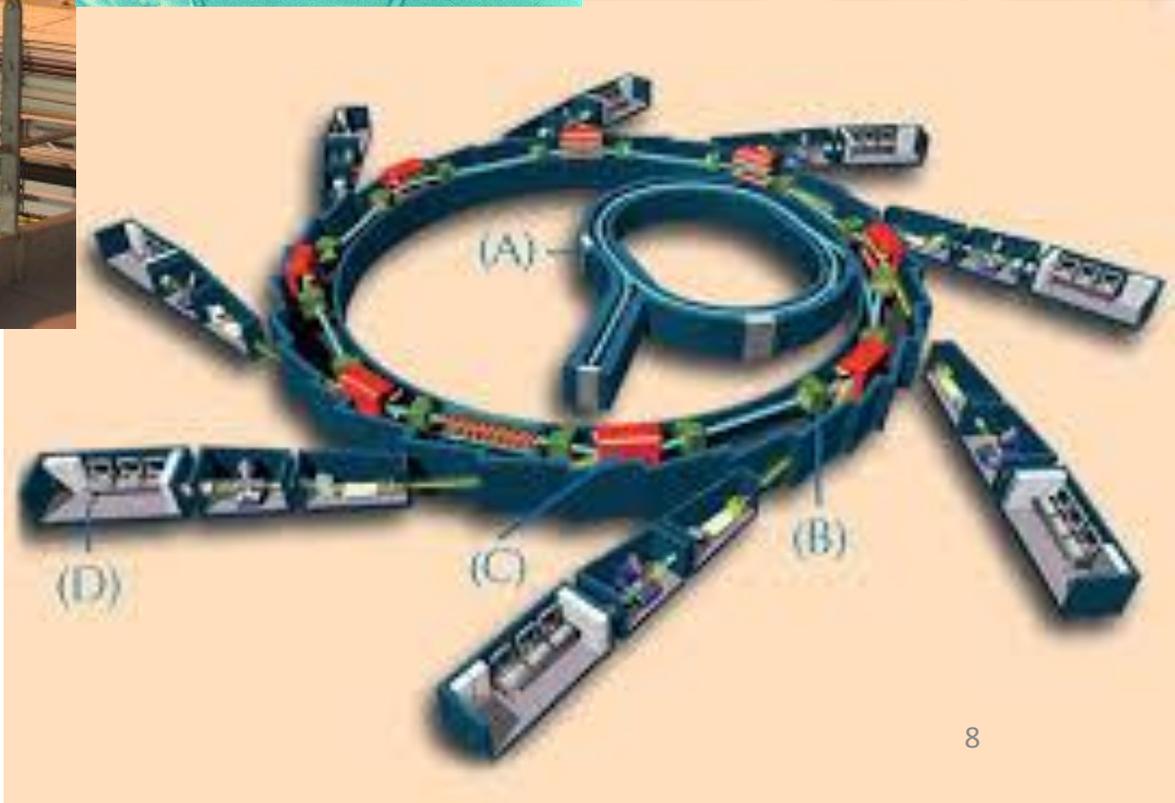


Рентгеноструктурный анализ проводится с использованием синхротронов. Синхротрон — это кольцевой ускоритель элементарных частиц. Длина ESRF, расположенного в Гренобле на слиянии двух рек, — 844 метра. Синхротроны используются для изучения широкого круга объектов, например, белков и их комплексов с нуклеиновыми кислотами, с помощью излучения, генерируемого при участии элементарных частиц.

<https://en.wikipedia.org/wiki/Synchrotron><sup>7</sup>



Для проведения рентгеноструктурного анализа белки предварительно нужно кристаллизовать. Кристалл белка помещают в рентгеновский пучок и получают дифракционную картину для последующего анализа.

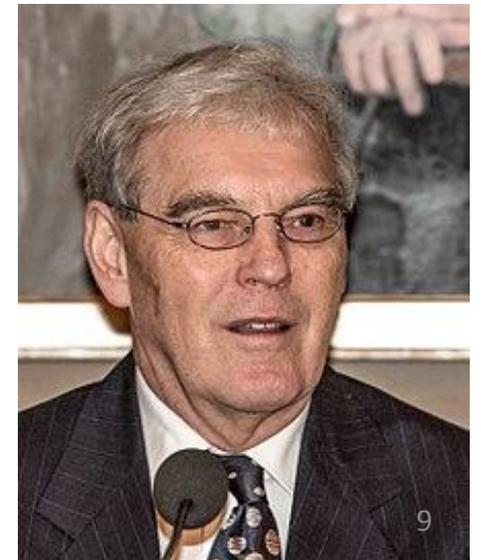
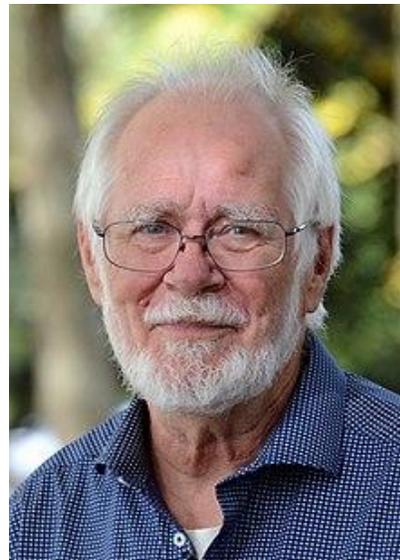


# Криоэлектронная микроскопия



Крио-ЭМ - метод электронной микроскопии, применяемый на образцах, охлажденных до криогенных температур и погруженных в среду стекловидной воды. Водный раствор образца наносят на специальную сетку и погружают в жидкий этан. Последние достижения в области детекторной техники и программных алгоритмов позволили определить биомолекулярные структуры с почти атомным разрешением.

В 2017 году Нобелевская премия по химии была присуждена Жаку Дюбоше, Иоахиму Франку и Ричарду Хендерсону «за разработку криоэлектронной микроскопии для определения структуры биомолекул в растворе с высоким разрешением».



Просвечивающий электронный микроскоп

## DUBOCHET'S VITRIFICATION METHOD

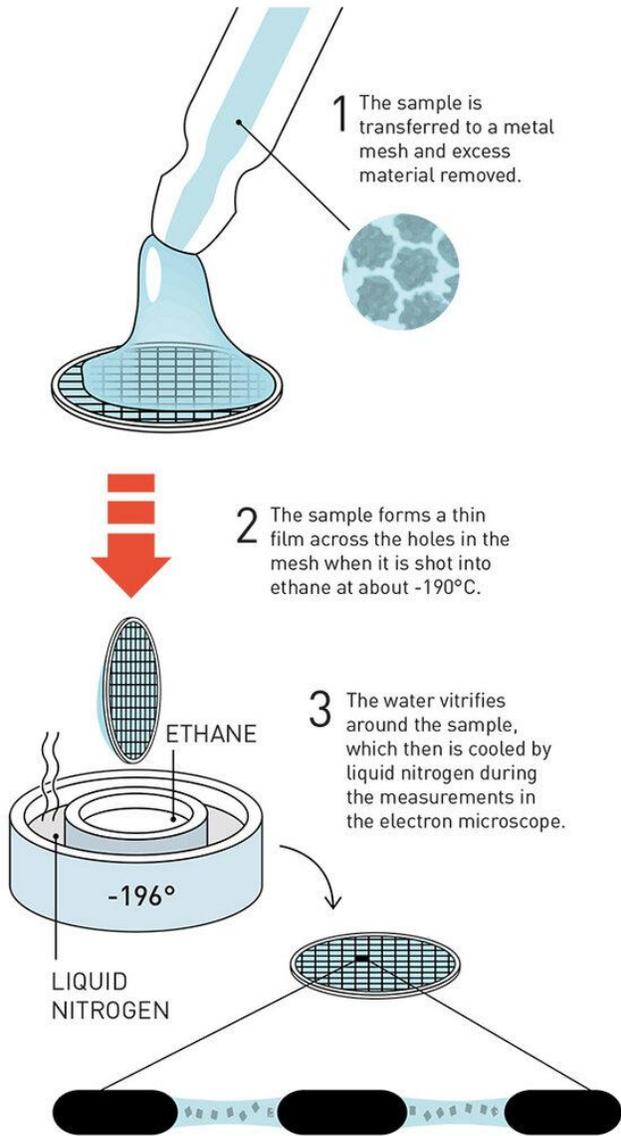


Illustration: ©Johan Jarnestad/The Royal Swedish Academy of Sciences

## FRANK'S IMAGE ANALYSIS FOR 3D STRUCTURES

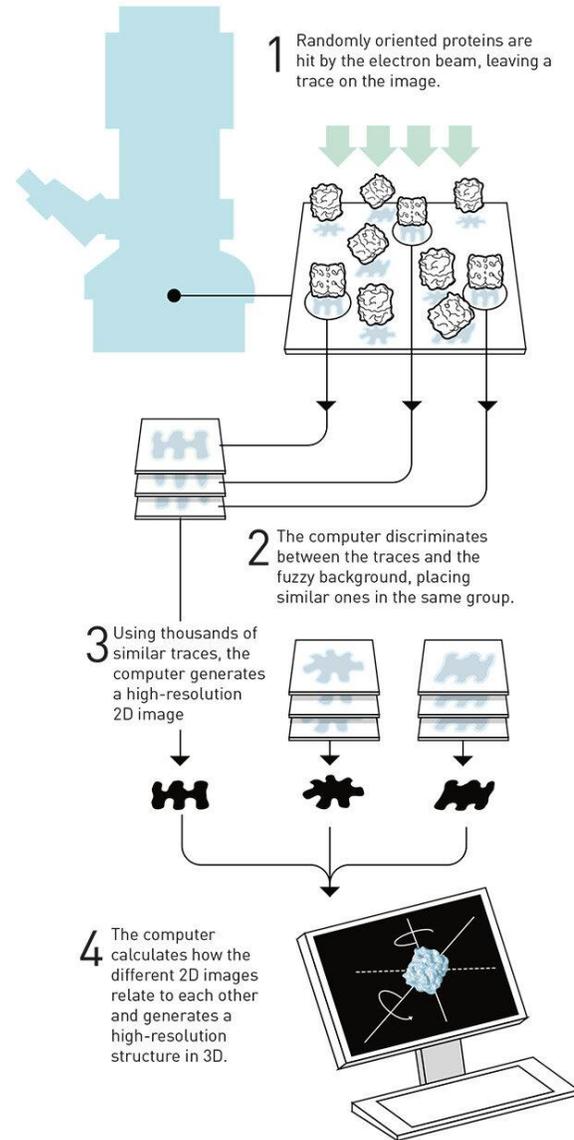
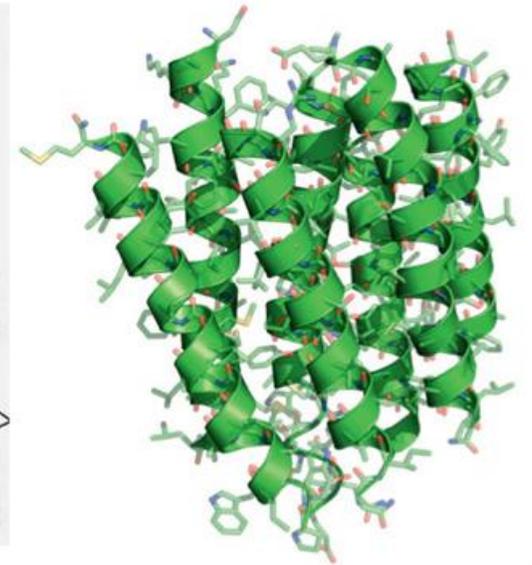
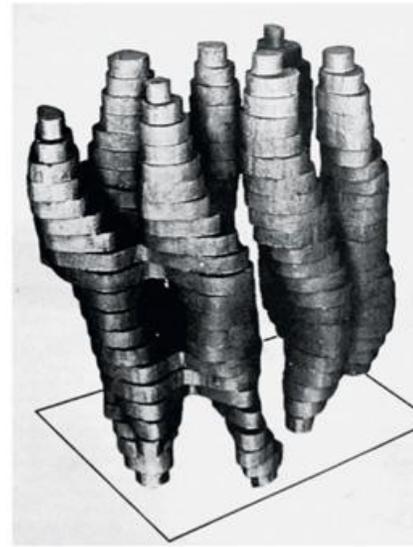
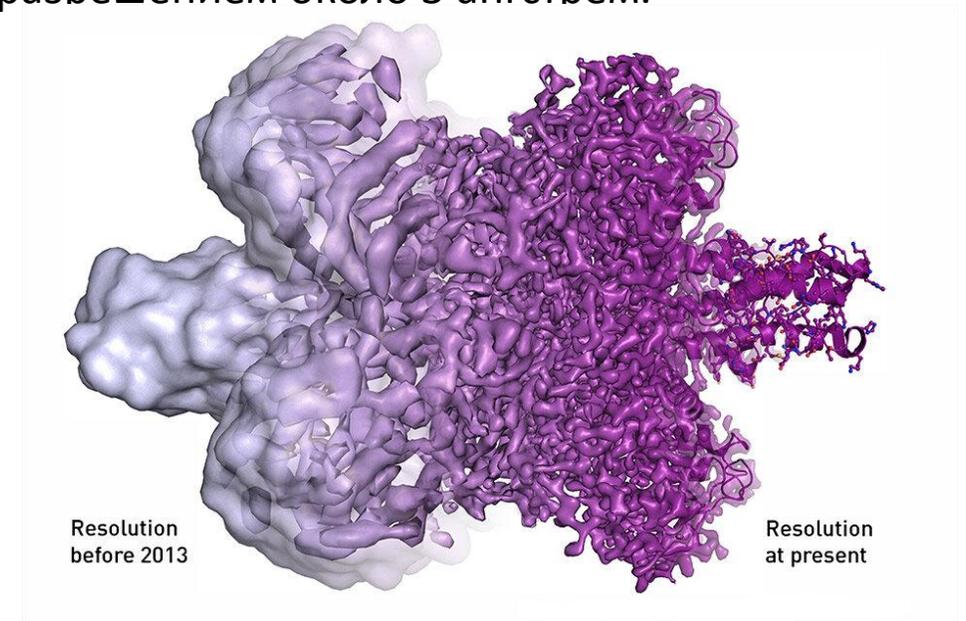


Illustration: ©Johan Jarnestad/  
The Royal Swedish Academy of Sciences



В 1990 году Ричард Хендерсон после 15 лет работы получил структуру бактериородопсина с разрешением около 3 ангстрем.





Molecular explorations  
through biology and medicine

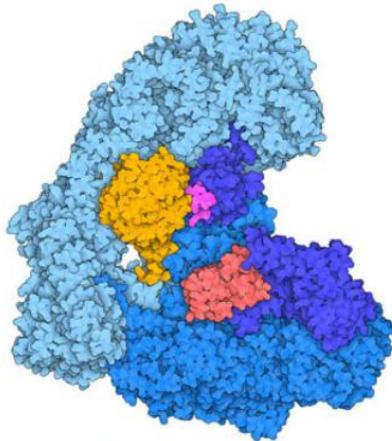


Training and outreach portal of PROTEIN DATA BANK



## Molecule of the Month

March 2023



### Anaphase-Promoting Complex / Cyclosome

APC/C guards the checkpoints that regulate key steps in the cell cycle

[More](#)

3D View: 5a31

Style

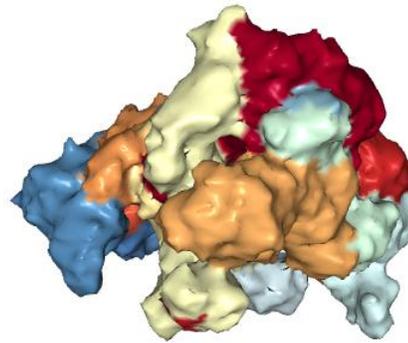
- Cartoon  
 Spheres  
 Surface

Color

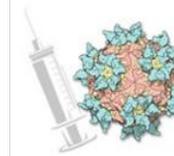
- Rainbow  
 Chain  
 Structure

Spin

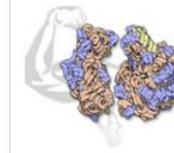
- On  
 Off



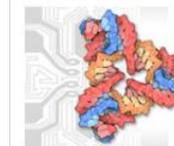
## Browse resources by category >



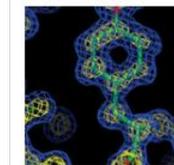
▾ Health and Disease



▾ Molecules of Life



▾ Biotech and Nanotech



▾ Structures and Structure Determination

Contact Us

All articles: [By Date](#) | [By Category](#) | [By Title](#)

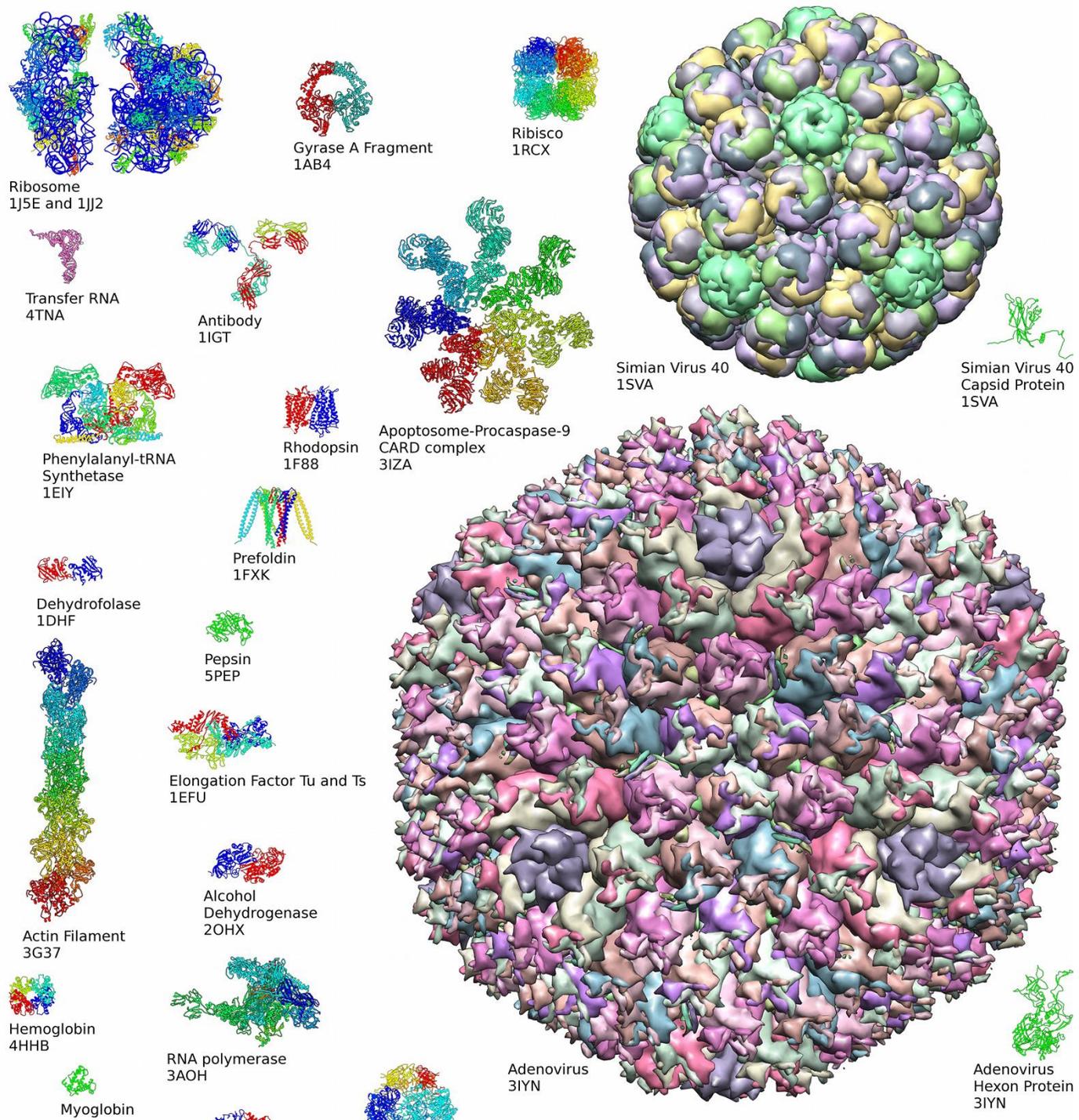
## Health Focus: Cancer Biology and Therapeutics

Cells in our bodies are carefully regulated so that they divide, grow, and die according to the best plan for keeping us alive. When they lose these controls, they can grow into deadly cancers. Atomic structures have revealed how cell

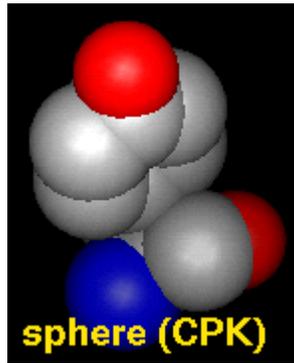
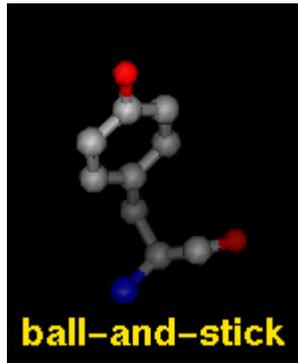
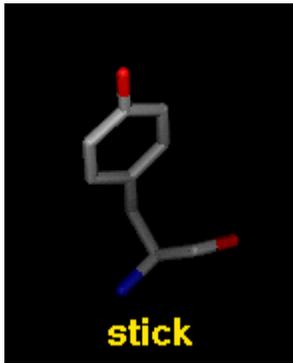
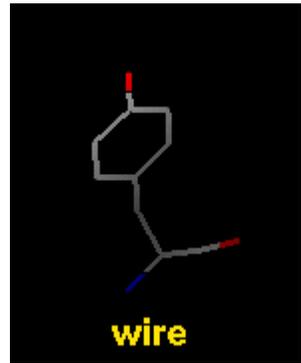
## News and Events



[Brain Awareness Week](#)

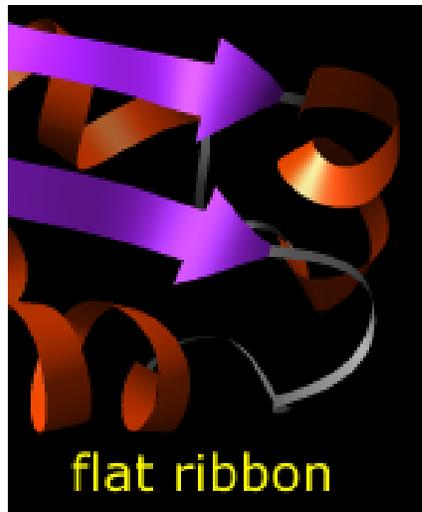


# Различные способы представления пространственных структур



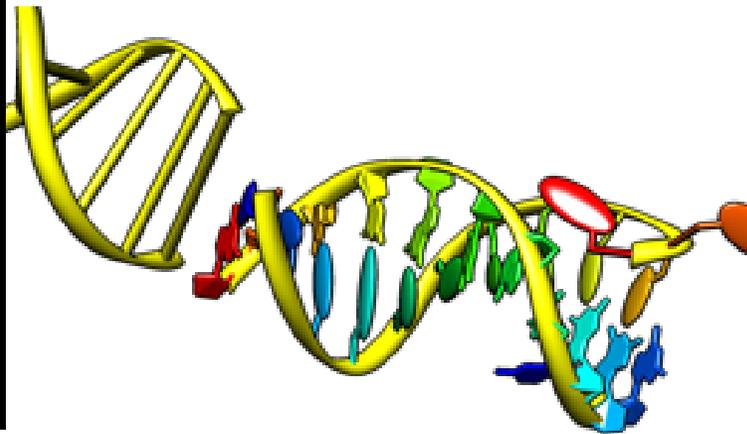
- **wire** (проволочная) - a simple line drawing with dot atoms and wire bonds
- **Stick** (палочковая) - “endcap” atoms and stick bonds
- **ball-and-stick** (шарико-палочковая) - ball atoms (atomic [VDW radius](#) × [ball scale](#)) and stick bonds
- **Sphere** (сферическая) - sphere atoms (full [VDW radius](#)) and wire bonds; also called Corey-Pauling-Koltun (**CPK**)

Cartoon

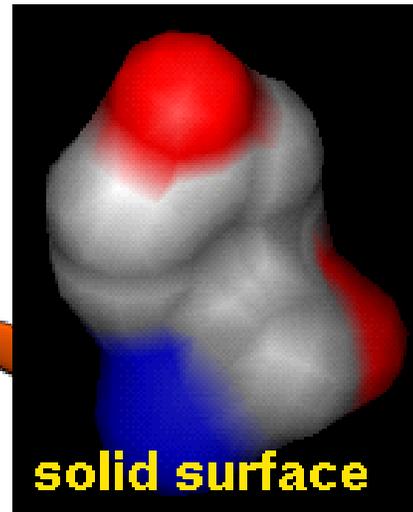


Ленточные модели

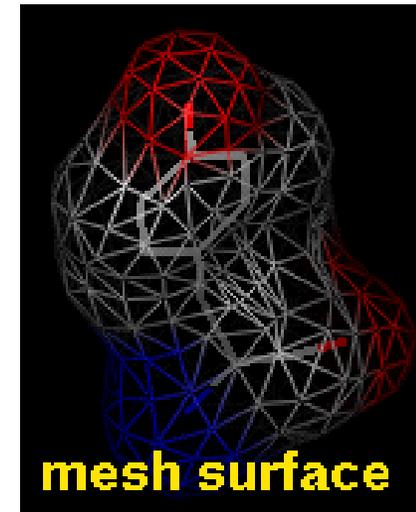
Нуклеиновые кислоты



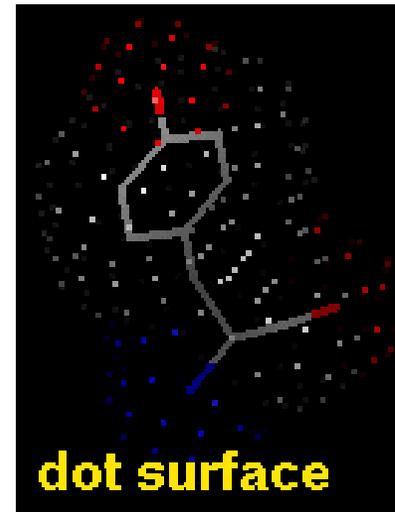
Молекулярные поверхности



твердая

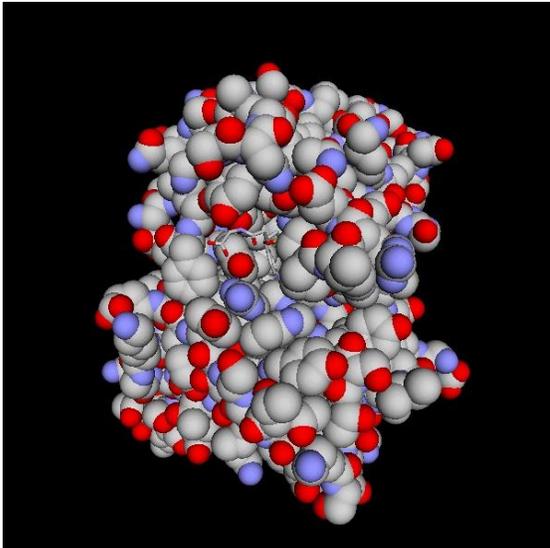


ячеистая



точечная

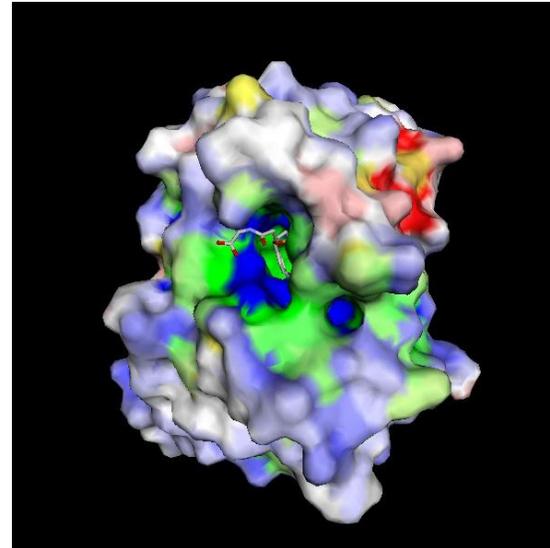
# Различные способы визуализации трехмерной структуры макромолекулы



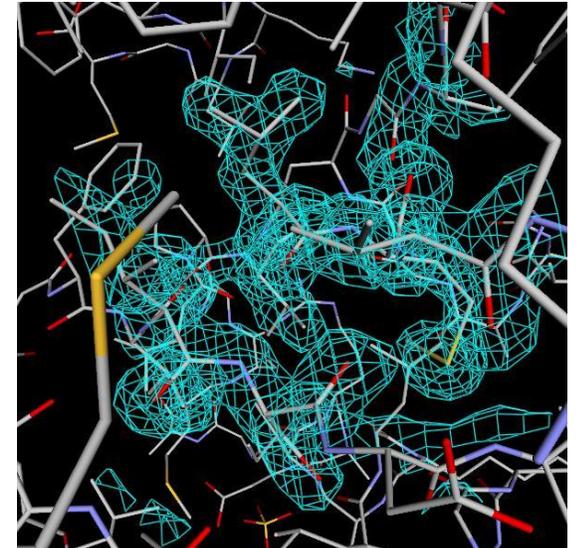
Spacefill expression



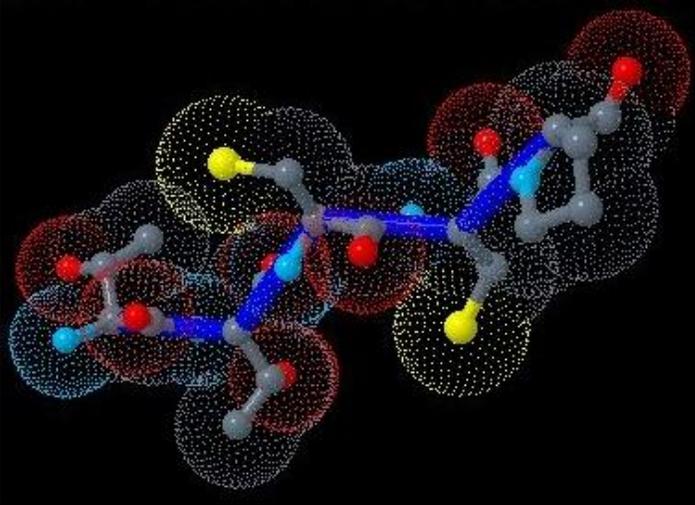
Cartoon expression



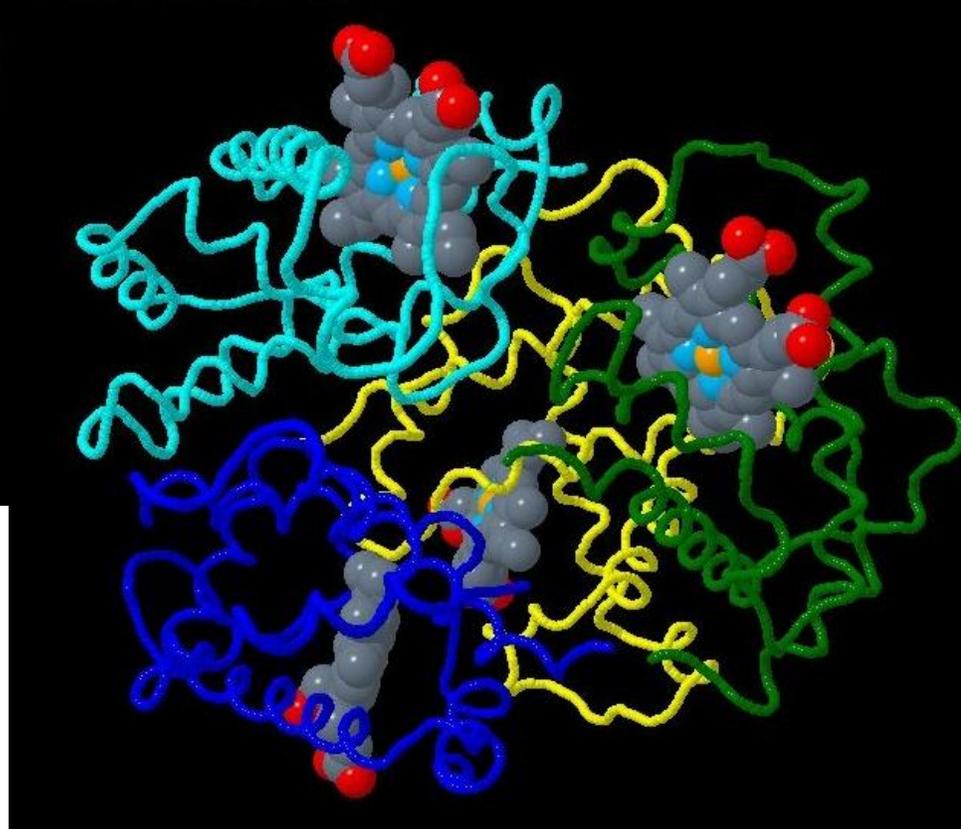
Molecular surface



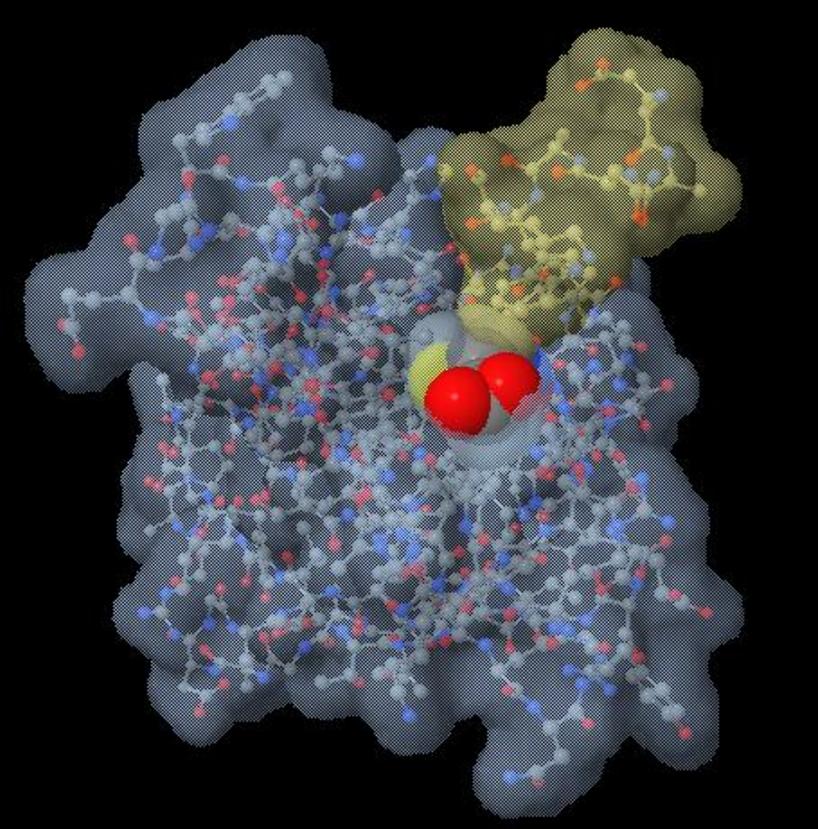
Electron density



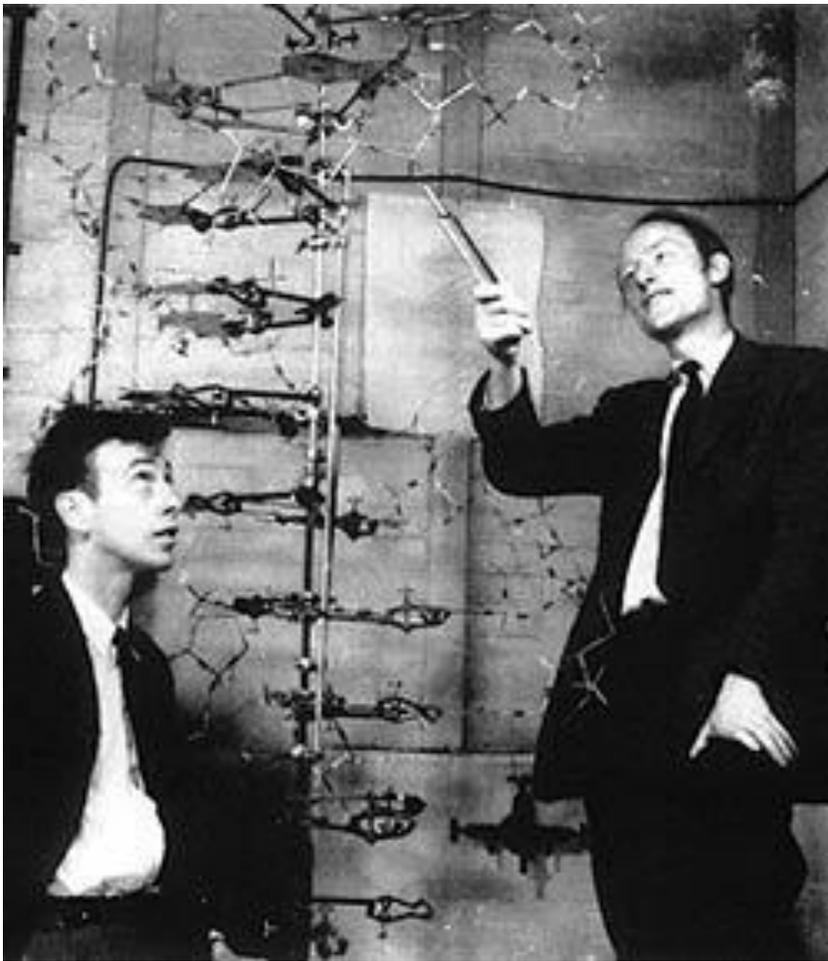
Участок белковой цепи



Гемоглобин. Окраска по цепям, гемы изображены в spacefilled mode.



Биотин (srk color, spacefilled) внутри участок связывания стрептавидина (cyan translucent surface) и фрагмент стрептавидина (yellow translucent surface).

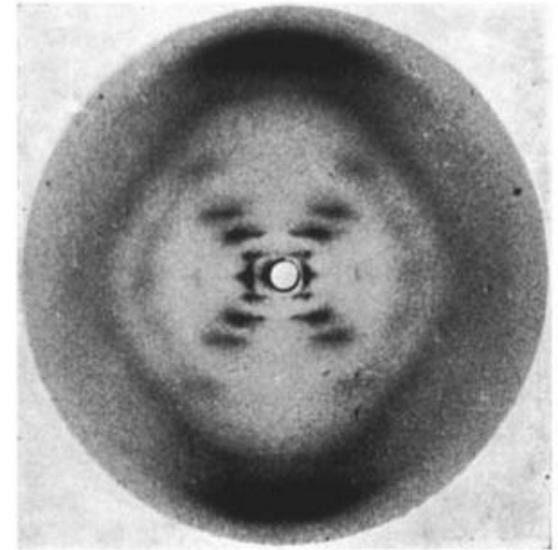


(Gonick & Whellis 1991)

### ERWIN CHARGAFF FOUND:

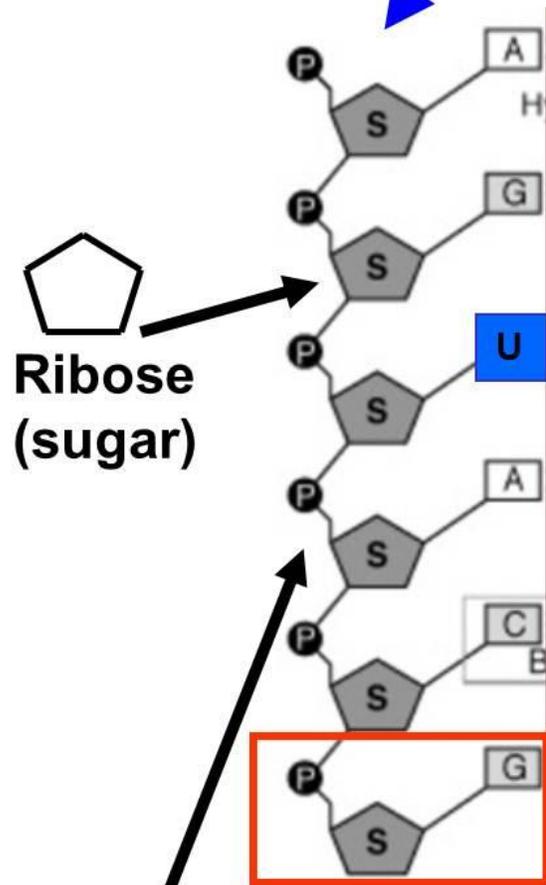
- ① THE COMPOSITION OF DNA VARIED FROM ONE SPECIES TO ANOTHER, IN PARTICULAR IN THE RELATIVE AMOUNTS OF THE BASES A, C, T, G.
- ② IN ANY DNA, THE NUMBER OF A'S WAS THE SAME AS THE NUMBER OF T'S; SIMILARLY, THE NUMBER OF C'S WAS EQUAL TO THE NUMBER OF G'S.

WHAT DID THIS MEAN?  
CHARGAFF COULDN'T SAY..



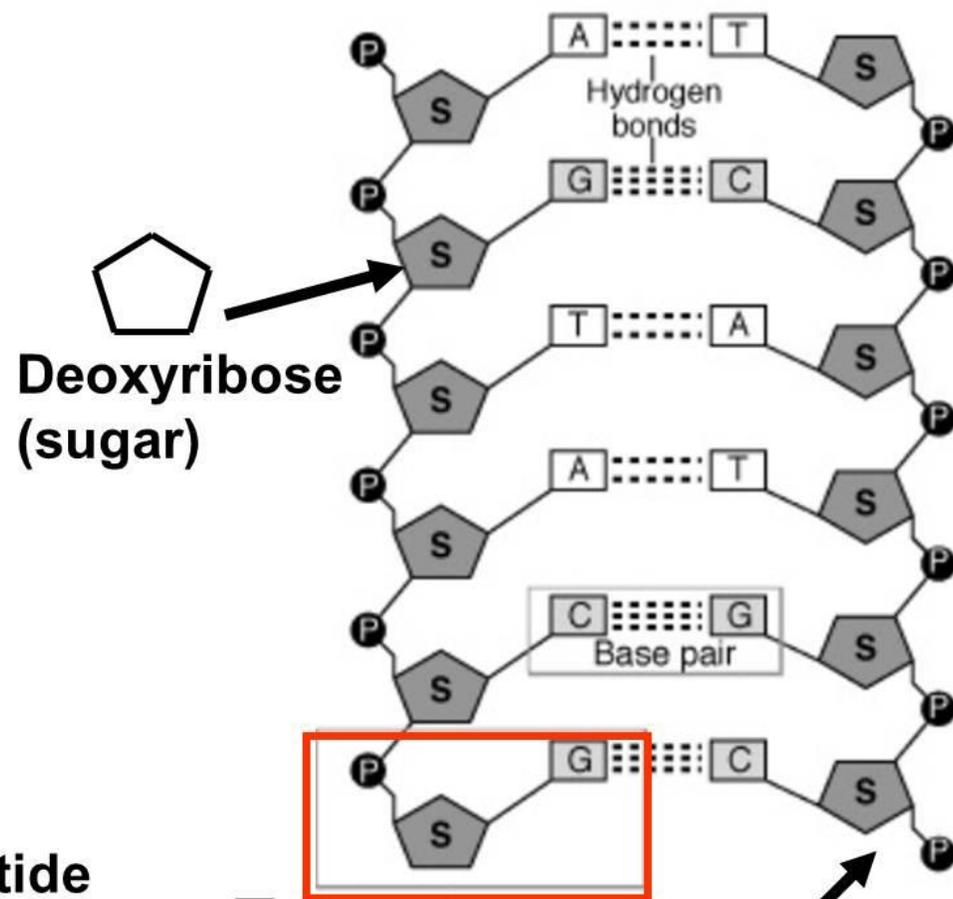
В 1952 году Джеймс Уотсон и Фрэнсис Крик стали работать над моделированием структуры ДНК. При этом они использовали шариковую модель. Используя правила Чаргаффа и рентгенограммы Розалинды Франклин и Мориса Уилкинса они построили двухспиральную модель ДНК. Результаты работы опубликовали 30 мая 1953 года в журнале Nature, а в 1962 году за эту работу им была присуждена Нобелевская премия.

RNA & DNA  
are chains of nucleotides



Ribose  
(sugar)

Sugar-phosphate  
backbone

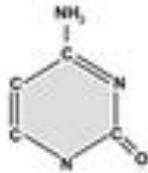


Deoxyribose  
(sugar)

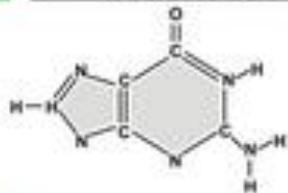
Nucleotide  
Unit

Sugar-phosphate  
backbone

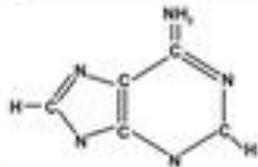
**Cytosine**



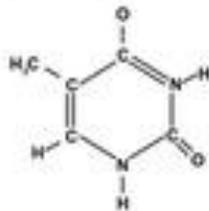
**Guanine**



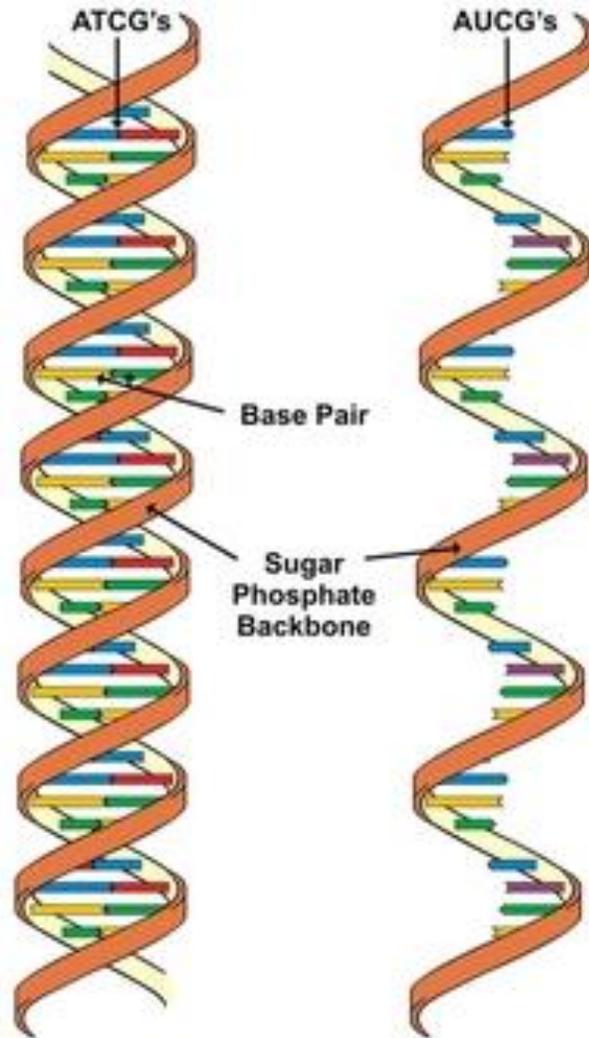
**Adenine**



**Thymine**



Nitrogenous  
Bases



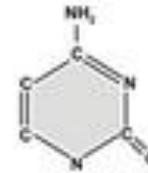
**DNA**

Deoxyribonucleic Acid

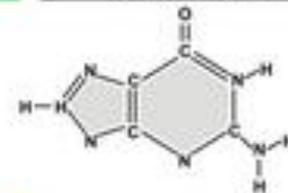
**RNA**

Ribonucleic Acid

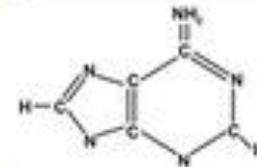
**Cytosine**



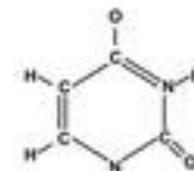
**Guanine**



**Adenine**

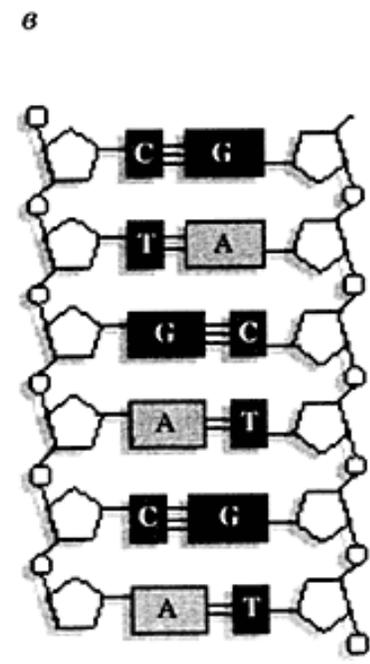
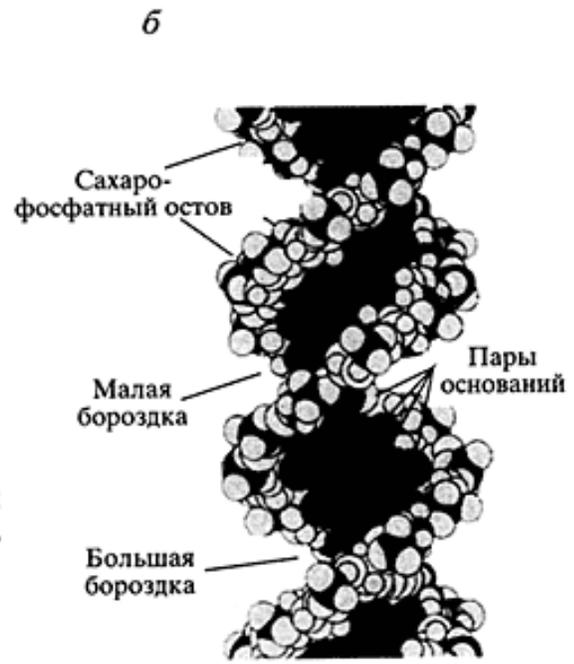
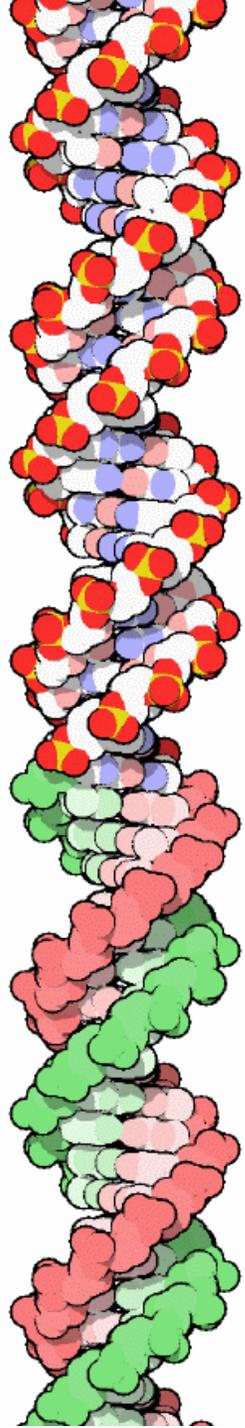


**Uracil**

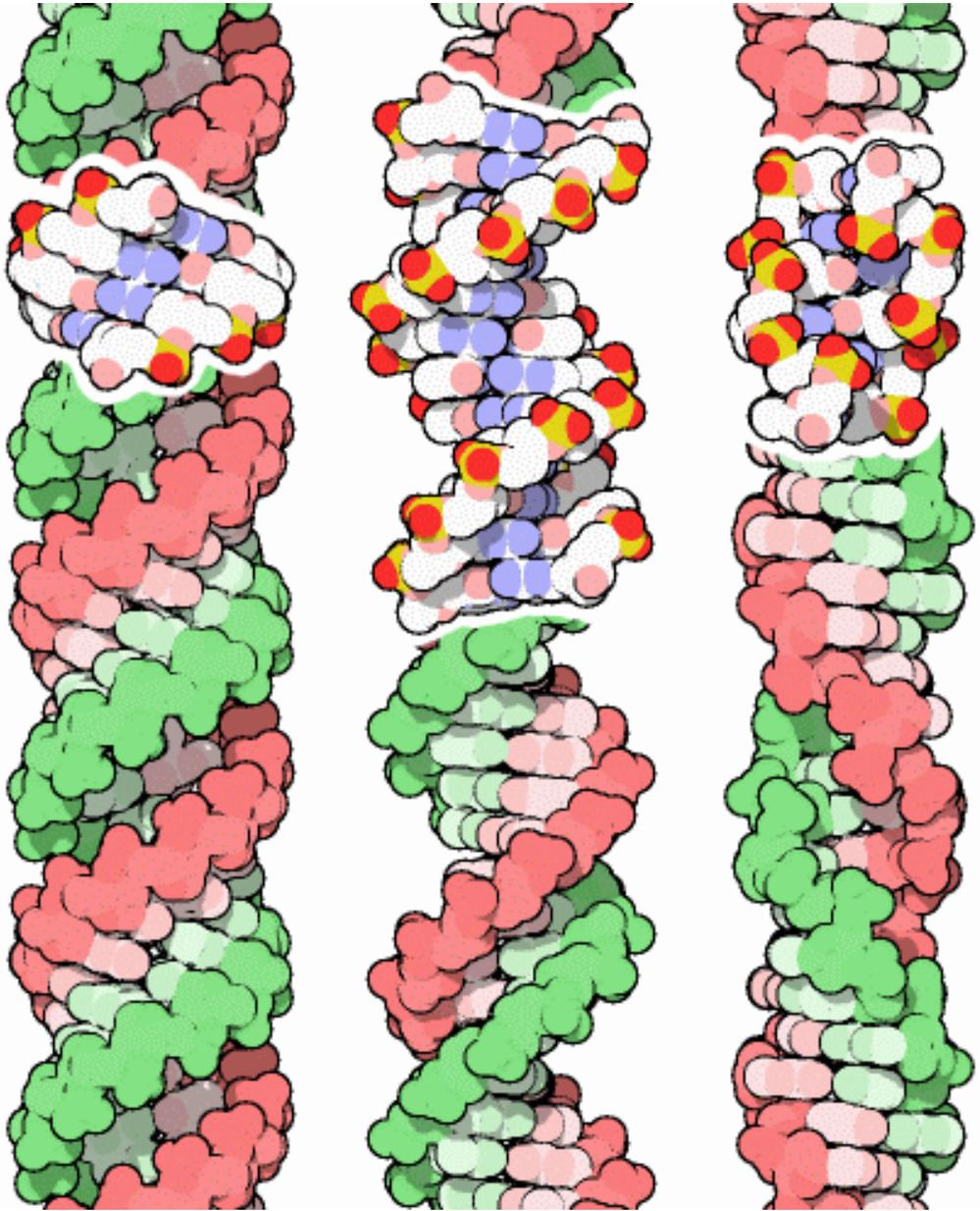


Replaces Thymine in RNA

Nitrogenous  
Bases



# Структура ДНК

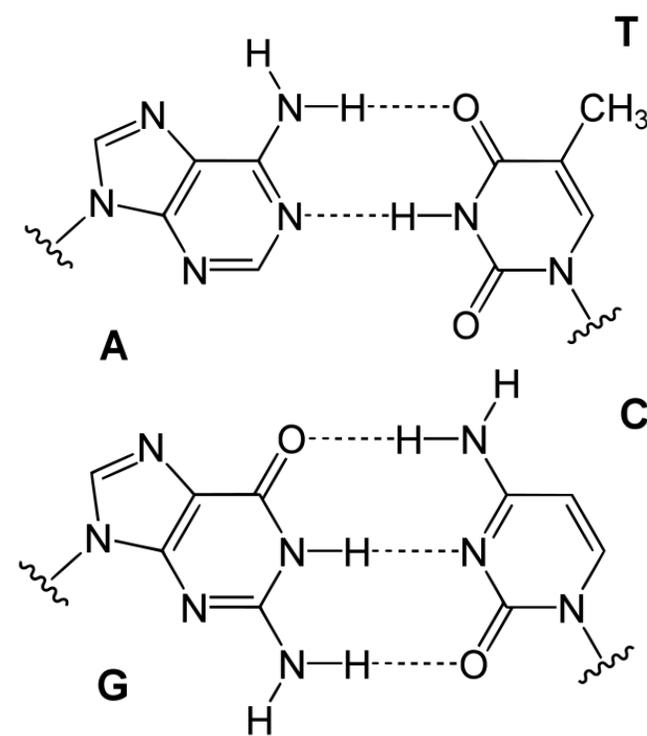
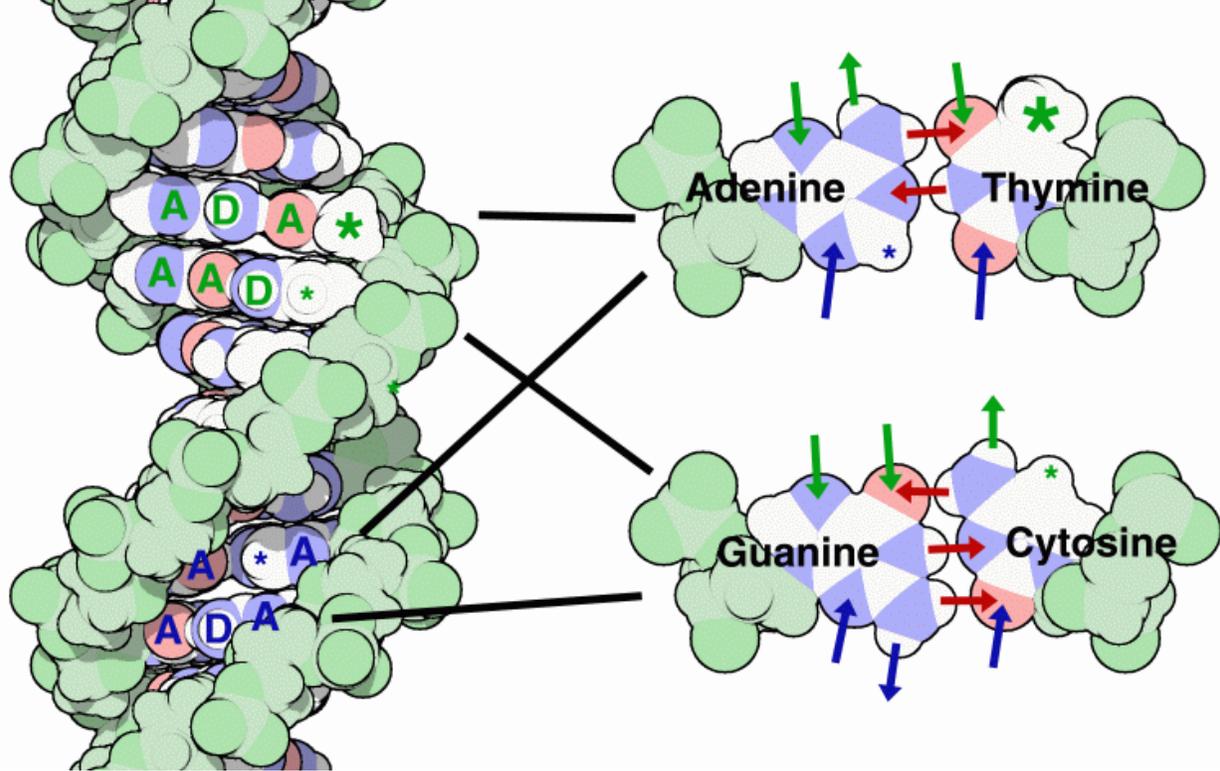


A

B

Z

В обычных условиях ДНК существует в виде В-спирали (в центре). Слева – А форма ДНК с наклоненными стопками оснований и глубокой большой бороздкой, которая получается в условиях дегидратации. Такая форма характерна для спиральных участков РНК. Справа – Z-форма, закрученная в противоположном направлении, она получается при высокой концентрации соли и требует определенных последовательностей с чередованием G-C и C-G пар.



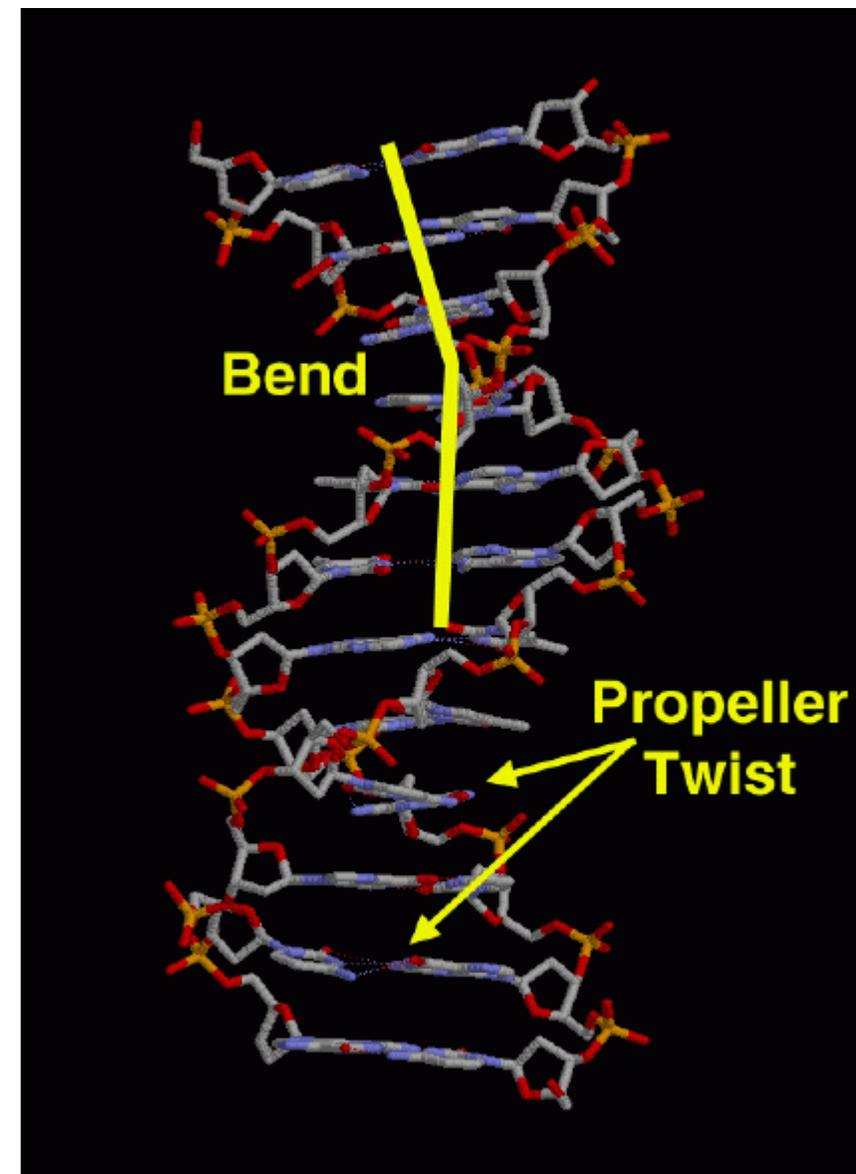
### Информация, хранимая в ДНК. Как ее считать?

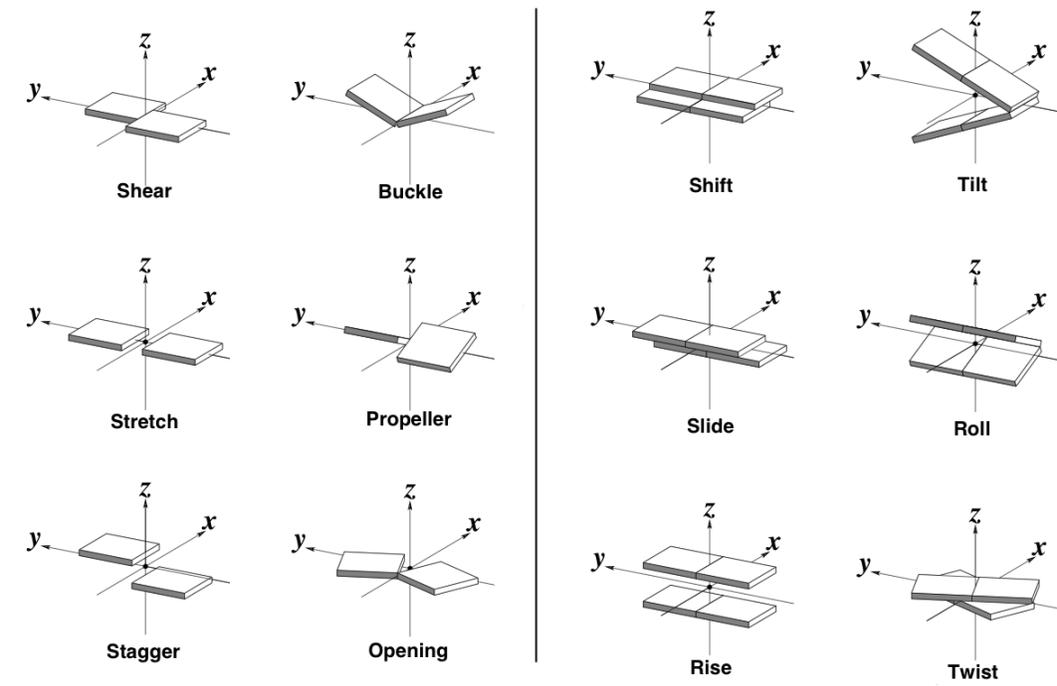
Механизм хранения генетической информации в ДНК и путь ее передачи в живой природе были открыты в 1950-х годах Уотсоном и Криком. Основная информация хранится в виде последовательности нуклеотидов, причем нуклеотиды одной цепи ДНК строго соответствуют нуклеотидам противоположной цепи, образуя комплементарные пары А-Т и G-С. В комплементарных парах нуклеотиды образуют между собой водородные связи (красные стрелки слева, пунктирные линии справа). Эта информация непосредственно считывается и воспроизводится в процессе репликации ДНК и транскрипции.

В процессе жизнедеятельности клетки с ДНК взаимодействует большое количество разнообразных белков, выполняющих структурные и регуляторные функции. Эти белки узнают ДНК, не раскручивая двойную спираль. Они считывают информацию с поверхности ДНК. В большую и малую бороздки ДНК обращены боковые стороны азотистых оснований, химические группы которых способны также быть донорами или акцепторами водородных связей, участвовать в образовании гидрофобных контактов. Белки также образуют контакты с фосфатами сахарофосфатного остова.

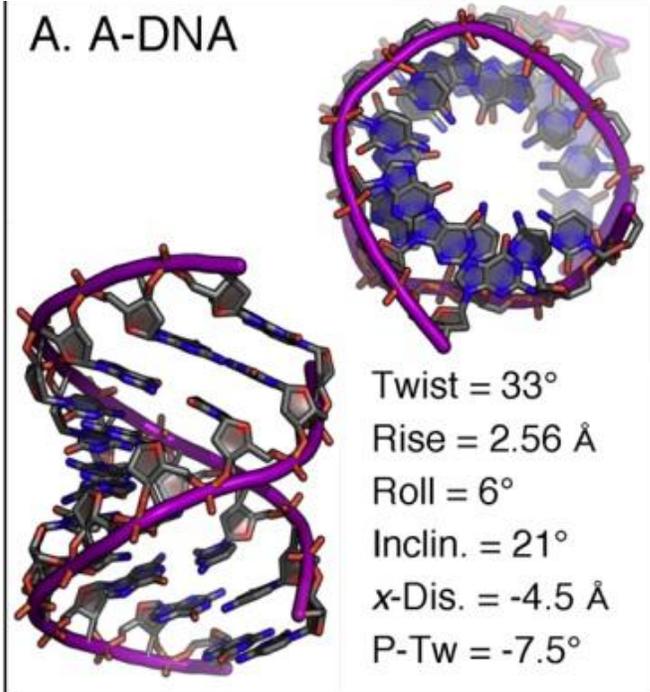
## Реальная структура ДНК

Молекулы ДНК, расшифрованные рентгеноструктурным анализом, как правило, не представляют собой совершенной двойной спирали, а представлены целым набором локальных конформаций. Данная ДНК – из PDB файла 1bna – в верхней части изогнута. В нижней части две пары оснований не лежат в одной плоскости, а образуют конформацию “propeller twist”. ДНК – динамичная молекула, претерпевающая различные изменения структуры в зависимости от условий окружающей среды и взаимодействующих с ней молекул.

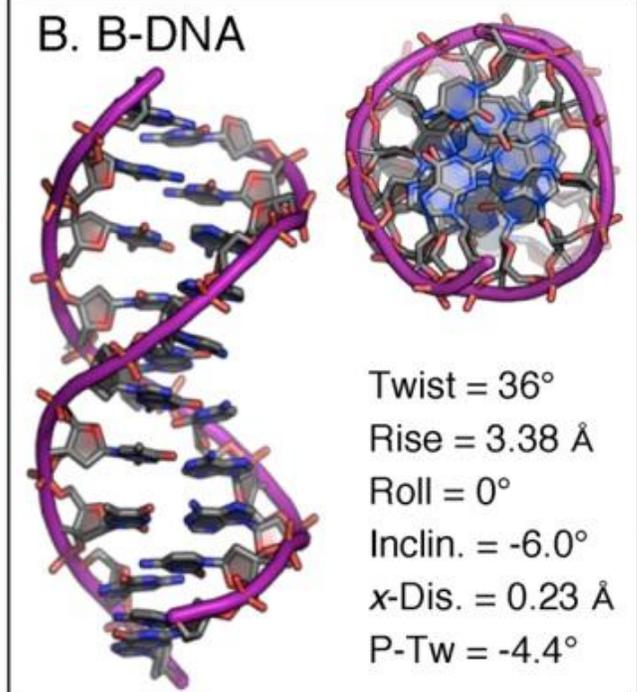




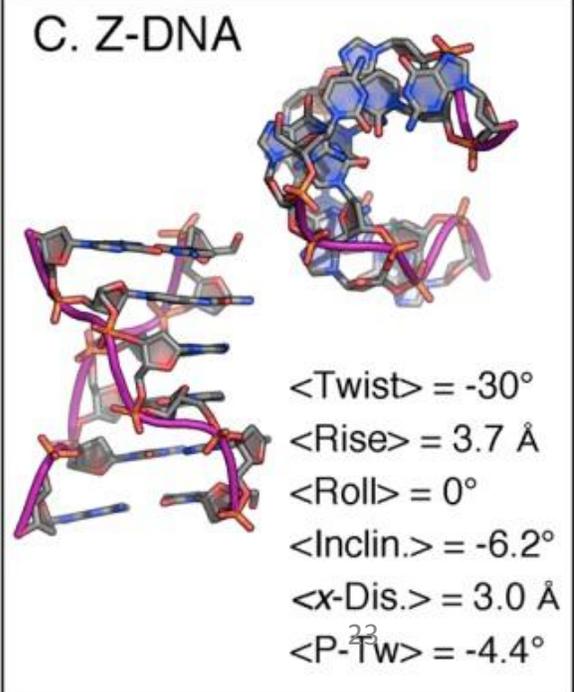
### A. A-DNA



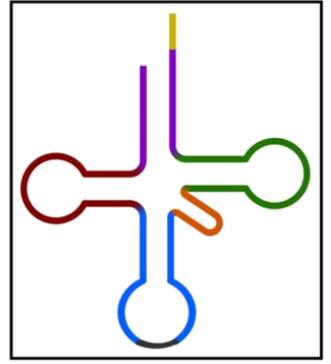
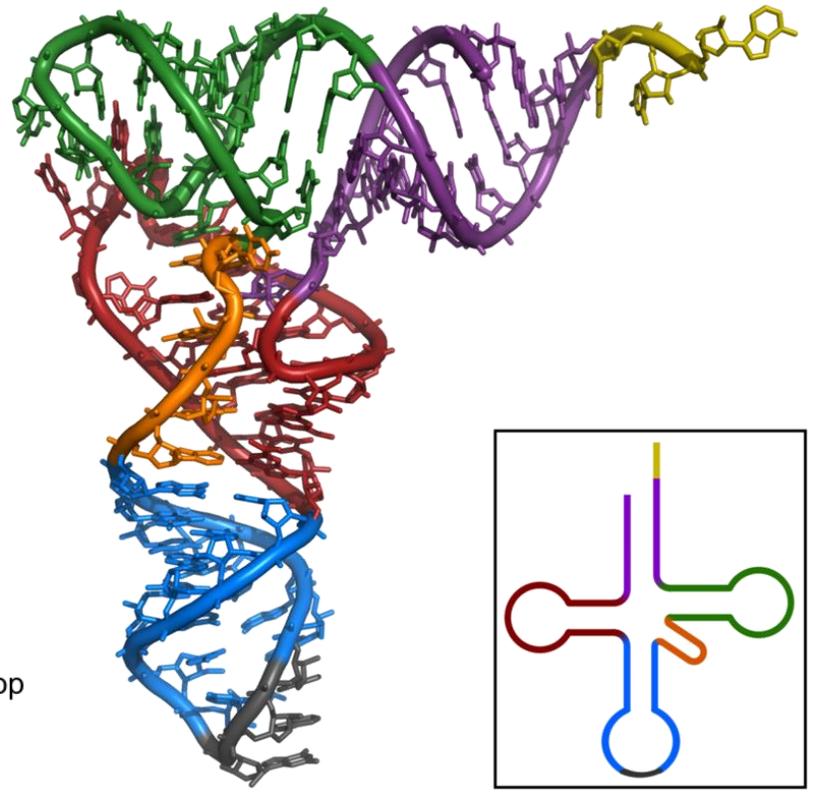
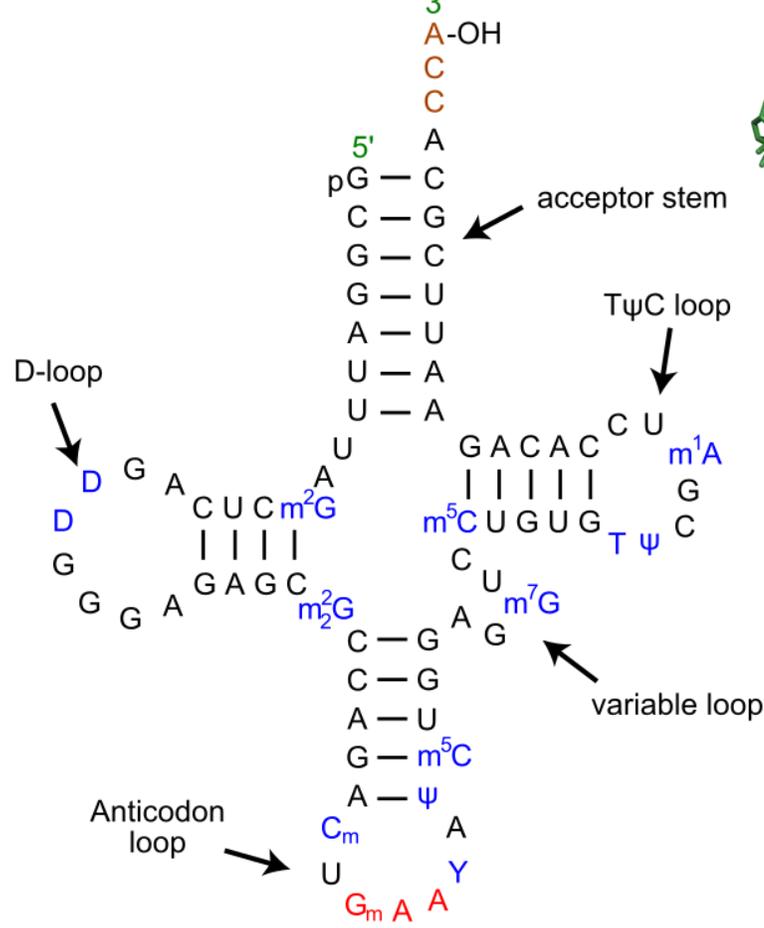
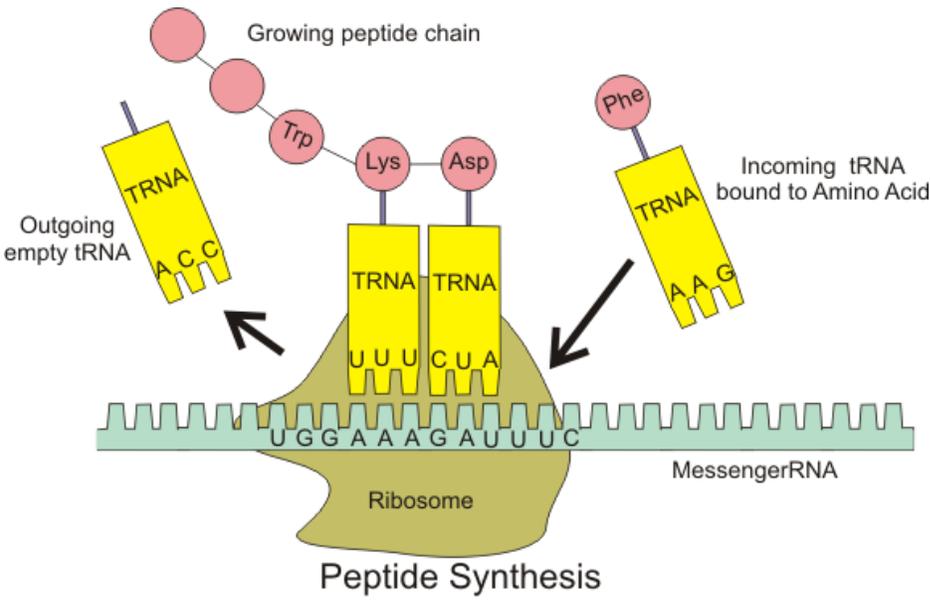
### B. B-DNA



### C. Z-DNA



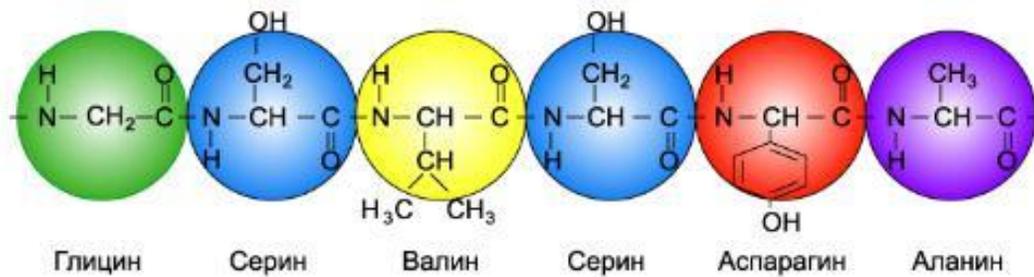
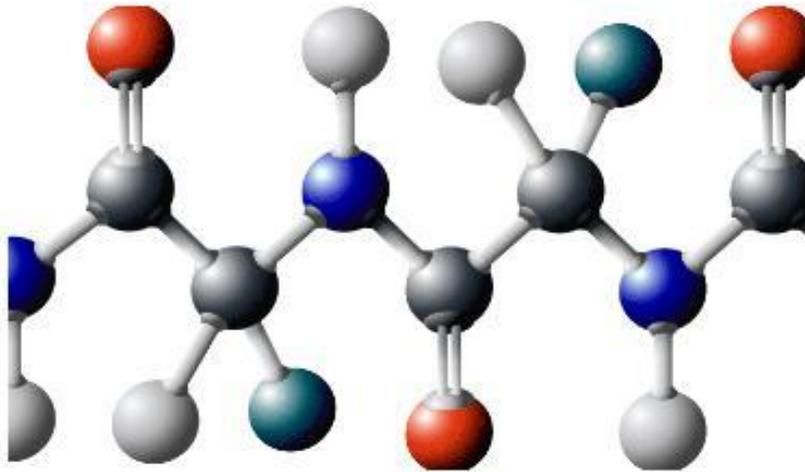
# Структура РНК



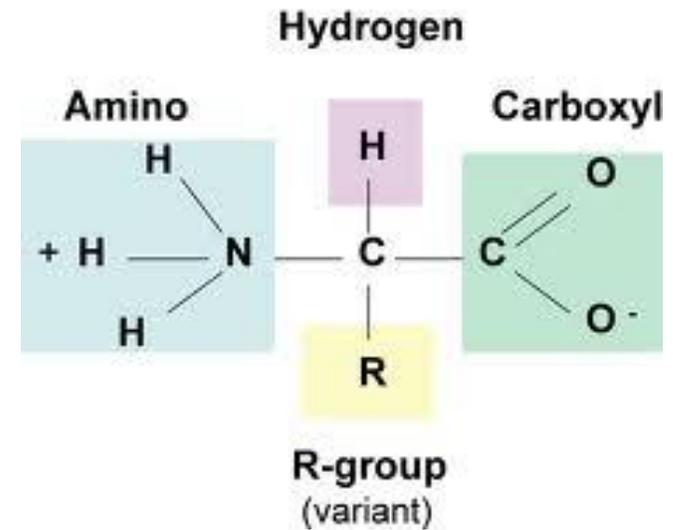
**тРНК** является одноцепочечной РНК, однако в функциональной форме имеет конформацию «клеверного листа». В ней выделяют 4 части, выполняющие различные функции. Первая часть-акцепторный "стебель", образованный 2-мя комплементарно соединёнными концевыми частями. Он состоит из 7 пар оснований. 3'-конец этого стебля несколько длиннее. Он формирует одноцепочечный участок, который заканчивается последовательностью ССА со свободной ОН-группой. К этому концу присоединяется транспортируемая аминокислота. 3 остальные части представляют собой комплементарно спаренные последовательности нуклеотидов, которые заканчиваются неспаренными участками, образующими петли. Средняя часть состоит из 5 пар нуклеотидов и содержит в центре своей петли антикодон. Аминокислота ковалентно присоединяется к 3'-концу молекулы с помощью специфичного для каждого типа тРНК фермента [аминоацил-тРНК-синтетазы](#). На участке С находится антикодон, соответствующий аминокислоте. тРНК содержат дигидроуридин, псевдоуридин, 7-метилгуанин, не Уотсон-Криковские пары.

# СТРУКТУРА БЕЛКА

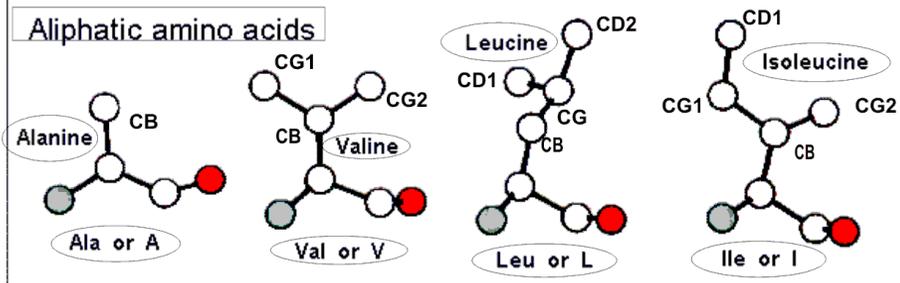
Первичная структура белка – это полипептидная цепочка из аминокислот.



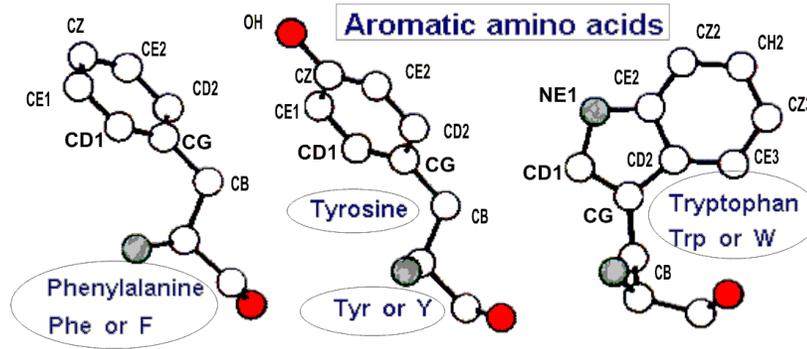
## Amino Acid Structure



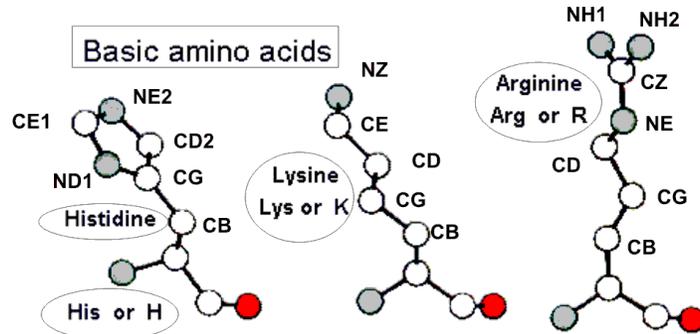
Aliphatic amino acids



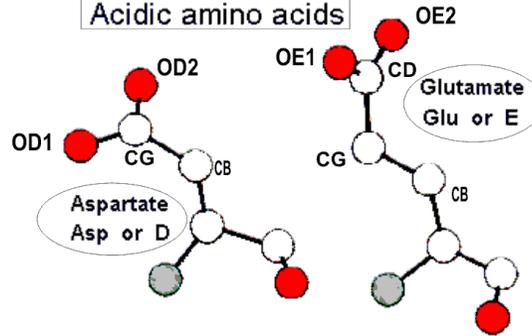
Aromatic amino acids



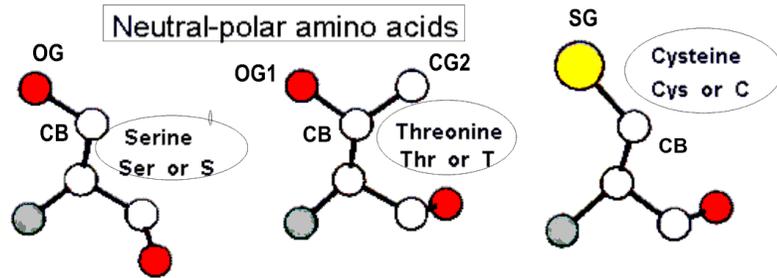
Basic amino acids



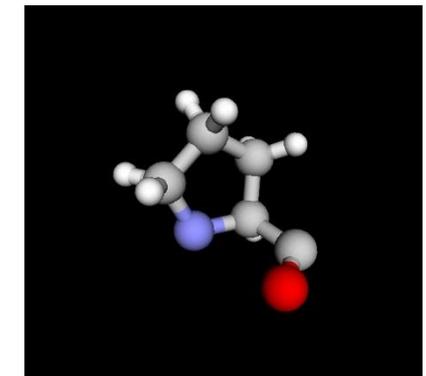
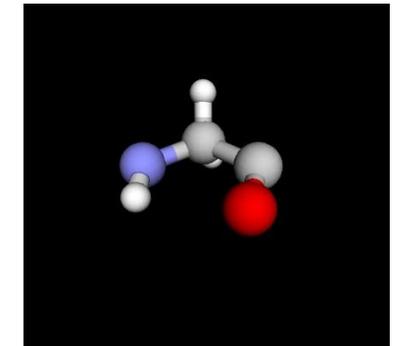
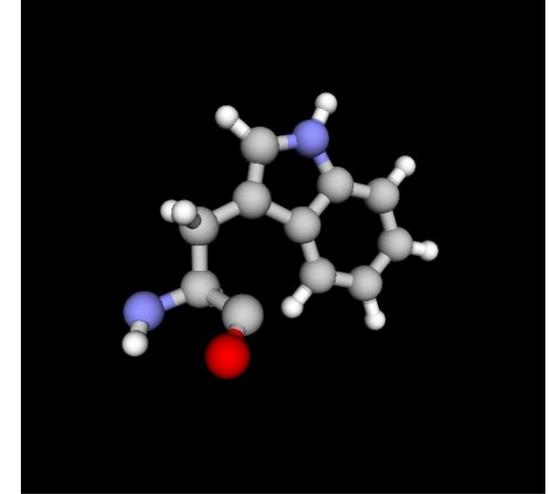
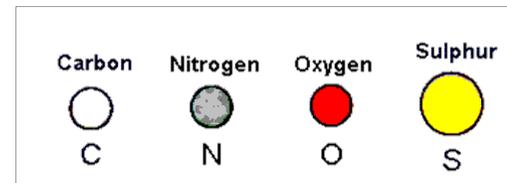
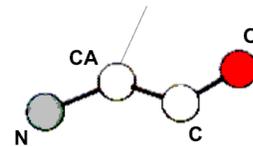
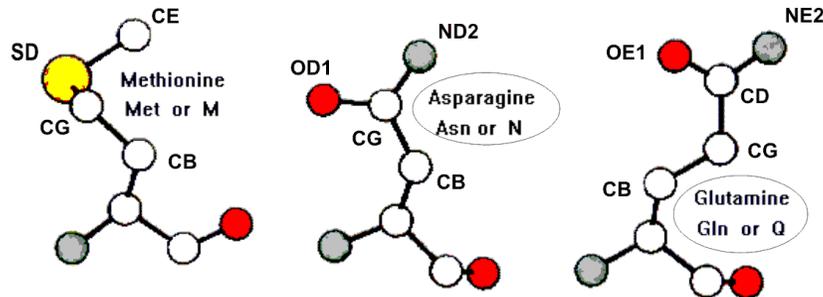
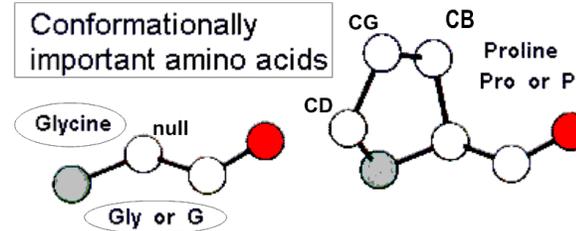
Acidic amino acids



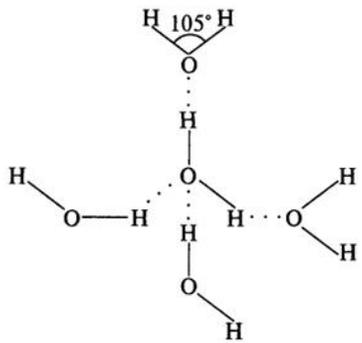
Neutral-polar amino acids



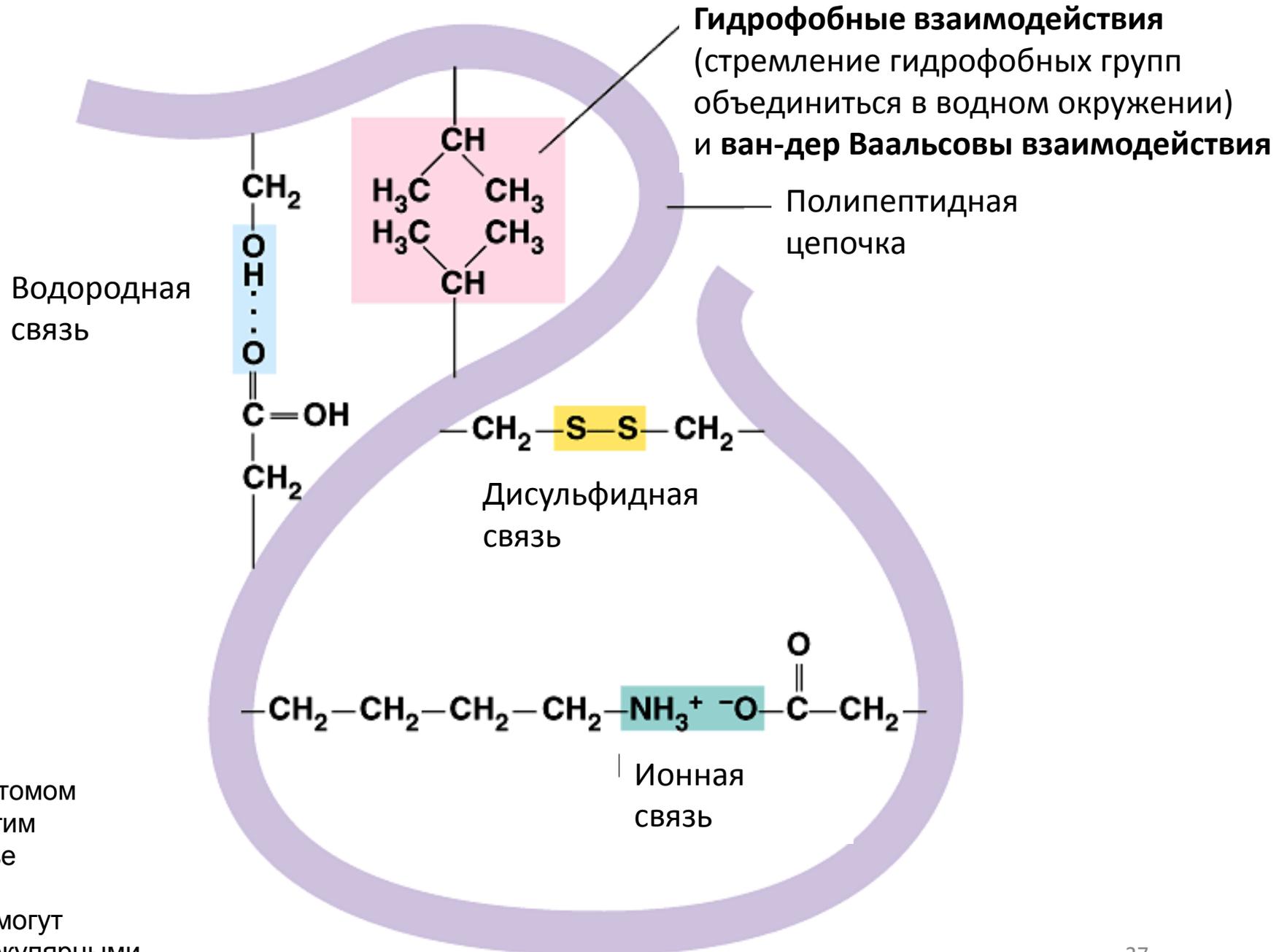
Conformationally important amino acids



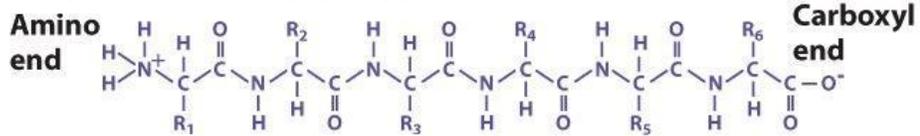
# Внутримолекулярные взаимодействия, способствующие укладке белковой цепи (фолдингу)



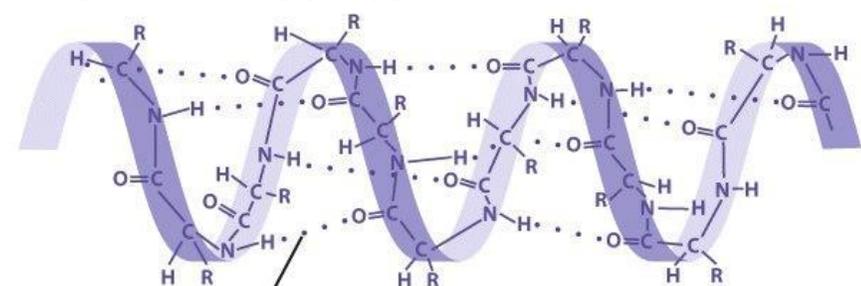
**Водородная связь** образуется между электроотрицательным атомом и атомом водорода **H**, связанным ковалентно с другим электроотрицательным атомом. В качестве электроотрицательных атомов могут выступать N, O или F. Водородные связи могут быть межмолекулярными или внутримолекулярными.



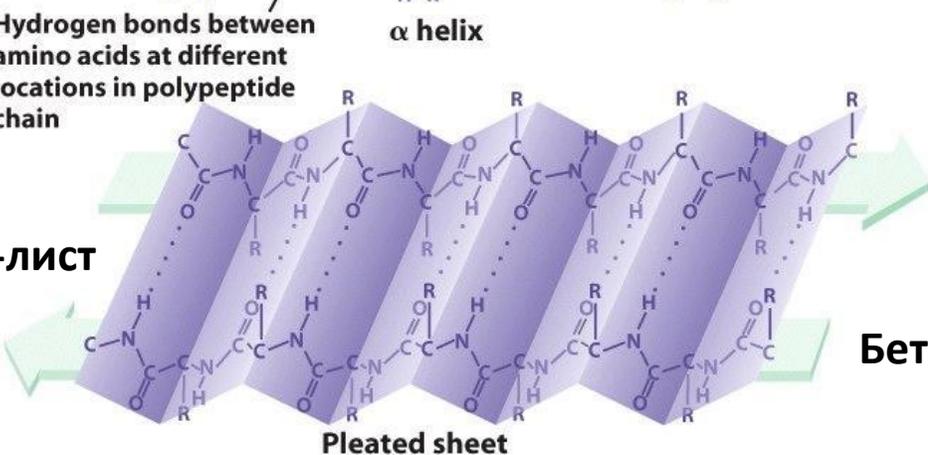
## Первичная структура



## Вторичная структура

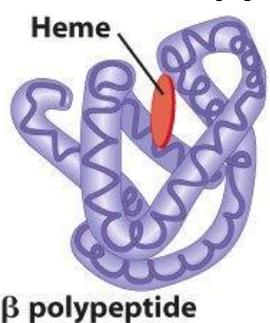


Альфа-спираль

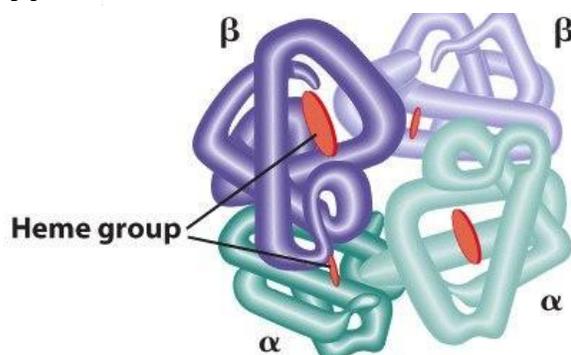


Бета-тяж

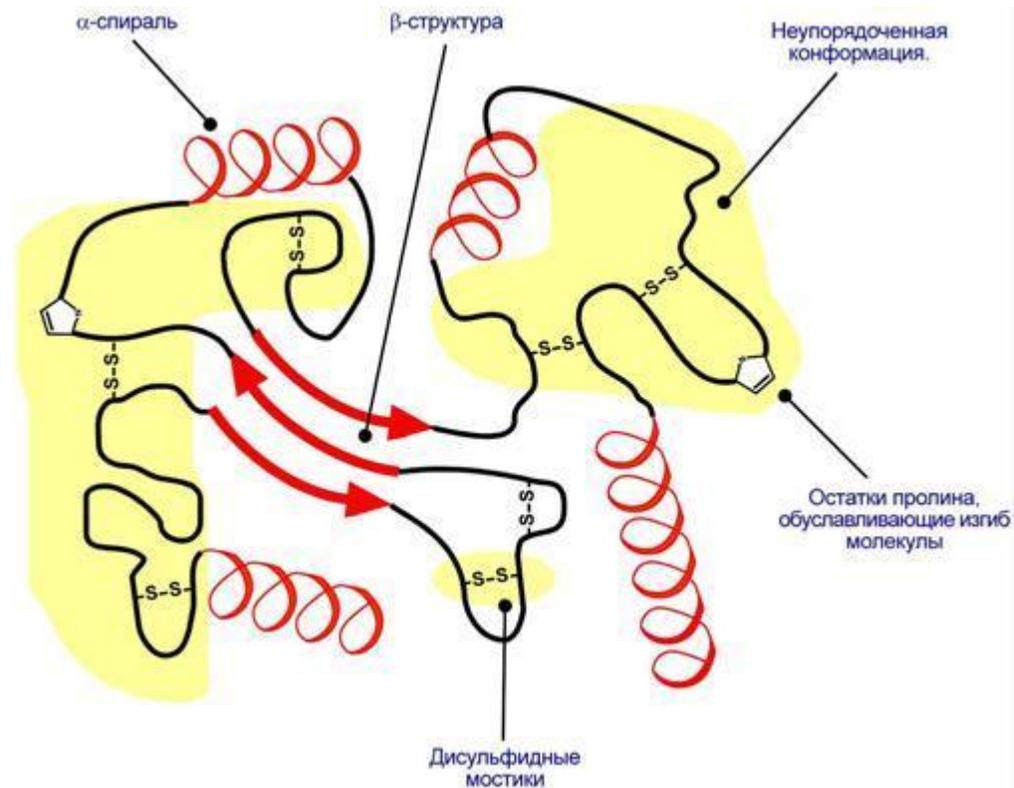
## Третичная структура

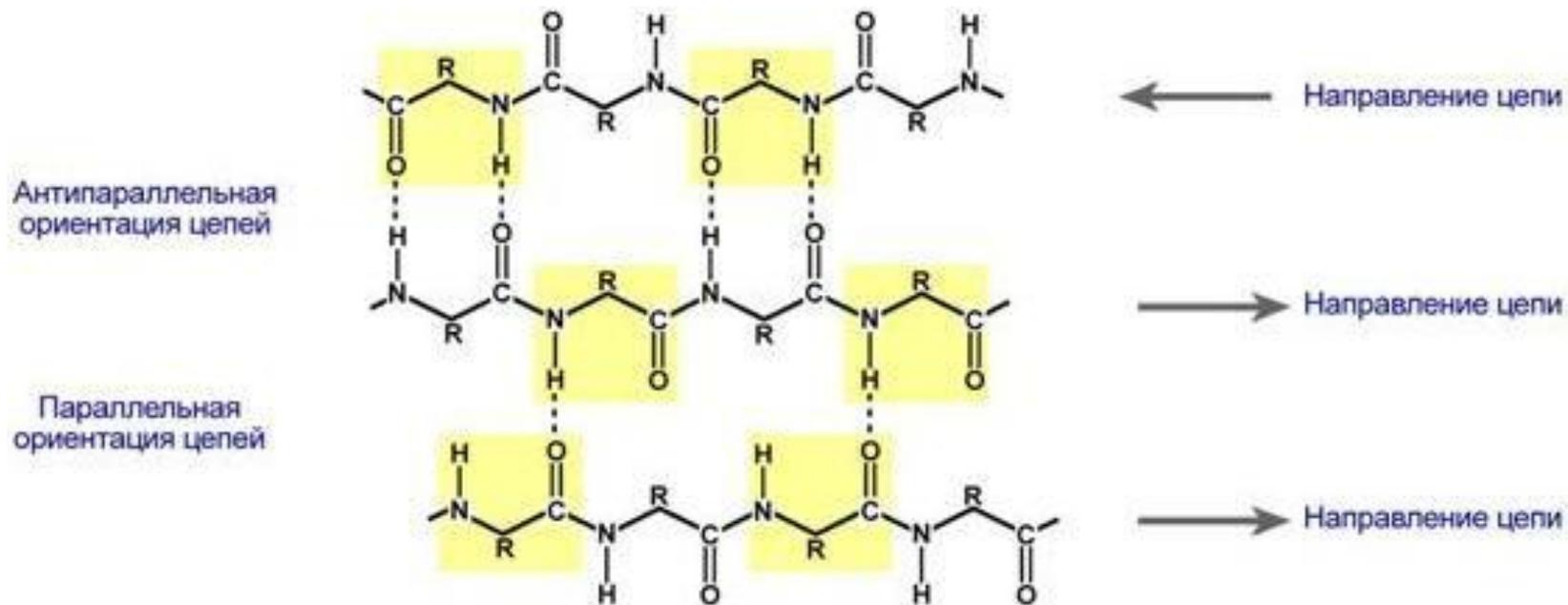
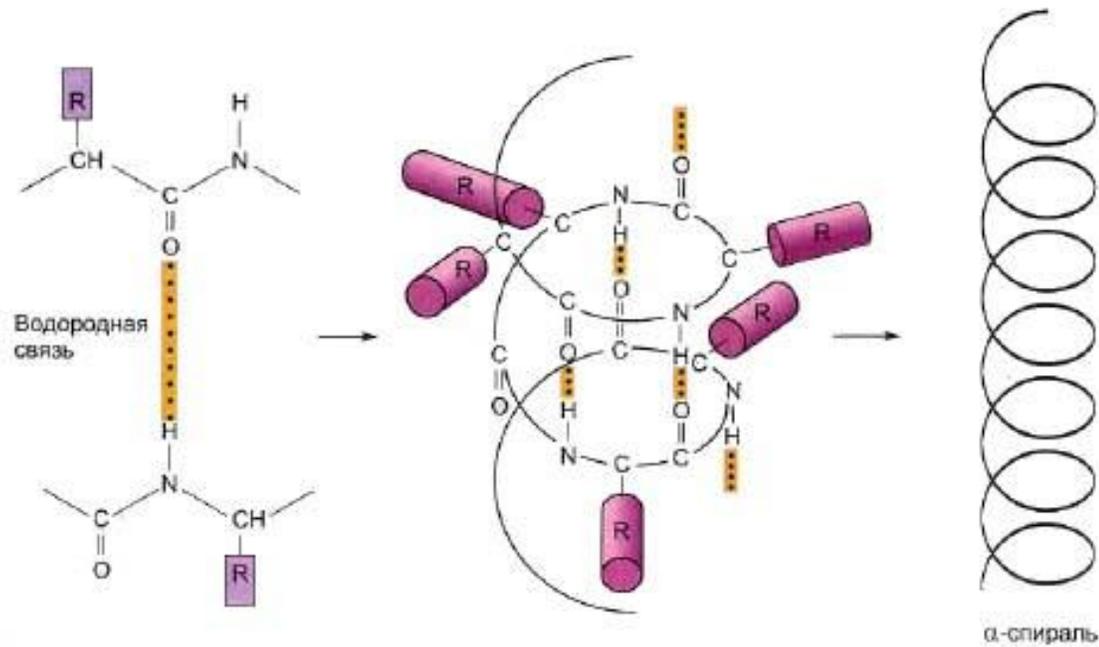


## Четвертичная структура

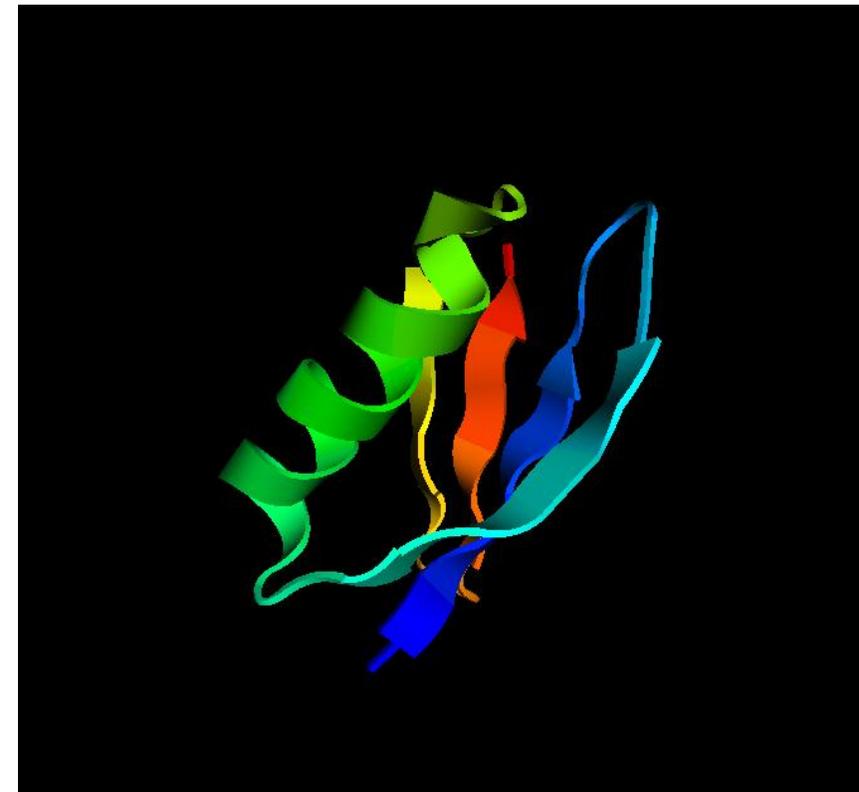
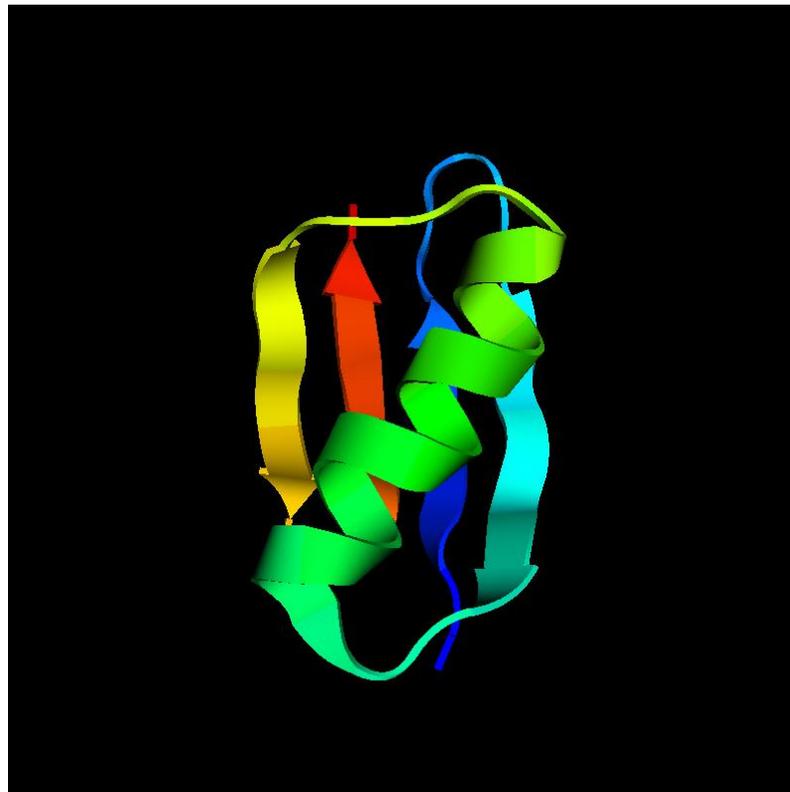
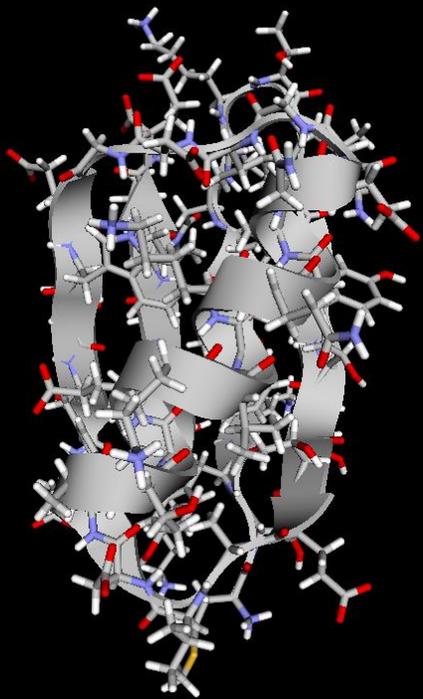


# Уровни структурной организации белка





Бета-лист



Структура маленького белка (protein G, [2GB1](#)) из 56 аминокислот. Остов белка показан в cartoon expression.

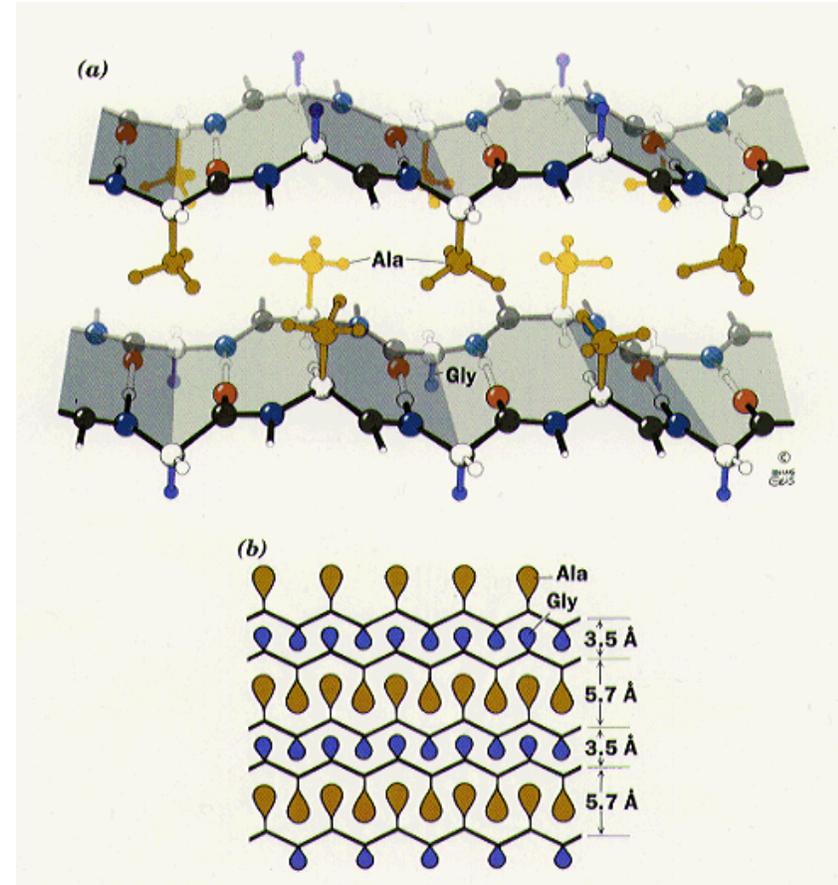
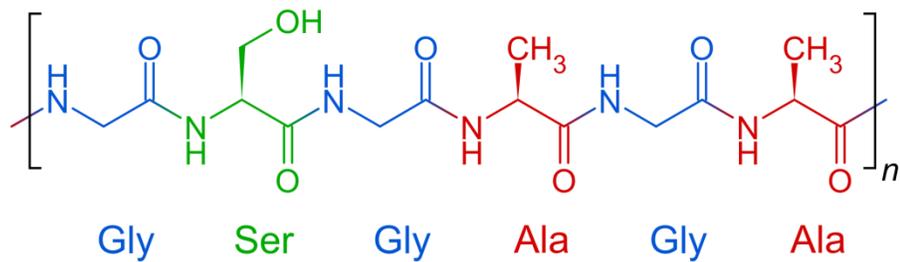
**Первичная структура:**

MTYKLILNGK TLKGETTTEA VDAATAEKVF KQYANDNGVD GEWTYDDATK TFTVTE

**Вторичная структура** образована двумя антипараллельными бета-тяжами, альфа-спиралью, и еще двумя антипараллельными бета-тяжами, между которыми находятся неструктурированные участки (петли).

**Третичная структура** представляет собой глобулярный белок, в котором альфа-спираль лежит поверх бета-слоя.

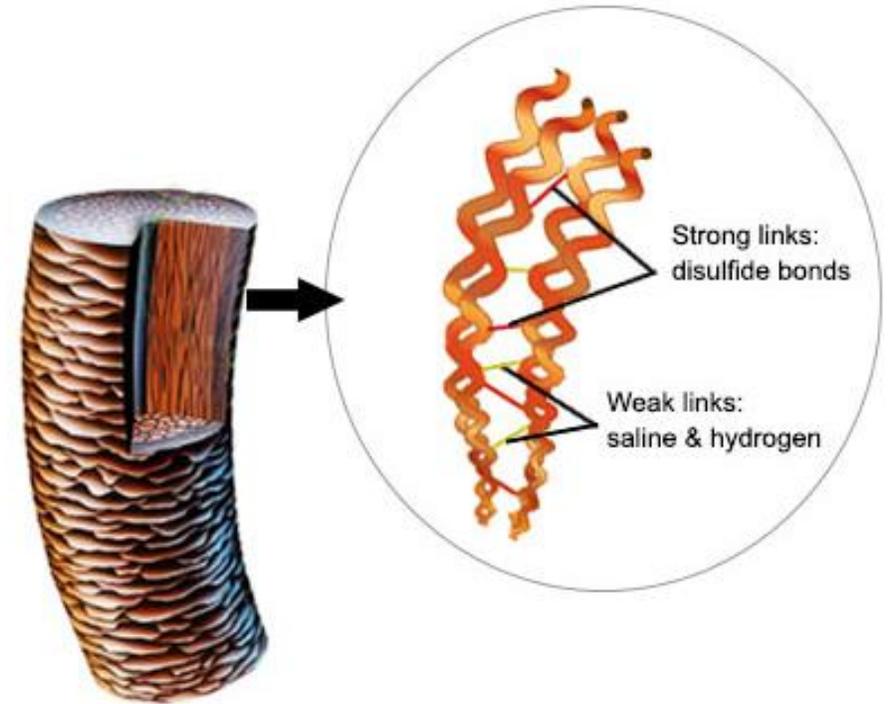
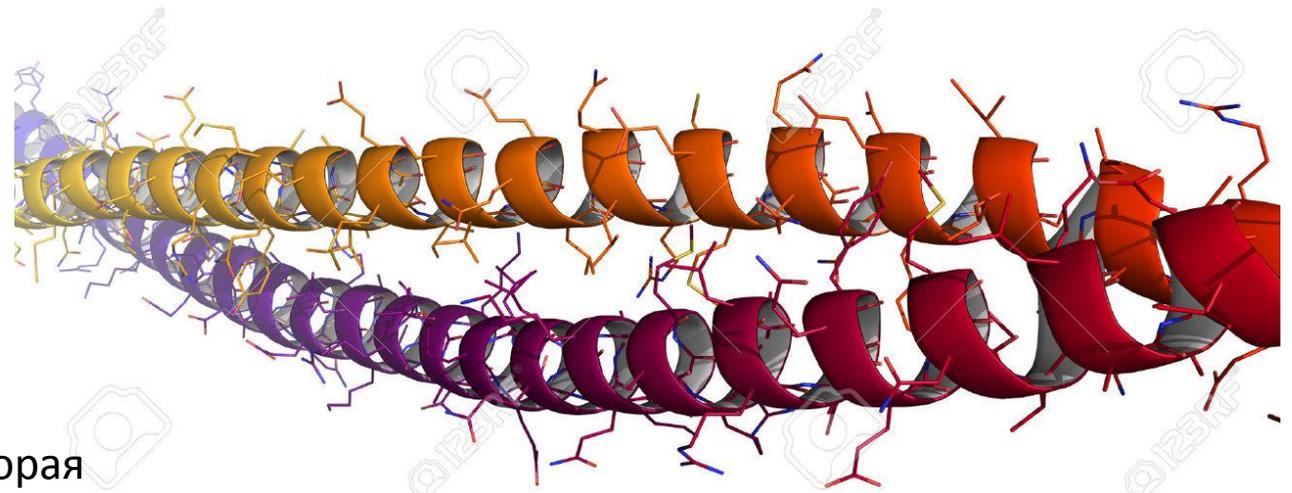
# Бета-структурные фибрилярные белки



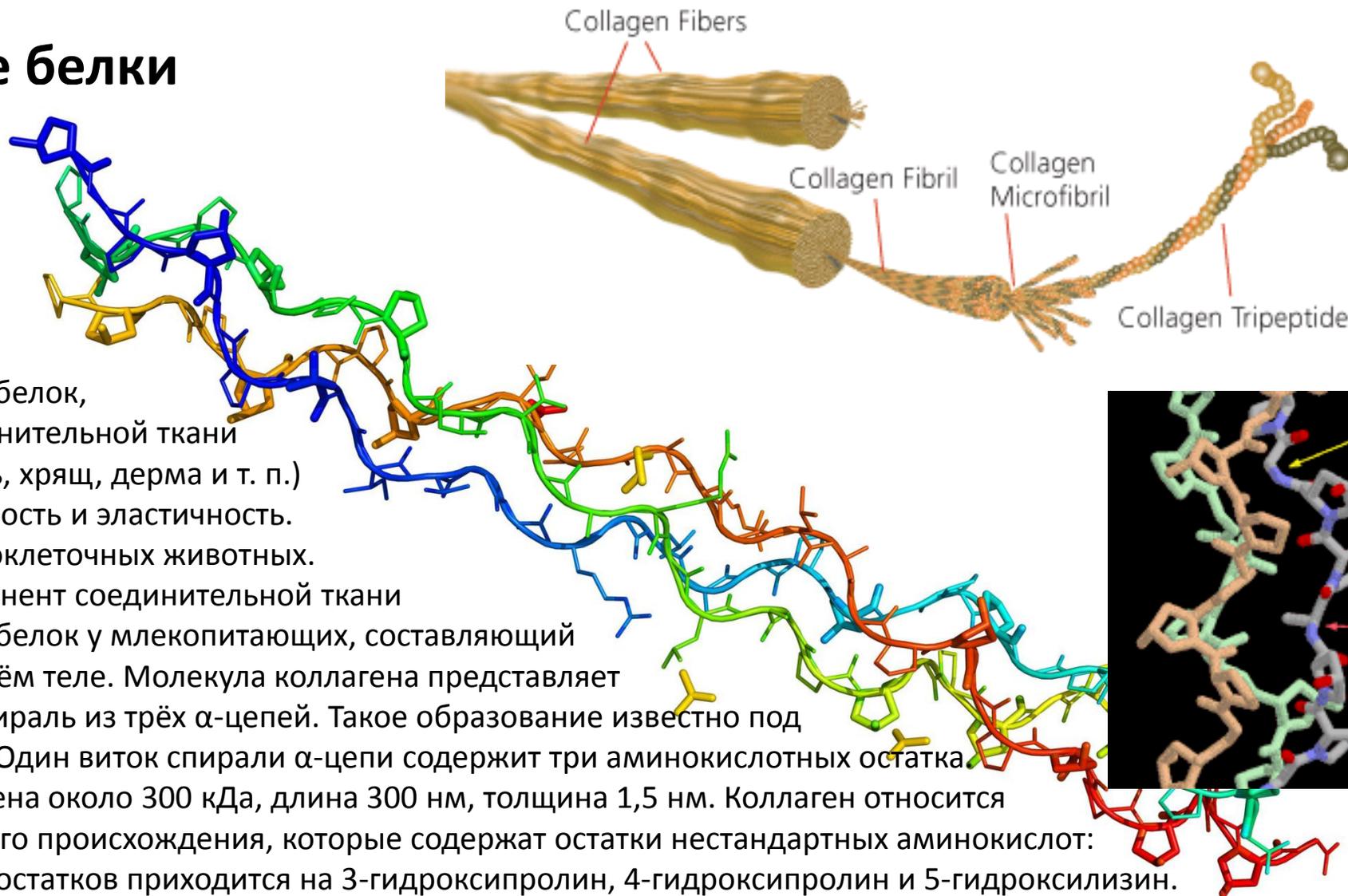
**Фиброин** – фибриллярный белок, выделяемый паукообразными и некоторыми насекомыми и составляющий основу нитей паутины и коконов насекомых, в частности шёлка тутового шелкопряда *Bombyx mori*. Его первичная структура состоит из повторяющейся аминокислотной последовательности (Gly-Ser-Gly-Ala-Gly-Ala)<sub>n</sub>. В свою очередь, повторяющиеся аминокислотные последовательности образуют антипараллельные складчатые β-слои, связанные водородными связями. Эта структура обуславливает высокий предел прочности нитей паутин и шелка. Антипараллельные бета-слои уложены друг на друга по принципу «лицом к лицу, спиной к спине»: двойной слой глицинов (расстояние между плоскостями 3.5 ангстрем) – двойной слой аланинов/серинов (расстояние между плоскостями 5.7 ангстрем) и т.д.

# Альфа-структурные фибриллярные белки

**Кератины** — семейство фибриллярных белков, обладающих высокой механической прочностью, которая среди материалов биологического происхождения уступает лишь хитину. В основном из кератинов состоят роговые производные эпидермиса кожи - такие структуры, как волосы, ногти, рога носорогов, перья и рамфотека клюва птиц и др. В это семейство входят также цитокератины, образующие наиболее прочные элементы внутриклеточного цитоскелета эпителиальных клеток. **α-кератины** являются основой волос (включая шерсть), рогов, когтей и копыт млекопитающих. Белки сложены из длинных перевитых альфа-спиралей (coiled coil). Содержат много глицина и цистеина, образующего дисульфидные связи. Человеческие волосы на 14% состоят из цистеина.



# Фибриллярные белки (коллаген)



**Коллаген** — фибриллярный белок, составляющий основу соединительной ткани организма (сухожилие, кость, хрящ, дерма и т. п.) и обеспечивающий её прочность и эластичность. Коллаген обнаружен у многоклеточных животных.

Коллаген — основной компонент соединительной ткани и самый распространённый белок у млекопитающих, составляющий от 25 % до 35 % белков во всём теле. Молекула коллагена представляет собой правозакрученную спираль из трёх  $\alpha$ -цепей. Такое образование известно под названием *тропоколлаген*. Один виток спирали  $\alpha$ -цепи содержит три аминокислотных остатка. Молекулярная масса коллагена около 300 кДа, длина 300 нм, толщина 1,5 нм. Коллаген относится к немногим белкам животного происхождения, которые содержат остатки нестандартных аминокислот: около 21 % от общего числа остатков приходится на 3-гидроксипролин, 4-гидроксипролин и 5-гидроксилизин.

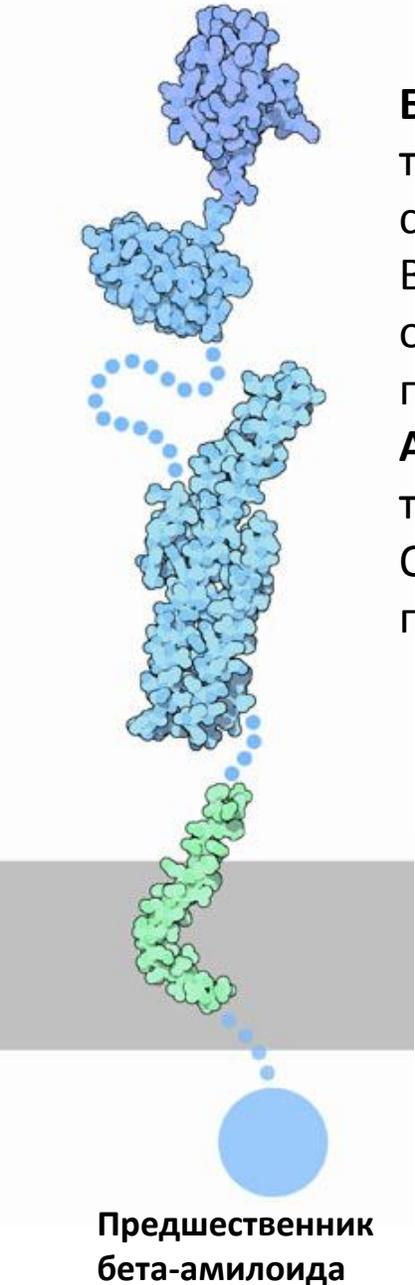
Каждая из  $\alpha$ -цепей состоит из триад аминокислот. В триадах третья аминокислота всегда глицин, вторая — пролин или лизин, первая — любая другая аминокислота, кроме трёх перечисленных. Мутации в генах коллагенов приводят к тяжелым генетическим заболеваниям, в частности, *osteogenesis imperfecta*, или болезнь «ломких костей».

Синтез нестандартных аминокислот происходит посттрансляционно. Для этого нужен витамин С. При нехватке витамина С нарушается синтез коллагенов. Развивающееся при этом заболевание называется цинга. Симптомы - выпадение зубов и слабость.

# Амилоидные фибриллы

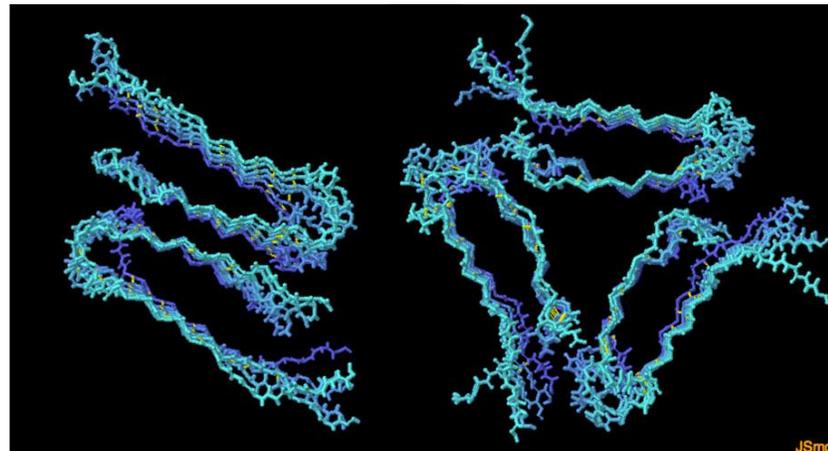
**Бета-амилоид** - пептид из 42 аминокислотных остатков, образующийся из трансмембранного белка — предшественника бета-амилоида. Пептид A $\beta$ 42 считается одним из основных факторов, провоцирующих болезнь Альцгеймера. В мозге пациента, страдающего болезнью Альцгеймера, этот пептид может образовывать так называемые амилоидные бляшки, состоящие из скоплений пептида, свёрнутого в виде бета-складчатых структур.

**Амилоидные фибриллы** могут формироваться и из других белков, например, таких «нормальных» глобулярных белков, как лизоцим, миоглобин и т.д. Образование амилоидных фибрилл также вызывает инфекционная форма прионов.

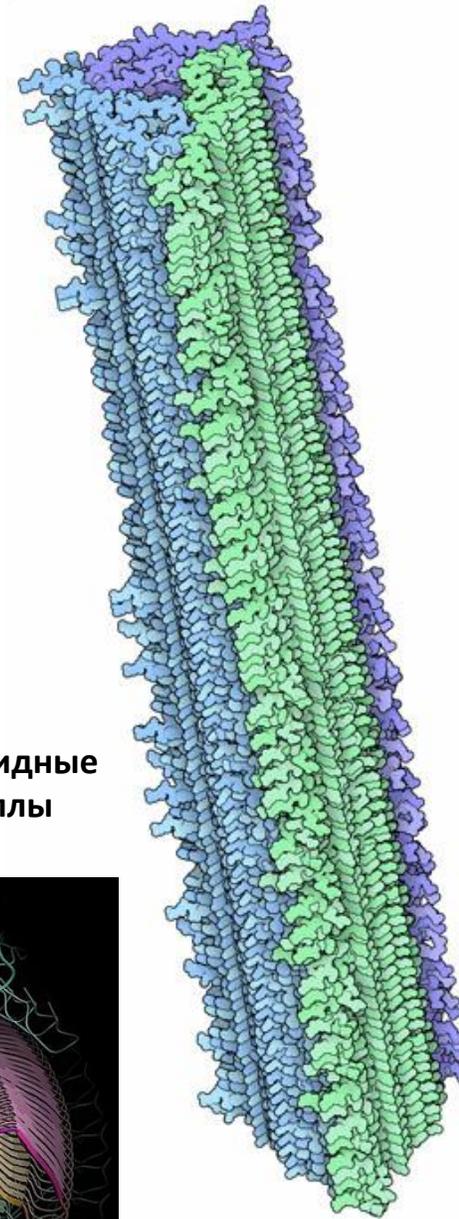
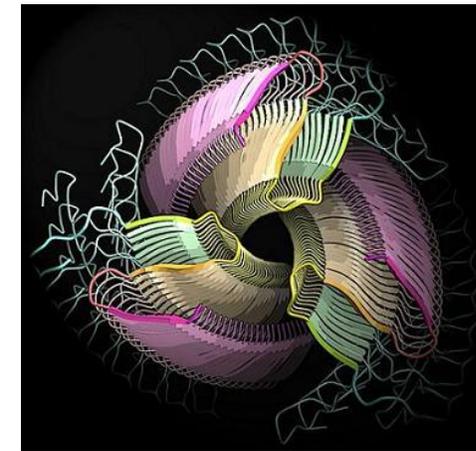


Бета-амилоид в альфа-спиральной конформации

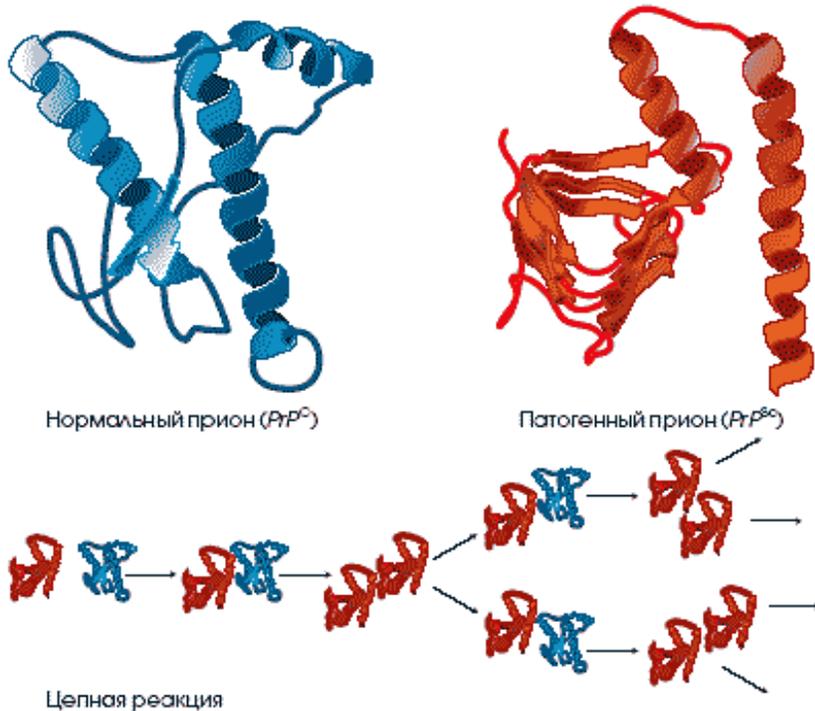
Бета-амилоид в бета-структурной конформации



Амилоидные фибриллы



# Прионы



**Прионы** (*prion* от *protein* — «белок» и *infection* — «инфекция») — особый класс инфекционных агентов, представленных белками с аномальной трёхмерной (третичной) структурой, способными катализировать конформационное превращение гомологичных им нормальных клеточных белков в себе подобные (Стенли Прузинер, Нобелевская премия 1997 года). Как правило, при переходе белка в прионное состояние его  $\alpha$ -спирали превращаются в  $\beta$ -слои. Появившиеся в результате такого перехода прионы могут в свою очередь перестраивать новые молекулы белка; таким образом, запускается цепная реакция, в ходе которой образуется огромное количество неправильно свёрнутых молекул. Все известные прионы вызывают формирование амилоидов — белковых агрегатов, включающих плотно упакованные  $\beta$ -слои. Прионы вызывают трансмиссивные губчатые энцефалопатии (ТГЭ) у различных млекопитающих, в том числе скрепи (почесуху) овец, губчатую энцефалопатию крупного рогатого скота («коровье бешенство»). У человека прионы вызывают болезнь Крейтцфельдта — Якоба, синдром Герстмана — Штраусслера — Шейнкера (замена пролина на лейцин в кодоне 102 PRNP) и фатальную семейную бессонницу (замена аспарагиновой кислоты на аспарагин в кодоне 178 гена PRNP), куру (Даниэль Карлтон Гайдузек — Нобелевская премия 1976 года). Все известные прионные заболевания поражают головной мозг и другие нервные ткани, в настоящее время неизлечимы и в конечном итоге смертельны. Прионные заболевания млекопитающих вызываются белком PrP. Его нормальная функция в организме до сих пор не установлена.<sup>35</sup>

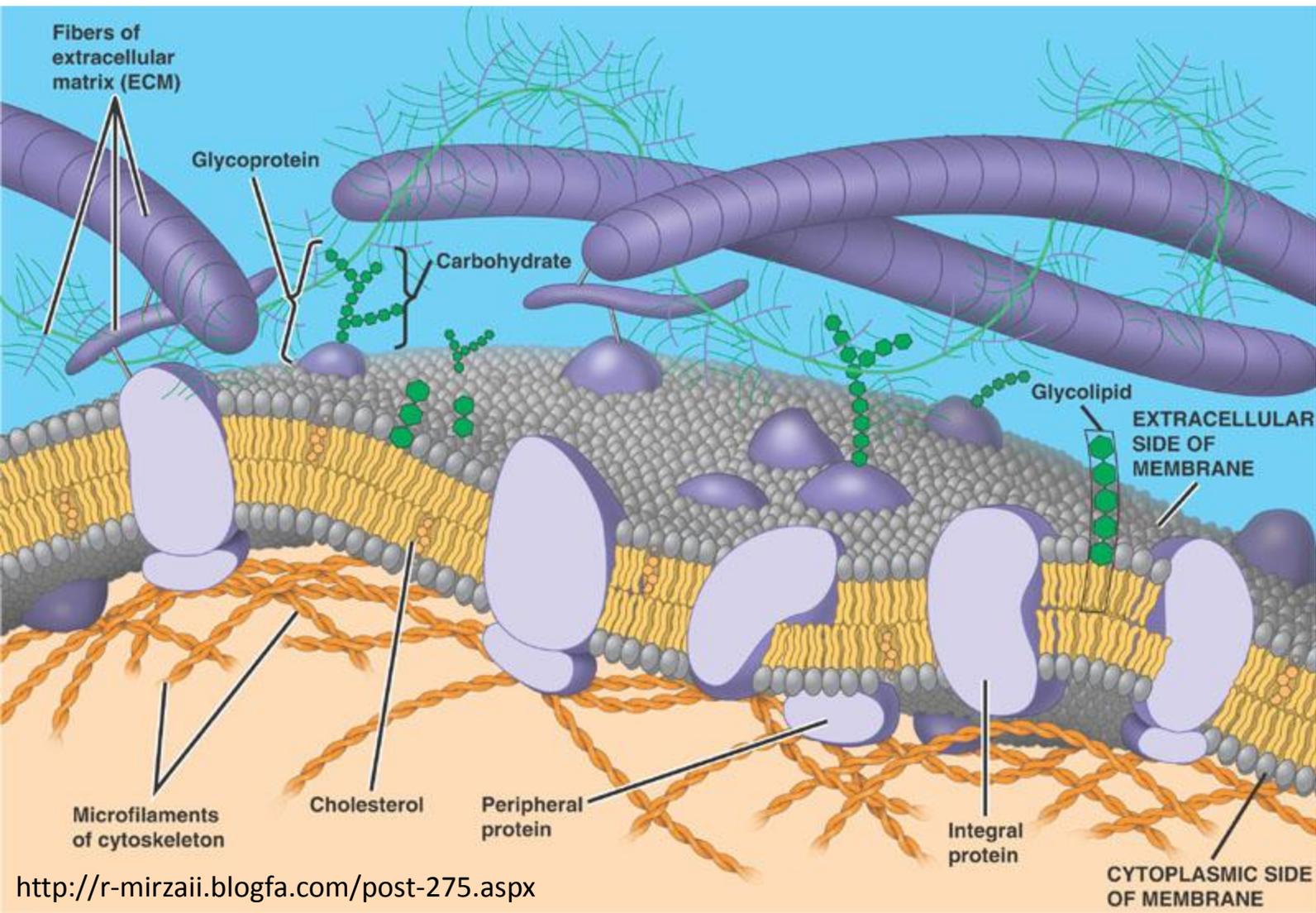
# Мембранные белки: классификация

Биохимическая классификация:

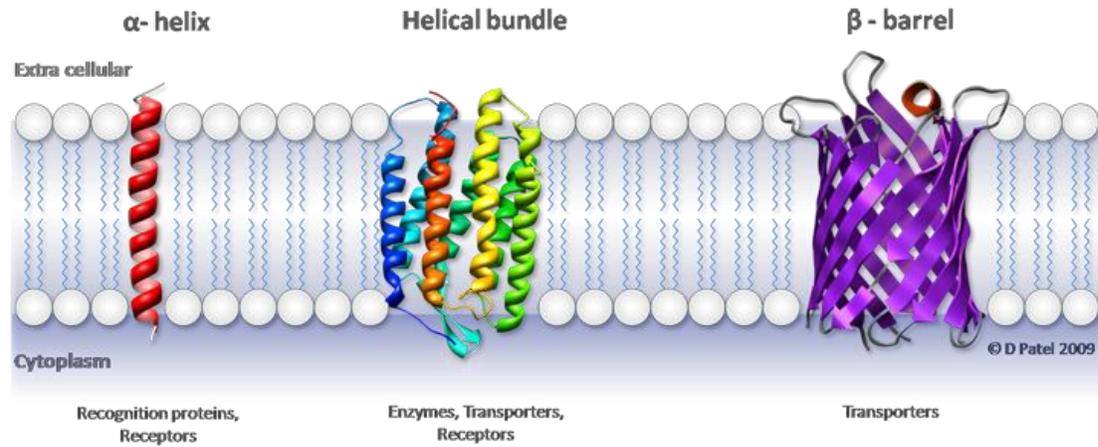
**Интегральные мембранные белки** прочно встроены в мембрану и могут быть извлечены из липидного окружения только с помощью детергентов или неполярных растворителей. По отношению к липидному бислою интегральные белки могут быть трансмембранными **политопическими** (выходить на обе стороны) или интегральными **монотопическими** (выходить на одну сторону).

**Периферические мембранные белки** являются монотопическими белками. Они либо связаны слабыми связями с липидной мембраной, либо ассоциируют с интегральными белками за счёт гидрофобных, электростатических или других нековалентных сил. Таким образом, в отличие от интегральных белков они диссоциируют от мембраны при обработке соответствующим водным раствором (например, с низким или высоким pH, с высокой концентрацией соли или под действием хаотропного агента). Эта диссоциация не требует разрушения мембраны.

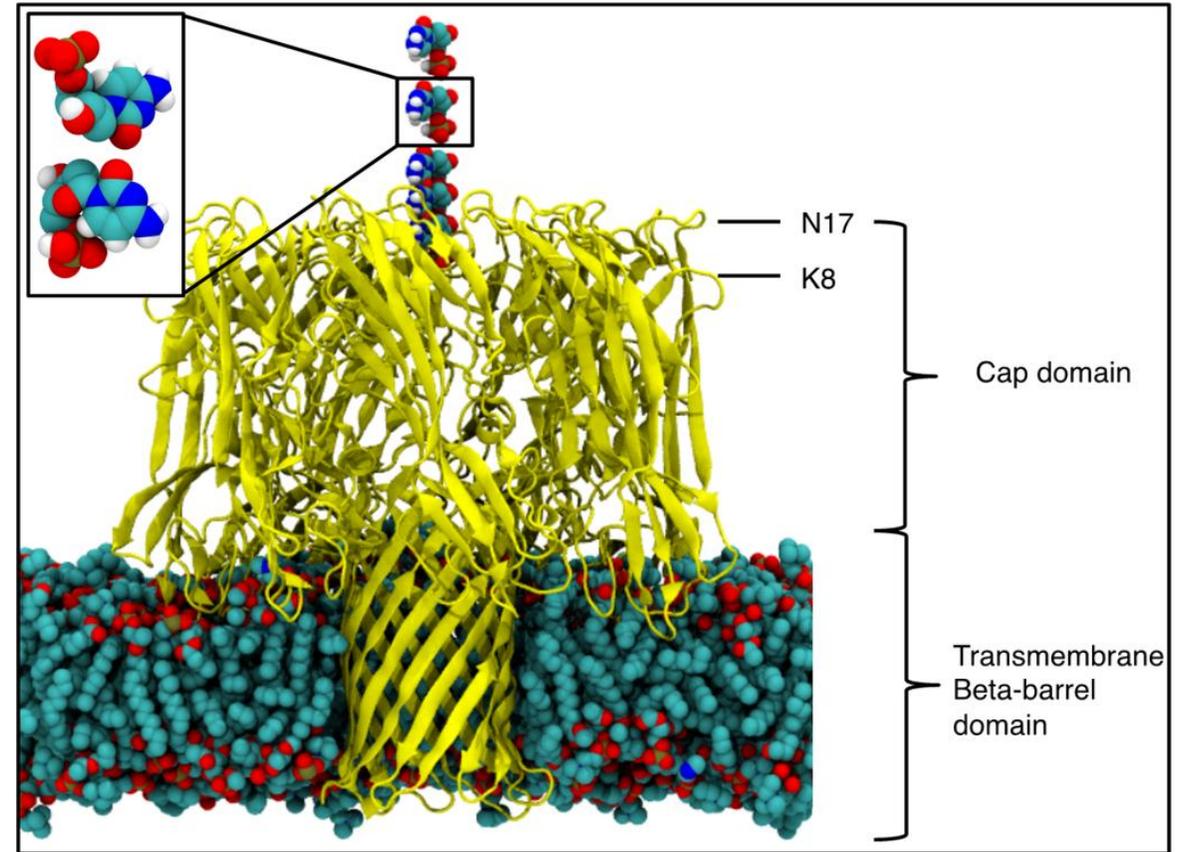
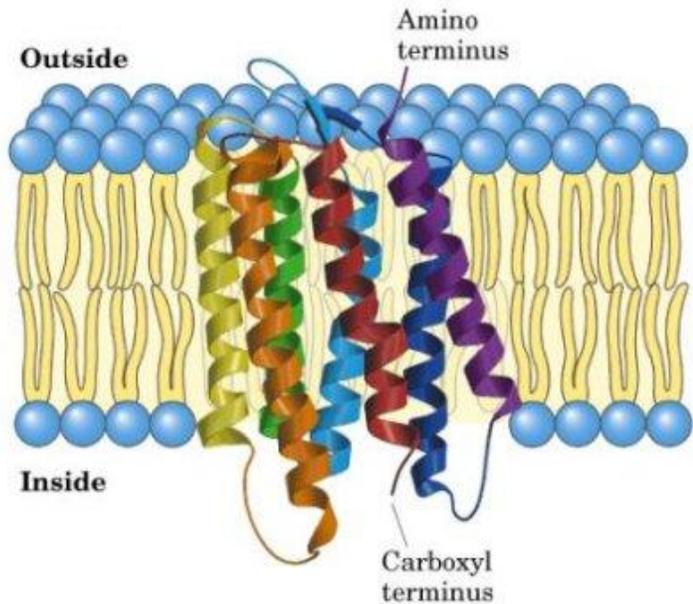
Мембранные белки могут быть встроены в мембрану за счёт [жирнокислотных](#) или пренильных остатков либо [гликозилфосфатидилинозитола](#), присоединённых к белку в процессе их [посттрансляционной модификации](#).



# Мембранные белки: примеры структур



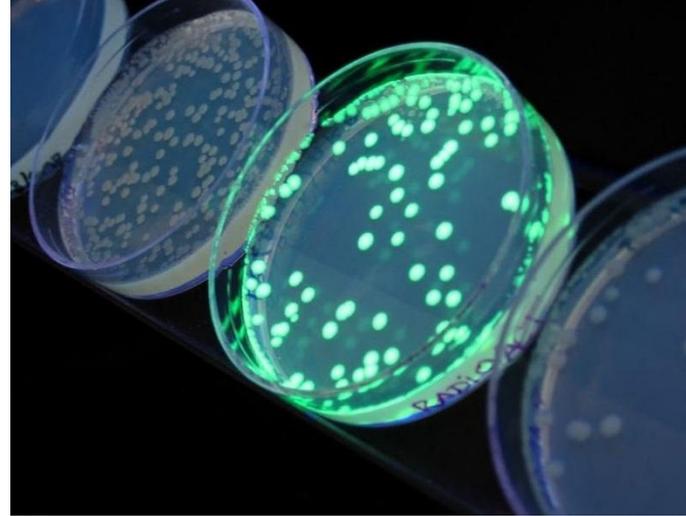
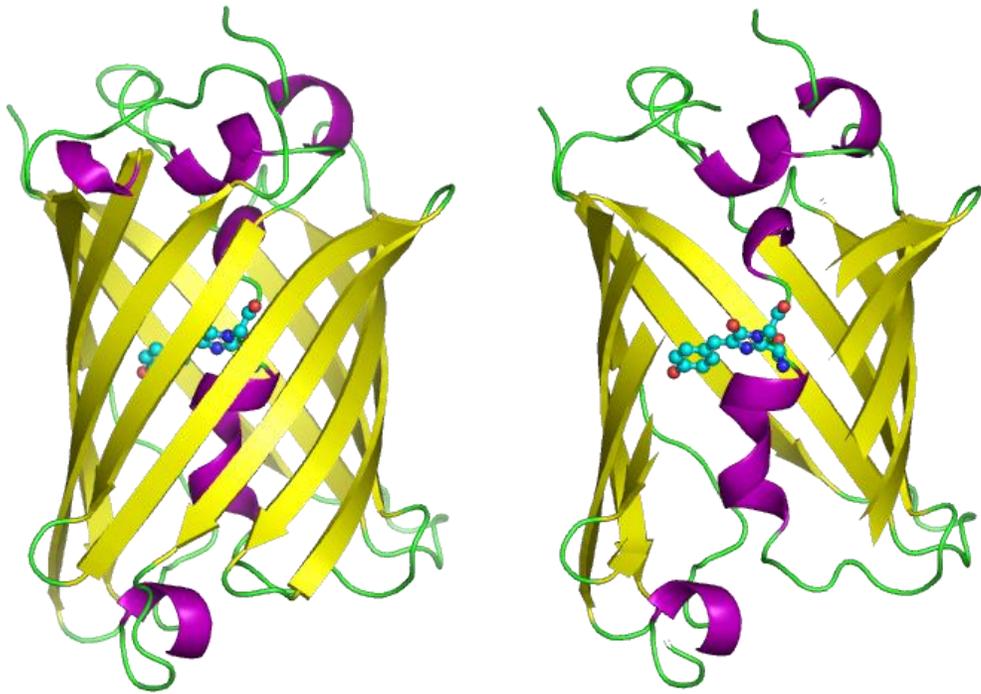
<http://www2.warwick.ac.uk/fac/sci/chemistry/research/dixon/dixongroup/members/msrhar/research/background/>



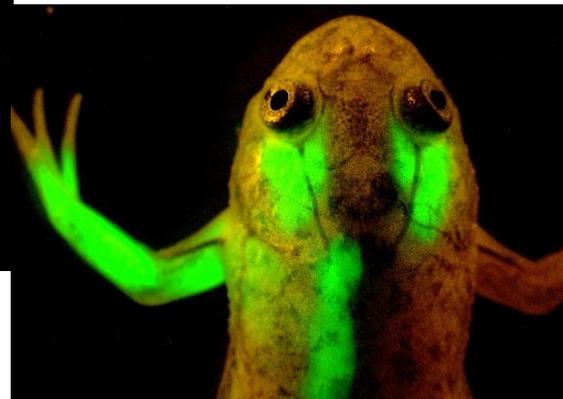
**Альфа-гемолизин** (*Staphylococcus aureus*, PDB код 7ahl) - токсин, формирующий поры в клеточных мембранах. Встроившись в мембрану, он создает в ней ионные каналы, что в итоге приводит к разрушению клетки.

**Бактериородопсин** (проводит протон через мембрану)

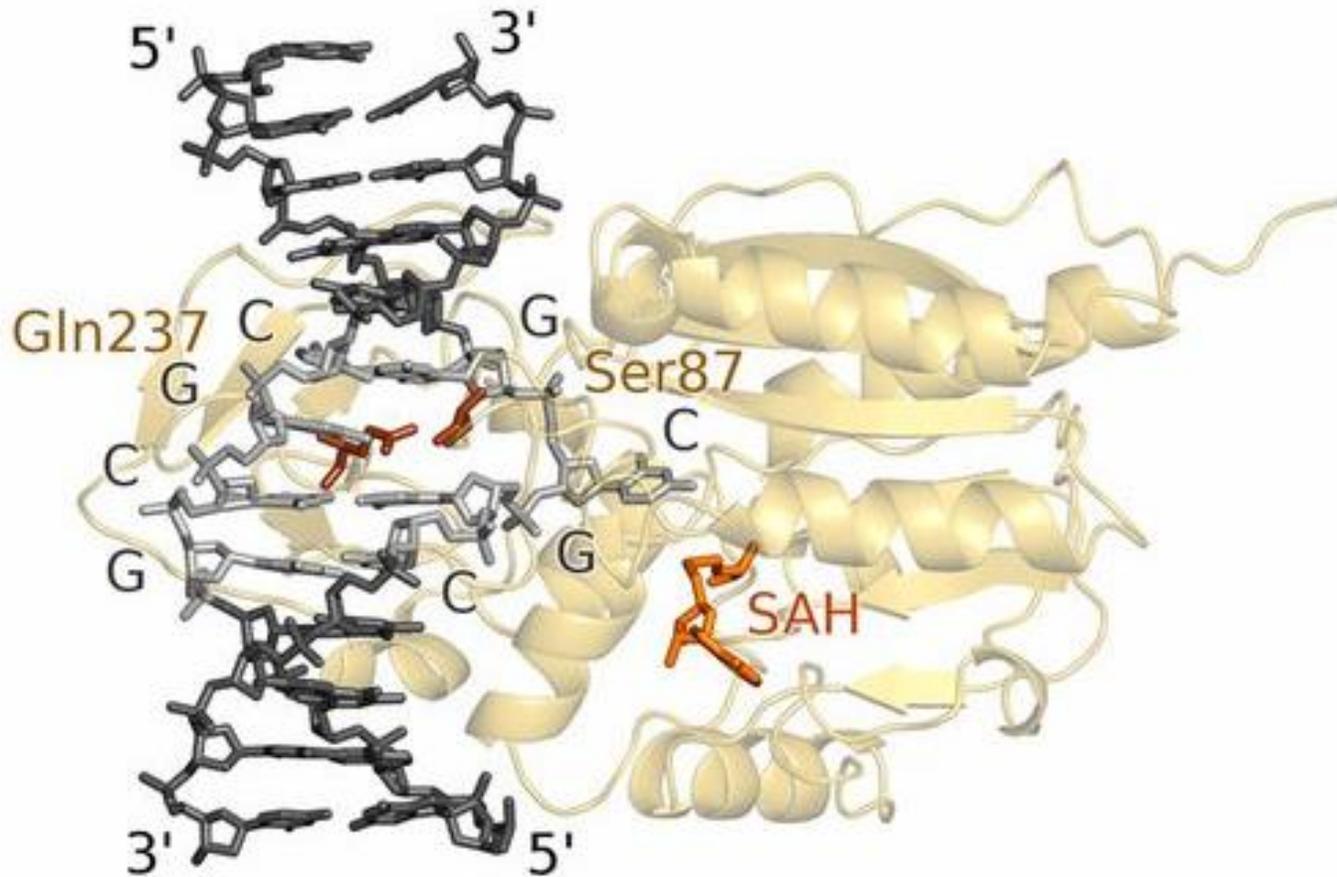
# Green Fluorescent Protein



*GFP* — белок, выделенный из медузы [\*Aequorea victoria\*](#), который флуоресцирует в зелёном диапазоне при освещении его синим светом. В 2008 году Осаму Симомура, Мартин Чалфи и Роджер Тсьен получили Нобелевскую премию по химии «за открытие и разработку зелёного флуоресцентного белка GFP». Белок состоит из 238 аминокислотных остатков с молекулярной массой 26,9 кДа. Белок представляет собой бета-баррель из 11 бета-тяжей, внутри которого находится флуорофор. Оболочка цилиндра защищает флуорофор от тушения его флуоресценции компонентами микроокружения. Кроме этого, внутренняя структура молекулы вызывает специфические реакции циклизации трипептида Ser65-Tyr66-Gly67, что приводит к образованию флуорофора.



# ДНК-белковые комплексы: ДНК-метилтрансфераза HhaI



M.HhaI

C5-Цитозин-ДНК-метилтрансферазы катализируют перенос метильной группы от S-аденозил-метионина на остаток [цитозина](#), находящегося в специфической последовательности в двухцепочечной [ДНК](#), с образованием [5-метилцитозина](#) и S-аденозилгомоцистеина.

Получение кристаллов ДНК-белковых комплексов очень сложный процесс. Удалось научиться кристаллизовать ДНК-метилтрансферазу HhaI с различными вариантами ДНК и донора метильной группы. Всего расшифровано 29 комплексов ДНК-метилтрансферазы HhaI с ДНК. Помимо этого удалось расшифровать лишь 2 комплекса для двух цитозиновых ДНК-метилтрансфераз бактерий.

# ДНК-белковые комплексы: рибосома

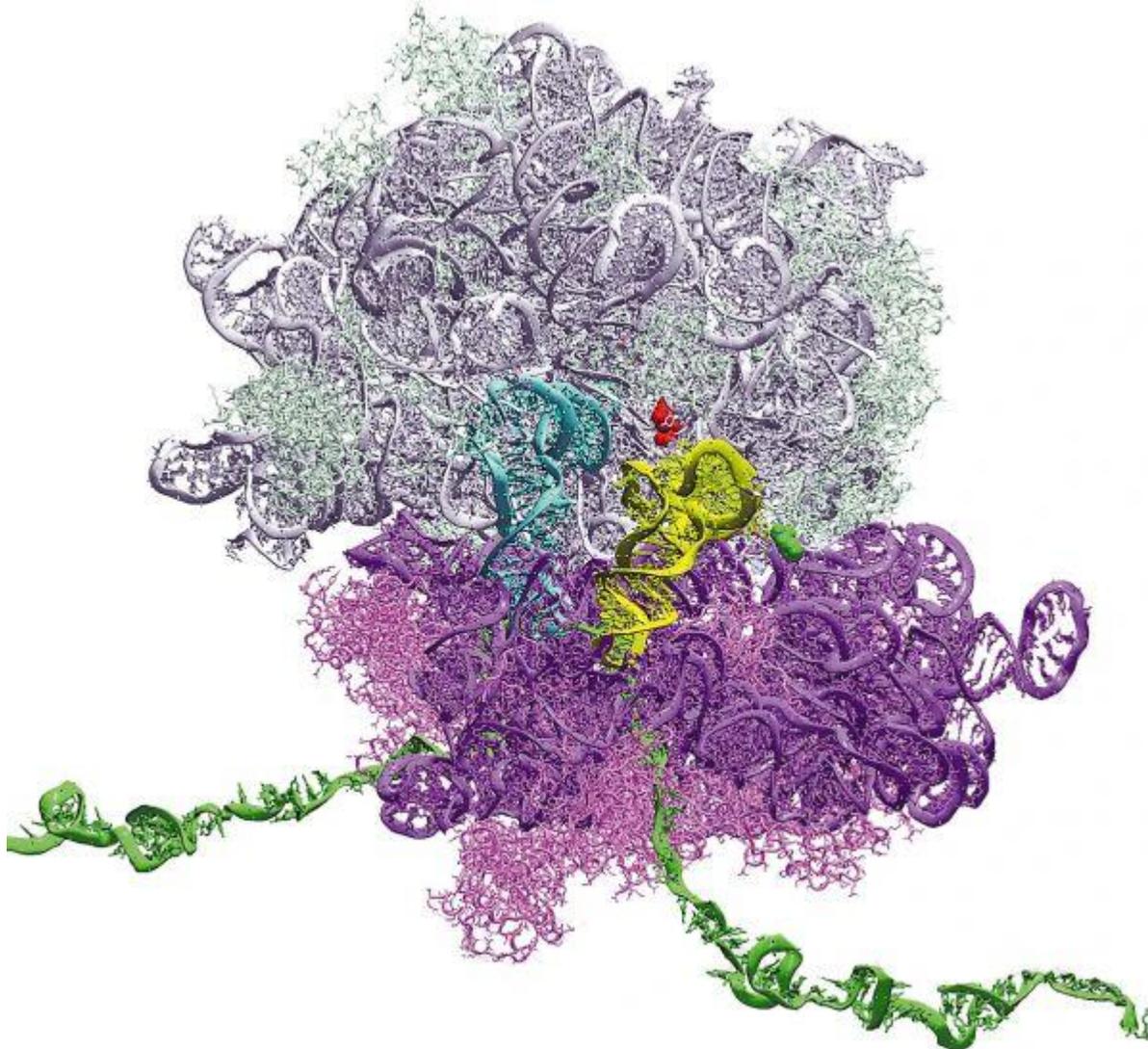
Эукариотическая рибосома (константы седиментации 80S (40S+60S)) включает 4 молекулы РНК и около 100 белков.

PDB содержит большое число расшифрованных структур различных рибосом и их компонентов, полученных разными методами:

X-ray (1840)

Electron Microscopy (688)

NMR (251)



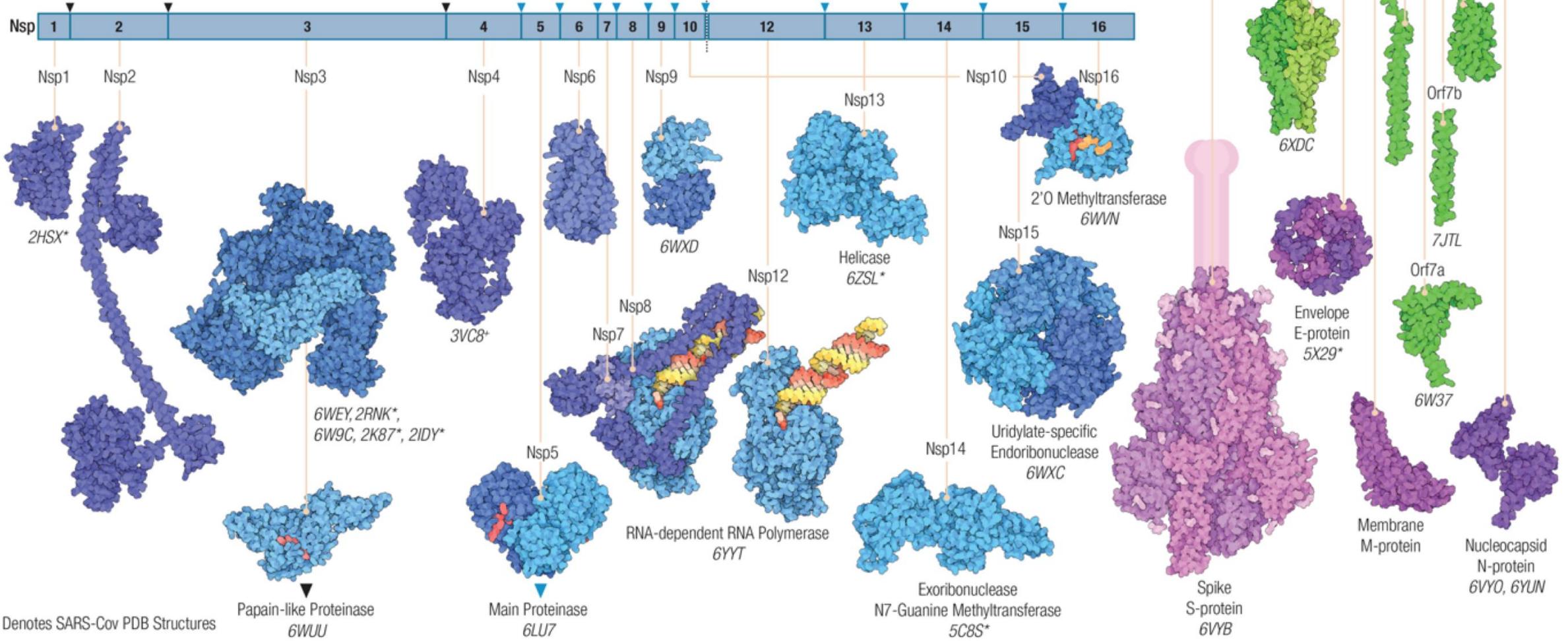
Genome Position



Open Reading Frames (Orf)



Non-Structural Proteins (Nsp)



\* Denotes SARS-Cov PDB Structures

\* Denotes Mouse Hepatitis Virus Structures

© Denotes N7-methyl Guanine Cap

AAAA Denotes Poly-A Tail

## Домашнее задание по Лекции 4 – Структура биополимеров

1. Найдите в PDB (<https://www.rcsb.org/>) структуру какого-нибудь белка из списка:

- Пепсин
- Трипсин
- Миоглобин
- Лизоцим
- Родопсин
- Инсулин
- Ещё какого-нибудь известного вам

В отчёте приведите: детали поиска (какое слово куда писали и что потом выбирали), сколько всего белков с подобным названием нашлось, код (4 символа) выбранной структуры, её краткое описание (желательно с переводом на русский), дату депонирования в PDB, из какого организма белок, ещё что-то, привлекшее ваше внимание на странице структуры. Из текста, приведенного в Википедии, составьте несколько предложений, описывающих функцию выбранного белка.

*Указания.* Выясните, как пишется название белка по-английски (хороший способ: найти статью об этом белке в русской Википедии и затем перейти по ссылке на соответствующую английскую статью). Зайдите на сайт PDB <https://www.rcsb.org/>. Внесите название в окошко поиска справа вверху или пройдите по ссылке *Advanced search*, далее действуйте по смыслу. На странице конкретного комплекса: краткое описание находится сразу под кодом, остальная информация чуть ниже.

2. На странице структуры найдите гиперссылку “3D View: Structure” и пройдите по ней (если такой гиперссылки вдруг нет, смените браузер :). Найдите окошко “Select a different viewer” и выберите JSmol. Создайте два изображения структуры на сером фоне: полноатомное с раскраской по химическим элементам; остовное с раскраской по цепям и приложите к отчёту графические файлы с изображениями.

*Указания.* Щёлкните правой кнопкой мыши по окошку JSmol и выберите Console. Чтобы создать полноатомное изображение, нужно выполнить в консоли следующие команды: `srk only, color srk, background grey`. Чтобы перейти от полноатомного изображения к остовному с раскраской по цепям, нужно выполнить следующие команды: `srk off, backbone 0.3` (вместо 0.3 может стоять 0.2 или 0.5, подберите толщину линий по своему вкусу), `color chain`. Чтобы сохранить изображение в графический файл, нужно (подобрав подходящий ракурс и величину изображения) щёлкнуть правой кнопкой мыши по окошку JSmol и в открывшемся меню выбрать File → Export.

3. (\* — дополнительно) Изучите руководство по Jmol <https://kodomo.fbb.msu.ru/wiki/Main/Jmol> и создайте, пользуясь возможностями Jmol или JSmol (интерфейс у них максимально сближен), какие-нибудь ещё изображения структуры, иллюстрирующие её особенности. Если будете делать это упражнение, обязательно снабдите картинки понятным описанием того, что именно на них изображено!

Выполненное задание присылать на адрес [akaryagina@gmail.com](mailto:akaryagina@gmail.com) в виде файла Word, вставив в текст рисунки (в подписях к рисункам необходимо привести название рисунка и расшифровать все обозначения). В теме письма указать «МФК Биоинформатика, задание по лекции 4, ФИО, факультет, курс».