

В.Ю.Лунин

Введение в кристаллографию макромолекул

Эти три лекции были прочитаны студентам второго курса факультета биоинженерии и биоинформатики МГУ осенью 2006. Цель лекций - неформальный рассказ о том, каким образом возникает информация о структуре белков, представленная в Protein Data Bank. Лекции не предполагали знаний математики, физики и химии, выходящих за пределы школьной программы.

Данный текст является записью звуковых комментариев к содержанию диапозитивов файла Lunin_MSU06_Pictures_2.pdf. Запись сделана в процессе подготовки лекций. Она не подвергалась литературной обработке и не является записью собственно лекции. Цель записи - обозначить круг вопросов, планируемых к обсуждению при показе того или иного слайда.

Цифры в тексте дают ссылку на номер слайда, соответствующего данному куску текста.

Лекция 2.

1

Напоминание о масштабе изучаемых объектов

2

Мы продолжаем сегодня обсуждать проблемы определения структуры белков при атомном разрешении. Характерный масштаб происходящих событий - это один ангстрем. Один ангстрем - это формально десять в минус десятой метра. Неформально - это расстояние между атомом водорода и соседним неводородным атомом. Длина валентной связи - порядка полутора ангстрем. 3 ангстрема - это длина водородной связи, расстояние между соседними цепочками в белковой глобуле. Пол-ангстрема - это радиус орбиты электрона в модели Бора для атома водорода. Если говорить о более мелких вещах, то размер ядра атома - это величины порядка десять в минус четвертой ангстрема, и на такой точности мы работать не будем, для нас ядро не имеет размеров. Длина волны рентгеновского излучения, которое мы используем для исследования объекта - это где-то в пределах от половины до полутора ангстрем. Поэтому один ангстрем - это такая характерная единица для тех объектов, с которыми мы работаем. Размер белковой глобулы для среднего размера белка равен в поперечнике где-то 30-40 ангстрем.

3

"С точки зрения рентгеновских лучей", объект задается распределением электронов в этом объекте. Для этого описания мы ввели функцию распределения электронной плотности, которая, грубо говоря, показывает для каждой точки, как часто в этой точке бывает электрон. Более формально - если мы имеем какой-то маленький объем с центром в этой точке, то произведение электронной плотности на этот объем дает средний по времени заряд, пробегающий через этот объем. Ну, а если мы хотим посчитать средний по

времени заряд в какой-то области, то надо, соответственно, посчитать интеграл по этой области от функции распределения электронной плотности. Поскольку электроны сосредоточены вблизи центров атомов, то понятно, что эта функция распределения электронной плотности для атомов структуры имеет большие значения вблизи центров атомов, и поэтому, если мы знаем эту функцию, мы можем определить положение атомов, глядя, где функция принимает большие значения. Вот здесь для такой функции распределения электронной плотности показана область, в которой функция принимает максимальные значения, и мы видим, что эта область распадается на такие шарики, отвечающие положениям атомов. Вообще говоря, если мы рассматриваем, как устроено облако электронной плотности, окружающее атомное ядро, то оно может быть устроено достаточно сложно. Области, в которой преимущественно находятся электроны, называется орбиталями. Эти орбитали для разных электронов могут быть разными, и устроены они достаточно сложно. Но в рамках метода рентгеноструктурного анализа мы пользуемся более простым описанием атома. Мы считаем, что электронная плотность в атоме является сферически симметричной, то есть значения плотности зависят только от расстояния от центра атома. И в этом случае область, в которой преимущественно находится электрон, имеет форму сферы..

4

Поскольку это распределение зависит только от расстояния от центра, то оно задается одномерной функцией. Вот здесь для атома серы нарисован график значения электронной плотности по мере убывания от центра атома.

5

Ну, если мы посмотрим другие атомы, то у них будут другие распределения электронной плотности. Мы видим, что для атома азота, кислорода и углерода (основных атомов, образующих белковую глобулу) они почти одинаковы. И мы видим, что для водорода электронная плотность совсем низкая, поэтому определение положения атомов водорода методами рентгеноструктурного анализа - это особенно сложная задача, и только в последнее время ее стали более или менее надежно решать.

6-7

Вид такой кривой наталкивает на мысль, что ее можно описать гауссовой функцией. Эту гауссову функцию принято записывать в таком виде. Оказывается, что одной такой гауссовой функции недостаточно, чтобы хорошо описать кривую, поэтому обычно она аппроксимируется пятью гауссовыми функциями, и соответствующие коэффициенты хранятся в кристаллографических таблицах.

Да, в рамках этого метода мы предполагаем, что для каждого атома мы знаем, как устроена его функция распределения электронной плотности. Эта информация получается другими методами, методами квантовой химии, например, ну, или некоторыми другими экспериментальными методами.

8

Рентгеновский эксперимент заключается в том, что мы помещаем исследуемый объект в пучок рентгеновских лучей и измеряем интенсивность электромагнитных волн, возникающих в разных направлениях, отличающихся от направления падающего пучка.

9

Направление падающего пучка мы будем характеризовать вектором единичной длины σ_0 , а вектором σ мы будем характеризовать направление на детектор, который измеряет интенсивность рассеянных лучей. Значение этой интенсивности зависит при фиксированном положении объекта от двух векторов (направления падающего пучка и направления на детектор). Ну, и, кроме того, поскольку в эксперименте объект вращается, то она зависит еще и от ориентации объекта.

Основные положения кинематической теории рассеяния

10

Мы говорили прошлый раз о том, что существует достаточно простая на сегодняшний день физическая теория, объясняющая, что происходит при попадании рентгеновских лучей на объект. Теория заключается в следующем: падающий рентгеновский пучок представляет собой плоскую электромагнитную волну. Когда эта волна попадает на электрон, электрон под воздействием волны начинает совершать колебания. Совершая колебания, электрон сам становится источником сферических электромагнитных волн, которые распространяются во всех направлениях. Значит, в результате попадания первичного пучка на электроны исследуемого объекта каждый электрон становится источником волны. И, если мы теперь где-то устанавливаем детектор, то этот детектор регистрирует сумму электромагнитных волн, сгенерированных всеми электронами (так называемая вторичная электромагнитная волна).

Сдвиг фаз

11

Дальше мы обсуждали, как выглядит эта суммарная волна, которая приходит на детектор. Мы говорили, что падающая волна является плоской электромагнитной волной, ее уравнение дается таким образом. Это величина вектора напряженности электрического поля для разных точек пространства в разные моменты времени. Эта волна записывается таким уравнением. Мы говорили, что волна, рассеянная одним электроном, имеет почти такой же вид. А если мы рассматриваем волну, рассеянную вторым электроном, то между этими двумя волнами возникает сдвиг фазы, который вызван вот чем: по отношению к падающему пучку первый и второй электроны находятся на разном расстоянии. Падающая волна приходит на второй электрон позже, чем на первый. Наоборот, если смотрим со стороны детектора, то волна от второго электрона к нам приходит раньше, чем от первого. Вот эти задержки, которые вызваны разностью положения электронов, они соответственно определяются проекцией вектора u на направление σ нулевое и проекцией вектора u на направление σ . Поскольку в одном случае у нас есть запаздывание, а в другом опережение, то они идут с разным знаком, и здесь мы получаем уравнение для волны, рассеянной вторым электроном - фазовый сдвиг, который вычисляется таким образом.

12

Теперь в зависимости от того, каким получился этот фазовый сдвиг, сумма двух электромагнитных волн может иметь очень разную форму. Ну, во-первых, существует общее правило, что сумма двух синусоид, различающихся только

сдвигом фазы, снова всегда будет синусоида, но амплитуда этой синусоиды будет разной в зависимости от того, какой здесь фазовый сдвиг. Если фазовый сдвиг равен нулю, понятно, что это будет просто удвоенная амплитуда, поскольку волны в фазе, они суммируются. Если сдвиг здесь будет равен половинке, то эти волны будут в противофазе, и они друг друга погасят, амплитуда суммарная будет равна нулю. Ну, а вот при промежуточных значениях этого фазового сдвига, соответственно, амплитуда будет меняться от удвоенной до нулевой.

Вектор рассеяния

Амплитуда суммарной рассеянной волны определяется распределением электронной плотности в образце, то есть она является результатом взаимодействия всех электронов, и она определяется направлением падающего пучка и направлением на детектор. На самом деле всегда эти направления присутствуют в виде из разности, поскольку нам важна разность фаз. Эта разность, деленная на длину волны, называется вектором рассеяния. Этот вектор рассеяния играет важную роль во всей дальнейшей теории.

Ну, поскольку я говорил - длина волны используемого рентгеновского излучения находится в районе одного ангстрема, то при первом слушании вообще можно считать, что длина волны у нас один ангстрем, и здесь - это просто разность этих величин.

13

Итак, в результате такого сложения от всех электронов мы получаем, что суммарная волна, получающаяся сложением всех электронов - это снова синусоидальная волна, которая имеет дополнительным множителем в амплитуде и сдвиг фазы. Вот этот дополнительный множитель и сдвиг фазы называются модулем и фазой структурного фактора. Модуль и фаза структурного фактора зависят от вектора рассеяния \mathbf{s} , то есть от этих двух направлений - падающий пучок и направление на детектор. И, кроме того, они зависят от того, как распределена электронная плотность в рассеивающем образце. То есть, мы говорили о том, что формулы для вычисления модуля и фазы структурного фактора получаются таким образом: мы должны просуммировать косинусы с весом $\rho(\mathbf{r})$, потом просуммировать синусы, и можем вычислить модуль таким вот образом.

То, что мы меряем в эксперименте - это интенсивность рассеянной волны, которая пропорциональна квадрату модуля структурного фактора.

Прямая и обратная задачи

Теперь, эти формулы дают нам возможность решать прямую задачу теории рассеяния, а именно: если мы знаем, как распределена электронная плотность в нашем образце, мы можем сказать, какова будет интенсивность рассеянной электромагнитной волны для любого направления \mathbf{s} .

14

Но нас больше интересует, естественно, обратная задача теории рассеяния: мы померили в эксперименте интенсивности рассеяния для разных направлений, и теперь хотим сказать, каково распределение электронной плотности в нашем образце. Один из возможных подходов к решению такой задачи - это метод

проб и ошибок. Он заключается в том, что выдвигается некоторая гипотеза о том, как расположены электроны (или как расположены атомы) в нашем образце. Это позволяет рассчитать функцию распределения электронной плотности. После этого по приведенным формулам мы можем рассчитать, какие в данном случае должны быть интенсивности рассеяния. И после этого мы их можем сравнить с экспериментом. И, если соответствие эксперименту хорошее, то все хорошо, а если нет - то надо, значит, выдвигать другую гипотезу. На самом деле такой метод определения структуры в реальности применялся на практике на заре кристаллографии, когда шла работа с небольшими соединениями, состоящими из нескольких атомов. Это выглядело так: специалисты по химии выдвигали гипотезу, как может быть устроена эта молекула, после чего проверялось, соответствует ли эта модель тому, что дал рентгеноструктурный анализ. Если эта модель не соответствовала, то химики говорили, что надо поправить, наверное, здесь вот так устроено. Таким образом, действительно, были решены структуры нескольких простейших веществ. Но понятно, что если, скажем, речь идет о молекуле белка, то сказать заранее, выдвинуть гипотезу, где находятся все десять тысяч атомов, входящих в состав молекулы белка - это занятие достаточно сложное. Поэтому здесь нужны какие-то другие методы.

Как фиксировать рассеянное излучение

15

Ну, прежде чем говорить, как все-таки мы можем практически решить эту обратную задачу, я хочу заметить следующую вещь:

хотя у нас и возникают во всех направлениях рассеянные волны, но интенсивность этих волн ничтожно мала. Если речь идет о рассеянии одной молекулой, даже молекулой белка, то никакая современная регистрирующая техника зафиксировать это не в состоянии. Как можно было бы пытаться бороться с этой проблемой?

Первое, понятно - увеличение мощности источника излучения, поскольку интенсивность рассеянной волны пропорциональна интенсивности падающей волны, но коэффициент этой пропорциональности очень маленький. В последние десятилетия в этом направлении был достигнут большой прогресс. Рентгеновские эксперименты стали проводить на синхротронных ускорителях, которые дают чрезвычайно мощный рентгеновский пучок. Но этого, оказывается, недостаточно.

Второй путь - это повышение чувствительности регистрирующего устройства. Такая работа тоже идет. Современные регистрирующие устройства - это очень сложная и дорогостоящая техника. Я приводил фотографию в прошлый раз. Но, тем не менее, даже этой пары - синхротронный ускоритель плюс очень чувствительный детектор - этого недостаточно для того, чтобы зафиксировать надежно интенсивность рассеянных лучей.

Следующий путь, по которому можно пойти для того, чтобы сделать это регистрируемым - это использовать для рассеяния одновременно большое число молекул исследуемого вещества. Понятно, что если у нас есть много молекул, то в принципе их рассеянные волны, складываясь, могут давать существенно большую интенсивность. Каким образом можно получить много молекул? Это может быть раствор, в котором содержится много молекул белка. Это могут быть газы. Это не белковая кристаллография, но для неорганических веществ

вещество может находиться в газообразном состоянии. Это порошки. И, наконец, это кристаллы.

Про кристаллы я скажу позже. Сейчас я просто замечу, что, когда мы говорим о растворах, газах или порошках, то в этом состоянии все молекулы находятся в разной ориентации. Их много, но они находятся в разной ориентации. Поэтому то, что мы можем измерить в эксперименте - это только такая некоторая усредненная характеристика, усредненная по всем возможным ориентациям молекулы. Такая информация тоже может быть очень важной. И существуют специальные методы исследования, которые занимаются именно исследованием веществ в растворах, газах, порошках. Но информация, которую можно извлечь - она достаточно грубая. Она сильно уступает тому, что удается получить в экспериментах с кристаллами.

Кристаллы

Что такое кристалл?

16

В рамках этого курса я буду говорить только о молекулярных кристаллах. В молекулярных кристаллах мы имеем много экземпляров одной и той же молекулы, расположенных в пространстве регулярным образом. Это называется идеальный монокристалл.

17-21

Вот мы имеем какую-то молекулу белка, и мы имеем еще много молекул белка, расположенных в одном направлении на равных промежутках друг от друга. Скажем, вектор **a** характеризует вот эту периодичность.

22-24

Дальше, такие ряды молекул расположены тоже регулярным образом во втором цикле. То есть здесь имеется второй вектор периодичности, и вот эти ряды также с равными промежутками расположены в этом направлении.

Ну и, вообще говоря, поскольку мы работаем с трехмерным объектом, то существует еще третье измерение, по которому такие вот уже плоскости молекул расположены тоже в регулярном порядке, но уже с третьим вектором периодичности **c**.

25-29

Надо заметить, что если мы имеем эти векторы периодичности **a** и **b**, то любая целочисленная комбинация этих векторов тоже будет периодом этого кристалла. Эти вектора **a** и **b**, вообще говоря, не обязаны быть перпендикулярными. То есть если мы имеем один слой, то второй слой может располагаться не прямо перпендикулярно первому, а под некоторым углом. Поэтому здесь, кроме величины векторов **a** и **b**, еще важен угол γ между ними. Ну и, в общем, в трехмерном случае эти три вектора могут быть не обязательно ортогональными. И полностью информация об этих периодичностях дается тремя длинами векторов и углами между этими векторами.

Элементарная ячейка

30

Параллелепипед, построенный на векторах периодичности, называется **элементарной ячейкой**. Длины векторов и углы называются параметрами

элементарной ячейки. Длины ребер элементарной ячейки обычно задаются в ангстремах, а углы задаются в градусах. Это важная информация о кристалле исследуемого вещества. И в .DB-файле она обязательно содержится. Для нее есть специальная карта CRYST1, где вот и приводятся размеры элементарной ячейки и углы между осями элементарной ячейки.

Следует заметить, что элементарная ячейка - это некоторый математический объект, который мы вводим для удобства описания. Если мы знаем распределение электронной плотности в элементарной ячейке, то мы дальше можем его восстановить во всем пространстве, поскольку, сдвигая на периоды этот элементарный кирпичик на вектора **a**, **b** и **c**, мы можем вымостить все пространство и получить распределение электронной плотности во всем пространстве. Но выбор элементарной ячейки, вообще говоря, неоднозначен. Во-первых, мы можем двигать начало элементарной ячейки как угодно, и это свойство будет сохраняться: размножая элементарную ячейку, мы будем получать все пространство. Дальше, как я уже говорил, выбор самих периодов решетки тоже может быть неоднозначен. Но здесь обычно стараются выбрать периоды решетки так, чтобы иметь элементарную ячейку минимального объема.

Различные виды координат

31

После того, как выбраны вот эти периоды и элементарная ячейка, можно ввести для каждой точки пространства ее относительные координаты как координаты в этой декартовой косоугольной системе координат. То есть каждый вектор может быть однозначно представлен в виде линейной комбинации векторов **a**, **b** и **c**. И вот коэффициенты этой комбинации называются **относительными координатами** точки. Понятно, что, пока мы находимся в пределах одной элементарной ячейки, относительные координаты всех точек лежат в пределах от нуля до единицы.

32

Но работать с такими относительными координатами достаточно неудобно. Особенно это было неудобно в докомпьютерную эпоху, когда вычисления приходилось производить вручную. Поэтому гораздо более удобными для описания структуры являются так называемые **абсолютные координаты**. Абсолютные координаты - это координаты в некоторой ортогональной декартовой системе координат, в которой все оси перпендикулярны друг другу. И в качестве единицы масштаба по всем осям выбран один ангстрем. В банке PDB координаты атомов приводятся именно в такой абсолютной системе координат. Существуют определенные правила, которые описаны в комментариях к формату PDB, как эта абсолютная система координат связывается с относительными координатами, с ребрами элементарной ячейки. Но, поскольку авторы иногда пытаются действовать не вполне каноническим образом, для этих целей в структуре PDB-файла предусмотрены дополнительные карты, записи которых описывают явным образом, каким образом от ортогональных можно перейти к относительным координатам.

33

Кроме того, чтобы сделать жизнь еще интересней, авторы иногда придумывают еще свою систему координат, которая им кажется более удобной, и депонируют координаты атомов в них. Чтобы опять-таки привязать это к некоторым

стандартным вещам, в PDB-файле могут присутствовать карты (они называются ORIX 1,2,3), которые уже показывают, как от депонированных перейти к стандартной абсолютной системе координат.

Надо заметить, что, как правило, структура депонируется в этих стандартных абсолютных координатах, и в этом случае здесь будет просто единичная матрица и нулевой вектор.

34-41

Следующее замечание заключается в том, что не надо думать, что приведенная в PDB-файле молекула всегда содержится **внутри** элементарной ячейки. Во-первых, как я уже сказал, начало элементарной ячейки мы можем сдвигать как угодно, и молекула может лежать вовсе не в центре ячейки, а где-то на краю. Но даже если бы мы хотели получить молекулу, лежащую строго в центре элементарной ячейки, мы в общем случае не можем сдвинуть начало координат элементарной ячейки так, чтобы удовлетворить этому требованию. Вот для примера давайте рассмотрим такую двумерную структуру. Вот она. Есть здесь кристаллическая решетка. И вот здесь мы имеем элементарную ячейку. Теперь, как бы мы эту элементарную ячейку ни двигали, при любом ее положении все равно у нас молекула никогда целиком в эту элементарную ячейку не попадет по очень простой причине - потому, что размер элементарной ячейки в этом направлении такой, а длина молекулы в этом направлении гораздо больше. Поэтому, как бы мы эту элементарную ячейку ни двигали, все равно у нас внутри нее всегда будут не одна молекула целиком лежать, а будут лежать куски разных молекул. Куски этих молекул будут лежать таким образом, что, сдвигая их на периоды, мы из них можем построить целую молекулу. Но, поскольку, естественно, было бы крайне неудобно работать со структурой белка, если бы она была приведена в виде трех кусков, относящихся к совершенно разным молекулам, то в PDB-файле приводятся всегда координаты одной молекулы белка, но при этом они вполне могут "вылезать" за пределы элементарной ячейки. Поэтому, если мы хотим посмотреть полностью содержимое молекулы, то нам часто надо брать для просмотра область большую, чем одна элементарная ячейка.

42

Суммируя информацию относительно элементарной ячейки, я еще раз хочу сказать, что элементарная ячейка - это математический объект. Она вводится для удобства работы. Такого физического объекта - элементарная ячейка - нет. Выбор элементарной ячейки и начала координат в значительной мере произвольный. При этом в начале координат вовсе не обязаны находиться какие-то атомы. И вообще молекула не обязана лежать в выбранной элементарной ячейке целиком.

Поскольку выбор элементарной ячейки и, тем самым, координат молекулы, которая дальше депонируется, достаточно произволен, то перед сравнением молекул двух веществ, решенных независимо и помещенных в PDB, эти структуры предварительно надо выровнять, приведя их к единой системе координат.

Преимущества работы с кристаллами

43

Теперь давайте обсудим вопрос, что же получается, если мы в пучок рентгеновских лучей помещаем кристалл. То есть мы помещаем много молекул,

которые уложены в строгом таком порядке, в одинаковой ориентации. Как мы говорили, для структурного фактора существует вот такая формула, по которой мы рассчитываем этот структурный фактор. Ну, давайте рассмотрим одну составляющую этой формулы, связанную с суммированием синусов.

Если у нас есть много атомов, расположенных на одинаковом расстоянии \mathbf{a} друг от друга, то в этом случае для того, чтобы просуммировать по этой цепочке атомов, мы должны просуммировать серию таких синусоид, которые имеют регулярный сдвиг друг относительно друга, и вот величина этого сдвига - она совпадает с проекцией вектора \mathbf{s} на направление \mathbf{a} . Вот этот сдвиг - это скалярное произведение вектора рассеяния и периода решетки \mathbf{a} . Что теперь происходит? Если, к примеру, этот сдвиг равен нулю, то у нас складывается сумма одинаковых синусоид, и мы получаем значительное усиление, мы получаем n -кратной высоты синусоиду, где n - число атомов в этой цепочке. Теперь, если вот это число δ - не целое, то у нас все синусоиды имеют какой-то сдвиг друг относительно друга, и если мы теперь для какой-то точки пытаемся просуммировать все эти синусоиды, то начинаем встречать положительные значения, отрицательные значения, близкие к нулю значения - то есть эти значения друг друга гасят, и мы, как правило, получаем в результате сложения достаточно малые числа, близкие к нулю. Я не буду выписывать сейчас формулу, просто покажу на рисунке, что происходит, если мы меняем эту величину дельта.

44

Давайте предположим, что атомы расположены в цепочку с шагом \mathbf{a} . Направление первичного пучка у нас совпадает с этим направлением \mathbf{a} . И мы начинаем детектор от направления первичного пучка двигать против часовой стрелки, увеличивая постепенно угол рассеяния. При этом у нас величина вектора рассеяния \mathbf{s} начинает возрастать, и, соответственно, его величина проекции на направление \mathbf{a} тоже начинает возрастать. Здесь вот показано, как меняется величина модуля структурного фактора при рассеянии такой цепочкой атомов по мере увеличения этого угла рассеяния. Здесь по оси отложена вот эта величина дельта - скалярное произведение вектора рассеяния на период решетки.

45-48

В случае рассеяния одним атомом мы имеем такую картину. Если мы в эту формулу добавляем второй атом, то зависимость становится вот такой. Если мы добавляем третий атом, то зависимость становится такой. Наконец, если мы включаем достаточно много атомов, то мы получаем вот такую картину. Как я говорил, если у нас дельта рано нулю, то у нас идет просто увеличение в n раз амплитуды синусоиды. Оказывается, это увеличение в n раз идет и для каждого из целочисленных значений. То есть для целочисленных значений величины модуля дельта мы получаем n -кратное усиление сигнала. А для значений, которые являются не целочисленными, мы сигнал на самом деле теряем. Мы видим, что вместо правильного рассеяния одним атомом здесь идет достаточно сумбурная осциллирующая картина низкой интенсивности. То есть если мы попробуем измерять рассеяние рассеянных волн такой цепочкой атомов, мы увидим, что в какие-то моменты мы получаем очень сильную рассеянную волну, причем интенсивность ее в n раз превосходит интенсивность рассеяния одним атомом. А если мы смотрим для других значений вектора рассеяния, то

мы не регистрируем никакую интенсивность, поскольку она идет на уровне экспериментальной ошибки.

49-50

Давайте еще раз рассмотрим эту картинку. Здесь есть цепочка атомов. Здесь отложено расстояние единица на a , и мы начинаем счетчик двигать по этой окружности против часовой стрелки. В тех случаях, когда направление на этот счетчик проекция на направление сигма ноль кратно величине единица, деленная на период решетки, у нас возникает вот это условие, что дельта - целое число, и мы получаем мощные рассеянные волны. А при других положениях счетчика мы никакой интенсивности зарегистрировать не можем, здесь интенсивность крайне низка.

Если мы теперь перейдем к трехмерной кристаллической решетке, то мы имеем периодичность во всех трех направлениях. И в результате мы получаем такую картину:

есть выделенные направления падающего и отраженного пучка (то есть выделенные значения вектора рассеяния s), при которых у нас происходит многократное усиление сигнала. Эти значения s , при которых у нас происходит усиление сигнала удовлетворяют трем уравнениям, которые называются уравнения Лауэ. И это условие заключается в том, что проекция вектора рассеяния на все три вектора элементарной ячейки должна давать целое число. Эти целые числа h, k, l называются индексами рефлекса. И они являются координатами того вектора рассеяния, для которого мы имеем сильный регистрируемый пучок. А для значений вектора рассеяния с нецелочисленными координатами интенсивность близка к нулю, и мы ее не можем зафиксировать. Поэтому на рентгенограммах мы видим такую вот картинку. Есть выделенные положения, при которых у нас возникает регистрируемая интенсивность, а в остальных случаях мы ничего измерить не можем.

51

Теперь давайте проведем небольшую оценку. Линейные размеры кристалла, которые используются в кристаллографии белка - это одна десятая миллиметра (характерный размер по всем направлениям), это десять в шестой ангстрем. Линейные размеры элементарной ячейки - ну, мы приблизительно можем их взять равными ста ангстрем, это белок среднего размера. Исходя из этого, мы можем посчитать, сколько копий молекулы находится в кристалле. Десять в четвертой в кубе - то есть в кристалле примерно десять в двенадцатой экземпляров молекулы. Поскольку мы регистрируем интенсивность, которая пропорциональна квадрату модуля структурного фактора, а модуль структурного фактора пропорционален числу атомов, то здесь мы получаем десять в двенадцатой в квадрате, то есть десять в двадцать четвертой раз. То есть использование такого кристаллика дает возможность для некоторых направлений рассеяния увеличить интенсивность рассеяния в десять в двадцать четвертой степени раз по сравнению с рассеянием одной молекулы. И вот эти десять в двадцать четвертой - это то, что дает возможность получить регистрируемую интенсивность рассеяния. Вот так бы мы не могли ее зафиксировать, и для других направлений мы ее не фиксируем, а здесь, в силу того, что у нас есть кристалл, мы можем в такое сумасшедшее число раз повысить интенсивность рассеяния и экспериментально ее зарегистрировать. Поэтому я хочу еще раз подчеркнуть, что в нашем случае кристалл используется как "умножитель" интенсивности рассеяния. Кристалл позволяет

многократно увеличить интенсивность рассеянных лучей. Причем коэффициент усиления сигнала одинаков для всех векторов \mathbf{s} с целочисленными индексами.

Недостатки работы с кристаллами

52

Недостатком является то, что мы можем увеличить интенсивность рассеянных лучей только в отдельных направлениях. При этом информация о рассеянии в направлениях с нецелочисленными индексами теряется, и ее мы получить никак не можем.

Однако, основной проблемой здесь является то, что для того, чтобы провести такой эксперимент, мы должны научиться получать кристаллы исследуемого объекта. Кристаллизация белка - это достаточно сложная процедура, потому что, вообще говоря, белок - это элемент живой природы, а кристаллическое состояние для живой природы - состояние крайне неестественное (хотя иногда и встречается). И поэтому, когда впервые начиналась белковая кристаллография, многие очень серьезные кристаллографы относились к идее белковой кристаллографии чрезвычайно скептически, замечая, что это совершенно бредовая идея - получать кристаллы белка. Но оказалось, тем не менее, что это возможно. Сейчас это получило широкое развитие. Десятки тысяч структур уже определены. Все-таки кристаллы белка получать можно. Хотя это проблема достаточно сложная и нерегулярная, и поэтому иногда даже приходится идти даже на всякие экстравагантные ухищрения, чтобы получить кристаллы белка, например, помещать раствор с молекулами белка в ультрацентрифугу, чтобы обеспечить сверхвысокое давление, или, наоборот, альтернативный вариант - поместить раствор с молекулами белка в невесомость, это кристаллизация на космических станциях. Туда отвозят белок в растворе, а оттуда привозят кристаллы, полученные в невесомости. Но это, так сказать, экзотика. А в нормальной жизни, в общем-то, кристаллизация идет при более или менее нормальных условиях. Хотя приходится перебирать очень много различных параметров, связанных с концентрациями используемых ингредиентов, и условиями, в которых идет эта кристаллизация.

Теперь второй вопрос, который сразу возник в связи с исследованием кристаллом - насколько структура белка в кристалле совпадает со структурой белка в растворе ну, и со структурой белка *in vivo*. Вопрос не праздный. Потому что, конечно, когда молекулы белка упаковываются в такую вот решетку, то на них действует некоторые силы кристаллические, которые заставляют их держаться в фиксированных позициях в этой решетке, и действие этих сил, вообще говоря, могло бы приводить к существенному искажению структуры. Но, по последним воззрениям, никаких существенных искажений структуры белка в кристалле не происходит. Во-первых, потому, что кристаллы белка - образования достаточно рыхлые. То есть в кристаллах белка примерно половина объема занята растворителем. Обычно это вода. То есть молекулы находятся в достаточно рыхлом состоянии, только чуть-чуть касаясь друг друга какими-то краями. Ну, и, кроме того, существуют эксперименты, в которых было продемонстрировано, что даже в кристаллическом состоянии молекулы ряда белков проявляют ферментативную активность. То есть, даже находясь в кристалле, они осуществляют ту биологическую функцию, которую они осуществляют в организме: катализируют какие-то реакции, что-то в таком духе делают. Поэтому в современном состоянии дел считается, что все-таки та

структура, которую мы получаем в исследовании кристалла - она достаточно близка к тому, как ведет себя белок *in vivo*.

53-54

Итак, суммируя сказанное:

- 1) рентгеновский эксперимент с монокристаллом позволяет измерить интенсивность волн, рассеянных в направлении, определяемых условиями Лауэ, то есть проекция на периоды элементарной ячейки должна быть целым числом;
- 2) вклады в амплитуду и фазу рассеянной волны могут быть рассчитаны по известным формулам; эти формулы я уже неоднократно приводил.

Описание структуры кристаллов

Центрированные ячейки

55

Теперь я хочу немного поговорить о науке под названием кристаллография, которая занимается описанием структуры кристаллов.

Вообще надо сказать, что использование монокристаллов сыграло решающую роль в рентгеноструктурном анализе, позволив определять атомные структуры вещества с использованием рассеяния рентгеновских лучей. Ввиду этой роли, которые сыграли кристаллы в этой науке, традиционно рентгеновское исследование структуры при атомном разрешении относится к кристаллографии. Хотя, я повторяю, кристалл здесь играет роль усилителя для рассеянных рентгеновских лучей. Но, поскольку это относится к кристаллографии, я немного поговорю о кристаллографии.

Как я уже говорил, периоды элементарной ячейки, вообще говоря, не обязаны быть перпендикулярны друг другу, и в общем случае мы имеем косоугольную ячейку, построенную на этих векторах.

56

Когда эта косоугольная ячейка выбрана так, что внутри этой ячейки никаких периодов не содержится, они совпадают с ребрами этой ячейки, мы говорим, что выбрана примитивная ячейка.

57

Однако бывает ситуация, когда мы можем выбрать более удобную для работы элементарную ячейку за счет того, что мы уже берем не минимально возможный объем элементарной ячейки, а мы берем увеличенный объем ячейки так, что некоторые периоды оказываются внутри этой ячейки, но зато ячейка становится ортогональной. Вот эта картинка соответствует тому же самому расположению атомов, но теперь мы взяли ячейку по-другому, и она стала ортогональной. Понятно, что в этом случае элементарная ячейка содержит избыточно количество информации. То есть на самом деле для того, чтобы восстановить распределение во всем пространстве, достаточно знать его только в половине точек элементарной ячейки. Но зато теперь мы можем работать с более простыми формулами, и описание структуры выглядит более просто.

58-60

То, что я нарисовал - это один из примеров. Такие ячейки называются центрированными. И, вообще говоря, существует некоторое количество ситуаций, в которых за счет выбора центрированной ячейки можно сделать ее

более удобной для работы. Такие типы центрирования имеют свои названия и свои буквенные обозначения - P, B, I, F. И существует всего 14 случаев (так называемые решетки Браве), когда за счет таких дополнительных трюков с искусственным увеличением элементарной ячейки удастся добиться от нее каких-нибудь удобных для работы свойств. Все эти случаи перечислены в соответствующих таблицах, я не буду на них останавливаться. Если нужно - их можно там найти. Я просто хочу еще раз обратить внимание, что каждый такой тип ячейки - он имеет свое буквенное обозначение, и этот тип ячейки - это информация, которая присутствует в карте CRYST1, описывающей кристаллическую структуру. Здесь, помимо длин ребер элементарной ячейки и углов, появляется еще эта буква, описывающая тип ячейки.

Симметрия в кристаллах

61

Следующее свойство кристаллов, на котором я хочу остановиться - это симметрия кристаллов.

Я первоначально сказал, что монокристалл - это когда все молекулы белка расположены в регулярном порядке и в одинаковой ориентации. "Это утверждение было не вполне корректно, потому что, вообще говоря, в кристалле молекулы могут находиться и в разных ориентациях.

62

Например, возможна такая ситуация, когда, кроме каждой молекулы, здесь содержится эта же самая молекула, повернутая на 180 градусов. То есть понятно, что ориентация этих молекул разная. В этом случае "кирпичиком", из которых построен кристалл, является пара таких молекул. То есть элементарная ячейка в этом случае содержит две молекулы, и вот внутри элементарной ячейки эти две молекулы связаны друг с другом поворотом на 180 градусов.

63

Соответственно, если мы теперь уже этот блок из двух молекул будет по периодам ячейки распространять, то мы покроем все пространство, и здесь "единицей" этого кристалла является эта пара молекул в разной ориентации, содержащаяся внутри элементарной ячейки. Мы можем это рассматривать как некоторый гиперобъект, состоящий из всех атомов двух молекул, и вот этот гиперобъект уже имеет такую регулярную периодичность, и дифракция от него устроена таким образом, что происходит многократное усиление рентгеновских лучей.

64-65

Такой элемент симметрии - поворот на 180 градусов вокруг оси, перпендикулярной плоскости, называется осью симметрии второго порядка. Она определяется таким значком. И эта ось симметрии означает, что для любой точки молекулы белка, если мы рассмотрим аналогичную точку, повернутую на 180 градусов, то мы получим совершенно аналогичный атом во второй молекуле белка. Но, поскольку у нас имеется кристалл, все свойства периодические, то, значит, если мы имеем ось симметрии здесь, то мы имеем такую же ось симметрии и во всех сдвинутых на периоды решетки точках пространства.

66

Ось симметрии называется кристаллографической осью симметрии, если она действует для всего кристалла. То есть если не только, повернув эту молекулу вокруг этой оси, мы получим идентичную молекулу, но и взяв любую молекулу как угодно далеко в кристалле и повернув ее вокруг этой оси на 180 градусов, мы снова попадем в точности на какую-то другую молекулу, лежащую в этом кристалле. Вот такая симметрия называется кристаллографической симметрией, и на самом деле вот такое условие, что она действует для всего кристалла, накладывает очень сильные ограничения на то, какие такие симметрии могут быть. Я скажу об этом чуть позже. Слово "кристаллографическая" здесь используется, чтобы подчеркнуть, что она действует именно для всего кристалла.

67-68

Вообще говоря, встречается ситуация, когда мы имеем симметрию, которая действует не для всего кристалла. Вот, к примеру, есть такой объект, который имеет ось симметрии пятого порядка. То есть если мы возьмем эту точку и повернем на 72 градуса, то мы получим идентичную точку. Но эта симметрия не действует, если мы возьмем точки на больших расстояниях от оси. Если мы возьмем вот эту точку повернем на 72 градуса, то мы уже получим точку с совсем другими свойствами. То есть на протяжении кристалла в целом такой симметрии нет. Она действует только локально, в пределах этого объекта. Такая симметрия называется некристаллографической симметрией. Мне больше нравится название локальная симметрия, поскольку это название позитивно - оно показывает, что это за симметрия. Но исторически так сложилось, что в названии присутствует негатив - некристаллографическая симметрия. И она тоже встречается в кристаллах.

69

Ну, вот, в качестве одного из последних примеров - это структура мембранного белка, магниевого канала. Здесь эта молекула состоит из пяти идентичных мономеров, которые связаны между собой осью симметрии пятого порядка. Поэтому, если мы посмотрим в торец на эту молекулу, то мы увидим такие картинки в зависимости от того, с какого конца на него смотреть. То есть здесь у нас есть ось симметрии пятого порядка. Но это ось симметрии некристаллографическая, то есть она действует только внутри этой молекулы.

70

Я возвращаюсь к кристаллографической симметрии. Сочетание того, что эта симметрия действует на протяжении всего кристалла, и периодичности электронной плотности, приводит к тому, что, помимо стартовой оси симметрии, здесь появляются дополнительные оси симметрии.

71

Нетрудно видеть, что, если бы мы поместили ось симметрии здесь, то, в силу периодичности, у нас ось симметрии появилась бы и здесь, поскольку вот эта молекула, в свою очередь, поворотом на 180 градусов вокруг этой оси связана вот с этой молекулой, которая получается отсюда просто переносом на период. И в результате у нас наличие одной оси симметрии плюс периодичность порождает еще много всяких осей симметрии, которые возникают в кристалле. Как я говорил, наличие периодичности накладывает сильные ограничения на возможные элементы симметрии. На самом деле в белковых кристаллах бывают только следующие виды симметрии:

72

это поворотная ось симметрии второго, третьего, четвертого или шестого порядка, и никакие другие (поворотная ось пятого порядка не может быть кристаллографической осью симметрии, равно как четырнадцатого или семнадцатого);

далее существуют такие элементы симметрии, как винтовые оси, то есть это когда мы поворачиваем на какой-то угол вокруг оси и сдвигаем вдоль этой оси. Оказывается, что винтовые оси симметрии тоже могут быть только ограниченного набора. Например, это поворотная ось на 180 градусов и сдвиг на полпериода. Или поворот на 120 градусов и сдвиг либо на треть периода, либо на две трети периода. Ну, и такие же винтовые оси четвертого или шестого порядка.

И больше на самом деле в кристаллах белков ничего быть не может.

Если рассматривать более общий случай - любых кристаллов, необязательно белковых, то здесь еще бывают дополнительные элементы симметрии - центр инверсии (это когда все отражается в некоторой точке), зеркальная плоскость (ну, как в зеркале структура отражается), и плоскость скольжения (это отражение в зеркальной плоскости плюс некоторый сдвиг вдоль этой зеркальной плоскости). Но эти последние три вида симметрии в кристаллах белков не встречаются. Как правило. Здесь опять-таки могут быть некоторые уточнения, но я о них сегодня не буду говорить.

73

Математически симметрия задается матрицей вращения и вектором трансляции и записывается вот таким образом. Координаты симметричной точки выражаются через координаты исходной точки посредством таких уравнений, где R и T - это матрица вращения и вектор трансляции.

Мы говорим, что распределение электронной плотности обладает кристаллографической симметрией, если электронная плотность для любой симметричной точки совпадает с электронной плотностью для исходной точки. И так для всех точек пространства.

Группы симметрии

Теперь, если мы рассмотрим все симметрии, которыми обладает конкретный кристалл, то эти симметрии образуют математический объект, который называется группой симметрии. Всего существует ровно 230 групп симметрии. Они все сведены в таблицы и описаны. Каждая из них имеет свое обозначение, по которому информацию о ней можно найти в соответствующих таблицах, и вот эта информация о группе тоже приводится в карте CRYST1 в соответствующей позиции.

74

Здесь полное название группы - это $P6_3$. Здесь P - это идентификация решетки, а 6_3 - это идентификация пространственной группы. Ну, и, кроме того, в карте CRYST1 приводится эта цифра, которая в нашем случае молекулярных кристаллов показывает количество молекул в элементарной ячейке. Здесь надо сделать маленькое замечание. Правильно говорить "количество мономеров", потому что под молекулой обычно имеют в виду то, что участвует в какой-то реакции, и молекула может иметь собственную ось симметрии, скажем, третьего порядка, которая является кристаллографической осью. То есть это на

самом деле количество мономеров, которые связаны между собой симметриями группы.

75

Теперь - что вносит кристаллографическая симметрия в рассеяние рентгеновских лучей? Ну, в общем, она вносит аналогичную симметрию в структурные факторы. Например, если у нас есть ось вращения второго порядка параллельно оси Y (то есть для симметричной точки координаты x и z просто меняют знаки, а координата y остается той же самой), то тогда возникает соответствующая симметрия в модулях структурных факторов, то есть для вектора рассеяния с h, k, l модуль будет таким же, как и для вектора рассеяния с моделями $-h, k, -l$. Что это позволяет? Ну, формально, с одной стороны, это позволяет сократить объем измеряемых экспериментальных данных, поскольку в таком случае не обязательно мерить все значения, поскольку часть может быть восстановлена по этой формуле. Ну, а на практике это измеряется для повышения точности измерений, поскольку, если мы измерили такой рефлекс, измерили такой рефлекс, и мы знаем, что они должны быть одинаковы, то мы можем взять среднее значение этих величин, поскольку при этом экспериментальная ошибка, вообще говоря, уменьшится.

Ряды Фурье

76

Теперь следующий раздел, к которому я перехожу - это ряды Фурье. Ряды Фурье сыграли поворотную роль в развитии рентгеноструктурного анализа. Как я говорил, изначально структуры решались методом проб и ошибок, то есть пытались угадать структуру такую, которая воспроизведет рентгеновский эксперимент. Использование рядов Фурье позволило обратить направление событий и позволило из эксперимента рассчитывать распределение электронной плотности, получать изображение структуры. Давайте поговорим немного о рядах Фурье.

77

Для функции одного переменного есть такие функции - $\sin(x)$ и $\cos(x)$ - которые отличаются только сдвигом вдоль оси x .

78-81

Теперь мы можем взять функцию $\sin(2\pi x)$ и $\cos(2\pi x)$, которые тоже синусоида, но просто которая колеблется в два раза чаще, на отрезке $0,1$ умещается два периода функции. Ну, соответственно, можно взять следующую функцию, которая еще быстрее колеблется, и так далее.

82

Замечательным фактом оказывается то, что любая функция на отрезке $0,1$ может быть представлена с какой угодно точностью в виде суммы таких синусов и косинусов. Что, вообще говоря, заранее является неочевидным. По крайней мере, многие столетия являлось неочевидным, поскольку это такие специфические функции. Но, тем не менее, оказывается, что любую функцию можно представить в таком вот виде.

83-86

То есть написать косинусы и синусы с какими-то коэффициентами, добавить косинусы и синусы удвоенного угла с какими-то коэффициентами, ну, и так далее. Давайте, я покажу на простом примере, что тут происходит.

Два примера

87-98

Давайте возьмем вот такую функцию, которая моделирует распределение электронной плотности в структуре, состоящей из трех атомов. Ну, если мы возьмем только один член в разложении, вот эту константу, мы получим какую-то постоянную функцию, которая понятно, что никакой еще информации о нашей функции не несет. Если мы добавим синус и косинус с такими коэффициентами, то мы получим вот такую функцию. Если мы добавим вторую пару синусов и косинусов вот с такими коэффициентами, то мы получим вот такую функцию, которая, конечно, еще очень далека от нашего настоящего распределения, но, тем не менее, уже передает его некоторые черты: мы имеем группу атомов, находящихся в этом месте, и не имеем ничего вот в этом участке элементарной ячейки. Добавим третью, четвертую. Добавление четвертой пары уже нам дает информацию, что на самом деле тут есть какое-то понижение в электронной плотности между этими двумя группами. Пять, шесть, семь... Восемь, десять, пятнадцать...

99

Ну, и вот мы видим, что здесь на самом деле мы уже с очень хорошей точностью описали наше исходное распределение электронной плотности при помощи вот этих функций синусов и косинусов. То есть я подчеркиваю, что здесь мы получили практически нулевую плотность, хотя функции синус и косинус на всем отрезке $0,1$ совершают колебания. Но за счет подбора коэффициентов мы сумели полностью подавить сигнал на этом участке и оставить правильный сигнал на этом. Но, впрочем, если посмотрим это в увеличенном виде, мы увидим, что здесь, конечно, идеального нуля у нас не получилось, но все равно, по сравнению с большими значениями, мы видим вполне удовлетворительное описание этой функции.

100-108

Давайте посмотрим другой пример аппроксимации более сложного контура в виде ряда Фурье. Вот мы наращиваем количество гармоник. Ну и, вот, получилось в целом уже вполне неплохо. Хотя, как вы видите, конечно, некоторые дефекты изображения здесь видны - вот эти колебания. Но это связано с тем, что этот базис функций из синусов и косинусов - он плохо приспособлен для того, чтобы передавать прямолинейные участки и вот такие вертикальные участки. Поэтому на самом деле для того, чтобы передавать изображение, где есть такие прямолинейные участки и горизонтальные или вертикальные участки, вообще говоря, существуют другие базисы функций, по которым тоже можно раскладывать исходную функцию.

Разложение произвольной функции

109

Но это я привел примеры. А теперь встает вопрос - если у нас есть какая-то функция и мы хотели бы ее таким образом представить в виде синусов и косинусов - что для этого надо сделать? Ответ для этого очень простой. Надо просто рассчитать соответствующие коэффициенты по известным формулам. Существуют достаточно простые формулы, которые показывают, как нам надо посчитать вот эти коэффициенты, с которыми брать синусы и косинусы, чтобы воспроизвести нашу функцию. То есть они просто берут и считают.

110

Теперь вот этот ряд Фурье для дальнейшего использования я напишу в несколько более компактном виде. Он записывается вот так в виде сумм синусов и косинусов, где коэффициенты A и B вычисляются по таким формулам. Ну, и вот эту пару синусов и косинусов можно свести к одному косинусу, введя такой модуль перед косинусом - модуль коэффициента Фурье и фазу коэффициента Фурье, введя ее под аргумент косинуса. Это общая формула. И в таком случае модуль и фаза вычисляются через коэффициенты A и B соответствующим образом. И любую функцию мы можем записать уже вот в таком виде.

111

Вообще говоря, вот эта идея с разложениями функции, аппроксимации функции по каким-то базисам - она очень широко применяется в технике для задач сжатия информации. В задачах сжатия информации возникают два типичных примера. Это передача сигналов на большие расстояния, когда стоимость передачи сигнала высока. Ну, например, какая-нибудь космическая станция получает снимок какого-то объекта, и если его передавать полностью, пиксель за пикселем на Землю, это будет очень дорого и долго. Поэтому на бортовом компьютере производится аппроксимация этой функции в виде ряда по какому-то базису, передается только небольшое число коэффициентов, а здесь на Земле можно при помощи другого компьютера восстановить исходное изображение, посчитав вот эту сумму ряда Фурье, например.

Ну, и второй пример - это хранение информации, когда у нас ограничено место для хранения информации. Вместо того, чтобы хранить информацию в исходном виде (ну, например, по пикселям), применяются форматы, которые позволяют хранить в гораздо более сжатом виде эту информацию. Ну, вот, из таких, с чем вы каждый день встречаетесь - это MP3-форматы для звуковых файлов. Или, скажем, JPEG-формат для графических файлов. Это не обязательно разложение именно по системе гармоник Фурье, но идея здесь та же самая - выбирается некоторый базис, и, вместо того, чтобы хранить исходную информацию, хранятся и передаются коэффициенты в разложении функции по этому базису.

112-115

Ну, это мы говорили пока про функцию одного переменного. А что, если у нас функция зависит от нескольких переменных, например, от трех пространственных переменных x , y , z ? Здесь ситуация на самом деле точно такая же. Только здесь мы уже должны взять такую трехмерную систему базисных функций. Мы должны, например, взять косинусы и синусы вот такие, где уже каждая из этих функций зависит от трех переменных x , y , z и определяется тремя индексами h , k , l . Ну, и по-прежнему функцию можно разложить по такой системе функций. Ну, или в более сжатом виде. Где коэффициенты эти вычисляются по похожим формулам, но только здесь уже идут интегралы по трехмерному пространству, соответственно, функций трех переменных.

116

Я напоминаю, что некоторое время назад мы говорили о том, что в процессе рентгеновского эксперимента мы измеряем интенсивность волн, рассеянных в направлениях, определенных условиями Лауэ, и что вклад в амплитуду и фазу

рассеянной волны могут быть рассчитаны по распределению электронной плотности по таким формулам.

117

Смотрите, что получилось. Формулы, которые мы получили для коэффициентов Фурье - это в точности те структурные факторы, которые у нас возникают в рентгеновском эксперименте. То есть рентгеновский дифракционный эксперимент с монокристаллом позволяет измерить модули коэффициентов в разложении функции распределения электронной плотности в ряд Фурье.

118

Если бы мы еще каким-то образом сумели измерить или определить фазы структурных факторов, то тогда проблема восстановления электронной плотности была бы крайне простой. Мы могли бы просто восстановить эту плотность, рассчитав ее как сумму соответствующего ряда Фурье. То есть вот с этой точки зрения мы можем сказать так: рентгеновский эксперимент шифрует информацию о распределении электронной плотности следующим образом - он посылает в разных направлениях рассеянные волны, при этом модуль коэффициента Фурье шифруется амплитудой рассеянной волны, а фаза шифруется фазой рассеянной волны. Если бы мы могли измерить всю эту информацию, которую нам посылает рентгеновский эксперимент - и модули, и фазы, - то мы могли бы дешифровать этот сигнал, просто посчитав сумму соответствующего ряда Фурье и получив синтез Фурье электронной плотности. Тогда мы бы могли посмотреть соответствующие карты - как он устроен, и интерпретировать его в терминах структуры (определить, где находятся атомы).

119

Вот такая сумма ряда Фурье, посчитанная по экспериментальным структурным факторам, называется синтезом Фурье электронной плотности. Главное ее отличие от настоящего распределения электронной плотности в том, что в настоящем распределении электронной плотности, вообще говоря, подразумевается бесконечная сумма слагаемых в этом разложении. А здесь, когда мы рассчитываем синтез Фурье, мы используем только то, что мы реально в состоянии померить в эксперименте. И, в зависимости от того, как много мы смогли померить вот этих структурных факторов, соответственно, точность воспроизведения распределения электронной плотности может быть очень разной.

Разрешение

120

Пока что я говорил про разложение в ряд Фурье функции, определенной на отрезке $0,1$. Но что происходит, если нас интересует некоторое распределение, скажем, не на отрезке $0,1$, а на отрезке $0,a$. Ну, в таком случае мы можем ввести относительную координату x , задав каждое значение r , как x , умноженное на a . При этом значение относительной координаты x у нас будет между нулем и единицей. И мы можем уже вот эту функцию $\rho(x)$ разложить на отрезке $0,1$. Теперь, когда мы говорим о синтезе Фурье электронной плотности, для нас очень важен вопрос, как много гармоник Фурье мы включили в расчет синтеза. (вот эти синусы и косинусы - они иногда называются гармониками Фурье), поскольку от этого зависит качество получаемого изображения. И здесь возникает важное понятие - разрешение синтеза Фурье.

121

Сначала мы поговорим о характеристике, которая относится к каждой конкретной гармонике Фурье - это разрешение, отвечающее этой гармонике. Давайте сначала посмотрим функцию одного переменного, которая устроена таким вот образом. Поскольку мы перешли к относительным координатам, она запишется вот таким образом.

Если мы нарисуем график этой функции, то вот это расстояние между соседними максимумами - оно равно a , деленное на h . И вот в белковой кристаллографии эта величина называется разрешением, отвечающим этой гармонике. То есть в макромолекулярной кристаллографии разрешение - это расстояние между двумя максимумами, и оно показывает, на самом деле, размер деталей, которые можно воспроизвести при помощи таких гармоник Фурье.

Теперь, если у нас здесь расстояние d , то на один ангстрем у нас будет приходится единица, деленная на d , таких атомов. В низкомолекулярной кристаллографии разрешением обычно называется эта обратная величина. Это величина - единица, деленная на d , т.е. число максимумов, приходящихся на один ангстрем. Это определение, ближе к тому, что применяется при обработке фотоизображений - количество пикселей на дюйм или чего-нибудь в таком духе. Здесь это количество максимумов на один ангстрем.

Но исторически так сложилось, что в кристаллографии белка, в макромолекулярной кристаллографии разрешением называется именно величина, измеряемая в ангстремах. Это расстояние между двумя соседними максимумами.

122

Теперь мы говорим, что синтез Фурье рассчитан с разрешением d минимальное, если все или почти все члены соответствующего ряда Фурье со значениями d , большими или равными d минимальному, включены в расчет.

Давайте на этом модельном примере посмотрим, какое разрешение дает какие картинки.

123

Вот здесь для определенности эти два атома взяты на расстоянии полтора ангстрема, а эти два на расстоянии 3 ангстрема. При разрешении 10 ангстрем мы видим вот такую картину. Это картина разрешения 5 ангстрем. Эта картина разрешения 2 ангстрема. Ну, наконец, эта картина разрешения 1.3 ангстрема. Мы видим, что при разрешении, скажем, 5 ангстрем мы видим отдельные пики, хотя и не очень ярко выраженные, две группы атомов, находящихся на расстоянии 3-4 ангстрема, но мы видим как слившийся пик атомы на расстоянии полтора ангстрема. То есть такая разрешающая способность синтеза - это примерно половина номинальной величины разрешения синтеза. То есть если у нас разрешение 5 ангстрем, то пики, которые находятся между собой ближе, чем два с половиной ангстрема, мы уже сидим как слившуюся картинку. А если расстояние между атомами больше, то мы видим их как отдельные разделенные пики, хотя и разной высоты.

124-125

Это была показана одномерная картинка. Когда мы говорим про трехмерное изображение, то мы работаем с такими гармониками Фурье, которые формально являются функциями трех переменных, но которые опять-таки, как электромагнитные волны, на самом деле являются одномерными. Здесь вот есть вектор рассеяния s , и вдоль этого вектора рассеяния вот эта функция меняется

как синусоидальная функция, и расстояние между этими двумя максимумами равно единице, деленной на модуль s . А в любой плоскости, перпендикулярной этому направлению, значения функции одни и те же. Поэтому можно представить, как выглядят эти слагаемые, отдельные гармоники Фурье. Это функции, которые постоянны на таких плоскостях, перпендикулярных направлению s , и меняются только вдоль направления s . Разрешение d - это расстояние между плоскостями, отвечающими двум максимумам - это соседние максимумы этой функции.

126

Здесь, на этой картинке, приведены примеры синтезов Фурье разного разрешения для небольшой молекулы: Protein G. В этом тесте по известной структуре белка были посчитаны абсолютно точные модули и фазы структурных факторов, и для них потом были построены синтезы Фурье, но при этом для разных картинок в расчет синтеза Фурье включалось разное количество гармоник Фурье именно в соответствии с заказанным разрешением. Если мы рассчитываем синтез Фурье разрешением 16\AA , то мы, когда смотрим на эти области высокой электронной плотности, мы видим, что все, что здесь на самом деле мы можем увидеть - это некоторую глобулу, которая передает нам место, положение молекулы в элементарной ячейке и какие-то совсем грубые внешние очертания этой молекулы. Если мы повышаем разрешение до восьми ангстрем - здесь обычно появляется возможность различить альфа-спирали. Они изображаются такими цилиндрическими вытянутыми областями. При разрешении 3.5\AA становится возможным проследить ход полипептидной цепи. Мы видим, здесь как идет полипептидная цепь, и появляется информация в виде таких вот облаков о положениях боковых групп. То есть опять-таки отдельные атомы мы при таком разрешении увидеть не можем, но мы видим, скажем, появившееся большое облако для триптофана. Или вот здесь появившееся облако для тирозина. Ну и, наконец, если разрешение синтезов удалось поднять, скажем, до разрешения 1 ангстрем, то на таком разрешении мы уже начинаем видеть каждый атом как отдельный шарик, и такие синтезы, конечно, совсем их легко интерпретировать, поскольку здесь все атомы становятся видны.

127

Ну, вот, я еще раз хочу показать примеры синтезов, которые я вам уже показывал. Первый синтез - это синтез белка альдоз-редуктаза, посчитанный при разрешении 0.9\AA . И здесь мы действительно видим атомы в виде шариков. А это стартовый синтез разрешения 3\AA для белка Protein G, с которого начиналась работа по определению этого белка. И здесь мы видим, что, конечно, качество синтеза гораздо более бедное. Но, тем не менее, мы видим ход полипептидной цепи и видим, по крайней мере, для больших боковых групп, места, где расположены эти боковые группы.