

Рентгеноструктурный анализ белков

Е. Горячева



*Екатерина Горячева
13-й выпуск биокласса
(Кроссастеры), школа
№ 520 (1991 г.), закончила
кафедру биофизики Биофа-
ка МГУ (1996 г.), к.б.н.,
сотрудник Института биоор-
ганической химии РАН (лаб.
рентгеноструктурного ана-
лиза), goryacheva@ibch.ru*

В природе встречается примерно 10^{12} различных белков, выполняющих самые разнообразные функции. Это и белки-ферменты, катализирующие большинство биохимических процессов в живой клетке; и белки-переносчики, позволяющие другим молекулам проходить через ядерные или клеточные мембраны или перемещаться между клетками всего организма (перенос кислорода гемоглобином); и иммуноглобулярные белки, отличающиеся высокой специфичностью взаимодействия с антигенами, что приводит к активации сигнальных путей, обеспечивающих иммунный ответ клеток. Это лишь несколько примеров уникальных свойств белковых молекул. По образному выражению Фрэнсиса Крика, белки важны прежде всего потому, что они могут выполнять самые разнообразные функции, причем с необыкновенной легкостью и изяществом.

При всем своем структурном и функциональном многообразии все природные белки построены из 20 аминокислот, соединенных в соответствии с кодом белкового синтеза. В зависимости от последовательности аминокислотных остатков в полипептидной цепи формируется определенная стабильная трехмерная структура белка, определяющая его структурные и функциональные свойства. Например, для каждого фермента характерна уникальная пространственная конформация его активного центра, обеспечивающего взаимодействие со строго определенными молекулами субстратов и осуществляющего каталитический акт. Причем, для эффективного образования фермент-субстратного комплекса большое значение имеет не только геометрическая комплементарность молекул фермента и субстрата, но и образование

водородных связей, электростатические и гидрофобные взаимодействия между атомами активного центра фермента и молекулы субстрата. Таким образом, любая белковая молекула характеризуется уникальностью структуры, которая определяет уникальность ее функции.

Выяснение пространственной организации белков является одним из основных направлений современной биохимии. Оно имеет важное научно-практическое значение для понимания огромного разнообразия функций белков, выполняемых ими в живых организмах. Во многих случаях знание структуры белка и его комплекса с ингибиторами является решающим фактором при создании лекарственных препаратов.

На сегодняшний день единственным экспериментальным способом с атомарной точностью узнать, что представляет собой трехмерная структура белка, является рентгеноструктурный, или кристаллографический, анализ, позволяющий определить пространственные координаты всех атомов исследуемого объекта. При наличии данных о положении отдельных атомов можно вычислить межатомные расстояния, валентные углы, углы вращения вокруг связей, распределение поверхностного заряда и другие детали молекулярной геометрии. Все эти данные особенно интересны химикам, биохимикам и биологам, изучающим зависимости между структурными характеристиками и функциональными свойствами, а также специалистам, занимающимся электронной структурой и взаимодействиями между молекулами.

Определение структур макромолекул имеет не только фундаментальное значение, но и большое прикладное применение. Рентгеновская кристаллография широко используется для дизайна и направленной модификации молекул с целью дальнейшего их использования в медицине и промышленности. Например, знание структуры белка и его комплекса с ингибиторами является решающим фактором при создании лекарственных препаратов.

Рентгеноструктурный анализ занимает особое место среди экспериментальных методов, и постоянный интерес к нему отражен в том факте, что с момента открытия рентгеновских лучей в 1901 г. по настоящее время работы в этой области 12 раз отмечались Нобелевскими премиями.

Применение рентгеноструктурного анализа для исследования сложноорганизованных биологических объектов началось после 1953 г., когда сотрудник Кавендишской лаборатории Кембриджского университета Макс Перутц нашел способ определения структуры крупных молекул, таких как миоглобин и гемоглобин. С тех пор рентгеноструктурный анализ молекул белка помогает нам понять химию биологических реакций.

На сегодняшний день известны структуры около 15 тыс. белков и их комплексов с биологически важными молекулами.

Рентгеновский эксперимент

Подобно радиоволнам и видимому свету, рентгеновские лучи являются электромагнитными волнами, которые занимают диапазон длин волн примерно 0.001–10 нм. С коротковолновой стороны они соседствуют с γ -лучами (длины волн примерно 0.01–0.0001 нм), с длинноволновой — с ультрафиолетовыми (длины волн примерно 10–380 нм).

Одним из необходимых условий проведения рентгеновского эксперимента является монохроматическое рентгеновское излучение (т.е. строго определенной длины волны). Для этой цели используются различные фильтры и монохроматоры.

Рассмотрим, каким образом происходит исследование структуры объекта в рентгеновском эксперименте. Обычно, когда нормальный человек слышит о рентгеновском исследовании, у него возникает образ рентгеновского кабинета в поликлинике. На самом деле рентгеноструктурный анализ не имеет ничего общего с такими исследованиями. Основой медицинской рентгеноскопии является различие в степени поглощения рентгеновских лучей различными тканями. В медицине рентгеновский снимок содержит **прямое** изображение исследуемого объекта, в то время как рентгеновская кристаллография основана на рассеянии рентгеновских лучей электронами атомов, и фотографические снимки, получаемые в нем, **не содержат никакого изображения чего бы то ни было**. Как же ставится рентгеновский эксперимент? Принципиальная схема проста (рисунок 1). Исследуемый объект помещают в пучок рентгеновских лучей, и измеряют интенсивность **рассеянного** в различных направлениях излучения. Самый простой способ — поместить на пути фотопленку и по степени потемнения пятна после проявления судить об интенсивности рассеяния в этом направлении. (Конечно,

на сегодняшний день существуют и более совершенные методы, но сейчас это неважно.) В данном случае важно то, что мы смотрим не на интенсивность лучей, прошедших сквозь объект, а на интенсивность лучей, возникших там, где их вроде бы и не должно было быть.

Итак, на входе мы имеем неизвестный объект, на выходе — набор интенсивностей, рассеянных в различных направлениях лучей. Дальнейшая наша задача — связать полученную в эксперименте информацию с атомной структурой исследуемого объекта. Перечислим основные положения, на которых строится простейшая *математическая модель рассеяния рентгеновских лучей*:

- 1) пучок рентгеновских лучей является плоской монохроматической электромагнитной волной;
- 2) под воздействием этой электромагнитной волны каждый электрон приходит в движение; при этом это движение может быть описано уравнениями для свободных зарядов;
- 3) движущийся электрон является, в свою очередь, источником новой рассеянной сферической электромагнитной волны, распространяющейся во всех направлениях;
- 4) эти новые волны суммируются и определяют интенсивность излучения в интересующем нас направлении.

Такая модель называется кинематической теорией рассеяния. Ее основной недочет заключается в том, что на электрон действует не только первичный пучок, но и рассеянные волны, и их влияние может изменять характер его движения. Попытка учесть эти поправки делается в более изощренной динамической теории рассеяния, однако для практических приложений более простая кинематическая теория рассеяния оказывается, как правило, вполне хороша.

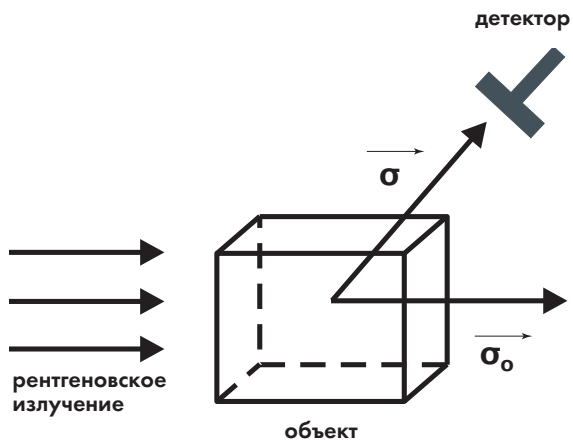


Рисунок 1. Схема рентгеновского эксперимента.

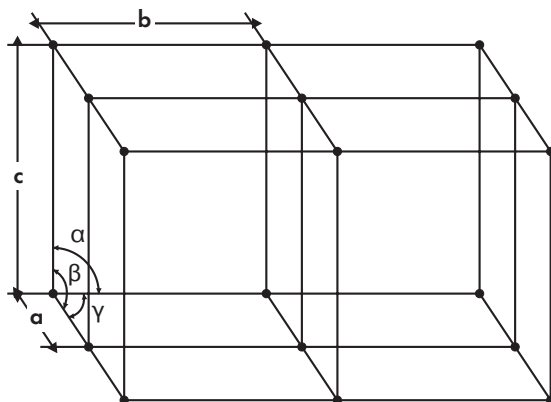


Рисунок 2. Модель кристаллической решетки.

Необходимость кристаллизации исследуемых объектов

Метод рентгеноструктурного анализа основан на дифракции рентгеновских лучей на кристаллической решетке и поэтому применим только к веществам в кристаллическом состоянии. Это связано с тем, что для регистрации дифракционной картины рассеяния необходимо иметь достаточное количество рассеивающих электронов. Но если образец состоит из большого числа идентичных молекул, расположенных в произвольной ориентации (раствор), то картина рассеяния будет определяться какими-то усредненными по всевозможным ориентациям характеристиками и вряд ли позволит получить детальную информацию об атомной структуре. Другое дело, если большое количество одинаковых экземпляров одной и той же молекулы помещены в одной и той же ориентации. Такую возможность дают нам кристаллические образцы. Говоря простыми словами (и не вдаваясь в сложные математические формулировки), кристалл — это такой образец исследуемого вещества, в котором много ($\sim 10^{12}$) идентичных молекул находятся в одинаковой ориентации, и их центры образуют правильную трехмерную решетку.

Основная особенность структуры каждого кристалла состоит в том, что он построен из регулярно расположенных в пространстве отдельных атомов или групп атомов. Если каждую повторяющуюся структурную единицу заменить точкой, или узлом, то получится трехмерная кристаллическая решетка (рисунок 2). Решетку можно представить себе как систему одинаковых параллелепипедов. Каждый такой параллелепипед носит название «элементарная ячейка кристалла», которая описывается шестью параметрами: длинами ребер (a , b , c) и углами между ними (α , β , γ).

Одной из основных претензий к методу рентгеноструктурного анализа с самого начала исследования структур белков являлась та, что в жизни белки находятся в растворе, а при исследовании мы их кристаллизуем. Задавался логичный вопрос: не деформируется ли структура белковой глобулы при кристаллизации настолько, что структура белка в кристалле перестает иметь что бы то ни было общее со структурой белка в функционально активном состоянии? В современных исследованиях принято считать, что сильных искажений все-таки не происходит. Некоторые доводы в пользу такой позиции следующие. Во-первых, ряд белков сохраняет ферментативную активность и в закристаллизованном состоянии, т.е. структура портится не настолько, чтобы белок стал «неработоспособен». Другое соображение — в кристаллах биомакромолекул значительный объем (от 30 до 80%) занимает растворитель, т.е. упаковка молекул белка в кристалле неплотная и вряд ли вызывает существенные искажения. Некоторые искажения в свободных петлях, наверное, возможны, но структура активного центра сохраняется. Еще одно подтверждение — альтернативное определение структур некоторых белков методом двумерного ядерного магнитного резонанса не дало существенных расхождений со структурами, расшифрованными рентгеновскими методами.

Результат эксперимента и его связь с исследуемой структурой

Монохроматическое рентгеновское излучение, проходя через кристалл, рассеивается, в основном, на электронных оболочках периодически повторяющихся атомов и образует дифракционную картину, или рентгенограмму (рисунок 3). Поэтому экспериментальные рентгеновские данные позволяют судить об особенностях расположения электронов в элемен-

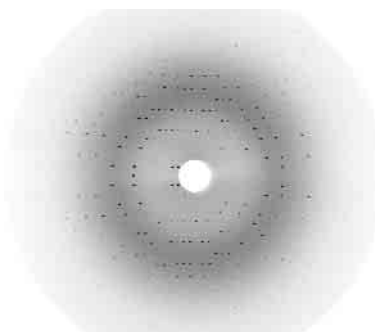


Рисунок 3. В дифракционной картине заключена вся информация о структуре белка.

тарных кристаллических ячейках. Электрон обладает волновыми свойствами, и его положение в пространстве характеризуется не точными координатами, а функцией распределения электронной плотности $\rho(r)$, которая дает среднее по времени число электронов, приходящееся на 1 \AA^3 (кубический Ангстрем). На основании этой функции можно судить о расположении атомов в элементарных ячейках, т.к. каждому атому соответствует сгусток электронной плотности определенной величины. Таким образом, при обработке данных рентгеновского эксперимента нужно решить две задачи:

1. Из данных рентгенограммы получить карту распределения электронной плотности $\rho(r)$ в кристалле исследуемого объекта. На этом этапе возникает принципиальная трудность (о которой речь пойдет ниже), связанная с невозможностью получить из эксперимента всю информацию, необходимую для восстановления исследуемой структуры. Для получения недостающей части информации используют различные обходные пути. Но универсального пути нет, и в каждом случае исследователь выбирает наиболее подходящий, основываясь на своем опыте и интуиции.
2. На основании карты распределения электронной плотности $\rho(r)$ определить положения атомов в исследуемом объекте. Для решения этой задачи структура многократно подвергается программной обработке и ручной доводке для достижения наилучшего совпадения с электронной плотностью.

Основные этапы определения структуры белка

Выделение, очистка

С этого этапа начинаются практически все экспериментальные исследования белковых структур. Для получения нужного белка используют различные биохимические методы. Последовательность операций по выделению белков обычно сводится к измельчению биологического материала (гомогенизация), извлечению белков, а точнее — переводу белков в растворенное состояние (экстракция), и выделению исследуемого белка из смеси других белков, т.е. очистке и получению индивидуального белка. На этом этапе наибольшая сложность заключается в наработке достаточного для эксперимента количества чистого белка.

Кристаллизация

Получение кристаллов, пригодных для рентгеноструктурного анализа, — зачастую (особенно для сложных соединений, таких как белки и

нуклеиновые кислоты) процесс трудоемкий и далеко не тривиальный. Наличие пересыщенного раствора — необходимое условие кристаллизации. Для получения такого раствора используют различные способы. Один из них заключается в постепенном удалении растворителя обычным испарением, что приводит к росту концентрации вещества в растворе, который в какой-то момент становится пересыщенным. Другой способ связан с использованием температурной зависимости растворимости. Например, если растворимость с увеличением температуры повышается, можно приготовить насыщенный раствор при более высокой температуре, а затем медленно охладить его. Благодаря понижению растворимости в процессе охлаждения получается пересыщенный раствор. Третий способ связан с введением в раствор какого-либо вещества, вызывающего понижение растворимости. В качестве таких веществ используют либо соли, либо органические растворители. Кроме того, растворимость белков и нуклеиновых кислот сильно зависит от рН раствора, это тоже можно использовать для получения пересыщенных растворов.

Это — в теории. На практике все намного сложнее. До сих пор не существует универсальных способов подбора оптимальных условий кристаллизации. Для каждого конкретного белка исследователь ищет эти условия, меняя тип буфера, значения рН, температуры, концентрации самого белка, осаждающей соли и т.д. В этой работе важно найти такие условия, при которых получится именно кристалл, а не выпадет соль. Поэтому выращивание биологических кристаллов — не только научное направление, но и большое искусство. Иногда, чтобы заставить белок кристаллизоваться, его помещают в экзотические условия: например центрифугируют или отправляют в невесомость.

Выбор кристаллов для рентгеновского эксперимента производят с помощью микроскопа. Для этой цели особенно полезен поляризационный микроскоп, позволяющий с помощью поляризационного света установить наличие дефектов в кристалле. Оптимальными считаются монокристаллы с размером каждой из сторон 0,2–0,6 мм. Кристаллы должны быть без дефектов и, по возможности, с хорошей огранкой. Наличие дефектов приводит к ошибкам при экспериментальном измерении дифракционной картины и, как следствие, к неточности (а часто и к невозможности) расшифровки кристаллической структуры. При повышении сложности исследуемого объекта требования к качеству кристаллов повышается. Как выглядят кристаллы белков, показано на рисунке 4.

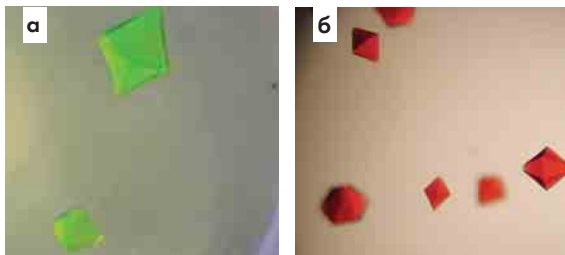


Рисунок 4. Кристаллы белков.
(а) — кристаллы зеленого флуоресцентного белка zGFP506; (б) — кристаллы мутанта белка zGFP506 с аминокислотной заменой N66D.

К сожалению, далеко не из всех белков удается вырастить кристалл, поэтому этот этап является главным ограничением метода рентгеноструктурного анализа макромолекул.

Рентгеновский эксперимент, обработка результатов

Это, пожалуй, самый дорогостоящий эксперимент в молекулярной биологии. В качестве источника рентгеновских лучей в настоящее время стараются использовать синхротронный ускоритель. Это довольно дорогое сооружение. Лабораторные рентгеновские установки тоже используются, но синхротрон имеет ряд преимуществ:

— мощность пучка. Здесь два плюса. Первый понятен — сокращается время эксперимента. Второй — биологические кристаллы имеют тенденцию разрушаться под действием рентгеновского излучения. Процесс разрушения занимает определенное время. Если пучок мощный — можно успеть зарегистрировать нужную картину, пока кристалл не разрушился;

— возможность получить желаемую длину волны. Рентгеновские трубки дают мощный пучок только фиксированной длины волны (обычно около 1.57Å), в то время как при проведении эксперимента зачастую необходимо иметь возможность выбора длины волны. Это позволяет сделать синхротрон.

Обработка результатов рентгеновского эксперимента базируется на мощном математическом аппарате, который я не буду приводить в данном обзоре. Остановлюсь лишь на основных моментах. Функция электронной плотности в кристалле является периодической, она может быть представлена в виде комплексного ряда Фурье:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{h}} F_{\mathbf{h}} e^{2\pi i(\mathbf{h} \cdot \mathbf{r})},$$

где \mathbf{r} — вектор координат, V — объем элементарной кристаллической ячейки, а трехмерный век-

тор $\mathbf{h} = (h, k, l)$ принимает все целочисленные значения. Комплексный коэффициент $F_{\mathbf{h}}$ называется структурным фактором. Поскольку $F_{\mathbf{h}}$ — вектор, его можно представить в виде:

$$F_{\mathbf{h}} = |F_{\mathbf{h}}| \cdot e^{i\phi_{\mathbf{h}}},$$

где $F_{\mathbf{h}}$ — модуль структурного фактора, а $\phi_{\mathbf{h}}$ — фаза.

Когда монохроматический рентгеновский луч падает на кристалл, определенным образом ориентированный относительно падающего луча, то рассеяние происходит в дискретных направлениях, определяемых кристаллической решеткой. Дифракционная картина, возникающая на пленке детектора (рисунок 3), представляет собой набор пятен, или рефлексов, пронумерованных с помощью векторов \mathbf{h} . Согласно кинематической теории рассеяния, интенсивность каждого отдельного рефлекса пропорциональна квадрату модуля структурного фактора. Измерив интенсивность рефлексов, можно получить значения модулей структурных факторов, и чтобы определить искомую функцию распределения электронной плотности $\rho(\mathbf{r})$, недостает только значений фаз. Если для какого-либо кристалла фазы определены, то расчет положений атомов этого кристалла не составляет принципиальных трудностей. Таким образом, центральная проблема метода рентгеноструктурного анализа, называемая «фазовой проблемой», заключается в невозможности получения всех необходимых для расчета данных непосредственно из эксперимента. На сегодняшний день существуют различные подходы к решению «фазовой проблемы», а также активно ведутся работы по развитию новых методов, позволяющих оценить значения фаз структурных факторов.

Решение «фазовой проблемы»

Общего решения «фазовой проблемы» на сегодня не существует. Каждый случай требует специального подхода. Здесь важно понимать, что новая информация не берется ниоткуда. Для того, чтобы получить новую информацию — значения фаз, — мы должны либо сделать какие-то новые предположения о структуре и особенностях объекта, либо провести новые эксперименты. И вот этот момент — откуда мы пытаемся извлечь новую информацию — нужно всегда четко осознавать. Ниже приведены основные подходы к решению «фазовой проблемы», применяемые в белковой кристаллографии.

Изоморфное замещение

Можно попытаться внедрить в молекулы кристалла некую тяжелую метку, которая представляет собой один или несколько тяжелых

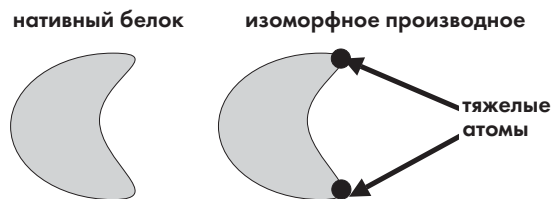


Рисунок 5. Изоморфное замещение белков.

атомов (например, ионы тяжелых металлов), которые могут быть либо добавлены к нативной структуре, либо могут замещать часть ее атомов (рисунок 5). Изоморфные производные должны быть закристаллизованы в той же самой пространственной группе, что и исходная макромолекула. Под изоморфным внедрением тяжелых атомов подразумевается, что они присоединяются к каждому экземпляру молекулы, в одном и том же месте, и структура молекулы белка при этом не изменяется. Затем, проведя дополнительно рентгеновский эксперимент с таким модифицированным соединением и играя на разностях интенсивностей рассеяния нативным белком и изоморфным производным, можно получить дополнительную информацию о значениях фаз. Трудность этого метода заключается в том, что не всегда удается получить хорошее изоморфное производное, а также в необходимости проведения дополнительного рентгеновского эксперимента.

Метод изоморфного замещения является основным методом решения фазовой проблемы при определении структуры биологических макромолекул. Сам этот метод возник достаточно давно, но именно при работе с белками он приобрел исключительно важную роль. Причин две:

- 1) долгое время он являлся единственным методом, позволяющим решать фазовую проблему для белков;
- 2) именно для белков удается «достаточно просто» получать изоморфные производные. Это связано с тем, что кристаллы белка довольно рыхлые — в них от 30 до 70% объема занято растворителем, т.е. в кристаллах есть «пустоты», куда могут поместиться дополнительные атомы.

Использование эффекта аномального рассеяния

При использовании этого метода длина волны падающего на кристалл луча варьируется вокруг значений, при которых происходит эффект резонанса, или аномальное рассеяние, для нескольких «специальных» атомов, содержащихся в структуре макромолекулы. Если аномально рассеивающих атомов в белке нет, иногда можно попытаться присоединить их химическим путем.

Дифракционные картины измеряют для нескольких значений длины волны падающего луча и на основании анализа разностей интенсивностей рефлексов этих картин оценивают значения фаз. Успех метода аномального рассеяния, как и изоморфного замещения, во многом зависит от возможности экспериментального получения производных с требуемыми свойствами.

Упомянутые два способа отвечают попытке решить фазовую проблему за счет дополнительной информации, получаемой из дополнительных экспериментов. Следующий способ применяют в ситуации, когда нам известна структура близкого (гомологичного) белка.

Метод молекулярного замещения

В биологии распространенной является ситуация, когда существуют ряды объектов, похожих друг на друга, или, как говорят, имеющих структурную гомологию. Такой гомологией могут обладать, например, белки одного типа (например, гемоглобины), выделенные из разных организмов. В этом случае можно надеяться, что фазы структурных факторов, рассчитанные по известной атомной модели гомологичного белка, будут достаточно хорошим начальным приближением к значениям неизвестных фаз, отвечающих исследуемому объекту. Комбинируя их далее с измеренными в эксперименте модулями структурных факторов для исследуемого объекта, мы можем получить хорошее приближение к искомому распределению электронной плотности. Однако, для того, чтобы надеяться на успех на этом пути, надо, как минимум, для начала «разместить» известный гомологичный объект на том же месте и в той же ориентации, что и глобула исследуемого белка. Вот эта процедура создания такого «компьютерного гибрида», в котором внутри элементарной ячейки кристалла одного белка размещается молекула другого, и получила название «метод молекулярного замещения». Судить о том, насколько полученное размещение близко к действительности, можно сравнивая рассчитанные по модели модули структурных факторов с величинами, полученными в эксперименте. Разумеется, такое замещение — всего лишь умозрительная процедура, и никакого химического замещения, конечно, не происходит.

«Прямые» методы

В отличие от предыдущих подходов, эти методы опираются не на дополнительный эксперимент или информацию о структуре гомологичного объекта, а на почти философскую идею об атомности изучаемого объекта. Под «прямыми» методами в кристаллографии понимаются стратегии определения структур, использующие

в качестве стартовой информации только набор интенсивностей рефлексов, полученный в рентгеновском эксперименте. Для определения фаз структурных факторов в них используют вероятностный подход. «Прямые» методы более объективны в том смысле, что они зависят только от применения математических соотношений. Один из двух ученых, получивших в 1985 г. Нобелевскую премию за развитие таких методов, — американский математик Герберт Хауптман. На основе «прямых» методов определяют структуры большинства низкомолекулярных соединений. Эти методы подкупают тем, что они не требуют ни дополнительных экспериментов и, что еще более ограничительно, тонкой биохимической работы по получению изоморфных производных, ни наличия известных гомологичных структур. К сожалению, в имеющемся варианте «прямые» методы неприменимы к структурам белков из-за принципиальных ограничений на число атомов исследуемой структуры. В настоящее время предпринимаются интенсивные попытки разработать аналогичные методы для белков. При этом работа идет как в направлении развития существующих подходов (например, в группе Г. Хауптмана, университет Баффало, штат Нью-Йорк, США), так и в поисках альтернативных подходов (сектор кристаллографии макромолекул в Институте математических проблем биологии, Пушино).

Расчет синтеза Фурье электронной плотности

Если нам известны и модуль, и фаза преобразования Фурье искомой функции $\rho(r)$, то мы можем восстановить $\rho(r)$, рассчитав обратное преобразование Фурье. Это несложная с современной точки зрения вычислительная задача, и этот шаг выделяется потому, что он подводит итог важного периода работ. Мы, наконец, получаем возможность «взглянуть» на интересующий нас объект. И по тому, насколько «четким» получилось изображение, — судить об успешности всех предыдущих этапов работы. А в случае неудачи — повторить всё сначала.

Интерпретация карт распределения электронной плотности

Этот этап заключается в построении приближенной атомной модели по рассчитанным картам распределения электронной плотности. Эта работа требует максимального использования интеллекта человека и осуществляется квалифицированными специалистами. С помощью специальных компьютерных программ, исследователь вручную вписывает атомы белковой структуры в полученную на предыдущем этапе карту электронной плотности (рисунок 6). При

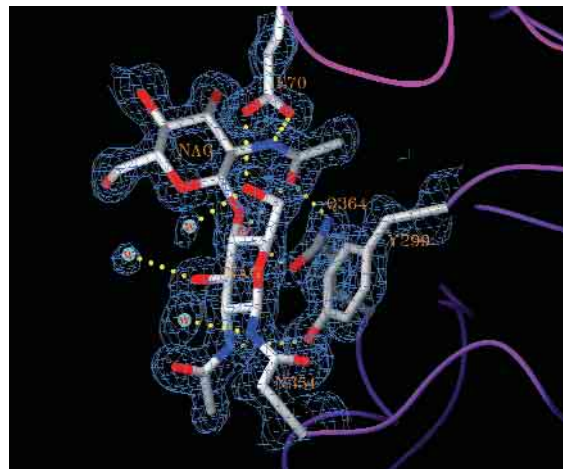


Рисунок 6. Модель, вписанная в карту электронной плотности.

этом необходимо учитывать характеристики связей в исследуемой молекуле (длина, углы) и факторы, влияющие на взаимное расположение аминокислотных радикалов (электростатическое притяжение или отталкивание, гидрофобные взаимодействия, водородные связи и др.). Эта работа выполняется на современных мощных графических системах, поддерживающих стереоизображение. Сложность работы заключается в том, что надо разместить в пространстве тысячи, а то и десятки тысяч атомов, сообразуясь с максимумами на картах функции распределения электронной плотности $\rho(r)$. Работа осложняется тем, что, как правило, карты не имеют атомного разрешения, зато облегчается знанием аминокислотной последовательности исследуемого белка, что позволяет строить модель блоками из известных аминокислот. Это очень ответственный этап, так как ошибки, заложенные на нем, потом бывает очень трудно исправить.

Уточнение структуры

Процедура уточнения — последний этап построения атомной модели исследуемой структуры. Цель этой процедуры — улучшить предварительную атомную модель (созданную на предыдущем этапе) и в результате получить такие координаты атомной модели, для которых рассчитанные модули структурных факторов были бы наиболее близки значениям, полученным в эксперименте. Математически уточнение представляет собой процедуру минимизации некоторой функции, являющейся мерой близости наборов рассчитанных и наблюдаемых модулей структурных факторов. При этом минимизируется функция большого числа переменных (координат атомов). Для решения этой задачи непосредственного участия исследователя не

требуется, все расчеты выполняются с помощью компьютерных программ. Необходимо поместить в программу координаты атомов, полученные на предыдущем этапе в результате ручной подгонки, и набор экспериментальных значений структурных факторов. Затем запускается расчет, который длится от нескольких минут до нескольких часов (в зависимости от сложности структуры и мощности компьютера).

В качестве одного из главных показателей, насколько модель соответствует экспериментальным данным и насколько успешно идет процедура уточнения, широко используют кристаллографический *R*-фактор:

$$R = \frac{\sum_{\mathbf{h}} |F^{\text{obs}}(\mathbf{h}) - F^{\text{calc}}(\mathbf{h})|}{\sum_{\mathbf{h}} F^{\text{obs}}(\mathbf{h})}$$

где $F^{\text{obs}}(\mathbf{h})$ — модули структурных факторов, измеренные в эксперименте, $F^{\text{calc}}(\mathbf{h})$ — модули структурных факторов, рассчитанные по координатам модели. Чем меньше *R*-фактор, тем лучше полученные координаты атомов соответствуют пространственной структуре исследуемой молекулы. Однако, уменьшить *R*-фактор до нуля не удастся, т.к. измеренные модули структурных факторов $F^{\text{obs}}(\mathbf{h})$ содержат экспериментальные ошибки. Обычно у расшифрованных структур *R*-фактор составляет 0,1–0,15.

Еще одним показателем качества модели является так называемый свободный *R*-фактор. Его отличие от обычного *R*-фактора заключается в следующем. До начала процедуры уточнения часть экспериментальных данных (например, 10%) должна быть объявлена контрольным набором, который не используется в процессе уточнения структуры. Если после проведенного уточнения окажется, что и для этих контрольных данных модули структурных факторов, рассчитанные по модели, близки к экспериментальным, то это позволяет с большим доверием относиться к полученной модели, нежели основываясь на степени совпадения с экспериментом тех величин, которые мы стремились к нему подогнать. Поэтому значение свободного *R*-фактора представляет собой более значимую характеристику, нежели обычный *R*-фактор.

На практике выполнение двух последних этапов рентгеноструктурного анализа (ручное вписывание координат атомов в карту электронной плотности и уточнение полученной модели) циклически повторяется. При этом в каждом цикле необходимо сравнивать *R*-факторы, полученные после уточнения модели с *R*-факторами, полученными для модели в предыдущем цикле. И если значения *R*-факторов (или одного из них)

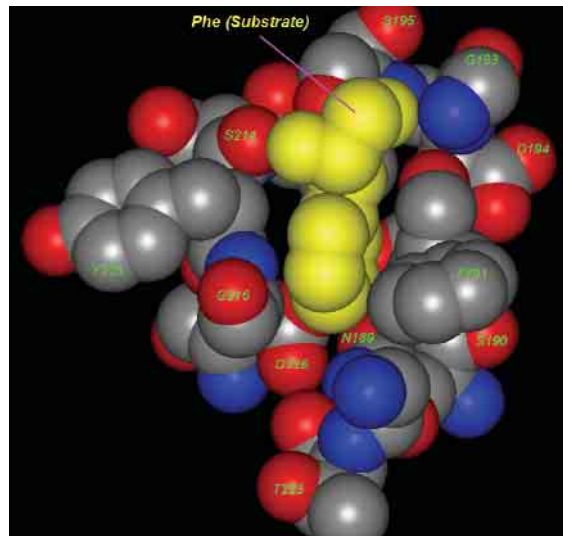


Рисунок 7. Структура фермента дуоденазы с молекулой субстрата.

сильно возросли, это значит, что, скорее всего, в модель были внесены неудачные изменения координат атомов и, следовательно, от этих изменений нужно отказаться и вернуться к модели предыдущего цикла.

Такие циклические повторения ручной правки и уточнения структуры осуществляются до тех пор, пока вносимые изменения улучшают структуру (т.е. улучшается ее соответствие карте электронной плотности, становятся более энергетически выгодным взаимное расположение атомных групп) и уменьшается *R*-фактор. Обычно на это уходит от нескольких месяцев до года. Если в результате этой работы так и не удастся получить значение *R*-фактора меньше 0,2, то приходится признать, что получены неудачные экспериментальные данные или допущены ошибки при их обработке. В этом случае остается только попробовать исправить эти ошибки или провести эксперимент заново. Однако такие неудачи случаются достаточно редко, и обычно после всех изложенных выше этапов работы мы получаем атомную структуру объекта, которая до этого была неизвестна. На рисунках 7 и 8 приведены иллюстрации структур, расшифрованных с помощью метода рентгеноструктурного анализа. Таким образом, атомная структура определена, и завершающий этап связан с оформлением полученных результатов в соответствии с определенными стандартами.

Помещение координат в банк белковых структур

Существует специальная база данных, в которой хранятся все расшифрованные на данный момент белковые структуры. Каждой струк-

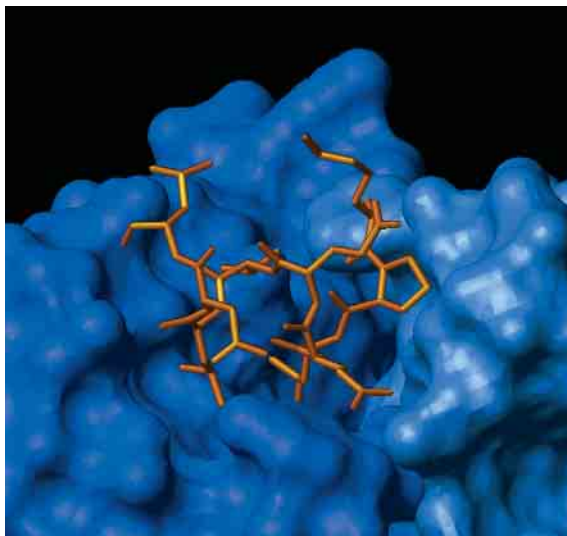


Рисунок 8. Комплекс Fab-фрагмента моноклонального антитела с синтетическим нонапептидом в активном центре.

туре присваивают уникальный идентификационный номер, по которому ее можно быстро найти. Эта база данных носит название Protein Data Bank (банк белковых структур). Она доступна через Internet по адресу <http://www.rcsb.org>). Разумеется, Protein Data Bank постоянно попол-

няется. Чтобы поместить в Protein Data Bank новую расшифрованную структуру, нужно отправить туда не только полученные для нее координаты атомов, но и множество дополнительной информации: параметры кристаллической ячейки; разрешение, при котором проводился эксперимент; полученные для конечной модели значения R-факторов; сведения об авторах, занимавшихся расшифровкой; статьях, в которых опубликованы результаты; и многое другое. Всю эту информацию отправляют в Protein Data Bank, там ее анализируют, и если возникают какие-то вопросы, их уточняют по электронной почте. В конечном итоге, структура оказывается в базе данных и становится доступной любому исследователю.

Литература

- Бландел Т., Джонсон Л.** 1979. Кристаллография белка. — М.: Мир.
- Кантор Ч., Шиммель П.** 1984. Биофизическая химия, Том 2. — М.: Мир.
- Ландау Л.Д., Лифшиц Е.М.** 1969. Краткий курс теоретической физики. Книга 1. Механика, Электродинамика. — М.: Наука.