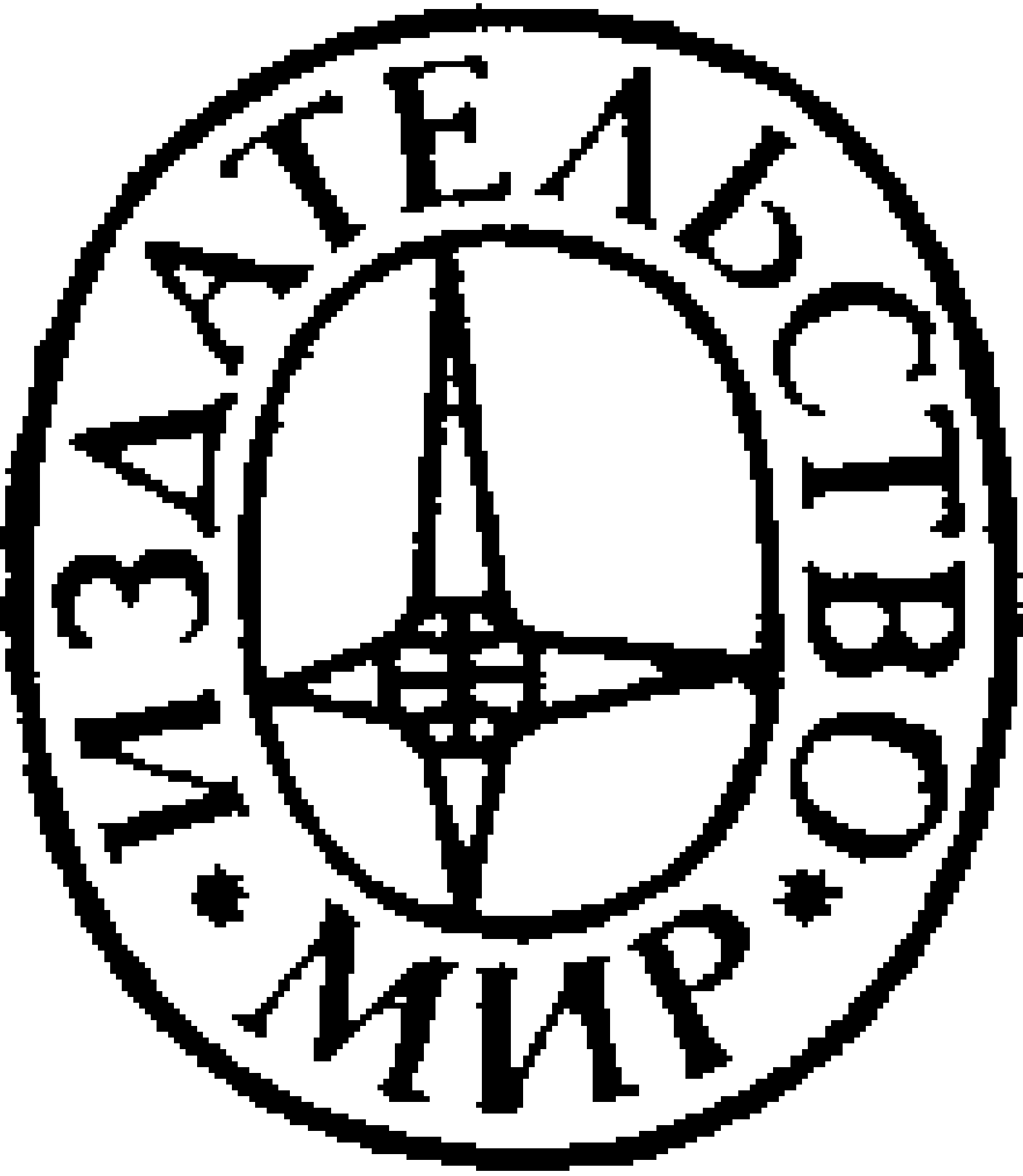

ЗДАТЕЛЬСТВО
МИР

SPRINGER ADVANCED TEXTS IN CHEMISTRY

Charles R. Cantor, Editor

G. E. Shulz
R. H. Schirmer

Principles of Protein Structure

Springer-Verlag
New York · Heidelberg · Berlin

Г. Шульц, Р. Ширмер

ПРИНЦИПЫ СТРУКТУРНОЙ ОРГАНИЗАЦИИ БЕЛКОВ

Перевод с английского
канд. хим. наук **В. Е. Шкловера**

под редакцией
доктора хим. наук, профессора **Е. М. Попова**

Наукова бібліотека
ім. М. Максимовича
КНУ
ім. ТАРАСА ШЕВЧЕНКА
1097728
Н - наукова
Ц: 0.00

ФОНД ЧИТАЛЬНИЙ

МОСКВА · МИР · 1982

УДК 547.96

В книге, написанной авторами из ФРГ, изложены современные представления о принципах, определяющих формирование пространственной структуры белков, причем вопрос о структурной организации этих важных биополимеров рассматривается в неразрывной связи с их биологическими функциями.

Книга предназначена для биохимиков, биологов, химиков — научных работников, преподавателей, аспирантов, студентов.

Редакция литературы по химии

1097728

2015000000

Ш $\frac{20505\text{-}086}{041(01)\text{-}82}$ 86-82, ч. 1

© Springer-Verlag New York Inc., 1979
All Rights Reserved
Authorized translation from English language edition published by Springer-Verlag Berlin-Heidelberg-New York
© Перевод на русский язык, «Мир», 1982

Предисловие редактора перевода

Все биологические процессы осуществляются при непременном участии белков. Они служат регуляторами генетической функции нуклеиновых кислот, в качестве ферментов участвуют во всех стадиях биосинтеза полипептидов, полинуклеотидов и других соединений, катализируют все метаболические процессы. Особые сократительные белки ответственны за клеточные и внутриклеточные движения. В комплексе с липидами белки входят в состав мембран, обеспечивая активный транспорт метаболитов в клетку и из нее. Белки служат для запасания и переноса кислорода. Низкомолекулярные полипептиды, гормоны, стимулируют функциональную активность в клетках других тканей и органов. Белки осуществляют иммунологическую функцию, защищая организм от чужеродных соединений. Они входят в состав кожи, волос, соединительных тканей, костей и т. д., выполняя динамическую опорную функцию, обеспечивая тем самым взаимосвязь органов, их механическую целостность и защиту. Это далеко не полный перечень осуществляемых белками функций.

Возможность реализации столь различных по своему характеру и чрезвычайно многообразных по спектру действия функций не может быть прямо связана с химическим составом белков, так как все они полипептиды и построены из одних и тех же аминокислотных остатков. Уникальные свойства белковых молекул объясняются прежде всего особенностями их пространственного строения, которое в зависимости от последовательности одних и тех же аминокислот может быть весьма разнообразным.

В основе биологических процессов, их высокой эффективности, строгой избирательности, взаимообусловленности и автоматизма лежат тонкие структурные соответствия взаимодействующих соединений, определяемые главным образом спецификой пространственной организации и динамическими конформационными свойствами белковых молекул. Поэтому решение практически любой биологической проблемы неминуемо сталкивается с необходимостью изучения трехмерных структур белков. Уровень наших знаний

в этой области в значительной степени определяет развитие многих направлений не только в биологии, но и в медицине, фармакологии и ряде других смежных наук. Это делает проблему пространственной организации белковых молекул самой актуальной, ключевой проблемой биологии.

Книга Шульца и Ширмера «Принципы структурной организации белков» — первая в мировой литературе монография, в которой рассматриваются общие вопросы пространственного строения белков и на этой основе их биологические функции. Основное достоинство книги — широкий охват рассматриваемых вопросов, причем главное внимание направлено на механизм свертывания полипептидной цепи в нативную конформацию белка и на структурно-функциональную зависимость. Обсуждение различных аспектов структурной проблемы белка проведено на высоком профессиональном уровне.

Монография построена как вводный курс по пространственной организации и функционированию белков. В ней кратко излагаются основные сведения об аминокислотных остатках, природе межостаточных и гидрофобных взаимодействий, структурной иерархии и т. д. вплоть до взаимодействия белков с другими молекулами, а также проблемы ферментативного катализа и механохимическая функция сократительных белков. Книга написана логично, простым, ясным языком, хорошо иллюстрирована и снабжена обширным библиографическим списком. Знакомство с этим материалом подготовит читателя к углубленному усвоению специализированной лнтературы о пространственном строении наиболее сложных из известных современной науке молекул.

Можно с уверенностью сказать, что выход в свет книги Г. Шульца и Р. Ширмера на русском языке будет встречен научной общественностью нашей страны с большим интересом. Книга безусловно окажется полезной многим исследователям, работающим в самых разных областях биохимии, биофизики и общей биологии.

Е. Попов

Предисловие редактора серии "Springer Advanced Texts in Chemistry"

Читатель постоянно снабжается новыми учебными пособиями по химии. Причем такие области, как, например, общая биохимия, теоретическая органическая химия (и прежде всего ее направление, изучающее механизмы органических реакций) и химическая термодинамика, хорошо обеспечены превосходными книгами, многие из которых после переработки и дополнений переиздаются; это позволяет учесть все последние достижения. Однако другие области химии, особенно те из них, которые изучаются как спецпредметы студентами выпускных курсов высших учебных заведений, страдают от фактического отсутствия современных учебников. Это относится прежде всего к быстро прогрессирующим областям. Авторами учебников по таким спецпредметам обычно должны быть ученые, которые непосредственно вносят вклад в исследования, продвигающие данную область. Не всегда легко убедить таких людей выделить время для распространения накопленных ими знаний. Наша цель — выявить области химии, развитие которых опережает содержание уже существующих учебников, найти специалистов по этим областям и убедить их написать не очень пространные, но содержательные пособия-введения в соответствующие области. В некоторых случаях появление таких пособий, в свою очередь, может оказать стимулирующее воздействие на развитие новых направлений.

Чарльз Р. Кантор

Нью-Йорк декабрь, 1978

От авторов

Из четырех основных типов биополимеров — нуклеиновых кислот, полисахаридов, липидов и белков — последние наиболее изучены к настоящему времени, поскольку они доступнее для выделения и анализа. Белки оказались на переднем крае исследования биополимеров более столетия назад, когда Кюне [1] выделил и охарактеризовал фермент трипсин, а Хоппе-Сейлер [2] получил кристаллы гемоглобина. Эти опыты показали, что белки имеют вполне определенную структуру, что, однако, не было общепризнанным до тех пор, пока не удалось непосредственно наблюдать детали их атомного строения с помощью дифракции рентгеновских лучей.

На сегодняшний день известны структуры более 70 белков*. С практической точки зрения эти структуры могут быть использованы биохимиками и фармакологами в качестве моделей при постановке и интерпретации их экспериментов. Цель данной книги — рассмотреть некоторые общие закономерности, которые следуют из анализа и сопоставления известных белковых структур. Эти закономерности необходимы для понимания взаимосвязи между информацией об аминокислотной последовательности белка, хранящейся в ДНК, и его функцией в организме. Кроме того, знание основ облегчит и более широкое использование трехмерных структур белков в научных и практических целях. Сама возможность сформулировать общие для всех белков структурные принципы свидетельствует о существовании однотипного механизма пространственной организации белковых молекул, несмотря на все разнообразие их функций. Для биологической медицины это могло бы означать, что глубокое понимание физиологических и патологических процессов на молекулярном уровне — не утопия и что вполне возможно строго целенаправленное влияние на эти процессы.

Главы книги вполне самостоятельны. Их можно изучать в произвольном порядке, но тем, кто не знаком со структурой белков, мы рекомендуем вначале ознакомиться с седьмой главой. Мы хотели бы поблагодарить наших коллег, предоставивших графический и цифровой материал, а также всех, кто помогал нам при подготовке рукописи.

Г. Е. Шульц
Р. Х. Ширмер

Гейдельберг декабрь, 1978

---

* За время, прошедшее с момента выхода **этой книги на** английском языке, получены кристаллографические структуры **еще приблизительно** для такого же количества белков. — *Прим. ред.*

Глава 1 Аминокислоты

1.1 20 канонических аминокислот

В настоящее время основную схему организации живой материи можно считать известной. Нуклеиновые кислоты несут всю генетическую информацию, которая заложена в последовательности четырех различных нуклеотидных оснований. Существуют нуклеиновые кислоты двух типов. Более стабильная дезоксирибонуклеиновая кислота (ДНК) является хранителем информации. Менее стабильная рибонуклеиновая кислота (РНК), транскрибирующаяся с ДНК, выполняет роль матрицы, которая транслирует нуклеотидный текст в аминокислотные последовательности белков с помощью рибосомного механизма. Белки участвуют фактически во всех типах деятельности организма.

В этой главе будут рассмотрены 20 канонических аминокислотных остатков, встречающихся в белках. Они приведены в табл. 1.1 в соответствии с их относительным содержанием; здесь же даны обычно используемые трехбуквенные и однобуквенные обозначения, а также облегчающие их запоминание мнемонические приемы. Молекулярные массы остатков варьируют от 57 до 186 (среднее значение $\sim$ 110). Таким образом, белок с молекулярной массой 33 000 содержит приблизительно 300 остатков. Аминокислоты можно классифицировать в зависимости от свойств их боковых цепей на полярные и неполярные. Полярные боковые цепи можно подразделить и далее на нейтральные, основные или кислотные. На рис. 1.1 показаны боковые цепи 20 наиболее часто встречающихся аминокислот, причем остатки с близкими свойствами помещены рядом.

1.2 Почему именно эти аминокислоты?

По-видимому, жизнь начиналась не с белков, а с автокаталитических нуклеиновых кислот. Ответить на вопрос, почему белки оказались построенными из аминокислот, приведенных в табл. 1.1, пока еще нельзя. Однако некоторые гипотезы могут приблизить нас к пониманию проблемы. Обсудим вначале основную организа-

Таблица 1.1

Обозначение и свойства 20 канонических аминокислотных остатков белков[a]

Аминокислота или ее остаток	Трехбуквенный символ	Однобуквенный символ	Мнемонический ключ к однобуквенному символу	Относительное содержание в белках *E. coli*, % [19]	Молекулярная масса остатка при pH 7	Значение pK боковой цепи [19]	Значение ΔG переноса боковой цепи из воды в этанол при 25°C, ккал/моль [16]
Аланин	Ala	A	**A**lanine	13,0	71		—0,5
Глутаминовая кислота	Glu	E	gluEtamic acid		128	4,3	
Глутамин	Gln	Q	**Q**-tamine	10,8	128		
Аспарагиновая кислота	Asp	D	asparDic acid	9,9	114	3,9	
Аспарагин	Asn	N	asparagiNe		114		
Лейцин	Leu	L	**L**eucine	7,8	113		—1,8
Глицин	Gly	G	**G**lycine	7,8	57		
Лизин	Lys	K	перед **L**	7,0	129	10,5	
Серин	Ser	S	**S**erine	6,0	87		+0,3
Валин	Val	V	**V**aline	6,0	99		—1,5
Аргинин	Arg	R	a**R**ginine	5,3	157	12,5	
Треонин	Thr	T	**T**hreonine	4,6	101		—0,4
Пролин	Pro	P	**P**roline	4,6	97		
Изолейцин	Ile	I	**I**soleucine	4,4	113		
Метионин	Met	M	**M**ethionine	3,8	131		—1,3
Фенилаланин	Phe	F	**F**enylalanine	3,3	147		—2,5
Тирозин	Tyr	Y	t**Y**rosine	2,2	163	10,1	—2,3
Цистеин	Cys	C	**C**ysteine	1,8	103		
Триптофан	Trp	W	t**W**o rings	1,0	186		—3,4
Гистидин	His	H	**H**istidine	0,7	137	6,0	—0,5
				Средняя масса	108,7		

[a] Дополнительные трехбуквенные символы: Asx, Glx — либо кислота, либо амид. Аминокислотный остаток —HN—CHR—CO есть часть аминокислоты (рис. 1.2,а), входящая и пептидное звено; R — боковая цепь (рис. 1.1). Дополнительные однобуквенные символы: B = Asx, Z = Glx, X = неизвестный или минорный аминокислотный остаток.

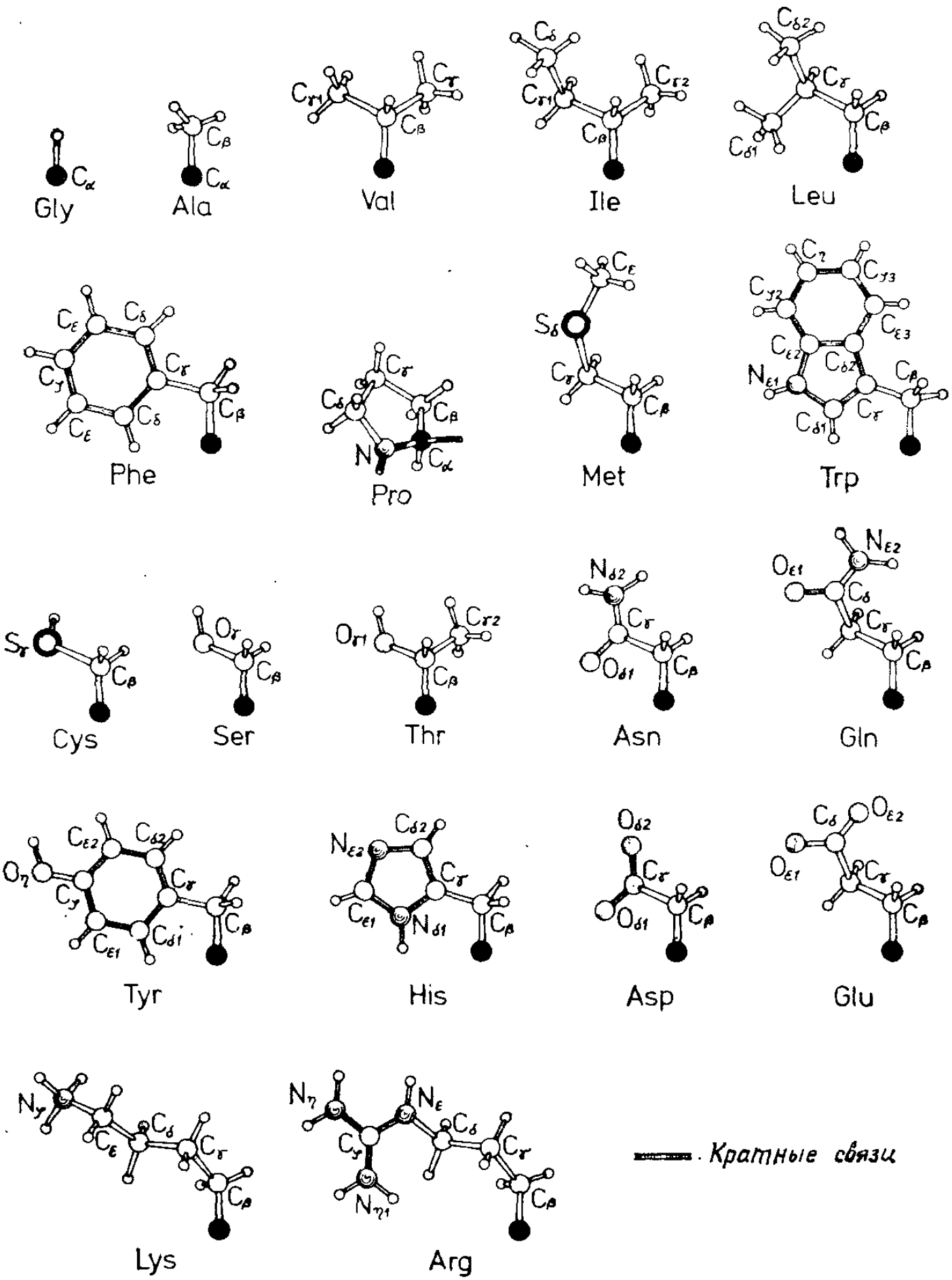


Рис. 1.1.

Боковые цепи 20 канонических аминокислот.

Для пролина показана часть основной цепи. Все остальные боковые цепи показаны отходящими от C_α -атомов главной цепи. Названия остатков даны в трехбуквенных обозначениях. Обозначения атомов даны согласно номенклатуре IUPAC—IUB от 1969 г. [21]. Основная цепь в Pro зачернена. Все C_α -атомы зачернены.

ционную схему происхождения жизни. Наиболее вероятно, что жизнь началась с простых автокаталитических, т. е. самовоспроизводящихся молекул [3]. Именно нуклеиновые кислоты, а не белки, способны выполнять эту роль [4—6], поскольку однонитевые нуклеиновые кислоты могут использовать спаривание оснований для выстраивания комплементарной последовательности нуклеотидов. При объединении выстроенных нуклеотидов две нити могут образовывать энергетически выгодную двойную спираль Уотсона — Крика. В следующем цикле, после развертывания двойной спирали, комплементарная нить может репродуцировать исходную. Оба фактора, точный отбор последовательности путем спаривания оснований и энергетически выгодное образование двойной спирали, безусловно, способствовали автокаталитическому воспроизведению. Напротив, исходить из белка вместо нуклеиновой кислоты было бы значительно сложнее с точки зрения единой организационной схемы, поскольку взаимного опознавания аминокислот не существует.

Наше представление о последующей молекулярной эволюции ограничивается пока что более или менее обоснованными догадками. Несмотря на то что самовоспроизведение вначале должно было быть малоэффективным, молекулы, находясь в жестких условиях эволюции, где выживают лишь наиболее приспособленные системы, должны были стремиться к усовершенствованию автокатализа. Можно предположить, что такое усовершенствование начиналось с «ферментов», образовавшихся из нуклеиновых кислот — осколков самих автокаталитических молекул. По-видимому, рибосомная и транспортная РНК — остатки таких катализаторов. Позже ферменты образовавшиеся из нуклеиновых кислот были вытеснены более эффективными белковыми ферментами.

Линейный набор стандартных элементов со стандартными связями. Белковые ферменты возникли по чрезвычайно простой организационной схеме. Аминокислотная последовательность полипептидной цепи белка есть коллинеарное и единственное представление нуклеотидной последовательности исходной нуклеиновой кислоты. Три соседних нуклеотида кодируют одну аминокислоту (рис. 1.5, *б*). Таким образом, полипептиды сходны с нуклеиновыми кислотами в том, что это *линейные* цепные молекулы, построенные из стандартных элементов с одной стандартной связью. Это обеспечивает простое и универсальное «считывание» с нуклеиновых кислот в процессе синтеза полипептидов. Простота линейных систем широко используется, в частности, в электронно-вычислительной технике, где хранение и вызов информации обычно осуществляются с помощью одномерных блоков, записанных в стандартной форме на линейных носителях, например магнитных лентах.

В нуклеиновых кислотах, как и в белках, стандартизованы не только связи, но и атомные группы, образующие скелет цепи;

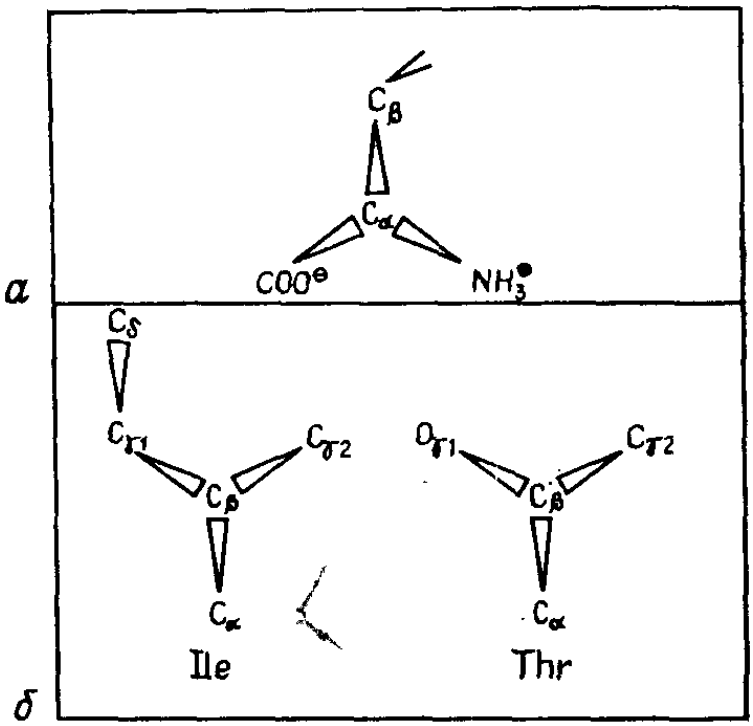


Рис. 1.2.
Асимметрические атомы углерода, встречающиеся в канонических аминокислотных остатках. Проекция от атома водорода к атому C_α или C_β соответственно, что противоположно направлению в схеме Кана — Ингольда — Прелога [41].

а — обычный L-изомер α-аминокислоты. Абсолютная конфигурация S при C^α в обозначениях Кана—Ингольда—Прелога. *б* — конфигурация при C_α -атоме изолейцина (S) и C_β -атоме треонина (R).

в полипептидах все 20 аминокислотных остатков относятся к α-типу и имеют L-конфигурацию C_α-атома (рис. 1.2,*а*). Все различия, а значит, и вся информация сводятся к довольно коротким боковым цепям.

Необходимость самопроизвольного свертывания. После синтеза на рибосоме полипептидная цепь самопроизвольно свертывается в определяемый аминокислотной последовательностью глобулярный белок, принимая состояние с наименьшей свободной энергией. Возможно, что свертывание начинается уже в ходе синтеза. Характер свертывания образующейся цепи определяет специфичность белка. Пространственная самоорганизация молекулы значительно упрощает общую схему, так как в противном случае потребовались бы «морфогенетические ферменты» или «ферменты, способствующие свертыванию». Поскольку возможно огромное число способов свертывания цепи, то возникла бы потребность в большом числе таких вспомогательных ферментов.

Допустим, что структуры строятся не путем спонтанного свертывания, а, например, из небольшой молекулы-ядра путем ковалентного присоединения последующих фрагментов. Даже в случае наиболее эффективной схемы, показанной на рис. 1.3, такой процесс потребовал бы двух ферментов для обработки каждого промежуточного продукта, т. е. в среднем по два фермента на образующуюся молекулу. Таким образом, для получения самоподдерживающейся системы все образующиеся на промежуточных стадиях молекулы использовались бы в качестве ферментов, каждый из которых должен был бы катализировать две различные синтетические реакции. Однако в существующих организмах у каждой молекулы наблюдается, вообще говоря, только одно ферментативное действие.

К тому же очевидно, что подобная система оказалась бы слишком чувствительной к ошибкам и почти не имела бы шансов эволюционировать в первичной химической массе [6]. Этот пример позволяет оценить простоту, присущую спонтанному свертыванию.

Однако не все полипептидные цепи свертываются самостоятельно. Некоторые из рибосомных белков очень вытянуты и имеют мало внутренних контактов [7]. В этом случае цепь принимает свою

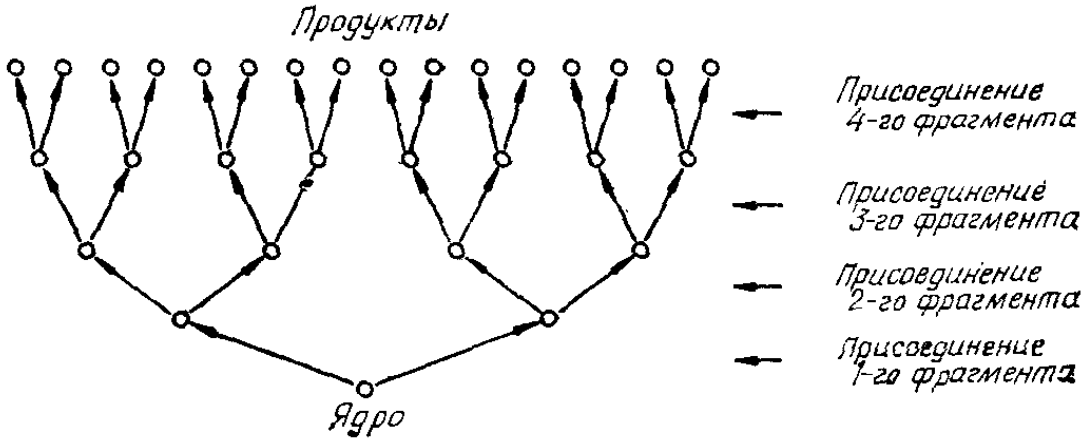


Рис. 1.3.

Гипотетическое образование молекулярной структуры путем последовательного присоединения различных фрагментов к молекуле-ядру или путем направленной модификации молекулы предшественника.

Чтобы получить различные структуры, каждое присоединение или модификация должны быть специфическими, что означает необходимость специфического ферментативного воздействия. На четвертой стадии, когда образуются $2^4=16$ различных структур, необходимы $2+4+8+16=30=2^5-2$ воздействий.

окончательную конформацию, прикрепляясь к другой молекуле, а именно к нуклеиновому ядру рибосомы. Эти, по-видимому, очень древние цепи отражают то положение, которое существовало в самом начале образования белков; рибосомы прекрасно сохранились в ходе эволюции. Подобная стабильность следует из центральной роли рибосом в жизнедеятельности организма. Любое изменение в рибосоме имело бы разрушительные последствия, оставляя очень мало шансов на выживание [3, 4].

Канонические аминокислоты могли быть отобраны на начальных стадиях жизни. Набор стандартных звеньев полипептидной цепи состоит из 20 аминокислотных остатков (рис. 1.1, табл. 1.1). Чтобы оказаться отобранными, эти аминокислоты должны были быть в среде, окружавшей первичные автокаталитические нуклеиновые кислоты. Так, по-видимому, и было, что подтвердили некоторые модельные эксперименты [3, 8], где газовую атмосферу, которая могла соответствовать реальной атмосфере на ранних стадиях развития Земли, подвергали воздействию электрической или радиационной энергии. При этом обычно получали рацемическую смесь L- и D-аминокислот (*рис. 1.2,а*) α-типа наряду с β-аминокислотами (рис. 1.4) и другими соединениями.

L-Конфигурация отобрана случайно. Какой смысл заложен в выборе именно L-конфигурации C_α-атома? Мы понимаем, что обе C_α-конфигурации равновероятны; но механизм синтеза цепи требовал вполне определенной конфигурации, иначе синтез оказался бы слишком сложным и не выдержал бы конкуренции. Однако яв-

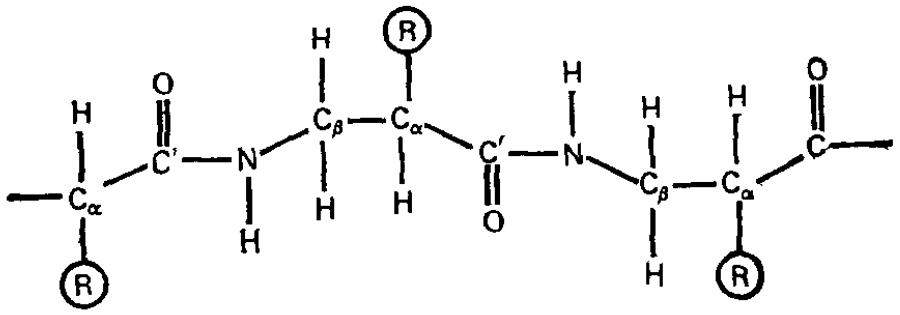


Рис. 1.4.

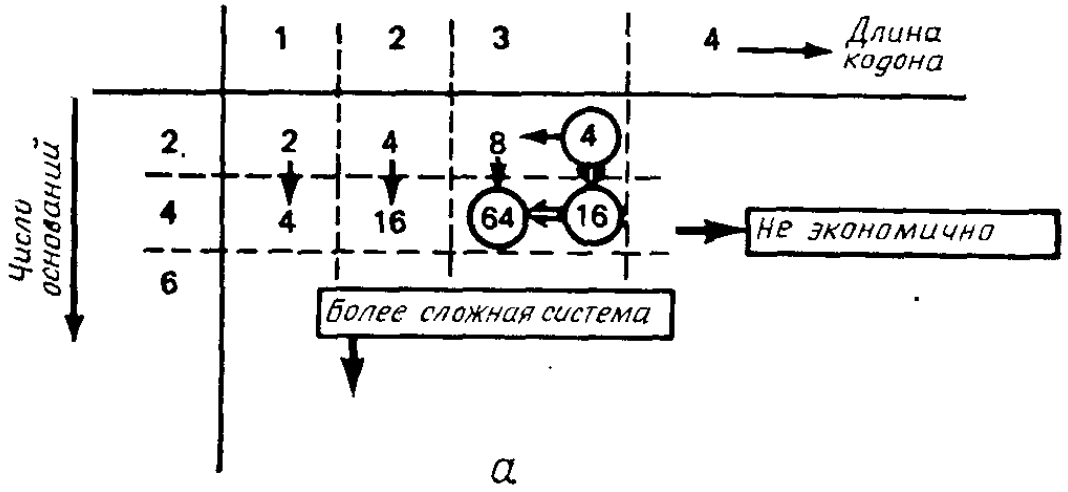
Гипотетическая основная цепь, состоящая из β-аминокислот. Вращение вокруг связей C_α—C_β и C_β—N достаточно свободно, что делает цепь слишком подвижной. Это также относится к случаю, когда боковые цепи R присоединены к C_β-атому, а не к C_α.

ной причины для предпочтения L-конфигурации ее зеркальному отображению, по-видимому, нет. Правда, было показано, что внутренняя асимметрия в распределении продуктов при β-распаде проявляется через молекулярную асимметрию путем предпочтительного разрушения D-аминокислот [9]. Но наблюдаемый эффект (несколько процентов) слишком слаб для объяснения отбора. По-видимому, выбор произошел случайно, а не вследствие немного более высокой концентрации среди исходных продуктов L-аминокислот. На последующих этапах D-система оказалась подавленной более развитой L-системой. Не исключено также, что обе системы развивались параллельно и L-система победила благодаря более удачным флуктуациям под действием внешних факторов, влияние которых, по-видимому, значительно превосходило влияние незначительного различия концентраций аминокислот.

β-Аминокислоты не подходят. Другой аспект указанных модельных экспериментов состоит в получении β-аминокислот (рис. 1.4) в количествах, сравнимых с аминокислотами α-типа [8]. Во многих отношениях β-аминокислоты сходны с α-аминокислотами, однако образующаяся полипептидная цепь будет содержать связи C_α—C_β, допускающие дополнительное свободное вращение. Как показано в разд. 8.1, такая подвижность препятствует самопроизвольному свертыванию цепи. Таким образом, смысл отбора α-аминокислот очевиден.

Трансляционный механизм фиксирует набор типов аминокислот. Почему все живое на Земле использует один и тот же набор аминокислот? Этот вопрос связан с трансляцией нуклеотидных последовательностей в аминокислотные последовательности, использую-

щей так называемый генетический код (рис. 1.5, *б*). Генетический код универсален, потому что он находится в самом центре организационной схемы жизни. Любое изменение в коде существенно нарушает схему и обычно приводит к вымиранию организмов. Поэтому код, а значит, и типы аминокислот исключительно тщательно со-



а

Второе основание

Первое основание	U	C	A	G	Третье основание
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr End End	Cys Cys End Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

б

Рис. 1.5.

Генетический код трансляции нуклеотидной последовательности в аминокислотную последовательность.

а — число кодируемых аминокислот как функция числа различных нуклеотидных оснований и длины кодона. Стрелки указывают возможные направления эволюции. Исходя из существующего кода, эволюция должна была следовать пути, отмеченому кружочками. *б* — таблица генетического кода, связывающая нуклеотидные триплеты с аминокислотами.

храняются. Нет никаких данных о том, что когда-либо существовали организмы с другими аминокислотами или с другими кодами.

Почему число отобранных типов аминокислот равно именно 20? Этот вопрос также связан с механизмом трансляции. На рис. 1.5,*а* в порядке дискуссии даны некоторые трансляционные схемы существующего генетического кода. При дублетном варианте (длина кодона равна двум нуклеотидам) с помощью четырех разных нуклеотидов можно закодировать $4^2 = 16$ аминокислот. Однако для длины кодона природа выбрала не два, а три нуклеотида. Для пояснения этого факта напомним, что длина кодона связана с решающим шагом в трансляции — опознанием нуклеотидной последовательности информационной РНК путем спаривания оснований нуклеотида с небольшой доставляющей аминокислоты транспортной РНК. Можно предположить, что при *дублетном* коде не оказалось оснований с достаточно большими константами ассоциации, и поэтому кодон должен был увеличиться до *триплета*, чтобы обеспечить специфическое узнавание. С помощью четырех различных нуклеотидов триплетный код может распознавать $4^3 = 64$ аминокислоты. Однако используются только 20 аминокислот. Для объяснения этого факта нужно предположить, что генетический код развивался и что его эволюция остановилась на полпути.

По-видимому, генетический код сформировался на очень ранних стадиях жизни. Как обсуждалось ранее, изменения генетического кода предельно заторможены. Так, совершенно невозможно изменение длины кодона, поскольку это уничтожило бы всю накопленную генетическую информацию и привело бы к катастрофе. Однако на очень ранних стадиях эволюции, когда организмы были значительно менее совершенны, чем теперь, изменение числа нуклеотидов или аминокислот было, по-видимому, реальным, при условии что оно происходило постепенно. Именно таким образом можно представить себе эволюцию кода в течение определенного периода.

Эта идея подтверждается значительной вырожденностью третьей позиции кодона (рис. 1.5, *б*). Наиболее вероятной начальной точкой существующего является, по-видимому, триплетный код, использующий два комплементарных нуклеотида. В этом случае значимы только первые две позиции каждого кодона, так что один триплет кодирует лишь четыре разные аминокислоты. На следующем этапе эволюции добавилась еще одна пара комплементарных нуклеотидов, что дало возможность кодировать 16 аминокислот. Наконец, приданием значимости третьей позиции кодона была введена некоторая вырожденность. Когда организмы стали настолько совершенными, т. е. настолько конкурентоспособными, что любое изменение типа хотя бы одной аминокислоты оказывалось опасным, а иногда и летальным, генетический код остановился в своем развитии. Таким образом, было зафиксировано число аминокислотных остатков, равное 20.

Сравнение существующих метаболических путей образования аминокислот с генетическим кодом показывает, что связанные метаболически между собой аминокислоты коррелируют также и в отношении их кодонов [10]. Это делает весьма привлекательной идею параллельной эволюции генетического кода и метаболизма, а также указывает на наличие исторической иерархии аминокислот. Более простые аминокислоты, как Gly, Ser, Ala, Asp и Glu, считаются ранними в отличие от более сложных аминокислот, например Met, His и Asn. Однако последовательное появление аминокислот не отражено в существующих белковых структурах, поскольку аминокислотные остатки белков в известной мере заменяемы; поэтому корреляция с ранними периодами жизни в настоящее время вряд ли правомерна.

1.3 Коллинеарная взаимосвязь между нуклеиновыми кислотами и полипептидами

Определение нуклеотидной последовательности ДНК может стать мощным методом определения аминокислотной последовательности белков. Генетический код является вырожденным в том смысле, что большая часть аминокислот описывается более чем одним кодоном. Поэтому нельзя установить нуклеотидную последовательность по коллинеарной аминокислотной последовательности. Однако удается извлечь информацию о неизвестной аминокислотной последовательности белка, анализируя исходную нуклеотидную последовательность. Реализация этого косвенного метода наталкивается на серьезное препятствие: экспериментальные ошибки, отвечающие делециям и вставкам отдельных нуклеотидов в полинуклеотидной последовательности, настолько нарушают порядок нуклеотидных триплетов, что правильное определение аминокислот оказывается пока не возможным.

Это наглядно иллюстрируется одной из последних работ [11] по белку оболочки вируса табачной мозаики, где определялась последовательность во фрагменте РНК; эта нуклеотидная последовательность была переведена в аминокислотную, которую в свою очередь можно сравнить с экспериментально найденными аминокислотами. Как видно из рис. 1.6, эти два полипептида никоим образом не идентичны ни по аминокислотной последовательности, ни по составу аминокислот.

Возможность определения последовательности полипептида по последовательности коллинеарной нуклеиновой кислоты представляет не только чисто теоретический интерес. Дело в том, что в настоящее время особенно велик прогресс как в накоплении генетического материала, так и в расшифровке последовательности нуклеиновых кислот; в то же время многие важные белки удается получить пока еще лишь в ничтожных количествах. Полный анализ

последовательности полинуклеотида, сопровождающийся исследованиями фрагментов последовательности коллинеарной полипептидной цепи для гарантии правильности фаз трансляции, может стать полезным методом определения первичной структуры белков [12].

Некоторые сегменты ДНК кодируют более чем один белок. Коллинеарность аминокислотных и нуклеотидных поледовательностей

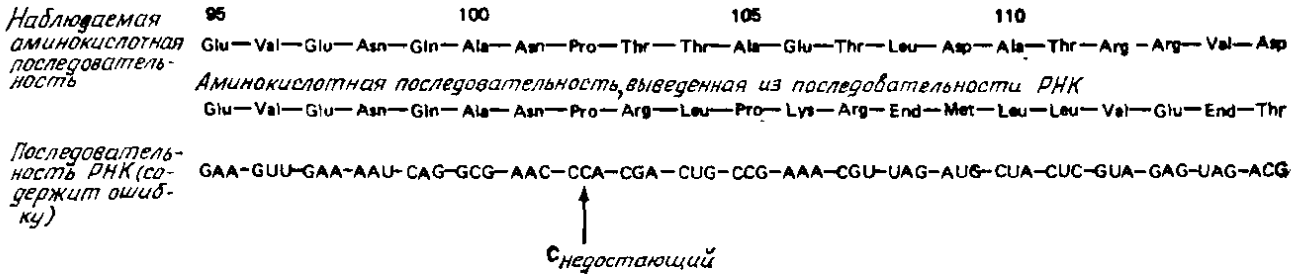


Рис. 1.6.

Установление аминокислотной последовательности по соответствующей нуклеотидной последовательности. В первой строке показан сегмент аминокислотной последовательности белка оболочки вируса табачной мозаики, а в третьей — соответствующая последовательность РНК [11].

Одна ошибка в последовательности ведет к неправильной аминокислотной последовательности (см. вторую строку), начиная от остатка 103. Ошибка, по-видимому, заключается в отсутствие С в отмеченном положении, так как при наличии С большая часть последующих аминокислот приобретает правильный порядок. Такие ошибки, связанные со «сдвигом рамки» часто обнаруживаются по появлению в неожиданном положении кодонов End (вторая строка).

привлекла особое внимание после сообщения о полной первичной структуре РНК, выделенной из вируса MS2 [13], и о ДНК, выделенной из вируса ϕX174 [12]. Бактериальные вирусы MS2 и ϕX174 — первые «живые» частицы, для которых была выяснена полная химическая структура генетического материала.

ДНК ϕX174 длиной в 5375 нуклеотидов содержит перекрывающуюся генную систему [14]: два белка считываются с одной и той же последовательности ДНК, но в различных фазах (рис. 1.7). Это означает, что общее число аминокислотных остатков в последовательности белка может быть больше, чем только одна треть

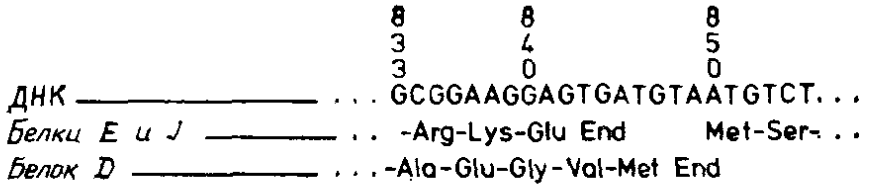


Рис. 1.7.

Перекрывание нуклеотидных последовательностей различных генов [12].

Показан сегмент 833—855 генома бактернофага φX174. В этом геноме последовательность 850—963 коднрует белок *J*, последовательность 392—847 кодируют белок *D* и последовательность 570—842 — белок *E*. Такнм образом, нуклеотндная последовательность, кодирующая белок E, полностью входит в область, коднрующую белок *D*, однако считывается в другой системе (строки 2 н 3). Кроме того, конец кодона *D* перекрывает начало кодона белка *J* в положении 850.

от числа кодирующих ее нуклеотидов. Вопрос, наблюдается ли перекрывание генов в других прокариотных или эукариотных организмах, остается пока открытым.

1.4 Свойства боковых цепей

Канонические аминокислоты отличаются только боковыми цепями. Каждая боковая цепь настолько специфична, что ее нелегко заменить какой-либо другой. Ниже будут рассмотрены свойства всех стандартных боковых цепей.

Gly увеличивает подвижность основной цепи. Это уникальная аминокислота, которая лишена боковой цепи и поэтому не проявляет асимметрии при C_α-атоме. Отсутствие боковой цепи позволяет остаткам Gly принимать необычные двугранные углы (рис. 2.3, *а*), что приводит к изгибам основной цепи. Кроме того, остатки Gly могут способствовать образованию плотной упаковки цепи белковой молекулы. Вероятно, по этим причинам остаток Gly практически не замещался в процессе эволюции.

В боковой цепи аланина находится одна метильная группа. Это небольшой неполярный остаток, не имеющий какой-либо особой предрасположенности находиться внутри или на поверхности белка. Ala — самая распространенная аминокислота, по-видимому, из-за ее простоты и доступности. Проще можно получить только Gly. Однако содержание Gly в цепи ограниченно: его избыток привел бы к чрезмерной лабильности основной цепи.

Жесткость разветвленных боковых цепей. Неполярные боковые цепи валина, изолейцина и лейцина разветвлены. Разветвление крупных боковых цепей определяет их ограниченную внутреннюю подвижность. Остаток Val содержит разветвление при C_β-атоме; его C_γ-метильные группы стерически взаимодействуют с главной цепью, уменьшая ее подвижность. Остаток Ile также разветвлен при C_β-атоме, причем ветви различаются между собой. Поэтому C_β в Ile является дополнительным асимметрическим центром. Так как все биологические реакции стереоспецифичны, используется только один стереоизомер (рис. 1.2,*б*). Присутствие Ile делает главную цепь более жесткой, как и присутствие Val. В случае Leu не возникает особых стерических взаимодействий с основной цепью, поскольку разветвление боковой цепи в этой аминокислоте находится при атоме C_γ. Жесткие боковые цепи легче фиксируются в определенном положении; понижение энтропии $\Delta S_{\text{цепи}}$ при этом не так велико (разд. 3.5), что способствует свертыванию цепи.

Все ароматические остатки содержат одно разделяющее метиленовое звено. Фенилаланин несет самую большую полностью неполярную боковую цепь. Как и в трех других ароматических боковых цепях, в этом случае между C_α и ароматическим кольцом имеется C_β-метиленовая группа. Даже при наличии метиленовой группы

подвижность боковой цепи заметно ограничена; без разделяющей метиленовой группы ароматическое кольцо создавало бы серьезные стерические затруднения при C_α-атоме и главная цепь оказалась бы слишком жесткой. Полностью неполярный пролин является весьма специфическим остатком, поскольку его боковая цепь ковалентно связана с основной цепью. В результате образуется почти жесткая структура, в которой возможны только незначительные искажения цикла. Кроме того, пролин фиксирует двугранный угол вращения вокруг связи между C_α и пептидным азотом в узком интервале $\pm 20°$.

Триптофан имеет самую объемную боковую цепь. Его небольшая полярность обусловлена наличием гетероциклического индольного цикла. Довольно подвижная боковая цепь метионина содержит один атом серы в тиоэфирной группе, что приводит к появлению дипольного момента. Все самые крупные неполярные остатки: Val, Ile, Leu, Phe, Pro, Trp и в меньшей степени Met — располагаются преимущественно внутри белковых молекул.

Полярные боковые цепи образуют водородные связи. Типичными полярными и нейтральными боковыми цепями обладают цистеин, серин, треонин, аспарагин, глутамин и тирозин. Из них Cys выполняет особую роль, поскольку он способен образовывать поперечные мостики (цистины) между различными частями основной цепи путем присоединения к другому остатку Cys (разд. 4.2). Остатки Ser и Thr несут гидроксильные группы, которые могут образовывать водородные связи. В Thr, имеющем асимметрический C_β-атом, активным является только один стереоизомер (рис. 1.2,*б*). Несущие амидные группы аспарагин и глутамин также способны к образованию водородных связей, причем амидные группы функционируют в качестве доноров водорода, а карбонильные — в качестве акцепторов. По сравнению с аспарагином у глутамина имеется лишнее метиленовое звено, придающее полярной группе большую подвижность и ослабляющее ее взаимодействие с основной цепью. Полярная гидроксильная группа Tyr, для которой $pK = 10{,}1$ может диссоциировать при высоких значениях pH. Поэтому Tyr до некоторой степени аналогичен заряженной группе: образованные им водородные связи довольно прочны. Нейтральные полярные остатки могут располагаться как на поверхности, так и внутри белковых молекул. Находясь внутри, они обычно образуют водородные связи между собой или с полипептидным остовом (разд. 3.6).

His подходит для каталитических целей. Гистидин содержит гетероциклическую ароматическую боковую цепь, для которой $pK = 6{,}0$ (табл. 1.1). В физиологическом интервале pH его имидазольное кольцо может либо оставаться незаряженным, либо присоединять ион H^+ из раствора. Поскольку такое присоединение происходит достаточно легко, гистидин может выполнять роль ката-

лизатора химических реакций. Установлено, что он присутствует в большинстве активных центров ферментов.

Заряженные боковые цепи обычно находятся на поверхности молекул. Остатки аспарагиновой и глутаминовой кислот при физиологических значениях pH отрицательно заряжены. Из-за короткой боковой цепи карбоксильная группа Asp довольно жестко фиксирована относительно главной цепи. Это может быть причиной того, что карбоксильные группы активных центров принадлежат главным образом остаткам Asp, а не Glu. Как правило, оба остатка находятся на поверхности белков.

Большая часть положительно заряженных остатков лизина и аргинина также находится на поверхности белковых молекул. Эти длинные и гибкие остатки обычно не имеют строго фиксированной конформации. Весьма подвижные в окружающей среде, они повышают растворимость белковых глобул. Однако в некоторых случаях боковые цепи Lys и Arg участвуют в образовании внутренних солевых мостиков, а также способствуют катализу. Благодаря своему расположению на поверхности остатки Lys и Arg чаще других подвергаются воздействию ферментов, которые либо модифицируют их боковые цепи, либо расщепляют полипептидную цепь со стороны карбонильного углерода (разд. 4.3).

Гидрофобность зависит от площади контактной поверхности и

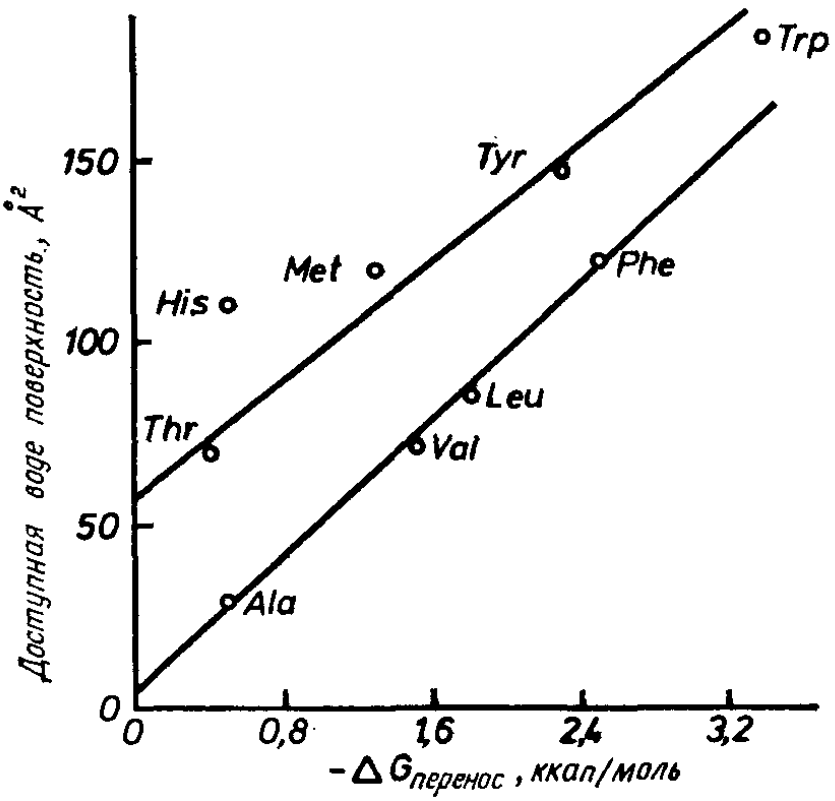


Рис. 1.8.

Доступная воде поверхность [17] боковой цепи аминокислоты как функция ее гидрофобности [16], т. е. как функция свободной энергии переноса из воды в этанол или диоксан, которые по предположению имитируют условия во внутренней области белка. Доступная воде поверхность определена на рис. 1.9.

числа диполей. Как указано в гл. 3 и 8, процесс свертывания полипептидной цепи зависит от гидрофобности (неполярности) боковых цепей, поскольку образование гидрофобного ядра глобулы, по-видимому, является одной из основных движущих сил свертывания. Козман [15] и позднее Носаки и Танфорд [16] определили эту гидрофобность путем измерения изменения свободной энергии при переносе аминокислот из воды в органические растворители. Чтобы получить разность свободных энергий $\Delta G_{\text{перенос}}$ для бо-

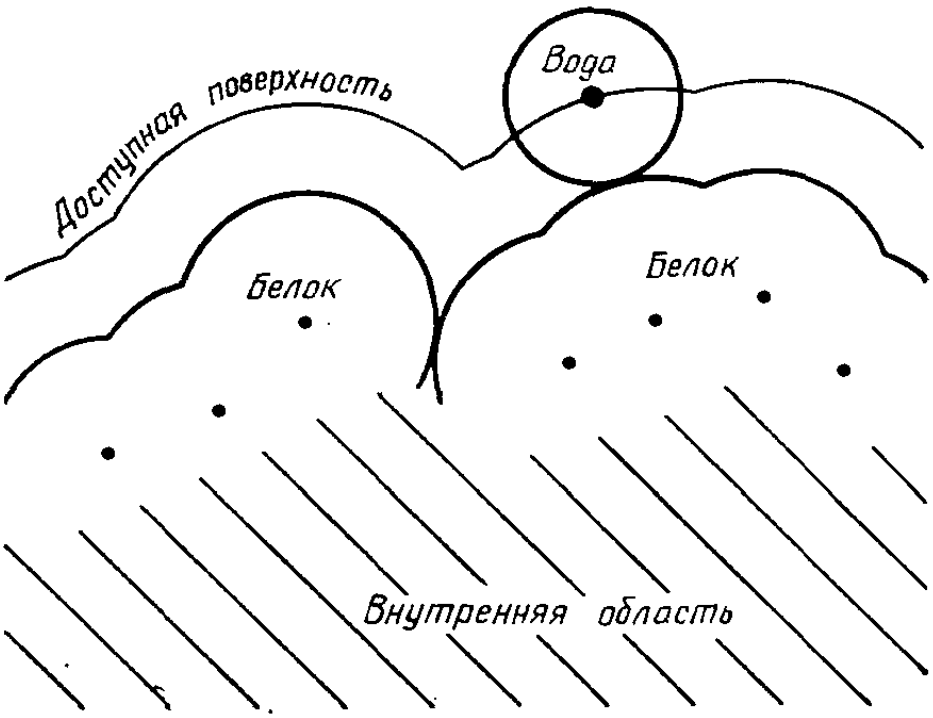


Рис. 1.9.

Доступная воде поверхность белка (или боковой цепи).

Поверхность белка описывается огибающей вандерваальсовых радиусов атомов, находящихся на поверхности (обозначены точками). Внутренняя область белка заштрихована. Молекулы воды представляются в виде сфер раднусом 1,4 Å. «Доступная воде поверхность» определяется как площадь, опнсываемая центром молекулы воды при ее скольжении по вандерваальсовой поверхности белка (или боковой цепи). Отметим, что в щелях существуют недоступные молекулам воды области белка (или боковой цепи).

ковой цепи, необходимо вычесть вклад главной цепи. Это можно сделать путем вычитания значения для Gly. Результаты приведены в табл. 1.1.

Как показал Чотиа [17] (рис. 1.8), для полностью неполярных боковых цепей существует линейная зависимость между $\Delta G_{\text{перенос}}$ и площадью поверхности, доступной воде. Площадь доступной воде поверхности определена на рис. 1.9. Аналогичная зависимость, но с уменьшением $\Delta G_{\text{перенос}}$ примерно на 1,5 ккал/моль найдена для остатков, содержащих один диполь, например Ser, Thr, Met, Tyr и Trp (разд. 3.5). Положение Trp на рис. 1.8 показывает, что индольный цикл представляет собой слабый диполь. Диполь Met, по-видимому, является более сильным. Это подтверждено обнаружением двух водородных связей с Met-180 в химотрипсине [18].

1.5 Эмпирические аналогии между аминокислотными остатками

Биологическое разнообразие облегчает анализ. Существующее множество биологических видов, наверное, никогда не сможет быть объяснено. Особенность каждого из них зависит от столь большого числа случайных событий (событий, которые мы не знаем и полагаем случайными), что в биологии приобрел большое значение чисто исторический подход. Хотя на первый взгляд такой подход выглядит весьма неподходящим для объяснения основных принципов, тем не менее на практике он оказался чрезвычайно полезным.

Частота замен выявляет сходство. В отношении боковых цепей, например, можно найти аналогию при анализе частоты замен аминокислот в последовательностях соответствующих белков гомологичных организмов. Для получения более точной меры подобия наблюдаемые частоты нормируют, т. е. относят к частоте замен, которые могут возникнуть случайно. Из-за неравномерного распределения аминокислот в природе последние характеристики специ-

Таблица 1.2

Матрица частот взаимных замещений на эволюционном пути в 256 точечных мутациях на 100 остатков (PAM)[a]

Gly	29																			
Pro	12	14																		
Asp	11	10	23																	
Glu	11	11	19	20																
Ala	13	14	12	12	14															
Asn	11	10	14	12	12	17														
Gln	10	11	14	14	11	12	21													
Ser	12	11	12	11	13	13	11	13												
Thr	10	10	10	10	12	12	11	13	18											
Lys	7	8	10	10	9	12	11	10	9	24										
Arg	4	4	7	7	5	10	13	7	6	22	75									
His	6	6	9	8	8	13	12	9	9	11	20	59								
Val	7	8	7	8	10	8	9	9	11	8	5	6	22							
Ile	6	6	6	7	8	7	8	8	10	7	5	7	21	25						
Met	5	6	6	6	8	7	8	8	9	8	10	6	16	16	26					
Cys	6	4	5	5	7	6	5	11	9	4	2	3	11	9	12	166				
Leu	4	4	5	6	6	5	6	6	7	6	4	6	15	18	21	5	40			
Phe	2	2	2	3	4	3	3	4	5	3	4	7	7	11	11	2	12	70		
Tyr	1	1	1	1	2	2	2	3	2	3	3	7	3	6	6	1	6	66	137	
Trp	1	1	1	1	1	2	2	2	2	2	3	15	2	4	4	1	3	41	46	414

[a] Данные взяты из работы [20]; единица PAM определена в разд. 9.1. Каждая величина есть отношение вероятности того, что данная пара аминокислот может находиться в данном месте белков данного рода, к вероятности того, что это может произойти случайно. Все величины умножены на 10. Например, в двух последовательностях на эволюционном расстоянии в 256 PAM можно найти Leu и Met в 2,1 раза чаще, чем это могло бы произойти случайно. Пара Asp и Tyr встречается всего в 0,1 раза чаще, чем случайно, т. е. для данной свернутой цепи замещение Asp → Tyr или Tyr → Asp весьма невыгодно, а значит запрещено. Аминокислоты расположены так, что величина $\sum_{ij} m_{ij} (i-j)^2$ минимальна. Таким образом, сближены аминокислоты, которые часто замещают друг друга в данном положении свернутой цепи.

фичны для каждой замены: замена Ala-Leu в $13,0 \cdot 7,8/3,3 \cdot 2,2 =$ $= 14$ раз более вероятна, чем замена Phe-Tyr (см. графу «Относительное содержание в белках *E. coli*» в табл. 1.1).

Группы взаимозаменяемых остатков. Наблюдаемые вероятности замен даны в табл. 1.2. Аминокислоты расположены таким образом, чтобы второй момент относительно диагонали, $\sum_{ij} m_{ij}(j-i)^2$, был минимален. Такое расположение выявляет обменивающиеся группы. Замена одной аминокислоты на другую происходит преимущественно внутри таких групп. Поэтому эти аминокислоты большей частью похожи одна на другую в отношении их общего влияния на структуры белков.

Эмпирические характеристики остатков в табл. 1.2 показывают, что ароматические аминокислоты Phe, Tyr и Trp образуют группу взаимозаменяемых остатков и что их роли близки. Кроме того, аналогичную группу образуют положительно заряженные остатки Lys, Arg и His. Отметим, что отрицательно заряженные остатки Glu и Asp заметно удалены от положительно заряженных остатков. Таким образом, замещения между противоположно заряженными остатками редки. Высокую частоту замещений следовало бы ожидать, если бы заряженные остатки всегда находились на поверхности молекулы и не выполняли бы иных функций, кроме увеличения растворимости белка.

Важен размер боковой цепи. Очевидно, что крупные алифатические неполярные остатки Val, Leu и Ile (но не Pro!) образуют группу взаимозаменяемых аминокислот, которая включает также слабополярные Met и Cys. Все малые остатки (Ser, Thr, Asp, Asn, Gly, Ala, а также Glu, Gln и Pro) образуют отдельную группу. Таким образом, оказываются объединенными такие не похожие остатки, как неполярный Pro и сильнополярный Asp. Следовательно, создается впечатление, что для белковых структур размер боковой цепи почти так же важен, как и ее химическая природа.

Заключение

В этой главе мы познакомились с каноническими аминокислотами, основными строительными блоками белка, которые являются частью основной схемы организации всех организмов. Исходя из аминокислот, можно объяснить некоторые аспекты этой универсальной схемы. Однако потребуется еще много времени, чтобы понять, почему была развита именно эта схема. Каждый аминокислотный остаток обладает уникальными свойствами и выполняет вполне определенную роль в данном месте белковой молекулы. Несмотря на это в ходе белковой эволюции происходили и происходят некоторые замены. Анализ частоты замен позволяет выявить аналогию среди аминокислот и подразделить все аминокислотные остатки на четыре основные группы.

Глава 2 Структурная роль пептидной связи

2.1 Синтез на рибосомах

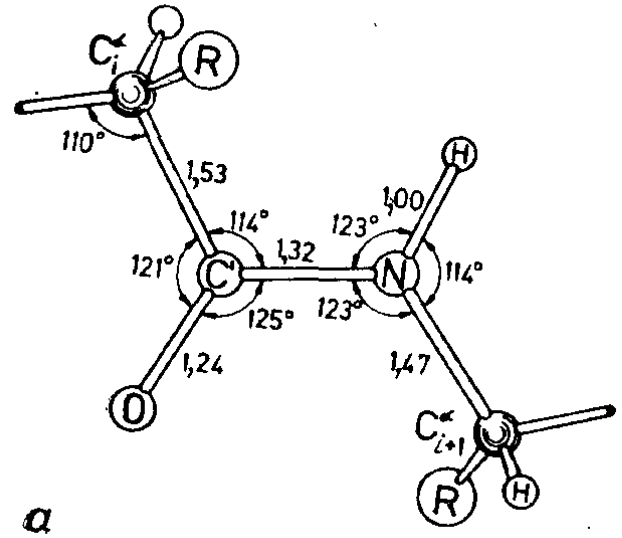
Полипептиды образуются в результате серии высокоспецифичных реакций. Аминокислоты встраиваются в полипептидную цепь на рибосомах в клетке. Полимеризация основана на образовании амидных связей, которые обычно называют «пептидные связи». Направление цепи определяют от аминоконца (N-конец) к карбоксильному концу (С-конец), как показано на рис. 2.2. Это совпадает с направлением синтеза цепи in vivo, которое в свою очередь отвечает направлению 5′ → 3′ информационной РНК.

Свободная энергия регулирует точность. При образовании одной пептидной связи расщепляются четыре связи аденозинтрифосфата (ATP) и гуанозинтрифосфата (GTP) [22] с выделением свободной энергии, ~25 ккал на моль аминокислоты [23]. Лишь часть этой энергии расходуется на образование пептидной связи в эндергонической реакции, стандартная свободная энергия которой в воде составляет ~5 ккал/моль [24]. Остальное затрачивается на трансляцию информационной РНК к полипептиду и на максимально точное проведение трансляции. Гидролитическому распаду пептидных связей препятствует высокий барьер энергии активации, который, однако, легко преодолевают многочисленные ферменты, расщепляющие белок (протеазы).

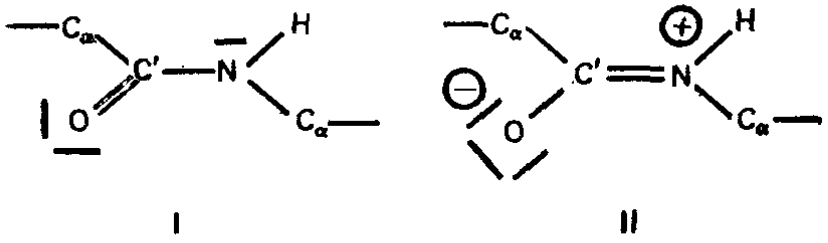
Малая энергия образования связи позволяет быстро адаптироваться к внешним изменениям. С организационной точки зрения пептидная связь очень удобна, поскольку ее можно получить и расщепить с относительно низкими энергетическими затратами. Это значительно увеличивает шансы организма на превращение его ресурсов в необходимую форму, т. е. помогает получать более удобную и надежную защиту от неблагоприятных изменений в окружающей среде. В случае более прочных связей, как, например, на алифатических концах жирных кислот, ответные изменения были бы медленнее, и поэтому приспосабливаемость организмов к конкуренции в процессе эволюции была бы слабее.

2.2 Параметры пептидной связи

Вращение вокруг пептидной связи заторможено. Геометрические параметры пептидной связи приведены на рис. 2.1, *а*. Эти данные получены Полингом и Кори [25] на кристаллических структурах молекул, содержащих одну или несколько пептидных связей. Наиболее неожиданным оказалось малое расстояние между атомами C′ и N, которое на 0,15 Å, или на 10%, короче обычного. Кроме того, двойная связь C′—O на 0,02 Å длиннее, чем в альдегидах и кетонах [26]. Полинг и сотр. объяснили этот эффект резонансом двух предельных структур, показанных на рис. 2.1, *б*. В структуре I

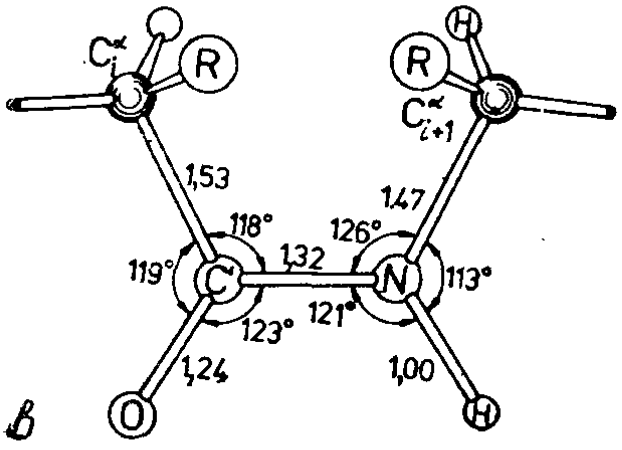


а



I II

б



в

Рис. 2.1.

Амидная или пептидная связь [29].

а — валентные углы и расстояния для *транс*-пептидной связи по даиным Полнига и Кори [25]. Атомы C′ и N находятся в плоскости. *б* — две предельиые электронные структуры связи; форма I допускает свободное вращение; у формы II, имеющей большой дипольный момеит, вращение заторможено. Гибридная структура состоит на 60% из формы I и иа 40% из формы II, что следует из длины связи C′—N. Эиергия резоианса около 20 ккал/моль. *в* — валентные углы и межатомиые расстояния в редко встречающейся *цис*-пептидной связи.

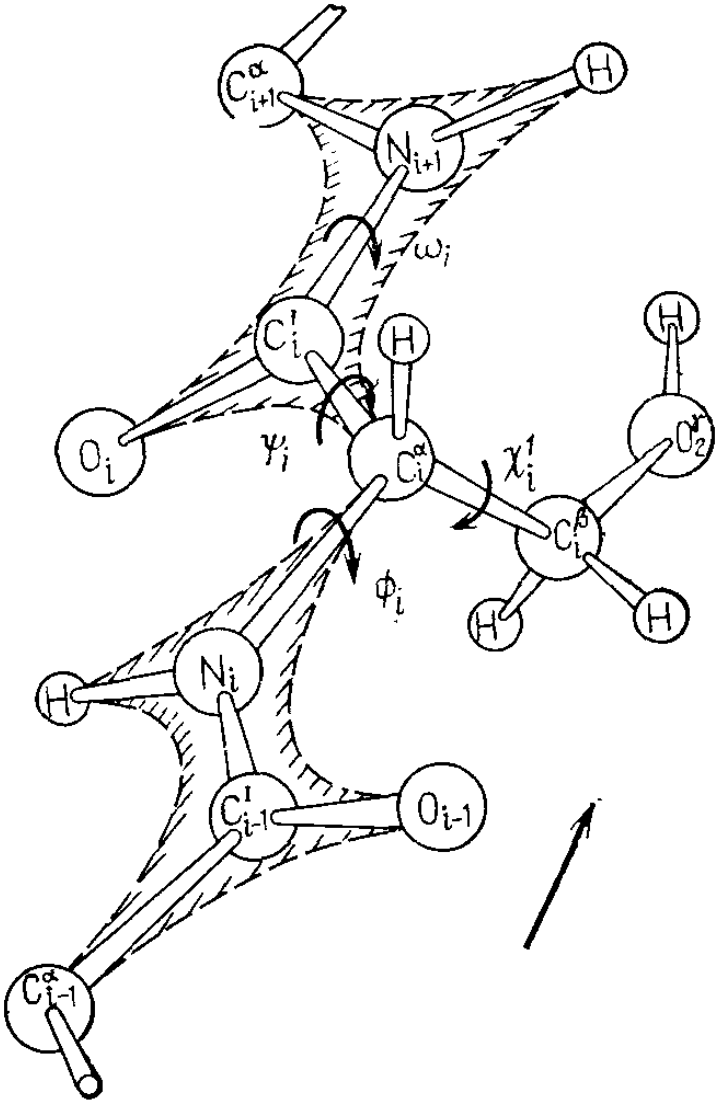


Рис. 2.2.

Определение двугранных углов в полипептидной цепи; направление цепи указано. Для иллюстрации определения двугранных углов боковой цепи использован остаток Ser.

Обозначение следует рекомендациям комиссии IUPAC—IUB от 1969 г. [21]. При обозначении прочих двугранных углов в боковых цепях следует обращаться к этим рекомендациям. Нулевые значения двугранных углов определены следующим образом:

$$\omega_i = 0 \text{ при } C_i^{\alpha} - C_i' \textit{ цис к } N_{i+1} - C_{i+1}^{\alpha}$$

$$\psi_i = 0 \text{ при } C_i^{\alpha} - N_i \textit{ транс к } C_i' - O_i$$

$$\Phi_i = 0 \text{ при } C_i^{\alpha} - C_i' \textit{ транс к } N_i - H_i$$

$$\chi_i^1 = 0 \text{ при } C_i^{\alpha} - N_i \textit{ цис к } C_i^{\beta} - O_i^{\gamma}$$

В приведенной основной цепи углы равны $\varphi_i = 180°$ и $\psi_i = 180°$. Как показано штриховкой, обычно шесть атомов C_i^{α}, C'_i, O_i, N_{i+1}, C_{i+1}^{α}, H_{i+1} расположены в одной плоскости, т. е. угол ω равен 180°. H_i и H_{i+1} — атомы водорода амидной группы. Указанные направления вращения ω, ψ и Φ положительны, если смотреть со стороны атома на N-конце данной пептидной связи и вращать С-конец этой связи в направлении, указанном стрелками (по часовой стрелке). Это же определение относится к χ, если смотреть от атома, ближайшего к C_α -атому.

связь C′ — N включает только σ-электроны, аксиальная симметрия которых допускает свободное вращение, тогда как в структуре II σ- и π-электроны связи C′— N приводят к большому дипольному моменту и к заторможенности вращения.

Судя по различию в длинах связей, результирующая гибридная структура включает формы I и II в отношении 3 : 2, с π-электронами, делокализованными по связям C′—O и C′—N. Термодинамические данные показывают, что энергия резонанса составляет около 20 ккал/моль [23]. Поскольку структура II плоская, гибридная структура также плоская, т. е. шесть атомов C^{α}_{i}, C'_{i}, O_{i}, N_{i+1}, H_{i+1}, C^{α}_{i+1} (рис. 2.2) находятся в одной плоскости.

Каждый остаток в основной цепи полипептида описывается двумя двугранными углами. При жесткой пептидной связи и довольно жестких длинах связей и валентных углах, конформация полипептидной цепи по существу описывается двугранными углами ϕ и ψ при C_{α}-атомах, как показано на рис. 2.2. Это описание соответствует номенклатуре IUPAC — IUB от 1969 г. [21]. Введен и торсионный угол ω, хотя вращение вокруг связи C′ —N заторможено. Все двугранные углы в боковых цепях обозначаются буквой χ с индексом, который может меняться от единицы до пяти. Показанная на рис. 2.2 боковая цепь Ser не имеет разветвлений, поэтому здесь потребовался только один индекс. Исчерпывающее описание обозначений можно найти в рекомендациях IUPAC — IUB [21].

Следует упомянуть, что ранее, в 1966 г., для двугранных углов ϕ и ψ основной цепи было полуофициально введено другое начало отсчета [27]:

$$\phi_{1969} = \phi_{1966} - 180^{\circ}$$
$$\psi_{1969} = \psi_{1966} - 180^{\circ}$$

Поскольку за время между этими двумя датами появилось довольно много публикаций по определению белковых структур, то старое обозначение все еще можно иногда встретить в литературе.

2.3 Стерические затруднения

Значения двугранных углов ссновной цепи ограничены. Конформационная подвижность полипептидной цепи ограничена. Жесткая пептидная связь заметно ограничивает подвижность полипептидной цепи. Кроме того, пептидное звено довольно громоздко и вызывает существенные стерические затруднения. Во время работы над моделями структуры коллагена Рамачандран и сотр. [28] исследовали вопрос, каким образом дальнее стерическое взаимодействие ограничивает свободу вращения вокруг связей N—C_{α} и C_{α}—C′ цепи, т. е. интервал допустимых углов ϕ и ψ. Для этой цели они определили по известным кристаллическим структурам контактные вандерваальсовы расстояния между интересующими их ато-

мами и получили величины, характеризующие так называемую нижнюю нормальную границу (табл. 2.1). Кроме того, они определили «нижнюю предельную границу», которая была меньше «нижней нормальной» на 0,10 Å [29]. Оказалось, что с погрешностью 0,05 Å расстояния, отвечающие контактам, равны суммам контактных радиусов, т. е. вклад данного атома в контактное расстояние не зависел от его партнера. Следовательно, используемые в расчетах атомы выбранных размеров прощупывали пространство

Таблица 2.1

Эмпирические нижние границы невалентных контактных расстояний [29]

Тип контакта	Нормальная граница, Å	Предельная граница, Å	Тип контакта	Нормальная граница, Å	Предельная граница, Å
H . . .H	2,0	1,9	O . . .N	2,7	2,6
H . . .O	2,4	2,2	O . . .C	2,8	2,7
H . . .N	2,4	2,2	N . . .N	2,7	2,6
H . . .C	2,4	2,2	N . . .C	2,9	2,8
O . . .O	2,7	2,6	C . . .C	3,0	2,9

по существу так же, как это происходит в моделях с заполнением пространства, использующих жесткие сферы контактного радиуса.

Конформационное пространство, интервал углов ϕ и ψ, допустимых для остатка без боковой цепи (Gly), показаны на рис. 2.3, *а*. Эта $\phi\psi$-карта построена для жесткой пептидной связи с параметрами, приведенными на рис. 2.1, *а*, и с учетом обеих, нижней нормальной и нижней предельной, границ контактных расстояний (табл.2.1). Разметка карты выполнена по старой номенклатуре [27]. При нанесении новых значений ϕ, ψ [21] начало сдвигается в центр карты. Остаток Gly имеет большую и непрерывную разрешенную область, занимающую около 50% пространства. Явно запрещенные контакты возникают между атомами H_{i+1} и O_{i-1} в центре, а также между атомами O_{i-1} и O_i. Менее заметные затруднения обнаружены между атомами H_i и H_{i+1}, O_{i-1} и C'_i, а также O_{i-1} и N_{i+1}. Наименее значимо стерическое взаимодействие между атомами N_i и H_{i+1}.

У всех остатков кроме Gly и Pro стерически разрешены по существу одни и те же области. При наличии C_β -атома, т. е. для всех остатков кроме Gly, пространство разрешенных конформаций резко суживается до областей, показанных на рис. 2.3, *б*. В модели жестких сфер с радиусами, отвечающими нормальным и предельно малым границам, разрешенными остаются соответственно только 8 и 22% пространства. На карте показаны контакты атомов C_β, ограничивающие три большие области пространства Gly. Две полностью разрешенные области включают соответственно конформа-

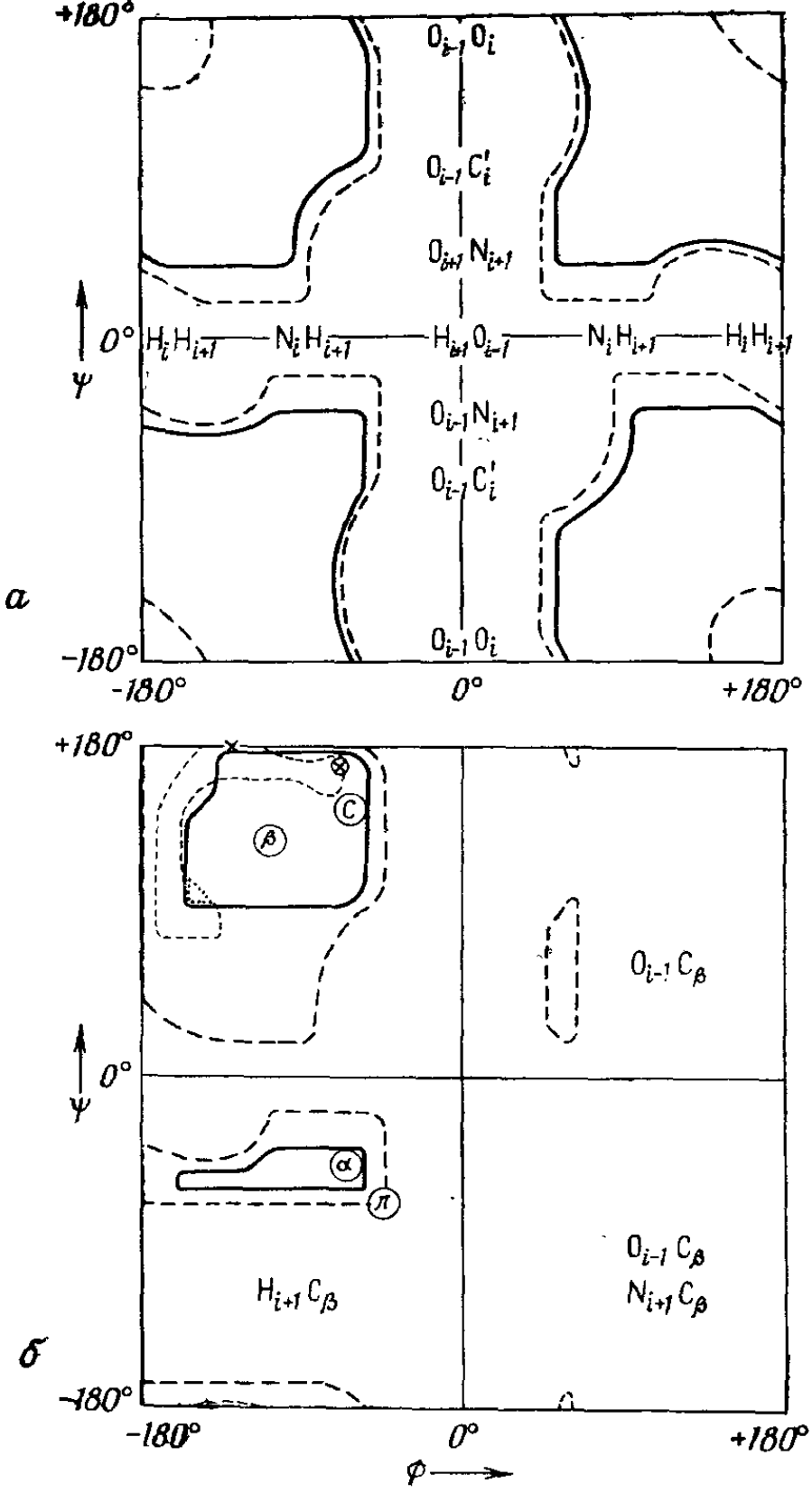


Рис. 2.3.

Разрешенные области для двугранных углов основной цепи. Приняты параметры пептидной связи Полинга и Кори, приведенные на рис. 2.1 [29]. Эти карты часто называют «картами Рамачандрана». *а* — карта глицина, не имеющего боковой цепи, с использованием модели жестких сфер с нормальными и с нижними граничными (— — —) контактными расстояниями, приведенными в табл. 2.1. В некоторых областях указаны запрещенные контакты. Обозначения атомов даны на рис. 2.2. б — карта аминокислот, содержащих C_β -атомы, построенная с использованием той же модели жестких сфер. Указаны конформации правой α-спирали, β-складчатого листа и коллагена. Отмечены стерические затруднения, вызванные C_β -атомом, которые сводят большую область глицина к указанной на рисунке меньшей области. Три разрешенные области иногда обозначают как α_R — правая α-спираль, область вокруг значка α, α_L — левая α-спираль, область $(\phi, \psi) \approx (+60°, +60°)$, ε — вытянутая цепь, область вокруг значка β. В предположении модели жестких сфер (- - - нормальные радиусы; — — — радиусы, отвечающие нижним границам) указаны области, разрешенные для поли-L-*цис*-алаинна. Конформация, приведенная на рис. 2.6, обозначена х, а конформация, отвечающая поли-L-пролину I (полностью *цис*), обозначена как х, обведенный кружком.

ции правой α-спирали и параллельной и антипараллельной β-структур. Условно разрешенная область (ϕ, ψ)= (+ 60°, + 60°) содержит левую α_L-спираль. Помимо Gly рис. 2.3, *б* не описывает только остаток Pro. Pro ведет себя иначе, поскольку его боковая цепь, связанная с пептидным азотом, фиксирует $\phi = -60° \pm 20°$. Разрешенная вариация углов ϕ в случае Pro отражает возможности изменения геометрии пирролидинового цикла.

2.4 Конформационная энергия

Модель жестких сфер — хорошее приближение. Данные Рамачандрана и сотр. [28, 29], выражаемые $\phi\psi$-картой (которую часто называют «карта Рамачандрана», см. рис. 2.3, *б*), подтверждаются результатами, полученными при исследовании кристаллических глобулярных белков. На рис. 2.4 сведены все ϕ и ψ-углы, найденные в 13 белках. Самая высокая плотность распределения экспериментальных точек наблюдается вблизи (—60°, —60°) в положении правой α-спирали, что отражает высокое содержание α-спиралей в глобулярных белках. Другой максимум в распределении находится вблизи (—90°, + 120°) и отвечает вытянутой цепи с остатками, образующими β-складчатый лист. Поскольку плотность вблизи (—90°, 0°) также довольно высока, отталкивание между N_i и H_{i+1} не настолько существенно, как это следует из модели жестких сфер.

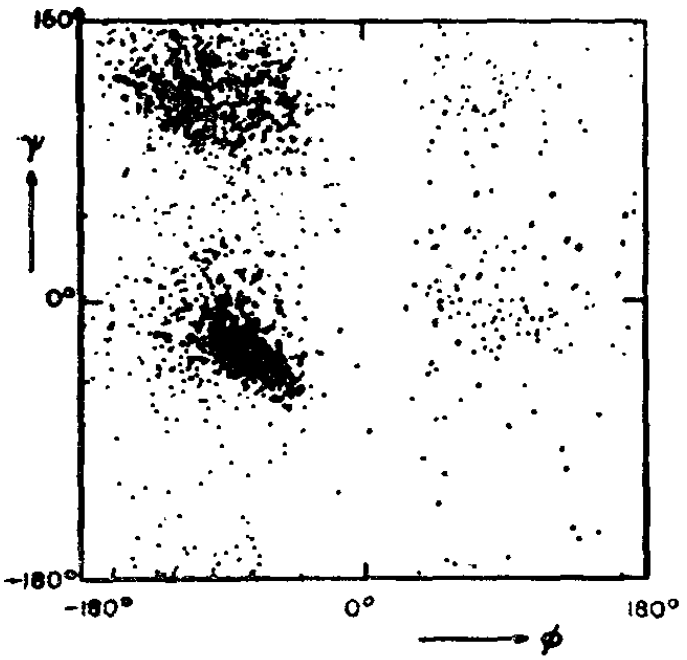


Рис. 2.4.

Двугранные углы основной цепи для приблизительно 2500 остатков 13 белков [30]. Эту карту иногда называют картой Рамачандрана.

Область, отвечающая левой α_L-спирали, заполнена довольно слабо по сравнению с соседними областями. Левую α_L-спираль пока еще не удалось наблюдать. Около 10% точек попадают в области, запрещенные для всех остатков кроме Gly, на правую и нижнюю левую части диаграммы. Исходя из частоты Gly, указанной в табл. 1.1, и допуская равномерное распределение остатков Gly по доступному для них конформационному пространству, можно было бы ожидать появления только половины из этих точек. Следовательно, около 5% всех остатков, содержащих C_β-атомы, находятся в запрещенной области. Этот факт подтверждают более точные картины распределения для восьми белков, на которых ϕ,

ψ-величины нанесены отдельно для каждого типа остатков [31]. Следует помнить, однако, что использованные экспериментальные данные также могут содержать ошибки, вызванные неправильной интерпретацией карт электронной плотности.

Карта потенциальной энергии более точно выявляет стерические препятствия. Модель жестких сфер представляет довольно грубый подход к описанию стерических затруднений. Ясно, что его можно улучшить, используя вместо жестких сфер набор функций потенциальной энергии. Карты потенциальной энергии были рассчитаны многими авторами [32—34] для всевозможных остатков. В качестве примера на рис. 2.5 приведена карта потенциальной энергии для аланина. Какого-либо качественного различия по сравнению с картой жестких сфер здесь нет. Потенциальные энергии правой α_R-спирали и левой α_L-спирали выше, чем энергия вытя-

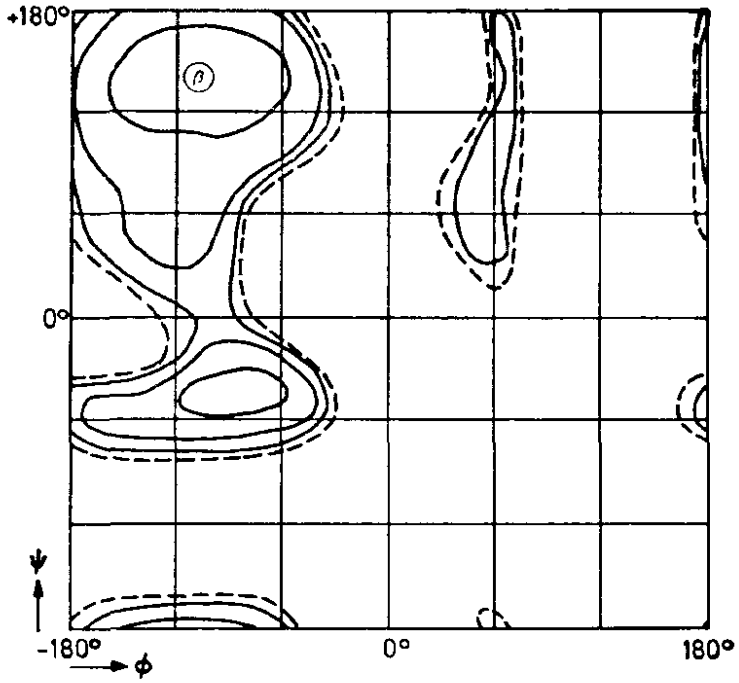


Рис. 2.5.

Распределение потенциальной энергии в плоскости (ϕ, ψ) для пары пептидных звеньев с остатком Ala между ними [34].

Карта построена с параметрами пептидной связи Полинга—Кори (рис. 2.1,*а*). Изоэнергетические линии проведены с интервалом 1 ккал/моль в отрицательную область от нуля. Нулевая эквипотенциаль показана штрихами. Отмечено положение скрученного β-листа. Потенциальная поверхность изменится, если учесть водородные связи с пептидными звеньями, удаленными по цепи (например, в α-спирали).

нутой цепи, примерно на 0,5 и 2,5 ккал/моль соответственно. В случае правой α_R-спирали это увеличение энергии кажется преодолимым благодаря выгодности прочного водородного связывания внутри α-спирали. У левой α_L-спирали нет явных структурных преимуществ по сравнению с правой. Обе спирали представляют собой цилиндры одинаковой формы. Поэтому если в структуре нужен цилиндр, нет причин для использования энергетически невы-

годной левой α_L-спирали, если другая спираль может с успехом выполнить то же назначение.

На карте потенциальной энергии виден низкоэнергетический перешеек между областью вытянутой цепи (ϕ, ψ) = (—120°, +120°) и областью α-спирали (—60°, —60°). Такого мостика нет в модели жестких сфер, даже если использовать предельно малые контактные радиусы. Поскольку эта так называемая мостиковая область, согласно данным по кристаллическим белкам (рис. 2.4), отличается высокой заселенностью, она и в действительности легко доступна. Модель жестких сфер в этом случае приводит к ошибкам. Стерические затруднения между атомами N_i и H_{i+1} (рис. 2.3, *а*), возникающие при этих углах, компенсируются благодаря диполь-дипольному взаимодействию, соответствующему слабой водородной связи. Подобные взаимодействия были найдены в кристаллах модельных соединений [35].

Пептидная связь обладает некоторой гибкостью. Конформационные карты на рис. 2.3 и 2.5 построены в предположении жесткой пептидной связи, имеющей параметры Полинга — Кори (рис. 2.1,*а*). Дальнейшее уточнение модели требует введения потенциалов изменения валентных углов, длин связей и торсионного вращения вокруг пептидной связи. Это, естественно, делает конформационное пространство одного остатка многомерным, а любое непосредственное использование или исчерпывающее описание — затруднительным. Приведем для оценки отклонения, отвечающие увеличению потенциальной энергии на 1 ккал/моль:

отклонение валентного угла	~ 5°
отклонение длины связи	~ 0,05Å
отклонение торсионного угла	~ 12°

Кроме того, следует также учитывать выходы атомов из плоскости [36]. В принципе, гибкость можно учесть на двумерной энергетической карте, если допустить возможность всех гибких деформаций и в каждом случае рассчитать минимальную энергию для данной пары углов ϕ, ψ. Отметим, что если принимать во внимание гибкость всех звеньев остова, его конформацию нельзя описать только значениями углов ϕ и ψ.

Таким образом, если пептидная связь обладает заметной гибкостью, то конформационная карта, построенная при фиксированных параметрах этой связи (рис. 2.5), может оказаться неточной в тех областях, которые отвечают слабым стерическим затруднениям, как, например, в областях, разрешенных только для Gly, поскольку их можно устранить небольшими отклонениями валентных и торсионных углов, а также длин связей. Этим объясняются те 5% случаев на рис. 2.4, в которых значения ϕ, ψ попадают в области, запрещенные, согласно модели жестких сфер, из-за стерического взаимодействия с C_β-атомом.

2.5 *Цис-* и *транс*-конфигурации

Одна из альтернатив должна быть универсальной. До сих пор мы предполагали, что все пептидные связи находятся в *транс-*, а не в *цис*-конфигурации (рис. 2.1, *в*). Очевидно, что пептидный синтез на рибосомах стереоспецифичен и поэтому должен приводить либо к *цис-*, либо к *транс*-связям. В любом случае одна из альтернатив должна быть универсальной, иначе механизм синтеза и генетический код усложнились бы, что неблагоприятно с организационной точки зрения. Точно так же невыгодна и пост-трансляционная изомеризация, поскольку для этого потребовалось бы много специальных ферментов. Спонтанная изомеризация в процессе свертывания полипептидной цепи также невозможна, поскольку энергетический барьер между *транс-* и *цис*-конфигурациями составляет $\sim$20 ккал/моль. Этот барьер ниже только для остатка Pro, для которого изменения геометрии пирролидинового цикла снижают барьер до 13 ккал/моль. В процессе свертывания остатки Pro могут претерпевать спонтанную изомеризацию (разд. 8.2). Энергетическое различие между *цис-* и *транс*-изомерами составляет $\sim$2 ккал/моль в пользу *транс*-изомера; это различие меньше всего для Pro, поскольку для него переход от *транс-* к *цис*-изомеру приводит к замене контакта между атомами C_i^α и C_{i+1}^α всего лишь аналогичным контактом между атомами C_i^α и C_{i+1}^δ (рис. 1.1).

В глобулярных белках найдены несколько *цис*-пептидов. *цис*-Связи обнаружены во многих циклических пептидах, в особенности со стороны N-конца перед остатками Pro [37, 38]. В глобулярных белковых структурах обнаружено только несколько *цис*-связей, например перед Pro-93 и Pro-114 в рибонуклеазе S [39], перед Pro-168 в субтилизине [40], перед Pro-116 в нуклеазе стафилококка [*Е. Хазен*, частное сообщение], перед Pro-8 и Pro-95 вариабельной цепи белка Бенс-Джонса [42] и между Ser-197 и Tyr-198 в карбоксипептидазе-А [43]. Молекулы синтетического поли-L-Pro I содержат только *цис*-связи. Таким образом, из всех типов стандартных аминокислот Pro *в наибольшей мере* способствует образованию *цис*-связи, что согласуется с энергетическими расчетами. Однако, как правило, *цис*-связи встречаются крайне редко.

В случае реализации полностью *цис*-белка конформационная свобода была бы чрезмерно ограничена. Если есть необходимость в одном универсальном изомере, то почему выбран *транс-*, а не *цис*-изомер? Убедительным аргументом в пользу *транс*-изомера служит строгое ограничение разрешенного конформационного пространства в полностью *цис*-полипептидной цепи. Для изолированного C_α-атома стерические затруднения в случае *цис-* и *транс*-изомеров не отличаются. Однако для гипотетической полностью *цис*-поли-

пептидной цепи отталкивание между соседними боковыми цепями резко возрастает. Это показано на рис. 2.6 для поли-L-*цис*-Ala. Цепь в этом случае образует резкие изгибы, сближая соседние C_α-атомы.

Рамачандран и Сасисекхаран [29] попытались найти конформации, разрешенные для такой цепи. В предположении модели жестких сфер с нормальными контактными радиусами они обнаружили только очень небольшую разрешенную область при (—155°, +105°). Даже для модели жестких сфер с радиусами, отвечающими нижним границам, разрешенная область оказалась очень малой, что можно видеть на рис. 2.3, *б*. В отличие от цепей, состоящих только из *цис*-пептидных групп, в случае *транс*-конфигурации разрешенные конформационные состояния цепи оказываются достаточными для об-

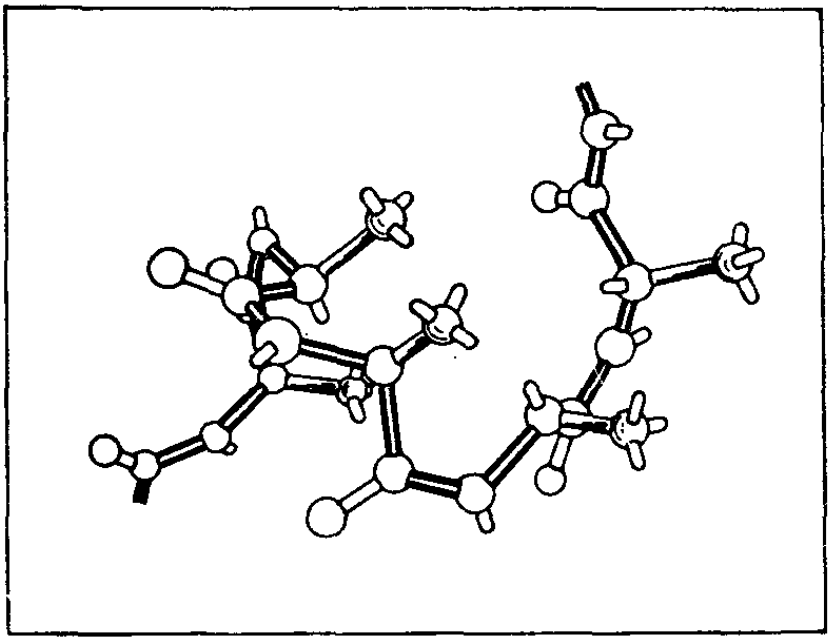


Рис. 2.6.

Поли-L-*цис*-аланин. Углы (ϕ, ψ) отмечены значком x на рис. 2.3, *б*. **Конформация разрешена** в предположении модели жестких сфер с контактными радиусами, отвечающими нижним границам (табл. 2.1).

разования множества глобулярных белков. С другой стороны, цепи со всеми *транс*-пептидными группами являются в меру жесткими, так что спонтанное свертывание не требует значительных усилий стягивания различных участков белка (разд. 3.5). Отметим в этой связи, что энергетическая предпочтительность в 2 ккал/моль *транс*-изомера по сравнению с *цис*-изомером не представляется существенной, поскольку цепь с пептидными группами в *цис*-конфигурации должна быть более жесткой и, следовательно, не требовать при формировании уникальной конформации белка такой же большой энергии связывания, как цепь с пептидными группами в *транс*-конфигурации.

Заключение

Пептидные связи синтезируются строго контролируемым путем на рибосомах. Они могут менять свою конфигурацию при довольно низких затратах свободной энергии. Это помогает организму быстро приспосабливаться к флуктуациям окружающей среды. Жесткость пептидной связи вносит существенный вклад в стабильность белка. Стабильность полипептидной цепи увеличивается в результате взаимодействий атомов основной цепи, а также атомов основной и боковых цепей. Геометрические условия удовлетворительно описываются с помощью карт потенциальной энергии, которые могут также учитывать гибкость основной цепи. *транс*-Пептидная связь имеет явные преимущества перед *цис*-связью, поскольку не вносит столь серьезных структурных ограничений, как *цис*-связь.

Глава 3 Взаимодействия, определяющие структуру белка

Невалентные взаимодействия были открыты Ван-дер-Ваальсом (1873 г.), пытавшимся объяснить отклонение поведения реального газа от закона идеального газа. Взаимодействия между валентно-несвязанными атомами имеют огромное значение для биологических организмов; в частности, они ответственны за спонтанное свертывание полипептидных цепей и цепей нуклеиновых кислот, а также спонтанное образование мембран. Благодаря им происходит узнавание комплементарных молекулярных поверхностей («система ключ — замок») [44]).

Типы различных невалентных взаимодействий перечислены в табл. 3.1; их энергия на один-три порядка меньше по абсолютной

Таблица 3.1

Типы взаимодействий, важных для белковых структур [a]

Тип взаимодействия	Пример	Энергия связи, ккал/моль	Изменение свободной энергии (вода→этанол), ккал/моль
Дисперсионные взаимодействия	Алифатический водород	—0,03	
	$-\overset{\vert}{\underset{\vert}{C}}-H\ldots H-\overset{\vert}{\underset{\vert}{C}}-$	—5	—1
Электростатическое взаимодействие	Солевой мостик $-\overset{-}{COO}\ldots H_3\overset{+}{N}-$		
	2 диполя $>\overset{\delta+}{C}=\overset{\delta-}{O}\ldots\overset{\delta-}{O}=\overset{\delta+}{C}<$	+0,3	
Водородная связь	Лед $\diagdown O-H\ldots O\diagup$	—4	
	Основная цепь белка $>N-H\ldots O=$	—3	
Гидрофобные взаимодействия	Боковая цепь Phe		—2,4

[a] При расчетах энергии электростатического взаимодействия диэлектрическая проницаемость принята равной 4.

величине энергии ковалентного связывания. Величину невалентных взаимодействий трудно измерить и столь же трудно рассчитать. Кроме того, в стабилизацию белка важный вклад энтропийного характера вносит окружающая среда. Поэтому необходимо рассматривать систему в целом, учитывая белок и растворитель.

3.1 Дисперсионные силы и отталкивание электронных оболочек

Дисперсионные силы притяжения действуют между любыми парами атомов. Дисперсионные силы существуют между любой парой атомов, даже если они совершенно неполярны. Если представить атом осциллирующим диполем, то в паре атомов каждый диполь будет поляризовать соседний атом. В результате между атомами возникнет сила притяжения, энергия которой в первом приближении обратно пропорциональна шестой степени расстояния между ядрами атомов [45] и пропорциональна поляризуемостям атомов [46].

Поляризуемость входящего в молекулу атома анизотропна, в связи с чем дисперсионные силы зависят от ориентации фрагментов молекулы друг относительно друга. Однако поскольку влияние ориентации мало и его трудно измерить, то им обычно пренебрегают и считают дисперсионные силы изотропными.

Электронные оболочки валентно-несвязанных атомов взаимно отталкиваются. Силы дисперсионного притяжения между парой валентно-несвязанных атомов уравновешиваются отталкиванием электронных оболочек. Это отталкивание было аппроксимировано Леннард-Джонсом введением слагаемого, обратно пропорционального m-й степени расстояния [47]. Слейтер [48] описал отталкивание экспоненциальной функцией, которая была модифицирована Букингемом [49] и носит его имя. В расчетах белков наиболее часто используют аппроксимацию Леннард-Джонса с $m = 12$. С учетом лондоновского члена для дисперсионных сил [45] это приводит к простому в вычислительном отношении потенциалу 6—12, пример которого приведен на рис. 3.1. Оба параметра, необходимых для определения этой функции, описаны в подписи к рисунку. Потенциал на рис. 3.1 имеет минимум в отрицательной области при расстоянии R_m. Поэтому атомы слабо «связаны» на этом расстоянии (E_m — энергия притяжения).

Энергетические параметры могут быть определены из данных по кристаллическим структурам. Абсолютную величину параметров потенциала 6—12 Леннард-Джонса можно определить по контактным расстояниям и контактным энергиям в кристаллах малых молекул, которые следуют из данных рентгеноструктурного анализа и, например, теплот сублимации. Эти же данные можно получить из экспериментов по рассеянию атомных и молекулярных пучков.

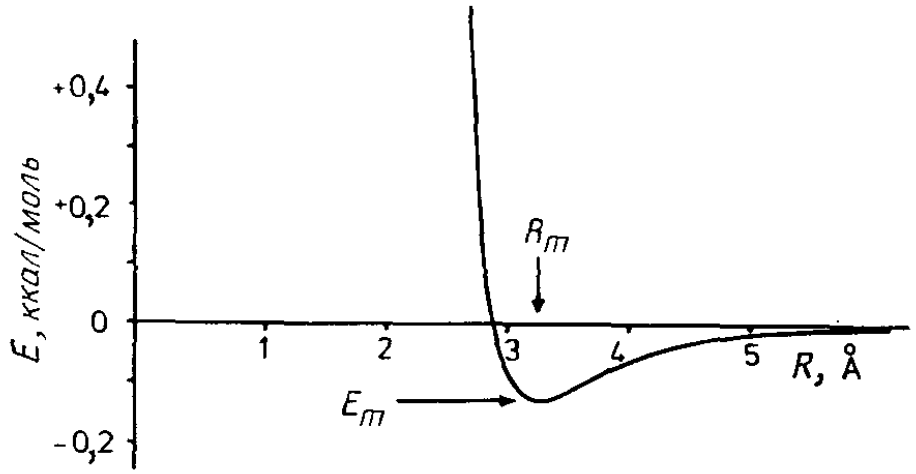


Рис. 3.1.

Потенциал 6—12 Леннард-Джонса, описывающий дисперсионные силы и электронное отталкивание ($R_m = 3{,}24$ Å, $E_m = -0{,}13$ ккал/моль).

Формула с двумя параметрами дана ниже в удобной для вычисления форме (A, B — параметры), а также в ее нормированном виде (E_m, R_m — параметры); величины A и B положительны. С уменьшением B и увеличением A E_m уменьшается, а соответствующее межатомное расстояние R_m увеличивается. Потенциальный барьер асимметричен. Отталкивание и притяжение взаимно уравновешиваются при расстоянии $0{,}89 \cdot R_m$; на расстоянии $1{,}5 \cdot R_m$ энергия притяжения составляет всего одну шестую E_m.

$$E = \frac{A}{R^{12}} - \frac{B}{R^6} = E_m \left[-\left(\frac{R_m}{R}\right)^{12} + 2\left(\frac{R_m}{R}\right)^6 \right] \qquad E_m = -\frac{B^2}{4A} \qquad R_m = \sqrt[6]{\frac{2A}{B}}$$

Примеры таких параметров даны в табл. 3.2. Различия между данными характеризуют величину возможных ошибок.

Как видно из табл. 3.2, доля энергии, приходящаяся на атомный контакт, очень мала. Однако количество контактов в белке велико. Например, при плотнейшей гексагональной упаковке одина-

Таблица 3.2

Параметры потенциала 6—12 Леннард-Джонса для отталкивания электронных оболочек и дисперсионных сил при невалентном контакте [a]

Взаимодействие	Данные Момани и сотр. [69]		Данные Лифсона и Уоршела [70, 71]	
	E_m, ккал/моль	R_m, Å	E_m, ккал/моль	R_m, Å
Алифатический H ... алифатический H	—0,04	2,92	—0,01	2,94
Алифатический C ... алифатический C	—0,04	4,12	—0,19	4,23
Карбонильный O ... карбонильный O	—0,20	3,12	—0,23	3,00
Амидный N ... амидный N	—0,11	3,51	—0,19	3,60

[a] Приведены контакты только между идентичными типами атомов. Для контакта между неидентичными типами i и j получаем $E_m^{(ij)} = (E_m^{(ii)} + E_m^{(jj)})^{12}$ и $R_m^{(ij)} = 1/2(R_m^{(ii)} + R_m^{(jj)})$.

ковых сфер каждая сфера образует контакты с 12 соседями. Поскольку контактные энергии складываются аддитивно, то общая энергия притяжения, приходящаяся на сферу, в этом случае в шесть раз больше энергии, приходящейся на один контакт.

3.2 Электростатические взаимодействия

Большинство ковалентно связанных атомов несет парциальные заряды. Поскольку ковалентные связи между разными типами атомов приводят к асимметричному распределению валентных электронов, большинство атомов молекулы несет парциальные заряды. Парциальные заряды некоторых аминокислотных остатков приведены в табл. 3.3. Поскольку суммарный заряд нейтральной молекулы равен нулю, то она может быть аппроксимирована набором диполей или мультиполей. Эти мультиполи взаимодействуют между собой по закону Кулона, как показано в табл. 3.4. Энергия взаимодействия зависит от диэлектрической проницаемости ε окружающей среды. Мы используем здесь величину $\varepsilon = 4$, макроскопическую диэлектрическую проницаемость амидного полимера [50]. Диэлектрическую проницаемость для микроскопических объектов рассчитать трудно, и используемые в расчетах значения ε могут варьироваться от 1 до 5 [51].

Примеры электростатических взаимодействий даны в табл 3.1. Внутренний ($\varepsilon = 4$) солевой мостик между Lys и Asp имеет связывающую энергию около 5 ккал/моль. Две карбонильные группы

Таблица 3.3

Парциальные заряды на атомах основной цепи полипептида и трех боковых цепей [a]

Пептид	N	−0,36	Tyr	O_η	−0,33
Пептид	H_N	+0,18	Tyr	H_η	+0,17
Пептид	C_α	+0,06	Asn	C_β	−0,12
Пептид	C′	+0,45	Asn	H_β	+0,06
Пептид	O	−0,38	Asn	C_γ	+0,46
Пептид	H_α	+0,02	Asn	O_δ	−0,38
Ser	C_β	+0,13	Asn	N_δ	−0,45
Ser	H_β	+0,02	Asn	H_δ	+0,20
Ser	O_γ	−0,31	Cys	S_γ	+0,01
Ser	H_γ	+0,17	Cys	H_γ	+0,01

[a] По данным Момани и сотр. [52], определивших парциальные заряды атомов остатков 20 природных аминокислот с помощью расчета ППДП/2 (*полное пренебрежение дифференциальным перекрыванием*), парциальные заряды были также получены из расчетов ab initio молекулярных орбиталей малых молекул [53], а также обработкой результатов рентгеноструктурного анализа [54].

Таблица 3.4

Электростатические взаимодействия

Закон Кулона:

$$E = \frac{332}{\varepsilon} \frac{q_1 \cdot q_2}{R_{12}}$$

q_i — парциальный заряд в единицах электрона (e_0).

R_{12} — расстояние между парциальными зарядами, Å
ε — диэлектрическая проницаемость
E — энергия отталкивания или притяжения, ккал/моль

Единицы энергии:
1 электронвольт = 1 эВ=23 ккал/моль = 96 500 Вт·с/моль = $9{,}65 \cdot 10^{11}$ эрг/моль
Единица дипольного момента:

$$1 \text{ дебай} = 1\text{Д} = 0{,}208 e_0 \text{ Å}$$

Энергия взаимодействия диполей с моментами μ_1 и μ_2 при условии точечности диполей и, следовательно, большого расстояния R_{12} по сравнению с длиной диполя.

$$E = \frac{332}{\varepsilon} \left[\frac{(\vec{\mu}_1 \cdot \vec{\mu}_2)}{R_{12}^3} - \frac{3(\vec{\mu}_1 \cdot \vec{R}_{12}) \cdot (\vec{\mu}_2 \cdot \vec{R}_{12})}{R_{12}^5} \right]$$

Энергия взаимодействия между беспорядочно ориентированными диполями равна нулю.

Энергия взаимодействия между двумя диполями с моментами 1 Д на расстоянии 5 Å в среде с $\varepsilon = 4$.

$\rightarrow \rightarrow$ $E = -1{,}32$ ккал/моль
$\rightarrow \leftarrow$ $E = +1{,}32$ ккал/моль
$\uparrow \uparrow$ $E = +0{,}66$ ккал/моль
$\downarrow \uparrow$ $E = -0{,}66$ ккал/моль

на контактном расстоянии R_m характеризуются энергией электростатического отталкивания 0,3 ккал/моль.

Электростатические взаимодействия в белках имеют локальный характер. Хотя расчет электростатических взаимодействий есть не что иное как непосредственное применение закона Кулона ко всем парциальным зарядам, он занимает очень много времени из-за большого числа атомов в белке и большого радиуса действия электростатических взаимодействий. Однако расчет можно упростить. Поскольку монополей, т. е. свободных зарядов, в белках нет (солевые мостики — это пары ионов или диполи), все частичные заряды образуют диполи или мультиполи. Взаимодействие между ними уменьшается пропорционально по крайней мере третьей степени расстояния (табл. 3.4). Поэтому реальный радиус влияния электростатических взаимодействий довольно мал и энергетические расчеты можно ограничить взаимодействиями между ближайшими соседями.

Можно возразить, что слабые взаимодействия на больших расстояниях могут складываться, если диполи одинаково ориентированы, как, например, диполи водородных связей в α-спиралях

и β-структурах. Однако в β-структурах соседние диполи антипараллельны и их поля на больших расстояниях компенсируют друг друга; в α-спиралях диполи располагаются линейно и также взаимно компенсируются, за исключением зарядов на обоих концах. Именно по этой причине антипараллельные спирали более выгодны, чем параллельные. Однако вклад электростатического взаимодействия даже в этом случае мал по сравнению с общей энергией взаимодействия между такими α-спиралями.

3.3 Потенциалы Ван-дер-Ваальса

Потенциалы включают отталкивание электронных оболочек, дисперсионные силы вандерваальсова притяжения и электростатические взаимодействия. Для удобства вычислений целесообразно объединить все три невалентные силы в одну простую потенциальную функцию (или силовое поле), которая традиционно называется потенциалом Ван-дер-Ваальса. Для этого необходимо еще более упростить представление об электростатическом взаимодействии. Прежде всего допускают, что контакты образуются только между ближайшими соседями и усредняют электростатические взаимодействия по всем относительным взаимным ориентациям, стерически допустимым для двух контактирующих групп: $>C=O$ и $H-C\langle$ и т. д. Таким образом, результирующий вклад зависит только от расстояния между контактирующими группами. При таком упрощении вандерваальсов потенциал изотропен:

$$E = \frac{A}{R^{12}} - \frac{B}{R^{6}} + \frac{q_i q_j}{R}$$

Он включает три параметра: A и B, как показано на рис. 3.1, и произведение эффективных зарядов $q_i q_j$ контактирующих атомов.

Эффективные заряды можно определить, исходя из парциальных зарядов отдельных атомов (табл. 3.3). Параметры A и B, описывающие отталкивание электронных облаков и дисперсионные силы, могут быть найдены из данных по кристаллической структуре [52]. Поскольку парциальные заряды удается рассчитать только с низкой точностью [53], Лифсон и сотр. [54] предприняли попытку получить их из данных по кристаллической структуре. Это было сделано путем одновременной вариации всех трех параметров потенциала 1—6—12 до получения наилучшего согласия с экспериментальными данными.

Потенциалы 6—12. При уточнении белковых структур с помощью процедуры минимизации энергии Левит [55] для удобства расчетов отказался от члена R^{-1}. Это привело к вандерваальсову потенциалу 6—12, который заметно отличается от потенциалов 6—12, описывающих только отталкивание электронных оболочек и дисперсионные силы притяжения (табл. 3.2). Примеры таких

потенциалов даны в табл. 3.5. Отметим, что, когда контакты образованы амидными атомами азота, действуют исключительно силы отталкивания. В этом случае электростатическое отталкивание между отрицательными зарядами на атомах азота преобладает над притяжением за счет дисперсионных сил, приведенных в табл. 3.2.

Таблица 3.5

Параметры вандерваальсова потенциала 6 — 12, полученные из данных по известным белковым структурам [55][a]

Взаимодействие	A, ккал/моль·Å^{12}	B, ккал/моль·Å^{6}	E, ккал/моль	R, Å
Алифатический C···алифатический C	2 750 000	+1425	—0,19	3,53
Карбонильный O···карбоинильный O	417 000	+108	—0,01	3,96
Амидный N···амидный N	417 000	0	Отталкивание	
Карбоинильный O···углерод в беизольиом цикле	695 000	—570	Отталкивание	

[a] A и B определены на рис. 3.1.

Вандерваальсовы радиусы приблизительно соответствуют контактным расстояниям. Вандерваальсовы потенциалы позволяют определить «расстояния, отвечающие вандерваальсовым контактам» между данными атомами. Нижние границы этих расстояний, полученные из кристаллических структур, определили Рамачандран и Сасисекхаран [29]. Они составляют около 75% равновесных расстояний R_m в табл. 3.2 и отвечают энергии отталкивания, равной приблизительно 1 ккал/моль. Эти контактные расстояния были использованы в модели жестких сфер для оценки стерических затруднений при C_α-атоме полипептидной цепи [28] (рис. 2.3).

Наблюдаемые вандерваальсовы контактные расстояния для пар

Таблица 3.6

Вандерваальсовы радиусы по Бонди [73]

Тип атома	Радиус, Å
Ароматический H	1,0
Алифатический H	1,2
O	1,5
N	1,6
C	1,7
S	1,8

атомов можно перевести в более общие «вандерваальсовы радиусы», если предположить, что каждое расстояние есть сумма радиусов двух определенных атомов. Это допущение верно лишь отчасти; результирующие радиусы представляют усреднение по многим типам контактов между различными атомами. Примеры вандерваальсовых радиусов даны в табл. 3.6. «Нижние нормальные границы» контактных расстояний, приведенные табл. 2.1, примерно на 10% меньше сумм соответствующих вандерваальсовых радиусов.

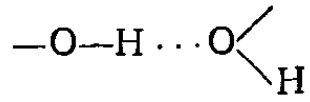
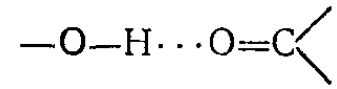
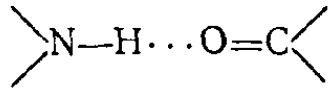
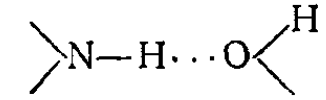
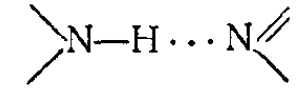
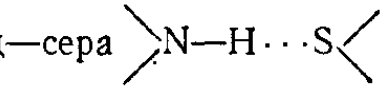
3.4 Водородные связи

Водородные связи — преимущественно электростатические взаимодействия. Как было показано, межатомные расстояния, отвечающие всем невалентным контактам атомов, приблизительно описываются соответствующими вандерваальсовыми радиусами. Это правило часто нарушается в случае контактов, образованных атомами водорода. Например, расстояние между амидным атомом H и карбонильным атомом O составляет всего 1,9 Å, вместо 2,7 Å, получающихся из расчета по вандерваальсовым радиусам из табл. 3.6. Опыт показывает, что этот эффект наблюдается всегда, когда атом H несет большой положительный парциальный заряд, а его партнер по контакту — большой отрицательный парциальный заряд (табл. 3.3). Эти заряды притягивают друг друга. Поскольку все электронное облако атома водорода (у водорода только один электрон!) сильно смещено к атому, с которым водород ковалентно связан, отталкивание оболочек контактирующих партнеров мало и притяжение зарядов может их еще более сблизить. Такое короткое расстояние обусловливает большую величину кулоновской энергии притяжения (табл. 3.4), а также высокую дисперсионную энергию [56]. Энергия взаимодействия имеет в этом случае значение, промежуточное по *абсолютной* величине между энергией вандерваальсовых контактов и энергией ковалентных связей. Такие контакты были выделены в особую группу «водородных связей», а атомы, участвующие в связи, получили название «донор — акцептор водородной связи».

Биологически важные водородные связи не приводят к большому перекрыванию волновых функций. Полемичен вопрос относительно существенности вклада в водородную связь перекрывания волновых функций [56]. В предельном случае иона $[F \cdots H—F]^-$ такое перекрывание несомненно имеется. В этом случае расстояние F—F составляет всего 65% суммы вандерваальсовых радиусов атомов H и F и длины связи H—F, а связывающая энергия (~50 ккал/моль) приближается к энергии ковалентных связей [57]. Однако в белках расстояния и связывающая энергия уменьшаются не так. Здесь вклад перекрывания волновых функций, по-видимому, значительно несуществен.

Таблица 3.7

Водородные связи в белках [58]

Тип водородной связи	Расстояние между донором и акцептором, Å	Уменьшение расстояния между донором и акцептором по сравнению с суммой вандерваальсовых радиусов, %	Примечание
Гидроксил — гидроксил 	2,8±0,1	25	Возможны разветвленные связи, например во льду. Для феиолов (Tyr) длина связи 2,7±0,1 Å
Гидроксил — карбонил 	2,8±0,1	25	
Амид—карбонил 	2,9±0,1	20	В белках между атомами основной цепи; эмпирическая длина в парвальбумине 2,9±0,3 Å [59]
Амид—гидроксил 	2,9±0,1	20	
Амид—азот имидазола 	3,1±0,2	15	
Амид—сера 	3,7	10	Для Met-180 химотрипсина [18] и для FeS-кластеров ферредоксина, рубредоксина HiPiP [72]; длиина 3,6±0,3 Å

Длина водородной связи пропорциональна ее энергии. Как видно из табл. 3.7, самые малые длины и самые высокие энергии найдены в водородных связях с участием атома кислорода, в частности между фенольными группами (Tyr—Tyr) и в воде. Несколько большие расстояния и меньшие энергии у водородных связей между атомами азота и кислорода, и значительно меньшие энергии имеют водородные связи между двумя атомами азота, а также между атомами азота и серы. По сравнению с суммой вандерваальсовых радиусов расстояния увеличиваются на 10—25%. Для грубой оценки энергии водородных связей проще всего воспользоваться теплотой сублимации льда, которая составляет 13 ккал/моль. Большая часть

этой теплоты приходится на водородные связи, поскольку теплоты сублимации H_2S и H_2Se (которые образуют значительно более слабые водородные связи) составляют всего 6 ккал/моль. Как показано на рис. 3.2, на молекулу H_2O приходятся две водородные связи, так что энергия одной водородной связи должна быть несколько больше половины разности теплот сублимации, а именно около 4 ккал/моль. Энергия водородных связей между амидной и карбонильной группами, которые часто образуются в основной цепи полипептидов, составляет около 3 ккал/моль (табл. 3.1).

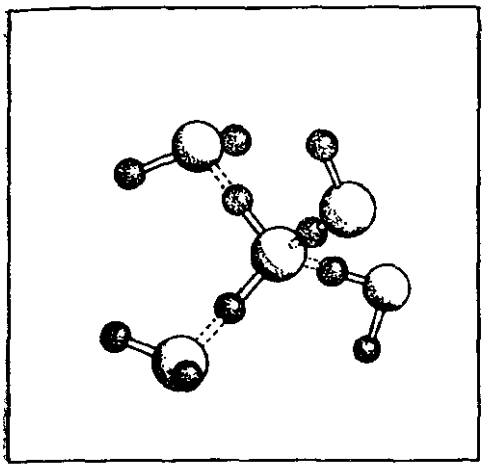
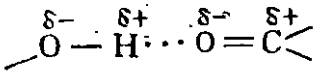


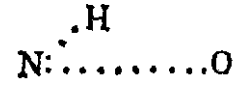
Рис. 3.2.
Структура льда по данным рентгеноструктурного анализа и нейтронографии.

Водородные связи линейны. Как упоминалось в предыдущем разделе, водородная связь включает большой положительный (Н-атом) и большой отрицательный (Н-акцептор) парциальные заряды, которые разделены коротким расстоянием. В общем случае эти заряды можно рассматривать как составляющие диполя (табл. 3.3). Ковалентно связанные соседние атомы (Н-донор и сосед Н-акцептора) несут противоположные заряды, например



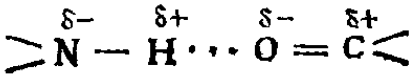
Таким образом, положительно заряженный Н-атом находится между двумя отрицательно заряженными атомами.

Положительный заряд обеспечивает атому Н, расположенному между двумя отрицательными зарядами, наиболее низкую потенциальную энергию, если все три заряда расположены на одной прямой. Как следствие водородные связи линейны, например в связи N—H · · · O атомы располагаются на одной прямой. Электростатические расчеты показывают, что



отклонение угла от нуля на 20° понижает энергию связи приблизительно на 10% [58]. В парвальбумине [59] среднее отклонение угла составляет именно 20°. Очевидно, в этой структуре (так же как и в других белковых структурах) устанавливается определенный компромисс между максимальной энергией водородной связи и другими геометрическими ограничениями.

Диполи, составляющие водородную связь, стремятся к линейному расположению. Линейность водородной связи не следует отождествлять с ориентацией составляющих ее диполей, как, например, с линейным расположением пептидного амида и карбонильной группы:



Как показано в табл. 3.4, диполи имеют минимальную энергию (самая прочная связь), если они ориентированы вдоль прямой*. Однако нарушение этой ориентации приводит к значительно меньшему снижению энергии, чем такое же по величине отклонение от линейности водородной связи [58]. Поэтому нарушение линейной ориентации диполей довольно обычно в белках, где водородные связи подчинены разнообразным геометрическим ограничениям. Для льда также характерно нелинейное расположение диполей (рис. 3.2), а в некоторых случаях, как, например, для связи между амидной и гидроксидной группами, линейная ориентация стерически невозможна (табл. 3.7).

3.5 Энтропия

Концепция свободной энергии

Связывающая энергия конкурирует в вакууме с энтропией цепи. До сих пор мы рассматривали только невалентную связывающую энергию, т. е. вандерваальсовы взаимодействия и водородные связи. Однако образование белков и их стабильность зависят от *свободной* энергии, а не только от связывающей энергии**. Кроме того, поскольку нативные белки находятся в водном окружении, то при строгом подходе необходимо учитывать свободную энергию всей системы, включающей белковую *цепь и растворитель*.

Чтобы выявить взаимосвязь между свободной энергией и связывающей энергией, рассмотрим вначале полипептидную цепь в вакууме. В этом случае разность свободных энергий свернутого

---

* На малых расстояниях, найденных в водородных связях, диполь — дипольные взаимодействия нельзя сводить к взаимодействию двух *точечных* диполей, как это сделано в табл. 3.4. Однако формула в табл. 3.4 применима к линейно расположенным диполям.

** Как принято в биологической литературе, под термином «свободная энергия» будет подразумеваться свободная энергия Гиббса. Термин «связывающая энергия» используется для обозначения «(связывающей) энтальпии». В большинстве биологических реакций изменениями объема и объемных энергий пренебрегают, поэтому свободная энергия Гиббса эквивалентна свободной энергии Гельмгольца, а (связывающая) энергия эквивалентна (связывающей) энтальпии.

нативного и вытянутого денатурированного состояний равна $\Delta G_{\text{цепь}} = G_{\text{нат}} - G_{\text{денат}}$, тогда как разность связывающих энергий равна $\Delta H_{\text{цепь}} = H_{\text{нат}} - H_{\text{денат}}$. Константа равновесия K = (число цепей в нативном состоянии)/(число цепей в денатурированном состоянии) задается уравнением

$$-RT\ln K = \Delta G_{\text{цепь}} = \Delta H_{\text{цепь}} - T\Delta S_{\text{цепь}} \qquad (3.1)$$

Правая часть уравнения (3.1) пояснена рис. 3.3, *а*. Изменение энтропии (беспорядка) $\Delta S_{\text{цепь}} = S_{\text{нат}} - S_{\text{денат}}$ играет решаю-

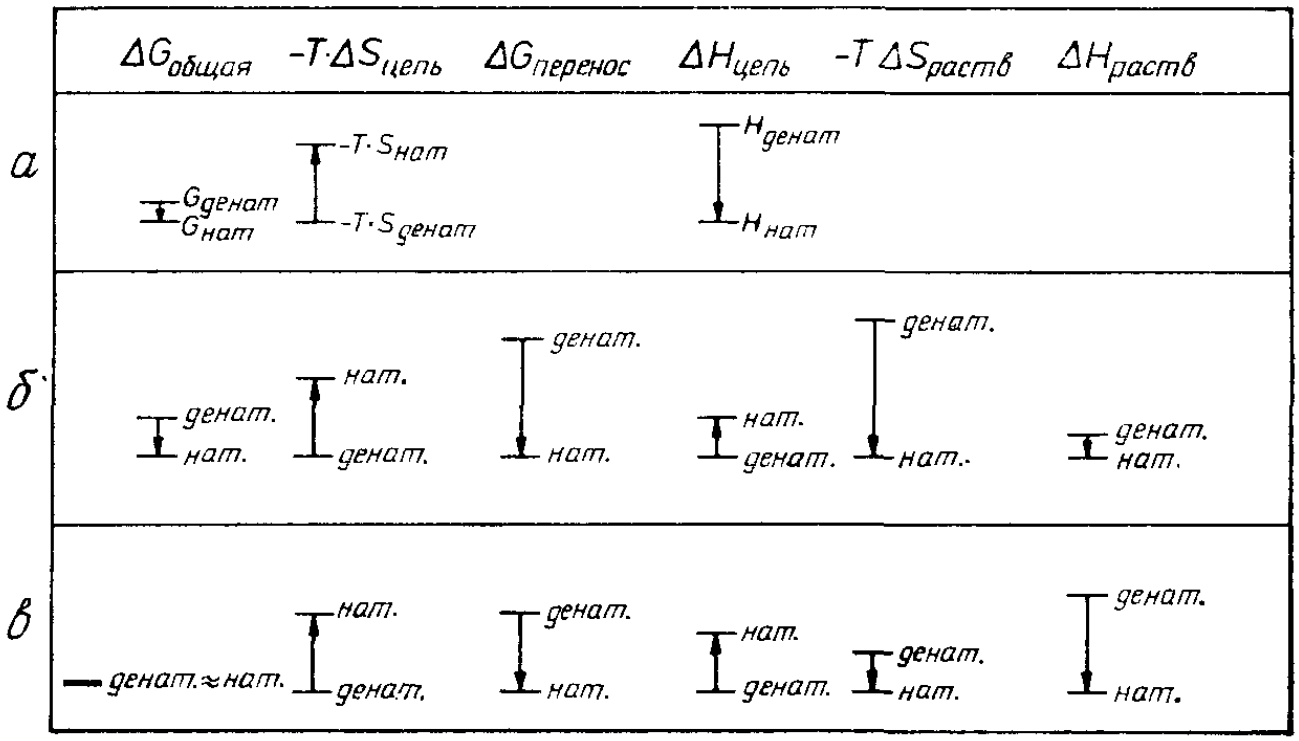


Рис. 3.3.

Энергетические вклады в стабильность белка на одну атомную группу.

Все значения даны приближенно. Стрелки указывают знаки вкладов: направление вниз означает предпочтительность нативного состояния. При разделении вкладов растворителя и цепи возникает также член, учитывающий энергию взаимодействия растворитель—цепь. Этот член включен в $\Delta H_{\text{цепь}}$. Величины $\Delta G_{\text{общая}}$ и $\Delta G_{\text{перенос}}$ определены соответственно уравнениями (3.2) и (3.3). *а* — полипептидная цепь в вакууме [уравнение (3.1)]. $\Delta G_{\text{общая}}$ благоприятствует нативному состоянию. *б* — неполярные группы цепи в водном окружении [уравнение (3.2)]. Для системы «масло—вода» индекс «цепь» следует заменить на индекс «масло». *в* — полярные группы цепи в водном окружении [уравнение (3.2)].

щую роль. Состояние денатурированных полипептидных цепей менее упорядочено (более высокая энтропия), чем состояние нативных цепей, поэтому $\Delta S_{\text{цепь}}$ отрицательна и энтропийный член $-T\Delta S_{\text{цепь}}$ положителен, что способствует переходу в денатурированное состояние. Величина $\Delta H_{\text{цепь}}$ отрицательна поскольку водородные связи и вандерваальсово притяжение увеличивают связывающую энергию нативного состояния, так что $H_{\text{нат}}$ ниже (более отрицательна), чем $H_{\text{денат}}$. Следовательно, при данной температуре T знак $\Delta G_{\text{цепь}}$ (а значит, величина K) зависит от соотношения абсолютных величин $\Delta H_{\text{цепь}}$ и $\Delta S_{\text{цепь}}$: «только при достаточной связывающей энергии цепь стремится выйти из денатуриро-

ванного состояния и принять нативную конформацию». Кроме того, «чем выше температура, тем выше должна быть связывающая энергия». Более строгий подход к ситуации в вакууме, основанный на статистической механике, дан в приложении и в разд. 8.1.

Растворитель очень важен. В применении к вакууму взаимосвязь между связывающей энергией и энтропией очевидна, однако такая ситуация мало связана с реальными биологическими условиями. Белкам необходим растворитель, более того, для стабильности белков предельно важны физические свойства растворителя. Так, почти все белки денатурируются в этаноле или в водных растворах, содержащих достаточные количества додецилсульфата натрия (ДСН) или мочевины. Рассмотрим поэтому систему из полипептидной цепи и растворителя; общая свободная энергия приобретает вид:

$$-RT\ln K = \Delta G_{\text{общая}} = \Delta H_{\text{цепь}} + \Delta H_{\text{раств}} - T\Delta S_{\text{цепь}} - T\Delta S_{\text{раств}} \quad (3.2)$$

Здесь энергия взаимодействия цепи с растворителем включена в член $\Delta H_{\text{цепь}}$.

Смесь масла и воды

Жидкая вода может образовывать локальные упорядоченные квазикристаллические структуры. Обсудим уравнение (3.2) в применении к водному раствору минерального масла, состоящего из углеводородов, вместо водного раствора полипептидов. «R-состоянием» назовем монодисперсный раствор масла в воде, а «N-состоянием» — раздельные фазы, т. е. каплю масла на поверхности воды. Величина $\Delta S_{\text{масло}}$ отрицательна, поскольку монодисперсный раствор менее упорядочен, чем раздельные фазы (рис. 3.3, *б*). Величина $\Delta H_{\text{масло}}$ положительна. Чтобы понять этот факт, следует учесть, что в N-состоянии большинство молекул масла окружено такими же молекулами, тогда как в R-состоянии все они окружены водой. В вандерваальсовом взаимодействии между молекулами масла преобладают дисперсионные силы, и поэтому оно слабо. Между молекулами масла и воды это взаимодействие сильнее, поскольку полярные молекулы воды индуцируют в углеводородах диполи, за счет чего заметно растет электростатический член [15]. В связи с этим величина $\Delta H_{\text{масло}}$ (которая представляет разность обоих состояний) благоприятствует образованию монодисперсного раствора. Однако эта величина сравнительно мала.

Экспериментально было показано [15], что величина $\Delta S_{\text{раств}}$ положительна, что благоприятствует N-состоянию. Это свидетельствует о том, что упорядоченность водной компоненты выше в монодисперсном растворе масла, где больше площадь контактной поверхности неполярных молекул и молекул воды. У находящейся в таких условиях воды имеются две возможности: либо сохранить свою жидкую структуру и лишиться водородных связей, либо сох-

ранить большую часть водородных связей и образовать вокруг неполярной молекулы локально упорядоченную, так называемую клатратную структуру, или «айсберг» [15]. Первая возможность отвечает увеличению $H_{\text{раств}}$ (уменьшение связывающей энергии), тогда как вторая означает уменьшение $S_{\text{раств}}$ (увеличение упорядоченности). Поскольку наблюдаемая величина $\Delta S_{\text{раств}}$ положительна, вода действительно образует на поверхности раздела локально упорядоченную квазикристаллическую структуру. Но даже при возникновении такой клатратной структуры вода лишается части водородных связей, что приводит к отрицательной величине $\Delta H_{\text{раств}}$, благоприятствующей N-состоянию. Заметим, однако, что абсолютная величина $\Delta H_{\text{раств}}$ сравнительно мала.

Энтропия воды способствует образованию капелек масла. Как мы установили, $\Delta H_{\text{масло}}$ и $\Delta S_{\text{масло}}$ благоприятствуют R-состоянию, тогда как $\Delta H_{\text{раств}}$ и $\Delta S_{\text{раств}}$ благоприятствуют N-состоянию. Поскольку разделение фаз масла и воды наблюдается в действительности, вклад последних величин превышает вклад первых. Поскольку величина $\Delta H_{\text{раств}}$ мала, основной «движущей силой», которая переводит молекулы масла в отдельную фазу или в каплю масла, является $\Delta S_{\text{раств}}$. Влияние $\Delta S_{\text{раств}}$ часто называют «энтропийными силами» или «гидрофобными силами». Эти энтропийные силы приводят также к спонтанному образованию мицелл, пузырьков и мембран из дифильных липидов.

Полипептидная цепь в воде

Природная полипептидная цепь включает специфическую последовательность полярных и неполярных групп. Рассмотрим теперь, пользуясь моделью вода — масло, полипептидную цепь в воде (рис. 3.3). Поверхность контакта полипептидной цепи и воды велика для беспорядочно вытянутой цепи и мала в ее нативной, свернутой конформации. Однако в отличие от совершенно неполярных молекул масла полипептидная цепь включает как полярные, так и неполярные группы. Например, амидные и карбонильные группы основной цепи, а также гидроксильные (в Ser, Thr, Tyr), карбоксильные (в Asp, Glu), аммониевые (в Lys) группы полярны и стремятся оставаться в воде (гидрофильность). Углеводородные боковые цепи Ala, Val, Leu, Ile и Phe неполярны.

На первый взгляд белки напоминают капли масла. Как было показано на примере системы масло — вода, все неполярные группы стремятся соединяться, чтобы уменьшить площадь поверхности между неполярной цепью и водой. Действительно, в нативных белках большая часть неполярных боковых групп удалена от воды и сосредоточена в гидрофобных ядрах. Это дало основание Козману [15] сравнивать белки с каплями масла.

Внутренняя часть белка также содержит много полярных групп. Поскольку глобулярные белки имеют диаметр около 30 Å, боковые цепи не могут располагаться внутри белка без того, чтобы там не находились также полярные амидные и карбонильные группы основной цепи. Однако полярные группы прекрасно сосуществуют с водой, и их локализация во внутренней части без потери свободной энергии возможна, только если они образуют водородные связи. В действительности такие связи наблюдаются для более чем 90% полярных групп, расположенных внутри белка.

Заряженные группы находятся на поверхности. Удаление заряженных групп (Asp, Glu, Lys, Arg) от воды энергетически очень невыгодно. Это может произойти только при сопутствующем образовании солевых мостиков, которые, однако, встречаются крайне редко. Почти все заряженные боковые группы находятся на поверхности белка, увеличивая его растворимость. Кроме того, находясь в растворе в процессе спонтанного свертывания и после него, они способствуют вполне определенному направлению свертывания полипептидной цепи.

Энтропия растворителя вносит большой вклад в стабильность белка. Обсудим теперь четыре члена уравнения (3.2). Как было отмечено выше, величина $\Delta S_{\text{цепь}}$ отрицательна для всех частей полипептидной цепи. Для неполярных частей абсолютные величины $\Delta H_{\text{цепь}}$ и $\Delta H_{\text{раств}}$ малы, причем $\Delta H_{\text{цепь}}$ положительна, а $\Delta H_{\text{раств}}$ отрицательна (рис. 3.3, *б*). Для неполярных частей $\Delta S_{\text{раств}}$ имеет большое положительное значение.

Полярные группы образуют водородные связи при контакте с водой и между собой внутри белка. В нашей системе подсчета (все взаимодействия между цепью и растворителем включены в величину $H_{\text{цепь}}$) величина $\Delta H_{\text{цепь}}$ положительна: одна водородная связь с водой рвется и только половина водородной связи образуется внутри белка. Величина $\Delta H_{\text{раств}}$ отрицательна: на полярную группу приходится половина водородной связи. На рис. 3.3, *в* абсолютная величина $\Delta H_{\text{цепь}}$ показана меньшей, чем абсолютная величина $\Delta H_{\text{раств}}$, поскольку в плотно упакованной внутренней части белка имеется больше вандерваальсовых контактов (разд. 3.6), чем на поверхности цепь — растворитель. Это благоприятствует нативному N-состоянию. Поскольку молекулы воды отчасти упорядочены вокруг полярных групп, $\Delta S_{\text{раств}}$ положительна, хотя и мала. В целом $\Delta G_{\text{общая}}$ для полярных групп близка к нулю. Сопоставляя все члены уравнения (3.2) (рис. 3.3, *б*, *в*) можно заключить, что $\Delta S_{\text{раств}}$ неполярной группы вносит самый большой вклад в стабильность белка. Большой вклад $\Delta H_{\text{раств}}$ для полярных групп на рис. 3.3, *в* в значительной степени компенсируется величиной $\Delta H_{\text{цепь}}$. Абсолютная величина $\Delta H_{\text{раств}}$ — всего лишь артефакт нашей системы подсчета, которая включает в величину $H_{\text{цепь}}$ взаимодействия растворитель — цепь.

Свободные энергии переноса боковых цепей дают возможность оценить вклады этих цепей в стабильность белка. Количественно оценить обсуждавшиеся выше вклады довольно сложно. В работе [15] были определены величины ΔH и ΔS переноса различных малых неполярных молекул из монодисперсного водного раствора в монодисперсный раствор в неполярном углеводородном растворителе. Если проводить аналогию между этими экспериментами и переходом неполярных боковых цепей из воды во внутреннюю часть белка, то

$$\begin{aligned} \Delta H_{\text{перенос}} &= \Delta H_{\text{цепь}} + \Delta H_{\text{раств}} \\ \Delta S_{\text{перенос}} &= \Delta S_{\text{раств}} \\ \Delta G_{\text{перенос}} &= \Delta H_{\text{перенос}} - T\Delta S_{\text{перенос}} \end{aligned} \qquad (3.3)$$

Было установлено, что величина $\Delta H_{\text{перенос}}$ мала и положительна, а величина $\Delta S_{\text{перенос}}$ велика и положительна, как и следует из обсуждавшегося ранее примера масло — вода. Нозаки и Танфорд [16] определили величины $\Delta G_{\text{перенос}}$ для переноса боковых цепей канонических аминокислот из воды в этанол или диоксан, которые имитируют полярность внутренней части белка лучше, чем углеводородные растворители. Не предпринималось никаких попыток разделить значения $\Delta H_{\text{перенос}}$ и $\Delta S_{\text{перенос}}$. Некоторые из этих данных приведены на рис. 1.8. Все гидрофобные боковые цепи имеют отрицательные значения $\Delta G_{\text{перенос}}$, что указывает на их стремление избежать водного окружения. Напротив, остаток глицина имеет $\Delta G_{\text{перенос}}$ около нуля. Таким образом, основная цепь молекулы (у Gly нет боковой цепи) не отдает предпочтения ни внутренней части, ни поверхности белка.

Величина $\Delta G_{\text{перенос}}$ пропорциональна контактной поверхности. Линейная зависимость между площадью доступной поверхности неполярной боковой цепи и величиной $\Delta G_{\text{перенос}}$ согласуется с гипотезой «айсберга», поскольку можно ожидать, что число упорядоченных молекул воды пропорционально поверхности контакта. Как видно из рис. 1.8, на каждый диполь приходится уменьшение абсолютной величины $\Delta G_{\text{перенос}}$ примерно на 1,5 ккал/моль. Однако эта величина была определена для переноса в этанол или диоксан, где водородные связи не столь прочны, как во внутренней части белка. В случае белка внутренние водородные связи частично компенсируют это уменьшение и значения для более полярных боковых цепей, например Thr и Tyr, находятся, как видно из рис. 1.8, ближе к прямой неполярных остатков. Таким образом, выигрыш в $\Delta G_{\text{перенос}}$ примерно одинаков для полярных и неполярных групп, а величина $\Delta G_{\text{перенос}}$ приблизительно пропорциональна общей площади доступной (полярной и неполярной) поверхности с коэффициентом пропорциональности, равным 0,025 ккал/моль · Å^2.

Солевые мостики

В белках существует ограниченное число типов солевых мостиков. Солевые мостики встречаются во многих белках, например в гемоглобине [60] их образуют Lys-40 (α_2-цепь)/α-карбоксил His-146 (β_1-цепь); в парвальбумине [61] — Glu-81/Arg-75; в химотрипсине [18] и в трипсине [62] — α-аммонийная группа Ile-16/Asp-194. Энергию электростатического притяжения рассчитать для них легко; она составляет около 5 ккал/моль для соседних карбоксильной и аммонийной групп (табл. 3.1) в среде с диэлектрической проницаемостью $\varepsilon = 4$.

Вклад свободной энергии солевого мостика незначителен. Чтобы оценить вклад солевых мостиков в стабильность белка в растворителе, следует сравнить их электростатическую энергию с энергиями притяжения обоих зарядов к молекулам растворителя, которые по нашей системе подсчета включены в параметр $H_{\text{цепь}}$. В растворителе заряд притягивает окружающие диполи (молекул H_2O) за счет электростатического монополь-дипольного взаимодействия. Поскольку соответствующая энергия намного превосходит энергию электростатического притяжения между ионами, результирующая величина $\Delta H_{\text{цепь}}$ близка к нулю. Кроме того, стабильность зависит от свободной энергии [уравнение (3.1)], а не только от связывающей энергии, в связи с чем необходимо учитывать также изменение энтропии. Поскольку притягивающиеся диполи ориентированы вокруг заряда радиально, вокруг заряженной группы наблюдается более высокая упорядоченность воды — точно так же вокруг неполярной группы наблюдается более высокоупорядоченная квазикристаллическая структура воды. Вклад величины $\Delta S_{\text{раств}}$, способствующей образованию солевого мостика, существен [15]. Результирующая общая свободная энергия $\Delta G_{\text{перенос}}$, благоприятствующая образованию солевого мостика, составляет около 1 ккал/моль [60]. Удаление одного заряда из воды без образования солевого мостика полностью нарушает $\Delta G_{\text{перенос}}$. Поэтому перенос отдельного заряда внутрь белка или в другие гидрофобные среды, например в мембраны, очень невыгоден при величине $\Delta G_{\text{перенос}}$ порядка 10 ккал/моль.

3.6 Молекулярная упаковка

Энергия водородных связей используется эффективно. После общего рассмотрения невалентных сил, способствующих стабильности свернутой полипептидной цепи, попытаемся оценить их действительную эффективность. Исследования белковых структур показывают, что около 90% всех внутренних полярных групп образуют водородные связи [17]. Это согласуется с большим содержанием вторичной структуры, наблюдающимся в белках (гл. 5), и

показывает, что фактически именно на ее образование расходуется почти вся возможная энергия водородных связей.

Высокая плотность упаковки согласуется со значительной ролью дисперсионных сил. Об эффективности участия дисперсионных сил можно судить по плотности упаковки внутри белка, поскольку плотность упаковки отражает количество имеющихся контактов. Плотность упаковки определяется как отношение собственного объема молекулы (объем, заключенный внутри вандерваальсовой оболочки молекулы, который в свою очередь определяется вандерваальсовыми радиусами атомов на поверхности) к объему, реально занимаемому молекулой в пространстве. В кристаллах малых молекул реально занимаемый объем задается элементарной ячейкой, а объем молекулы внутри вандерваальсовой оболочки можно определить по координатам атомов в кристалле и их вандерваальсовым радиусам.

Чтобы определить общую плотность упаковки молекулы белка, вандерваальсов объем вытянутой цепи следует отнести к объему свернутой цепи, т. е. к объему молекулы нативного белка. Однако целесообразно вначале определить локальную плотность упаковки, которая в белке может варьировать в зависимости от определенных структурных особенностей. Интегрируя по всем локальным плотностям упаковки, можно затем получить среднюю плотность упаковки белка.

Полиэдр Вороного позволяет определить локальную плотность упаковки. Локальные плотности упаковки были впервые определены Ричардсом [63]. Для этой цели полипептидная цепь была разделена на небольшие атомные группы, такие, как метил, метилен, амид, гидроксил, и другие, содержащие один более тяжелый атом и до трех атомов водорода. Ароматические циклы и гуанидиниевая группа Arg также рассматривались как отдельные группы. Из рентгеноструктурных данных были определены центры этих групп. Затем пространство было разбито на полиэдры Вороного, как показано на рис. 3.4. Локальная плотность упаковки при данном центре есть отношение объема атомной группы к соответствующему занимаемому объему, т. е. к объему окружающего полиэдра Вороного. Чтобы избежать ошибок, связанных с учетом поверхности, в рассмотрение был включен монослой молекул H_2O, находящихся на поверхности. Таким образом, в полный полиэдр Вороного были включены только атомы белка, но не фактические молекулы H_2O. Это поясняется на рис. 3.4, где A — центр атома белка, а B, C и D — центры фактических молекул H_2O (см. также рис. 1.9). Эта процедура позволяет попутно определить площадь молекулярной поверхности. Лучше всего воспользоваться тем монослоем поверхностных молекул воды, который содержит их максимальное количество. Однако поскольку определить такой оптимальный слой трудно, то выбранный слой обычно содержит меньшее по сравнению с

оптимальным число молекул воды. Вводимая за счет этого ошибка невелика.

Белки упакованы так же плотно, как хорошие молекулярные кристаллы. Наблюдаемые локальные плотности упаковки в белках варьируют от 0,68 до 0,82. Низкая плотность найдена в активных центрах [63, 64], что подтверждает предположение о подвижности активных центров. Высокую плотность имеют гидрофобные ядра в центре белка*. Средняя плотность упаковки белка составляет около 0,75 (плотность упаковки правильных твердых сферических тел составляет 0,74). Для кристаллов малых молекул, связанных вандерваальсовыми силами, характерны значения от 0,70 до 0,78, в среднем 0,74. Стекла, масла или исключительно мягкие вандерваальсовы кристаллы (или некоторые кристаллы, построенные за счет направленных связей, например водородных связей обычного льда,

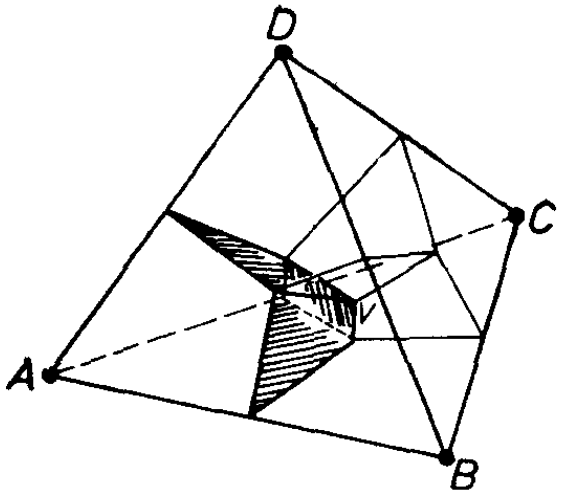


Рис. 3.4.

Разбиение пространства на полиэдры вокруг атомов.

Для данного набора атомов пространство можно разбить на неправильные полиэдры с атомами в их вершинах; один из таких полиэдров приведен на этом рисунке. Как показал Вороной (1908), четыре биссекторные плоскости между каждой парой атомов пересекаются в одной точке — вершине *V*. Эти плоскости делят тетраэдр на четыре неравные части. Часть, прилежащая к атому *A*, выделена заштрихованными плоскостями. Данный атом *A* может принадлежать произвольному числу тетраэдров. Чтобы определить общий полиэдр вокруг *A*, приведенное построение необходимо выполнить для всех таких тетраэдров и все части, прилежащие к *A*, нужно сложить. После повторения этой процедуры для всех атомов (атомы на поверхности набора исключаются) пространство полностью разделено. Плотность упаковки данного атома определяется как отношение объема внутри вандерваальсовой оболочки к объему окружающего полиэдра. Ричардс [63] перенес эту процедуру на белковые молекулы, которые содержат атомы нескольких типов с разными вандерваальсовыми радиусами. В этом случае процедура Вороного, которая дает точные результаты для одинаковых атомов, в своем первоначальном виде неприменима. В ее модифицированной форме, дающей удовлетворительные результаты, биссекторные плоскости проводятся через точки, которые делят расстояние между атомами, например между *A* и *B*, пропорционально соотношению вандерваальсовых радиусов *A* и *B* или, если *A* и *B* связаны ковалентно, в соответствии с отношением их ковалентных радиусов. Вследствие таких смещений четыре плоскости уже не пересекаются в точке *V*, а образуют небольшой тетраэдр вблизи *V*. Относительный объем этого «ошибочного» тетраэдра всего около 1%, и им можно пренебречь, если отношение радиусов не превышает 1,5. Чтобы эти отношения оставались на уровне 1,0, Ричардс ввел в атомные группы все (малые) H-атомы.

---

* Плотность упаковки нельзя смешивать с электронной плотностью или с плотностью массы. Гидрофобные ядра, как и все углеводороды, имеют значительно более низкую плотность массы, чем полярные участки.

рис. 3.2) имеют плотность упаковки ниже 0,70 и даже ниже 0,60. Таким образом, белки действительно упакованы столь же плотно, как малые молекулы в молекулярных кристаллах; естественно, что для них должна играть большую роль дисперсионная энергия.

Наблюдаемые плотности упаковки показывают, что метафору Козмана о капле масла не следует понимать буквально. Внутренняя часть белка более похожа на кристалл. Это подтверждается также слабой сжимаемостью белков по сравнению с маслом [65, 66]: $k_{\text{масло}}/k_{\text{белок}} \approx 20$.

Упаковка сохраняется в процессе эволюции. Концепция плотной упаковки белка и, следовательно, связанных с ней стерических ограничений, должна быть согласована с образованием большого числа направленных водородных связей практически всеми полярными группами внутри его молекулы. Напомним в связи с этим, что в обычном льде именно из-за образования линейных водородных связей плотность упаковки составляет всего 0,58. Высокая структурная оптимизация в белке подтверждает предположение о том, что эволюционные изменения гораздо реже касались внутренней части белка, чем его поверхности. Кроме того, имеется определенная тенденция ко взаимной компенсации внутренних изменений [67], например изменение Ile → Val может сопровождаться соседним изменением Gly → Ala, которое восстанавливает потерянную метильную группу.

Плотность упаковки может служить критерием в предсказании структуры. Плотность упаковки предоставляет возможность независимо проверить результаты расчетного моделирования свертывания цепи (разд. 8.6) и предсказания структур по известным аминокислотным последовательностям путем сравнения с филогенетически связанными белками (гл. 9). Этот критерий не сводится только к сравнению со средними плотностями упаковки. Объем, занимаемый боковыми цепями (сумма объемов всех составляющих данную цепь атомных групп), специфичен для каждой боковой цепи [63, 68] и варьирует всего лишь приблизительно на 5%. Поэтому в качестве еще одного критерия в предсказании белковой структуры можно воспользоваться объемом, занимаемым каждой боковой цепью.

Заключение

Как обсуждалось в гл. 1, образование и стабильность белковой структуры в значительной мере зависят от невалентных сил. Их подразделяют на (1) отталкивание между валентно несвязанными атомами, (2) взаимодействие между осциллирующими диполями, приводящее к дисперсионным силам притяжения, (3) электростатическое притяжение и отталкивание парциальных зарядов, а также притяжение целочисленных зарядов в солевых мостиках, и

(4) водородные связи. Поскольку белки существуют только в водных средах, следует рассматривать систему в целом: и полипептидную цепь, и растворитель. В уравнении (3.2) энтропия и члены, отвечающие связывающей энергии, рассматриваются отдельно. Наибольший вклад в стабильность белка вносят гидрофобные силы. Высокая плотность упаковки, наблюдающаяся в белках, свидетельствует об эффективном использовании всех невалентных сил в организации их структуры. Внутренняя часть белка упакована так же плотно, как хороший кристалл.

Глава 4 Ковалентная структура белков

Общее положение о функциональной форме белка, получаемой путем синтеза полипептидной цепи на рибосомах и последующего спонтанного свертывания, с которого был начат разд. 1.1, нуждается в значительной детализации. В этой главе обсуждается ряд важных биологических фактов и явлений, позволяющих уточнить это представление.

4.1 Ансамбли цепей

Функциональные и структурные домены

Основная идея, заключающаяся в том, что *один структурный ген** отвечает *одной полипептидной цепи*, *эквивалентной одному функциональному звену* [74], требует уточнения. В настоящее время представляется нерациональным отождествлять полипептидную цепь с функциональной единицей. Начиная с 1969 г. [75], появилось большое число данных, указывающих [76], что основной единицей следует считать «домен», а не полипептидную цепь и что все остальные категории следует сводить к этому базису.

Функциональные домены и функционально автономные области полипептидной цепи. Домены были впервые обнаружены при выявлении корреляции генной и полипептидной структур [77]. Домены представляют собой субобласти полипептидной цепи (закодированные в субобластях структурного гена), которые автономны

---

* В высших организмах менее 10 % общей нуклеотидной последовательности структурного гена, так называемые структурные зоны [78, 79], кодируют белок. Исследования генов, кодирующих глобин, овальбумин, а также некоторые белки SV40 и вируса полиомы, подтверждают мозаичный характер эукариотического структурного гена. Нуклеотидные последовательности ДНК, которые транслируются в аминокислотные последовательности, не монотонны по строению, а прерываются участками нетранслирующейся ДНК. Первичная матрица РНК содержит внутренние области, которые должны быть затем исключены, и окончательная мРНК представляет собой «перекроенный» продукт (см. краткий обзор [78]).

в том смысле, что они обладают всеми характеристиками всего глобулярного белка. Многие такие субобласти были обнаружены как продукты неполного протеолиза [76]. Это означает, что индивидуальные субобласти часто можно выделить из полипептидной цепи без потери ими их свойств.

Значимость представления о доменах наиболее очевидна для некоторых ферментов, участвующих в биосинтезе His и Trp (разд.

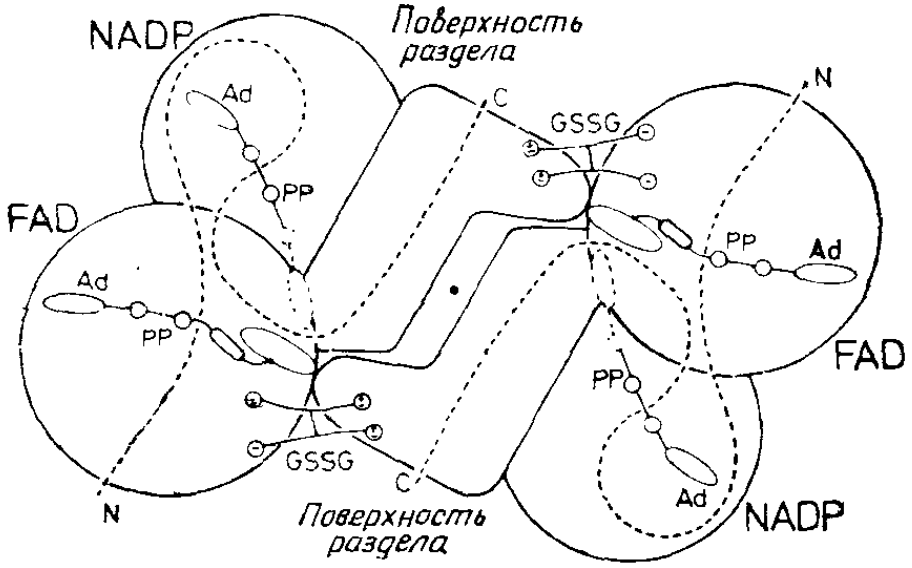


Рис. 4.1.

Схема димерного фермента глутатионредуктазы.

Каждая субъединица состоит из трех доменов; один из них присоединяет FAD (флавинадениндинуклеотид), другой присоединяет NADP (никотинамидадениндинуклеотидфосфат), третий образует поверхность раздела. Общий ход полипептидной цепи показан пунктирной линией. Кофакторы FAD и NADP присоединяются в вытянутых конформациях. Субстрат глутатион (GSSG) помещается между субъединицами. Каждый активный центр образован четырьмя доменами [124].

9.4). В этом случае функция белка требует большого числа доменов. Оказалось, что эту функцию могут независимо выполнять домены, которые расположены или на одной полипептидной цепи, или на отдельных полипептидных цепях, связанных невалентными силами, или на совершенно разных цепях.

Структурные домены — геометрически обособленные образования. Поскольку описанные выше субобласти идентифицируются по наблюдаемым свойствам цепи (например, лигандприсоединяющая или ферментативная активность), они представляют собой «функциональные домены» [76]. По мере развития структурного анализа белка было показано, что функциональные домены состоят из одного или более «структурных доменов». Структурные домены были обнаружены при изучении многих трехмерных белковых структур, в частности глутатионредуктазы (рис. 4.1). Это геометрически обособленные образования с молекулярной массой около 20 000. Почти все глобулярные белки можно подразделить на такие субобласти. По-видимому, большинство функциональных доменов с молекулярной массой свыше 20 000 состоят из более чем одного структур-

ного домена. «Активный центр» такого функционального домена обычно расположен на границе двух структурных доменов.

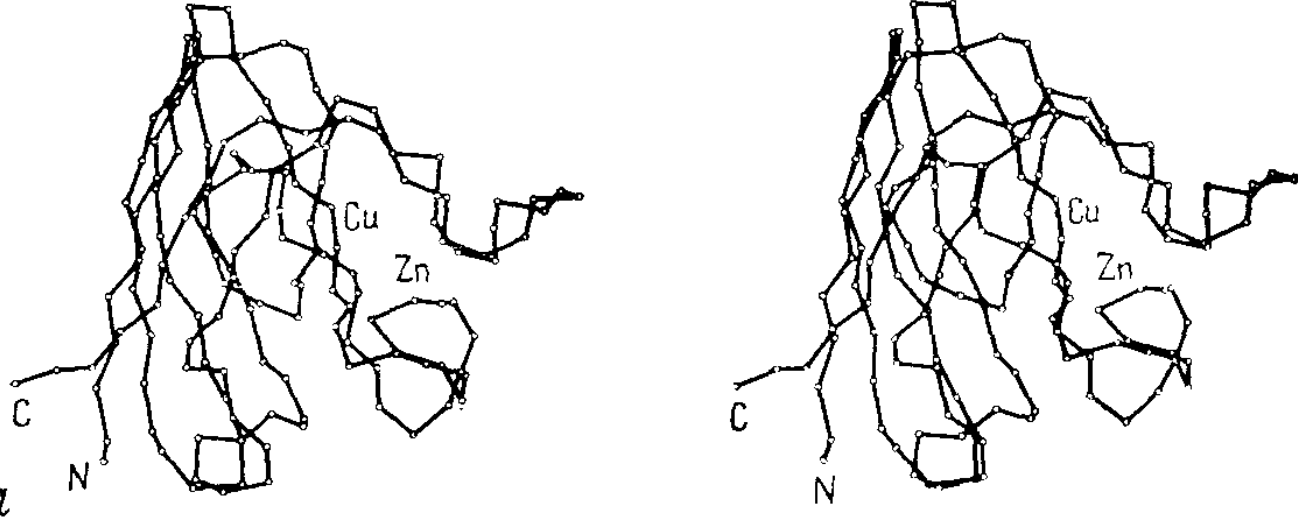
Белки могут конструироваться с использованием модульных систем. Важная роль структурных доменов как основных единиц следует из сравнения трехмерных белковых структур. В разных белках встречаются одни и те же структурные домены, обнаруживаемые по характеристическому свертыванию цепи. Типичными примерами повторения структурных доменов являются домены иммуноглобулина (рис. 4.2), NAD-связывающий домен (рис. 5.17, *б*) или TIM-полость (рис. 5.17, *д*), имеющаяся в триозофосфатизомеразе и пируваткиназе [80].

Подразделение глобулярных белков на структурные домены приблизительно одинаковой величины и присутствие одинаковых структурных доменов в разных белках привели к гипотезе, согласно которой большинство белков сконструировано как модульные системы из структурных доменов в качестве модулей (см. также разд. 9.4.). В соответствии с такой гипотезой не следует придавать слишком большого значения всей полипептидной цепи, а концентрировать внимание на функциональных и, насколько возможно, структурных доменах. Соответственно, основная концепция [74] должна быть приведена к следующему виду: «одна полипептидная цепь = =один или несколько функциональных доменов, один функциональный домен = один или несколько структурных доменов».

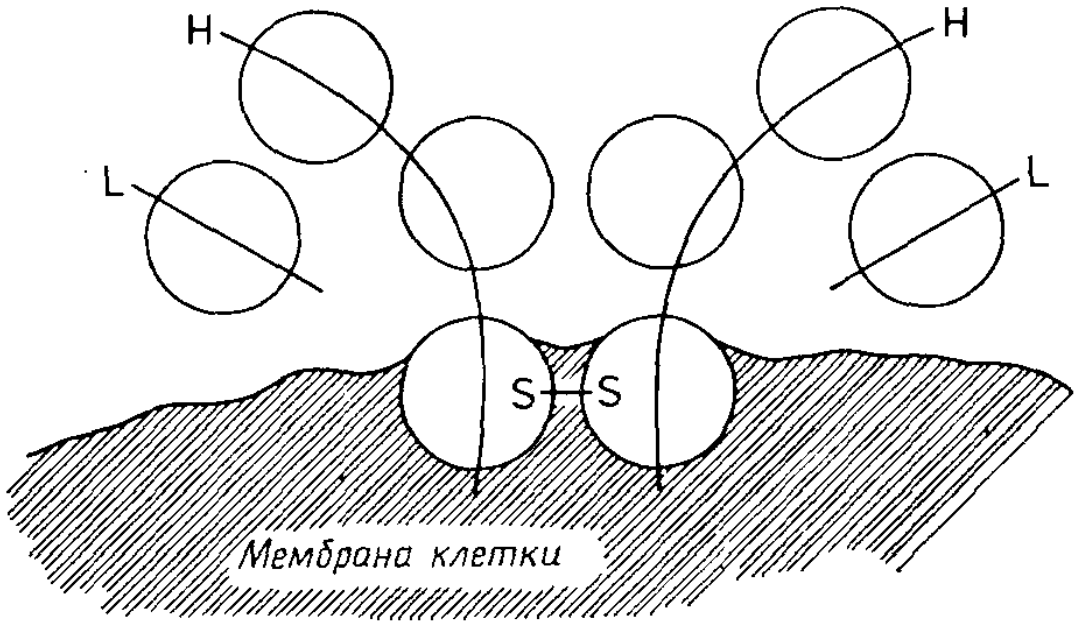
Влияние окружения

Как было показано в предыдущем разделе, основной единицей является домен. Ассоциации таких доменов имеют второстепенное значение. В последующем изложении рассматривается, каким образом окружение, в котором находится белок, влияет на характер объединения доменов.

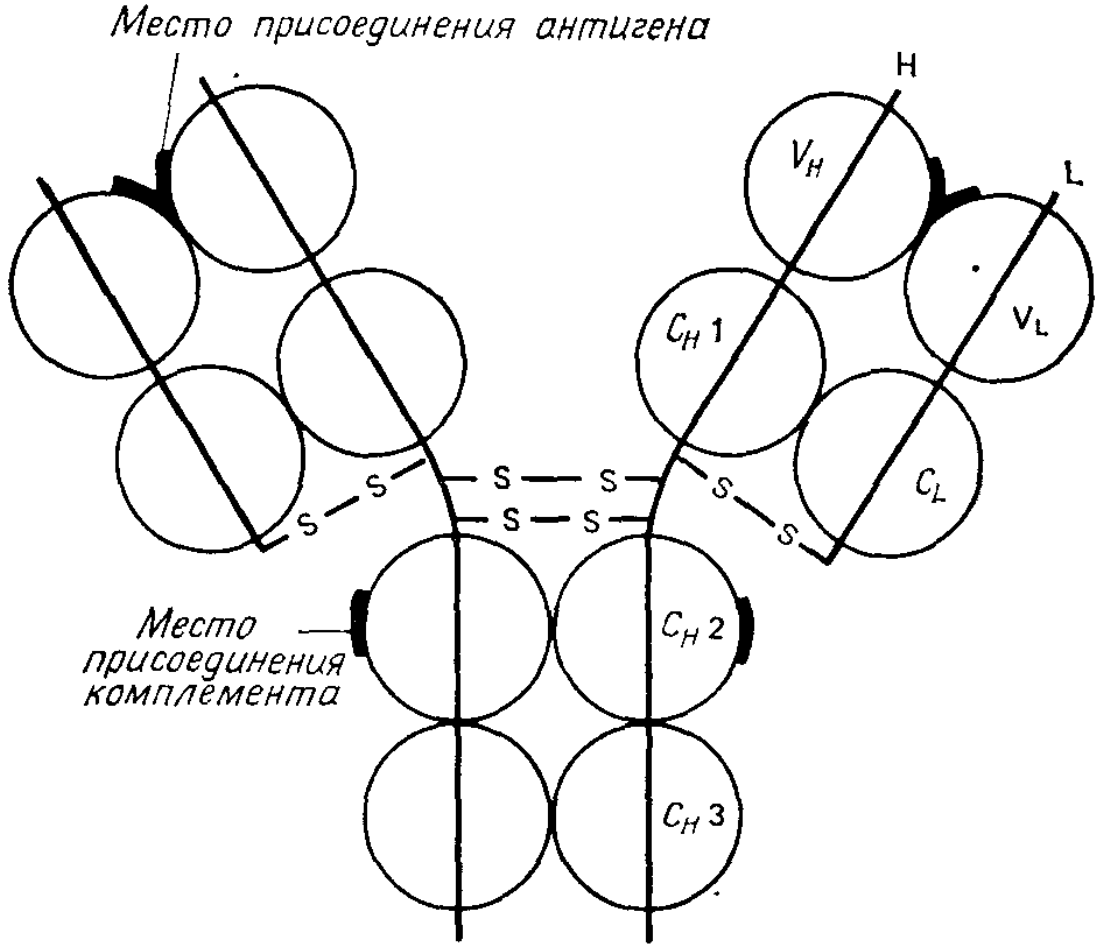
Олигомеры в отличие от мономеров могут диссоциировать. Белки обычно подразделяют на мономеры и олигомеры. Согласно определению Клотца и сотр. [81], белок представляет собой «мономер», если он состоит только из одной полипептидной цепи или если он построен из нескольких цепей, связанных ковалентно (например, дисульфидными мостиками). По этой номенклатуре такие белки, как инсулин, α-химотрипсин и иммуноглобулины, представляющие собой образования из валентно-связанных цепей, должны быть отнесены к мономерам. Отличительная особенность «олигомерных» белков состоит в том, что они построены из так называемых субъединиц, т. е. из связанных невалентными силами более мелких образований (рис. 4.1 и 5.18). Как указывалось выше, мономеры могут состоять из нескольких функциональных доменов или из еще большего числа структурных доменов. Это относится и к субъединицам олигомеров, хотя субъединица часто эквивалентна функциональному домену.


Cu
Zn
C
N
Cu
Zn
C
N

а


H
H
L
L
S
S
Мембрана клетки

б


Место присоединения антигена
H
V_H
L
C_H 1
V_L
C_L
S
S
S
S
S
S
S
S
Место присоединения комплемента
C_H 2
C_H 3

в

Три основных физиологических вида белковых молекул

Типы связей между функциональными доменами (ковалентная, невалентная), так же как и их отсутствие, можно коррелировать с физиологической средой, в которой существует белок. Большинство внутриклеточных белков олигомерны, белки плазмы крови — крупные мономеры, состоящие из нескольких функциональных доменов, а белки, действующие вне организма*, представляют собой небольшие мономеры. Попытаемся пояснить некоторые аспекты такого распределения.

Внеэпителиальные белки — небольшие мономеры. Вне организма судьба отдельной молекулы (например, пищеварительного белка или лизоцима человека) очень неопределенна. Поэтому целесообразно образовать из выделяемого на эти белки вещества как можно больше независимых единиц, чем и объясняются их малые размеры. Кроме того, эти белки мономерны, поскольку олигомеры склонны к диссоциации при растворении.

Белки плазмы крови — крупные мономеры. Существенное свойство специфических белков плазмы состоит именно в том, что они способны задерживаться в плазме крови. Только белки с молекулярными массами около 60 000 и более достаточно велики для того, чтобы они не проникали с легкостью во внешнее пространство, как это происходит с более мелкими молекулами ($< 20\,000$). Кроме того, небольшие белки фильтруются, обрабатываются и выделяются почками.

Крупные белки плазмы крови могут быть либо олигомерами, либо мономерами. Однако олигомеры способны к диссоциации, а субъединицы могут фильтроваться тем же путем, что и мелкие мономеры. Это видно на примере β_2-микроглобулина (рис. 4.2, *б*), небольшой растворимой субъединицы белков HL-A (которые

---

* Границами между внутренней средой животных и окружающей средой служат эпителиальные поверхности; например, содержимое желудочно-кишечного тракта относится к окружающей среде.

Рис. 4.2.

Белки, состоящие из Ig-доменов.

а — субъединица пероксид-дисмутазы (детальная структура) [800]. Свертывание цепи показано в виде стереоизображения C_α-остова (разд. 7.2); отмечены N- и C-концы обеих цепей. Субъединица включает Ig-домен и медьсодержащий каталитический центр с Zn-ионом вблизи него. *б* — антиген тканевой совместимости (human leukocyte antigen = = белок HL-A). Молекула симметрична; каждая половина состоит из тяжелой цепи H и легкой цепи L. Легкая цепь, которая представляет один домен Ig, называется также β_2-микроглобулином. Обе тяжелые цепи соединены дисульфидным мостиком, который, по-видимому, находится в присоединяющейся к мембране части молекулы. *в* — иммуноглобулин G. Молекула симметрична; каждая половина состоит из тяжелой цепи H и легкой цепи L. Каждая из цепей связана с по меньшей мере одной другой цепью одним или бо́льшим числом дисульфидных мостиков. Использованы стандартные обозначения индивидуальных доменов вдоль L-цепи и вдоль H-цепи.

Каждый большой круг на *б* и *в* означает один домен Ig. Углеводные части белков не показаны.

присоединяются к поверхности некоторых клеток крови, а также некоторых других клеток). Известно, что эта субъединица олигомерного белка проникает через стенки капиллярных сосудов; к тому же при некоторых условиях (например, при отравлении кадмием) большие количества микроглобулина выделяются с мочой. В связи с этим очевидно, что истинными белками плазмы должны быть крупные мономерные, а не олигомерные белки. Типичным примером может служить сывороточный альбумин (рис. 7.2, *б*), построенный из набора функциональных доменов [76, 82] в одной полипептидной цепи. Некоторые плазменные полипептиды заметно увеличивают свой размер за счет ковалентно присоединенных углеводных остатков.

Внутриклеточные белки, как правило, — олигомеры. Рассмотрим теперь внутриклеточные белки, которые представляют собой преимущественно олигомеры. Небольшие мономерные белки, как, например, изоферменты аденилаткиназы (M = 22 000) [83, 84], встречаются в клетках редко. Олигомеры в этом случае имеют ряд преимуществ перед крупными одиночными полипептидными цепями (разд. 4.1). В отличие от плазмы крови в клетках олигомеры весьма эффективны, поскольку клеточная мембрана непроницаема даже для небольших белков, так что потери олигомеров в форме их субъединиц исключены.

Образование олигомеров понижает осмотическое давление во внутриклеточном пространстве. Кроме того, отношение поверхность/объем для олигомера меньше, чем для мономера. Поэтому олигомер связывает меньше молекул воды и вносит меньший вклад в вязкость внутри клетки. Наконец, олигомерные белки обычно кооперированы и хорошо регулируются эффекторами. Хорошо известным примером физиологической оптимальности этих свойств является тетрамерный гемоглобин эритроцитов млекопитающих [85]. Однако такие свойства существенны и для менее специализированных клеток, содержащих самые разнообразные белки.

Иммуноглобулины и белки HL-A

Ковалентная связь между цепями иммуноглобулина необходима для их функции. Иммуноглобулины плазмы прекрасно иллюстрируют многие аспекты корреляций между структурой белка, его функцией и окружением [81]. В иммуноглобулине (рис. 4.2, *в*) каждый из двух центров присоединения антигена образован полипептидными цепями двух типов. Объединение двух идентичных легких цепей и двух идентичных тяжелых цепей существенно для функции молекулы. Неконтролируемый обмен цепей в тысячах различных иммуноглобулинов плазмы крови воспрепятствовал бы эффективному взаимодействию этих антител с антигенами, ко-

торые вызвали их синтез. Существенное для специфичности ковалентное соединение четырех цепей предохраняет иммуноглобулины от случайных нарушений их структуры [81].

Белки HL-A должны быть олигомерными. В отличие от иммуноглобулинов биологическая специфичность белка HL-A (рис. 4.2, *б*) определяется только двумя (идентичными) тяжелыми цепями; легкие цепи всех белков HL-A в организме идентичны [86, 87]. Поэтому нет необходимости в ковалентной связи легких цепей с тяжелыми; такая связь должна быть только между тяжелыми цепями. Диссоциация и реассоциация легких цепей не могут нарушить молекулярной специфичности. Трехмерная структура белков HL-A была предсказана на основании данных по их специфичности и известной структуре иммуноглобулинов. Таким образом, физиологические данные способствуют появлению рабочих гипотез при исследовании структуры белка.

С одной стороны, олигомерную структуру белка HL-A в принципе можно рассматривать как недостаток, поскольку она приводит к фильтрации и деградации легкой цепи (β_2-микроглобин, рис. 4.2, *б*) в почках [88]. Однако компенсирующим преимуществом является то, что легкая цепь белков HL-A синтезируется специализированными клетками, и, поскольку они распределены, как и гормоны, по всему организму, частичные потери в почках не имеют серьезного значения. По-видимому, целесообразно начать систематические исследования клеток, образующих легкую цепь. Таким путем результаты изучения белковых структур также могут привести к появлению рабочих гипотез в общей физиологии.

Олигомерные белки

Олигомеры проще, чем одиночные гигантские полипептидные цепи. Олигомерные белки имеют некоторое принципиальное преимущество перед гигантскими полипептидными цепями, состоящими из отдельных функциональных доменов. По-видимому, этим объясняется тот факт, что большинство крупных белков олигомерны [81].

(а) *Если некоторые или все субъединицы идентичны, для кодирования олигомерного белка требуется значительно меньше генетического материала, чем для кодирования отдельной гигантской полипептидной цепи.* Несмотря на то что изолированные субъединицы многих олигомерных белков нефункциональны [89, 90], такие субъединицы несут всю информацию, необходимую для самообъединения и образования функциональной единицы. Это позволяет сократить участок ДНК (или РНК), кодирующий белок данной функции. Эффект хорошо иллюстрируется олигомерами средней величины, например тетрамерной дегидрогеназой [81, 91] и особенно такими биологически важными структурами, как белок оболочки вируса табачной мозаики, который состоит из 2130 самообъединяющихся

идентичных субъединиц с 158 аминокислотными остатками в каждой из них.

(б) *Вероятность возникновения дефекта у крупных полипептидных цепей значительно выше, чем у олигомерных белков той же величины.* Допустим, что при данном механизме синтеза белка возникает 1 ошибка на 1000 аминокислотных остатков [81] белка; тогда для отдельного полипептида, состоящего из 1000 остатков, вероятность дефекта велика [из распределения Пуассона следует, что среднее число правильных копий составляет $\exp(-1000 \cdot \frac{1}{1000}) = 37\%$]. Напротив, синтез четырех субъединиц, каждая из которых состоит из 250 остатков, дает большее количество недефектных субъединиц, а именно $\exp(-250 \cdot \frac{1}{1000}) = 78\%$*.

(в) *Олигомерные ансамбли значительно лучше приспособлены к молекулярной эволюции* [92]. Они значительно быстрее приспосабливаются к окружающим изменениям, чем крупные одиночные полипептидные цепи.

(г) *Образование олигомеров—наиболее простое решение проблем, стоявших перед первичными клетками.* Поскольку клеточные мембраны весьма несовершенны, утечка молекул с $M = 20\,000$ через поры в мембранах оказывалась существенной. Агрегация была, по-видимому, более простым способом увеличения молекул, чем синтез больших цепей [93].

4.2 Дисульфидные связи

Большинство дисульфидных связей in vitro образуется спонтанно. Как уже упоминалось, дисульфидные мостики между парами цистеиновых остатков могут объединять различные участки белка, приводя к образованиям из ковалентно связанных цепей. Дисульфидные связи встречаются также в одиночных полипептидных цепях. Остаток цистеина полипептидной цепи может объединяться только с вполне определенным остатком цистеина, поэтому набор дисульфидных мостиков строго индивидуален для данного белка [94—100]. Кроме того, как правило, эти связи образуются спонтанно in vitro, не требуя внешних воздействий, например, ферментов [94]. Некоторые известные исключения, например дисульфидные связи инсулина и химотрипсина [101], обсуждаются в разд. 8.2.

---

* Можно возразить, что это преимущество невелико, поскольку введение одной дефектной субъединицы в тетрамерный белок может подавить функцию всего олигомера. Но обычно этого не происходит; олигомерные белки находятся в динамическом состоянии диссоциации и реассоциации; субъединицы обмениваются между олигомерами и дефектные устраняются. С помощью такого диспропорционирования могут быть исправлены природные дефекты и химические изъяны индивидуальных субъединиц [81].

Действительным исключением представлялся случай иммуноглобулина А, в котором для образования некоторых дисульфидных связей необходима *J*-цепь; однако *J*-цепь вряд ли можно рассматривать как внешний фактор [102].

In vivo мостики S—S образуются в восстановительных условиях. Цистеиновые остатки легко окисляются in vitro в остатки цистина, однако до сих пор неясно, каким образом дисульфидсодержащие белки образуются in vivo и когда они возникли в процессе эволюции белков. Эти две проблемы взаимосвязаны. Предполагается, что предшественники современных белков эволюционировали в восстановительной атмосфере; в существующих на сегодня клетках восстановительная среда поддерживается глютатионовой системой (5 мМ восстановленного глютатиона и 0,1 мМ окисленного глютатиона). При таких условиях образование дисульфидных мостиков в белках может происходить и не спонтанно. Поэтому предположили [103], что система, которая включает окисленный цистамин и связывающийся с мембраной фермент, может обеспечить введение связей S—S в белки. Эти ферменты, естественно, отличаются от дисульфидизомеразы белка [104—107], которая перестраивает имеющиеся в белках дисульфидные связи до получения (*мета*) стабильных структур.

Общая функция дисульфидных связей состоит в повышении стабильности свернутых белков (см. разд. 8.4 и работы [94] и [108]). Поэтому не вызывает удивления тот факт, что разрушение некоторых мостиков S—S не влечет за собой нарушения функций белка (табл. 4.1). Например, в случае α-амилазы можно восстановить все три дисульфидных звена и это не уменьшит ферментативной активности [109]. Однако известно много примеров белков, в которых дисульфидные связи несут весьма специфические функции.

Таблица 4.1

Восстановление дисульфидных мостиков

Белок	Литература	Число		
		мостиков S—S	остатков	S—S-мостиков, разрушающихся без потери активности
Ингибитор трипсина поджелудочной железы быка	452	3	58	1
Рибонуклеаза	109	4	124	1
Лизоцим	161, 162	4	129	Один из 3
Трипсин	109	6	224	3
Химотрипсин	109	5	242	2
Амилаза	109	3	410	3

Роль мостиков S—S во внеклеточных и внутриклеточных белках

Дисульфидные мостики определяют механические свойства внеклеточных белков. Дисульфидные мостики обычны в белках, которые переносятся или действуют во внеклеточном пространстве; типичными примерами служат змеиные яды и другие токсины, пептидные гормоны, пищеварительные ферменты, белки комплемента, иммуноглобулины, лизоцимы и белки молока. Кроме того, эти мостики играют важную роль в некоторых крупных структурах. Свойства вязкости и эластичности различных природных продуктов по крайней мере отчасти определяются дисульфидными мостиками между структурными белками [110]. Поперечные связи между молекулами кератина придают эластичность шерсти и волосу [110], когезионноэластичный характер теста из пшеничной муки определяется дисульфидами глютенина, а трехмерная сеть дисульфидов глютенина создает трудности при влажном помоле зерна. Таким образом, оказывается, что успехи в таких древних занятиях, как помол зерна, обработка шерсти и даже парикмахерское искусство, зависят от сложных конструкций дисульфидных связей [110].

Дисульфидные мостики определяют химические свойства внутриклеточных белков. Из тех белков, которые никогда не покидают своей клетки, лишь немногие содержат дисульфидные мостики. Во внутриклеточных белках дисульфиды, как правило, выполняют функции, не связанные со стабилизацией структуры. Так, в глутатионредуктазе дисульфид играет активную роль в катализе [111] (рис. 11.4); в треониндезаминазе дисульфиды (рис. 8.3) стабилизируют нужную конформацию ассоциированных белковых субъединиц [112]. Тиолдисульфидный обмен (рис. 4.3) может играть функциональную роль в регуляции ферментативной активности [113—116]; обнаружено также, что обмен по этому механизму может происходить при взаимодействии связывающегося с мембраной сократительного белка и микротубулярного белка в процессе цикла деления яйца морского ежа [117]. Предположение, что реакции тиол-сульфидного взаимодействия могут играть роль в специфических взаимодействиях между мембранными белками и внеклеточными белками, было отвергнуто в результате изучения двух известных систем: инсулин — мембранный рецептор [118, 119] и белки комплемента — мембранные рецепторы [120].

Своеобразная конструкция, основанная на образовании дисульфидных мостиков, обнаружена в ферментах триптофаназе [121] и треониндезаминазе [122] (рис. 8.3), выделенной из *Salmonella*. Здесь 4 идентичные полипептидные цепи попарно связаны дисульфидными мостиками, так что белки содержат по 2 (мономерные) субъединицы. Такая асимметричная агрегация, вероятно, связана с тем обстоятельством, что эти ферменты содержат по 2 центра связывания пиридоксальфосфата. Аналогичную функцию выполняют

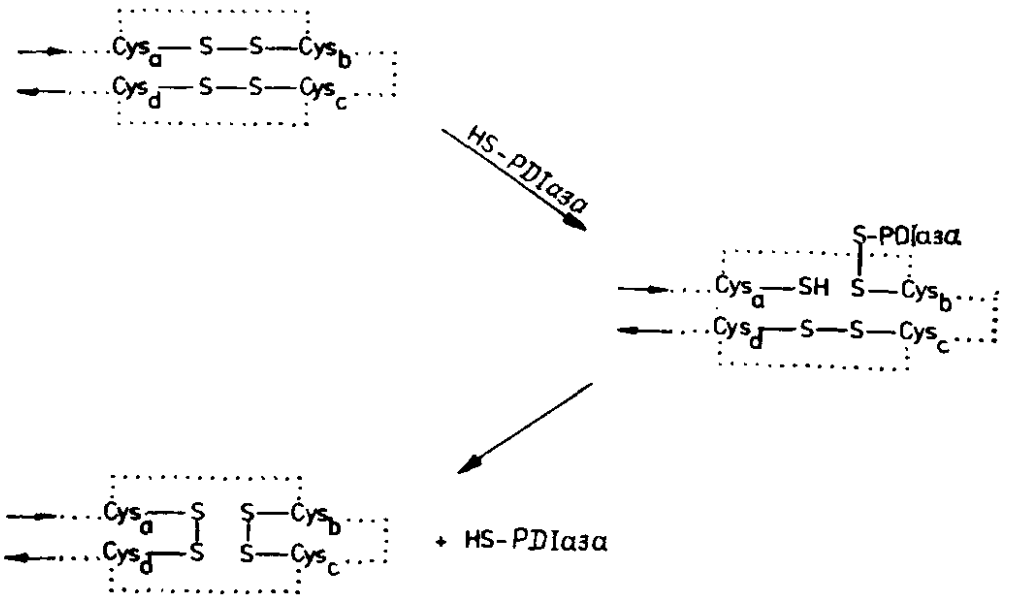


Рис. 4.3.

Перегруппировка дисульфидных мостиков белка в результате реакций тиол-дисульфидного обмена.

Cys_a, Cys_b и т. д. — дисульфидобразующие остатки полипептидиой цепи, ход которой намечен точками. Первым шагом перегруппировки дисульфида является катализ ферментом дисульфидизомеразой (PDIазой, ее каталитической тнольиой HS-группой); при этой реакции образуется смешанный дисульфид белка и фермента [109]. Ряд последующих тиолдисульфидных взаимодействий приводит к иовому набору дисульфидных мостиков в белке и к регенерации каталитической тиольной группы PDIазы. Дисульфидную изомеризацию белков может катализировать не только фермент дисульфидизомераза [104—106], но и другие S—H-соединеиия, иапример восстановленный глютатион или дитиотрейтол (при pH>7).

дисульфидные мостики между легкой и тяжелой цепями в иммуноглобулинах, где они обусловливают специфичность центра связывания.

Дисульфидиые мостики как составиые части структуры белков

Дисульфидные мостики, приводящие к образованию петель в полипептидной цепи, обнаружены в нескольких белках (пепсине, тиоредоксине, А-цепи инсулина, фиброине шелка [145], липоамидной дегидрогеназе и других пиридиннуклеотиддисульфидных оксиредуктазах [111]). Между мостиковыми цистеиновыми остатками в полипептидной цепи находится 2—4 остатка. Рассмотрение моделей, а также рентгеноструктурный анализ показывают, что такие петли имеют уплощенную жесткую структуру. В глутатионредуктазе и родственных ферментах в петле участвует изоаллоксазиновое кольцо FAD [123, 124].

Другой часто встречающийся структурный элемент — это последовательность — Cys—Cys—, когда оба остатка образуют дисульфидные мостики с другими цистеиновыми остатками [82]; геометрия пептидного звена не допускает образования дисульфидного мостика между соседними остатками. Таким образом, последова-

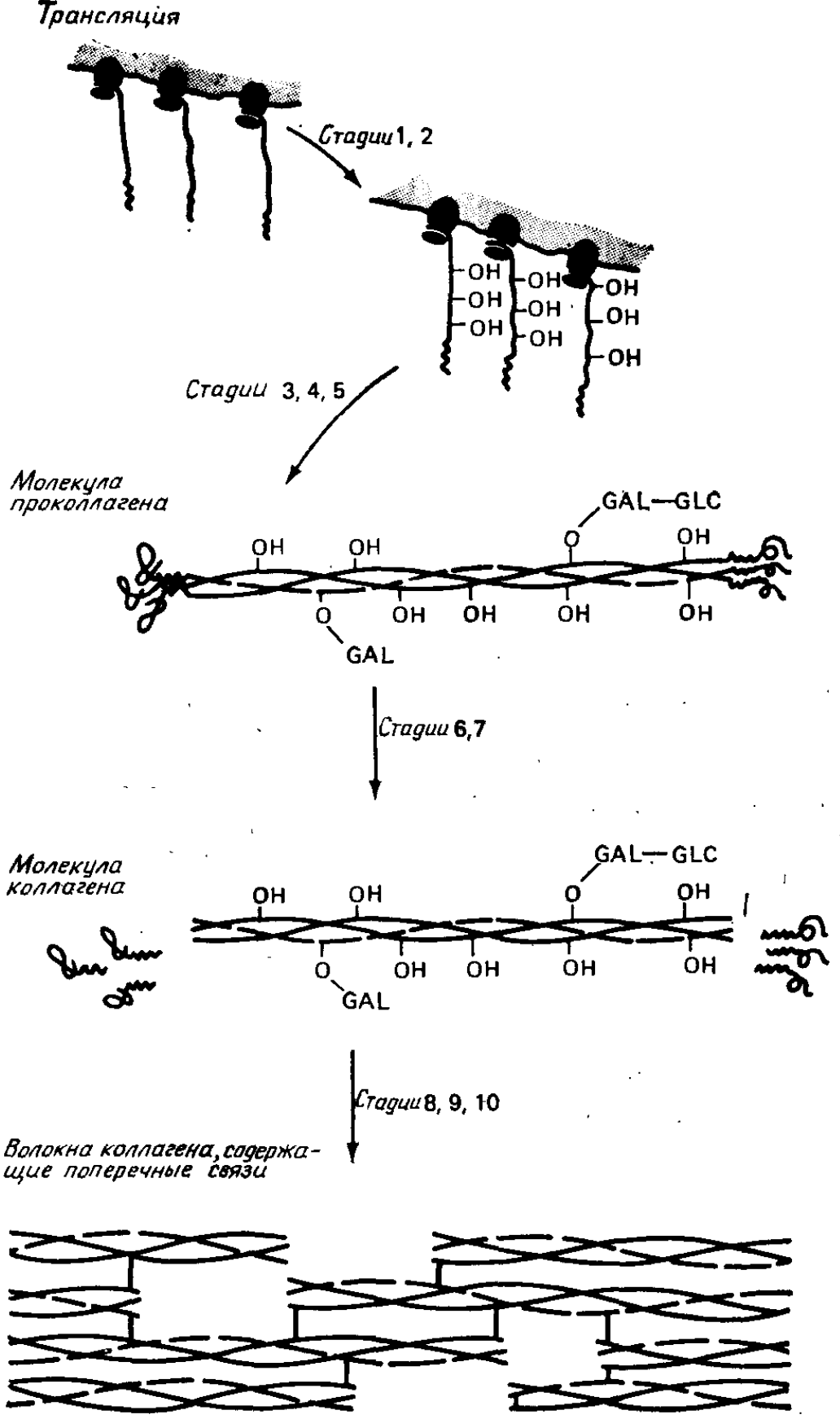


Рис. 4.4.

Биогенез коллагена [131].

Идентифицированы следующие стадии: *Стадия 1.* Синтез пред-α_1-цепей и пред-α_2-цепей в отношении 2 : 1. Каждая цепь содержит 1300 остатков. *Стадия 2.* Гидроксилирование некоторых остатков Pro и Lys. *Стадия 3.* Присоединение сахаров (Glc—Gal) к гидроксилированным остаткам. *Стадия 4.* Образование тримера, связанное с образованием межцепных дисульфидных мостиков, которые, по-видимому, расположены на концевых участках цепей. *Стадия 5.* Образование тройной спирали в центре молекулы проколлагена. *Стадия 6.* Выделение проколлагена в межклеточное пространство. *Стадия 7.* Удаление глобулярных концевых фрагментов коллагена. *Стадия 8.* (Спонтанное) самоагрегирование молекул коллагена (образование волокон). *Стадия 9.* Дезаминирование Lys- и гидроксилированных остатков с образованием альдегидов. *Стадия 10.* Образование перекрестных связей между полипептидными цепями в волокне за счет реакций, использующих эти альдегиды, а также боковые цепи Lys и His (подробности см. в работах [127] и [131]).

тельность —Cys—Cys— создает возможность связывания и сближения трех сегментов цепи. Маловероятно, чтобы присутствие фрагментов —Cys—Cys—, соединенных S—S-связями, в 15 белках было случайным или что оно обусловлено общим происхождением этих белков. Поэтому было высказано предположение, что последовательность —Cys—Cys— представляет особенно полезную деталь белковой структуры и что многие белки в процессе эволюции независимо выработали эту последовательность (рис. 7.2).

4.3 Модификации главной цепи, контролируемые ферментами

Для понимания многих аспектов организации структуры белка важно иметь представление о его физиологической судьбе на протяжении всего цикла от синтеза до разрушения. Это особенно важно для тех белков, которые претерпевают посттрансляционные (вторичные) модификации, т. е. изменения в системе ковалентных связей после их синтеза на рибосоме.

Классическим примером того, насколько значительно модифицируются и обрабатываются белки во время синтеза и после него, является биогенез коллагена [85, 125—131]. По-видимому, он включает большее число реакций, чем показано на рис. 4.4. Очевидно, однако, что путь коллагена весьма примечателен. В последующих разделах описаны некоторые общие аспекты модификации белка in vivo.

N- и C-концы

N-концевые аминогруппы часто модифицируются. N-Концевая α-аммониевая группа (pK 8) многих белков претерпевает вторичную модификацию (см. обзорную работу [132]). Она ацетилирована в белке оболочки вируса табачной мозаики, в большинстве цитохромов *c*-типа, в белках мышц, актине, миозине, тропомиозине, парвальбумине, аденилаткиназе и лактатдегидрогеназе. Другими примерами могут служить кератин шерсти и α-меланоцитстимулирующий гормон [132]. Аналогичная модификация α-аминогруппы — формилирование обнаружена в пчелином яде мелиттине и в гемоглобине миноги. Существенно иным изменением N-конца полипептида является превращение N-концевой глутаминовой кислоты в пирролидонкарбонильную группу. Такая модификация обнаружена во многих белках, функционирующих во внеклеточном простран-

стве, как, например, в легкой и тяжелой цепях иммуноглобулинов человека, в гормонах, выделенных из различных источников, а также в змеиных ядах [132—134].

Биологическое значение маскировки α-аминогрупп недостаточно ясно; возможно, она защищает белок от атаки аминопептидаз или способствует закреплению N-концевой части полипептида в аполярном окружении либо на молекуле рецептора, либо внутри самого белка, чтобы препятствовать его контакту с раствором. Это предположение не относится к метилированию α-аминогруппы, обнаруженному в рибосомных белках, выделенных из *Escherichia coli* [135], поскольку метилирование не элиминирует заряд. Физиологическая роль ацетилирования α-аминогруппы совершенно ясна в случае некоторых гемоглобинов рыб: такая модификация помогает сохранять способность к связыванию кислорода независимо от pH-среды, что предотвращает выделение избыточного кислорода в плавательный пузырь [136] (разд. 10.3).

С-концевые модификации редки. Для полипептидов природного происхождения единственным известным заместителем на С-конце является амидная группа. Амидированные С-концевые группы найдены, в частности, в гормонах и в пчелиных ядах [132]. В функцию амидной группы, по-видимому, входит защита пептидной цепи от атаки карбоксипептидазы или облегчение аккомодации отрицательно заряженного С-конца (величина pK α-CO_2H-группы около 3) в неполярном окружении. Как правило, С-концевые модификации значительно более редки, чем N-концевые.

Специфическое расщепление полипептидной цепи

Специфический протеолиз — удобный процесс для образования сложных белковых структур. Во многих случаях белки модифицируются путем расщепления одной или нескольких пептидных связей. Для обозначения этого типа катализируемых ферментами реакций, которые играют доминирующую роль во многих физиологических процессах [137—139], используются термины «ограниченный протеолиз» или «специфический протеолиз» (табл. 4.2). Хорошо известными примерами специфического расщепления полипептидов являются активация предшественников пищеварительных ферментов, морфогенетические процессы в бактериальных вирусах и каскадные процессы коагуляции и комплементного действия крови [138, 140]. Недавно было показано, что механизмы посттрансляционного расщепления имеют место также при образовании таких разных белков, как инсулин, коллаген и специфичные белки вирусов. Кроме того, высокоспецифичное протеолитическое расщепление ферментов важно при инактивации и активации специфических внутриклеточных ферментов (табл. 4.2).

Таблица 4.2

Аспекты специфического протеолиза

Биологический процесс	Пример ограниченного протеолиза	Функциональные аспекты
Образование вирусспецифичных белков в вирусах животных [163, 164]	Предшественник полиовирусных белков → полиовирусные белки	Адаптация вирусов к особенностям синтеза белка в клетках животного
Морфогенез вирусных частиц [140]	Две независимые субструктуры вируса Т4 → → образование из двух субструктур	Расщепление полипептидной цепи создает новые центры связывания; необратимость протеолитических реакций способствует упорядочению образований из вирусных частиц
Образование соединительной ткани [125; 127, 131]	Проколлаген с присоединенными фрагментами на N- и С-концах → тройная спираль коллагена + крупные полипептиды	Отщепление фрагментов от коллагена после его обработки и транспортировки в виде растворимого предшественника
Пищеварение [138]	Трипсиноген → трипсин+ небольшой пептид	Активация протеазы в соответствующее время и в физиологически обусловленном месте
Коагуляция крови [148]	Неактивный фактор коагуляции А → активный фактор коагуляции А+ +пептид; затем А активирует неактивный фактор коагуляции В и т. д.	Защитная триггерная система многих проферментов; усиление отклика с использованием каскада ферментов; контроль на каждом уровне каскада
Комплементное действие [147]	Фактор комплемента → → один или два активных белка комплемента	Защитная триггерная система по меньшей мере из 18 сывороточных белков; серией специфических протеолитических реакций ограничивается действие комплемента во времени и пространстве. Растворимые белки присоединяются к мембране
Действие токсина [165]	Дифтерийный токсин → → ферментативно активный фрагмент +проникающий фрагмент	Протеаза пораженной клетки выделяет летальный фрагмент токсина, обладающий ферментативной активностью

Продолжение табл. 4.2

Биологический процесс	Пример ограниченного протеолиза	Функциональные аспекты
Разложение белка [154—156, 166]	Сериндегидратаза (апофермент) → крупный фрагмент + небольшой пептид	Инактивация и распад белка часто инициируются одной специфической протеолитической стадией
Образование гормона [167]	Проглюкагон → глюкагон + октапептид	Экстрапептид есть С-концевая часть прогормона
	Проинсулин → инсулин + + С-пептид	Расщепление правильным образом свернутого профермента автоматически приводит к функционирующему белку
Выделение белка [153]	Протрипсиноген → трипсиноген + гидрофобный пептид	Удаление сигнальной последовательности, способной передавать информацию в среду

Известен также процесс, противоположный специфическому протеолизу, а именно наращивание белков [141]. Примером присоединения С-конца одного белка к N-концу другого является процесс сборки бактериальных вирусов [142].

Активация трипсиногена

Трипсиноген превращается в трипсин, который в свою очередь активирует многие другие ферменты. Во многих системах специфические протеазы не были обнаружены. Типичным примером достаточно подробно изученного процесса ограниченного протеолиза является активация в пищеварительном тракте зимогенов [1, 138, 139], неактивных предшественников ферментов [143]. Ключевым пищеварительным ферментом считается протеаза трипсин [1], не только ввиду его собственной активности по отношению к перевариваемому белку, но и потому, что он является единственным активатором других зимогенов, в частности химотрипсиногенов, прокарбоксипептидазы, проэластазы и профосфолипазы А. Сам трипсин выделяется в двенадцатиперстной кишке в виде неактивного предшественника трипсиногена (рис. 4.5).

Активатор трипсиногена узнает специфическую последовательность. Физиологический активатор трипсиногена — энтеропептидаза (называемая также энтерокиназой) находится на поверхности мембраны эпителиальных клеток в двенадцатиперстной кишке. Исключительно узкую специфичность энтеропептидазы харак-

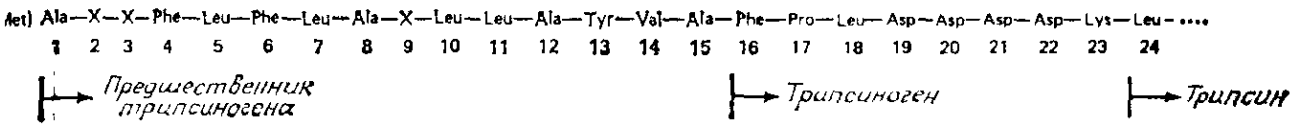


Рис. 4.5.

Аминокислотная последовательность протрипсиногена собаки (предварительные данные [153]).

Предполагается, что концевой аминокислотой всех образующихся полипептидных цепей эукариот служит Met (который затем отделяется). Пока не ясно, какой из остатков протрипсиногена, 16 или 17, отвечает N-концевой аминокислоте трипсиногена собаки. Использованный метод не позволяет различить остатки Leu и Ile (например, в положении 24).

теризует тот факт, что этот фермент не гидролизует ни одной связи ни в одном из исследованных нативных или денатурированных белков, кроме единственной связи в трипсиногене (или в синтетических пептидах), которая находится перед последовательностью Asp-Asp-Asp-Asp-Lys (Нейрат, частное сообщение, а также [138]). Таким образом, этот фермент специфичен для последовательности, но не для третичной структуры.

Если пренебречь поправочными факторами, учитывающими природное содержание отдельных аминокислот [144, 145], последовательность Asp-Asp-Asp-Asp-Lys встречается один раз на 20^5 комбинаций. Это показывает, что уникальная аминокислотная последовательность более надежна в качестве места специфичного опознания белка, чем какая-либо уникальная особенность третичной структуры, а надежность весьма существенна для любого инициатора ферментативного каскада.

Активация трипсиногена характерна для других систем. Система трипсиноген — энтеропептидаза представляет собой, по-видимому, типичную систему специфического протеолиза (табл. 4.2) в следующих аспектах:

а) в большинстве случаев трансформация (*пробелок* → *белок*) может быть осуществлена трипсином in vitro, однако с относительно низкой скоростью и с относительно низкой специфичностью [146]. Следует подчеркнуть, что активный трипсин встречается in vivo только в двенадцатиперстной кишке.

б) Фермент, ответственный за превращение in vivo, более эффективен и более специфичен; каждая из приведенных в табл. 4.2 реакций, по-видимому, требует отдельного фермента. Можно ожидать, что, как и энтеропептидаза, специфичные протеазы могут узнавать не только основной остаток, но более широкое структурное звено (такое, как последовательность Asp-Asp-Asp-Asp-Lys в трипсиногене).

в) Специфичные протеазы часто присоединены к мембранам.

Действие комплемента — результат защитной системы пробелков

«Комплемент» — это часть защитной системы организма, кото-

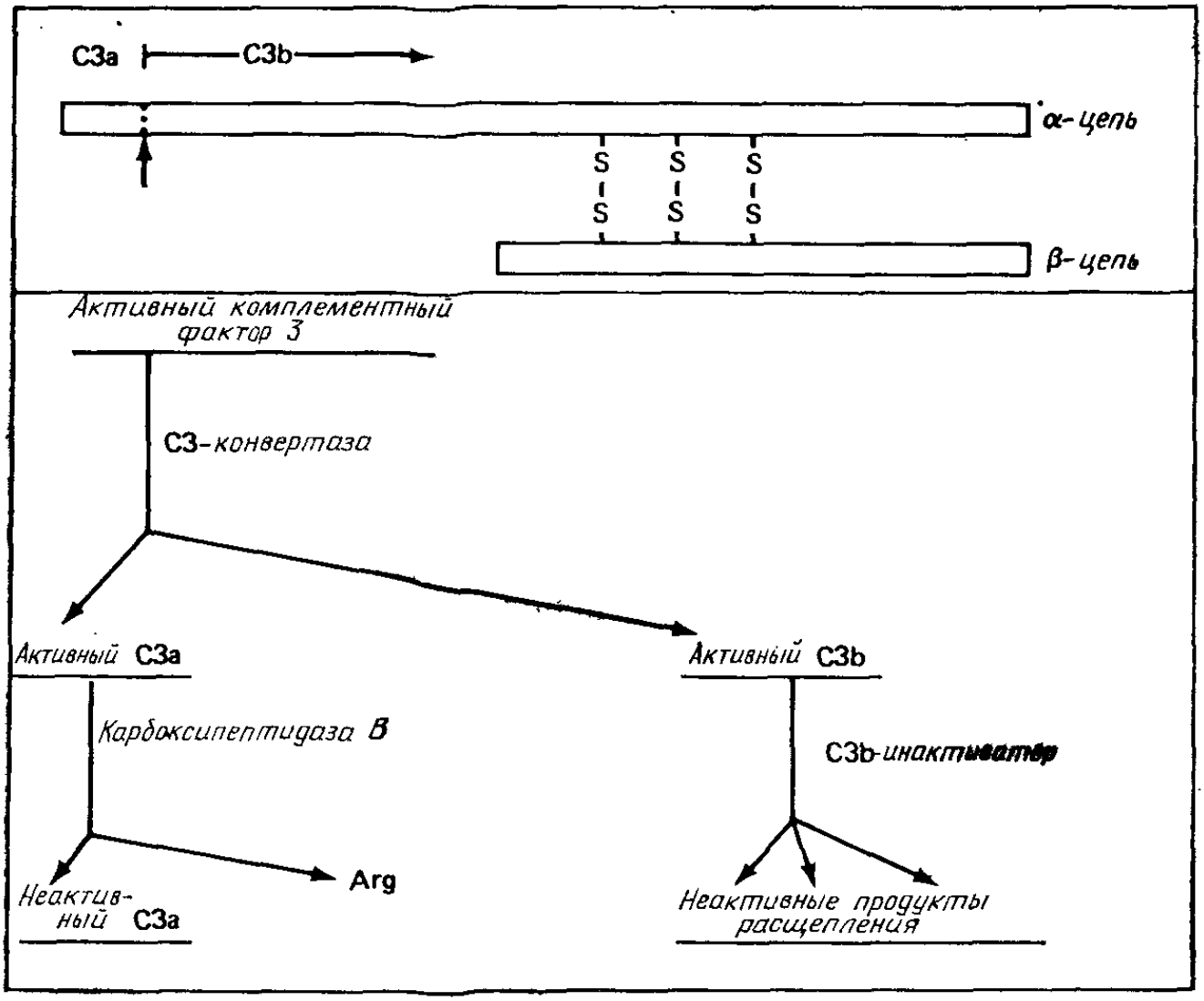


Рис. 4.6.

Многообразные функции белка С3 комплемента и продуктов его расщепления [147].

Точка расщепления С3 конвертазой С3 отмечена стрелкой, направленной вверх. С3 (*M* 180 000) участвует в «пропердиновом (альтернативном) пути», а также является физнологическим предшественником двух активных ферментов, малого С3а и большого С3b. С3а участвует в реакциях возбуждения, а С3b представляет субъединицу или модулятор четырех различных ферментов комплемента. Следует отметить, что в основе ннактивацин С3а и С3b также лежит ограниченный протеолиз.

рая устраняет клетки, например бактерии или раковые клетки. Сам комплемент построен по меньшей мере из 18 различных плазменных белков [147]. В функциональном отношении большинство из этих белков представляет пробелки (зимогены), поскольку они являются предшественниками последовательно или одновременно действующих ферментов, т. е. усилительного ферментативного каскада. Вся система комплемента со всеми ее компонентами в оптимальных концентрациях всегда готова к действиям и может быть запущена в любое время. В соответствии с назначением (осуществлять защиту и запускать многих предшественников активных белков) комплемент включает и систему свертывания крови [148].

Протеолитический запуск имеет физиологические преимущества. На молекулярном уровне белки комплемента демонстрируют многие особенности специфических протеолитических процессов [138].

а) Подобно отдельным протеолитическим реакциям, протекающим количественно и необратимо, весь процесс комплемента может идти только в одном направлении.

б) Активация и инактивация белков, т. е. быстрые изменения активированных центров присоединения, накладывают жесткие ограничения во времени и пространстве; поэтому действие комплемента ограничивается ближайшим окружением иммуноглобулинов, которые прикрепляются к инородным клеткам.

в) Как данный белок, так и продукты его расщепления могут выполнять физиологические функции; специфический протеолиз все эти функции подавляет. Это видно на примере фактора С3 комплемента (рис. 4.6).

г) Действие комплемента включает морфогенетический аспект. Серия специфических протеолитических реакций приводит к агрегированию, в результате которого пять различных растворимых белков присоединяются к мембране [147].

Сигнальные последовательности в пробелках

Наращивание с N-конца цепи, по-видимому, используется в качестве метки. Многие только что синтезированные белки выводятся посредством секреторного аппарата клетки; другие белки действуют внутри клетки: в цитозоле, в мембранах или в окружающей их среде. Назначение полипептида может быть закодировано в его последовательности, по крайней мере это представляется вероятным для выводящихся из клетки белков. Согласно «сигнальной гипотезе» [149], секреторный белок синтезируется в виде «пробелка», который представляет собой предшественник с короткоживущим фрагментом, присоединенным к N-концу.

Удлиненные, так называемые сигнальные пептиды, содержащие около 15 преимущественно гидрофобных аминокислотных остатков, обнаружены в предшественниках некоторых легких цепей иммуноглобулина [134, 150, 151], в пропаратироидальном гормоне [152] и различных белках поджелудочной железы [153] (рис. 4.5). Поскольку сигнальный пептид появляется от рибосомы, можно предполагать, что он прикреплен к микросомной мембране; это первый шаг в последовательности стадий, приводящих к переносу возникающей цепи через мембрану. В процессе или непосредственно после переноса сигнальная последовательность отделяется с помощью присоединенной к мембране протеазы.

Расщепление белка

Расщепление так же важно, как и синтез. Восприимчивость к инактивации и расщеплению — важное биологическое свойство всех белковых структур. Полагают [154—157], что первый шаг в

расщеплении белка состоит в специфическом разрыве одной или нескольких пептидных связей. Затем следует обширный неспецифический протеолиз, обычное завершение распада всех белков [156]. Расщепление необходимо для удаления дефектных или не нужных в данный момент белков. Другой аспект — обычный белковый обмен, наиболее существенный для пищеварительных ферментов. Если бы эти белки не расщеплялись до аминокислот, т. е. если бы протеазы, например трипсин и химотрипсин, не действовали как «молекулярные каннибалы», человек терял бы 100 г аминокислот в день в виде пищеварительных ферментов. Это количество соответствует содержанию белка в суточном рационе человека.

4.4 Контролируемые ферментами модификации боковых цепей

Поперечные связи на основе модифицированных остатков Lys

Образование ε-(γ-глутамил)лизиновых звеньев* существенно, например, при свертывании спермы грызунов в процессе ее ферментативной коагуляции [158] или крови при стабилизации хрупких фибриновых агрегатов [159]. Поперечные связи между полипептидными цепями в эластине (рис. 4.7) и в коллагене (рис. 4.4) также образованы боковыми цепями лизиновых остатков. В эластине, резиноподобном белке, поперечные связи, по-видимому, важны для обеспечения возврата волокон эластина в их исходное состояние после растяжения. Поперечные связи между молекулами коллагена способствуют устойчивости тройной спирали; эта структура термолабильна и расплавляется в желатин приблизительно при 40°C [125, 127]. Таким образом, внутри- и межмолекулярные поперечные связи между полипептидными цепями в коллагене предохраняют этот наиболее часто встречающийся в организме человека белок от превращения в аморфную желатину при температуре приближающейся к 40°C, когда человек болен. Во всех упомянутых выше белках поперечные ковалентные связи существенны для стаби-

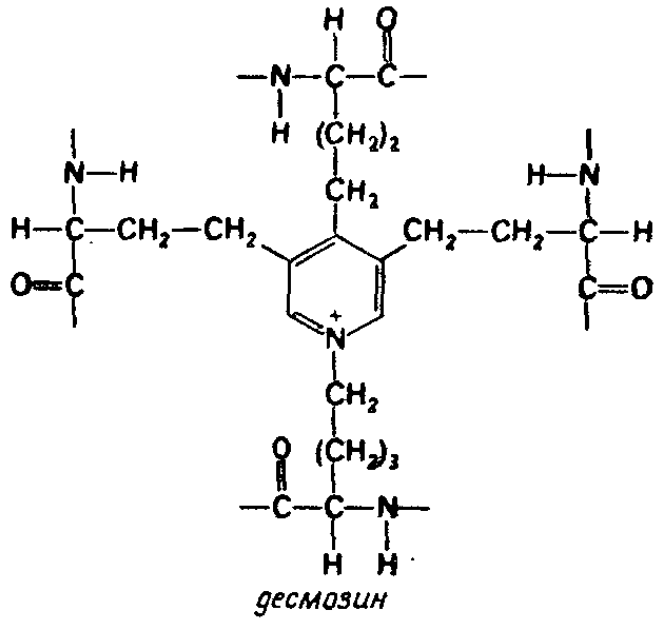


Рис. 4.7.

Десмозин, образованный за счет объединения четырех боковых цепей остатков лизина; встречается в эластине.

* В соответствии с номенклатурой рис. 1.1, ξ-(δ-глютамил)лизин.

Таблица 4.3

Примеры модификаций боковых цепей

Модификация	Модифицируемые боковые цепи	Один конкретный, необязательно типичный пример	Функциональный аспект
Аденилирование [168]	Tyr	Глутаминсинтетаза	Аденилированием одного Tyr на субъединицу фермент переводится из формы, в которой его биосинтетическая активность зависит от Mg^{2+} (оптимально pH 7,6), в форму, зависящую от Mn^{2+} (оптимально pH 6,5)
ADP-рибоксилирование [169]	Lys или Arg	Фактор элонгации EF2	Эта модификация катализируется дифтерийным ферментом [165], подавляет биосинтез и приводит к гибели клетки
Карбоксилирование [170, 171]	Glu	Протромбин и другие кровосвертывающие факторы (VII, IX, X)	Модифицированные остатки глутаминовой кислоты необходимы для нормального связывания Ca^{2+} протромбином, т. е. для осуществляемого через Ca^{2+} присоединения протромбина к фосфолипидным поверхностям
Глюкозилирование [172, 173]	Asn, Thr, Ser	Коэролоплазмин	Потеря концевого звена карбогидрата в боковых цепях сывороточных гликопротеинов — сигнал к поглощению и разложению этих белков в печени
Гидроксилирование [125, 127]	Pro, Lys	Коллаген	Ухудшение гидроксилирования препятствует созреванию коллагена; в этом состоит молекулярная сущность цинги
Метилирование [174]	Asp, Gln, His, Lys, Arg	ε-N-Метиллизин в ответвлении белка	Несмотря на широкую распространенность этих реакций, примеров, доказывающих прямую взаимосвязь между метилированием белка и определенной функцией, нет
Фосфорилирование [175]	Ser, Thr	Фосфорилаза	Неактивная фосфорилаза *b* превращается в активную фосфорилазу *a* путем фосфорилировании определенного остатка серина каждой субъединицы

лизации, но не для образования белковых структур — особенность, которая свойственна и некоторым дисульфидным мостикам (разд. 4.2).

Другие разновидности реакций модификации

Модификация боковых цепей расширяет спектр белковых свойств. Поскольку модификация боковой цепи наделяет белок уникальными свойствами, по этому поводу можно сделать несколько лишь самых общих замечаний. В большинстве случаев модифицируется только одна или ограниченное число боковых цепей белка. Эта специфичность определяется структурой белка и модифицирующего фермента, а также природой боковой цепи. Примеры модификаций боковых цепей и некоторые функциональные аспекты даны в табл. 4.3. Стреер [85] составил рисунки, иллюстрирующие химизм этих реакций.

Относительно модифицирующей группы можно сделать следующие замечания:

а) Она наделяет боковую цепь функцией, которая не может быть выполнена ни одной из канонических аминокислот.

б) Она не меняет основных свойств белка; присоединяющаяся группа имеет молекулярную массу клеточного метаболита, т. е. меньше 1000.

в) Модифицирующую группу можно отделить от белка путем расщепления пептидной связи. Это свойство позволяет отличить модификацию боковой цепи от «простетической группы», которая может ковалентно присоединяться к боковой цепи, но которая прочно связана невалентными силами также и с белком. Примерами простетических групп, ковалентно присоединяющихся к некоторым белкам, могут служить FAD [160] и группа гема. Для этих белков они, конечно, являются в некоторых отношениях модификациями боковых цепей. Следует отметить, что любая биологическая классификация страдает некоторой неопределенностью.

Заключение

Для получения более ясной картины обычно стремятся к упрощениям, что и было сделано в данной главе. Необходимо, однако, представлять себе истинную сложность проблемы. Тот факт, что полипептидная цепь, построенная из 20 аминокислот, самопроизвольно свертывается, а затем функционирует как глобулярный белок, далеко не исчерпывает всей сложности этого вопроса. Обычно, например, цепи объединяются и функционируют в виде агрегатов. Изученные свойства таких агрегатов заставляют предполагать, что основными структурными единицами белков служат не отдельные цепи, а отдельные домены. В белках с неизвестной структурой

можно выявить только функциональные, но не структурные домены. Вместе с тем по данным известных белковых структур можно заключить, что структурный домен — более основополагающее понятие. Известные примеры белковых агрегатов настолько сложны, что даже выделение мономеров и олигомеров представляет проблему. Мы используем физиологические критерии, согласно которым мономеры могут содержать более чем одну полипептидную цепь, если полипептидные цепи ковалентно связаны.

Достаточно развернутое описание белков в предыдущих главах не включает поперечных связей и эпигенетических модификаций. Наиболее обычной поперечной связью является дисульфидный мостик, который служит как механическим, так и химическим целям. Механически важные поперечные связи часто образуются с использованием ε-аминогруппы Lys. В процессе эпигенетических модификаций главная цепь часто расщепляется. Это очень важный физиологический инструмент, поставляющий необходимый белок в нужное место и в нужное время. Распространены также модификации боковых цепей, которые наделяют ферменты новыми свойствами. Обо всех этих явлениях следует помнить при попытках вывести конкретные заключения из довольно общих принципов, изложенных в начале этой главы.

Глава 5 Способы свертывания и ассоциации полипептидных цепей

Шесть уровней структурной организации белков. Согласно Линдерстрем — Лангу [176], можно выделить четыре уровня структурной организации белков: первичный, вторичный, третичный и четвертичный. Эти термины означают соответственно аминокислотную последовательность, упорядоченное строение основной цепи полипептида, трехмерную структуру глобулярного белка и структуры белковых агрегатов. На основании имеющихся у нас теперь знаний мы можем добавить еще два уровня: сверхвторичные структуры для обозначения энергетически предпочтительных агрегатов вторичной структуры и домены для обозначения тех частей, которые представляют собой достаточно обособленные глобулярные области. Схема структурной организации приведена на рис. 5.1.

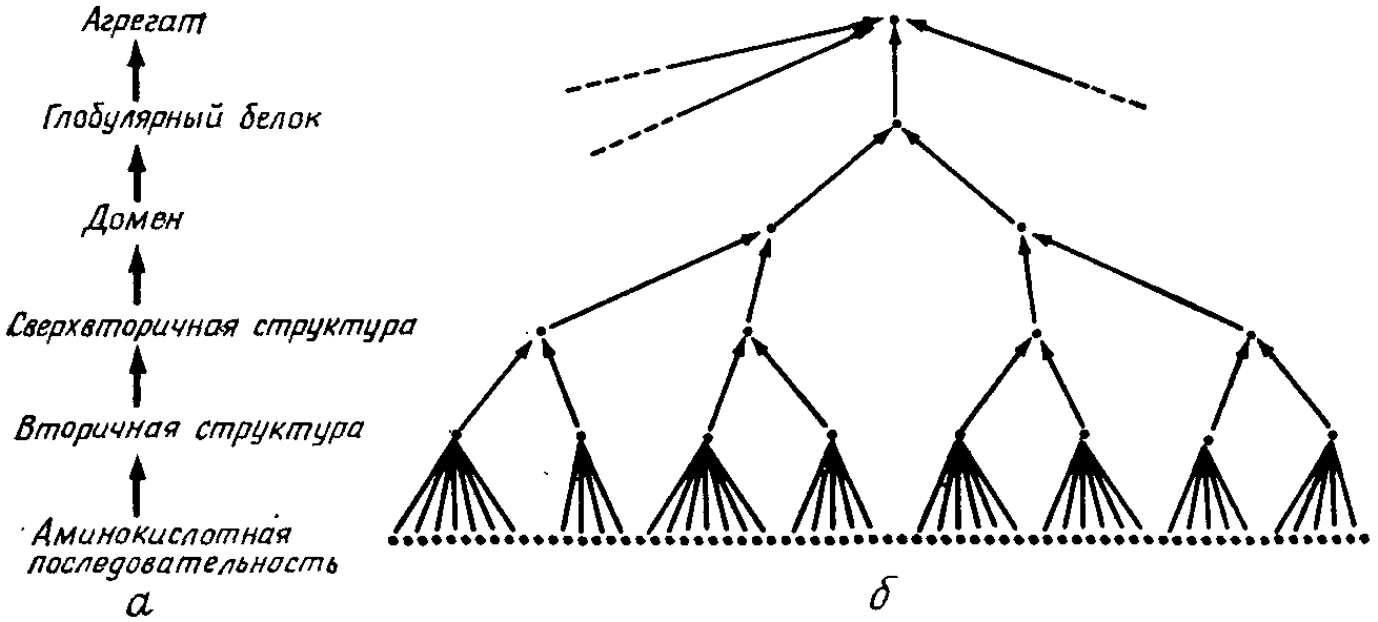


Рис. 5.1.
Уровни структурной организации глобулярных белков.

Эксперименты по ренатурации белков показали, что в аминокислотной последовательности заключена полная структурная информация [177]. Поэтому между всеми уровнями имеется взаимосвязь, так что элементы более низких уровней определяют элементы более высоких уровней.

Эта взаимосвязь показана в виде предельно упрощенной схемы на рис. 5.1, *б*, построенной в предположении, что взаимодействия всех видов четко разграничены, нет взаимодействий между несоседними в полипептидной цепи элементами или между элементами разных уровней. Такое построение в какой-то мере сходно с иерархической социальной системой. В этой главе будут рассмотрены все уровни структурной организации белков, а также обсуждена взаимосвязь между ними.

5.1 Вторичная структура

Вторичные структуры — это упорядоченные расположения основной цепи полипептида, безотносительно к типам боковых цепей и их конформациям. Все вторичные структуры стабилизированы водородными связями между пептидными амидными и карбонильными группами.

Линейные группы

Основная цепь полипептида образует линейную структуру, если ее двугранные углы повторяются. Любая линейная группа представляет собой спираль. Если упорядоченность такова, что все последовательно расположенные пептидные звенья имеют идентичные взаимные ориентации, т. е. если все углы (ϕ, ψ) одинаковы (рис. 2.2), основная цепь полипептида образует линейную группу. Любая линейная группа есть спираль, которая может быть описана смещением вдоль винтовой оси d, приходящимся на элемент, числом n элементов на виток и расстоянием r от данной точки заданного элемента (здесь — C_α-атома) до оси спирали (рис. 5.2). Поскольку

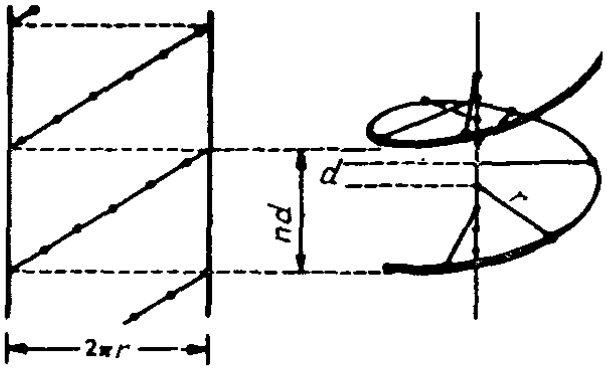


Рис. 5.2.

Спираль, параметры спирали и развертка поверхности спирали.

n — число остатков на виток, d — аксиальное смещение на остаток, nd — шаг спирали, r — радиус спирали. Для получения развертки поверхности спираль проецируют на коаксиально расположенный цилиндрический лист бумаги, этот лист разрезают параллельно оси спирали и затем разворачивают. Развертку поверхности называют также цилиндрической диаграммой. Она делает наглядными геометрические соотношения между остатками. Изменение лицевой стороны обрезанного листа ведет к зеркальному изображению той же цилиндрической диаграммы.

величину d всегда считают положительной, хиральность спирали определяется знаком n. Каждая спираль имеет направленность, так как направлено образованное из N- и C-концов аминокислот пептидное звено.

На рис. 5.3 приведены кривые зависимости между углами (ϕ, ψ) (рис. 2.2) и параметрами спирали d и n. Спираль с $|n|<2$ невозможна. Имеется всего лишь несколько линейных групп, в которых нет стерических затруднений и которые стабилизированы водородными связями либо в пределах цепи (например, α-спирали), либо между соседними цепями (например, β-складчатый лист, коллаген). Параметры наиболее важных линейных групп даны в табл. 5.1.

Линейные группы обнаружены также в сахарах, нуклеиновых кислотах и белковых агрегатах. Линейные группы были обнаружены в полисахаридах, элементом которых являются моно- или олигосахариды с боковыми цепями или без них [178], а также в

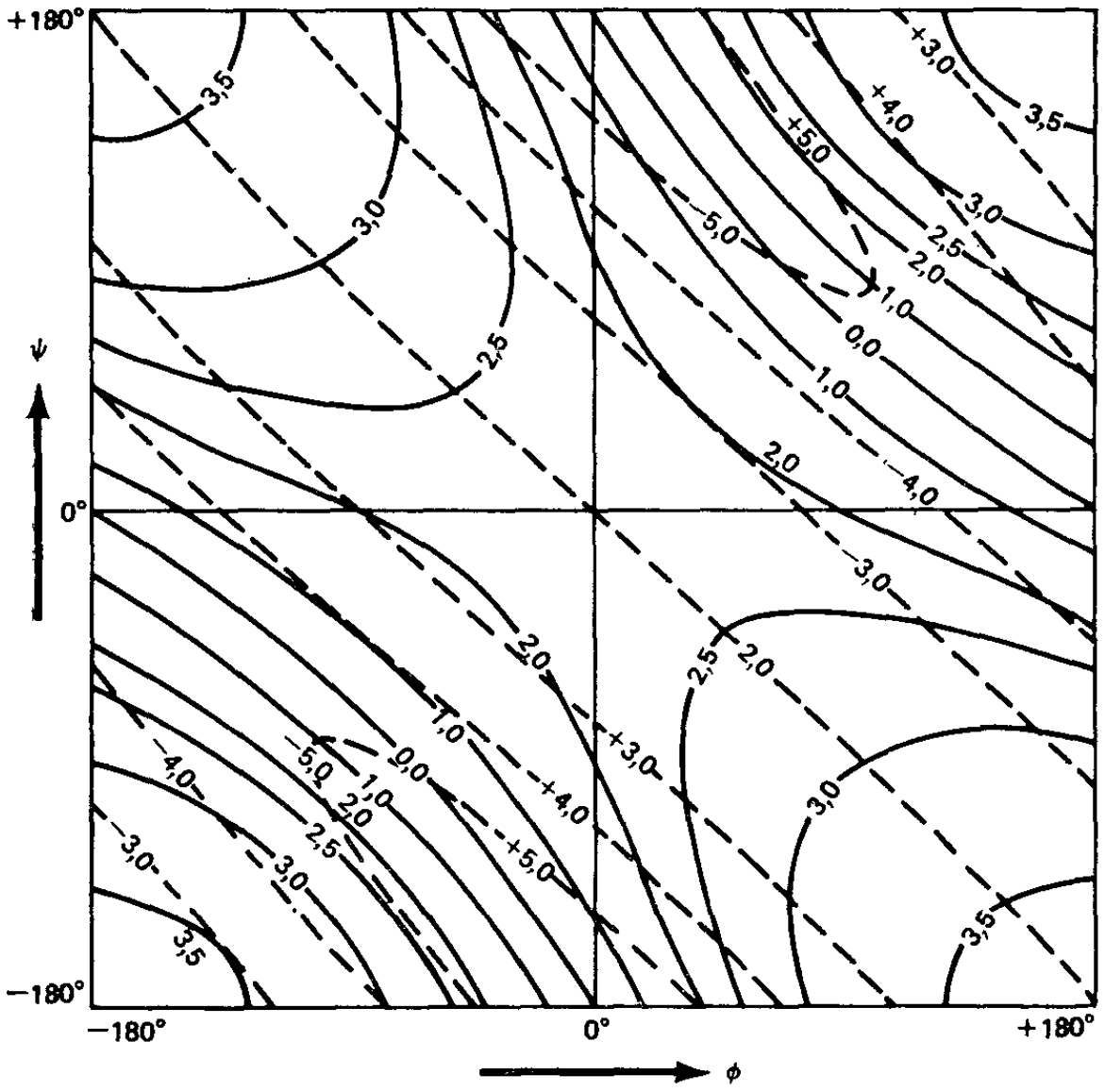


Рис. 5.3.

Параметры спирали d (сплошные линии) и n (штриховые линии) как функции двугранных углов ϕ и ψ [29].

У одной спирали значения ϕ и ψ всех остатков одинаковы.

Таблица 5.1

Линейные группы, образуемые полипептидными цепями

Линейная группа	Наблюдается	Число остатков на виток n и хиральность[a]	Смещение на остаток d, Å	Радиус спирали r, Å
Плоская параллельная β-структура	Редко	±2,0	3,2	1,1
Плоская аитипараллельиая β-структура	»	±2,0	3,4	0,9
Скручеииая параллельная или антипараллельная β-структура	Часто	—2,3	3,3	1,0
3_{10}-Спираль	Небольшие участки	+3,0	2,0	1,9
α-Спираль (правая)	Часто	+3,6	1,5	2,3
α_L-Спираль (левая)	Гипотетически	—3,6	1,5	2,3
π-Спираль	»	+4,3	1,1	2,8
Спираль коллагеиа	В волокиах	—3,3	2,9	1,6

[a] Плюс и минус означают правую и левую спирали соответственио.

РНК (ДНК), где рибозафосфатдиэфирные элементы обычно образуют в результате переплетения двух антипараллельных отдельных спиралей двойную спираль [179]. Вирус табачной мозаики состоит из белковой оболочки и РНК. Оболочка представляет собой линейную группу, элементом которой служит один белок, присоединенный к тринуклеотиду [180]. Цилиндрические конструкции из белков, найденные, например, в микротубулах [181], в хвостах фага Т4 [182] или в филаментах F-актина [183], часто описывают как спирали или линейные группы. Тем не менее следует помнить, что подобные агрегаты не содержат линейных цепей, таких, как, например, РНК, которая соединяет белковые субъединицы в вирусе табачной мозаики и образует особую спираль. Для описания таких моно- и полиспиральных образований используются различные способы.

3_{10}-, α- и π-Спирали

Правильные спирали полностью описываются парой двугранных углов. Строение этих спиралей показано на рис. 5.4, а их параметры приведены в табл. 5.1. 3_{10}-, α- и π-Спирали стабилизированы водородными связями между пептидными амидной и карбо-

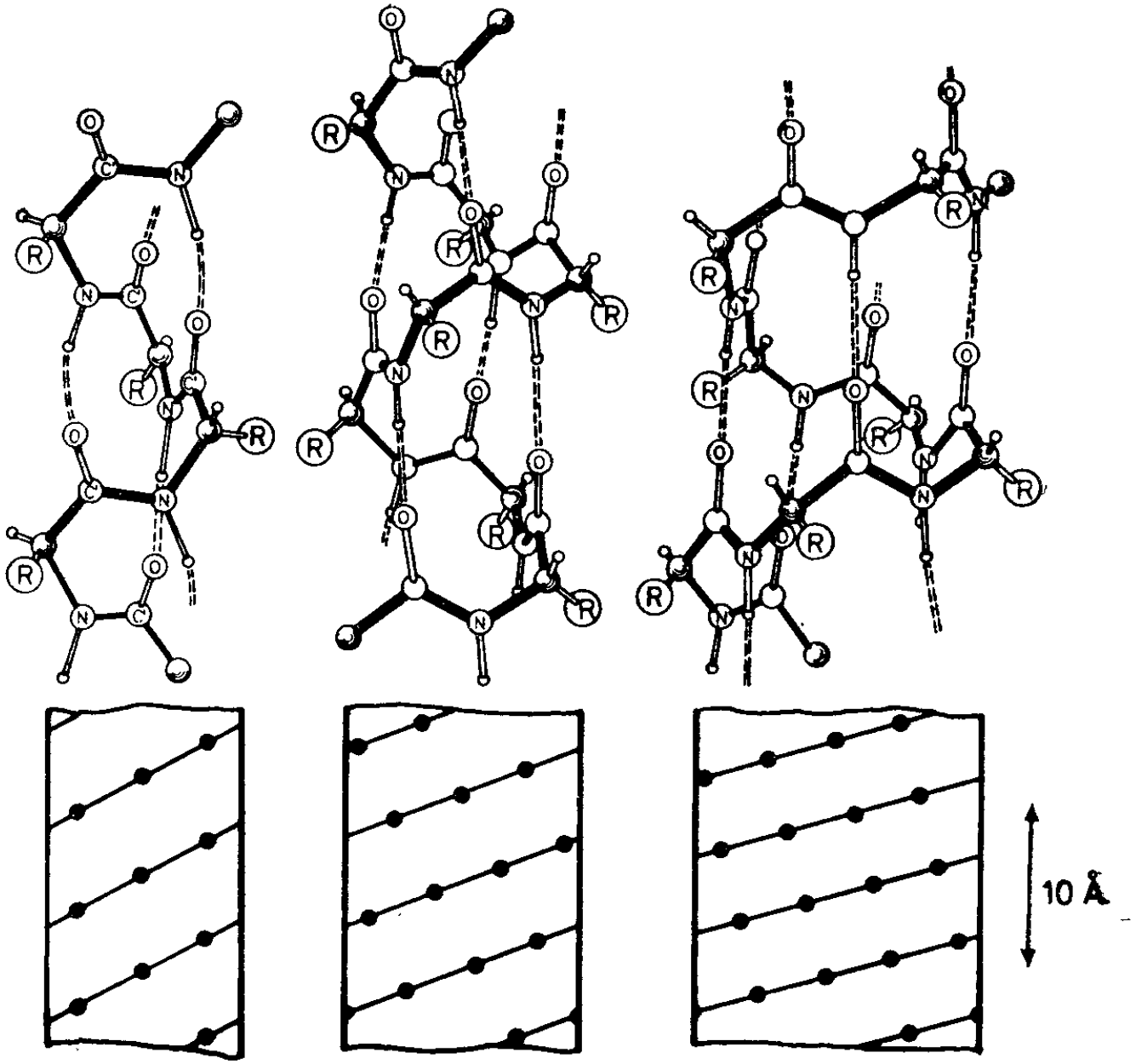


Рис. 5.4.

Спирали полипептидных цепей с внутренними водородными связями.
Слева направо: 3_{10}-, **α- и π-спираль. Внизу: цилиндрические диаграммы этих спиралей; показан ход цепей и положение** C_α **-атомов.**

нильной группами остатков $(i, i+3)$, $(i, i+4)$ и $(i, i+5)$ соответственно. Следовательно, имеется возможность образовывать спирали с системами последовательных водородных связей между близкоотстоящими элементами цепи. Являясь линейными группами, спирали полностью описываются одной точкой на карте $\phi\psi$ (рис. 2.3). Эти точки отвечают локальным энергетическим минимумам; возможны небольшие отклонения от этих минимумов.

α-Спираль очень стабильна и потому встречается наиболее часто. α-Спираль — наиболее часто встречающаяся в белках вторичная структура, в связи с чем ее конформация должна быть весьма стабильной. С этим согласуется расположение α-спирали в центре разрешенной области на карте $\phi\psi$ (рис. 2.3), а также тот факт, что диполи ее водородных связей имеют линейное расположение, также отвечающее минимуму энергии (разд. 3.4). Кроме того, радиус спи-

рали (табл. 5.1) благоприятствует дисперсионному притяжению между остатками, расположенными по разные стороны от оси спирали.

α-Спираль была впервые постулирована Полингом и сотр. [25]. В то время идея о спирали с нецелочисленным значением n была встречена скептически, поскольку она не согласовывалась с предположением о высокой степени упорядоченности молекул. Непосредственное подтверждение постулата было получено из данных рентгеноструктурного анализа кристаллического гемоглобина [183]. Позднее α-спирали были обнаружены во многих волокнах [184], в частности в α-кератине и парамиозине и почти во всех глобулярных белках. В самом деле, обозначения α и β для α-спирали и β-структуры появились в связи с тем, что все рентгенограммы волокон подразделяли на α- и β-рентгенограммы, которые получали в основном от образцов α-кератина и β-фиброина шелка.

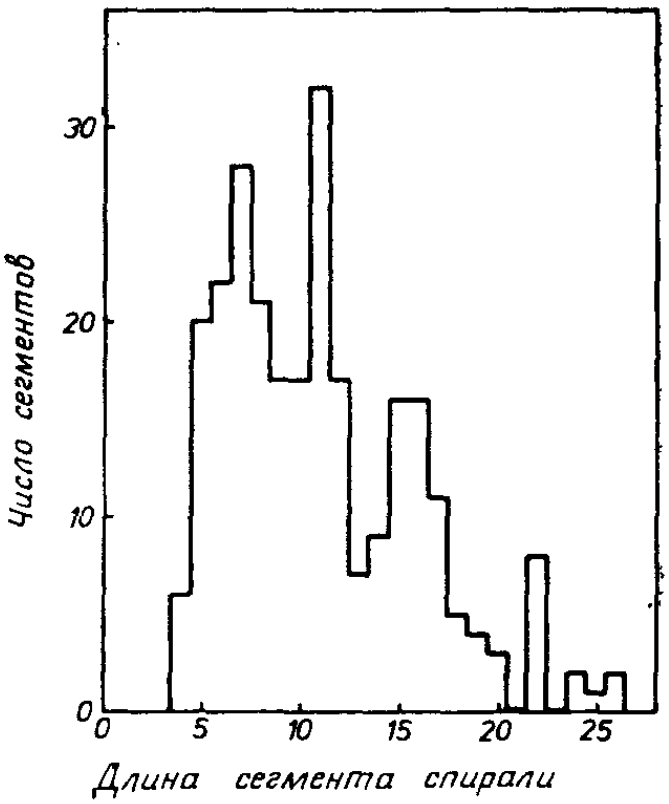


Рис. 5.5.

Распределение длин α-спиральных сегментов для 23 белков (данные взяты из работы [323]).

Длины выражены в числе остатков.

На рис. 5.5 показано распределение наблюдаемых длин α-спиралей в глобулярных белках. Средняя длина составляет около 17 Å, что отвечает 11 остаткам на 3 витка спирали. Значительный разброс в распределении показывает, что предпочтительных длин не существует, что согласуется с ожидаемыми свойствами линейной группы. Относительные максимумы для длин из 7, 11 и 15 остатков, отвечающие 2, 3 и 4 полным виткам, меньше ожидаемой ошибки распределения Пуассона ($\approx 25^{1/2} = 5$), в связи с чем их значимость довольно сомнительна. α-Спирали обеспечивают белковые структуры довольно прочными «стержнями». Изгибы этих стержней на углы около 20° часто осуществляются путем включения остатков Pro, как это можно видеть в миоглобине [185] и аденилаткиназе [186].

3_{10}-Спирали встречаются редко и образуют только очень короткие участки. 3_{10}-Спираль получила название по числу остатков на виток $n = 3$ и по числу атомов (10) в цикле, образованном водородной связью [187]. Диполи, образующие водородные связи,

не находятся на одной линии, что не отвечает минимуму энергии (разд. 3.4). Упаковка боковых цепей в этом случае также довольно невыгодна. Как видно из рис. 5.4, боковые цепи имеют одинаковое азимутальное положение, тогда как в α-спирали они находятся в скрещенном положении. На карте (ϕ, ψ) (рис. 5.7) 3_{10}-спираль попадает на край разрешенной области, что свидетельствует о наличии небольших стерических трудностей. Энергетически невыгодное пространственное строение объясняет малую распространенность 3_{10}-спиралей в белках; обычно ее участки образуют примерно один виток (в гемоглобине морской миноги имеются две 3_{10}-спирали, каждая длиной по два витка [188, 189]). Эти участки обычно находятся при N- и C-концах α-спиралей.

π-Спираль невыгодна. π-Спираль ни разу не удалось наблюдать на опыте, так что она представляет пока чисто академический интерес. Эта спираль находится на краю разрешенной области (рис. 2.3), что указывает на наличие в ней стерических трудностей [29, 190]. Диполи, образующие водородные связи, выстроены благоприятно. Вследствие большого радиуса r (табл. 5.1), атомы основной цепи, находящиеся по разные стороны от оси спирали, уже не образуют стабилизирующих контактов друг с другом. Возникает аксиальная полость*, заметно понижающая энергию вандерваальсова притяжения. Как видно из цилиндрической функции на рис. 5.4, боковые цепи π-спирали расположены не так выгодно, как в α-спирали, но значительно благоприятнее, чем в 3_{10}-спирали. Отсутствие π-спиралей в реальных системах, по-видимому, следует объяснять стерическими трудностями, а также ослаблением вандерваальсова притяжения противостоящих элементов основной цепи.

Левые 3_{10}-, α- и π-спирали не были обнаружены. Если рассматривать только основную цепь полипептида, каждая спираль должна иметь энергетически эквивалентное зеркальное отображение. Однако взаимодействия боковых цепей приводят к энергетической невыгодности левой α_L-спирали, что и объясняет ее отсутствие в глобулярных белках. Это же относится и к левым 3_{10}- и π-спиралям.

Спираль коллагена

Структура коллагена представляет собой суперспираль, образованную тремя параллельными, вытянутыми левыми спиралями. Спираль коллагена показана на рис. 5.6. Это правая спираль, состоящая из трех одиночных левых спиралей, расположенных па-

---

* Поскольку полость не может быть заполнена водой, связывающая энергия понижена на $p\Delta V$ (p — давление, ΔV — свободный объем полости). Однако при атмосферном давлении это понижение составляет всего 0,001 ккал/моль для π-спирали из 50 остатков, т. е. пренебрежимо мало.

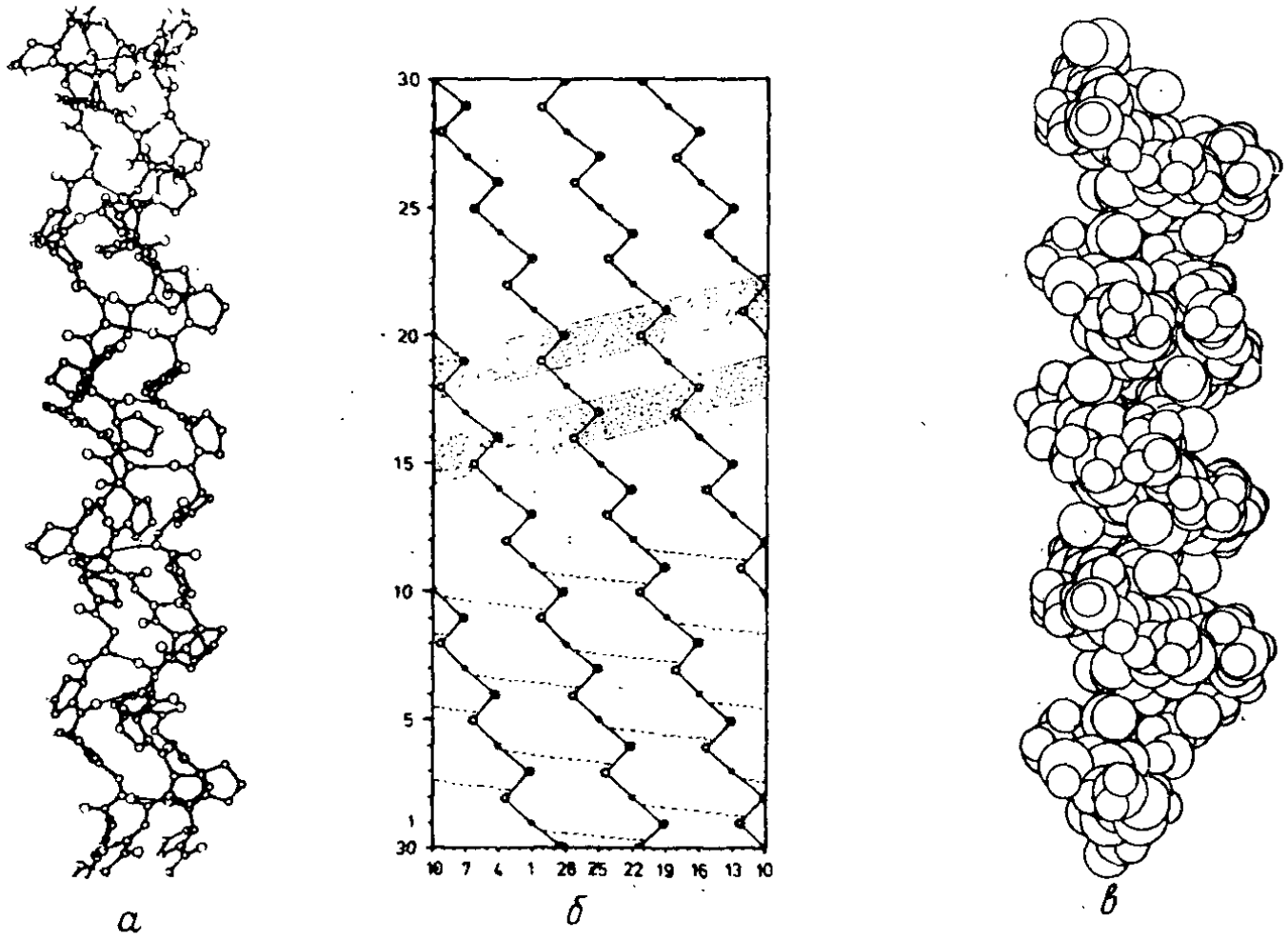


Рис. 5.6.

Структура коллагена.

а — рисунок ORTEP [324] тройной спирали коллагена с последовательностью $(Gly\text{-}Pro\text{-}Pro)_n$ и координатами из работы [193]. Приведены только те атомы водорода, которые участвуют в связях (тонкие, почти горизонтальные линии). *б* — цилиндрическая диаграмма полного периода идентичности тройной спирали коллагена [325]. Обозначены ход спирали и C_α-атомы. Три C_α-атома в последовательности $(Gly\text{-}X\text{-}Y)_n$ отмечены точками: (• O ●). Водородные связи указаны пунктиром. *в* — рисунок ORTEP тройной спирали коллагена, соответствующий *а*, но атомы заданы их вандерваальсовыми радиусами и включены все атомы водорода. Заметны спирально расположенные с шагом 9 Å выступающие остатки Pro. На рис. *б* они показаны затененными полосами.

раллельно. Такое образование называют суперспиралью. Рентгеноструктурный анализ коллагена представляет значительные трудности [191—194]: наиболее важная информация для расшифровки рентгенограммы коллагена может быть получена из сходных с ней картин рентгеновской дифракции синтетических полимеров — поли-Gly-II и поли-L-Pro-II (поли-L-Pro-I содержит *цис*-пептидные связи). Особенно актуально построение моделей для Pro-гомополимеров, поскольку у них двугранный угол ϕ фиксирован вблизи —60°, а образование структуры, подобной α-спирали (ψ = —60°), невозможно по стерическим причинам; приемлемыми оказываются лишь значения в области ($\phi\psi$) = (—60°, + 140°) (рис. 2.3), хорошо описывающие строение отдельной спиральной нити коллагена.

Элементом группы суперспирали является триплет. Для детального анализа пространственного строения коллагена решающим

оказалось знание аминокислотной последовательности [195—197]. Около 96% цепи может быть описано формулой $(Gly\text{-}X\text{-}Y)_m$. Положение X часто занимает Pro, а положение Y — Hyp (оксипролин = Pro с OH-группой при C_γ-атоме; см. рис. 1.1). Содержание каждого из остатков Pro и Hyp составляет 11%. До недавнего времени обсуждались три модели, построенные из тройных спиралей, образованных параллельными пептидными цепями [191, 192]. С наиболее часто встречающимся триплетом (Gly-Pro-Hyp) согласуется структура так называемого коллагена-II, все остальные модели в конце концов пришлось исключить [194].

Суперспираль лучше всего описывается с помощью цилиндрической развертки, показанной на рис. 5.6, *б*. Это линейная группа с триплетом (Gly-X-Y) в качестве элемента группы. В одном витке суперспирали на каждую полипептидную цепь приходится 10 триплетов; шаг суперспирали равен 86 Å. Три цепи стабилизированы вандерваальсовыми силами, а также водородными связями (по одной на триплет). Диполи, образующие водородные связи, расположены линейно и приблизительно перпендикулярно к оси большой спирали (рис. 5.6, *б*).

Отдельная полипептидная цепь левой спирали недостаточно упорядочена. Элементом группы левой спирали одиночной пептидной цепи следует считать триплет, поскольку он является элементом группы суперспирали. Идентичны не все, а только каждые третьи пары значений углов (ϕ, ψ). Однако отклонения в углах (ϕ, ψ) для всех остатков невелики и составляют около 10°. Поэтому структуру отдельной цепи можно аппроксимировать спиралью (линейной группой) с одним остатком в качестве элемента группы, которой отвечает одна конформационная точка на карте (ϕ, ψ) (рис. 2.3). Параметры спирали одиночной полипептидной цепи приведены в табл. 5.1. При значении спирального параметра $d = 2{,}9$ Å цепь настолько вытянута, что ее геометрия стабильна только при агрегации с другими цепями.

Остаток Gly в каждом третьем положении необходим для стабильности. Поскольку параметр $n = -3{,}3$ для левой спирали отдельной цепи нецелочислен и правая суперспираль содержит 10 триплетов на виток, то каждый третий C_α-атом оказывается вблизи оси большой спирали. Этим объясняется необходимость именно последовательности $(Gly\text{-}X\text{-}Y)_m$, т. к. компактную, образованную за счет водородных связей структуру можно получить только при отсутствии боковых цепей у остатков вблизи оси; единственным таким остатком является Gly.

Боковые цепи остатков коллагена обращены в сторону окружающей среды. Внешний вид спирали коллагена создают боковые цепи остатков в положениях X и Y. Это особенно заметно на пространственной модели $(Gly\text{-}Pro\text{-}Hyp)_m$. Как видно из рис. 5.6, *в*, боковые

цепи Pro и Hyp образуют расположенные на спиральной цилиндрической поверхности выступы с шагом 9 Å.

Коллаген — наиболее распространенный белок млекопитающих. Коллаген, состоящий из сильно вытянутых полипептидных цепей, может выдерживать значительное механическое напряжение вдоль оси, поскольку оно по направлению совпадает с ориентацией ковалентных связей его пептидного остова. Поэтому структура коллагена хорошо соответствует характеру выполняемых им функций — передавать усилия в сухожилиях, образовывать защитный слой кожи и других органов. Коллаген безусловно является одним из самых распространенных в природе белков.

Тройные спирали коллагена большинства позвоночных состоят из двух α_1-цепей и гомологичной α_2-цепи. Пока известна аминокислотная последовательность только α_1-цепи, которая включает 1052 остатка [195—197]. За исключением 16 N-концевых и 25 C-концевых остатков, состав $(Gly\text{-}X\text{-}Y)_m$ в ней строго соблюдается. Пользуясь этой формулой, можно легко выявить в различных аминокислотных последовательностях коллагеноподобные структуры. В глобулярных белках такие структуры пока еще не были обнаружены. Однако весьма вероятно, что они присутствуют в компоненте C1q системы комплемента человека [198], которая узнает антитела, соединенные с антигенами. Как установлено по электронным микрофотограммам, это белок содержит пучок из 18 параллельных цепей, организованных в шесть коллагеноподобных волокон, которые входят в шесть глобул. Возможно, что в дальнейшем будут выявлены и другие смешанные белки, содержащие коллагеноподобные структуры.

Реверсивные повороты пептидной цепи

Пептидная цепь может образовывать резкие повороты, включающие водородную связь. При исследовании предпочтительных конформаций трех последовательно расположенных пептидных звеньев (от C_α^i до C_α^{i+3}) Венкатачалам [199] обнаружил три структуры с водородной связью между атомами O_i и N_{i+3}. Они были названы реверсивными поворотами цепи I, II и III. Поворот III представляет собой участок описанной выше спирали 3_{10} (рис. 5.4). Поворот I — это деформированная спираль 3_{10}, в которой углы (ϕ, ψ) правильной спирали приобретают несколько различные значения для C_α^{i+1} и C_α^{i+2} (рис. 5.7, *а*). У поворота II плоскость пептидной группы между остатками в положениях $i+1$ и $i+2$ резко меняет свою ориентацию (рис. 5.7, *б*) и углы (ϕ, ψ) при атомах C_α^{i+1} и C_α^{i+2} приобретают существенно отличные значения (рис. 5.7, *а*). Это ведет к серьезным стерическим затруднениям между атомом O_{i+1} и боковой цепью остатка $i+2$; напряжения в цепи в случае

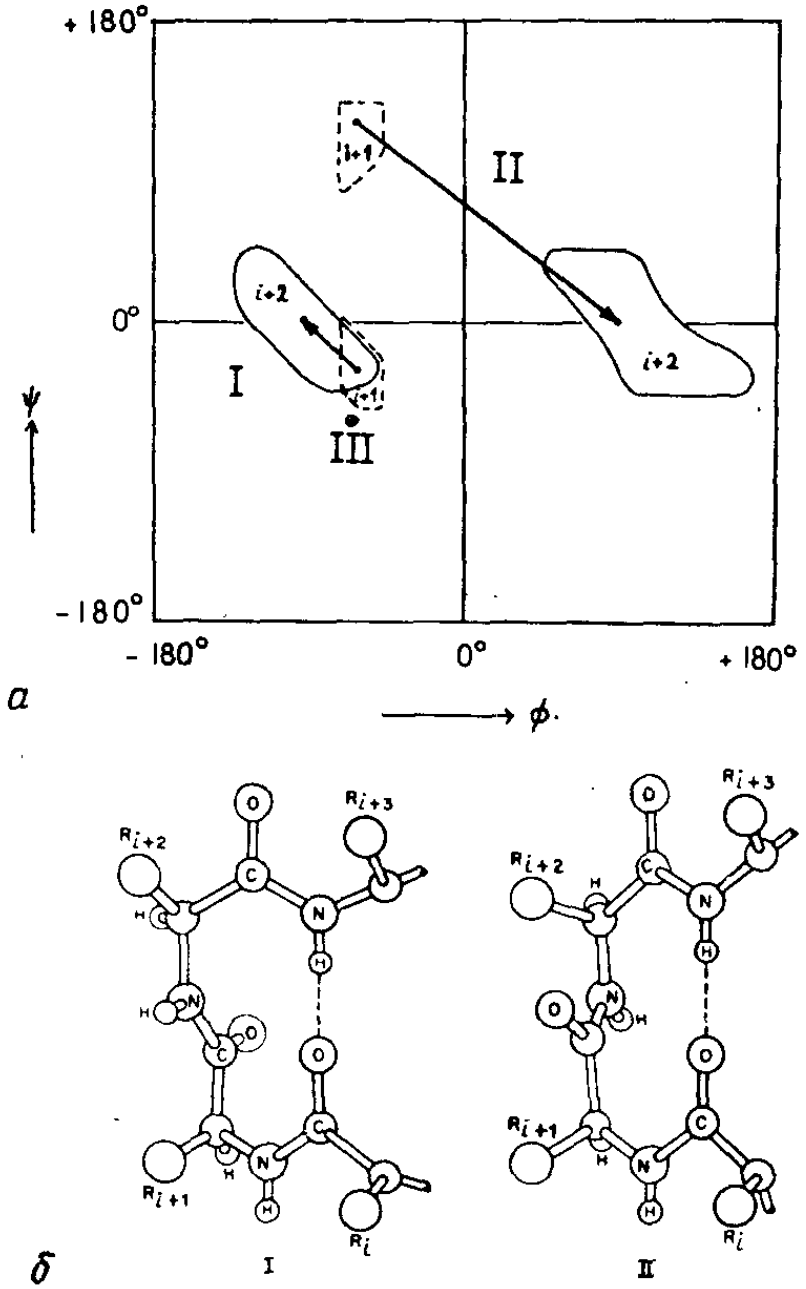


Рис. 5.6.

Реверсивные повороты цепи.

а — области двугранных углов основной цепи остатков $i+1$ (пунктирная линия) и $i+2$ (сплошная линия) в реверсивных поворотах типа I, II и III [199]. Соответствующие области соединены стрелками. В случае поворота III обе области совпадают между собой и с положением 3_{10}-спирали (рис. 5.4). *б* — повороты типов I и II [326]. В типе II имеются стерические затруднения между боковой цепью R_{i+2} и O_{i+1}, так что в положении остатка $i+2$ должен находиться Gly. Поворот типа III см. на рис. 5.4.

поворота II не возникают, если в положении $i+2$ находится остаток Gly. Зеркальные отображения реверсивных поворотов I, II и III, называемые поворотами I′, II′, и III′, невыгодны по стерическим причинам.

Реверсивные повороты цепи часто встречаются в белках. Анализ поворотов цепей в глобулярных белках показал, что они очень распространены и в их образовании участвует около одной четверти всех остатков [200, 201]. Далее, было выяснено, что повороты I, II и III составляют соответственно 35, 15 и 15% от их общего числа [202]. Как и следовало предполагать, почти все найденные повороты II содержат в положении $i+2$ остаток Gly. Общая распростра-

ненность зеркальных отображений (повороты I′, II′ и III′), как и ожидалось, низка и составляет 10%. Нестандартные реверсивные повороты, не содержащие водородных связей, встречаются с частотой 25%. Они идентифицируются с помощью более общего критерия, согласно которому расстояние между атомами C_α^i и C_α^{i+3} должно быть меньше 7 Å при иной, чем α-спираль, конформации основной цепи. Такой критерий включает все стандартные повороты цепи, однако допускает возможность и других реверсивных конформаций. Иногда нестандартные повороты содержат *цис*-Pro; в частности Pro-93 и Pro-114 в рибонуклеазе [39], Pro-168 в субтилизине [40] и Pro-116 в нуклеазе стафилококка [41].

Большинство поворотов находится на поверхности белка. Анализируя расположение поворотов в белках, Кунтц [203] обнаружил, что они сконцентрированы на поверхности. В соответствии с таким расположением они содержат в основном гидрофильные остатки. Предполагается, что в процессе свертывания повороты играют только пассивную роль, образуя участки наименьшего сопротивления невалентным силам, стремящимся изогнуть цепь. Это предположение подтверждается большим разнообразием наблюдаемых поворотов, ни один из которых не имеет особенно стабильной конформации.

β-Складчатый лист

Вытянутые полипептидные цепи могут взаимодействовать между собой посредством водородных связей и образовывать слоистые структуры. Кроме α-спиралей в качестве возможных упорядоченных структур полипептидной цепи, образованных водородными связями, Полинг и Кори [204] постулировали «плоские» параллельный и антипараллельный β-складчатые листы (рис. 5.8). И в том, и в другом типах β-структур цепь образует линейную группу с одним остатком в качестве элемента группы, (спиральные) параметры которой приведены в табл. 5.1. Углы (ϕ, ψ) в обоих случаях находятся в разрешенной области (рис. 2.3), а образуемые водородными связями диполи находятся на одной линии. Расположение водородных связей схематически показано на рис. 5.8, *б* и 5.8, *в*. Если смотреть вдоль полипептидного остова, видно, что боковые цепи ориентированы поочередно то по одну, то по другую стороны средней плоскости складчатого листа, причем связи C_α—C_β приблизительно перпендикулярны плоскости (рис. 5.8, *а*). Возможны смешанные параллельно-антипараллельные слои, для чего требуется некоторое изменение углов (ϕ, ψ).

Складчатые структуры наблюдаются очень часто. Складчатые структуры были обнаружены в β-фиброине шелка и в растянутом волосе [184]. У глобулярных белков эти структуры составляют

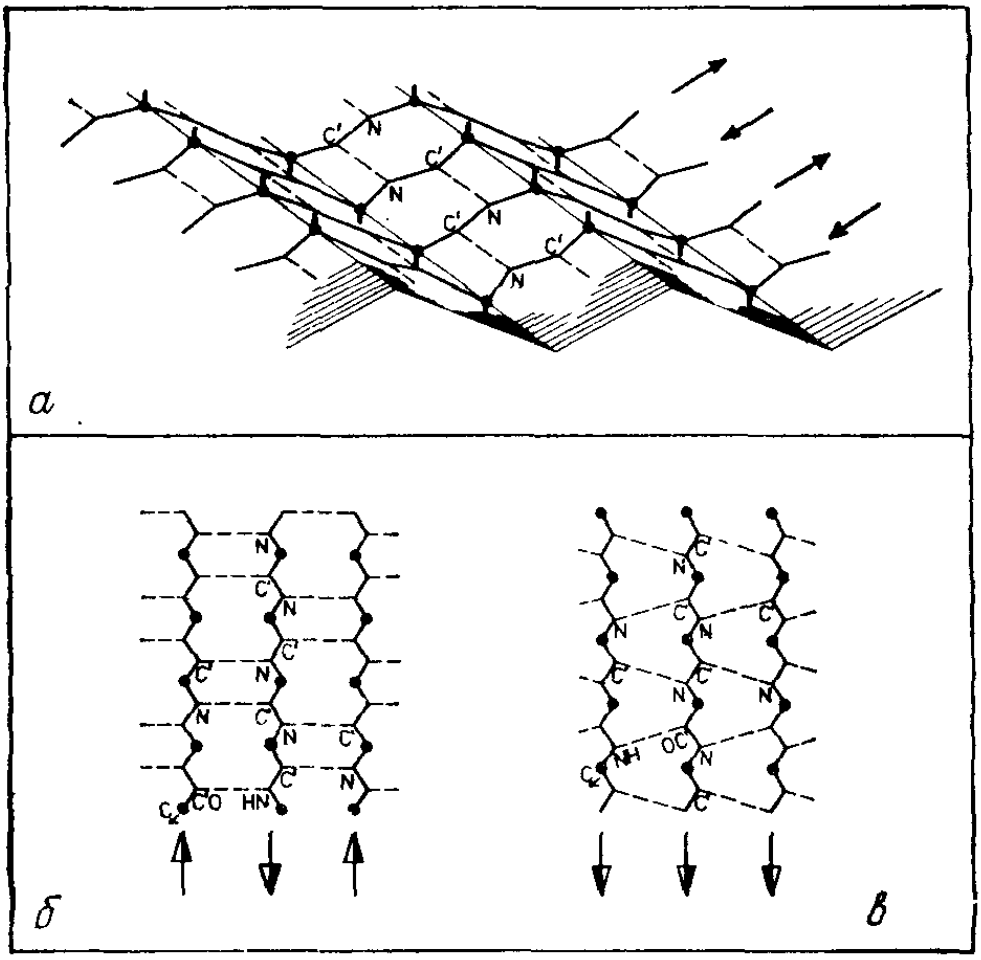


Рис. 5.8.

β-Складчатый лист. Водородные связи указаны пунктирными линиями, а направления цепей — стрелками. C_α-атомы обозначены точками.

а — складчатая структура (нескрученного или плоского) антипараллельного листа. Направления связей $C_\alpha - C_\beta$ показаны короткими линиями при C_α-атомах; они перпендикулярны листу. *б* — расположение водородных связей антипараллельной трехцепочечной β-структуры. *в* — расположение водородных связей параллельной трехцепочечной β-структуры.

примерно 15% [201]. Гистограммы длины и ширины β-структур, встречающихся в глобулярных белках, приведены на рис. 5.9. Средняя длина примерно из шести остатков (около 20 Å) приблизительно отвечает диаметру домена (см. ниже). Как видно из рис. 5.9, *а*, вытянутые цепи, длины которых выражаются нечетным числом остатков, встречаются чаще других. Так происходит потому, что при соответствующей β-структуре системе водородных связей (рис. 5.8, *б* и 5.8, *в*) последние не образуются между концевыми остатками при их четном числе в цепи. Большинство складчатых листов содержит менее шести цепей, причем ни одно из наблюдающихся образований не имеет какого-либо преимущества. Ширина шестицепочечного β-складчатого листа составляет около 25 Å, что соответствует диаметру типичного домена. Не было отмечено предпочтительности ни для параллельной, ни для антипараллельной β-структур, однако параллельные структуры с числом нитей, меньшим четырех, встречаются редко. Складчатые листы с обоими направлениями цепей встречаются часто, но все-таки в меньшей сте-

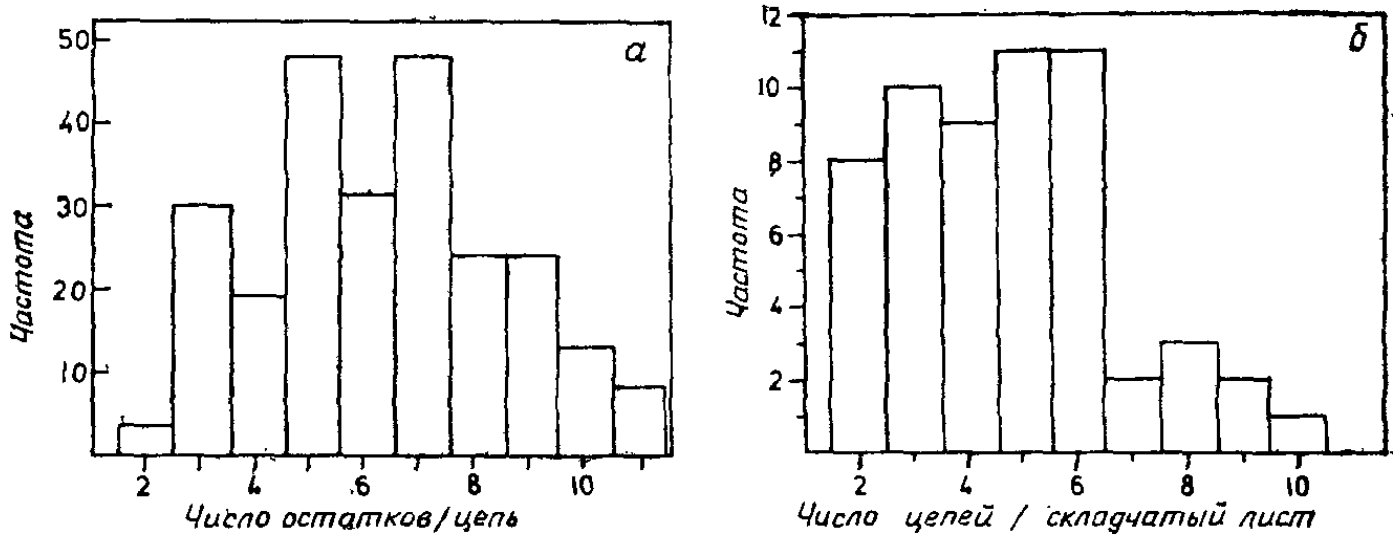


Рис. 5.9.

Наблюдаемые β-структуры [327].

а — распределение по длинам и *б* — распределенне по ширине β-складчатых листов.

пени, чем можно было бы ожидать при случайном выборе направления. Таким образом, цепи в складчатых листах проявляют кооперативные свойства, т. е. обнаруживают тенденцию образовывать либо параллельные, либо антипараллельные β-структуры.

Большинству β-складчатых листов свойственна левая закрутка цепей. При $n = \pm 2{,}0$ (табл. 5.1) пептидные цепи, образующие параллельные и антипараллельные β-структуры, постулированные Полингом и Кори [204], имеют общую среднюю плоскость. Такая «плоская» (антипараллельная) β-структура была найдена, например, в глутатионредуктазе [124]. Однако большинство складчатых листов являются *неплоскими* [43, 205]; они характеризуются левой закруткой, если смотреть вдоль плоскости листа перпендикулярно его вытянутым цепям, как показано на рис. 5.10, *г* (если смотреть по направлению цепей, то скручивание будет считаться правым). Отдельную цепь скрученного листа можно в хорошем приближении описать линейной группой с одним остатком в качестве элемента. Это очень растянутая левая спираль, углы (ϕ, ψ) и спиральные параметры которой приведены соответственно на рис. 2.3 и в табл. 5.1. Как схематично показано на рис. 5.10, *б*, такая левая спираль отвечает правому повороту карбонильной и амидной групп примерно на 60° на два остатка. Поэтому водородные связи между соседними цепями могут образоваться только в том случае, если направления цепей образуют друг с другом угол *около* 25° (рис. 5.10, *в*). Это и приводит к скручиванию слоя. Длина скрученного листа неограниченна. β-фиброин шелка содержит, по-видимому, очень длинные скрученные ленты β-складчатого листа.

Структура широкого (например, шестицепочечного) скрученного листа геометрически соответствует большому желобу двойной спирали ДНК [206], а структура скрученного антипараллельного двухцепочечного листа — малому желобу РНК [207] и ДНК [208].

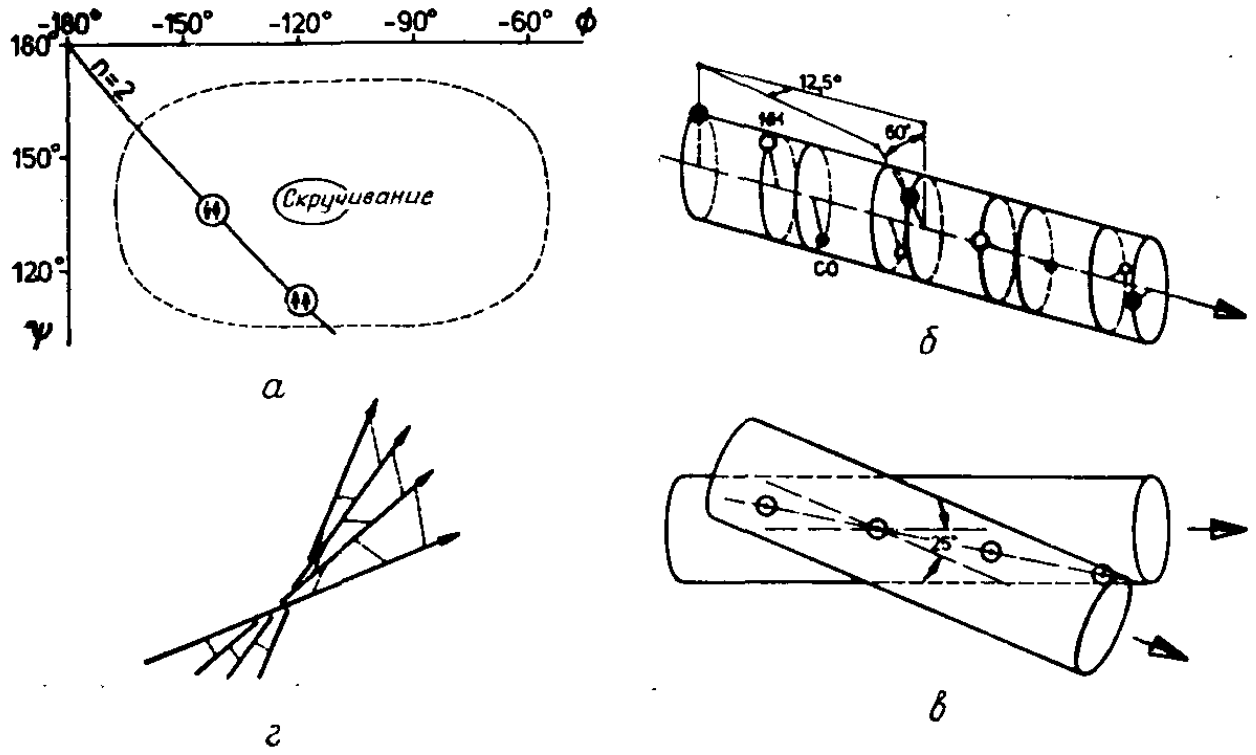


Рис. 5.10.

Скручивание β-складчатых листов.

а — область карты (ϕ, ψ) с двугранными углами основной цепи нескрученных параллельной (↑↑) и антипараллельной (↓↑) β-структур и левой скрученной β-структуры. Пунктирная линия соответствует эквипотенциали 1 ккал/моль (рис. 2.5). *б* — в скрученной β-структуре скручена каждая цепь. Ход основной цепи показан пунктирной линией. Прослежены направления карбонильных (●) и амидных (○) групп. *в* — скрученные, подобно (*б*), две параллельные цепи могут образовывать между собой водородные связи (◎). Для этого угол между цепями должен быть равен 25°. *г* — схема левой скрученной параллельной β-структуры.

Таким образом, скрученные листы могут играть важную роль во взаимодействиях белка с двойной спиралью ДНК (РНК) (разд. 10.6).

Часто скручивание заканчивается так называемым β-выступом в том месте, где в одной из β-цепей находится дополнительный остаток. Такая структурная особенность была обнаружена только в антипараллельных складчатых листах [209].

Скручивание цепей — следствие локальной оптимизации. Преимущество скрученных листов перед «плоскими» легко объяснимо. На карте (ϕ, ψ) (рис. 5.10) все три типа β-структуры расположены в разрешенной области. Как показано на рис. 5.10, *а*, энергия E представляет довольно пологую функцию углов (ϕ, ψ).* Водородные связи между цепями не изменяют заметно вида этой функции, поскольку их реализация возможна при вариации значений ϕ, ψ в широкой области. Поэтому при нормальной температуре статистические веса (пропорциональные $\exp(-E/RT)$, см. Приложение) всех конформационных состояний цепи в этой области сравнимы по абсолютной величине. При отсутствии серьезных ограниче-

* Энергия E относится к $H_{\text{цепь}}$ в вакууме; объяснение дано на рис. 3.3 (см. уравнение (А.18)).

ний со стороны основной цепи конкретные значения углов ϕ, ψ в широкой разрешенной области будут в каждом случае определяться взаимодействиями боковой цепи данного остатка с ближайшим окружением. Количество возможных здесь ситуаций практически бесконечно, и ни одна из них не обладает предпочтительностью. Поэтому распределение соответствующих конформационных точек ϕ, ψ в разрешенной области носит случайный характер; состояние, отвечающее средней скрученной β-структуре, находится в центре разрешенной области. Такое объяснение подтверждается очень широким у β-структур разбросом значений (ϕ, ψ), покрывающим большую часть области, показанной на рис. 5.10, *а*.

5.2 Сверхвторичные структуры

В предыдущих главах были рассмотрены упорядоченные участки основной цепи белка, образующие энергетически предпочтительную вторичную структуру. Следующий, более высокий уровень сложности занимают ансамбли взаимодействующих между собой вторичных структур. Пространственное строение таких агрегатов, так называемую сверхвторичную структуру, предсказать трудно*. Их можно выделить лишь в том случае, если они достаточно часто встречаются в белках. Сам факт возникновения таких структур указывает на то, что они либо предпочтительны с точки зрения кинетики процесса свертывания, либо имеют энергетическую предпочтительность в уже свернутом белке.

Суперспирализация α-спиралей

В своей наиболее упорядоченной форме суперспирализованная α-спираль встречается в фибриллярных белках. Хорошим примером сверхвторичной структуры является суперспирализованная α-спираль, постулированная Криком [210]. В этой структуре две α-спирали скручены друг относительно друга, образуя левую суперспираль с периодом идентичности около 140 Å (рис. 5.11, *а*). Суперспирализованные α-спирали обнаружены в фибриллярных белках α-кератине [211, 212], тропомиозине [213], парамиозине [214] и легкой цепи меромиозина [215]. Короткие участки такой сверхвторичной структуры наблюдались в глобулярных белках, содержащих α-спирали, упакованные приблизительно параллельно или антипараллельно. Наиболее известными примерами упаковки спиралей, приближающейся к линейной, являются гемеритрин [216,

* В более широком смысле определение сверхвторичной структуры должно охватывать и β-складчатые листы, которые представляют собой ансамбли регулярных цепей. Однако обычно к сверхвторичной структуре относятся только системы регулярно расположенных в β-листах пептидных цепей.

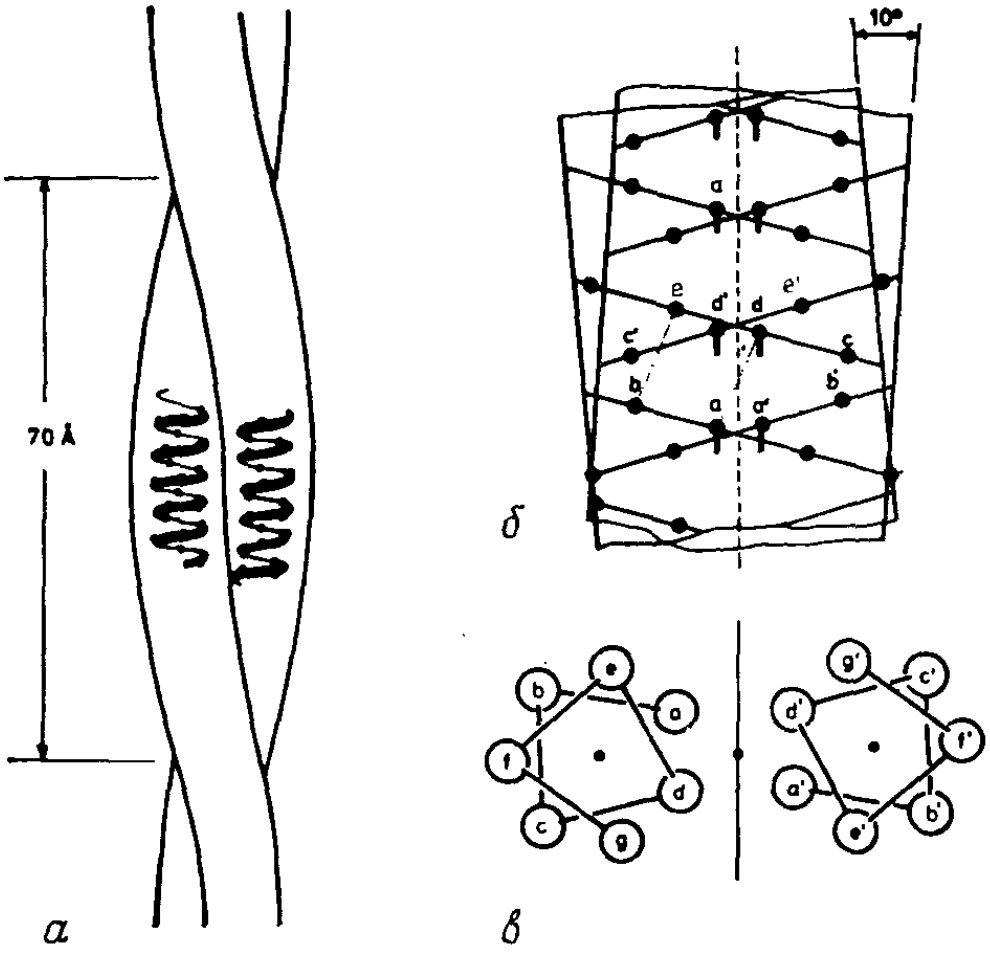


Рис. 5.11.

Скрученная двухцепочечная α-спираль.

а — общий вид левой двухцепочечной суперспирали с периодом идентичности 140 Å. Показаны размеры отдельных α-спиралей. Обе спирали параллельны. *б* — суперпозиция цилиндрических диаграмм двух параллельных α-спиралей. Пунктиром показана центральная ось суперспирали. C_α-Атомы обозначены жирными точками. При наклоне в 10° все контакты *aa'*, *dd'* и т. д. располагаются вдоль оси суперспирали. Боковая цепь *d'* и др. (обозначены короткими штрихами, указывающими противоположные спирали направления связей $C_\alpha - C_\beta$) располагаются между боковыми цепями *a*, *e*, *d* и боковой цепью *a* следующего витка. Таким образом, все боковые цепи в областях контактов между *a*, *d* и т. д. и *a'*, *d'* образуют плотную упаковку по типу выпуклость—впадина. *в* — сечение параллельной двухцепочечной α-спирали. Положения атомов C_α обозначены в алфавитном порядке вдоль полипептидных цепей; положения *a* и *a'* находятся на одной высоте.

217], белок оболочки вируса табачной мозаики [180, 218], бактериородопсин [219], оболочка бактериофага fd [220] и тирозил-тРНК-синтетаза [221].

В суперспирали возможна плотная упаковка боковых цепей без существенных искажений α-спиралей. Структуру суперспирали можно представить путем наложения цилиндрических разверток двух соседних α-спиралей. Это может быть сделано таким образом (рис. 5.11, *б*), что образуется контактная линия, по которой осуществляются взаимодействия боковых цепей чередующихся остатков обеих спиралей. Такое взаимодействие может продолжаться беспрепятственно, если обе спирали образуют левую суперспираль с контактной линией в виде прямой оси. Шаг суперспирали можно определить с помощью контактной линии на одной из цилиндрических разверток, а именно по участку развертки, который отсекается двой-

ным пересечением с ней контактной линии. При угле между α-спиралями около 10° (рис. 5.11, *б*) шаг составляет приблизительно 140 Å. Значения углов (ϕ, ψ) несколько отклоняются от величин, отвечающих идеальной одиночной α-спирали. Повторяющейся структурной единицей в каждой цепи является гептапептид (рис. 5.11, *в*) — седьмой остаток оказывается в эквивалентном положении относительно оси суперспирали (рис. 5.11, *б*). Такая повторяемость в 7,0 остатка несколько меньше удвоенной высоты витка одиночной α-спирали (2· 3,6 = 7,2 остатка на виток). Если представить аминокислотную последовательность цепи формулой $(abcdefg)_m$, то все остатки *a* (или соответственно *b*, *c*, *d*, *e*, *f*, *g*) будут находиться в структурно эквивалентных положениях (рис. 5.11, *в*).

Суперспирализация α-спирали энергетически выгодна, поскольку упаковка боковых цепей способствует образованию дополнительных благоприятных вандерваальсовых контактов между α-спиралями. Если взаимодействующие боковые цепи гидрофобны, то уменьшение свободной энергии такой структуры будет особенно эффективным, поскольку, располагаясь вдоль оси суперспирали, в этом случае боковые цепи экранированы от контактов с молекулами растворителя. Действительно в тропомиозине [222] и α-кератине [212], аминокислотные последовательности которых известны, положения *a* и *d* внутри суперспирали заняты гидрофобными остатками (рис. 5.11, *в*). Полярные остатки обычно располагаются на внешней поверхности, в положениях *b*, *c* и *f*. В тропомиозине положения *e* и *g* часто заняты заряженными остатками, которые образуют солевые мостики с находящимися напротив остатками *g'* и *e'* соответственно.

Модель сверхспирали пригодна как для параллельных, так и для антипараллельных α-спиралей. Все приведенные выше соображения отвечают параллельным и антипараллельным α-спиралям. Однако способу упаковки, приведенному на рис. 5.11, *б*, удовлетворяет только параллельное расположение; поскольку связь C_α—C_β α-спирального остатка не перпендикулярна оси α-спирали, а ориентирована под углом около 45° в направлении, обратном ходу спирали, то боковые цепи имеют ту же ориентацию. Они хорошо совмещаются только в случае параллельных спиралей (рис. 5.11, *в*). Для антипараллельных спиралей система совмещения будет выглядеть несколько иначе, чем показанная на рис. 5.11, *б*. Для устранения стерических отталкиваний между боковыми цепями требуется относительный сдвиг спиралей вдоль контактной линии примерно на 1 Å. Если, например, инвертировать показанную на рис. 5.11, *в* спираль с остатками *abcde*, отталкивания между боковыми цепями остатков *a* и *d'* устраняются при перемещении спирали на 1 Å вниз. Происходящее при этом сближение боковых цепей остатков *a'* и *d* не приведет к стерическим затруднениям, так как эти

цепи направлены в противоположные стороны. Таким образом, ни одно из направлений спирали, по-видимому, не является предпочтительным.

С помощью химических данных, а также результатов рентгеноструктурного анализа и электронной микроскопии было показано, что в тропомиозине [213, 223, 224] и в легком меромиозине [215] α-спирали параллельны. По-видимому, это относится и к α-кератину, поскольку длинная цепь α-кератина может быть полностью синтезирована и стабилизирована, прежде чем сможет образоваться суперспираль из антипараллельных α-спиралей. Напротив, в глобулярных белках гемеритрине [216, 217] и оболочке вируса табачной мозаики [180, 218] упаковка спиралей антипараллельна.

Были предложены три другие предпочтительные относительные ориентации с различными контактами между двумя α-спиралями [225]. Они получены из рассмотрения цилиндрических разверток и также соответствуют благоприятной упаковке боковых цепей. Однако они пока еще не подкреплены надежно экспериментальными данными.

βξβ-Структурное образование

βξβ-Звено обычно правое. Другим распространенным элементом, который обычно относят к сверхвторичной структуре, является βξβ-звено (обозначаемое также βxβ). Оно представляет собой два параллельных β-листа с сочленением ξ между ними. Если сочленение имеет вид неупорядоченной цепи (клубка), звено обозначают βcβ; в случае α-спирали складчатого или листа другой β-структуры звено обозначают βαβ или βββ. Как видно из рис. 5.12, *а*, такие звенья хиральны. В глобулярных белках встречаются почти исключительно правые βξβ-звенья [226—228]. Это безусловно указывает на предпочтительный путь свертывания цепи и на предпочтительную форму агрегации вторичной структуры, т. е. сверхвторичной структуры.

Предпочтительную хиральность можно объяснить асимметричными стерическими затруднениями в неупорядоченной полипептидной цепи. Правая хиральность настолько выгоднее, что этому можно попытаться дать объяснение. Как уже отмечалось выше, статистические веса конформационных состояний полипептидной цепи благоприятствуют углам (ϕ, ψ) β-складчатого листа (рис. 5.10, *а*). Поэтому вытянутая цепь в растворе обычно будет иметь те же значения (ϕ, ψ) и приобретет правый поворот, показанный на рис. 5.10, *б*. Если такая цепь образует петлю, то при общем правом направлении закрутки цепи ее скручивание будет увеличиваться при левой петле и уменьшаться при правой. Эта взаимосвязь проиллюстрирована на рис. 5.13. Дополнительное скручивание означает увеличение углов ϕ, ψ, что видно из сравнения плос-

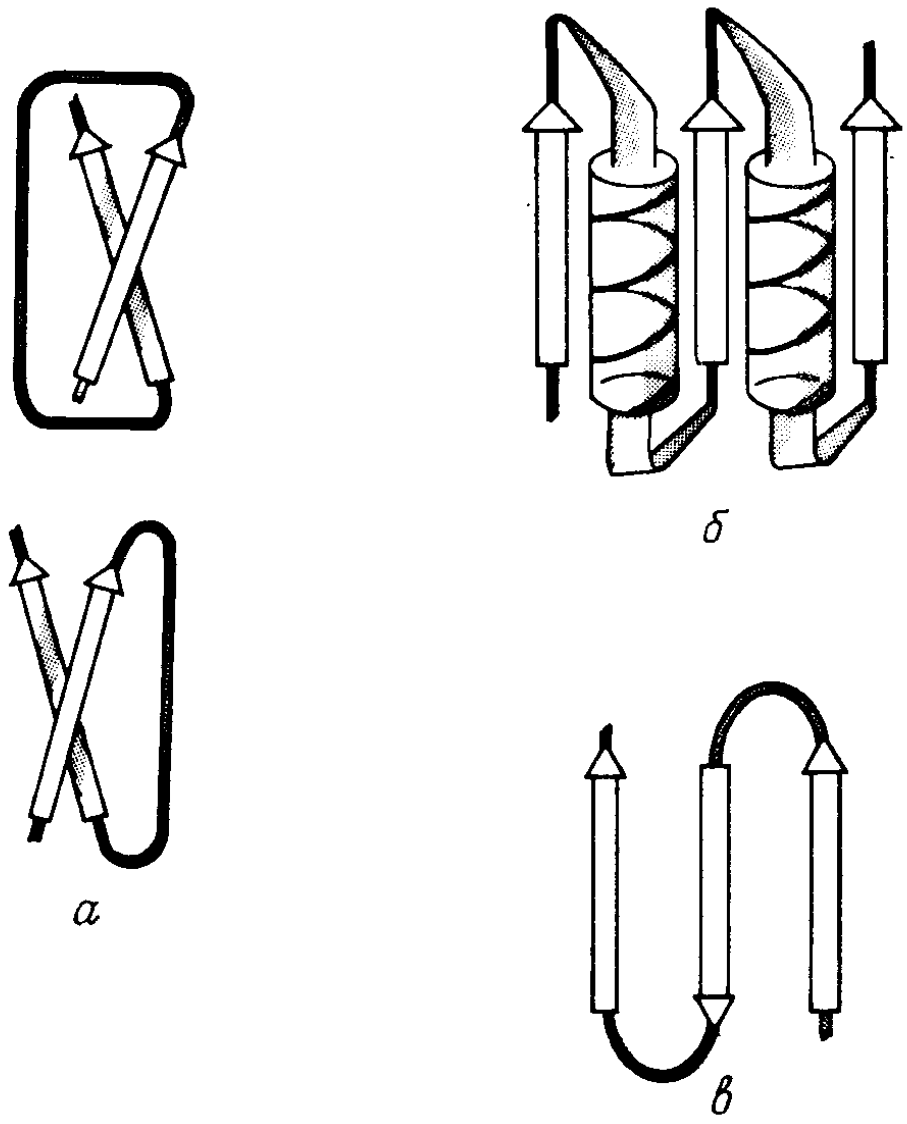


Рис. 5.12.

Сверхвторичные структуры, включающие β-складчатые листы. Стрелки обозначают складчатые листы.

а — βξβ-звено. Левое (сверху) и правое (снизу) сочленения ξ двух параллельных складчатых листов, образующих левую скручениую β-структуру (рнс. 5.10, *г*). Обычно наблюдается правое сочленение. *б* — укладка цепи по Россману, представляющая собой два последовательно соединеиных правых участка βαβ. *в* — β-зигзаг, являющийся антнпараллельной трехцепочечной β-структурой.

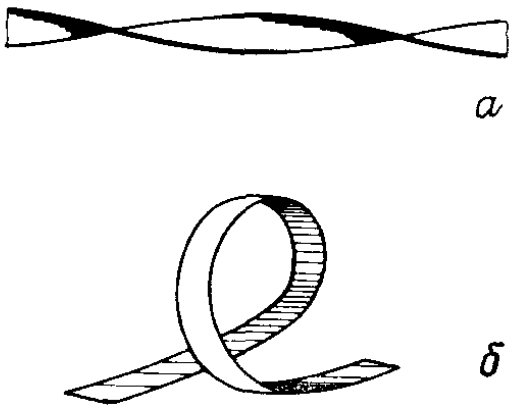


Рис. 5.13.

Схема свертывания полипептидной цепи. Правое скручивание полипептидной цепи невозможно при образованнн правой петли. В этом можно убедиться, скручивая ремень или полоску бумаги.

а — полоска бумагн с правой закруткой. *б* — иенапряженная левая петля образуется, если сближать концы полоски бумагн, показанной на рнс. *а*.

кого и складчатого листов. Если исходить из углов (ϕ, ψ) в складчатом листе, то увеличение ϕ и ψ (перемещение к верхней правой части карты) значительно более затруднено (рис.2.5), чем уменьшение этих углов (перемещение к нижней левой части). Таким образом, в общем случае благоприятно уменьшение скручивания и петля становится правой. Очевидно, что если в процессе свертывания началось образование правой петли, то изменить направление скручивания уже трудно. Таким образом формируется правое βξβ-звено.

Сочетание двух последовательных βαβ-звеньев. К сверхвторичной структуре можно отнести также «свертыванием по Россману» [229] (рис. 5.12, *б*). Оно представляет собой частный случай βξβ-звена и состоит из двух последовательных βαβ-звеньев. Такое βαβαβ-звено, которое может содержать между слоями и спиралями гидрофобное ядро, было обнаружено в ряде белков [186, 230—240].

β-Зигзаги и другие особенности β-структур

β-Зигзаг содержит столько же водородных связей на остаток, как и α-спираль средней длины. В некоторых белках, например Т4-лизоциме [241], стафилококковой нуклеазе [242], Ser-протеазе [18, 243—246] и каталитическом домене лактатдегидрогеназы [232], β-структура реализуется в виде трех близко расположенных антипараллельных пептидных цепей, соединенных между собой короткими участками. Такой структурный элемент называют β-зигзагом; он показан схематически на рис. 5.12, *в*. Часто встречающийся в белках β-зигзаг можно классифицировать как сверхвторичную структуру, он был обнаружен в половине всех известных β-структур из трех последовательно расположенных антипараллельных цепей [247]. Такая ситуация обусловлена большой согласованностью взаимодействий между антипараллельными цепями (разд. 5.3, см. также рис. 5.15, *а*). Кроме того, поскольку β-зигзаг включает на обоих концах повороты цепи, он обеспечивает образование примерно двух третей всех возможных водородных связей между атомами остова в данном участке цепи. Это соответствует числу водородных связей в α-спирали со средней длиной в три витка (рис. 5.5). Таким образом, по своей свободной энергии и с точки зрения взаимодействий между соседними остатками β-зигзаг сопоставим с α-спиралью, чем, по-видимому, и объясняется его распространенность.

В β-структурах предпочтительны вполне определенные расположения цепей. Хорошая корреляция между взаимодействиями остатков (рис. 5.15, *а*) в β-структурах отражает наличие предпочтительности некоторых расположений цепей в β-листах. Особенности β-структур были подробно рассмотрены Ричардсоном [247], который проанализировал частоту всех структурных элементов из двух и трех последовательных β-цепей, встречающихся в β-

структурах из четырех и более цепей. Соответствующее распределение неравномерно; β-зигзаги оказались наиболее часто встречающимся структурным элементом в образованиях из трех последовательных пептидных цепей.

Больший интерес однако представляет распределение β-листов по топологии*. Достаточно надежное эмпирическое распределение пока еще не получено из-за малого объема экспериментальных данных. Поэтому для всех пяти известных топологий β-структур такое распределение было рассчитано путем их построения из всех возможных сочетаний двух или трех последовательных β-цепей с последующим анализом частоты их реализации. На основании полученного неравномерного распределения можно предсказать, что некоторые варианты топологии β-листа должны встречаться намного чаще, чем другие. Этот результат важен для идентификации сверхвторичных структур или структурной предпочтительности, а также для оценки значимости сходства между структурами (разд. 9.6).

5.3 Структурные домены

Структурные домены четко ограничены на картах электронной плотности. Анализ карт электронной плотности, рассчитываемых в ходе рентгеноструктурного анализа, показал, что многие белки состоят из нескольких глобулярных областей, довольно слабо связанных между собой. Эти области, четко ограниченные на картах распределения электронной плотности, получили несколько неопределенное название «структурные домены». Очевидно, что определение домена весьма нестрого и здесь может возникать много спорных ситуаций. Среди глобулярных белков четко определенные домены были обнаружены в иммуноглобулинах. Схематически они изображены на рис. 4.2, *в*; в этом случае домены располагаются вдоль полипептидной цепи, как жемчужины в ожерелье.

Корреляция между близко расположенным в последовательности остатками

Остатки, далеко отстоящие по цепи, обнаруживают тенденцию к удаленному расположению и в трехмерной структуре. Как видно из рис. 4.2, *в*, полипептидную цепь можно разделить на несколько последовательно расположенных участков, которые принадлежат последовательно расположенным доменам [248]. Таким образом, остатки, которые в цепи находятся далеко друг от друга, оказы-

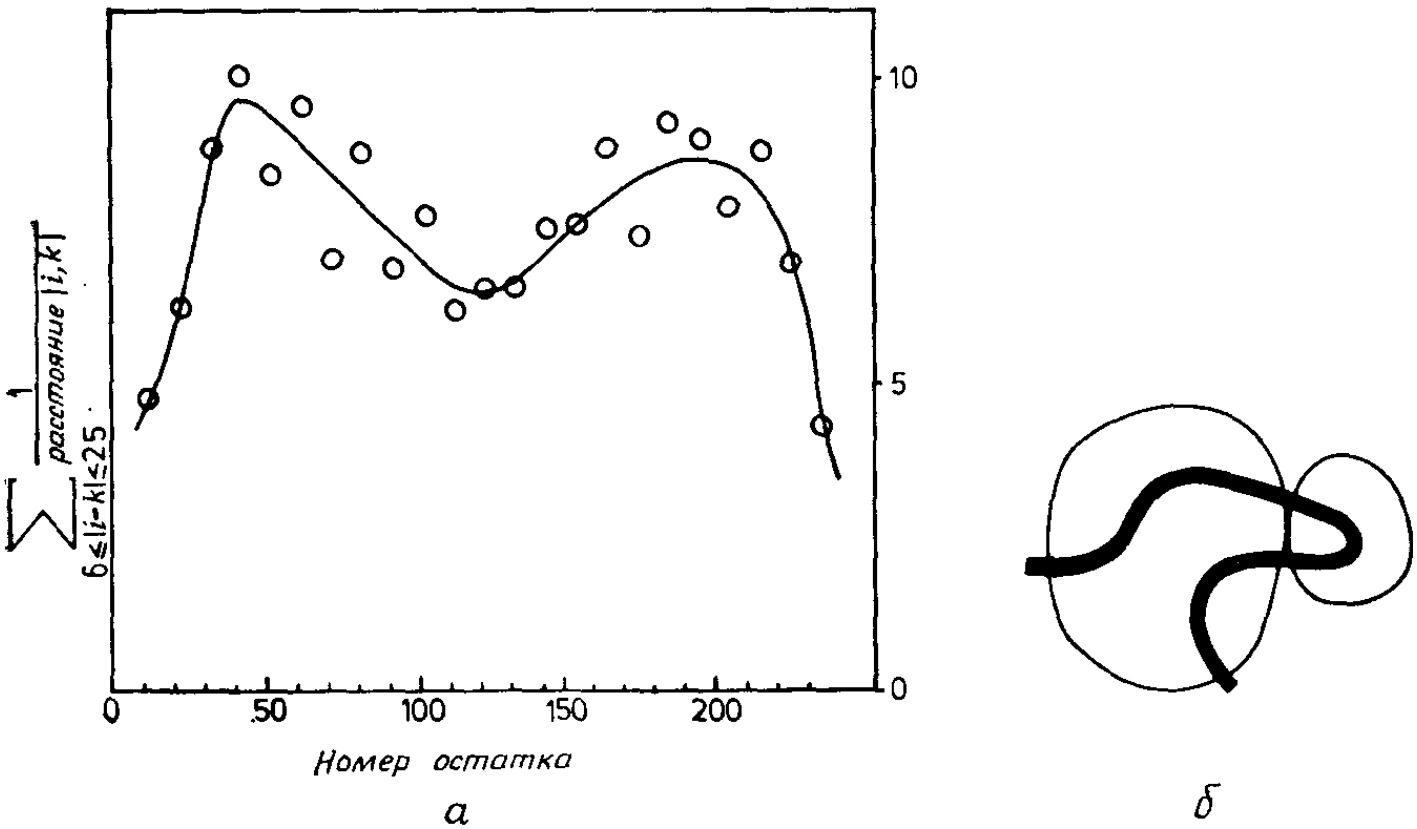
* Топология β-структуры определяется порядком чередования цепей согласно аминокислотной последовательности, направлением этих цепей в листе и хиральностью ββ-звеньев.

ваются значительно удаленными и в пространстве. Этот принцип соблюдается во всех глобулярных белках, состоящих из четко определенных доменов (табл. 5.2). Иначе говоря, доменная структура указывает на высокую степень «корреляции между близко расположенными по цепи остатками», расстояние между остатками в цепи коррелирует с их расстоянием в пространственной структуре.

Наличие корреляции между близкими в цепи остатками способствует выявлению структурных доменов. Такой способ был количественно оценен для химотрипсина, деление которого на два домена не очень ясно проявлялось на карте электронной плотности. В качестве критерия была избрана сумма всех величин, обратных расстояниям между парами остатков, которые разделены 6—25 остатками в цепи:

$$\text{корреляция между близкими по цепи остатками } (i) = \sum_{6 \leqslant |i-k| \leqslant 25} \frac{1}{\text{расстояние } (i, k)} \qquad (5.1)$$

Результаты приведены на рис. 5.14, *а*. Два пика с минимумом между ними указывают на наличие двух областей, в которых остатки, близкие по цепи, в результате процесса свертывания оказались сближенными друг с другом и в пространстве. Эти об-



а *б*

Рис. 5.14.

Корреляция между соседними остатками полипептидной цепи.

а — мера корреляции между соседними остатками как функция номера остатка *i* в химотрипсине. Для сглаживания кривой значения по 10 остаткам усреднены. Отчетливо выделяется двухдоменная структура, схематически показанная на рис. 5.17, *г*. *б* — доменная структура аденилаткиназы. Малый и большой структурные домены связаны двумя полипептидными цепями. Корреляция между соседними остатками в малом домене выражена сильнее, чем в большом.

ласти следует классифицировать как домены. Сравнение с трехмерной структурой химотрипсина показывает, что каждый домен содержит одну из показанных на рис. 5.17, *г* β-структур.

Последовательные цепи β-структур стремятся к близкому расположению. Отмеченная корреляция между близкими по цепи остатками проявляется не только на уровне доменов, но также и внутри их. Наглядным количественным примером может служить гистограмма расстояний в β-структурных цепях, приведенная на рис. 5.15, *а*. В частности, в цепях антипараллельной β-структуры, где чисто пептидная цепь образует два или более последовательных β-зигзага, проявляется заметная корреляция между близкими по цепи остатками. В параллельных слоях такая

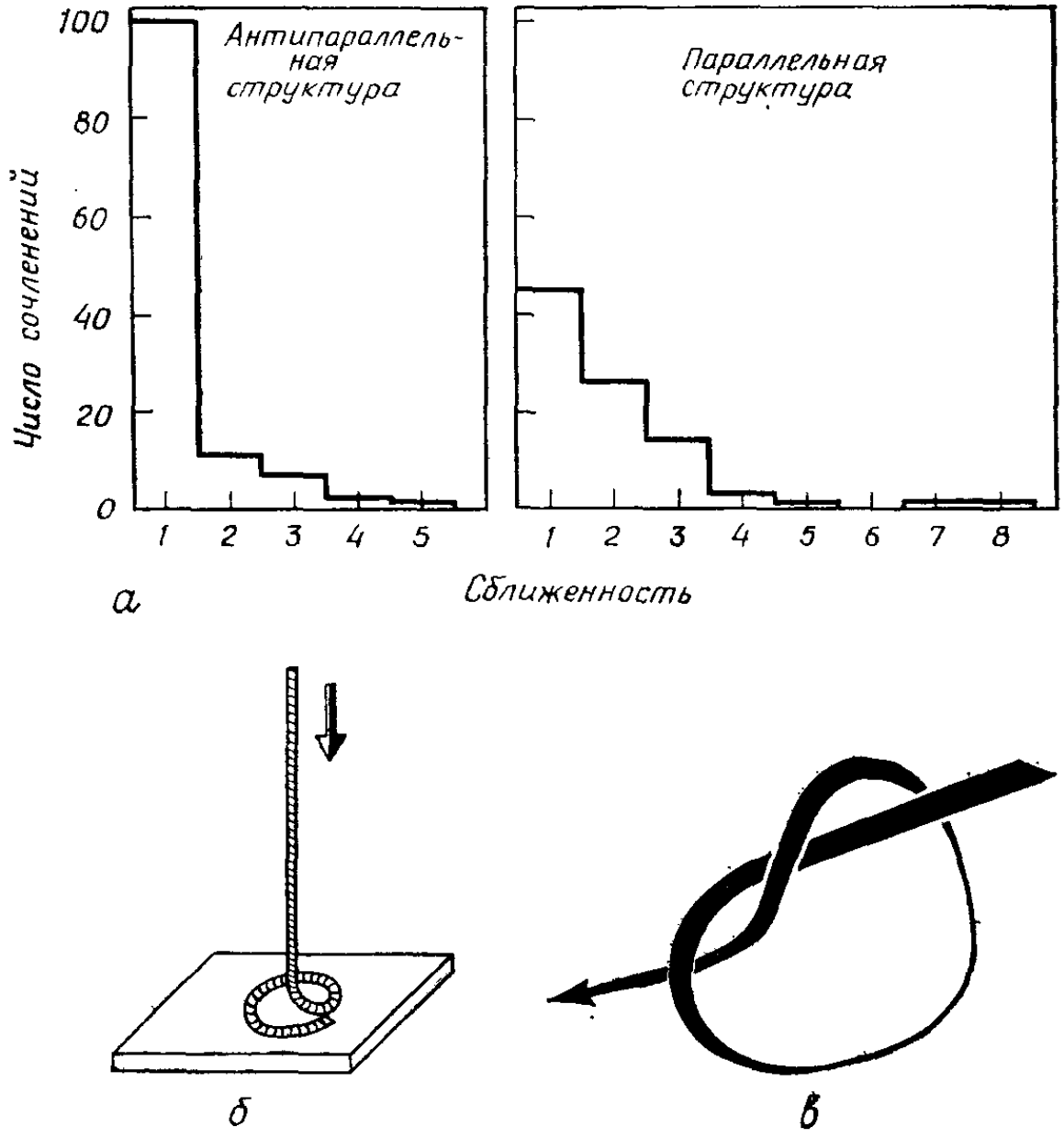


Рис. 5.15.

Предпочтительные структуры глобулярных белков.

а — корреляция между соседними остатками, наблюдаемая в антипараллельной и параллельной β-структурах. Приведено число контактов между одной цепью в β-структуре (вдоль цепи) и ее первой, второй, ... соседними цепями [327]. *б* — удерживаемой вначале за верхний конец веревке предоставляют возможность упасть. В свернутой форме веревки отсутствуют узлы и отчетливо проявляется корреляция между ее соседними частями. *в* — узел полипептидной цепи; экспериментально не обнаружен.

корреляция слабее, однако и там она имеет достаточно важное значение.

Корреляцию между близкими по цепи остатками можно интерпретировать, пользуясь как кинетическим, так и термодинамическим подходами. Корреляция между близкими по цепи остатками возникает, по-видимому, в процессе свертывания цепи. Поскольку конформаций, сближающих удаленные по цепи остатки, значительно больше, чем конформаций, сближающих близкие по цепи остатки, то процесс свертывания должен существенно зависеть от взаимодействий между последними. В противном случае уже на начальной стадии потребовалось бы опробывание слишком многих конформационных состояний, отделение правильных состояний от неправильных стало бы исключительно сложным и процесс свертывания цепи имел бы много шансов пойти по неверному пути.

С точки зрения термодинамики изложенные выше соображения выглядят следующим образом. При свертывании энтропия цепи должна понижаться (разд. 3.5). Первоначальное понижение энтропии минимально, если свертывание цепи приводит к конформации, близкой к среднестатистической конформации цепи в растворе, т. е. если оно приводит к структуре с сильной корреляцией между близкими по цепи остатками. При таком незначительном понижении энтропии не возникает первоначального (активационного) энтропийного барьера, который было бы трудно преодолеть за разумное время. Таким образом, процесс свертывания цепи может происходить постепенно, не требуя слишком большой связывающей энергии и гидрофобных сил (см. также разд. 8.3).

Учитывая корреляцию между близкими по цепи остатками, можно сравнить свертывающуюся цепь с веревкой, которую опускают вниз способом, показанным на рис. 5.15, *б*. Образующийся на полу клубок не беспорядочен, но характеризуется корреляцией между близкими участками, он не запутан, и веревку можно легко поднять за один конец. Эта аналогия подтверждается отсутствием «узлов» во всех известных структурах. Термин «узел» использован здесь в его обиходном смысле (рис. 5.15, *в*), а не в математическом (математические узлы могут существовать только в замкнутых системах).

Структурные домены, по-видимому, служат единицами свертывания белка. Корреляция между близкими по цепи остатками дает основание рассматривать домены как такие участки цепи, которые свертываются независимо друг от друга. Если это так, то структурные домены можно определить как единицы свертывания цепи, т. е. более конкретно, чем ранее. Строение химотрипсина в известной степени подтверждает такое определение [18]. Этот белок содержит 13 молекул воды между двумя доменами внутри молекулы (рис. 5.14, *а*). По-видимому, домены свертываются отдельно и молекулы воды задерживаются в процессе последующей ассоциации

доменов. Это соображение подкрепляется тем фактом, что *функциональные* домены (как было выяснено биохимическими испытаниями) действительно свертываются независимо друг от друга [76, 77].

Совершенно очевидно, что большинство крупных белков можно разделить на несколько структурных доменов, содержащих 100—150 остатков, которые отвечают глобуле диаметром около 25 Å (табл. 5.2). С точки зрения свободной энергии такое ограничение величины домена является неожиданным, поскольку одна крупная глобула характеризуется значительно меньшим отношением поверхность/объем, чем несколько более мелких. В большой глобуле должно образоваться крупное гидрофобное ядро и многочисленные внутренние водородные связи: и то, и другие энергетически выгодно. Вероятно, ограничение величины необходимо для простоты процесса свертывания путем соблюдения достаточно малой длины независимо свертывающегося звена.

В некоторых случаях, как, например, в аденилаткиназе [186], имеются две области, которые можно называть доменами и которые сочленены двумя цепями вместо одной. Как видно из рис. 5.14, *б*, корреляция между соседними по цепи остатками в этих двух доменах заметно отличается. По-видимому, домен с сильной корреляцией должен свертываться первым, а затем использоваться в качестве своеобразной матрицы, способствующей свертыванию другого домена.

Структурные классы

Структурные домены можно классифицировать по их вторичной структуре. Известные структуры доменов можно разбить на пять приведенных в табл. 5.2 классов в соответствии с их типом, распространенностью и организацией их вторичной структуры [249, 250]. Наибольший интерес представляют домены, состоящие только из одного типа вторичной структуры. Среди них для α-спиральных доменов можно ожидать наиболее простого процесса свертывания. Образование спирали, по-видимому, происходит на начальной стадии свертывания, после чего единственной задачей остается упаковка α-спиральных цилиндров энергетически выгодным способом.

Все чисто β-структурные домены (класс № 2) содержат антипараллельные β-структуры, обеспечивающие сильную корреляцию между близкими по цепи остатками (рис. 5.15, *а*). Для них характерна высокая упорядоченность. Менее упорядоченные структуры входят в класс № 3, состоящий из доменов, в которых α- и β-структуры разделены в пептидной цепи. Класс № 4 образуют преимущественно параллельные β-структуры в центре доменов и окружающие их α-спирали, так что α- и β-структуры обнару-

Таблица 5.2

Структура доменов и структурные классы глобулярных белков[a]

Структурный класс	Часть глобулярного белка, занимаемая структурным доменом	Число остатков в домене	Белок	Число доменов в глобулярном белке	Число соединяющих домены цепей
№ 1	Весь белок	104	Цитохром *c* [271—273]	1	
Только α-спирали	»	158	Белок оболочки ВТМ [180, 218]	1	
	»	108	Парвальбумин [59]	1	
	»	~ 145	Гемоглобин [188, 274—276]	1	
	»	153	Миоглобин [185, 277]	1	
	»	~ 120	Гемеритрин [216, 217]	1	
	»	51	Инсулин [259]	1	
	»	29	Глюкагон [278]	1	
	С-конец	111	Папаин [279]	2	1
	»	160	Термолизин [280]	2	1
	»	90	Т4-лизоцим [241]	2	1
	N- и С-концы	~ 100, ~ 100	Туг-тРНК-синтетаза [221]	3	
№ 2	Весь белок	237	Конканавалин А [281, 282]	1	
Почти исключительно β-структура	»	127	Проальбумин [206, 283]	1	
	»	54	Рубредоксин [284, 285]	1	
	»	151	Пероксид-дисмутаза [286, 800]	1	
	»	62	Эрабутоксин *b* [287]	1	

	Весь белок	113	Ингибитор субтилизина [798]	1	
	N-конец	110	Папаин [279]	2	(1)
	N- и С-концы	128, 114	Химотрипсин [18]	2	1
	»	110, 119	Эластаза [243]	2	1
	»	107, 116	Трипсин [244, 245]	2	1
	»	80, 106	Протеиназа-В [246]	2	1
	»	175, 145	Кислая протеаза [288—290]	2	1
	N-конец	175	Алкогольдегидрогеназа [234]	2	1
	Весь белок	~110	Иммуноглобулин [291—294]		1
№ 3	Весь белок	258	Карбоангидраза [295, 296]	2,4	
α-Спираль и β-структура, разделенные вдоль цепи	»	85	Цитохром b_5, выделенный из печени теленка [297, 298]	1	
	»	85	Железосодержащий белок с высоким окислительно-восстановительным потенциалом [299]	1	
	»	129	Лизоцим белка куриного яйца [260, 300]	1	
	»	56	Панкреатический трипсиновый ингибитор [269]	1	
	Весь белок	124	Рибонуклеаза [39, 301, 302]	1	
	»	149	Стафилококковая нуклеаза [242]	1	
	»	~330	Белок бактериохлорофилла [303]	1	
	N-конец	156	Термолизин [280]	1	
	»	74	Т4-лизоцим [241]	2	(1)
	С-конец	150	Лактатдегидрогеназа [232]	2	(1)
	»	170	s-Малатдегидрогеназа [233]	2	1
	»	118	Глутатионредуктаза [124]	2	1
	Весь белок	194	Аденилаткиназа [186]	3	
№ 4				1	
α-Спираль и β-структура, альтернирующие вдоль цепи	»	307	Карбоксипептидаза [43]	1	
	»	140	Флаводоксин [237, 238]	1	
	»	275	Субтилизин [239, 240]	1	
	»	108	Тиоредоксин [304]	1	

Продолжение табл. 5.2

Структурный класс	Часть глобулярного белка, занимаемая структурным доменом	Число остатков в домене	Белок	Число доменов в глобулярном белке	Число соединяющих домены цепей
	Весь блок				
	»	247	Триозофосфатизомераза [305]	1	
	»	~240	Бактериальная альдолаза [306]	1	
	»	~230	Фосфоглицеромутаза [307]	1	
		159	Дигидрофолатредуктаза [308]	1	
	N-конец	181	Лактатдегидрогеназа [232]	2	(1)
	»	~150	s-Малатдегидрогеназа [233]	2	(1)
	С-конец	199	Алкогольдегидрогеназа [234]	2	(1)
	N- и С-концы	147, 187	Глицеральдегид-3-фосфатдегидрогеназа [230, 231]	2	1
	»	~150, ~150	L-Арабинозасвязывающий белок [309]	2	(1)
	»	~180, ~175	Фосфоглицераткиназа [235, 310, 311]	2	
	»	~250, ~150	Пируваткиназа [80, 312]	3	
	»	~130, ~130	Роданеза [257, 799]	2	
	»	~360, ~150	Глюкозофосфат-изомераза [313]	2	
	»	~230, ~230	Гексокиназа [261]	2	
	N-конец и центральная часть	210, 134	Глутатионредуктаза [124]	3	
	Центральная часть	~200	Туг-тРНК-синтетаза [221]	3	
№ 5	Весь белок	54	Ферредоксин [314]	1	
Ни α-спирали,	»	124	Фосфолипаза [315]	1	
β-слоя	»	41	Агглютинин, выделенный из зерна пшеницы [316]	4	1

[a] Если ход цепи еще не подтвержден данными об аминокислотной последовательности, то число соединяющих домены цепей в мультидоменных белках не приводится. Число цепей обозначается как (1), если около десяти остатков, находящихся на N-или С-концах цепи, перекрываются с другим доменом. Классы № 3 и № 4 называют также α+β и α/β соответственно [249].

живают тенденцию к альтернированию вдоль цепи. К этому классу принадлежит свертывание цепи по Россману (рис. 5.12, *б*). Домены без выраженной вторичной структуры сведены в класс № 5.

Очевидно, что такая классификация может быть полезной при трактовке процесса свертывания. Эта задача усложняется при увеличении номера класса, т. е. по мере снижения структурной регулярности цепи.

Симметрия

В пределах одной полипептидной цепи часто встречаются повторяющиеся структурные единицы. Повторяющиеся структурные единицы существуют во многих белках. Такой единицей может быть домен, сверхвторичная структура или какой-либо другой структурный элемент. Примеры таких структур приведены в табл. 5.3. В некоторых белках повторяющиеся единицы расположены друг относительно друга несимметрично, как, например, домены Ser-протеазы (рис. 5.17, *г*) или кофактор, — связывающие домены глутатионредуктазы [124]. Обращает на себя внимание, однако, то обстоятельство, что часто повторяющиеся единицы имеют *приблизительно* симметричное расположение; это дает основание полагать, что симметрия является характерным признаком образуемой ими структуры.

Для большого числа идентичных единиц состоянием с наименьшей свободной энергией является кристалл. Предпочтительность симметрии в белках можно уяснить, рассматривая кристаллы, симметрия которых является следствием выгодной упаковки. Упаковка зависит от формы повторяющихся единиц и от профиля их поверхности. Поскольку среди всех возможных схем упаковки имеется одна энергетически наиболее выгодная, она будет отвечать минимальной свободной энергии, в случае если она адекватна для всех звеньев. В любом случае это состояние соответствует кристаллу, принадлежащему к одной из 65 пространственных групп симметрии, допустимых для асимметрических элементов, поскольку кристаллизация остается единственным способом агрегирования бесконечного числа единиц, при котором все образующиеся контакты идентичны.

Для небольшого числа ассоциированных единиц энергетически выгодна симметрия точечной группы. При построении структуры ограниченной величины используется аналогичный принцип. Среди всех контактов между идентичными единицами имеется один, энергетически наиболее выгодный. Если реализуется только этот контакт, то образуется спираль, т. е. линейная группа неограниченного размера. Ее размеры окажутся ограниченными, если на первом витке возникнут стерические затруднения (рис. 5.16, *б*). Однако в этом случае последний контакт, а значит и вся структура,

Таблица 5.3

Повторяющиеся структурные элементы глобулярных белков [a]

Белок	Повторяющаяся структура	Число повторяющихся структур	Приближенная симметрия
Химотрипсин эластаза трипсин	Полость, образованная шестицепочечной антипараллельной β-структурой	2	Отсутствует
протеиназа-В	β-Зигзаг внутри полости		2
Парвальбумин	Ca^{2+}-связывающий центр и две α-спирали	2 +1	2 Отсутствует
Ферредоксин	FeS-кластер и 27 остатков	2	2
Глицеральдегид-3-фосфатдегидрогеназа лактатдегидрогеназа s-малатдегидрогеназа алкогольдегидрогеназа фосфоглицераткиназа фосфорилаза	Свертывание по Россману из трех параллельных β-структурных цепей и двух соединительных фрагментов	2	2
Триозофосфатизомераза пируваткиназа	βαβ-Структурное образование	8	8
Аглютинин пшеничного зерна	41 остаток с 4 мостиками S—S	4	Линейная группа
Гемеритрин	Две антипараллельных α-спирали	2	2
Роданеза	Один домен	2	2
Иммуноглобулин	*V*-домен	4 на молекулу	2
	Домены C_H 1 и C_H 3	6 на молекулу	2
Глутатионредуктаза	Четырехцепочечная параллельная β-структура и β-зигзаг	2	Отсутствует
Кислая протеаза	Один домен	2	2
L-Арабиноза-связывающий белок	»	2	2

[a] См. литературу в табл. 5.2.

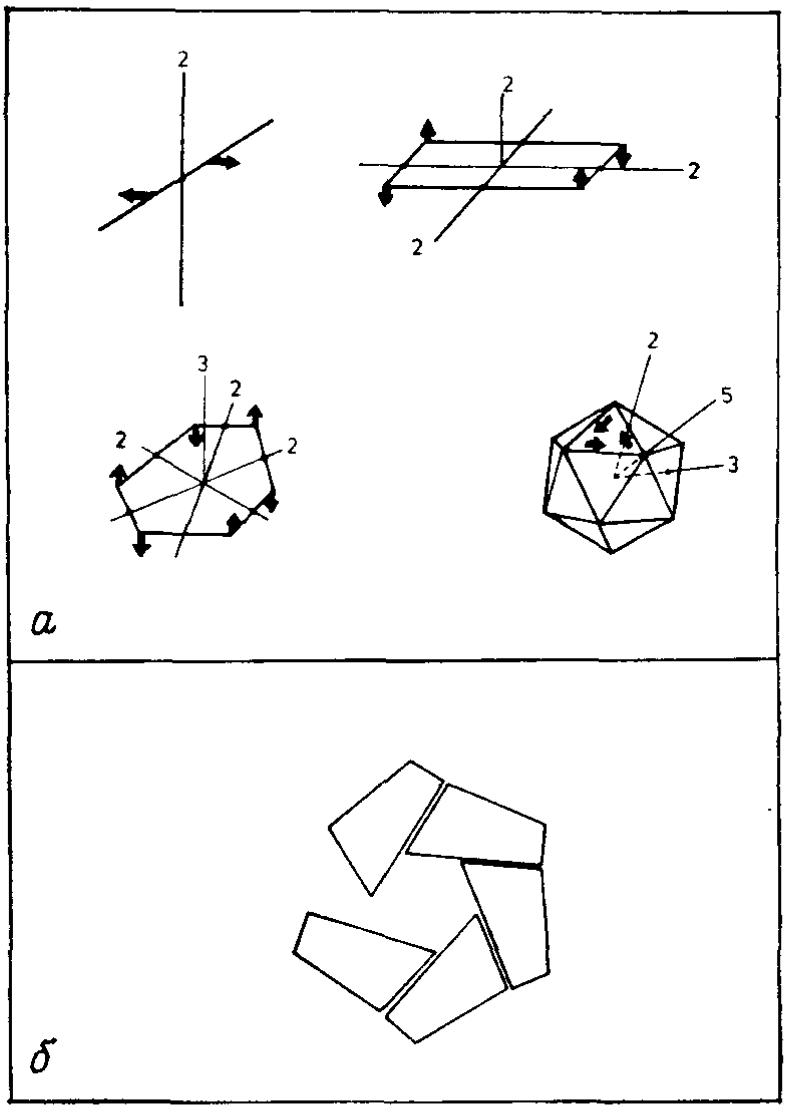


Рис. 5.16.

Точечная симметрия.

а — изображение точечных групп 2 (C_2), 222 (D_2), 32 (D_3) и 532 (Y, икосаэдр). Цифрами обозначены поворотные оси 2-го, 3-го и т. д. порядков. *б* — образующие предпочтительный контакт асимметричные элементы. В общем случае такой предпочтительный контакт не приводит к точечной группе. Образование линейной группы также может быть запрещено из-за стерических трудностей. В приведенном примере отсутствует связывающая энергия между первым и пятым звеньями и поэтому пятый контакт не вносит вклад в стабильность. Из-за дополнительной связывающей энергии предпочтительны полностью замкнутые, высокосимметричные структуры.

энергетически невыгодны*. Энергетический оптимум достигается тогда, когда цикл строго замкнут, т. е. если выдерживается симметрия. Поэтому симметричные образования предпочтительны также для агрегатов ограниченной величины. Приведенный аргумент проиллюстрирован на рис. 5.16, *б* для точечной группы n, т. е. для симметричного расположения вокруг точки с n-й поворотной осью в качестве оператора симметрии. Этот случай можно

* В агглютинине зерна пшеницы обнаружена открытая структура из четырех единиц, образующих три очень сходных контакта [253]. Это показывает, что особенно прочный трижды повторенный, контакт может быть выгоднее, чем альтернативный контакт, удовлетворяющий симметрии и потому повторенный четыре раза.

обобщить для всех точечных групп, образованных симметричными элементами: *n*, *n2*, *23*, *432*, *532* [251, 252]; некоторые из них показаны на рис. 5.16, *а*.

На основании симметрии можно выявить структурные единицы, которые свертываются отдельно. Полипептидная цепь представляет структуру ограниченной величины. Если в процессе

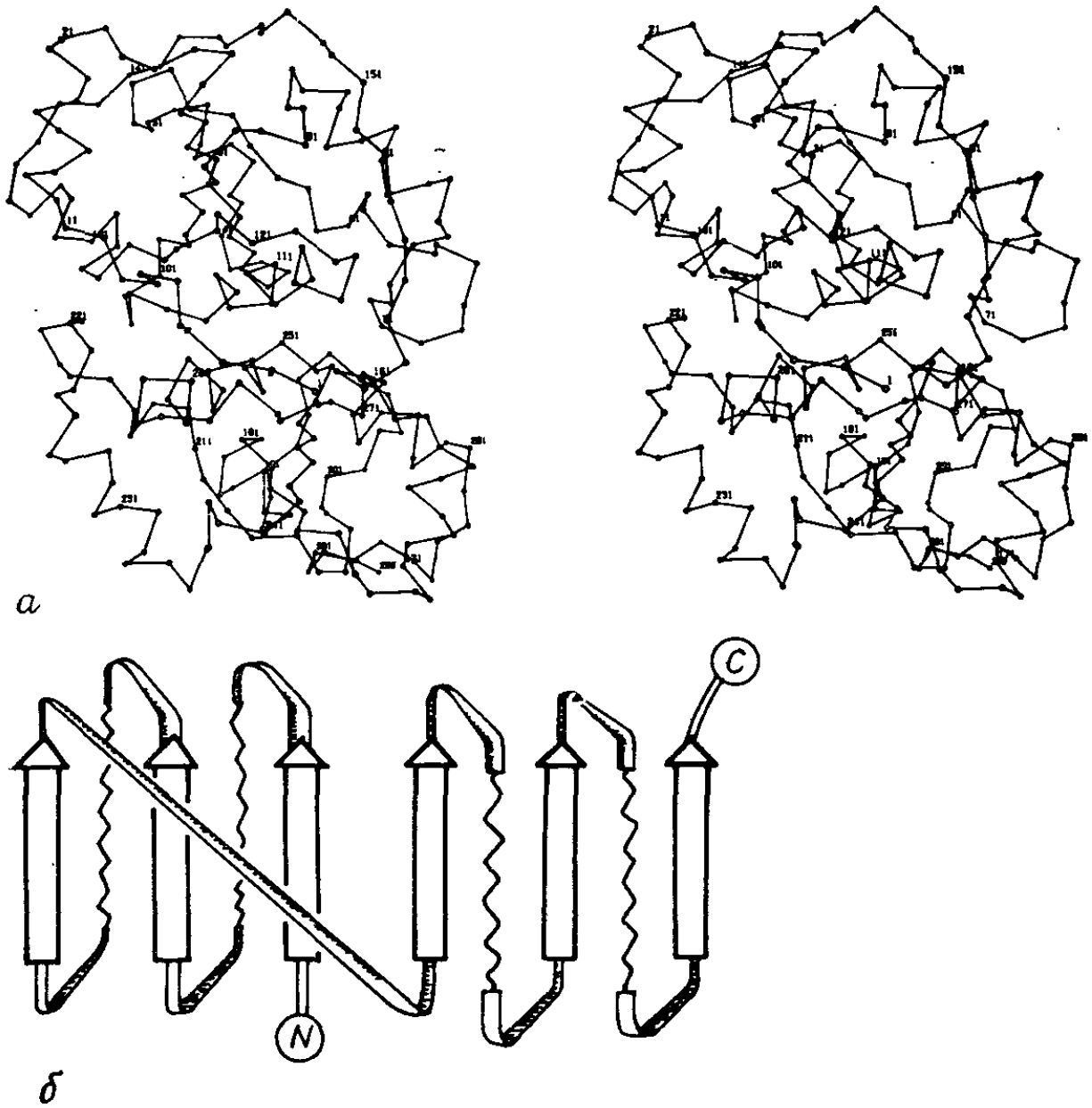
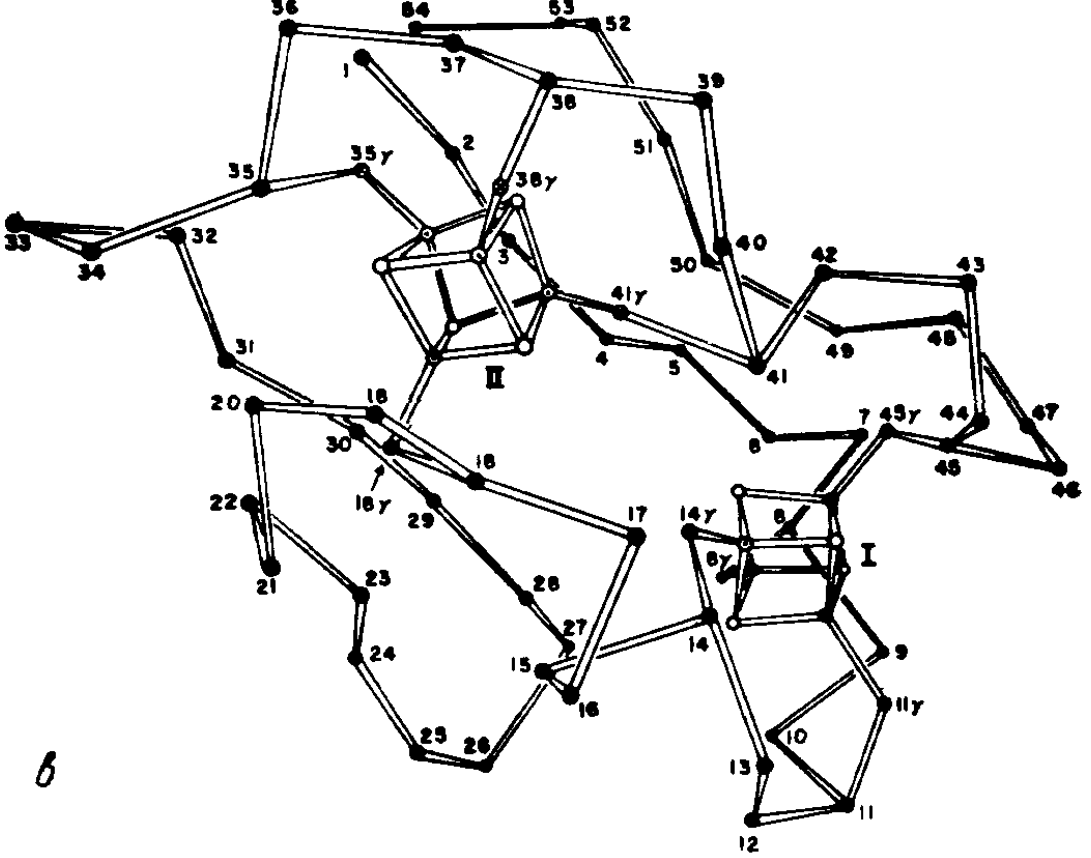


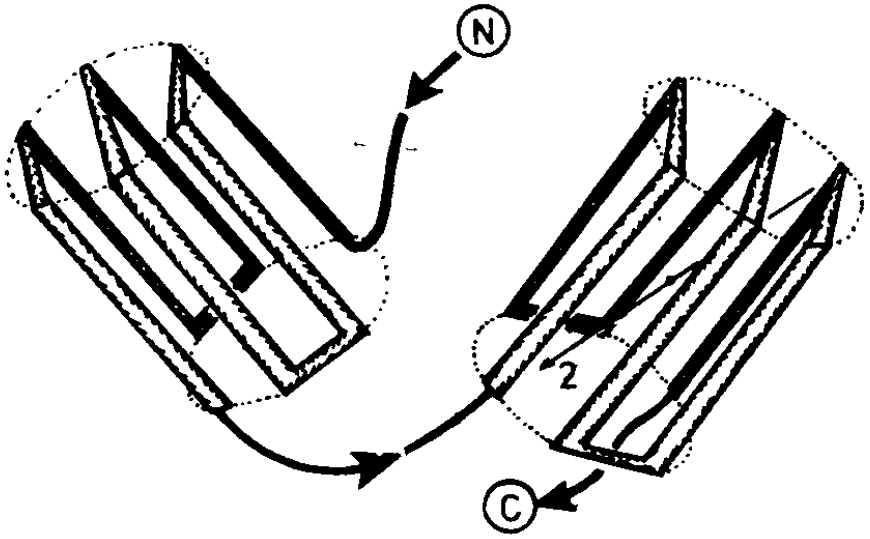
Рис. 5.17.

Симметрия одиночных полипептидных цепей.

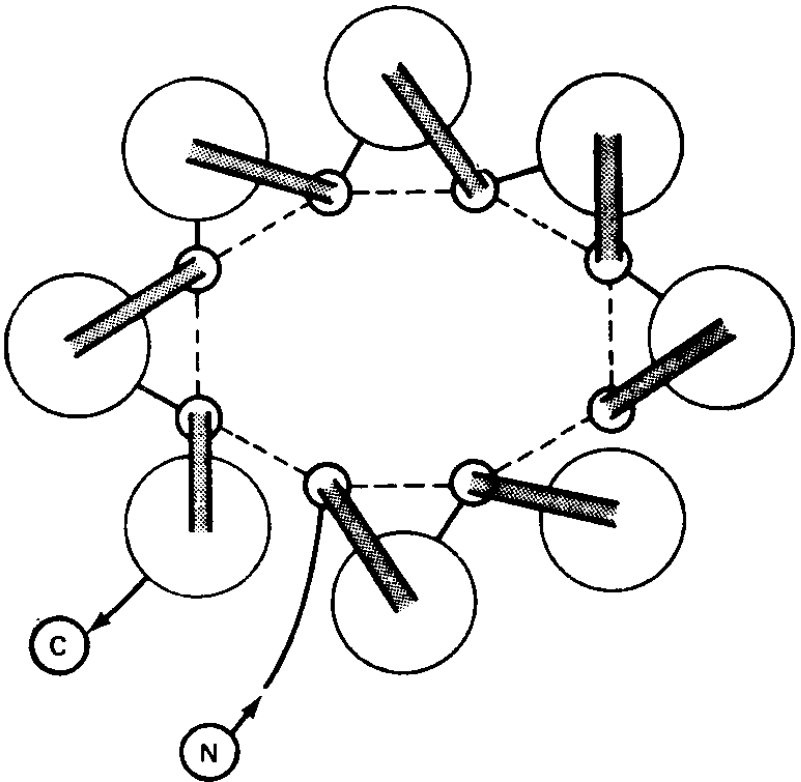
а — стереопроекция основной цепи роданезы [799]. Если не учитывать короткого промежуточного участка, связь между верхней и нижней половинами цепи почти точно передается поворотной осью 2-го порядка, перпендикулярной плоскости рисунка. *б* — топология β-структуры нуклеотидсвязывающего домена NAD-зависимой дегидрогеназы. α-Спирали показаны волнистыми линиями. Структура содержит два участка цепи, свернутых по Россману (рис. 5-12, *б*), связанных вертикальной осью 2-го порядка. *в* — схема основной цепи ферредоксина [314] с FeS-кластерами. Некоторые аминокислоты, FeS-кластеры и часть свернутых участков цепи имеют в качестве элемента симметрии ось 2-го порядка; она приблизительно перпендикулярна плоскости рисунка. *г* — топология двух цилиндрических образований из β-структур в химотрипсине и в некоторых других трипсиноподобных протеазах. Как видно, каждый цилиндр имеет приближенную внутреннюю симметрию поворотной оси 2-го порядка. Вся молекула лишена какой-либо симметрии. *д* — топология β-структуры в триозофосфатизомеразе [305], включающая восемь параллельных складчатых листов и окружающие спирали. Если пренебречь эллиптичностью сечения полости и различиями спиралей, можно допустить симметрию поворотной оси 8-го порядка.



б



г



д

свертывания структурные домены, сверхвторичные структуры и другие структурные элементы образуются независимо на начальной стадии, а затем агрегируются, они подчиняются правилу ассоциации отдельных единиц и будут стремиться образовывать симметричные расположения.

Поскольку структурные домены, по-видимому, являются единицами свертывания, то симметричное расположение доменов, наблюдаемое в роданезе (рис. 5.17, *а*) и иммуноглобулинах (табл. 5.3), объяснимо. Можно предположить, что в симметрически связанных сверхвторичных структурах домены образуются преимущественно на начальной стадии процесса свертывания, так как они сами по себе достаточно стабильны. Последующая агрегация создает симметрию. Симметричные сверхвторичные структуры обнаружены в виде βαβαβ-звена дегидрогеназы (рис. 5.12, *б* и 5.17, *б*), в виде βαβ-звеньев триозофосфатизомеразы и пируваткиназы (рис. 5.17, *д*), а также в виде β-зигзагов в полости β-структуры Ser-протеаз (рис. 5.17, *г*) и в виде антипараллельных пар α-спиралей в гемеритрине.

Полипептидную цепь ферредоксина (рис. 5.17, *в*) можно разбить на две симметрично связанные половины. Каждая половина цепи свернута вокруг одного FeS-кластера. Кластеры связаны между собой симметрией. Симметричная структура, образованная FeS-кластерами и окружающими их полипептидными цепями, довольно стабильна. Аналогичная ситуация обнаружена в парвальбумине, который образован двумя симметрично связанными структурными элементами, каждый из которых состоит из Ca^{2+}-связывающего центра и двух α-спиралей, расположенных от него в противоположных направлениях [59]. На первый взгляд, такая вытянутая структура сама по себе не должна быть стабильной. Однако симметричное расположение дает основание предположить, что энергия на участке связывания Ca^{2+} достаточна для стабилизации этого структурного элемента в растворе. Таким образом, симметрия позволяет выявить стабильные структурные элементы, а в некоторых случаях указывает на определенную последовательность образования структур в процессе свертывания.

5.4 Глобулярные белки

Глобулярные белки состоят из одного или большего числа структурных доменов. Глобулярным белком принято называть продукт свертывания одиночной полипептидной цепи. В связи с этим глобулярные белки обычно состоят из одного или большего числа структурных доменов. В некоторых случаях (Ser-протеазы, иммуноглобулины, роданеза и др.) домены настолько подобны, что можно предположить дупликацию гена. Каждый домен может быть отнесен к одному из пяти классов, приведенных в табл. 5.2.

Отсутствует предпочтительность между структурными доменами одного глобулярного белка, принадлежащими одному классу (табл. 5.2). Какой-либо заметной корреляции между структурным классом и функцией домена или глобулярного белка пока не обнаружено. Однако нуклеотидсвязывающие домены обычно принадлежат к классу № 4. Кроме того, было отмечено, что все ферменты гликолитического пути, структуры которых известны, относятся к тому же классу № 4.

Обычно активные центры ферментов включают части всех структурных доменов глобулярного белка. Активные центры всех известных мультидоменных белков (табл. 5.2) расположены между доменами (рис. 4.1). Эти домены определяются не только как глобулярные области, разделенные полостью активного центра, но имеют и другое характерное для доменов свойство — они связаны между собой только одной пептидной цепью (табл. 5.2). Субстраты и кофакторы обычно присоединяются к разным доменам. В случае NAD связывающий кофактор домен всегда имеет ту же самую с довольно развитой открытой поверхностью топологию и NAD присоединяется в эквивалентных положениях (рис. 5.17, *б*), что является результатом эволюции [254, 255]. Кроме того, этот домен обнаружен на N-конце трех дегидрогеназ и одной киназы [230—233, 235], а также на С-концевой половине четвертой дегидрогеназы [234] и в средней части фосфорилазы [236], что указывает на возможность дупликации соответствующего гена и его переноса в другое место генома. Все эти факты, включение в активный центр частей различных доменов, наличие кофакторспецифичных доменов и возможность переноса домена дают основание предположить, что ферменты конструируются с использованием модульной системы: кофактор и субстратспецифичные домены, необходимые для обеспечения заданной функции, отбираются и объединяются в одной цепи глобулярного белка [124, 256].

5.5 Агрегаты глобулярных белков

Не существует принципиальных различий между агрегированием структурных доменов и агрегированием глобулярных белков. Различия между ассоциацией структурных доменов одной цепи и ассоциацией глобулярных белков, т. е. разных цепей, весьма расплывчаты. Так, в роданезе (рис. 5.17, *а*) домены одной цепи при агрегировании образуют систему с высокой симметрией [257], как и субъединицы димерного белка. При образовании оболочек вируса полиомиелита [163] и вируса лесов семлики [258] большое число белковых глобул формируется из одной полипептидной цепи, а затем уже разделяется протеазой. После расщепления белок симметрично агрегируется с образованием оболочки вируса. Это показывает, что отдельные домены мультидоменного глобулярного

белка могут быть стабильны сами по себе, а также могут выполнять роль независимых глобулярных белков. Однако в общем случае субъединицы крупных белковых агрегатов синтезируются и свертываются раздельно.

Причин агрегирования довольно много. Среди многообразных функций, выполнению которых способствует агрегирование, можно упомянуть следующие: (1) создание сложного полифункционального механизма; (2) сближение ферментов одного метаболического пути в целях исключения потерь метаболических промежуточных продуктов; (3) построение структур заданной геометрии, например длинных каналов; (4) снижение осмотического давления или (5) увеличение числа возможных проявлений ферментативной активности, а значит и новых типов регуляции.

Симметрия агрегатов

Асимметричные агрегаты выполняют сложные функции. Все агрегаты могут быть подразделены на асимметричные и симметричные. Типичным асимметричным агрегатом (который входит также и в РНК) является рибосома. Функции рибосом в синтезе белка настолько сложны, что для их выполнения необходима большая группа различных белков, а неидентичные белки могут агрегироваться только асимметрично.

Агрегаты могут иметь пространственную или линейную симметрию, а также симметрию точечной группы. Симметричные агрегаты можно разделить на агрегаты, обладающие пространственной или линейной симметрией, а также симметрией точечной группы. Симметрия пространственной группы обнаружена в кристаллах инсулина, которые образуются в поджелудочной железе и обеспечивают форму, которая может сохраняться при пренебрежимо малом осмотическом давлении [259]. Симметрия такого же типа наблюдается в поперечнополосатых мышцах позвоночных и насекомых [215]. Линейные группы были найдены в микрокапиллярах [181], вирусе табачной мозаики [180] и нитевидных фагах [220]. Симметрия точечной группы очень распространена. Симметрия аминокислот исключает точечные группы, содержащие центры инверсии или отражения, так что возможны лишь группы, n, $n2$, 23, 432, 532 при $n = 1, 2, 3 \ldots$ [252, 260]. Примеры всех этих групп, за исключением 23, приведены в табл. 5.4.

Как уже обсуждалось в предыдущих разделах, симметричные образования энергетически предпочтительны. По-видимому, только в случае гексокиназы наблюдается асимметричная агрегация идентичных единиц, и то лишь в кристаллическом состоянии [261]. В некоторых случаях может реализоваться «псевдосимметричное» расположение. Обычный гемоглобин человека, например, имеет идеальную симметрию 2 и псевдосимметрию 222. Псевдо-

Таблица 5.4

Симметрия агрегатов глобулярных белков [a]

Белок	Литература	Число субъединиц	Точечная симметрия	
			кристаллографическое обозначение	обозначение по Шенфлису
Алкогольдегидрогеназа	234	2	2	C_2
Иммуноглобулин	294	4	2	C_2
s-Малатдегидрогеназа	233	2	2	C_2
Пероксид-дисмутаза	286	2	2	C_2
Триозофосфатизомераза	305	2	2	C_2
Фосфорилаза	236	2	2	C_2
Щелочная фосфатаза	317	2	2	C_2
6-Фосфоглюконат-дегидрогеназа	318	2	2	C_2
Агглютинин зерна пшеницы	316	2	2	C_2
Глюкозофосфатизомераза	313	2	2	C_2
Туг-тРНК-синтетаза	221	2	2	C_2
Глутатионредуктаза	124	2	2	C_2
Альдолаза	306	3	3	C_3
Белок бактериохлорофилла	303	3	3	C_3
Глюкагон	278	(3)	3	C_3
ВТМ, диск белка	218	17	17	C_{17}
Конканавалин А	281, 282	4	222	D_2
Глицеральдегид-3-фосфатдегидрогеназа	230, 231	4	222	D_2
Лактатдегидрогеназа	232	4	222	D_2
Проальбумин	206	4	222	D_2
Пируваткиназа	80	4	222	D_2
Фосфоглицератмутаза	307	4	222	D_2
Гемоглобин человека	274	2+2	Псевдо-222	Псевдо-D_2
Инсулин	259	6	32	D_3
Аспартат-транскарбамилаза	319	6+6	32	D_3
Гемеритрин	217	8	422	D_4
Апоферритин	320	24	432	O
Белок оболочки вируса кустистого карликового томата	263	180	532	Y

[a] Во многих случаях при агрегации или вновь образуется [259], или продолжается [206, 281, 282] β-структура.

симметрия оказалась бы точной в случае идентичности β- и α-субъединиц. Если все субъединицы химически идентичны, но имеют две несколько различающиеся конформации, т. е. образующие довольно близкие контакты, то появляется так называемая квазисимметрия. Она часто встречается в сферических вирусах. При наличии квазисимметрии белок одного типа может образовать икосаэдрическую оболочку [262], которая приближенно аппроксимируется сферой. В вирусе кустистого карликового томата субъединица содержит два структурных домена. Конформационное различие вызвано относительной подвижностью этих доменов [263].

Контакты между субъединицами в дегидрогеназах

В 222-тетрамере необходимо образование только двух типов контактов из трех возможных. Рассмотрим типы контактов, которые могут возникать в симметричном олигомере, на примере четырех дегидрогеназ с известной структурой. s-Малат- и алкогольдегидрогеназы представляют собой димеры идентичных субъединиц с точечной группой 2, а лактат- и глицеральдегид-3-фосфатдегидрогеназы — тетрамеры идентичных субъединиц с точечной группой 222 (рис. 5.16, *а*). Димеры образуют контакт только одного типа вблизи поворотной оси 2-го порядка (рис. 5.18, *а*). В тетрамерах имеются три типа контактов, по одному на пересечении с каждой осью второго порядка (рис. 5.18, *б*). Отметим, что для стабилизации тетрамера необходимы только два из трех контактов.

Контакты 222-тетрамера можно представить тригональной диаграммой. Контакты тетрамера можно условно описать тетраэд-

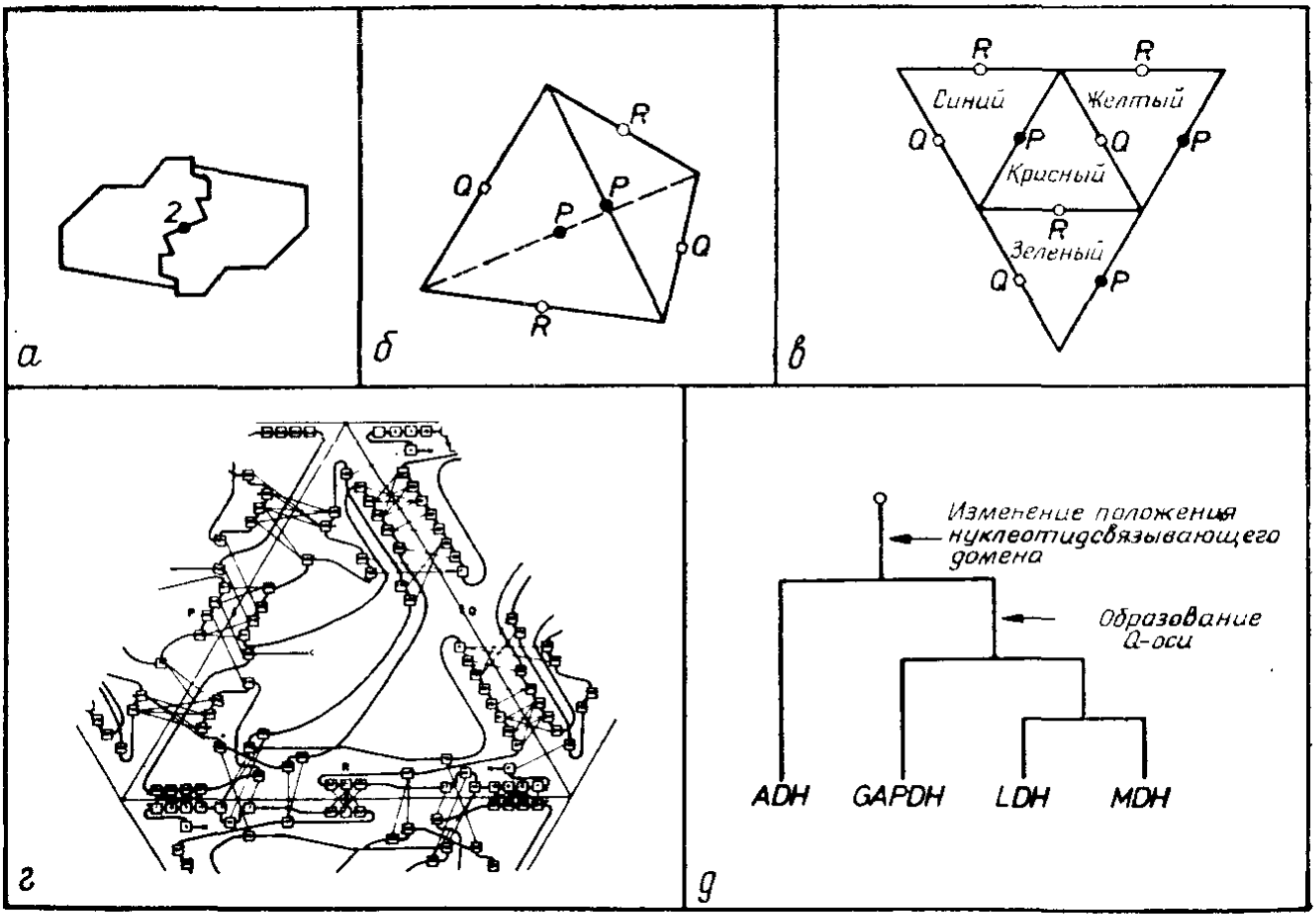


Рис. 5.18.

Олигомерные ферменты — дегидрогеназы.

а — схема симметричного димера, например алкогольдегидрогеназы; имеется поверхность раздела только одного типа. Контактирующие поверхности комплементарны (система ключ—замок [44]). *б* — тетрамер симметрии 222 представлен тетраэдром. Указаны три поворотные оси 2-го порядка: *P*, *Q* и *R*. Каждая субъединица представлена треугольником, а каждый контакт между субъединицами — ребром. *в* — плоское представление тетраэдра и контактов между субъединицами. *г* — использование плоского представления для идентификации атомных контактов между субъединицами лактатдегидрогеназы [264]. *д* — дерево эволюции дегидрогеназ, построенное на основе анализа контактов между субъединицами [264]. ADH — алкогольдегидрогеназа, GAFDH — глицеральдегид-3-фосфатдегидрогеназа, LDH — лактатдегидрогеназа, MDH — *s*-малатдегидрогеназа.

ром, показанным на рис. 5.18, *б*, а также двумерным графиком на рис. 5.18, *в*. Для идентификации субъединиц использован цветовой код, а поворотные оси второго порядка обозначены *P*, *Q* и *R*. «Красная» субъединица считается базовой. Она образует контакты с «голубой», «желтой» и «зеленой» субъединицами, пересекающиеся соответственно осями *P*, *Q* и *R*. Как видно из рис. 5.18, *г*, двумерный график удобно использовать для подсчета взаимодействий между остатками «красной» субъединицы с остатками других субъединиц.

Контакты выявляют эволюционную взаимосвязь. В работах Россмана и сотр. [264, 265] было показано, что четыре дегидрогеназы эволюционно связаны и что сравнение контактов между субъединицами позволяет выявить некоторые детали этой взаимосвязи. Эти контакты можно сопоставить, поскольку все четыре дегидрогеназы имеют одинаковый нуклеотидсвязывающий домен (рис. 5.17, *б*), составляющий приблизительно половину субъединицы. Такой домен можно использовать в качестве базового. Для лактат- и s-малатдегидрогеназы характерно почти одинаковое свертывание цепей,что свидетельствует о родственности этих белков. Глицеральдегид-3-фосфат, лактат- и s-малатдегидрогеназы образуют *Q*-осевой контакт (показано на рис. 5.18, *г*). Соответственно димер «красное» — «желтое» должен был хорошо сохраниться в процессе эволюции белка. Что касается *P*- и *R*-осевых контактов дегидрогеназ, то они не проявляют никакого сходства. Кроме того, нет подобия между контактом в алгокольдегидрогеназе и каким-либо другим контактом*. Построенная на основании этих данных схема эволюционной взаимосвязи приведена на рис. 5.18, *д*.

Поверхности контакта

Чем больше контактов, тем сильнее взаимодействие. Подробный анализ поверхностей контакта белок — белок был проведен для белков, перечисленных в табл. 5.5. Данные этой таблицы, свидетельствуют о наличии большого числа вандерваальсовых контактов, а также об образовании нескольких водородных связей и солевых мостиков. Очевидно, что с ростом числа контактов между субъединицами, увеличивается белок-белковое взаимодействие, что проиллюстрировано следующими четырьмя примерами.

В гемоглобине наиболее сильным является контакт $\alpha_1\beta_1$. Как видно из табл. 5.5, в окси-, а также в дезоксигемоглобине контакт $\alpha_1\beta_1$ ($\alpha_2\beta_2$) сильнее, чем контакт $\alpha_1\beta_2$ ($\alpha_2\beta_1$). Известно, что коо-

---

* Особое положение алкогольдегидрогеназы отражает также то обстоятельство, что ее нуклеотидсвязывающий домен находится на С-конце полипептидной цепи, тогда как в трех других дегидрогеназах этот домен является их N-концевой частью.

Таблица 5.5

Контакты между белками

Белок	Литература	Симметрия	Наименование контакта	Число		
				вандерваальсовых контактов	водородиых связей	солевых мостиков
Оксигемоглобин лошади	321	2	$\alpha_1\beta_1$	110	5	0
			$\alpha_1\beta_2$	80	1	0
			$\alpha_1\alpha_2$	0	0	0
			$\beta_1\beta_2$	0	0	0
Дезоксигемоглобин лошади	322	2	$\alpha_1\beta_1$	98	5	0
			$\alpha_1\beta_2$	69	1	1
			$\alpha_1\alpha_2$	0	0	2
			$\beta_1\beta_2$	0	0	0
Конканавалин А	282	222	I—II	250	14	0
			I—III	156	4	2
			I—IV	16	2	0
Инсулин	259	32	*OP*	111	4	0
			OQ	99	2	1
α-Химотрипсин	18	2	*A*	143	8	1
			B	57	6	0
Комплекс трипсина с ингибитором	267, 269	отсутствует		Много	7	1

ративное изменение гемоглобина при насыщении кислородом состоит в основном в повороте на 13° $\alpha_1\beta_1$-димера относительно $\alpha_2\beta_2$-димера. Контакты $\alpha_1\alpha_2$ и $\beta_1\beta_2$ (в гемоглобине, имеющем точную симметрию 22, контакт $\alpha_1\alpha_2$ должен быть эквивалентен контакту $\beta_1\beta_2$) очень слабы. Как отмечалось выше, этот третий тип контактов в 222-тетрамере не является необходимым для удержания субъединиц вместе.

Контакты часто представлены β-складчатыми листами. В конканавалине А один (I—IV) из трех типов контактов также намного слабее, чем два других. Особенно прочен контакт I—II, включающий 14 водородных связей. В этом случае широкая (шестицепочечная) β-структура пересекает ось второго порядка, так что образуется регулярная β-структура из 12 цепей. Две цепи, связанные пересекающими ось второго порядка водородными связями, антипараллельны, таким образом, эта ось перпендикулярна плоскости листа. Подобные продолжения β-структур на поверхности раздела белок — белок обычны для агрегатов субъединиц и структурных доменов.

Контакты в кристаллах конканавалина А, инсулина и α-химотрипсина помогают в определении структуры димеров в растворе. В растворе тетрамерный конканавалин А находится в равновесии

с димерами и мономерами. Поскольку димер проявляет физиологическую активность, его идентификация представляет интерес, Так как из трех типов димеров наиболее стабильным является, по-видимому, I — II (табл. 5.5), то наиболее вероятно, что именно этот димер существует в растворе.

Аналогичная проблема возникает в случае инсулина. В растворе гексамер инсулина находится в равновесии с димерами и мономерами. Поскольку димер, по-видимому, важен для связывания с рецепторами гормона, его необходимо идентифицировать. Судя по величинам, приведенным в табл. 5.5, прочность обоих возможных димеров не должна сильно различаться. Однако димер *OP* помимо несколько более прочного контакта образует антипараллельную β-структуру. Поэтому можно считать, что в растворе присутствует именно этот димер [259].

Кристаллический α-химотрипсин образует два типа контактов *A* и *B* с осями симметрии 2-го порядка. Контакт *A* намного прочнее контакта *B*. В растворе при значении pH кристаллов (pH 4) α-химотрипсин образует димеры. Данные, приведенные в табл. 5.5, позволяют заключить, что эти димеры скорее всего относятся к типу *A*.

Свободная энергия ассоциации

Свободная энергия ассоциации может быть оценена из структуры контактной поверхности. На основании приведенных в табл. 5.5 точечных взаимодействий можно сделать сравнительную оценку энергии связывания, т. е. свободной энергии ассоциации, и, как показано выше, идентифицировать наиболее стабильные олигомеры. Однако поверхности раздела можно анализировать и более детально и определять абсолютную величину свободной энергии ассоциации $\Delta G_{\text{acc}}^{\text{расч}}$ по кристаллической структуре. С помощью уравнения

$$\Delta G_{\text{acc}}^{\text{эксп}} = G_{\text{олиг}}^{\text{эксп}} - G_{\text{мон}}^{\text{эксп}} = -RT\ln K_{\text{acc}}^{\text{эксп}}$$

рассчитанные величины $\Delta G_{\text{acc}}^{\text{расч}}$ можно затем сравнивать с экспериментальными величинами $\Delta G_{\text{acc}}^{\text{эксп}}$, полученными из константы $K_{\text{acc}}^{\text{эксп}}$. Рассмотрим теперь способы расчета $\Delta G_{\text{acc}}^{\text{расч}}$.

Термодинамика ассоциации сходна с термодинамикой свертывания. В результате ассоциации контактные поверхности субъединиц оказываются скрытыми. Это может быть представлено как перенос поверхности (атомов) из воды во внутреннюю часть белка. Параллельно происходит понижение энтропии системы, поскольку ассоциированные мономеры (олигомеры) характеризуются более высокой упорядоченностью, чем свободные. Следовательно, свободная энергия ассоциации имеет вид:

$$\Delta G_{\text{асс}}^{\text{расч}} = \Delta G_{\text{перенос}} - T \cdot \Delta S_{\text{асс}}$$

Это уравнение можно сравнить с уравнениями (3.2) и (3.3): $\Delta G_{\text{асс}}^{\text{расч}}$ и $\Delta S_{\text{асс}}$ соответствуют $\Delta G_{\text{общ}}$ и $\Delta S_{\text{цепь}}$, величина $\Delta S_{\text{перенос}}$ остается той же. Таким образом, мы можем пользоваться рисунками 3.3,*б* и 3.3, *в*.

Основной движущей силой ассоциации служит энтропия. Для неполярных поверхностей величина $\Delta G_{\text{перенос}}$ пропорциональна площади поверхности, доступной воде (рис. 1.8). Для полярных поверхностей величина $\Delta G_{\text{перенос}}$ меньше, однако остается величиной того же порядка (рис. 3.3). Таким образом, во всех случаях значение $\Delta G_{\text{перенос}}$ примерно пропорционально общей скрытой поверхности. Чотиа и Джанин [266] определили площади скрытой контактной поверхности белков, перечисленных в табл. 5.6. Как видно из этих данных, две трети скрытых поверхностей образованы неполярными группами (табл. 5.6). Поэтому авторы не внесли большой ошибки, пользуясь константами пропорциональности для неполярных поверхностей (рис. 1.8). Для получения величины $\Delta S_{\text{асс}}$ были использованы теоретические данные для свободных частиц [267]. Рассчитанные таким способом значения $\Delta G_{\text{асс}}^{\text{расч}}$ даны в табл. 5.6. Они приблизительно согласуются с экспериментальными данными, что указывает на адекватность представлений о характере, сил, направляющих процесс ассоциации.

Таблица 5.6

Свободная энергия ассоциации белка [a]

Белок	Общая доступная воде поверхность, экранированная при ассоциации, Å^2	Доля неполярной поверхности, %	$\Delta G_{\text{перенос}}$, ккал/моль	$-T \cdot \Delta S_{\text{асс}}$, ккал/моль	$\Delta G_{\text{асс}}^{\text{расч}}$, ккал/моль	$\Delta G_{\text{асс}}^{\text{эксп}}$, ккал/моль
Оксигемоглобин лошади, контакт $\alpha_1\beta_1$	1720	84	—43	+27	—16	<-11
Инсулин, контакт *OP*	1130	78	—28	+23	—5	—7
Комплекс трипсина с ингибитором	1390	68	—35	+27	—8	—18

[a] Площадь поверхности определена по даиным о кристаллической структуре, взятым из работы [266]. Площадь доступной поверхности определена на рис. 1.9.

Наибольший интерес представляет изменение свободной энергии, сопровождающее конформационные изменения агрегирующейся структуры. Во многих олигомерах абсолютная свободная энергия ассоциации значительно менее значима, чем ее изменение при

конформационных перестройках, вызванных присоединением лиганда. Основываясь на таких изменениях свободной энергии, можно попытаться объяснить явления кооперативного действия и регуляции. На сегодняшний день наиболее хорошо известным примером индуцированных лигандом изменений агрегированной структуры является гемоглобин. Сравнение скрытых поверхностей его окси- и дезоксиформ (табл. 5.5) позволяет выявить вклад свободной энергии при изменении агрегированной структуры [268]. Следует учесть, однако, что в этом случае мы имеем дело с небольшой разностью между двумя большими величинами; для ее оценки требуется высокая точность, которую пока довольно трудно достичь.

Специфичность взаимодействия белок — белок

Комплементарность поверхностей необходима. Суммируя результаты расчетов предыдущего раздела можно сделать следующие выводы в отношении взаимодействий белок — белок. Прежде всего мы убедились, что удаление от воды неполярных поверхностей дает наибольший выигрыш свободной энергии (табл. 5.6); кроме того, важнейший вклад вносит энтропийный фактор. Абсолютная величина такого вклада зависит от комплементарности контактирующих поверхностей: площадь контактной поверхности заметно уменьшается, если имеется возможность доступа к ней воды. Плотность упаковки (разд. 3.6) на всех изученных в настоящее время таких поверхностях раздела столь же велика, как и во внутренней части белка [267].

Поверхность раздела обнаруживает сходство с внутренней частью белка и в отношении водородных связей, которые образуются почти всеми ее полярными группами [266]. Кроме того, все заряженные группы связаны солевыми мостиками. Напомним, что рис. 3.3, по которому величина $\Delta G_{\text{перенос}}$ полярных групп приблизительно равна $\Delta G_{\text{перенос}}$ неполярных групп, основан на данных об образовании эффективных водородных связей и солевых мостиков во внутренней части белка.

Специфичность достигается комплементарностью профилей контактных поверхностей, а также структурным соответствием доноров и акцепторов водородных связей и остатков, образующих солевые мостики. Для биологических процессов важна не только эффективность взаимодействия, но и его специфичность. Каким образом достигается специфичность во взаимодействиях белок — белок? Как отмечалось выше, для получения большого значения величины $\Delta G_{\text{перенос}}$ контактирующие поверхности должны быть комплементарны. Кроме того, должны быть пространственно сближены и соответствующим образом ориентированы доноры и акцепторы водородных связей; в противном случае энергия ассоциации резко ослабляется. В еще большей мере это относится к скрытым солевым

мостикам. Несмотря на то что их «активный» вклад мал, им свойственно большое «пассивное» влияние: если заряды не скомпенсированы, энергия связывания ослабляется настолько сильно (разд. 3.5), что белок теряет способность к ассоциации. Таким образом, специфичность поверхности определена ее формой, характером доноров и акцепторов водородных связей, а также зарядами. Если судить по числу различных мест присоединения антигена в иммуноглобулинах, поверхность площадью 100 Å^2 может принимать порядка тысячи конформаций, проявляющих различные связывающие свойства.

5.6 Иерархия уровней

Иерархическая взаимосвязь облегчает анализ. Уровни структурной организации белка и схема их иерархии приведены на рис. 5.1. Если следовать схеме, анализ процесса свертывания значительно упрощается, так как он распадается на несколько четко разделенных стадий (это можно сопоставить с факторизацией при решении дифференциальных уравнений). Так, на первой стадии по аминокислотной последовательности должна быть установлена вторичная структура, на второй — по вторичным структурам определена сверхвторичная структура и так далее. Однако как ни заманчиво решать проблему свертывания белка таким образом, реальная ситуация намного сложнее и предложенная схема применима, по-видимому, лишь в немногих случаях.

Иерархия уровней структуры белка во многих случаях не соблюдается. Иерархия уровней установлена на примере глобулярных белков и их агрегатов. Действительно, в большинстве случаев глобулярный мономер достаточно стабилен и почти не изменяется при агрегировании*, в связи с чем комплексная структура по существу определяется характеристиками поверхности известной мономерной структуры. Такой подход был применен при исследовании серповидно-клеточного гемоглобина [270]. Однако иерархические зависимости неприменимы к рибосомам. Рибосомные белки настолько вытянуты [7], что структура каждой отдельной единицы не стабильна и стабилизируется только в процессе агрегирования.

Иерархию проявляют также уровни, соответствующие аминокислотной последовательности и вторичной структуре (в особенности α-спиралей), поскольку образование вторичной структуры сильно зависит от конечного свертывания цепи. Это заключение следует из наблюдаемых корреляций между аминокислотной последовательностью и вторичной структурой, описанных в следующей главе.

---

* Подробный анализ небольших изменений при ассоциации трипсина и панкреатического трипсинового ингибитора быка проведен Хубером и стр. [269].

Ассоциация доменов в мультидоменных глобулярных белках также, по-видимому, имеет иерархическую природу. В частности, преобладающее одноцепочечное сочленение последовательно расположенных доменов и наблюдаемое симметричное расположение доменов указывают на то, что домены образуются до их ассоциации в белковой глобуле и что они сами по себе весьма стабильны.

Менее ясен переход от вторичной структуры к сверхвторичной структуре и от сверхвторичной структуры к доменам. Из всех вторичных структур лишь α-спирали достаточно стабильны и могут полностью формироваться до последующей стадии ассоциации. Однако единичные β-складчатые листы и реверсивые повороты цепи не обладают такими свойствами. Что касается сверхвторичных структур, то обнаруженная в ряде доменов симметрия позволяет предполагать, что сверхвторичные структуры сами по себе достаточно стабильны и образуют отдельный уровень организации белков.

Заключение

Анализ известных белковых структур дает ценные сведения для понимания механизма свертывания и стабильности белков. В структурах этих белков обнаруживаются шесть уровеней организации. На первом уровне находится аминокислотная последовательность, которая целиком определяет окончательную структуру белка. В структурах белков можно выделить несколько типов упорядоченности формы основной цепи. Это так называемые вторичные структуры, которые составляют второй уровень. Две из таких регулярных структур (α-спираль и β-складчатый лист) были предсказаны на основе ковалентного строения основной цепи как наиболее простые. Следующие два уровня, сверхвторичные структуры и структурные домены, гораздо более сложны и пока не предсказуемы. На этих уровнях также проявляются вполне определенные закономерности, например такие, как корреляция между близкими по цепи остатками. Эти закономерности не выражаются в каких-либо определенных структурах, а носят весьма общий характер. На двух самых высоких уровнях организации, занимаемых глобулярными белками и агрегатами, сейчас уже делаются попытки некоторых структурных предсказаний. Возможность таких предсказаний основана на том, что нижние структуры, домены для глобулярных белков и глобулярные белки для агрегатов предполагаются внутренне стабильными (в некоторых случаях это подтверждено экспериментом). Характер агрегатов можно предсказать с помощью анализа контактной поверхности глобулярных белков. Это же относится и к предсказаниям строения глобулярных белков по их доменам. Кроме того, свойства поверхности, как это следует из изучения поверхностей раздела белок — белок, имеют важное значение для белкового узнавания. В главе обсуждены некоторые законо-

мерности такого узнавания. Общим свойством, проявляющимся на всех уровнях организации выше аминокислотного, является симметрия. Во вторичных структурах и в агрегатах найдены линейные группы симметрии. Симметрия точечных групп проявляется на всех уровнях выше вторичной, а симметрия пространственных групп — в агрегатах. Наличие симметрии и иерархии уровней белковых структур позволяет более ясно представить некоторые аспекты сложнейшей проблемы пространственной организации белковых молекул.

Глава 6 Предсказание вторичной структуры по аминокислотной последовательности

В первом приближении вторичная структура сегмента цепи является функцией входящих в его состав аминокислот. Число известных аминокислотных последовательностей намного превышает число известных трехмерных структур (см., например, Атлас Дейхофа [20]). Поскольку в порядке расположения аминокислот заложена полная информация о структуре белка (разд. 1.2, а также работа [177]), то в принципе можно установить пространственное строение по аминокислотной последовательности без рентгеноструктурного анализа. Первым шагом в этом направлении могло бы послужить более тщательное исследование взаимосвязи между двумя низшими уровнями структурной организации белков (рис. 5.1), т. е. между аминокислотной последовательностью и вторичной структурой. Хотя такая организация и носит иерархический характер, однако иерархия здесь далеко не однозначна в том смысле, что образование вторичной структуры в данном сегменте полипептидной цепи зависит не только от аминокислотной последовательности сегмента, но и от влияния других сегментов, удаленных от исходного по цепи*.

Первые данные о взаимосвязи получены на синтетических гомополимерах. Первая корреляция между аминокислотной последовательностью и вторичной структурой была установлена Блоутом и сотр. [328]. На основе опытов с синтетическими гомополимерами — полипептидами, поли-Glu, поли-Lys и т. д., они исследовали способность к образованию и к нарушению α-спиральной структуры аминокислотных остатков семи типов (табл. 6.1). К спиралеобразующим были отнесены остатки, которые принимали α-спиральную конформацию, а к спираленарушающим — те, которые такую конформацию не принимали.

---

* В таком аспекте рассматриваемая функциональная зависимость в известной мере сходна с преобразованием Фурье, в котором все точки основного пространства вносят вклад в данную точку преобразованного пространства, и наоборот.

Дэвис [329] применил эти данные к нативным глобулярным белкам и обнаружил явную антикорреляцию между содержанием спиралей, найденным измерениями дисперсии оптического вращения (ДОВ) [330] и содержанием аминокислотных остатков (Ser + Val + + Cys + Thr + Ile). Как видно из табл. 6.1, первые три из них, по Блоуту и сотр. [328], являются спиралеразрушающими остатками. Остатки Thr и Ile были выбраны в связи с их сходством соответственно с Ser и Val.

Таблица 6.1

Тенденция встраиваться в α-спираль[a]

		Блоут и сотр., 1960 [328]	Котельчук и Шерага, 1968 [363]	Льюис и сотр., 1970 [368]	Робсон и Пэйн, 1971 [346]	Чоу и Фасман, 1974 [340]	Финкельштейн и Птицын, 1976 [371]
A	Ala	(H)	H	I	+0,09	1,45	1,08
C	Cys	C	H	I	+0,03	0,77	0,95
D	Asp	H	C	B	—0,02	0,98	0,85
E	Glu	H	H	H	+0,12	1,53	1,15
F	Phe	(H)	H	H	+0,03	1,12	1,10
G	Gly	—		B	—0,05	0,53	0,55
H	His	(H)	H	I	+0,08	1,24	1,00
I	Ile	(C)	H	H	+0,07	1,00	1,05
K	Lys	(H)	C	I	—0,03	1,07	1,15
L	Leu	H	H	H	+0,11	1,34	1,25
M	Met	H	H	H	+0,10	1,20	1,15
N	Asn	(C)	C	I	—0,4	0,73	0,85
P	Pro	—		B	—	0,59	—
Q	Gln	(H)	H	I	+0,07	1,17	0,95
R	Arg	(H)	H	I	+0,02	0,79	1,05
S	Ser	C	C	B	—0,07	0,79	0,75
T	Thr	(C)	H	I	—0,01	0,82	0,75
V	Val	C	H	I	0,04	1,14	0,95
W	Trp	(H)	C	H	+0,10	1,14	1,10
Y	Tyr	(H)	C	H	—0,02	0.61	1,10

[a] Эти свойства найдены по наблюдаемым частотам встречаемости. Аминокислотные остатки даны в алфавитном порядке в соответствии с однобуквенными символами. Символы «H», «I» и «B» использованы для обозначения соответственно спиралеобразующих, индифферентных и нарушающих спираль. Буквой «C» обозначен беспорядочный клубок.

Корреляционный анализ расширяется по мере получения новых экспериментальных данных. Первоначальный успех Дэвиса стимулировал дальнейшие исследования корреляций в глобулярных белках. Однако перспективы таких работ оставались ограниченными до тех пор, пока оставалась малой экспериментальная основа, т. е. число известных трехмерных белковых структур. Поэтому ранние результаты Гуццо [331], Протеро [332], Хавенстина [333], Кука [334], Перити и сотр. [335], Данила [336] и Лоу и сотр. [337] представляют в основном исторический интерес. По мере расширения экспериментальной базы точность корреляции существенно улучшается.

Предлагаемые корреляционные методы, или, как их обычно называют, «методы предсказания вторичной структуры по аминокислотной последовательности» (сокращенно «методы предсказания»), можно разбить на две категории: вероятностные и физико-химические. К первой категории относятся методы, устанавливающие закономерности исключительно на основе статистического анализа исходных рентгеноструктурных данных. Физико-химические методы используют дополнительно (или исключительно) иную структурную информацию. Очевидно, что различие между этими двумя категориями не может быть очень четким и возможны промежуточные случаи, отнесение которых к первой или второй категории произвольно.

В этой главе будут обсуждены все применяющиеся в настоящее время методы, поскольку ни один из современных подходов, по-видимому, не имеет перед другими явных преимуществ. С другой стороны, даже методы, которые представляются сейчас совершенно неудовлетворительными, могут содержать идеи, полезные для дальнейшего усовершенствования подхода в целом.

6.1 Вероятностные методы

Склонность остатка встраиваться во вторичную структуру

Простейший подход состоит в рассмотрении каждого остатка отдельно, без учета его ближних или дальних соседей. Простейший, чисто статистический подход был применен Дирксом [338, 339]. Он определил частоту встречаемости каждого из 20 остатков в α-спиралях, β-структурах и реверсивных поворотах цепей (*rt*), используя кристаллографические структуры белков. Эти частоты были затем приняты в качестве характеристик, отражающих «склонности остатков встраиваться в α-, β- и *rt*-конформации». В качестве примера в табл. 6.2 приведены данные Диркса, касающиеся тенденции остатков встраиваться в β-структуру.

Для предсказания α- (β-, *rt*-) структуры остатка в каждом положении цепи рассчитывался «потенциал* α-(β-, *rt*-)» как средневзвешенное α-(β-, *rt*-)-склонностей m ближайших по цепи остатков. Как только потенциал α-(β-, *rt*-) превышал некоторое пороговое значение, для остатка предсказывалась α-(β-, *rt*-)-конформация. Таким образом, такая потенциальная функция была предназначена для ответа на вопрос — «да» или «нет». Схема определения средневзвешенного фиксировалась. Три пороговые величины, а также значение m были выбраны по наилучшему совпадению предсказанной и наблюдаемой вторичной структуры в исходном наборе. Оптималь-

* Термин «потенциал» здесь использован, разумеется, не в физическом, а в более широком смысле. — *Прим. ред.*

Таблица 6.2

Относительная частота встречаемости остатков в β-структурах и реверсивных поворотах *rt*[a]

		Склонность к β-структуре			*rt*-склонность		
		Чоу и Фасман, 1974 [340]	Бэджис и сотр., 1974 [31]	Бегин и Диркс, 1975 [339]	Льюис и сотр., 1971 [326]	Кунтц, 1972 [203]	Чоу и Фасман, 1974 [340]
A	Ala	0,97	0.29	0,37	0,22	(T)	0,15
C	Cys	1,30	0,53	0,84	0,20	—	0,31
D	Asp	0,80	0,27	0,97	0,73	T	0,33
E	Glu	0,26	0,26	0,53	0,08	T	0,12
F	Phe	1,28	0,32	0,53	0,08	—	0,19
G	Gly	0,81	0,31	0,97	0,58	T	0,45
H	His	0,71	0,20	0,75	0,14	T	0,18
I	Ile	1,60	0,41	0,37	0,22	—	0,15
K	Lys	0,74	0,27	0,75	0,27	T	0,27
L	Leu	1,22	0,40	0,53	0,19	—	0,14
M	Met	1,67	0,38	0,64	0,38	—	0,18
N	Asn	0,65	0,23	0,97	0,42	T	0,45
P	Pro	0,62	0,34	0,97	0,46	T	0,41
Q	Gln	1,23	0,33	0,64	0,26	T	0,15
R	Arg	0,90	0,36	0,84	0,28	T	0,27
S	Ser	0,72	0,35	0,84	0,55	T	0,41
T	Thr	1,20	0,39	0,75	0,49	T	0,27
V	Val	1,65	0,50	0,37	0,08	—	0,08
W	Trp	1,19	0,23	0,97	0,43	—	0,30
Y	Tyr	1,29	0,43	0,84	0,46	T	0,33

[a] Наблюдаемые частоты встречаемости использованы в качестве характеристик склонности. В номенклатуре Кунтца T означает, что соответствующий остаток стремится встраиваться в повороты цепи.

ные значения m равны 17, 11 и 3 соответственно для α-, β- и *rt* предсказаний. Понятие «склонность остатка», имеющее основное значение, и следуемое из него понятие «потенциал остатка» будут использоваться в этой главе именно в определенном выше смысле.

В реверсивных поворотах цепи следует учитывать относительное положение остатка. Информация только об отдельных остатках (синглетах) была использована также для предсказания реверсивных поворотов. Льюис и сотр. [326] определили реверсивные повороты пептидной цепи как фрагменты из четырех остатков i, $i+1$, $i+2$, $i+3$, в которых расстояние между C_α-атомами в положениях i и $i+3$ меньше 7Å, а цепь при этом не находится в α-спиральной конформации. Однако в отличие от Диркса четыре положения в повороте не считались эквивалентными. Частоты встречаемости, определенные для каждого остатка в положениях i, $i+1$, $i+2$ и $i+3$, были названы «склонностью данного остатка встречаться в данном месте поворота». *rt*-Потенциал квартета остат-

ков затем определялся как произведение соответствующих склонностей. Пороговое значение потенциала было установлено по наилучшему совпадению с известными экспериментальными данными (рис. 6.2).

Подобные частоты встречаемости использованы Кроуфордом и сотр. [200] и позднее Чоу и Фасманом [340] для предсказания реверсивных поворотов цепи, идентифицированных согласно данным Льюиса и сотр. [326]. В табл. 6.2 приведены частоты встречаемости, найденные Льюисом и сотр. [326], а также Чоу и Фасманом [340]. Однако чтобы придать таблице компактный вид, в ней перечислены только усредненные значения частот встречаемости остатков в реверсивных поворотах цепей. Для некоторых остатков различия положений заметны: Pro находят почти исключительно в положении $i + 1$, а Trp — только в положении $i + 3$.

Простой подход с использованием синглетных склонностей применили также Птицын и Финкельштейн для α- и β-предсказаний [341, 342]. По утверждению этих авторов [343] никакой дополнительной информации из встречаемостей пар остатков (дублетов) получить нельзя. Однако это положение спорно, поскольку использованные исходные данные касались всего 9 белков и поскольку разбиение 20 остатков на четыре категории было произведено довольно произвольно.

Склонность двух остатков одновременно встраиваться во вторичную структуру

Склонности дублетов включают взаимодействия остаток — остаток. Строгое предсказание основано на частотах встречаемости дублетов. Вопреки Финкельштейну и Птицыну [343] многие методы предсказания используют информацию о дублетах, поскольку такая информация отражает взаимодействие между близкими по цепи остатками. Перити [344] воспользовался дублетами для предсказания α-спиралей и β-структур чисто вероятностным методом. Он рассмотрел 27 дублетов, в каждом из которых остатки находились в положениях $i \pm 1$, $i \pm 2$, ..., $i \pm 6$, полагая, что при более далеких расстояниях между остатками взаимодействия отсутствуют. Всего было получено 10 800 дублетов различных типов*. Для каждого остатка Перити рассматривал возможность реализации состояний α, β и клубка (состояние, не являющееся ни α и ни β). Таким

---

* Для остатка в положении i имеются семь дублетов вида $(i, i + 6)$, $(i - 1, i + 5)$, ..., $(i - 6, i)$, шесть дублетов вида $(i, i + 5)$, ..., $(i - 5, i)$,... и два дублета $(i, i + 1)$, $(i - 1, i)$, что в сумме составляет $7 + 6 + 5 + 4 + 3 + 2 = 27$ различных дублетов. Каждый дублет состоит из двух остатков. Для 20 различных типов остатков возможны $20 \cdot 20 = 400$ различных комбинаций. Таким образом, число дублетов различных типов составляет $27 \cdot 20 \cdot 20 = 10\,800$.

образом, по экспериментальному набору данных была составлена таблица из 32 400 = 3 · 10 800 частот встречаемости, которые были интерпретированы как склонности. Для предсказания вторичной структуры данного остатка в конкретной аминокислотной последовательности по таблице оценивались и комбинировались по стандартной статистической методике склонности к α- (β-, клубок-) структуре 27 типов дублетов. Таким путем был получен α-(β-, клубок-)потенциал. Затем α-(β-, клубок-)потенциалы сопоставлялись и самый высокий из них выбирался для определения состояния. Следует отметить, что этот метод не содержит изменяющихся параметров. Кроме того, он не позволяет отдельно предсказать α- и β-структуры, а объединяет их довольно искусственным способом.

Неизменные потенциалы и варьируемые пороги. Для предсказания α-, β- и *rt*-конформаций Робсон и сотр. [345—352] использовали синглеты совместно с 16 дублетами, образованными анализируемым остатком и остатками, находящимися в положениях $i \pm 1 \div i \pm 8$. Склонности синглетов и дублетов были определены по соответствующим частотам встречаемости в базовом наборе, а затем преобразованы в α-, β- и *rt*-потенциалы для данного положения остатка с помощью стандартных методов теории информации. Эти потенциалы сопоставлялись с тремя различными порогами, по одному на каждый тип вторичной структуры. Была проведена также раздельная коррекция всех трех порогов по лучшему соответствию между предсказываемой и наблюдаемой в базовом наборе вторичной структурой. Предсказания α-, β- и *rt*-структур проводились независимо одно от другого.

Варьируемые потенциалы и варьируемые пороги. Нагано [228, 353—356] использовал те же дублеты, что и Перити [344], с тем различием, что в качестве максимального расстояния было принято не 6, а 7 остатков. Это дало 35 дублетов, влияющих на анализируемый остаток, поэтому таблица наблюдаемых частот встречаемости (склонностей) содержала 20 · 20 · 35 · 3 = 42 000 значений. Таким путем были определены склонности всех типов дублетов данного остатка конкретной аминокислотной последовательности, а затем их линейной комбинацией были получены α-, β- и *rt*-потенциалы.

Аналогично Робсону и сотр., Нагано обосновывал свое предсказание на раздельном сопоставлении потенциалов с тремя разными порогами. Каждый порог подгонялся до достижения наилучшего соответствия в базовом наборе. Кроме того, задавая потенциалы в виде линейных комбинаций, содержащих 3·35 = 105 коэффициентов, по три (α, β, *rt*) на каждый из 35 типов дублетов, Нагано ввел дополнительно большое число подобранных параметров. Все эти параметры варьировались до получения наилучшего соответствия в базовом наборе. Таким образом, Нагано использовал большее число подобранных параметров, чем Перити, а также Робсон

и сотр., которые определяли свои потенциалы по склонностям с использованием стандартных вероятностных процедур.

Нагано также расширил использующийся базовый набор путем добавления (с малым весом) информации, полученной из аминокислотных последовательностей белков, гомологичных белкам с известной структурой, и в своей более поздней работе [356] он предпринял попытку учета частот встречаемости и склонности триплетов. Следует отметить, что в методе Нагано в неявном виде включались и синглеты, поскольку некоторые линейные комбинации склонностей дублетов эквивалентны склонности синглета.

Предсказания остатков, прерывающих вторичную структуру. Воспользовавшись только одним типом дублетов, а именно парой $(i-1, i+1)$ двух соседних остатков анализируемого остатка i, Кабат и Ву [357—359] попытались развить метод «отрицательного» предсказания, с помощью которого можно было бы идентифицировать остатки, прерывающие α-спираль и β-структуру. Склонности к α- и β-нарушениям были получены из частот встречаемости в базовом наборе. Потенциал рассматривался как характеристика соответствующей склонности остатка. Пороги потенциалов подбирали по наилучшему совпадению с имеющимися экспериментальными данными.

Склонность трех остатков встраиваться во вторичную структуру

Попытка предсказать общее свертывание цепи группы родственных белков с помощью триплетов. Кабат и Ву использовали триплеты соседних в последовательности остатков для предсказания не только вторичной структуры [360], но также и свертывания всей цепи [361, 362]. Для предсказания укладки цепи углы (ϕ, ψ) центральных остатков всех триплетов в белках с известными трехмерными структурами, образующими базовый набор, были записаны в виде «таблицы (ϕ, ψ)», которая включала $20 \cdot 20 \cdot 20 = 8000$ позиций, по одной на каждый тип триплета.

Метод был применен к предсказанию общего свертывания цепи группы родственных белков (цитохромы *c* или иммуноглобулины, разд. 9.1). В такой группе многие аминокислотные последовательности характеризуются аналогичной укладкой цепи, т. е. идентичными углами (ϕ, ψ). Поэтому для данного остатка в положении i можно составить триплеты $(i-1, i, i+1)$ по всем гомологичным белкам, а затем с помощью таблицы (ϕ, ψ) получить их углы $\phi\psi$*. Средние из этих углов (φ, ψ) выбираются в качестве предсказате-

* Рассмотрим следующий гипотетический случай: в четырех родственных белках имеются следующие последовательности в триплетных позициях $(i-1, i, i+1)$: Ala-Gly-Thr, Val-Gly-Pro, Leu-Ala-Ser и Leu-Ala-Thr. Для каждого типа триплета соответствующая позиция таблицы включает набор углов ϕ, ψ, а именно набор частот встречаемости рассматриваемого

льных для остатка i. Предсказанное свертывание цепи получают, повторяя эту процедуру для всех остатков.

6.2 Физико-химические методы

Физико-химические методы в отличие от вероятностных основываются не только на известных корреляциях между аминокислотной последовательностью и структурой в базовом наборе глобулярных белков, а используют также другие экспериментальные и теоретические данные. Эти методы можно подразделить на применяющие стереохимический анализ и статистическую механику. В Приложении рассмотрены некоторые важные аспекты статистики полипептидных цепей, необходимые для понимания методов статистической механики и физических основ укладки цепи.

Методы, основанные на статистической механике

Расчеты энергии

Статистические веса можно рассчитать на основании геометрии и вандерваальсовых сил. Идея использования для предсказания методов статистической механики наиболее широко разрабатывалась Шерагой и сотрудниками. Котельчук и Шерага [363, 364] основывали свои предсказания на упрощенном методе, описанном в приложении при выводе уравнения (А.4). Они рассчитали статистические веса $z^{\alpha R}$, $z^{\alpha L}$ и z^{ε} для каждого остатка* кроме Gly и Pro, которые были рассмотрены отдельно. В расчетах они учитывали вандерваальсовы взаимодействия (гл. 3) между различными частями цепи. Поскольку рассматривалась система, состоящая из цепи и растворителя, был включен также член, учитывающий вклад свободной энергии растворителя. Для каждого остатка конформация с наибольшей величиной z принята за его склонность (табл. 6.1). Поскольку предпочтения состояния α_L обнаружено не было, то выявились только «спиральные» ($\equiv\alpha_R$) и «случайные» (клубок $\equiv$ $\equiv$не α_R) склонности. Отметим, что сделанные с помощью этих расчетов оценки хорошо согласуются с экспериментальными данными по синтетическим полипептидам [328] и с эмпирическими данными,

---

типа триплета в базовом наборе. Затем средние значения всех углов ϕ, ψ, найденные в четырех позициях таблицы, используются в качестве предсказанных величин (ϕ, ψ) для положения i-го остатка цепи, свернутой общим для данной группы белков образом.

* Статистические веса $z^{\alpha R}$, $z^{\alpha L}$ и z^{ε} — значения вероятности того, что данный остаток принимает в основной цепи конформацию, находящуюся соответственно в α_R-, α_L- и ε-(растянутая цепь)-областях. Эти области определены на рис. 2.3.

полученными из рассмотрения базового набора по глобулярным белкам (табл. 6.1).

Расчет спиральных потенциалов по склонностям может учитывать кооперативность. Исходя из склонностей, Котельчук и Шерага [364] предложили простой алгоритм предсказания, согласно которому спирали инициируются в том случае, когда располагаются подряд четыре остатка со склонностью к спирали; спирализация в сторону С-конца продолжается до тех пор, пока процесс не затормаживается двумя последовательно расположенными остатками со склонностью к беспорядочному свертыванию. Направленное наращивание явилось следствием расчетной схемы, согласно которой влияние данного остатка на соседа осуществлялось только в направлении С-конца, но не N-конца. Использование такого алгоритма предполагает отказ от простой схемы невзаимодействующих остатков уравнения (А.4). Вводятся инициация спирали и ее роста и тем самым в неявном виде учитывается взаимодействие между соседними остатками.

Идея инициации спирали четырьмя остатками и рост спирали в одном направлении вплоть до появления двух последовательных остатков с тенденцией встраиваться в беспорядочный клубок была также использована Леберманом [365], который несколько модифицировал методику для получения лучшего соответствия в базовом наборе.

Эмпирическая функция распределения

Склонности остатков образовывать спирали, основанные на частотах встречаемости и найденные с помощью эмпирических функций, эквивалентны. Поскольку объем данных о структуре глобулярных белков в настоящее время уже достаточно велик, становится возможным нахождение статистических весов $z^{\alpha R}$ и z^{ε} по уравнению (А. 4). Для этой цели Бэджисом и сотр. [31] уравнение (А.4) было представлено в виде:

$$z^{x} = \int_{x} \left[\int \exp \frac{-E(\phi, \psi, \varkappa)}{RT} \, d\varkappa \right] d\phi d\psi \qquad (6.1)$$

где $\varkappa = \alpha_R$ или ε; очень малой величиной $z^{\alpha L}$ пренебрегали. Интеграл в скобках представляет собой вероятность того, что у рассматриваемого остатка будут наблюдаться данные углы ϕ и ψ на карте (ϕ, ψ) (рис. 2.3). Это есть не что инсе, как распределение Ельцмана, усредненное по всем ориентациям боковой цепи. Поскольку функция $E(\phi, \psi, \varkappa)$ не известна с достаточной точностью, рассчитать распределение вероятностей невозможно. Однако они должны коррелироваться с наблюдаемыми частотами встречаемости остатков данного типа на карте (ϕ, ψ) (рис. 2.4). Поэтому выражение в скобках

можно удовлетворительно описать непрерывной функцией ϕ и ψ, подогнанной к наблюдаемым частотам встречаемости. Такие эмпирические функции могут быть затем использованы для расчета интеграла (6.1) по α_R- и ε-областям, что дает вероятность или частоту встречаемости α_R и ε-конформаций $z^{\alpha R}$ и z^{ε}. Полученные таким способом величины z^{ε} приведены в табл. 6.2*. Однако основанный на статистической механике подход к определению склонности остатка принимать то или иное конформационное состояние в действительности близок к чисто статистическому анализу.

Описанный выше способ не позволяет определять склонности к поворотам цепи, поскольку при соответствующих конформациях углы (ϕ, ψ) остатков не находятся в определенной области, а рассеяны по карте (ϕ, ψ). Для этого случая Бэджис и сотр. [31] избрали статистический подход Льюиса и сотр. [326], использовав расширенный базовый набор.

Склонности нонапептидов были использованы для нахождения потенциала центрального остатка встраиваться в спираль. При предсказаниях вторичной структуры Шерага и сотр. [31, 366] рассмотрели комбинации склонностей остатков от $i-4$ до $i+4$ и применили их для расчета α_R-, ε- и rt-потенциалов (см. определение в разд. 6.1) остатка i. Причина рассмотрения именно девяти остатков связана с тем обстоятельством, что, согласно более ранним энергетическим расчетам [367], нонапептид представляет как раз тот фрагмент, который существенно определяет конформацию центрального остатка. Затем для всех сегментов, включающих более четырех остатков, была предсказана конформация спирали ($\equiv\alpha_R$), если α_R -потенциал был выше, чем ε- и rt-потенциалы, а также больше порогового значения. Этот порог был определен как средневзвешенное по всем α_R -потенциалам цепи. Таким же способом были сделаны ε- и rt-предсказания.

Модель Зимма — Брэгга в применении к гетерополимерам

Модель Зимма — Брэгга применима к природным белкам. Алгоритм предсказания спирали, основанный на модели Зимма — Брэгга, был предложен Льисом и сотр. [368, 369]; при этом в уравнении (А.9) учитываются индивидуальные относительные статистические веса s_i для каждого положения i цепи в отличие от равенства всех величин s_i в уравнении (А.8) для гомополимера. Вероятность P_i нахождения остатка в положении i в α-конформации задается выражением:

$$P_i = \frac{\partial \ln Z}{\partial \ln s_i} = \frac{s_i}{Z} \cdot \frac{\partial Z}{\partial s_i} \qquad (6.2)$$

* В табл. 6.2 ε-склонности обозначены как β-склонности. Шерага и сотр. отметили, что можно предсказать только ε-, а не β-структуру, поскольку конформация остова не определяет полностью водородных связей.

с помощью которого из всех возможных 2^N конформаций цепи отбираются только те, которые являются спиральными в положении i (уравнение (А.6)). Нормирование выполняется, как обычно, путем деления на Z, характеризующую распределение возможных конформационных состояний полипептидной цепи (разд. А.1). В качестве примера получим P_1 для случая $N = 2$ (см. уравнение (А.7)):

$$P_1 = \frac{s_1}{Z} \cdot \frac{\partial Z}{\partial s_1} = \frac{s_1}{Z} \cdot \frac{\partial(1 \cdot 1 + 1 \cdot s_2 + s_1 \cdot \sigma \cdot 1 + s_1 \cdot s_2)}{\partial s_1} =$$

$$= \frac{s_1 \cdot \sigma + s_1 \cdot s_2}{1 + s_2 + s_1 \cdot \sigma + s_1 \cdot s_2}$$

Остатки с низкими относительными статистическими весами значительно укорачивают среднюю длину спирали. Чтобы оценить спиральный потенциал данного белка, было использовано одно значение параметра инициации $\sigma = 5 \cdot 10^{-4}$ (разд. А.4). Кроме того, были введены три различные значения s для всех типов остатков. Так, $s = 0{,}385$ соответствовало остаткам, прерывающим спираль (В), $s = 1{,}00$ — индифферентным к спирали (I) и $s = 1{,}5$ — образующим спираль (Н) (табл. 6.1)*. Значения σ и s получают по наклонам и температурным переходам зависимостей, описывающих переходы спираль — клубок в синтетических полипептидах, используя уравнения (А.18) и (А.20). Спиральная конформация предсказывается для всех положений остатков i, для которых P_i больше средней величины P_i. В результате получаются непрерывные потенциальные функции, поскольку уравнение (6.2) учитывает кооперативность модели Зимма — Брэгга, согласно которой спирали должны иметь определенную длину (рис. А. 1). Этот метод предсказания дает спиральные сегменты длиной около 10 остатков, что намного меньше длины, ожидаемой для данного значения σ гомополимеров при $s = 1$, т. е. $1/\sqrt{5 \cdot 10^{-4}} \cong 40$ (уравнение (А.17)). Такое укорочение спирали является следствием включения остатков с низкими значениями s.

Модифицированная модель Зимма— Брэгга. Аналогичным образом модель Зимма — Брэгга была применена Птицыным и сотр. [370], которые использовали для всех типов остатков один параметр инициации $\sigma = 5 \cdot 10^{-4}$ и шесть различных значений s_i, базирующихся на экспериментальных данных по синтетическим полипептидам. Значения s_i остатков, для которых отсутствовали экспериментальные данные, были выбраны по методу Льюиса и сотр. [368]. В последующих работах [371—374] для определения величин s (табл. 6.1) привлекались также стереохимические данные. Модель

* Как отмечено в разд. А.2, относительные статистические веса s соответствуют склонностям. В уравнении (А.8) s есть склонность к спирализации.

Зимма — Брэгга была модифицирована путем учета влияния соседних остатков (например, остатков i и $i + 3$ в α-спирали). Кроме того, авторы ввели индивидуальные значения σ для каждого типа остатка и расширили алгоритм, сделав попытку предсказать уже не две (спираль и клубок), а до шести различных конформаций.

Параметры Зимма—Брэгга, полученные из данных по глобулярным белкам и синтетическим полипептидам

Вероятностный характер метода предсказания Чоу и Фасмана; учет кооперативности. Метод предсказания Чоу и Фасмана [201, 340] во многих отношениях основывающийся на модели Зимма — Брэгга, имеет вероятностную природу. Авторы определяли склонности к α-, β- и *rt*-конформациям по частотам встречаемости в базовом наборе. Для заданной аминокислотной последовательности они последовательно проанализировали точки инициации α- и β-конформаций путем комбинирования соответственно α- и β-склонностей в гекса- и пентапептидах. От каждой найденной точки инициации затем производилось наращивание цепи в обоих направлениях до тех пор, пока не встречались остатки с малыми склонностями. Для предсказания реверсивных поворотов авторы пользовались схемой Льюиса и сотр. [326]. Применение метода Чоу и Фасмана описано в разд. 6.3, где приведены дальнейшие подробности.

Склонности к спирали, полученные с помощью статистической механики синтетических полипептидов, соответствуют склонностям, основанным на частотах встречаемости в глобулярных белках. В своих последующих работах Чоу и Фасман [201] сопоставили склонности к спирали, определенные но наблюдаемым частотам встречаемости в глобулярных белках, с данными, полученными на основании температур θ переходов спираль — клубок синтетических полипептидов, согласно модели Зимма — Брэгга. Как показано в разд. А.5, по температуре перехода θ можно определить относительные статистические веса s, а следовательно и склонности к спиралеобразованию. Чоу и Фасман показали, что величины s семи типов остатков, для которых имеются данные по синтетическим полипептидам, в пределах 10% согласуются со склонностями к спирали, полученными по частотам встречаемости в глобулярных белках. Это соответствие было более подробно исследовано Судзуки и Робсоном [352].

Параметры инициации, полученные из данных по синтетическим полимерам, хорошо коррелируют с частотами встречаемости остатков на концах спирали. Чоу и Фасман [201] сравнили также параметры инициации σ, полученные при рассмотрении переходов спираль — клубок (разд. А.5), с частотами встречаемости остатков на концах спиралей в глобулярных белках. Как и в работах Лифсона и Ройга [375], в данном случае было проведено усреднение по обоим

концам спирали. Кроме того, рассматривались частоты встречаемости трех остатков на каждом конце (а не только одного остатка по модели Зимма — Брэгга), поскольку каждый из трех концевых остатков α-спирали образует только одну водородную связь. При сопоставлении найденных частот с шестью значениями σ, полученными при изучении переходов спираль — клубок, несмотря на различие абсолютных величин, была обнаружена хорошая корреляция (коэффициент корреляции между логарифмами составлял +0,75). Поскольку переходы беспорядочный клубок — β-складчатый лист в синтетических полимерах значительно менее изучены, то таких корреляций для β-структур пока установить не удалось.

Методы, основанные на стереохимических данных

Реверсивные повороты содержат главным образом полярные остатки. Простой стереохимический подход для предсказания поворотов цепи был предложен Кунтцем [203]. Автор обратил внимание на то, что повороты располагаются почти исключительно на поверхности белка. В связи с этим он высказал предположение, что все триплеты полярных остатков (обозначенные буквой Т в табл. 6.2) имеют конформацию реверсивного поворота.

Спирали, находящиеся между гидрофобным ядром и растворителем, можно узнать по расположению в них неполярных остатков. С помощью графов, содержащих пять витков спирали, Шиффер и Эдмундсон [376, 377] показали, что неполярные остатки группируются на одном краю спирали, образуя неполярные дуги. Пример такого «α-спирального колеса» показан на рис. 6.1. Для предсказания спирали авторы исследовали наличие неполярных триплетов в положениях i, $i+3$, $i+4$ (относительные положения 1-4-5), а также в положениях i, i—3, i—4 (1-2-5) данной аминокислотной последовательности. Затем эти положения были приняты за спиральные нуклеации и относительно них были построены «спиральные колеса». Предполагалось, что спираль будет наращиваться в обоих направлениях до тех пор, пока неполярная дуга не будет прервана полярными остатками или остатком Pro. Этот метод предназначался для поиска спиралей, экранирующих гидрофобное ядро от растворителя, как в случае миоглобина. Обнаружить с его помощью спирали в других положениях более трудно.

Расположение неполярных остатков в α-спиралях было рассмотрено также Палау и Пигдоменечом [378], которые использовали значительно больший базовый набор. Эти авторы показали, что частота встречаемости неполярных триплетов в относительных положениях 1-2-5 и 1-4-5 в α-спиралях значительно выше, чем можно было ожидать по аминокислотному составу этих спиралей. Очень часто эти триплеты располагались в виде цепей 1-2-5-″1″-″4″-″5″ → → 1-2-5-8-9, т. е. образовывали дуги на «α-спиральном колесе»

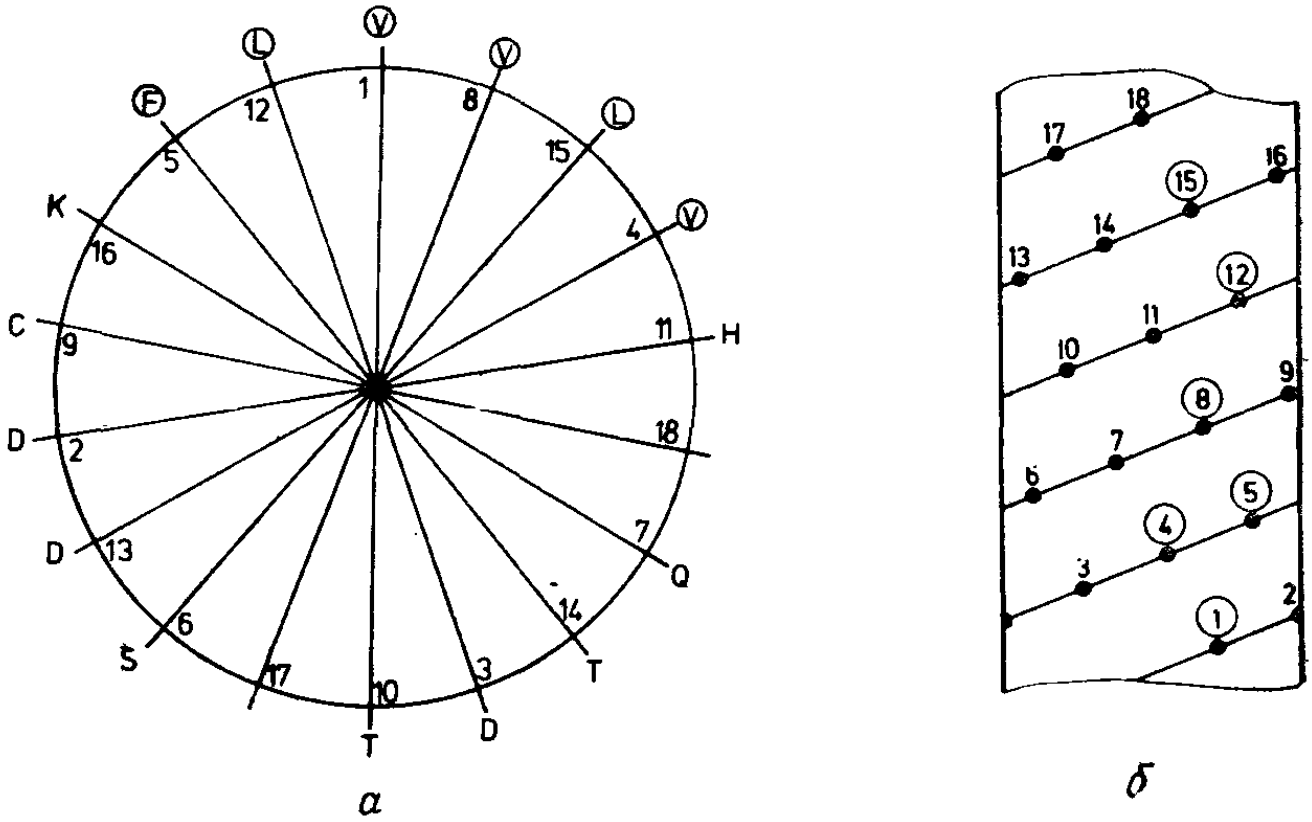


Рис. 6.1.

Спиральный круг С-концевой α-спирали аденилаткиназы.

а — круг представляет собой проекцию положеиий всех боковых цепей вдоль оси спирали на плоскость. Остаток в положении 1—Val-179, а последний остаток в положении 16—Lys-194. Неполярные остатки (обведены кружками) располагаются с одной стороны спирали, которая в трехмерной структуре [186] обращена к гидрофобному ядру; другая сторона спирали составляет часть внешней поверхности молекулы. *б* — положения боковых цепей на цилиндрической диаграмме α-спирали (рис. 5.2). Остатки, образующие неполярные дуги, обведены кружками.

(рис. 6.1). Таким образом, наблюдения Шиффера и Эдмундсона, основывающиеся только на данных по трем белкам, подтвердились для большего базового набора.

Метод, выявляющий расположение остатков в спиралях и β-складчатых листах. Рассмотрение нуклеаций неполярных остатков было использовано также Лимом в его предсказаниях вторичной структуры [379, 380]. Аналогично Палау и Пигдоменечу [378], он обратил внимание на расположение триплетов в относительных положениях 1-2-5 и 1-4-5, а также дублетов в положениях 1-4 и 1-5. Если боковые цепи были слишком объемными, образование кластеров не обсуждалось. Учитывалось также влияние полярных остатков в том аспекте, что крупные полярные остатки могут стабилизировать соседние неполярные кластеры путем экранирования их от растворителя. В предсказании β-структуры различались внутренние, полуповерхностные и поверхностные β-листы. Поверхностные листы имеют полярную и неполярную стороны, поэтому в них должно наблюдаться альтернирование полярных и неполярных остатков. Напротив, во внутренних листах неполярные остатки расположены по обе стороны, и, таким образом, эти структуры можно идентифицировать по непрерывному чередованию неполярных остатков. Такие сегменты обычно экранируются крупными поляр-

ными остатками на концах цепи. В полуповерхностных листах все боковые цепи расположены наклонно к поверхности, как, например, в случае С-концевого β-листа центральной β-структуры аденилаткиназы, показанной на рис. 7.6. Такие β-листы могут не содержать крупных неполярных остатков, поскольку нет необходимости в их экранировке от растворителя.

Лим сформулировал свои соображения в виде 22 правил для α-предсказаний и 14 правил для β-предсказаний. Он составил таблицу 17 антиспиральных дублетов и 26 антиспиральных триплетов и сгруппировал типы аминокислотных остатков 14 различными способами. β-Конформация была постулирована только для областей, в которых не предсказывалась α-спираль. Следует отметить, что правила, предложенные Лимом, содержат большой объем информации, т. е. варьируемых параметров, и, таким образом, их всегда можно согласовать с базовым набором, даже если этот набор велик. Таким образом, ценность этого метода можно проверить только путем опробования на неизвестных структурах.

6.3 Применение предсказательных методов

Простые методы предсказания наиболее популярны. Предсказательные методы обычно требуют громоздких расчетов с использованием больших таблиц (например, метод Нагано [353]), так что для их воспроизведения необходимо получить от авторов программу для ЭВМ, таблицы и параметры, выведенные из базового набора. Некоторые методы можно применять, не прибегая к вычислительной технике, но они настолько сложны, что требуют огромных затрат времени (например, метод Лима [380]. Обычно авторы не пользуются приближениями, которые могли бы упростить применение предсказательного метода. Однако существует несколько простых и в связи с этим наиболее популярных методов. Из них чаще всего, по-видимому, применяется метод Чоу и Фасмана [340]. В качестве простого примера рассмотрим предсказания этого метода для первых 24 остатков аденилаткиназы.

Первым этапом метода Чоу и Фасмана является обнаружение мест инициации спирали и слоя. Предсказания спирали и β-структуры проводятся независимо и выполняются в связи с этим одновременно. Аминокислотная последовательность, величины, характеризующие склонности остатков к спирали и β-структуре (табл. 6 1 и 6.2), а также обозначения H, h, I, i, b, B даны на рис. 6.2.

Первый этап заключается в нахождении центров инициации. Для инициации спирали необходимо наличие гексапептида, содержащего четыре спиралеобразующих остатка h^{α} или H^{α} (I^{α} эквивалентно половине h^{α}), а также не более одного прерывания спирали b^{α} или B^{α}. Четыре таких пептида, определяющие две нуклеации, показаны на рис. 6.2. Точно так же инициация β-структуры

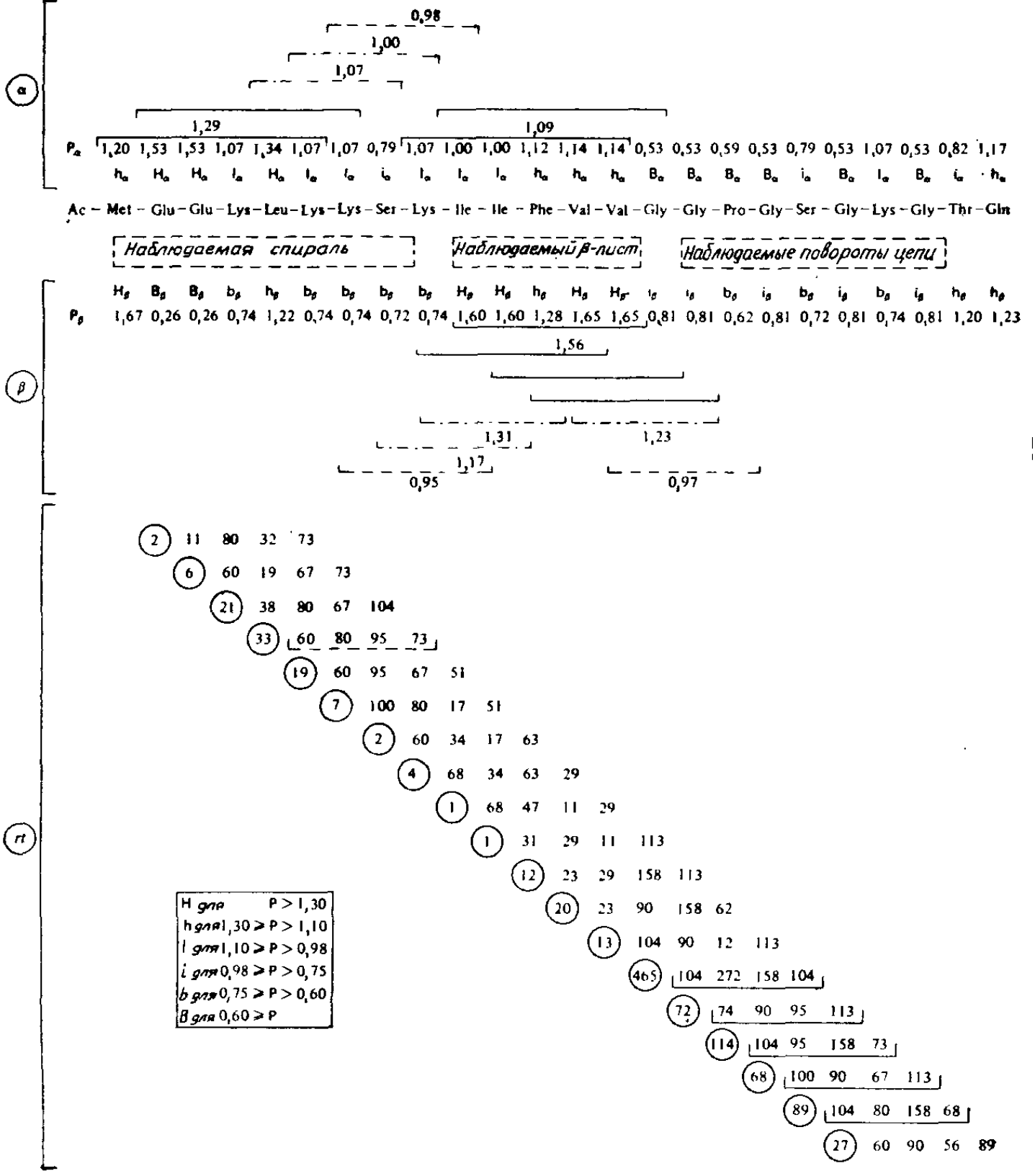


Рис. 6.2.

Предсказание α-спирали, β-структуры и реверсивного поворота (*rt*) для 24 N-концевых остатков аденилаткиназы [389] с помощью процедуры Чоу и Фасмана [340].

Параметры P_α и P_β, характеризующие склонность остатков входнть в α-спираль и β-структуру (соответствеино α- и β-потенциалы остатков), взяты нз табл. 6.1 и 6.2. Критерии оценкн приведеиы на рнсуике в рамке. При предсказаиии α-спирали гексапептиды со средней склоиностью выше 1,00 выделеиы сплошиыми лиииями. Также обозиачены пентапептиды со средней β-склонностью выше 1,00. Тетрапептидные концевые участки спиралей н складчатых листов со средней склонностью выше 1,00 отмечеиы штрихпунктнрными, а ниже 1,00 — пуиктирными линиями. Склоииость остатка входить в поворот цепи зависит от его положения в рассматриваемом тетрапептиде. Таким образом, у каждого остатка выявляются четыре склонности. Все оии умножаются иа 10^3. *rt*-Потенциал является произведением склоииостей остатков в тетрапептиде. На рисунке оии обведены кружками и приведены в тех же строках, что и *rt*-склонности тетрапептида. Все потеициалы умножеиы на 10^{-6}. Тетрапептиды с потенциалами выше $50 \cdot 10^{-6}$ подчеркнуты сплошнымн линиями, а несколько ниже этой величины — пунктирной линией.

происходит в том случае, если пентапептид содержит три остатка h^3 или H^3 и не более одного остатка b^3 и B^3. Как видно из рис. 6.2, имеются четыре таких пептида, определяющих один центр инициации. Оба центра (спирали и листа) находятся около остатка 12. Для разрешения этой неопределенности сравнивают самые высокие средние склонности. Поскольку склонность к β-структуре (1,56) выше, чем склонность к спирали (1,09), этот сегмент считают центром инициации β-структуры.

Нуклеации спирали и β-структуры наращиваются в обе стороны. На следующем этапе нуклеация спирали с центром на остатке 4 и нуклеация складчатого листа с центром на остатке 12 наращиваются в обоих направлениях до тех пор, пока значение средней склонности тетрапептида не упадет ниже 1,00. Правила не указывают точно, какой из остатков нарушающего тетрапептида следует включать соответственно в спираль и β-структуру. По-видимому, целесообразно включать только остатки H, h и I. Таким образом, спираль предсказывается для остатков 1—7, а β-структура — для остатков 10—14.

Предсказание реверсивного поворота. Для предсказания поворота Чоу и Фасман [340] применили метод Льюиса и сотр. [326] с расширенным базовым набором. Потенциал реверсивного поворота тетрапептида определялся как произведение склонностей к повороту четырех остатков в положениях i, $i+1$, $i+2$, $i+3$ (разд. 6.1). Соответствующие величины показаны на рис. 6.2. С использованием порога $50 \cdot 10^{-6}$ повороты были предсказаны для остатков 16—23.

При сравнении предсказанной и наблюдаемой вторичной структуры в данном случае обнаруживается хорошее соответствие: предсказанная спираль на один остаток короче, а β-лист и поворот предсказаны точно. Однако приведенный пример не очень показателен в том отношении, что в рассматриваемом случае потенциалы α, β и rt весьма высоки и явно различаются. Обычно в реальных структурах ситуация менее определена и предсказания значительно менее точны.

6.4 Оценка предсказательных методов

Произвольность выбора параметров

Вводя достаточное число параметров можно получить прекрасное соответствие с любым базовым набором. Попытки улучшить предсказания должны всегда опираться на оценку точности метода. Имея это в виду, многие авторы опробовали свой метод на базовом наборе белков с известной структурой. Однако такая процедура имеет весьма ограниченную ценность. Совершенно очевидно, что чем больше внутренней информации (правил и параметров) вклю-

чает метод, тем лучше будет соответствие с базовым набором, однако это вовсе еще не определяет предсказательную ценность метода для белков с неизвестной структурой.

Чтобы как-то решить эту проблему, некоторые авторы вначале исключают некий белок из базового набора и проводят параметризацию с сокращенным базовым набором. Затем найденные параметры используются для предсказания вторичной структуры выделенного белка. Таким образом, абсолютные величины подбираемых параметров как будто становятся независимыми (склонности, данные в табл. 6.1 и 6.2). Однако правила использования параметров, которые и составляют сущность каждого метода, по-прежнему основываются на всех известных белках, включая и предсказываемый.

Некоторые методы не определены достаточно четко. Помимо трудностей, связанных с независимой проверкой предсказательности метода, в некоторых случаях оказывается, что метод просто невоспроизводим [381, 382]. Этими причинами и объясняется большой скептицизм в отношении методов предсказаний вторичной структуры. Нужно отметить однако, что хорошая воспроизводимость и соответствие с базовым набором еще не адекватны успешным предсказаниям *неизвестных* структур белков. Это происходит главным образом потому, что авторы не определяют свой метод с достаточной четкостью, оставляя тем самым возможность большого произвола в его применении. Критическое отношение к предсказательным методам до известной степени поколебалось после сравнительного опробования на аденилаткиназе [383] и Т4-лизоциме [384].

Критерии оценки методов

Различные критерии качества метода отражают разные аспекты проблемы предсказания. Для того чтобы сопоставить различные методы, предложены разнообразные критерии — показатели качества. С их помощью оценивается результативность предсказаний для отдельных положений остатков в данном белке по параметрам w, x, y, z (см. определение в табл. 6.3). Каждый из этих параметров отражает различные аспекты предсказаний, их арифметические комбинации сопоставлять довольно трудно. Полный перечень способов оценки дан в табл. 6.4.

Наиболее распространенный показатель дает завышенные значения. Если по показателю Q_1 можно судить о завышениях в предсказаниях вторичных структур, то показатель Q_2 вовсе их не учитывает; в экстремальном случае, когда каждый остаток предсказывается спиральным ($w = 1$), показатель Q_2 принимает значение, отвечающее оптимальному соответствию ($Q_2 = 1$). Наиболее распространенным является, по-видимому, показатель Q_3, который, однако, дает завышенные значения. В этом случае положительные и от-

Таблица 6.3

Четыре возможных случая при отнесении остатка к α-спирали (или другим типам вторичной структуры)

Конформационное состояние остатка		Оценка предсказания	Процентное содержание случая в данном белке
предсказание	*наблюдение*		
Спиральное	Спиральное	Положительное правильное предсказание	w
Неспиральное	Неспиральное	Отрицательное правильное предсказание	x
Неспиральное	Спиральное	Недопредсказание	y
Спиральное	Неспиральное	Сверхпредсказание	z
			$w + x + y + z = 1$

рицательные правильные предсказания (табл. 6.3) входят с равными весами. Поскольку, однако, среднее содержание спиралей в белках, как известно, приближается к 35% [201], то вероятность отрицательного правильного предсказания спирали приблизительно в два раза выше вероятности положительного правильного предсказания. Следовательно, наибольший вклад в показатель Q_3 вносят отрицательные предсказания; эта ситуация иллюстрируется примерами 1 и 2 табл. 6.4. Пример 1 иллюстрирует неверное предсказание, когда не было найдено ни одного спирального остатка. Тем не менее величина показателя Q_3 в этом случае достигает 0,70. В менее экстремальном примере 2 только половина спиральных остатков определена правильно, а другая половина — неправильно. Полученный показатель ($Q_3 = 0{,}80$) здесь близок к значению для оптимального соответствия. Для β-структур преобладание отрицательных правильных предсказаний выражено еще более резко, поскольку среднее содержание этих структур в белках составляет всего около 15% [201].

Показатели, занижающие отрицательные правильные предсказания. Таким показателем, занижающим вклад отрицательного правильного предсказания, является Q_4. В примерах 1 и 2 он более правильно отражает ситуацию, чем показатель Q_3. В показателе Q_5 отрицательные правильные предсказания учитываются путем не простого прибавления x, как в Q_3, а путем умножения на $1/(1 - x)$. В связи с этим он дает более разумные значения для примеров 1 и 2. Однако этот показатель не получил сколько-нибудь широкого распространения. Показатель Q_6 представляет собой монотонную функцию Q_3 и не дает дополнительной информации. Показатель Q_7 есть не что иное как общий статистический критерий. Поскольку в этом случае используются средние значения [384], вклад отрицательных

Таблица 6.4

Критерии оценки качества предсказания[a]

Описание	Авторы критерия и авторы, его использующие	Показатель качества	Формула	Пределы	Пример 1 $w=0$ $x=0,7$ $y=0,1$ $z=0,2$	Пример 2 $w=0,1$ $x=0,7$ $y=0,1$ $z=0,1$
$\frac{\text{Положительное правильное предсказание—сверхпредсказание}}{\text{наблюдаемое}}$	Протеро [332]	Q_1	$\frac{w-z}{w+y}$	$(-\infty, +1)$	—	—
$\frac{\text{Положительное правильное предсказание}}{\text{наблюдаемое}}$	Льюис и Шерага [369] Нагано (% правильных отнесений третичных структур) [353] Чоу и Фасман (%α) [340]	Q_2	$\frac{w}{w+y}$	(0,1)	—	—
Положительное и отрицательное правильные предсказания	Котельчук и Шерага [364] Льюис и сотр. [368] Робсон и Пэйн [346] Леберман [365] Нагано (% правильно предсказанных остатков) [353] Чоу и Фасман (%α+$n\alpha$) [340]	Q_3	$w+x$	(0,1)	0,70	0,80
Среднее значение для $\frac{\text{Положительное правильное предсказание}}{\text{наблюдаемое}}$ +	Птицын и Финкельштейн [342]	Q_4	$\frac{1}{2}\left(\frac{w}{w+y}+\frac{w}{x+z}\right)$	(0,1)	0,39	0,69

отрицательное правильное предсказание / ценаблюдаемое	Чоу и Фасман (Q_α) [340]					
Положительное правильное предсказание / 1—отрицательное правильное предсказание	Нагано (% правильных отнесений вторичных структур) [353]	Q_5	$\frac{w}{1-x}$	(0,1)	0,00	0,33
Все правильные предсказания / все неправильные предсказания	Кабат и Ву [359]	Q_6	$\frac{w+x}{y+z} = \frac{Q_3}{1-Q_3}$	(0, ∞)	—	—
Коэффициент корреляции между положительными правильными предсказаниями и наблюдаемыми состояниями, а также между отрицательными правильными предсказаниями и не наблюдаемыми состояниями	Мэтьюз [384] Аргос и сотр. [381] Ленстра [382]	Q_7	$\frac{w \cdot x - x \cdot z}{\sqrt{(x+y)\cdot(x+z)\cdot(w+y)\cdot(w+z)}}$	(−1, +1)	−0,17	+0,38

[a] Два примера иллюстрируют чувствительность критериев оценки.

правильных предсказаний уже не будет завышенным; так, например, отрицательные правильные предсказания α-спирали для данного положения остатка получатся с низким весом, если в белке всего несколько спиралей, и с большим весом, если спиралей много.

Предсказание всех трех типов вторичной структуры должно производиться одновременно. Ни один из обсуждаемых показателей качества не учитывает α-, β- и *rt*-структуры одновременно. Очевидно, что все три предсказания не независимы (никто не предсказывает β-структуру там, где обнаружен высокий α-потенциал) и что общий результат предсказания данным методом может быть надежно описан только объединенным показателем. Простым объединенным показателем может служить величина

$$Q_3^{\text{общ}} = w_\alpha + w_\beta + w_{rt} + w_{\text{клубок}}$$

где w_α, w_β и w_{rt} — процентное содержание правильно (положительно) предсказанных α-, β- и *rt*-остатков, а $w_{\text{клубок}}$ — отрицательное правильное предсказание, соответствующее величине x в выражении для Q_3. Поскольку белки содержат в среднем около 35% спиралей, 15% β-структур, 25% поворотов и 25% клубков, то все четыре вклада приблизительно сбалансированы без преобладания вклада отрицательного предсказания.

Современное состояние предсказательных методов

В двух белках с неизвестной структурой более чем двум третям остатков была приписана правильная вторичная структура. Предсказательная способность методов охарактеризована в табл. 6.5, где приведены результаты опробования на двух белках: аденилаткиназе [383] и Т4-лизоциме [384], а также проиллюстрирован уровень соответствия в базовом наборе. Для каждого белка было сделано по нескольку предсказаний, однако в табл. 6.5 входят только средние значения по двум лучшим предсказаниям. Очевидно, что в параметрах $Q_{3\beta}$ и Q_{3rt} преобладают отрицательные правильные предсказания и что они дают завышенную оценку. Более достоверные коэффициенты корреляции Q_7 имеют величину 0,45; они свидетельствуют о наилучших предсказаниях для α- и наихудших для *rt*-структур. Значение $Q_3^{\text{общ}}$ достигает 68%, что указывает на правильное предсказание более чем двух третей остатков. Для аденилаткиназы показатели Q_3 и Q_7 выше, чем для T_4-лизоцима, однако значения $Q_3^{\text{общ}}$ близки. Это несоответствие объясняется тем, что для аденилаткиназы число сверхпредсказаний больше, чем для T_4-лизоцима, а сверхпредсказание сильнее сказывается на значении $Q_3^{\text{общ}}$, чем на показателе Q_3 или Q_7.

Обычно качество предсказания не зависит от класса белка. Содержание спиралей в обоих рассмотренных белках составляет

около 55% по сравнению со средней величиной 35% [201]. Следовательно, общую правильность полученных результатов нельзя было предвидеть априори. Однако анализ качества предсказаний по базовому набору [382], выраженный в значениях Q_7, показал, что оно очень близко к полученному для обоих белков (табл. 6.5). Таким образом, современные методы позволяют достичь коэффициентов корреляции около + 0,45 в отношении α-, β- и *rt*-предсказаний, что намного выше случайного набора. В своем обзоре Ленстра [382] подразделил базовые данные по белкам на четыре класса в соответствии с табл. 5.2. За исключением α-предсказания в классе 2 (β-белки), которое представляется предельно сложным, качество предсказания практически одинаково для всех классов белков.

Если имеется несколько предсказаний, их можно усреднить. При оценке предсказаний для аденилаткиназы был предложен дополнительный метод коллективного предсказания [383]. Он состоит в усреднении (без относительных весов) всех индивидуальных предсказаний. В случае аденилаткиназы такое усреднение лучше согласовывалось с наблюдаемыми данными, чем любое из индивидуальных предсказаний. Такой подход был применен также Аргосом и сотр. [381], которые усреднили результаты пяти методов предсказания для всех известных белковых структур.

Оценка индивидуальных результатов — ключ к дальнейшим улучшениям. Вообще говоря, для улучшения предсказаний необходимо анализировать все индивидуальные результаты отдельно. Для этого можно воспользоваться обоими рассмотренными выше случаями аденилаткиназы и Т4-лизоцима, поскольку они позволяют читателю самому оценить возможности методов. Но опять-таки следует иметь в виду, что опробование было проведено на белках, которые представляют только два из пяти классов, перечисленных в табл. 5.2. Этот недостаток удалось обойти Ленстра [382], который сравнил результаты трех индивидуальных и одного коллективного методов предсказания по всему базовому набору, используя в качестве критерия показатель Q_7. Он обнаружил отчетливые различия в качестве предсказаний. По-видимому, наилучшие результаты дает коллективный метод [383].

Самосогласованность можно использовать в качестве критерия качества. Существенно иной подход к оценке качества метода предсказания состоит в применении его к нескольким гомологичным белкам. Можно полагать, что такие белки свертываются одинаковым образом и имеют одинаковую вторичную структуру. Следовательно, предсказания должны быть инвариантны по отношению к наблюдаемым заменам аминокислот: чем меньше вариаций, тем лучше метод предсказания. Такая проверка качества трех методов предсказания была проведена на 24 гомологичных последовательностях рибонуклеазы поджелудочной железы [385].

Таблица 6.5

Показатели качества при опробовании на двух белках и при соответствии в базовом наборе[а]

	Структурный класс белка (табл. 5.2)	Число остатков	α	β	rt	Сумма	$Q_{3\alpha}$	$Q_{3\beta}$	Q_{3rt}	$Q_{7\alpha}$	$Q_{7\beta}$	Q_{7rt}	$Q_3^{общ}$
Аденилаткиназа	4	194	0,54	0,12	0,24	0,90	0,77	0,92	0,86	0,56	0,58	0,60	0,67[б]
T_4-лизоцим	3	164	0,57	0,07	0,15	0,79	0,69	0,87	0,75	0,42	0,28	0,20	0,29[в]
			0,55	0,10	0,20		0,73	0,90	0,81	0,49	0,43	0,40	0,68
Базовый набор	1									0,50	—	0,41	
Базовый набор	2									0,22	0,45	0,37	
Базовый набор	3									0,58	0,51	0,45	
Базовый набор	4									0,51	0,50	0,43	
										0,45	0,49	0,42	

[а] Для обоих белков [383, 384] усреднены два лучших значения. Базовый набор включает все белки с известной структурой, данные взяты из работ Нагано [356] и Аргоса и сотр. [381] в том виде, как они представлены Ленстра [382].
Число под чертой есть среднее значение верхних величин.

[б] Среднее из предсказаний Нагано [353] и Чоу и Фасмана [340].

[в] Среднее из предсказаний Птицына и сотр. [371, 380] и Бэджиса и сотр. [31].

Возможные улучшения методов, основанных на стереохимических данных. Возможно ли дальнейшее улучшение предсказаний? Расширение базового набора в этом плане вряд ли может оказаться полезным. Это совершенно очевидно для склонностей синглетов и дублетов, поскольку для них наблюдаемые частоты встречаемости уже достаточно высоки для получения хороших средних значений. Таким образом, в рамках вероятностных методов вряд ли возможны существенные улучшения. В этом отношении перспективы методов, основывающихся на стереохимических данных, представляются более благоприятными. Сейчас, по-видимому, уже нецелесообразно раздельно рассматривать α-, β- и *rt*-структуры. Наступило время комбинированных методик. Нужно также стараться предсказывать класс белка. Конечная цель — предсказание укладки цепи — по-прежнему выглядит отдаленной.

6.5 Замечания общего характера

Качество предсказания и локальные взаимодействия

При разном содержании α- и β-структур в белках трудно сопоставлять качество предсказаний. Обычно считают, что спирали предсказываются значительно лучше, чем β-структуры. Такое положение объясняется в основном тем, что в спиралях взаимодействие остатков через водородные связи имеет в отличие от структур локальный характер. Соответственно показатели качества для спиралей должны быть выше. Однако из табл. 6.5 не следует какого-либо заметного различия в величинах Q_7. Такое несоответствие между интуитивными соображениями и величинами Q_7 легко устранить, если учесть, что в белках β-структур меньше, чем спиралей, что приводит к большему отрицательному правильному предсказанию для β-структур (табл. 6.3). Хотя отрицательное правильное предсказание несколько занижается в значении Q_7, оно значительно увеличивает значение $Q_{7\beta}$ по сравнению с $Q_{7\alpha}$. Таким образом, несмотря на меньшее положительное правильное предсказание для β-структур, что и создает впечатление более низкого качества предсказания, значения $Q_{7\beta}$ и $Q_{7\alpha}$ сопоставимы.

Реверсивные повороты не определяются локальными взаимодействиями. В табл. 6.5 наиболее неудовлетворительными выглядят предсказания поворотов. Это объяснимо в том случае, если считать, что повороты цепи, несмотря на их геометрически локальную природу, определяются нелокальными взаимодействиями. При таком допущении повороты следует рассматривать как точку наименьшего сопротивления при свертывании цепи, т. е. точку вынужденного изгиба цепи. Энергетические соображения приводят к тому же заключению, так как поворот содержит только одну водородную

связь. Кроме того, повороты обычно находятся на поверхности, поэтому их водородная связь подвергается действию воды, которое понижает ее вклад в свободную энергию.

Успех предсказания и центры свертывания

Центры инициации свертывания цепи можно обнаружить по наиболее удачным предсказаниям. Было высказано предположение [383], что наилучшим образом предсказанные вторичные структуры имеют наивысшую вероятность оказаться первоначальными нуклеациями процесса свертывания, поскольку хорошее предсказание свидетельствует о преобладании локальных взаимодействий, которые, по-видимому, и направляют механизм укладки цепи. Такой центр свертывания, состоящий из спирали, был обнаружен в аденилаткиназе [383]; он соответствует малому домену на рис. 5.14, *б*.

Эта идея подтверждается и тем фактом, что при предсказаниях аденилаткиназы и Т4-лизоцима в обоих случаях N-концевая половина цепи была предсказана лучше, чем С-концевая [383, 384]. Действительно N-концевая часть синтезируется на рибосоме первой и, по-видимому, первой свертывается. На это обратили внимание Аргос и сотр. [381], которые с помощью коэффициента корреляции Q_7 показали, что усредненные предсказания значительно лучше для N-концевой половины белка.

Предсказание как аналитический метод

Современные методы, по-видимому, позволяют предсказывать структурный класс белка. Достигнутая в настоящее время точность предсказаний позволяет отнести белок к одному из пяти классов, приведенных в табл. 5.2. Характер вторичной структуры в этих классах настолько различен, что даже при правильности предсказания 68% ($Q_3^{\text{общ}}$ в табл. 6.5) их можно отличить друг от друга.

Предсказания помогают распознавать новые члены в группе родственных белков. Помимо установления структурного класса можно попытаться выявить цепи с определенной, уже известной топологией. Применив в глютаматдегидрогеназе метод Чоу и Фасмана [340], Вутон [386] выявил два домена, характер вторичной структуры которых аналогичен структуре нуклеотидсвязывающего домена других, уже исследованных дегидрогеназ [254]. После установления общей структуры фрагмента он обнаружил, что аминокислотная последовательность гомологична последовательности одной из этих дегидрогеназ (глицеральдегид-3-фосфат-дегидрогеназе). Это обстоятельство свидетельствует о большой «значимости подобия» (разд. 9.6) и позволяет выявить эволюционную взаимо-

связь [387], отраженную в существовании нуклеотидсвязывающего домена с той же топологией цепи. Для второго из рассматриваемых доменов гомологии последовательности обнаружить не удалось, хотя по характеру вторичные структуры соответствовали друг другу. В этом случае «значимость» была слишком низка для выявления топологического подобия.

Аналогичным путем было проведено исследование Ca^{2+}-связывающих белков Аргосом [388], который показал, что предсказания могут способствовать определению аминокислотной последовательности, особенно в сомнительных случаях. Кроме того, с помощью предсказаний он попытался определить положение Ca^{2+}-связывающих центров в ацилпереносящем белке синтетазы жирных кислот *E.coli*, а также в белках кровесвертывающего каскада.

Заключение

Опыты на пептидных гомополимерах позволили установить, что некоторые аминокислотные остатки обнаруживают свойство встраиваться в α-спираль. Эти данные были затем применены к глобулярным белкам с известным пространственным строением. Оказалось, что остатки, которые образуют спирали в гомополимерах, стремятся встраиваться в спирали также и в глобулярных белках. Такое соответствие послужило основой для многочисленных попыток установить корреляции между аминокислотной последовательностью белка и наличием спиралей в этом белке. Позднее такие корреляции были распространены и на другие вторичные структуры. Эти попытки интересны с той точки зрения, что они являются некой основой пока еще не известного будущего метода, с помощью которого можно будет устанавливать трехмерную структуру белка только по его аминокислотной последовательности. Это и послужило причиной подробного описания большинства из существующих методов «предсказания» вторичной структуры по аминокислотной последовательности. Сведения, необходимые для понимания методов, основанных на статической механике, даются в приложении.

С помощью существующих методов можно предсказать правильную вторичную структуру приблизительно для двух третей всех остатков белка. Не существует больших различий в качестве предсказаний между вероятностными методами, основанными только на математическом анализе структурных данных, и физико-химическими методами, которые используют также дополнительную стереохимическую информацию. Это свидетельствует о том, что пока еще мы недостаточно хорошо понимаем физические основы корреляций. До последнего времени сопоставлять различные предсказательные методы было довольно трудно, поскольку авторы пользовались разными показателями качества. Взаимосвязь между

этими показателями и их относительные достоинства обсуждаются в этой книге. Такая сравнительная оценка методов может способствовать их дальнейшему развитию. Хороший предсказательный метод может оказаться полезным не только при определении трехмерных структур, но может также дать некоторую информацию о процессе свертывания, выявляя сегменты цепи, которые сворачиваются первыми.

Глава 7 Модели, изображение и документация белковых структур

7.1 Данные о строении белка

Во многих случаях известно только ковалентное или только пространственное строения белка. Сведения о структуре белка могут включать известную аминокислотную последовательность (ковалентную структуру), известную трехмерную (геометрическую) структуру или все эти данные одновременно. В настоящее время известно значительно больше ковалентных структур, чем пространственных. С другой стороны, имеется более десятка геометрических структур, аминокислотные последовательности которых пока еще не установлены [80, 124, 217, 221, 236, 261, 303, 306, 307, 310, 313, 316]. Все расшифрованные ковалентные структуры приводятся в Атласе аминокислотных последовательностей и структур белков [20]. Данные об известных пространственных структурах сведены в табл. 5.2.

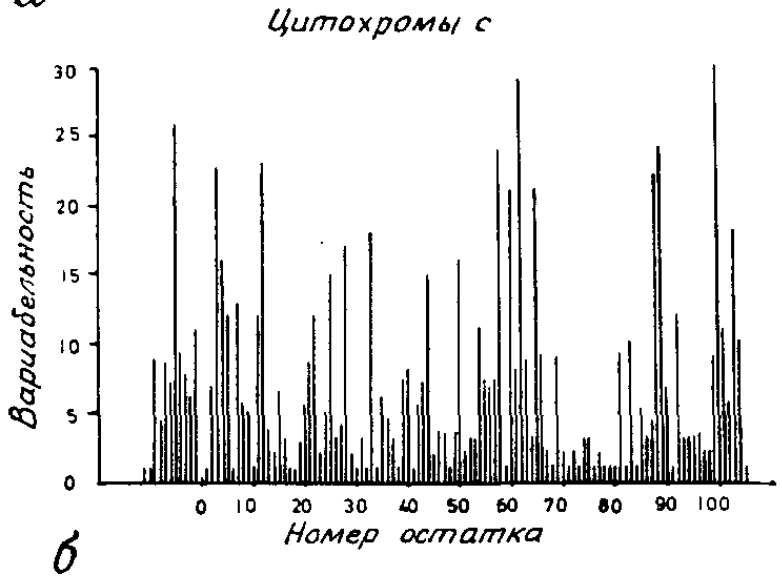
Ковалентное строение

Поскольку белки состоят из линейных полипептидных цепей, документация ковалентной структуры проста: ее можно изобразить линейной последовательностью аминокислотных остатков, записанных сокращенно в виде однобуквенных символов (табл. 1.1). Такую запись легко читать. Для удобства каждому остатку присваивается порядковый номер вдоль цепи.

Обычно для последовательностей гомологичных белков используется одинаковая нумерация. Запись гомологичных последовательностей несет значительно больше информации, если она выполнена в виде единой схемы, как это представлено на рис. 7.1, *а*. Здесь для всех последовательностей используется одинаковая нумерация, что возможно, однако, лишь в том случае, если допустимы пропуски аминокислот — делеции. Трудности возникают тогда, когда при последующих исследованиях обнаруживают гомологичные последовательности, содержащие один или более дополнительных остатков. Публикуя эти данные, обычно сохраняют установленную схему нумерации, а дополнительным остаткам приписывают

	3 4 ...7 8 9 0 1 2 3 4 5 6 7 8 9 0 1...
Лошадь	CHTV------EKGGK
Neurospora crassa	CHTL------EEGGG
Ginkgo biloba	CHTV------ZKGAG
Euglena gracilis	CHSA------QKG-V
Crithidia oncopelti	CHTG------AKGGA
Rhodospirillum rubrum	CHTF------DQGGA
Rhodopseudomonas palustris	CHRA-------D---K
Rhodopseudomonas sphaeroides	CHVIVDDSGTTIAGR
Rhodopseudomonas capsulata	CHSIIAPDGTEIV-K
Paracoccus denitrificans	CHMIQAPDGTDI--K

а



б

Рис. 7.1.

Представление аминокислотных последовательностей.

а — сопоставление положений остатков 27—41 в последовательностях нескольких цитохромов *с* [145]. Прочерки означают пропуски (делеции). *б* — гистограмма вариабельности последовательностей цитохрома *с* [408]. Вариабельность определяется числом различных аминокислот, встречающихся в данном положении, отнесенным к частоте появления обычно наблюдаемой в этом положении аминокислоты.

номера предыдущих остатков (например, 27) с буквами в алфавитном порядке (например, 27А, 27В, ...). Каждые несколько лет в Атласе аминокислотных последовательностей и структур белков [145] производится перенумерация всех последовательностей и дополнительные позиции становятся узаконенными. Однако изменение привычной нумерации остатков приводит к значительным неудобствам при чтении и записи белковых структур. По-видимому, было бы более целесообразно после очередного изменения нумерации сохранять все схемы нумерации неизменными в течение значительного периода времени, например 10 лет.

Малая частота замен остатков в данном положении указывает на структурную и функциональную важность этого положения. Сопоставление структур родственных белков показывает, что чувствительность данного положения в последовательности к замене аминокислот резко различается. Это наблюдение иллюстрируется на рис. 7.1, *б*. Высокая частота замен в определенном положении

свидетельствует о том, что остаток в этом положении не очень существен для структуры или функции белка. Поскольку большинство замен происходит на поверхности белка, характер такой структурной вариабельности может дать также некоторую информацию о том, где находится данный остаток, на поверхности или

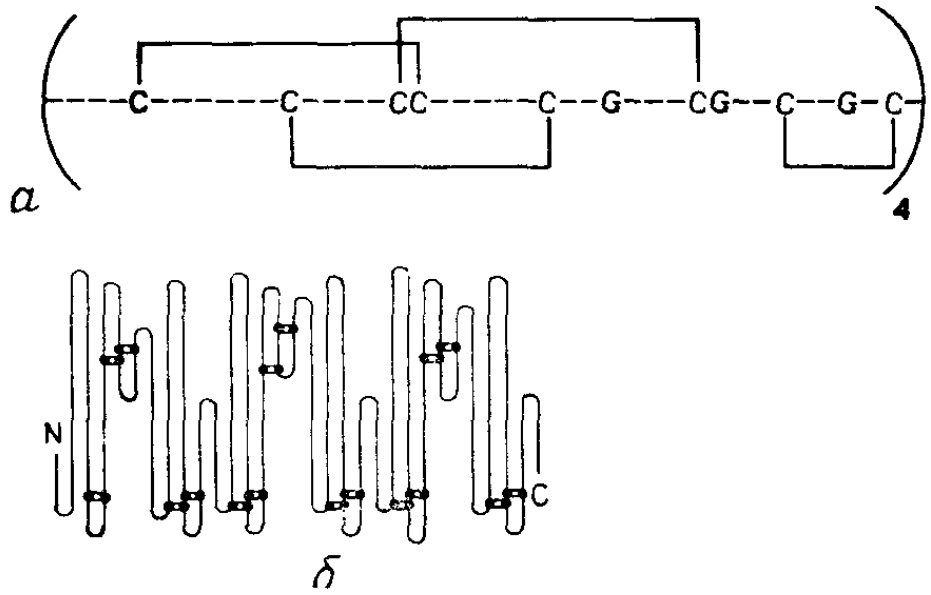


Рис. 7.2.

Представление дисульфидных связей в белках.

а — система связей S—S в агглютинине зериа пшеницы, повторяющаяся четыре раза в одиой полипептидной цепи. Иивариаитные остатки Gly выделеиы. Четыре одинаковые пространствеиные формы обнаружеиы также в трехмерной структуре [316]. *б* — система связей S—S в альбумине из сыворотки человека [409]. Связи повторяются, но не очень точно; амииокислотиая последовательность также обнаруживает повторяющиеся участки. Трехмерная структура белка еще ие известиа.

внутри белка. Это полезно в тех случаях, когда пространственная структура белка еще не известна.

Расположение дисульфидных мостиков выявляет эволюционные связи. Многие белки, в частности внеклеточные, содержат дисульфидные мостики, ковалентно связывающие удаленные части полипептидной цепи. Способы изображения таких связей показаны на рис. 7.2. Мостики S—S обнаруживают тенденцию сохраняться в процессе эволюции, поэтому они характерны для данного семейства гомологичных последовательностей. Более того, по дисульфидным мостикам можно выделить структурные повторения в одной цепи, как в случае агглютинина пшеничного зерна (рис. 7.2, *а*), строение которого было выяснено по карте электронной плотности, так и сывороточного альбумина человека (рис. 7.2, *б*), строение которого установлено химическими методами.

Пространственное строение

У белка, ковалентная структура которого неизвестна, можно проследить укладку основной цепи по карте электронной плотности при эффективном разрешении 3 Å. Без данных о ковалентной струк-

туре результаты рентгеноструктурного анализа кристалла белка (карта электронной плотности) имеют ограниченное значение. При номинальном разрешении 3 Å, т. е. при эффективном* разрешении несколько хуже 3 Å, распознавание полипептидной цепи на карте электронной плотности часто неоднозначно и может привести к неверным выводам. Этим объясняются различия (часто неявные) в геометрической структуре белка, приведенной в разных публикациях. При повышении разрешения, и в особенности после уточнения структуры, цепь можно проследить однозначно. Карта электронной плотности в этом случае до некоторой степени передает индивидуальность остатков, что значительно облегчает анализ последовательности. Даже при эффективном разрешении 3 Å геометрическую структуру можно использовать для идентификации пептидных фрагментов [289], что уже весьма полезно для анализа последовательности.

При эффективном разрешении 3 Å карта электронной плотности позволяет с достаточной точностью найти положения C_α-атомов. Обычно координаты C_α передаются в «банк» данных о белках [390], откуда их можно получить, послав соответствующий запрос. Эти координаты используются также для записи свертывания цепи с помощью точных стереоизображений (разд. 7.2). Документация некоторых способов свертывания цепей выполняется только с помощью структурных эскизов.

Полные структуры

Полная структура представляется в виде таблицы с обозначениями атомов и их координат. Полная структура объединяет данные о ковалентном и пространственном строениях. Ковалентная структура определяет обозначения данного атома, а геометрическая — его координаты. Атомы водорода с их единственным электроном проявляются обычно на карте электронной плотности настолько слабо, что их можно идентифицировать лишь в очень редких случаях. Однако положения атомов водорода определяются в достаточно хорошем приближении по координатам неводородных атомов, к которым они присоединены.

---

* Следует различать номинальное и эффективное разрешения. Номинальное разрешение означает предел разрешения структурных факторов в синтезе Фурье, тогда как эффективное разрешение относится к результирующей карте электронной плотности. Эффективное разрешение всегда ниже номинального. Различие может быть заметным, если точность фаз структурных факторов низка или если набор структурных факторов неполон. К сожалению, оценки эффективного разрешения не существует. В литературе обычно приводится номинальное разрешение.

Полную структуру обычно записывают в виде таблицы, включающей все атомные координаты. В этой таблице каждый атом описывается своим внутренним обозначением в пределах данного остатка, а также названием остатка и его порядковым номером (например, $O_{\varepsilon 2}$ Glu-131). В некоторых случаях указывается также подвижность атомов в кристалле белка. Все эти сведения собраны в банке данных о белках [390], откуда они предоставляются по запросам в виде записей на магнитной ленте. Чтобы использовать эти данные, обычно нужно составить собственную программу для ЭВМ.

7.2 Представление полной структуры

Огромные таблицы, в которых записана полная структура, воспринимать очень трудно; очевидно, что они должны быть представлены в другой, более доступной форме. В этом разделе описаны все использующиеся сейчас способы представления полной белковой структуры.

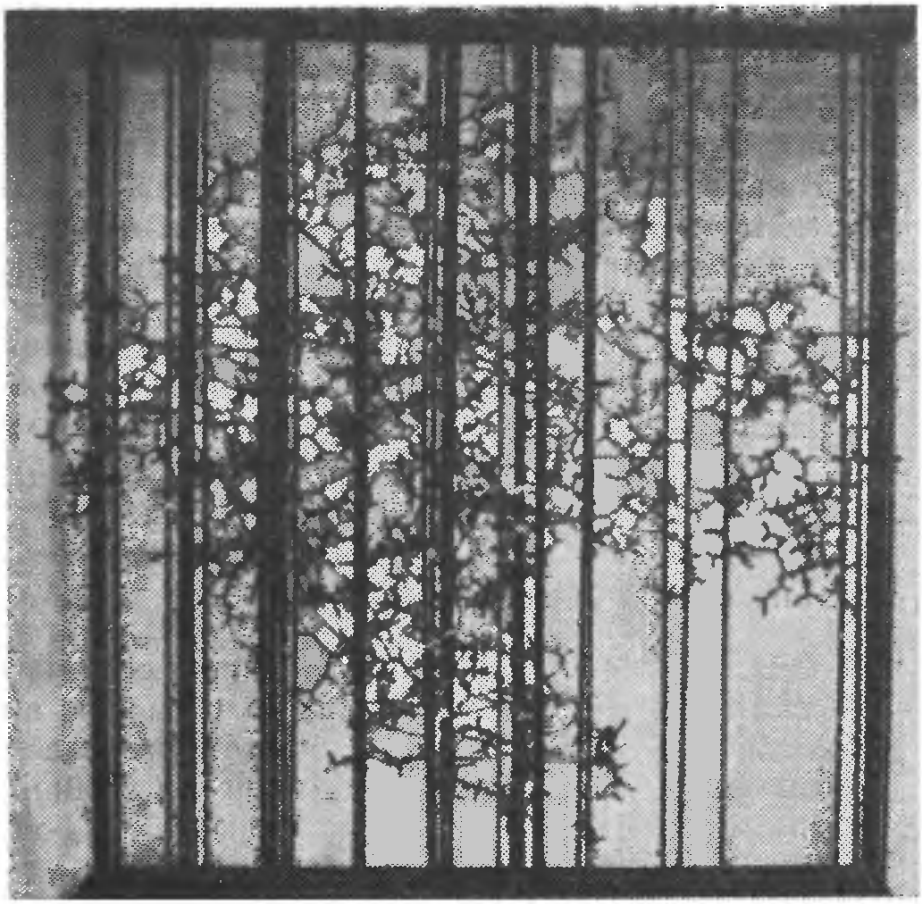
Пространственные модели

Точные молекулярные модели громоздки, но удобны. Самым удобным представлением молекулярной структуры пока еще остается трехмерная модель. Такая модель незаменима при постановке любых экспериментов, основанных на знании полной структуры белка. В частности, она наиболее адекватна для поиска центров связывания субстратов и эффекторов. Однако такие модели громоздки и к тому же стоят очень дорого, чем и объясняется их сравнительно малая распространенность.

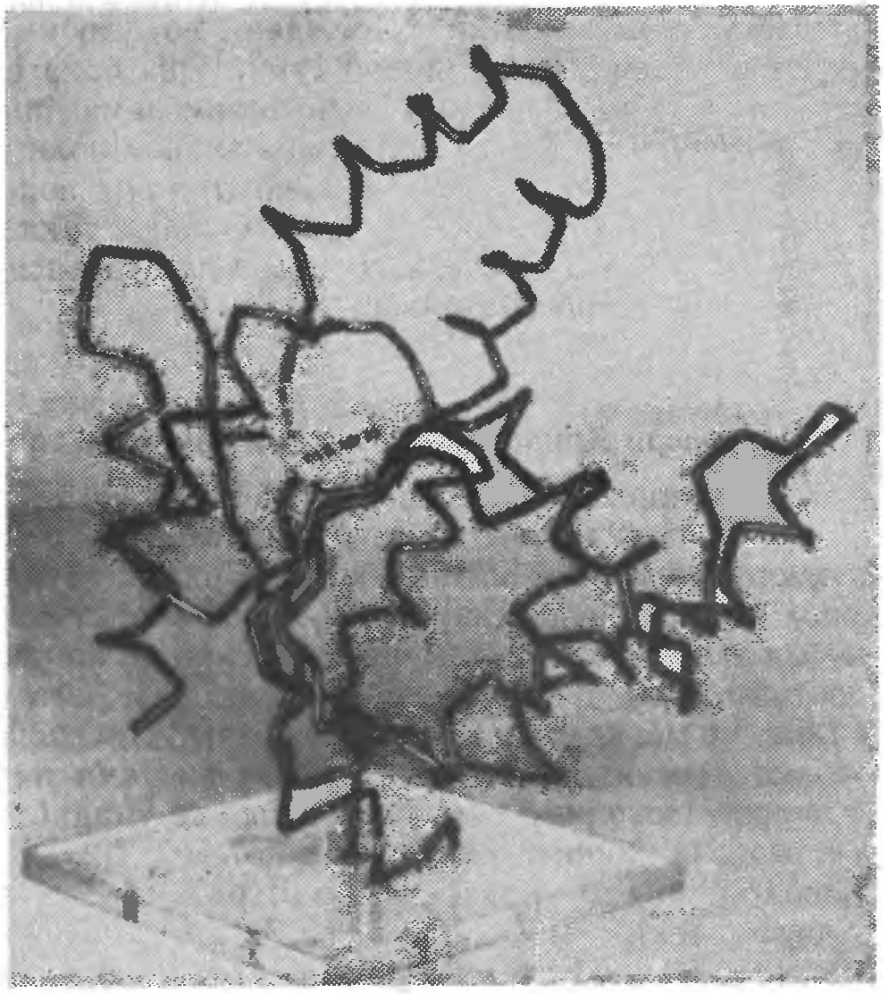
Проволочные модели

Проволочные модели Кэндрью — Уотсона наиболее точны. Проволочные модели Кэндрью — Уотсона [391] в масштабе 1 Å = 20 мм и диаметром проволоки 2 мм повсеместно используются для наиболее точного трехмерного представления белка. Вид такой модели показан на рис. 7.3, *а*. Положения атомов, которые специально не обозначены, соответствуют точкам разветвлений и концам проволочек. Валентные углы, длины связей и пептидные звенья в этой модели жесткие, но могут быть модифицированы. Двугранные углы можно свободно изменять: они фиксируются винтами на соединительных втулках. Поскольку используют только тонкие проволочки и винты, модели хорошо просматриваются.

Как правило, модели Кэндрью — Уотсона строят непосредственно по карте электронной плотности. Для этой цели карту вычерчивают в той же шкале (1 Å = 20 мм) на полупрозрачных пла-

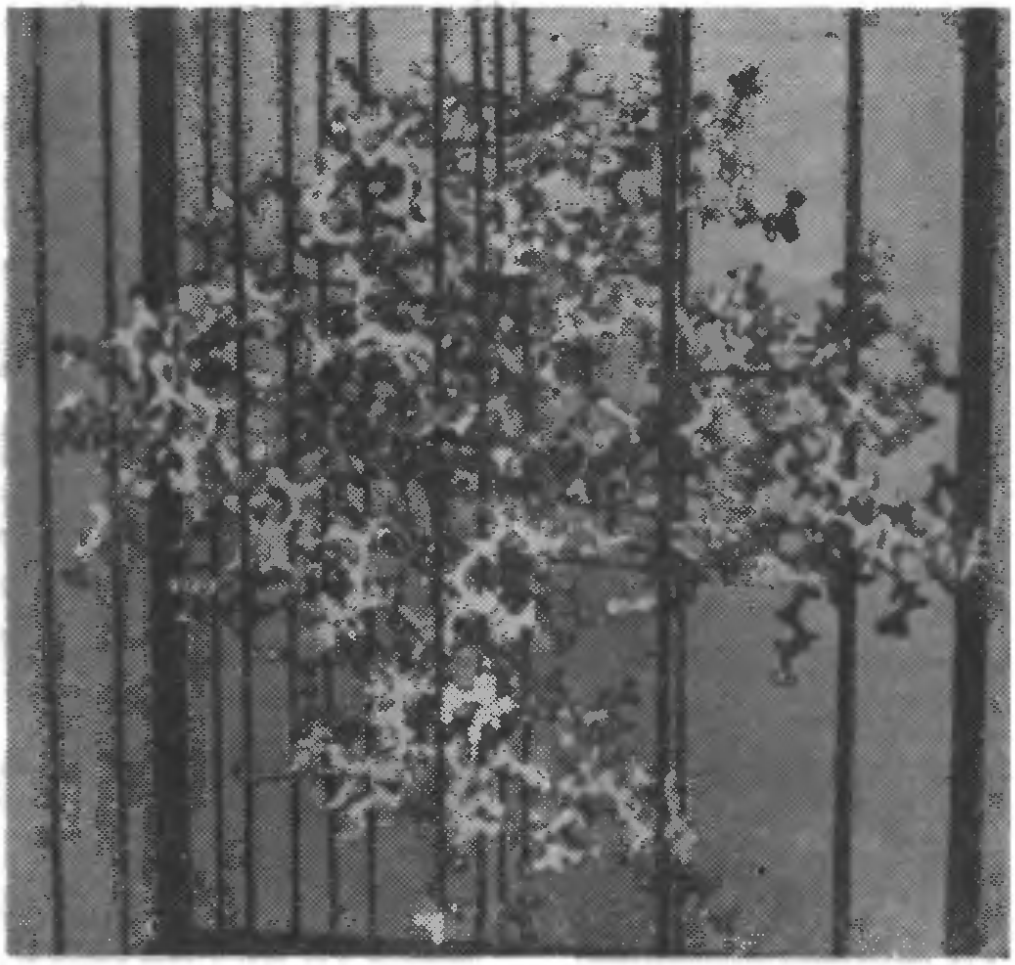


а

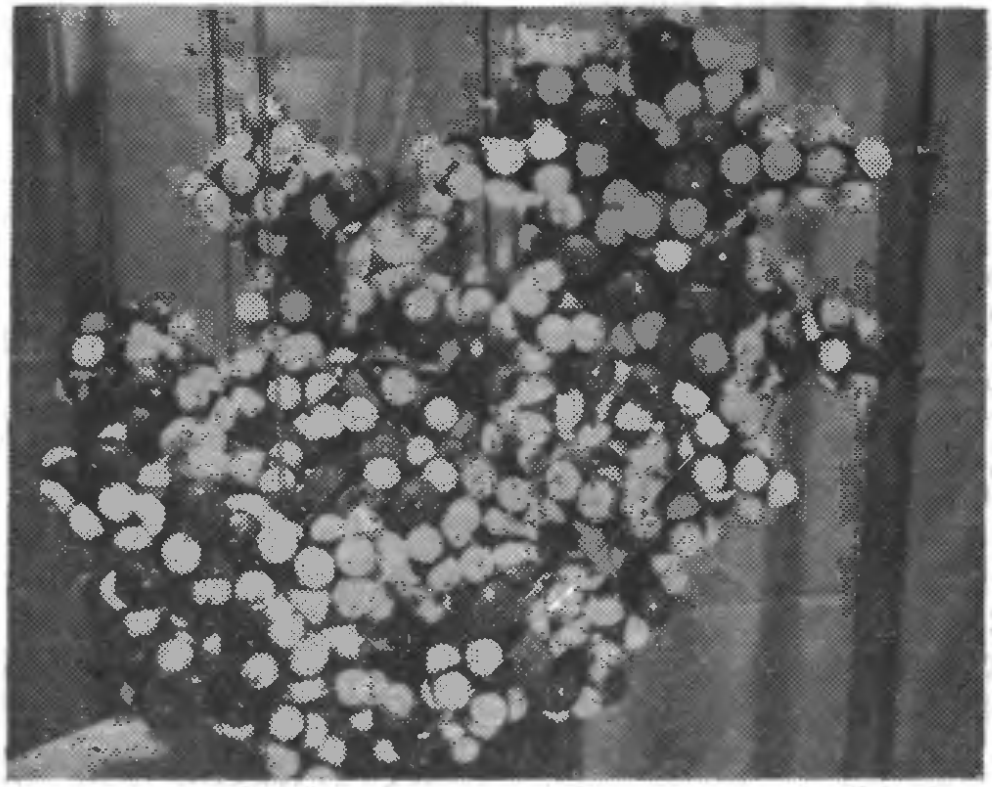


б

Рис. 7.3 (см. с. 164).



в



г

Рис. 7.3 (см. с. 164).

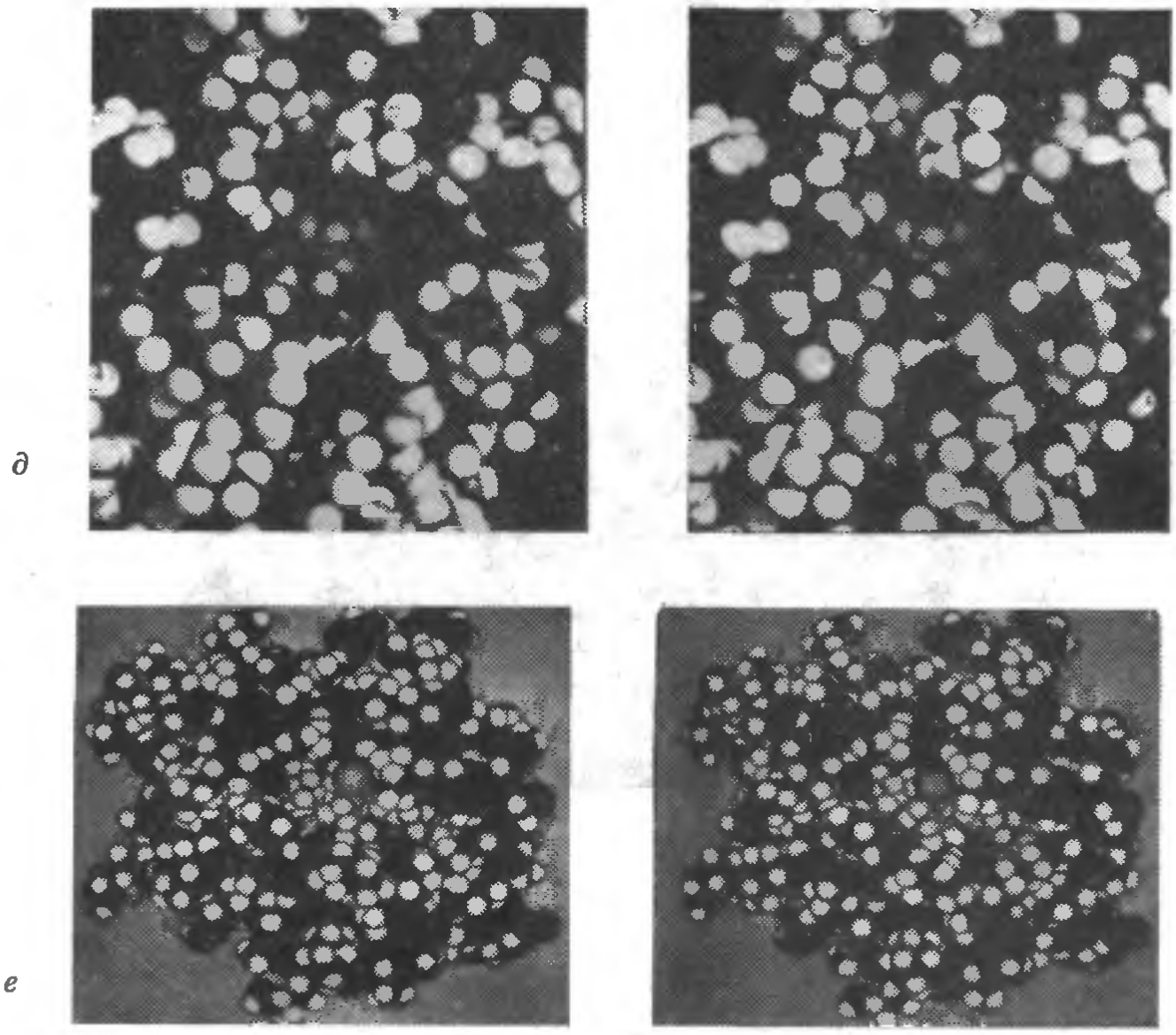


Рис. 7.3.

Представление белковых структур.

а — аденнлаткиназа, построенная с помощью проволочной модели Кэндрью—Уотсона [391] (масштаб 20 мм=1 Å), *б* — проволочная модель Бнрона [393] основиой цепи аденилаткиназы; отмечены положення трех боковых цепей (масштаб 5 мм=1 Å). *в* — аденилаткиназа, построенная из деталей модели Николсона [394] (масштаб 10 мм = 1 Å). *г* — объемная модель (деталн КПК, масштаб 12,5 мм = 1 Å [395]) рибонуклеазы S (представлена Викофом). *д* — стерическое изображение объемиой модели (КПК) активного центра карбоксипептидазы (предоставлеио В. Н. Липскомбом). *е* — стерическое изображение объемиой модели поверхиости молекулы цитохрома *с* (предоставлено Фельдманом и Портером [399]); изображение получено с помощью ЭВМ и выведено иа цветной телевизиоиный экраи.

стиковых пластинах. Затем модель строят «в карте», т. е. модель и карту совмещают с помощью полупрозрачного зеркала при соответствующем освещении [392]. Эта процедура схематически представлена на рис. 7.4. «Прозрачность» модели позволяет видеть в зеркале каждую ее часть, даже когда построена вся молекула. На конечном этапе с помощью тубусов, перемещающихся по трем взаимно перпендикулярным осям, с модели считываются координаты атомов.

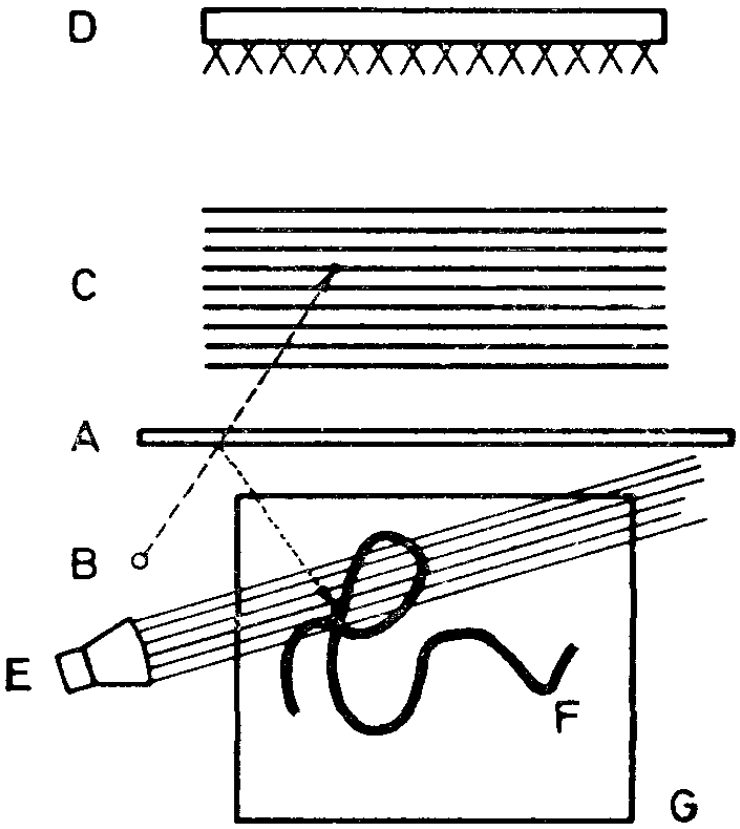


Рис. 7.4.
Оптический компаратор Ричардса (вид сверху) [392].
A — полупрозрачное зеркало, *B* — глаз наблюдателя. Модель и карта совмещаются при наблюдении в зеркале. *C* — карта электронной плотности, нанесенная на прозрачные пластины. *D* — рассеянный свет для освещения карты. *E* — сфокусированный свет, высвечивающий строящуюся часть модели. *F* — строящаяся модель. *G* — жесткий каркас модели.

Скелетные проволочные модели просты для построения. Если нужно получить приблизительное представление о характере свертывания цепи, можно воспользоваться линейной проволочной моделью с небольшими изгибами в положениях C_α. Углы изгиба проволоки (виртуальные валентные и двугранные углы, рис. 7.10) рассчитывают по положениям C_α и переносят на проволоку диаметром 3 мм (1/8 дюйма) с помощью специального устройства [393]. Для такой модели удобно пользоваться масштабом 1 Å = 5 мм. Важные боковые цепи можно присоединить после построения пептидного остова. Модель наглядна, и ее очень легко изготовить. На рис. 7.3, *б* показана такая модель аденилаткиназы.

Пластмассовые модели из легко скрепляемых деталей

Пластмассовые модели позволяют наглядно представлять химические свойства. Другая весьма распространенная модель собирается из вставляемых одна в другую пластиковых деталей, предложенных Николсоном [394]. В такой модели учитываются все атомы: неводородные атомы представлены шариками диаметром 8—10 мм, а атомы водорода — связями, ведущими к ним (диаметр связи 4 мм). Вид модели показан на рис. 7.3, *в*. В масштабе 1 Å = = 10 мм модели Николсона значительно более компактны, чем проволочные модели Кэндрью — Уотсона.

Пластмассовые детали этих моделей окрашены, что позволяет более наглядно представлять свойства отдельных групп белка:

положительно заряженные группы (аммониевые и гуанидиниевые) имеют голубую окраску, отрицательно заряженные (карбоксильные) и гидроксильные группы — красную, сера — желтую, ароматические части — фиолетовую. На этом фоне легко прослеживается основная цепь, окрашенная в белый цвет. В связи с этим модели Николсона очень удобны для тех, кто хочет быстро разобраться в молекулярной структуре.

Модели Николсона далеко не столь точны, как модели Кэндрью — Уотсона. Кроме того, по ним трудно определять водородные связи. Поскольку размеры шариков и палочек сравнительно велики, то модель недостаточно «прозрачна». Поэтому модели Николсона лучше всего строить по координатам атомов, а не по карте электронной плотности (рис. 7.4). Однако, несмотря на эти недостатки, пластмассовые модели позволяют быстро и легко получать довольно подробную структурную информацию; они удобны также и для обучения.

Объемные модели

Объемные модели непосредственно передают поверхность молекулы. Объемные модели очень трудно построить для всей молекулы белка. Поскольку при построении всей модели белка внутренняя часть молекулы становится невидимой, то модели такого типа иллюстрируют только свойства поверхности. В связи с этим их используют преимущественно при выявлении центров присоединения субстрата или эффектора, т. е. для детализации взаимодействий белок — субстрат и белок — эффектор. Вид модели показан на рис. 7.3, *г* и 7,3, *д*.

К сожалению, детали наиболее распространенных объемных моделей, предложенные Кори, Полингом и Колтеном (КПК) [395], построены в масштабе 1 Å = 12,5 мм, который не соответствует ни модели Николсона, ни модели Кэндрью — Уотсона. Поэтому эти модели не могут быть использованы совместно, что безусловно оказалось бы полезным во многих случаях.

Модельные карты

Модельные карты соответствуют картам электронной плотности. Первоначально данные рентгеноструктурного анализа получают в виде карты электронной плотности, представляющей собой стопку прозрачных пластин с контурами плотности, которые затем интерпретируются. Очевидно, что и исследуемую молекулярную структуру можно представить таким же образом. Такой опыт был проделан для рибонуклеазы [396]: все связи отпечатывались в виде полос, а вандерваальсовы огибающие полости — в виде соединяю-

щихся кругов. Однако модельные карты не получили распространения. Очевидно, что по сравнению с другими типами моделей они значительно менее информативны, особенно при рассмотрении взаимодействий белка с другими молекулами.

Компьютерная графика

Проекция вращающегося объекта на телевизионный экран воспринимается как трехмерное изображение. В течение длительного времени графические системы ЭВМ не могли конкурировать с твердыми пространственными моделями. Ограничения, вносимые двумерным экраном катодно-лучевой трубки, были слишком жесткими для удовлетворительного изображения трехмерной структуры. Положение изменилось с появлением «быстрых» катодно-лучевых трубок и «быстрых» электронных устройств, предназначенных для расчета изменения координат в процессе вращения [397]. Такое устройство позволяет при вращении большого набора векторов, рассчитывать проекцию этого набора на данную плоскость (определяющую изображение) и одновременно давать изображение на экране. При скоростях вращения начиная с одного оборота за десять секунд и выше изображение на экране близко к изображению, которое мы получаем рассматривая объект, поворачивающийся в наших руках. Поэтому вращающийся объект (или набор векторов) на телевизионном экране воспринимается в трех измерениях.

Набор векторов на телевизионном экране может передавать и ковалентную структуру, если, например, векторы отвечают проволочкам модели Кэндрью — Уотсона; набор векторов может также изображать карту электронной плотности. Части модели и распределение электронной плотности можно показывать и одновременно, проверяя таким путем соответствие модели карте. Более того, поскольку эти системы взаимосвязаны, можно манипулировать моделью до получения наилучшего соответствия с картой [398]. При этом можно сохранить трехмерное восприятие, вращая модель и карту на экране.

В будущем компьютерная графика может заменить твердые пространственные модели. По своим возможностям графическая система ЭВМ приближается к оптическому компаратору (рис. 7.4). Однако графическая система позволяет показывать карту электронной плотности в любом нужном направлении или на любом уровне, а не только при фиксированных направлениях и уровнях, как это происходит в компараторе. Такие изменения направлений очень важны при рассмотрении распределения электронной плотности.

Объемные модели также можно пресбразовывать на телевизионный экран [397, 399], однако в этом случае расчет внутренних частей занимает столько времени, что не позволяет вращать модель

со скоростью, достаточной для создания трехмерного восприятия. Пример такого дисплея дан на рис. 7.3, *е*.

В настоящее время графические системы применяются лишь в больших научных центрах, которые могут оплачивать и оборудование, и персонал, необходимый для обслуживания и разработки соответствующих программ. Однако стоимость электронного оборудования уменьшается настолько быстро, что, по-видимому, наступит день, когда графические системы заменят оптические компараторы, а может быть, и молекулярные модели.

Двумерные представления

Стереоизображения

Стереоизображения — информативные двумерные представления трехмерного объекта. Для публикации в периодической печати необходимы двумерные представления структуры. Для этой цели весьма удобны стереоизображения. Обычно такие стереоизображения состоят из двух расположенных рядом рисунков, центры которых разнесены на расстояние, равное расстоянию между зрачками (65 мм). Каждый из рисунков нужно рассматривать соответствующим образом; можно применять и специальное стереоустройство [400]. Расстояние между рисунками (65 мм) налагает ограничение на размеры рисунков, так что они не превышают 50 × 50 мм. Очевидно, что при ограниченном разрешении печати не всегда можно вместить необходимый объем информации в такой формат. Можно несколько увеличить объем считываемой информации, печатая и рассматривая стереоизображения в красном и зеленом цветах. Однако взаимное исключение красного и зеленого цветов обычно неполно, к тому же цветная печать требует больших затрат, поэтому красно-зеленые стереорисунки применяются сравнительно редко.

Линейные изображения

Стереоскопическое линейное изображение хорошо описывает структуру. Чтобы представить свертывание цепи в журнальной статье, можно подготовить стереорисунки, в которых все C_α-атомы связаны прямыми линиями (рис. 4.2, *а*). Иногда C_α-атомы обозначаются маленькими кружками и порядковым номером. На таком стереорисунке исключительно хорошо прослеживается аминокислотная цепь, даже если она состоит из более чем 800 остатков, как в случае фосфорилазы [236].

Если нанести на рисунок не только C_α-атомы и виртуальные связи между ними, но и все связи между неводородными атомами, стереоизображение всей молекулы окажется слишком громоздким и потеряет свою наглядность. Поэтому такие детали приводят

только на специальных изображениях отдельных участков молекулы. Постоянно пополняющаяся коллекция стереопроекций различных частей всех известных полных структур белков приводится в атласе микрофишей AMSOM [401]. Несмотря на большую детализацию (около 10^3 изображений на молекулу), Атлас AMSOM дает только стандартные проекции. Более разносторонняя информация

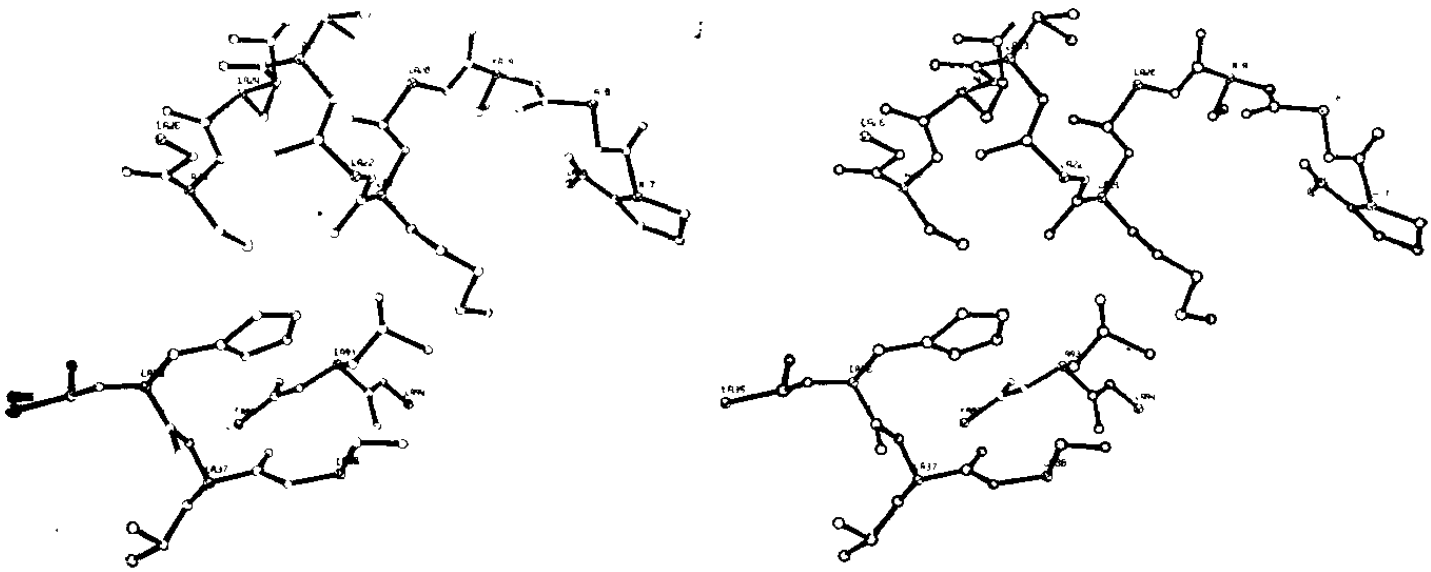


Рис. 7.5.

Стереоизображение ORTEP [324] активного центра аденилаткиназы. Приведены только неводородные атомы; выделены атомы C_α.

получается при изображении молекулы на дисплее катодно-лучевой трубки. Стереоскопического эффекта можно достичь путем разделения экрана на две половины или вращая рассматриваемую молекулу с определенной скоростью [397] (разд. 7.2).

Графическую программу ORTEP можно использовать для построения как гибких, так и объемных моделей. Весьма изящное изображение молекулы может быть получено с использованием программы ORTEP, которую составил Джонсон [324] для изображения малых молекул. Программа ORTEP позволяет получить точное моно- или стереоизображения моделей из шариков и палочек. Она была использована для изображения структуры коллагена (рис. 5.6, *а*). Программа оказывается полезной при анализе функциональных элементов белковых молекул, например активного центра аденилаткиназы на рис. 7.5. Если задать каждый атом сферой с соответствующим вандерваальсовым радиусом, ORTEP дает прекрасное изображение поверхности молекулы в виде фотографии объемной модели (рис. 5.6, *в*). При этом, однако, возможны небольшие ошибки в местах пересечений сфер валентно связанных атомов.

Макеты структур

Макеты отображают лишь существенные части структуры. Поскольку объем графической информации по белку очень велик,

читатель может потерять интерес, если эта информация будет представлена во всех деталях. Во многих случаях бывает вполне достаточно общего наброска, который выявляет лишь существенные части структуры. Этим объясняется довольно широкая распространенность так называемых структурных макетов (наиболее удачные см. в [794]). На таких макетах α-спирали обычно изображают цилиндрами, а β-складчатые листы — стрелками (рис. 7.6). Чтобы избежать грубых ошибок, эти макеты строятся обычно после того, как схема основной цепи выполнена компьютером.

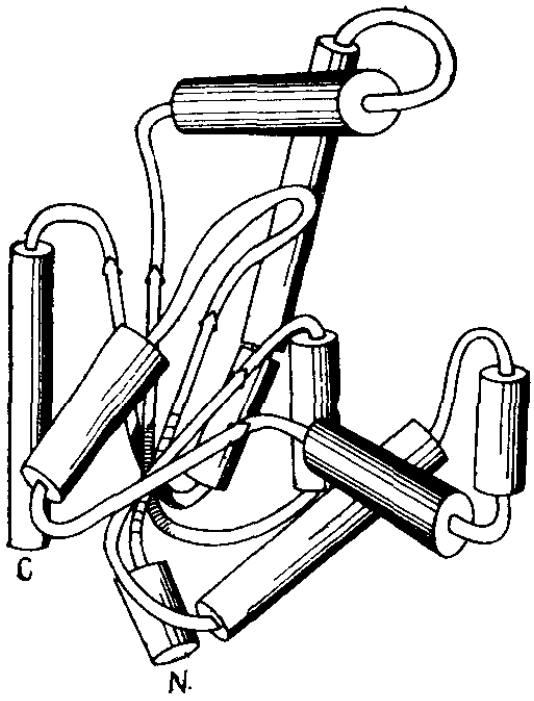


Рис. 7.6.
Структурный макет аденилаткиназы. Спирали изображены в виде цилиндров, а β-складчатые листы — в виде стрелок.

В последнее время широкое распространение [235, 247, 255] получил своего рода молекулярный стриптиз, в результате которого все детали структуры сводятся к топологии β-листов. При этом характер свертывания цепи определяется по направлению и соединениям β-складчатых листов. Как показано на рис. 7.7, в результате получается по существу двумерное изображение. На более детальных схемах [235, 249] в местах соединения между складчатыми листами обозначены α-спирали (рис. 5.17, *д*). С помощью таких макетов удалось выявить ряд структурных особенностей, в частности преобладание правого β⁀β-звена (разд. 5.2). Отметим, что топология β-листа легко совмещается с данными об аминокислотной последовательности в виде простого графа [186].

Структура водородных связей

Макеты можно эффективно использовать для представления стереохимических данных. Помимо общей информации, т. е. обозначений атомов и их координат, при описании полной структуры белка обычно приводят целый ряд важных дополнительных сведений, касающихся водородных связей и других видов невалентных взаимодействий между атомными группами скелета и боковыми цепями, а также данных о любом взаимодействии полипептидной цепи (включая ковалентное) с простетическими группами, кофакторами, субстратами, металлами и другими лигандами, молекулами воды и т. д. Обычно эти сведения объединены в длинный перечень, из которого по мере необходимости извлекается нужная информация. Значительно лучшей формой записи в некоторых случаях могут, однако,

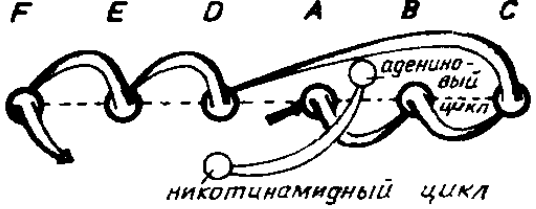


Рис. 7.7.

Топология β-структуры лактатдегидрогеназы и других чувствительных к NAD дегидрогеназ [255].

Скрученная параллельная β-структура (рис. 5.10) не показана. Шесть цепей (круги *A—F*) обращены С-концами к наблюдателю. Все промежуточные участки, согласно определению (рис. 5.12, *а*), являются правыми.

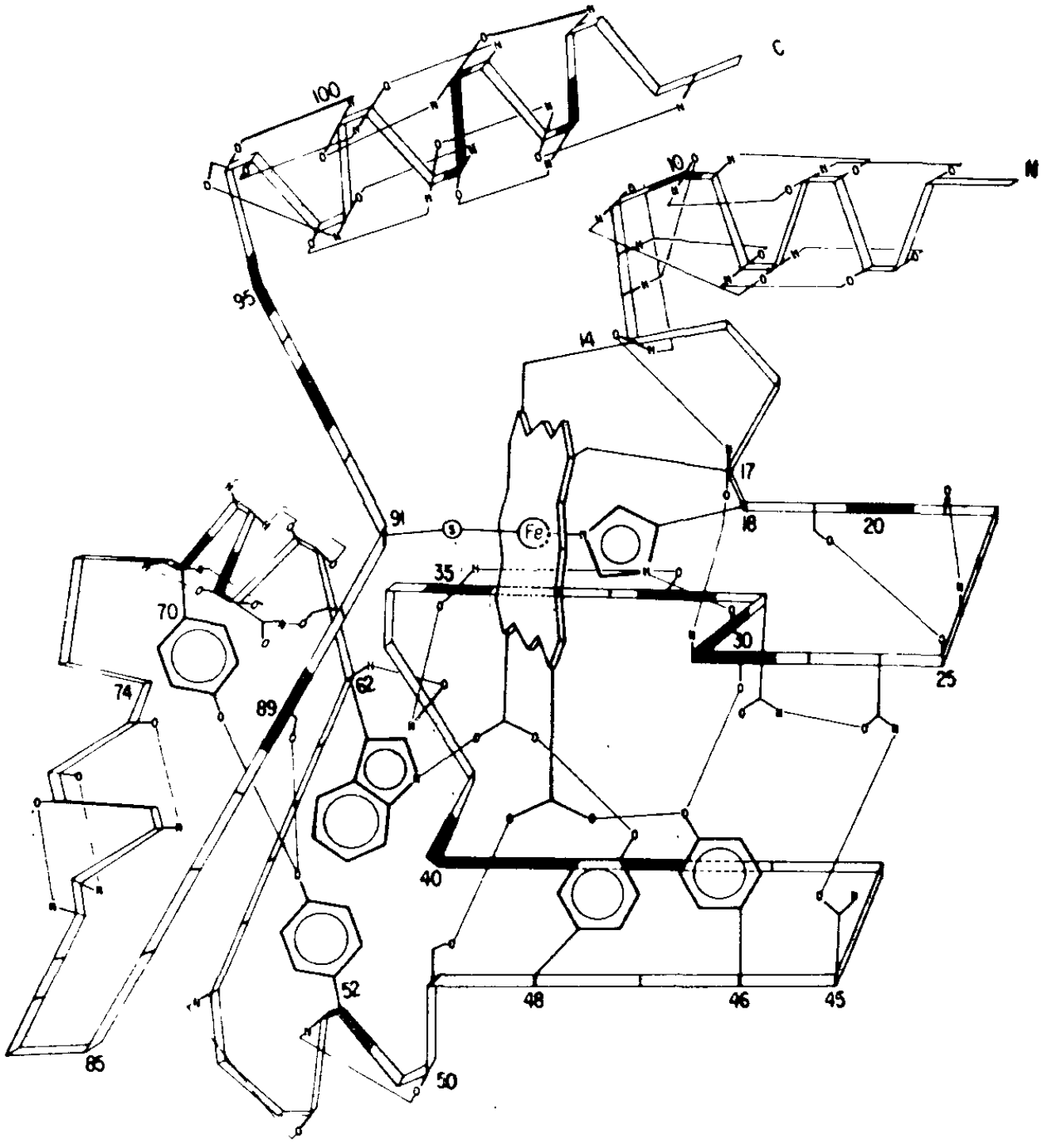


Рис. 7.8.

Схематическое изображение цитохрома c_2 из *Rhodospirillum rubrum* [272] показывает систему водородных связей. Остатки на поверхности полости, в которой находится гем, зачернены. Схема нумерации не совпадает с использованной Дайхофом [145].

служить макеты структур, аналогичные показанному на рис. 7.8, которые предназначены для представления таких химических данных.

7.3 Абстрактные предстaвления о свертывании цепи

До сих пор мы обсуждали запись молекулярной структуры с помощью декартовых координат ее атомов. Однако для некоторых целей, в особенности при теоретическом анализе белковых структур, необходимо привлекать также иные, более абстрактные представления.

Графики для углов осиовной цепи

Основная цепь полностью описывается набором углов (ϕ, ψ). Как обсуждалось в гл. 2, свертывание цепи может быть полностью описано двугранными углами основной цепи ϕ и ψ при каждом C_α-атоме (только ψ для N-концевого C_α-атома и только ϕ для C_α-концевого C_α атома). Эти углы можно представить в виде точек на карте (ϕ, ψ) (рис. 2.4), однако для идентификации углов все точки должны быть снабжены порядковыми номерами C_α-атомов. Такая карта оказывается громоздкой, трудно читаемой, и потому она неудовлетворительно представляет структуру.

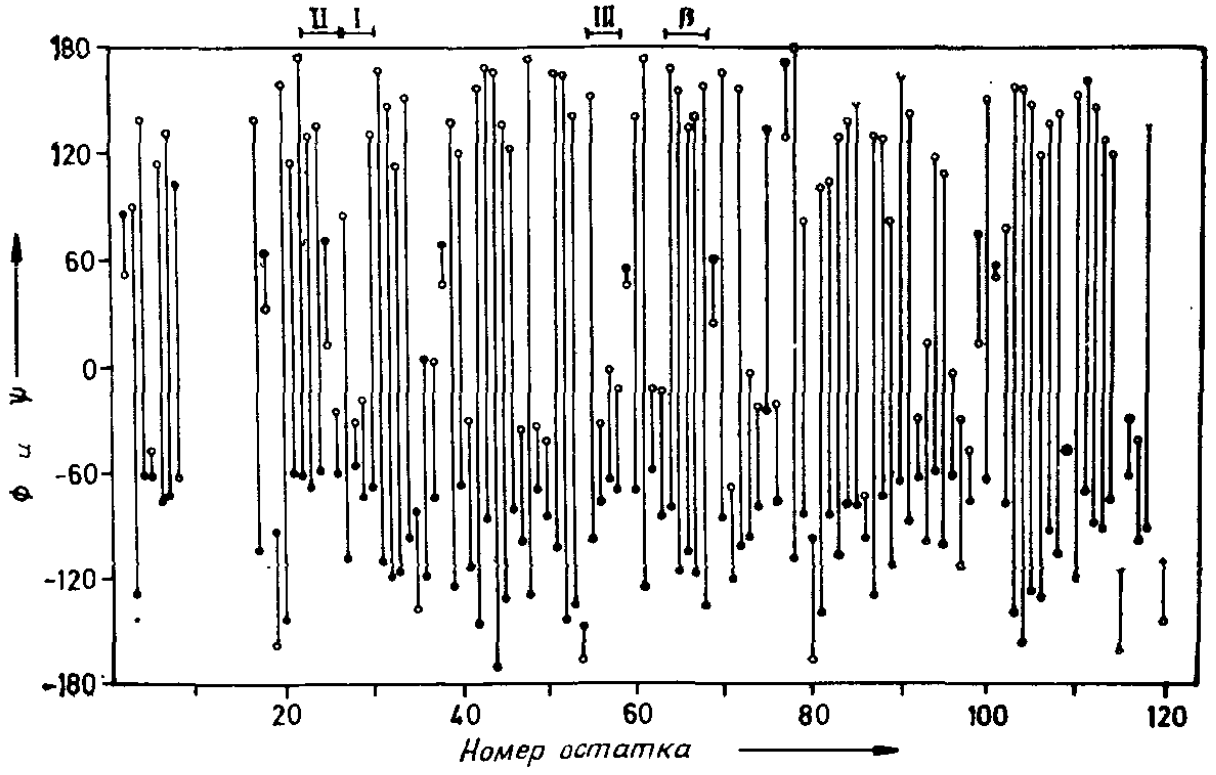


Рис. 7.9.

Линейное представление двугранных углов основной цепи N-концевой части химотрипсина [402].

Черная точка отвечает значению ϕ, кружок — ψ; для каждого остатка значення двух углов соедннены. Отчетливо проявляются вторичные структуры. На рисунке обозначены реверсивные повороты цепи тнпов I, II, III и β-структура. α-Спирали обнаружнваются последовательнымн короткнмн лнниямн вблнзн —60°.

Более удобный способ был предложен Баласубраманианом [402]. На основании карты (ϕ, ψ) он получил линейный график зависимости углов ϕ и ψ от порядкового номера остатка, как это показано на рис. 7.9. График имеет то преимущество, что с его помощью можно легко установить углы (ϕ, ψ) данного остатка. Кроме того, он эффективно выявляет области вторичной структуры.

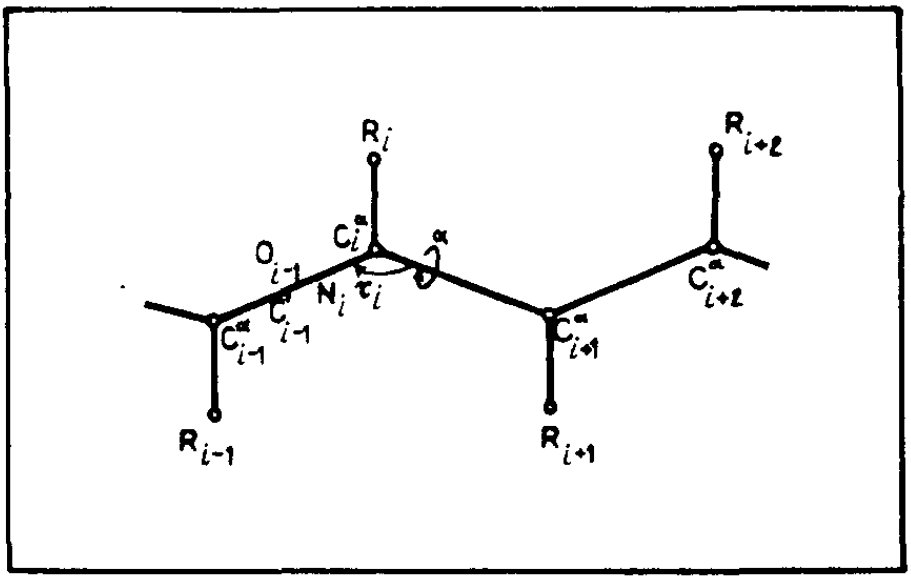


Рис. 7.10.

Представление основной цепи виртуальными связями между атомами C_α. Цепь полностью описывается значеннями виртуальных двугранных углов α_i, определяемыми атомами C_{i-1}^α, C_i^α, C_{i+1}^α, C_{i+2}^α, и значениями виртуальных валентных углов τ_i. Такое опнсание использовано для построения проволочной модели основной цепи (рис. 7.3, *в*). В упрощенном представлении τ_i можно рассматривать как функцию α_i [30] и осиовная цепь в этом случае будет опнсываться только одним независимым параметром на остаток.

Виртуальные связи C_α—C_α упрощают представление свертывания цепи. При другом способе представления свертывания цепи пептидный остов обозначается только C_α-атомами и виртуальными связями между ними, как показано на рис. 7.10. Укладка цепи в этом случае описывается виртуальными валентным и двугранным углами [30, 403]. Геометрия пептидной связи ограничивает значение виртуального валентного угла интервалом приблизительно от 80° до 160°, тогда как на виртуальный двугранный угол первоначально не налагается никаких ограничений. Однако из-за стерических трудностей может реализоваться примерно половина возможных комбинаций углов. Это представление было использовано для описания свертывания цепи [404]; оно полезно также при изготовлении проволочных моделей (рис. 7.3, *в*). Более подробное представление основной цепи с использованием виртуальных C_α-связей обсуждается в [403] и [405].

Карты расстояний между C_α-атомами

Свертывание цепи можно представить картами расстояний C_α—C_α. Координаты атомов C_α, определяющие свертывание цепи, полностью задаются картой расстояний между атомами C_α, поскольку все координаты (кроме трех) можно легко получить из этих расстояний тригонометрически*. На карте взаимных C_α-расстояний

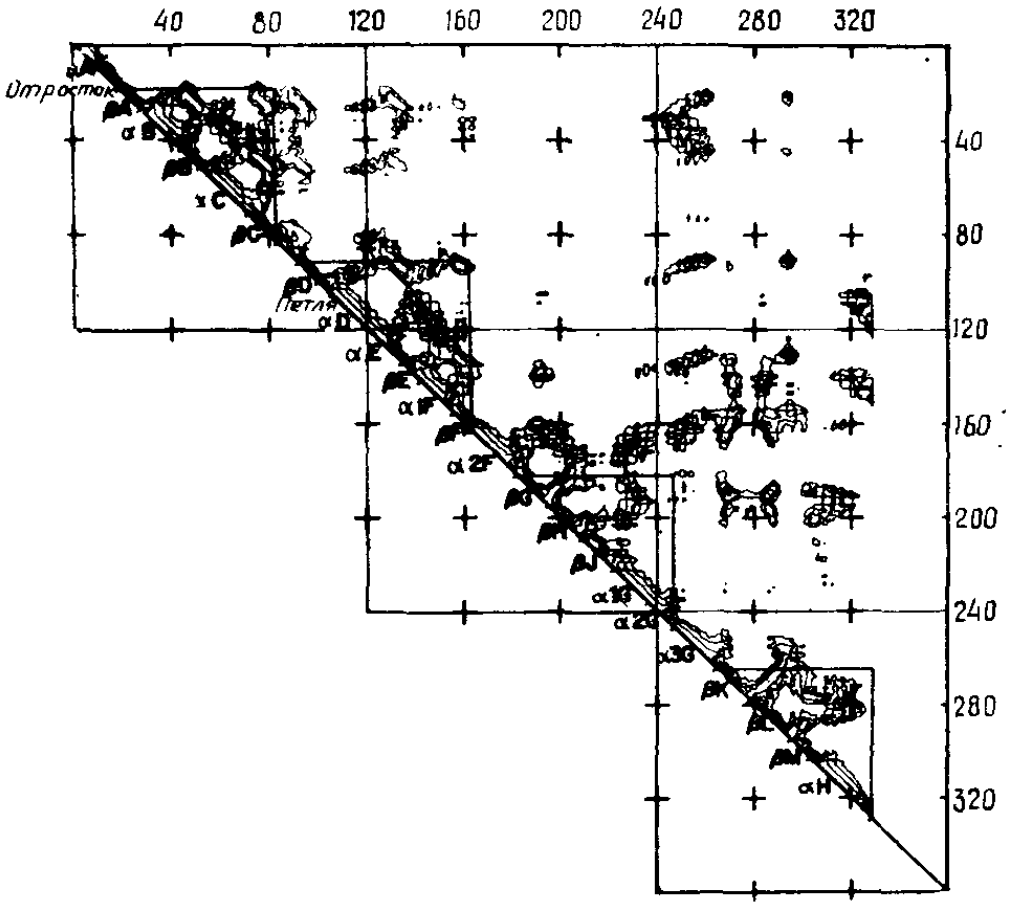


Рис. 7.11.

Диаграмма межостаточных контактов лактатдегидрогеназы [407].

Расстояния между C_α -атомами i-го и j-го остатков представляются треугольной матрицей. Результаты даны не в цифровом, а в контурном виде; контуры проведены при 16, 12, 8 и 4 Å. Различные структурные элементы идентифицируются относительно диагонали; треугольники отвечают четырем геометрически хорошо выделяющимся полудоменам. Два треугольника со стороны N-конца представляют βαβαβ-звенья.

в определенных областях точки располагаются чрезвычайно близко друг к другу. Поэтому можно использовать контурные изолинии равных расстояний, аналогичные контурам на карте электронной плотности. Таким способом можно представить свертывание цепи с помощью двумерного графа.

Вторичные структуры обнаруживаются по характерным областям на карте расстояний. α-Спиралям соответствуют малые рас-

---

* При наличии $N \cdot (N + 1)/2$ расстояний, (т. е. уравнений) для $3 \cdot (N - 1)$ неизвестных координат задача может быть решена однозначно для трех C_α-атомов и оказывается переопределенной для четырех и более C_α-атомов.

стояния (много контурных линий) вдоль диагонали, параллельным β-складчатым листам — малые расстояния вдоль линий, параллельных диагонали (рис. 7.11). Для антипараллельных β-складчатых листов, а также реверсивных поворотов характерны области малых расстояний в направлениях, перпендикулярных диагонали [406]. Расположение линий на карте C_α-расстояний является своеобразным «отпечатком пальца» укладки цепи. Более того, включения и делеции лишь частично нарушают общий рисунок за счет включения или вычеркивания строчки матрицы. Поэтому такие карты можно использовать для сопоставления свертывания цепей [407].

Другие зависимости. График корреляции соседних остатков на рис. 5.14, *а* также представляет собой не что иное, как весьма своеобразный график расстояний C_α. Однако с его помощью невозможно получить исходные координаты.

Льюис и сотр. [202] построили график зависимости расстояний между C_α^i и C_α^{i+3} от порядкового номера i. Такой график позволяет выявить все повороты полипептидной цепи (разд. 5.1); он имеет также то преимущество, что указывает повороты, которые не отвечают типам, перечисленным Венкатачаламом [199]. Однако в этом случае также нельзя вернуться к исходным координатам, поскольку имеется только N—3 расстояний для $3N$—3 координат.

График радиальных расстояний от данной точки молекулы, например остатка Trp, может представить интерес при соответствующем спектральном исследовании. На таком графике цепь развернута по замкнутому кругу и расстояния от всех атомов C_α^i до выбранной точки наносят радиально.

Схематичный график расстояний, представляющий перечень контактов между атомами различных субъединиц олигомерного белка, показан на рис. 5.18, *г*.

Все рассмотренные в этом разделе графики расстояний C_α легко запрограммировать, пользуясь C_α-координатами, предоставляемыми, например, «банком» данных о белках [390].

Заключение

Структуры белков содержат громадный объем информации, записанной в виде наименований и координат многих тысяч атомов. Такое количество данных может быть обработано только с помощью электронно-вычислительной техники. Однако поскольку исследователь должен оценивать, сопоставлять и усваивать обработанные структурные данные, необходимы наглядные изображения пространственных структур. Такие представления должны передавать в основном только те структурные особенности, которые важны при конкретном исследовании. Предельно упрощенной моделью пространственного строения белка является структурный макет.

Применяющиеся описания структур можно классифицировать в соответствии с их размерностью. Трехмерные модели расположения атомов и связей наиболее удобны при исследованиях взаимодействий белка с лигандами, например субстратами и кофакторами или другими молекулами. Очевидно, что, если в течение длительного времени ведется работа с определенным белком, структура которого известна, имеет смысл построить модель этого белка по координатам атомов. При предоставлении материала в печати используются двумерные изображения. С помощью некоторых двумерных изображений можно получить трехмерное представление, пользуясь, например, стереорисунками или проецируя вращающийся объект на телевизионный экран. Помимо моделей, фиксирующих положение атомов и цепей белка в пространстве, существует целый ряд представлений, большей частью двумерных или даже одномерных, с помощью которых можно описывать различные аспекты строения или функционирования белка.

Глава 8 Термодинамика и кинетика свертывания полипептидной цепи

В предыдущих разделах рассматривались изолированные полные структуры белков. Обсудим теперь спонтанное образование этих структур в процессе или после синтеза полипептидных цепей. Хотя некоторые аспекты процесса свертывания можно описать с помощью равновесной термодинамики, такое описание не исчерпывает всех сторон этой сложной проблемы. Для установления корреляций между аминокислотной последовательностью и трехмерной структурой необходимо понимание также и кинетики свертывания. В этой главе обсуждаются экспериментальные и теоретические данные, касающиеся термодинамического и кинетического аспектов образования белковых структур.

8.1 Термодинамические аспекты

Переход между двумя термодинамическими состояниями цепи

Процесс свертывания может быть описан с помощью обычных понятий физической химии. В процессе синтеза на рибосоме или после его завершения полипептидная цепь свертывается в свою нативную глобулярную структуру. Как правило, процесс свертывания происходит самопроизвольно в том смысле, что для его осуществления не требуется действия дополнительных факторов, например ферментов или присутствия рибосом. Наиболее убедительное подтверждение именно такого характера свертывания было получено после полного химического синтеза фермента рибонуклеазы [410, 411].

Принцип спонтанности был первоначально сформулирован на основе экспериментов по ренатурации, которые свидетельствовали о том, что после денатурации нативного белка и последующего удаления денатурирующего агента цепь повторно и самопроизвольно свертывается в исходную конформацию [94, 412]. Однако обоснованность такого заключения подвергалась определенным сомнениям вследствие того, что денатурация белка могла быть неполной.

Поведение синтетической рибонуклеазы и денатурированных белков показывает, что самопроизвольно свертываются не только образующиеся, но уже полностью синтезированные цепи. В связи с этим процесс свертывания можно рассматривать как переход системы «цепь плюс растворитель» из состояния с более высокой свободной энергией (R, случайная конформация) в состояние с более низкой свободной энергией (N, нативная конформация). Для описания этого процесса достаточно обычных понятий физической химии.

При образовании нативной структуры в вакууме изменение связывающей энергии превосходит изменение энтропии. Переход из состояния R в состояние N описывается уравнением (3.2), приведенным в разд. 3.5. Рассмотрим вначале соотношение между энтропией цепи и связывающей энергией цепи, $\Delta S_{\text{цепь}}$ и $\Delta H_{\text{цепь}}$ (уравнение (3.2)). Возьмем полипептидную цепь из 100 остатков, образующую α-спираль. Энтропия цепи $S_{\text{цепь}} = R \cdot \ln W$, где W — вероятность возникновения данной конформации, а R — газовая постоянная. Следовательно, различие энтропии между состояниями N и R связано с вероятностями конформаций, реализующихся в этих системах:

$$\Delta S_{\text{цепь}} = S_{\text{N}} - S_{\text{R}} = R \cdot \ln W_{\text{N}} - R \cdot \ln W_{\text{R}} = R \ln \frac{W_{\text{N}}}{W_{\text{R}}}$$

Для сравнения случайного и α-спирального состояний требуется оценить отношение $W_{\text{N}}/W_{\text{R}}$. В R-состоянии двугранные углы при данном C_α-атоме могут принимать любые значения в разрешенных областях карты (φ, ψ) (рис. 2.3, *б*), тогда как в N-состоянии возможны лишь значения (φ, ψ) в узкой области (—60°, —60°). Конформации боковых цепей не учитываются, т. е. полагают, что в обоих состояниях они имеют близкие значения энтропии. Если оценить α-спиральную область как 10% общей разрешенной области (рис. 2.3, *б*), то вероятность реализации у данного остатка α-спиральных углов составит 10%. Для всей цепи, состоящей из 100 остатков, вероятность будет равна приблизительно $(0{,}1)^{100}$, что дает

$$\Delta S_{\text{цепь}} = R \cdot \ln \left(\frac{1}{10}\right)^{100} = -0{,}46 \text{ ккал} \cdot \text{град}^{-1} \cdot \text{моль}^{-1} \qquad (8.1)$$

Если пренебречь величиной $\Delta G_{\text{растворитель}}$, т. е. рассматривать полипептидную цепь *в вакууме*, то получим

$$\Delta G_{\text{цепь}} = \Delta H_{\text{цепь}} + 143 \text{ ккал/моль (при } T = 37^\circ\text{C)}$$

Поэтому образование спирали индуцируется (величина $\Delta G_{\text{цепь}}$ становится отрицательной) в том случае, когда приходящаяся на остаток связывающая энергия превышает 1,43 ккал/моль, т. е. энергию, легко получаемую при образовании водородной связи. В этом примере изменение температуры на 1°C отвечает изменению кон-

станты равновесия K между нативной и случайной конформациями в 2,1 раза, что следует из уравнения (3.1). Таким образом, даже малые изменения температуры оказывают сильное воздействие на положение равновесия в том смысле, что при высоких температурах цепи легче компенсируется связывающая энергия и принимается случайная конформация. Такое положение свойственно системе с большими, но хорошо сбалансированными вкладами энтропии и связывающей энергии.

Высокоподвижные цепи не могут образовывать определенные структуры. Предшествующие замечания позволяют нам обсудить роль β-аминокислот при образовании белковых структур (разд. 1.2). В β-аминокислотах возможно свободное вращение вокруг связи C_α —C_β (рис. 1.4). Поэтому для одного остатка вероятность принятия им вполне определенного строения в основной цепи составляет не 0,1, а только около 0,003 (если рассматривать поворот на приблизительно каждые 10° вокруг связи C_α—C_β как отдельную структуру), и величина $\Delta S_{\text{цепь}}$ для 100 остатков становится равной около—1,2 ккал · град$^{-1}$ · моль$^{-1}$, что в 2,5 раза больше, чем для α-аминокислот в уравнении (8.1). Компенсировать такую большую энтропию цепи почти невозможно как одной величиной нековалентной связывающей энергии *в вакууме*, так и в совокупности с энтропией растворителя.

Таким образом, цепь β-аминокислот, по-видимому, не способна самопроизвольно принимать определенную структуру. Эти же соображения применимы и ко всем другим типам цепных молекул; высокая подвижность препятствует образованию определенной структуры.

Баланс энергетических вкладов в глобулярном белке

Глобулярные белки — точно сбалансированные системы. Приведенные выше оценки дают представление о балансе термодинамических вкладов. Однако этот баланс неполон, поскольку отсутствует вклад $\Delta G_{\text{раств}}$. Для рассмотренного примера величина $\Delta G_{\text{раств}}$ довольно мала, так как образование спирали приводит к удалению от растворителя только полярного остова, но не крупных гидрофобных боковых цепей (разд. 3.5).

Рассмотрим теперь «средний» глобулярный белок, состоящий из 150 аминокислотных остатков. Как следует из уравнения (8.1), величина $T \cdot \Delta S$ составляет в этом случае несколько сотен килокалорий на моль. Для приведения пептидной цепи и боковых цепей в нативную конформацию эту энергию должны компенсировать величины $\Delta G_{\text{раств}}$ и $\Delta H_{\text{цепь}}$ (уравнение (3.2)). Судя по числу остатков, удаляемых из растворителя, и по вкладам свободной энергии этого процесса (рис. 1.8), величина $\Delta G_{\text{раств}}$ составляет по крайней мере 100 ккал/моль. Вклад $\Delta H_{\text{цепь}}$ обеспечивается водородным

связыванием во внутренней части белка (90% всех внутренних полярных групп которого образуют водородные связи [171) и вандерваальсовыми взаимодействиями. Результирующее значение $\Delta G_{общ} = \Delta H_{цепь} - T\Delta S_{цепь} + \Delta G_{раств}$ составляет лишь небольшую долю от величин компенсирующихся членов, или приблизительно 10 ккал/моль.

Такой порядок абсолютной величины $\Delta G_{общ}$ процесса свертывания был найден экспериментально для большого числа белков, содержащих около 150 остатков [413]. Для оценки $\Delta G_{общ}$ привлекались скорости водородного обмена в состояниях N и R [414—416], калориметрические данные процесса деструктурирования цепи, данные, полученные из кривых денатурации (см. обзор в [413]), а также константы равновесия K между нативной и случайной конформациями, найденные при иммунологических исследованиях [418]. Поскольку величина $\Delta G_{общ}$ мала, любые энергетические расчеты, преследующие цель установления корреляции между ковалентной и геометрической структурами белка, должны быть исключительно точными, чтобы дать значимые результаты. Поскольку невалентные связывающие силы в белке поняты еще ие достаточно, достичь такую точность пока еще трудно (гл. 3). Нужно отметить также, что константа равновесия очень чувствительна к температурным изменениям.

Тепловые флуктуации в белковых структурах

Изолированные молекулы белка претерпевают большие структурные флуктуации. При любом термодинамическом описании структуры белка необходимо иметь в виду, что отдельная белковая молекула с окружающим ее растворителем представляет собой очень малую систему [419]. При данной температуре среднеквадратичное значение флуктуаций энергии δH системы составляет согласно [420]:

$$\sqrt{\overline{\delta H^2}} = \sqrt{kmCT^2}$$

Для макроскопических систем эти флуктуации по сравнению с внутренней энергией в самом деле очень малы. Однако для отдельной белковой молекулы с массой m около $5 \cdot 10^{-20}$ г и удельной теплоемкостью $c = 0{,}32$ кал $\cdot$ град$^{-1}$г$^{-1}$ [421, 422] при 37°С флуктуации составляют $7 \cdot 10^{-20}$ кал (k-постоянная Больцмана), что соответствует приблизительно 40 ккал/моль при пересчете в молярном базисе. Такие флуктуации значительно превосходят величину $\Delta G_{общее}$. Однако времена релаксации составляют всего наносекунды, что недостаточно для индуцирования значительного развертывания нативной структуры.

Тем не менее такие флуктуации могут наблюдаться экспериментально. Одним из примеров служит динамическое поведение боковых

цепей Tyr в ингибиторе трипсина поджелудочной железы. Согласно данным рентгеноструктурного анализа, эти боковые цепи жестко упакованы, однако спектры ядерного магнитного резонанса свидетельствуют о колебаниях этих боковых цепей между двумя состояниями, отличающимися поворотом на 180° вокруг связи $C_\beta - C_\gamma$ [423, 424]. Эта же методика использовалась для доказательства конформационной лабильности на первый взгляд жестких остатков в лизоциме [425], рибонуклеазе [426] и миоглобине [427]. Кроме того, было установлено, что флуоресценцию скрытого в рибонуклеазе Trp можно подавить действием O_2 [428, 429], что свидетельствует о способности O_2 проникать внутрь молекулы. Аналогичные выводы получены из опытов по релаксации флуоресценции [430], фосфоресценции [431] и водородному обмену [414]. Все эти, а также некоторые другие данные [432] показывают, что молекула белка достаточно подвижна. Атомы могут смещаться относительно средних положений, устанавливаемых рентгеноструктурным анализом. Отметим, что олигомерные белки (и белковые кристаллы, по которым получена большая часть данных) имеют большую массу m, чем мономерные. Поэтому в них меньше флуктуации, приходящиеся на единицу массы.

Нативное состояние: глобальный или локальный энергетические минимумы?

Если белки только метастабильны, их структуры должны сильно отличаться от наиболее стабильных структур. В течение длительного времени в литературе обсуждается вопрос, отвечает ли нативная структура абсолютному (глобальному) минимуму свободной энергии (термодинамическая гипотеза свертывания белка [433]) или только локальному минимуму (кинетическая гипотеза свертывания белка [248, 434]), т. е. метастабильному состоянию. Предполагается, что самым стабильным состоянием должен был бы быть, например, сложный узел (рис. 5.15, *в*), который цепь самопроизвольно в действительности образовать не может. На языке термодинамики это означает, что цепь не в состоянии преодолеть высокий барьер $\Delta G^{\neq}$, свободную энергию активации. Этот барьер включает большой энтропийный вклад из-за исключительности конформаций, допускающих образование узла путем прохождения одного конца цепи через петлю, образованную другим концом. К тому же такие конформации цепи оказались бы довольно неустойчивыми, поскольку потеря энтропии не будет скомпенсирована связывающей энергией или свободной энергией растворителя [уравнение (3.2)], как это имеет место в нативной структуре. Поэтому барьер $\Delta S^{\neq}$ почти полностью определяет барьер $\Delta G^{\neq}$. Наконец, нативная структура хорошо описывается метастабильным состоянием с очень большим временем жизни. Однако ни один экспериментальный метод не в состоянии различить стабильное и метастабильное состояния. Более того, это не имеет и какого-либо биологического значения. Однако метастабильное

состояние должно *значительно отличаться* по своей структуре от наиболее стабильного состояния, чтобы противостоять описанным выше большим энергетическим флуктуациям.

Доказать термодинамическую гипотезу свертывания очень трудно. Обычно полагают, что термодинамическая гипотеза подтверждается экспериментами по повторному свертыванию панкреатической рибонуклеазы [435]. В этих опытах восстановленная несвернутая рибонуклеаза повторно окислялась в присутствии 8 М мочевины, причем образовывалась случайная система водородных связей. В результате получалось около 100 различных продуктов, каждый из которых характеризовался своим набором из 4 дисульфидных связей (возможны $(2\cdot4)!/2^4\cdot4! = 105$ систем связей S—S [436, 437]; кроме того, существуют различные теоретические возможности образования петель. Удаление мочевины и добавление меркаптоэтанола (рис. 4.3) приводило к постепенному и в конечном счете количественному образованию нативной структуры. Этот факт показывает, что исходя из 100 различных групп, находящихся в исходных конформациях, можно получить нативную структуру. Хотя это число и невелико, данный опыт свидетельствует об уникальности продукта свертывания, однако он не позволяет решить вопрос о локальном или глобальном минимуме.

8.2 Быстрота, безошибочность и ограничения свертывания in vitro

Свертывание может происходить значительно быстрее, чем синтез цепи. Свертывание in vitro осуществляется чрезвычайно быстро, по крайней мере для малых белков, не содержащих дисульфидных мостиков. Нуклеаза стафилококка повторно свертывается в течение 1 с [438], а метмиоглобин — в течение 10 с [439]. Если эти величины применимы также и к условиям in vivo, свертывание цепи может происходить по крайней мере в 10 раз быстрее, чем биосинтез аминокислотной последовательности. Дисульфидсодержащие белки, например панкреатическая рибонуклеаза, повторно свертываются за время от 1 до 10 с, если дисульфидные связи не были разорваны в процессе предшествующей денатурации [440]. Однако если такие белки развернуты и восстановлены, последующее свертывание цепи (которое включает образование правильной системы дисульфидных связей) продолжается при оптимальных условиях в течение многих минут.

Безошибочность свертывания цепи in vitro [94] проверяют путем сопоставления свойств нативного и ренатурированного белков в отношении биологической активности и специфичности [441]. Например, денатурированный различными способами переносчик кислорода — гемоглобин может быть вновь переведен в нативный белок, который а) имеет ту же растворимость, что и исходный белок, б) способен кристаллизоваться, в) имеет спектр поглощения, харак-

терный для нативного гемоглобина, г) может обратимо присоединять кислород, д) обладает тем же относительным сродством к двуокиси углерода и к кислороду и е) с трудом расщепляется трипсином, что указывает на компактную структуру ренатурированного белка.

Повторное свертывание белков, обработанных ферментами, обычно невозможно. Если белки подвергались ферментативному воздействию in vivo, их безошибочное повторное свертывание часто становится невозможным. Примером такого белка является коллаген. Одна из многих проблем повторного свертывания коллагена в условиях in vitro состоит в несовпадении взаимного расположения трех цепей, которые должны при взаимодействии образовывать триплетную спиральную молекулу (рис. 5.6). В случае in vivo проблема снимается обработкой растворимого предшественника коллагена перед его выделением. Глобулярные образования на обоих концах каждой полипептидной цепи обеспечивают правильное расположение вытянутых участков трех цепей коллагена относительно друг друга (рис. 4.4).

Другим примером является инсулин, который не удается ренатурировать, если его нативные дисульфидные связи были разрушены тиолами или если их структура менялась при ферментативных воздействиях [101]. Этот факт стимулировал поиски предшественника, который был действительно обнаружен в форме проинсулина [442]. Проинсулин стабилен к действию фермента дисульфидизомеразы (рис. 4.3); в опытах по денатурации — ренатурации он самопроизвольно повторно свертывается [443]. Протеолитическое расщепление проинсулина in vivo приводит к инсулину, стабильность которого обеспечивается энтропийным вкладом его нативной системы связей S—S (разд. 8.3). Лабильность структуры инсулина имеет, по-видимому, физиологическое значение [444], поскольку скорость инактивации является фактором, контролирующим степень и продолжительность действия гормона.

Повторное свертывание модифицированных белков дает информацию о процессе свертывания. Исследования повторного свертывания нативных белков были дополнены опытами по повторному свертыванию модифицированных белков. В ранних исследованиях [445] было показано, что рибонуклеаза поджелудочной железы, иодинированная по расположенному на поверхности нативной структуры Tyr-115, после денатурации теряет способность к повторному свертыванию. Это показывает, что состояние остатка Tyr-115 имеет важное значение для процесса свертывания. Для того чтобы установить, можно ли модифицировать белок (путем расщепления цепи и делеций в последовательности или путем присоединения объемных групп к боковым цепям) без потери им способности к свертыванию, был проведен ряд систематических исследований нуклеазы стафилококка и панкреатической рибонуклеазы.

Кроме того, об участии фрагментов полипептидов в процессе

свертывания можно судить, используя методы комплементации белка. «Комплементацией» называют восстановление биологической активности с помощью невалентных взаимодействий различных (поли)пептидных цепей [446]. Классический пример комплементации in vitro представляет панкреатическая рибонуклеаза. После того как так называемый S-пептид (включающий 20 N-концевых аминокислотных остатков) отщепляется от белка, последний теряет свою ферментативную активность и способность к повторному свертыванию восстановленной цепи. В результате комплементации белка S-пептидом способность к ферментативному действию и свертыванию восстанавливается, хотя ковалентные связи между двумя пептидными цепями и не образуются [94]. Таким образом, эксперименты по ренатурации (см. обзоры [94, 177]) модифицированных и комплементированных форм белка дают существенную информацию о механизме его свертывания.

8.3 Структурные элементы в несвернутых цепях

До сих пор по отношению к несвернутой цепи употреблялся термин «термодинамическое состояние». Однако в отличие от нативной структуры несвернутому состоянию соответствует множество конформаций цепи. Тем не менее можно предположить, что некоторый порядок все-таки существует и у несвернутой цепи. Выделение у нее структурированных элементов и тем самым сужение конформационного пространства полезно для рассмотрения процесса свертывания; в этом случае можно выделить зародышевые формы — нуклеации, с которыми взаимодействуют остальные участки цепи.

Эксперименты по повторному свертыванию выявляют быстро и медленно свертывающиеся цепи. Подразделение несвернутых цепей на два класса, быстро и медленно свертывающихся, было установлено при исследовании рибонуклеазы [440, 447, 448]. В опытах по повторному свертыванию рибонуклеазы, денатурированной без нарушения системы дисульфидных связей, было обнаружено, что 20% цепей повторно свертываются в пределах 0,1 с (быстро свертывающиеся цепи.) Процентное содержание таких цепей не зависит от способа денатурации (гуанидин гидрохлорид [449], нагревание [450] и т. д.). Остальные 80% цепей (медленно свертывающихся) превращаются за время от 10 до 100 с в быстро свертывающиеся, которые затем укладываются моментально. Имеется сообщение об аналогичном, хотя и менее количественном наблюдении в отношении ингибитора трипсина поджелудочной железы быка (BPTI), в котором 15% цепей свертывались значительно медленнее, чем остальные [451]. Это явление пока еще не объяснено. Согласно одной из гипотез [449, 452], в быстро свертывающихся цепях все пептидные связи представлены правильными *цис-транс*-изомерами, тогда как медленно свертывающиеся цепи содержат неправильные изомеры;

время расходуется на конверсию неправильных изомеров. *Цис-транс*-превращение происходит наиболее легко с N-концевой стороны остатка Pro, поскольку в этом случае барьер энергии активации составляет всего около 13 ккал/моль по сравнению с 20 ккал/моль для других пептидных связей (разд. 2.5). И BPTI, и рибонуклеаза содержат четыре остатка Pro, в нативной рибонуклеазе два из них существуют в виде *цис*-изомеров.

Антитела можно использовать для идентификации зародышей свертывания. Другой путь поиска следов структуры в несвернутых полипептидных цепях состоит в расщеплении нативных полипептидных цепей и в проверке с помощью специфичных антител, который из фрагментов принимает свою нативную конформацию. Большинство из этих экспериментов было выполнено на нуклеазе стафилококка [418; 453]; они показали, что фрагменты величиной от 17 до 120 остатков свертываются в нативную конформацию с константой равновесия $K = [N]/[R] \approx 10^{-3}$. Согласно уравнению (3.2), это отвечает невыгодному изменению свободной энергии приблизительно на 4 ккал/моль. В принципе этот метод можно использовать для поиска зародышей свертывания; для этой цели необходимо выделить или синтезировать необходимые фрагменты полипептидной цепи [454]. Стабильные фрагменты, которым соответствуют высокие значения K, наиболее вероятно представляют нуклеации свертывания.

Такие нуклеации должны характеризоваться высокой степенью корреляции с соседними по цепи остатками (раздел 5.3), поскольку именно взаимодействия с ними повышают стабильность промежуточного продукта. Это явствует из формулы Флори для оценки вклада дисульфидных мостиков в энтропию цепи [455]:

$$\Delta S_{\text{цепь}}^{(SS)} = 0{,}75 \cdot \nu \cdot R \cdot \ln(n' + 3) \qquad (8.2)$$

где ν — число соединенных мостиками цепей (в два раза больше числа мостиков), n' — число статистических сегментов между мостиками, за которое в первом приближении [456] можно принять число остатков между мостиками.

Чем короче цепи между мостиками, тем меньше понижение энтропии при их образовании. Это понижение энтропии вызвано не химическими свойствами, а только уменьшением числа возможных конформаций при образовании дисульфидных связей. Формулу Флори в связи с этим можно использовать для оценки понижения энтропии, необходимого для сближения двух элементов цепи.

Формула Флори количественно показывает, что сильная корреляция с соседними по цепи остатками выгодна. Как следует из уравнения (8.2), необходимое изменение энтропии тем меньше, чем короче цепь между сближаемыми элементами, т. е. чем сильнее корреляция с соседними остатками в соответствующем сегменте цепи.

И наоборот, энтропия активации $\Delta S^{\neq}_{\text{цепь}}$ при переходе от несвернутой цепи к нативному состоянию с сильной корреляцией соседних остатков намного меньше, чем $\Delta S^{\neq}_{\text{цепь}}$ перехода к нативному состоянию со слабой корреляцией. В соответствии с уравнением (3.2) меньшее значение $\Delta S^{\neq}_{\text{цепь}}$ отвечает более низкому барьеру свободной энергии, а следовательно, и более стабильной структуре промежуточного состояния. Таким образом, для свертывания цепи в структуру с сильной корреляцией по соседним остаткам требуется намного меньший вклад связывающей энергии и энтропии растворителя, чем для перехода к структуре со слабой корреляцией. Сильная корреляция по соседним остаткам облегчает процесс свертывания и тем самым заметно увеличивает его скорость.

Вторичные структуры, например α-спирали и β-слои, являются потенциальными зародышами свертывания, поскольку им свойственна большая связывающая энергия и сильная корреляция с соседними по цепи остатками. В связи с этим нужно отметить, что при повторном свертывании BPTI первый стабильный дисульфидный мостик [451, 457] связывает Cys-30 β-складчатого листа с Cys-51 α-спирали. На этом участке наблюдается самая сильная корреляция по соседним остаткам (кратчайший сегмент между образующими мостик остатками) в нативной структуре (рис. 8.1).

8.4 Пути свертывания

Аналогично тому как шар скатывается с горы вниз до самой глубокой точки долины, так и полипептидная цепь свертывается в конформацию с наименьшей свободной энергией. В обоих случаях и шар, и полипептидная цепь следуют определенному пути и доходят до самой нижней, кинетически допустимой точки (локальный минимум), которая не обязательно отвечает абсолютно низшей точке (глобальный минимум). Вопрос заключается в том, проделывает ли шар (цепь) всегда один и тот же путь, после того как оставляет вершину горы (которую можно рассматривать как широкое плато, отвечающее большому числу конформаций несвернутой цепи), т. е. существует ли расщелина, направляющая шар на его пути вниз.

Дисульфидные связи — специфические индикаторы свертывания. Поиск такого пути представляет весьма трудную задачу, поскольку для этой цели необходимо охарактеризовать конформации цепи в растворе. Все спектроскопические методы дают при описании свертывания цепи лишь самые общие результаты (круговой дихроизм, например, позволяет оценить количество α-спиралей, усредненное по всем имеющимся конформациям). Наиболее точную, хотя и ограниченную, информацию о структуре можно получить, исследуя образование связей S—S, поскольку эти связи появляются только при соответствующей ориентации двух остатков Cys.

В настоящее время наиболее подходящим объектом для таких экспериментов является ингибитор трипсина поджелудочной железы быка (BPTI) [793]. BPTI — это небольшой белок известной структуры, содержащий 58 аминокислот в одной цепи и три дисульфидных мостика. Для изучения образования связи S—S и ее перегруппировки в процессе свертывания существенно, чтобы реакции с уча-

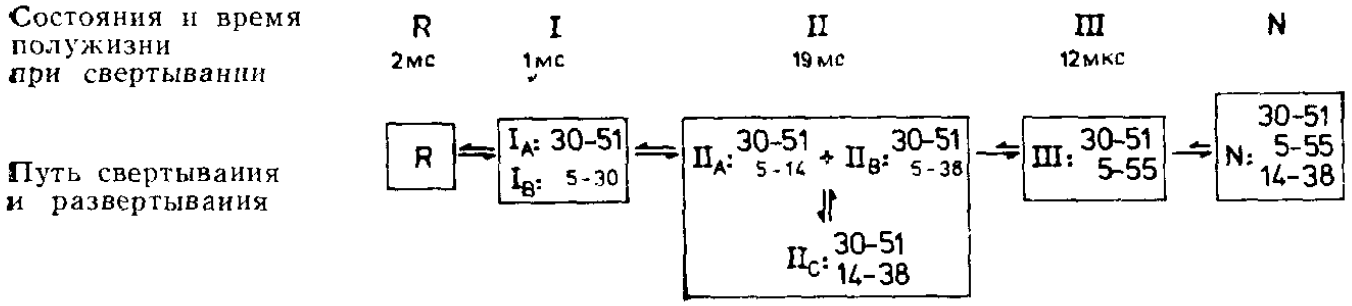


Рис. 8.1.

Ковалентные промежуточные состояния на пути свертывания бычьего панкреатического трипсинового ингибитора (BPTI) [451, 793].

Обозначения пояснены в тексте. «Правильные» дисульфидные связи показаны в крайней рамке справа. В промежуточных состояниях, например II_A и II_B, «правильные» дисульфидные связи изображены тем же шрифтом, что и в N-состоянии. Приведенный на рисунке путь свертывания относится к быстро свертывающимся цепям, т. е. к 85% несвернутых молекул.

стием остатков Cys тщательно контролнровались. Для этого используются низкомолекулярные тиолдисульфидные системы с известным окислительно-восстановительным потенциалом [451], напрнмер восстановленная и окнсленная формы дитиотрейтола высокой чистоты. Очищенные реагенты, например иодоацетат, используют и для связывания свободных тиолов. Прн этом нет надобности в денатурирующих агентах, поскольку полностью восстановленный BPTI при физиологических условнях существует в развернутой форме.

Добавление дисульфнда (например, окисленного дитиотрейтола) к раствору восстановленного BPTI означает начало свертывания. В этом процессе отчетлнво выделяют три промежуточные стадии (I, II и III [451]), которые показаны на рис. 8.1. Переходы от одной стадии к следующей характеризуются существенно различными временными постоянными, что указывает на различня активационных барьеров. На стадии I имеются два преобладающих вещества I_A и I_B соответственно с правильным и неправильным дисульфидными мостиками. Поскольку образование связи S—S и ее разрыв происходят в течение микросекунд, относительные частоты этих однодисульфидных молекул отражают их относительные термодинамические стабильности. Конформацня I_B далее не развивается; только конформация I_A переходит на следующую стадию, на которой образуются три преобладающие формы с двумя мостиками в каждой: II_A, II_B и II_C. Все они сохраняют первый правильный мостик, в II_A и II_B добавляется неправильный мостик, а в II_C — правиль-

ный. Интересно, что форма II_C не получает развития, хотя она и содержит два правильных мостика. Только формы II_A и II_B переходят на стадию III, на которой имеется одна форма, содержащая первоначальный правильный мостик и еще один дополнительный правильный мостик. Это превращение происходит очень медленно, поскольку для него необходимо конформационное изменение с высокой энергией активации. Переход от стадии III к нативному белку, напротив, происходит очень быстро.

Для свертывания необходимы промежуточные состояния, не входящие в нативную конформацию. Наиболее неожиданный результат заключается в том, что свертывание не есть последовательное все большее и большее приближение к нативной структуре, но что для него необходимы состояния II_A и II_B, которые не входят в окончательную нативную структуру. Этот аспект еще более отчетливо проявляется в случае рибонуклеазы, повторное свертывание которой также было прослежено аналогичным образом [458]. Как и BPTI, этот белок можно ренатурировать только путем образования дисульфидных мостиков. Но в отличие от BPTI на ранней стадии свертывания рибонуклеазы требуются по крайней мере три мостика, причем они могут быть случайными, что опять подтверждает необходимость ненативных интермедиатов. Отсюда следует вывод, что укладка белка начинается с образования нуклеаций на локальных участках цепи и что, прежде чем продолжится дальнейшее свертывание [459], необходимый вклад в свободную энергию ($-\Delta S_{\text{цепь}}$) должны внести дисульфидные мостики. Процесс свертывания in vitro происходит очень медленно и продолжается около 20 мин.

В случае лизоцима первые две нативные дисульфидные связи образуются на порядок быстрее, чем две последующие, что указывает на наличие предпочтительного пути свертывания [460]. Однако этот путь не обязательно единственный, что и было показано другими исследованиями лизоцима [162]. В этих экспериментах удалось ренатурировать восемь возможных изомеров (восстановленного лизоцима), содержащих по одному постоянно блокированному остатку Cys на молекулу белка. Все изомеры представляли собой ферментативно активные структуры. Этот факт исключает уникальную роль в процессе свертывания какой-либо из возможных дисульфидных связей. Следует особенно подчеркнуть, что ни одна из четырех нативных дисульфидных связей не обязательна для образования трех других правильных связей S—S.

Свертывание кооперативно. Как и переходы спираль — клубок (см. приложение), свертывание и развертывание белковой цепи — кооперативный процесс. Конкретизируем это общее положение для BPTI. Все промежуточные формы в этом случае нестабильны; положение конформационного равновесия смещается от развернутого восстановленного белка до нативного состояния с тремя связями S—S. Как видно из рис. 8.2, молекулярная заселенность существен-

на только для этих двух состояний и BPTI можно в хорошем приближении представить системой из двух состояний.

Свертывание может существенно отличаться у белков с дисульфидными связями и без них. Свертывание цистинсодержащих белков происходит на один-два порядка медленнее, чем свертывание белков той же величины без дисульфидных связей (разд. 8.2). Поскольку само образование дисульфидной связи не является стадией, определяющей скорость процесса, то эти белки по характеру свертывания, по-видимому, принадлежат к различным классам. По этой причине исследование механизма укладки цепи дисульфидсодержащего белка может иметь ограниченное значение для свертывания белков без таких мостиков.

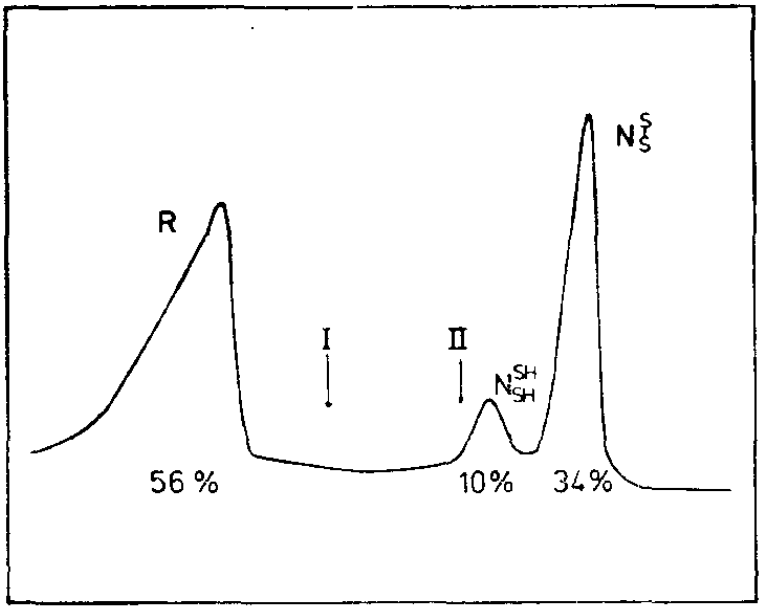


Рис. 8.2.

Модель свертывания белка с двумя состояниями [451].

Бычий панкреатический трипсиновый ингибитор (BPTI) был выдержан в растворе, содержащем 1 мМ восстановленного и 30 мМ окисленного дитиотрейтола. Образующаяся при этих условиях полностью восстановленная форма BPTI развернута. После установления равновесия перехода свертывание — развертывание белок извлекали с помощью иодоацетата, разделяли гель-электрофорезом и количественно охарактеризовывали. Согласно электрофоретическим данным в равновесии находятся только несвернутое состояние R и нативные состояния N^{SH}_{SH} (=III; рис. 8.1) и N^{S}_{S} (=N; рис. 8.1). Интермедиаты I и II процесса свертывания отсутствуют.

Медленное и затрудненное развертывание и повторное свертывание белков, содержащих связь S—S, можно объяснить с точки зрения физиологии. Такие белки обычно являются внеклеточными, где они должны быть стабильными при различных меняющихся воздействиях окружающей среды. Кинетическая стабильность позволяет им сохранять квазинативное состояние в условиях денатурации до тех пор, пока временные константы воздействия среды остаются меньше, чем константы процессов развертывания. Время «полуразвертывания» BPTI (стадия III → II, рис. 8.1) составляет несколько часов [451].

В некоторых случаях медленное свертывание и медленное раз-

вертывание цепи могут определять продолжительность жизни белка. После завершения свертывания в клетке такой белок должен выделяться во внеклеточное пространство. Как только он развертывается, он становится подверженным действию протеаз, поскольку повторное свертывание происходит через образование подвижных промежуточных форм с относительно большими временами «полужизни».

Обычно в дисульфидсодержащих белках часть связей S — S необходима для термодинамической стабильности. Рассмотрим еще раз вопрос о том, насколько необходимы дисульфидные связи для стабильности нативной структуры. Набор поперечных сшивок, уникальный и характеристичный для упорядоченного состояния цепи, стабилизирует это состояние в связи с понижением энтропии несвернутой формы цепи с тем же набором поперечных связей. Для белка из 120 остатков вклад четырех связей S—S в стабильность нативной конформации составляет приблизительно $4 \times 2 = 8$ ккал/моль (уравнение (8.2)). Из данных о том, что большинство внеэпителиальных белков имеет величину $\Delta G_{\text{общ}}$ около 10 ккал [413] и что они денатурируются, если все дисульфиды восстановлены (табл. 4.1), следует, что для термодинамической стабильности этих белков необходим энергетический вклад одной или большего числа дисульфидных связей; конкретное число таких связей варьируется в зависимости от белка (табл. 4.1).

8.5 Влияние лигандов

Как правило, лиганды стабилизируют нативную структуру. Для некоторых белков образование нативной конформации зависит от присутствия специфических лигандов, часто ионов металлов или других простетических групп [94, 413]. Наличие лиганда не обязательно приводит к существенно иной конформации свернутого белка, но способствует стабилизации его нативной формы. На основании опытов по водородному обмену [415], а также с помощью кривых денатурации мочевиной [461] апомиоглобина, глобулярного белка с $\Delta G_{\text{общ}} \approx 9$ ккал/моль, установлено, что этот белок должен быть на 4 ккал/моль менее стабилен, чем миоглобин. Противоположным примером является цитохром *c*, присутствие в котором ковалентно связанной группы гема абсолютно необходимо для стабильности белковой структуры. При удалении этой группы белок перестает быть глобулярным [462]. Если удалить только ион железа, белок остается глобулярным, хотя и менее стабильным, например, к термической денатурации [463].

Более подробно влияние лигандов тимидин-3,5-дифосфата (TDP) и Ca^{2+} было изучено для нуклеазы T, протеолитического производного нуклеазы стафилококка [94]. Константу равновесия системы нативная конформация — развернутая конформация можно

выразить как $K = k_{+1}/k_{-1}$, где k_{+1} — константа скорости свертывания, а k_{-1} — константа скорости развертывания. Присутствие лигандов не влияет на скорость свертывания нуклеазы Т [461], однако скорость развертывания в присутствии Ca^{2+} и TDP уменьшается в 20 раз.

Для всех белков, содержащих функционально важные лиганды, пока еще не разрешены два существенных вопроса: на какой стадии процесса свертывания начинает принимать участие лиганд и каким образом полипептидная цепь в условиях in vivo отбирает свою специфическую простетическую группу, например данный металл при наличии многих альтернатив.

Свертывание пируваткиназы обеспечивается L-валином. Влияние аминокислоты L-валина на ренатурацию пируваткиназы [466] дрожжей и треониндезаминазы *E. coli* [468] иллюстрирует две различные функции, которые могут выполнять лиганды в процессе образования активных олигомерных белков. Пируваткиназа [467] — это тетрамерный фермент, построенный из четырех идентичных субъединиц, каждая из которых содержит одну молекулу невалентно связанного L-валина. Субъединицы диссоциируют и развертываются при действии 6 М гидрохлорида гуанидина. При выдерживании в ренатурирующей среде L-валин служит специфичным инициатором процесса повторного свертывания. Он индуцирует ренатурацию с константой скорости псевдопервого порядка по отношению к мономеру, а это означает, что L-валин влияет на свертывание мономерной формы в ее нативную конформацию и что завершающим процессом является спонтанное образование тетрамерного фермента. L-Валин остается составной частью всей структуры молекулы нативного белка [466].

Лиганды нндуцируют формирование ферментов. В отличие от рассмотренного выше случая для треониндеаминазы, димерного фермента с молекулярной массой 360 000, содержащего две пиридоксальфосфатные группы (рис. 8.3), L-валин (или L-изолейцин) служит только в качестве катализатора и не присоединяется к молекуле фермента [468]. Неактивный олигомер при этом проходит дополнительное формирование, в результате которого получается ферментативно активный препарат. При pH 7,5 и выше для этого процесса, являющегося непременным условием активации [469], L-валин совершенно необходим.

Существуют и другие примеры олигомерных ферментов, образование которых обязательно проходит через каталитически неактивные промежуточные формы. Эти формы претерпевают конформационное превращение на более поздней стадии, давая термодинамически стабильный активный фермент. Такой процесс конформационного «дозревания» имеет лимитирующую скорость и представляет собой практически необратимую стадию в образовании фермента. Предполагают [472], что в некоторых олигомерных ферментах алло-

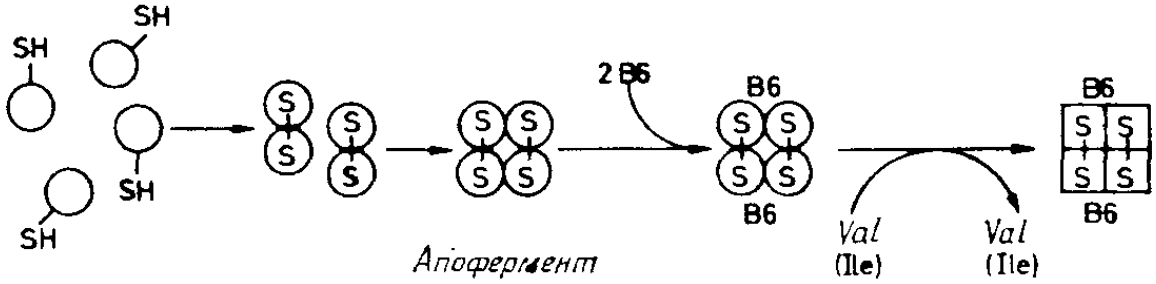


Рис. 8.3.

Образование активной треониндезаминазы из свернутых мономеров [122]. Кругами обозначены четыре мономера, каждый из которых включает SH-группу, использующуюся для образования дисульфида с идентичной SH-группой другого мономера. Две пары объединенных субъединиц образуют апофермент; присоединение кофактора пиридоксальфосфата (витамин B_6) приводит к неактивному холобелку, который переходит при каталитическом воздействии Val (или Ile) в активный фермент [468]. Последняя стадия названа в [472] «прекондиционированием».

стерические лиганды могут специфически ускорять формирование. Подобное воздействие аллостерических лигандов отличается от их влияния на активность фермента, когда они стабилизируют или активную, или неактивную конформацию белка при условии легко обратимого равновесия.

8.6 Моделирование процесса свертывания

Чтобы определить трехмерную структуру по аминокислотной последовательности, необходимо понять механизм укладки цепи. Трехмерная структура белка определяется невалентными взаимодействиями между аминокислотными остатками цепи, а также между этими остатками и растворителем (гл. 3). В принципе, если учесть все эти взаимодействия, можно рассчитать нативную конформацию по известной ковалентной структуре. Однако поскольку нативная конформация может не отвечать глобальному энергетическому минимуму, то расчет энергии всех возможных конформаций цепи может не привести к правильному ответу. Наиболее существенно, однако, что для рассмотрения всех возможных конформаций цепи потребуется машинное время, намного превышающее возраст Земли. Очевидно поэтому такие расчеты можно осуществить лишь в том случае, если не делать попытки охватить все статические структуры, а попытаться смоделировать процесс свертывания, следуя природному пути свертывания данной цепи. Если этот путь однозначен (аналогия с глубокой расселиной и шаром; разд. 8.2), то расчеты умеренной точности смогут привести к правильному решению задачи. Но если путь определен недостаточно хорошо, требуется высокая точность расчетов.

Моделирование свертывания основано на иерархии уровней структуры белка. Было предпринято несколько попыток промоделировать процесс свертывания, например, апомиоглобина [470], BPTI [30, 404] и парвальбумина [471]. Все они основаны на предположении о строгой иерархии уровней белковой структуры (разд. 5.6)

при допущении, что α-спирали нативной структуры уже сформированы в еще не полностью свернутой цепи. Такое предположение отчасти подтверждается данными иммунологических исследований, по которым можно судить о содержании α-спиралей в несвернутой цепи [418, 454], а также достаточно удовлетворительными результатами методов предсказания α-спирали (разд. 6.5). Моделирование свертывания цепи выполняется a posteriori для белков, структуры которых определены с помощью рентгеноструктурного анализа. Применение такого моделирования к белку с неизвестной трехмерной структурой потребует информации об α-спиральных участках цепи; отчасти эта информация может быть получена с помощью методов предсказания (гл. 6).

Свертывание апомиоглобина рассматривалось как сборка α-спиралей. Моделирование свертывания апомиоглобина [470] было выполнено без применения ЭВМ, поскольку использовалась предельно простая модель, в которой были заранее заданы девять спиралей, наблюдающихся в нативной структуре (рис. 8.4, *а*). Это существенно сузило рассматриваемое конформационное пространство. На начальной стадии процесса свертывания были определены четыре структурные нуклеации, каждая из которых содержала три соседние по цепи α-спирали. В каждой нуклеации спирали располагались различным образом. Свободная энергия этих спиральных образований рассчитывалась исключительно по гидрофобным взаи-

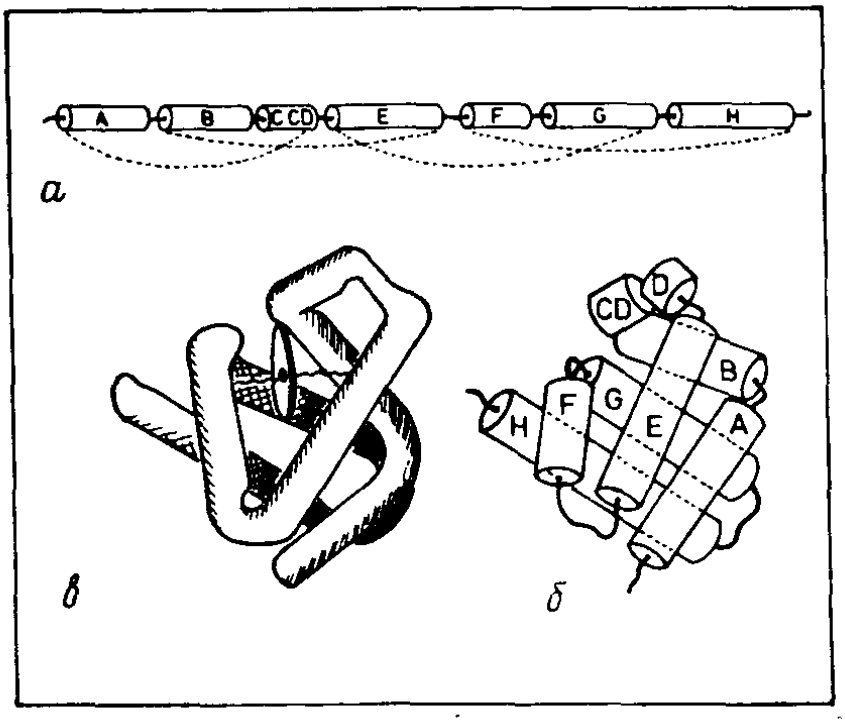


Рис. 8.4.

Моделирование свертывания апомиоглобина [470].

а — α-спирали, наперед заданные в начале моделирования процесса свертывания. На начальной стадии учитывалнсь четыре отмеченные на рисунке нуклеацни, каждая из которых включает три α-спирали. *б* — результат моделирования свертывания цепи, отвечающий минимальному рассчитанному значению $\Delta G_{\text{общ}}$. *в* — модель молекулы миоглобина по данным рентгеноструктурного анализа.

модействиям (разд. 3.5) крупных неполярных боковых цепей; предполагалось, что ни одна из заряженных боковых цепей не находится внутри белка. Из большого числа полученных спиральных конструкций последующему анализу подвергнуты только 40 наиболее предпочтительных по энергии структур, попадающих в интервал 4 ккал/моль, считая от уровня, соответствовавшего форме с самой низкой свободной энергией. На последующих стадиях моделирования к этим нуклеациям присоединяли другие спирали и комбинировали структуры до тех пор, пока вся цепь не оказалась свернутой. Начиная с некоторой стадии, учитывали только такие конформации, свободная энергия которых превышала минимальную не более чем на 4 ккал/моль. Эта процедура привела к 19 различным способам свертывания цепи. Как и ожидалось, расположение спиралей с минимальной свободной энергией (рис. 8.4, *б*) отвечало нативной структуре миоглобина (рис. 8.4, *в*). Однако различия вычисленных энергий для этой формы и структуры со следующей по величине свободной энергией составили всего 3%, что намного меньше ожидаемой ошибки.

Рассматривая этот метод, необходимо сделать следующие замечания: а) метод применим только к белкам класса I (табл. 5.2), т. е. к белкам, состоящим преимущественно из α-спиралей; б) на начальной стадии расчета допускается существование крупных нуклеаций, составляющих одну треть молекулы, хотя экспериментальные данные о таких нуклеациях пока отсутствуют; в) предполагалось, что в такие нуклеации входят соседние по цепи спирали, что отвечает низкой энтропии цепи и наличию сильной корреляции по соседним остаткам (разд. 8.3); г) считалось, что нативная структура свертывается прямым путем: все промежуточные структуры отбрасывались как неспособные к перегруппировке. Последнее предположение находится в противоречии с опытами по белкам, содержащим дисульфидные связи, которые выявили существование промежуточных форм с геометрией, существенно отличающейся от нативной (рис. 8.1). Однако процесс свертывания апомиоглобина происходит намного быстрее, и поэтому можно предположить его большую направленность по сравнению со свертыванием дисульфидсодержащих белков (разд. 8.2).

В BPTI и парвальбумине свертывание упрощенной цепи производилось путем спуска по градиенту энергии. Несколько иной метод был использован для моделирования свертывания цепи BPTI [30, 404] и парвальбумина [471]. В этом случае полипептидная цепь рассматривалась более детально, чем при моделировании апомиоглобина, однако аппроксимация оставалась грубой, поскольку в качестве независимой переменной использовался только угол вращения вокруг виртуальной связи C_α—C_α (рис. 7.10). Виртуальный валентный угол при C_α-атоме считался функцией торсионного C_α-угла; боковые цепи были заданы жесткими сферами (рис. 7.10).

Это упрощение снизило число независимых параметров до 1 на остаток.

Моделирование процесса свертывания было начато с цепи, имеющей вытянутую квазибеспорядочную конформацию; исключение было сделано для нативных α-спиралей, которые считались уже сформированными. Затем выполнялся спуск по градиенту свободной энергии. Кинетическая энергия цепи в этом процессе не принималась во внимание. Для учета температурных флуктуаций (разд. 8.1) необходимо «встряхивать» цепь через короткие интервалы времени. Поскольку это заняло бы слишком много машинного времени, «возбуждение» цепи проводили, только если она попадала в минимум энергии, с тем чтобы вывести ее из возможного локального минимума. В отличие от расчетов апомиоглобина в этом методе рассматриваются как гидрофобные силы, так и связывающая энергия. Кроме того, не производилось фиксации каких-либо субансамблей и допускались непрерывные перегруппировки во всех частях цепи. В результате было показано, что с вероятностью 50% цепь BPTI свертывается в нативную конформацию со среднеквадратичным отклонением C_α-атома около 6 Å. Однако даже самые лучшие результаты привели к неправильной топологии β-структуры, за что и были подвергнуты критике [801]. Следует отметить, что это моделирование не учитывало специфического образования дисульфидных связей, происходящего в процессе свертывания BPTI (рис. 8.1). Использование этой информации может оказаться полезным в дальнейших исследованиях. Однако экспериментальные данные по ренатурации дисульфидсодержащих белков показывают, что путь свертывания таких белков может быть значительно сложнее свертывания других белков (см. с. 189), в связи с чем дисульфидсодержащие белки являются, по-видимому, не самыми удачными объектами для развития теоретического анализа.

В случае парвальбумина [471], α-спирального белка класса I (табл. 5.2), моделирование свертывания оказалось менее успешным. Отчасти это можно объяснить тем, что не была учтена возможность связывания в двух местах иона Ca^{2+}. Поскольку эти центры расположены в коротких петлях между спиралями [59], можно предполагать, что они влияют на относительную ориентацию соседних спиралей, а значит, и на процесс свертывания.

Заключение

В настоящее время многие принципы и детали процесса свертывания полипептидной цепи остаются непонятыми. Мы знаем, что этот процесс происходит самопроизвольно и что, следовательно, он основан на простых физико-химических принципах. Это позволяет надеяться, что свертывание удастся объяснить и промодели-

ровать с помощью теоретических подходов. Сейчас уже сделаны первые попытки в этом направлении. Интерес к процессу свертывания объясняется не простым научным любопытством, а имеет значительно более глубокие причины. Чтобы проникнуть в сущность биологических процессов, мы должны знать трехмерную структуру различных белков, в том числе и тех, которые невозможно закристаллизовать, а значит, сделать доступными для рентгеноструктурного анализа. Единственная надежда на раскрытие таких структур связана с возможностью их определения по аминокислотным последовательностям. Такое установление должно следовать по пути процесса естественного свертывания: отбор всех возможных конечных структур данной полипептидной цепи по их свободной энергии столь же безнадежен, как и попытка ответить на вопрос, отвечает ли нативная структура локальному или глобальному минимумам энергии.

Наиболее сложной проблемой динамики свертывания является описание промежуточных состояний. Единственные сведения о структуре этих состояний пока могут быть получены только по образующимся в них дисульфидным связям. К сожалению, вполне возможно, что процесс свертывания белков, содержащих связь S—S, качественно отличается от свертывания других белков, и таким образом, мы получаем здесь только ограниченную информацию.

Большие усилия направляются на поиск нуклеаций — зародышей процесса свертывания, которые, как предполагают, служат основой дальнейшего свертывания. Такие зародышевые образования должны иметь исключительно стабильную структуру. Возможно, что их можно будет обнаружить с помощью антител или экспериментов с повторным свертыванием модифицированных белков. Локализация зародышей должна в большой мере облегчить моделирование свертывания. В любом случае очевидно, что понадобится еще очень большое число экспериментов, прежде чем загадка процесса свертывания будет разрешена.

Глава 9 Эволюция белков

Биология — естественная наука, имеющая исторические корни. Все существующие биологические виды постепенно развились из одного или из очень ограниченного числа предшественников. Поскольку в событиях прошлого важную роль мог играть случай, мы должны смириться с тем обстоятельством, что направления эволюции, по-видимому, никогда не удастся полностью объяснить в терминах «причины» и «следствия». Исследователям, склонным к построению стройных систем на основе развития неких общих принципов, это представляется серьезнейшим препятствием. Однако квазислучайное развитие привело к огромному разнообразию биологических видов, которое можно рассматривать как результат многочисленных экспериментов природы. Мы должны быть признательны за эти эксперименты, и должны анализировать и интерпретировать их результаты.

В процессе эволюции белков происходят замены отдельных остатков, вставки и делеции нескольких остатков, удвоение и слияние генов. Для белков основные этапы исторического* процесса знаменуются заменами аминокислотных остатков в полипептидной цепи. С течением времени эти замены накапливаются, так что в конечном счете какое-либо сходство между исходной** и окончательной аминокислотными последовательностями может исчезнуть. Однако, как правило, даже после того как исчезнет сходство аминокислотных последовательностей двух гомологичных белков, сохраняется соответствие в укладке их цепей. Тенденции к замещениям заметно отличаются у остатков в разных положениях в цепи (рис. 7.1, *б*). Различия гомологичных белков не исчерпываются за-

---

* Применяя термин «история» мы обычно имеем в виду каким-то образом зарегистрированные события. В этом смысле использование этих понятий для белков означает, что мы рассматриваем нуклеотидную и аминокислотную последовательности как запись событий прошлого.

** Следует учитывать, что нам известна структура только существующих в настоящее время белков. Поэтому мы можем судить об эволюционных изменениях в прошлом лишь по имеющимся различиям между гомологичными белками.

мещениями отдельных аминокислот. Более значительными событиями являются вставки и утраты (пропуски, делеции) отдельных остатков или целых групп остатков (рис. 7.1, *а*). Наличие таких вставок и делеций значительно затрудняет сопоставление последовательностей отдаленно родственных белков [803].

Существенную помощь в подобных случаях оказывает знание трехмерной структуры белков. Имеющиеся в настоящее время данные показывают, что обычно остатки, находящиеся во внутренней части белка, мало подвержены изменениям и что все различия между гомологичными белками (замены аминокислот, делеции или вставки петель в цепи) касаются поверхности молекул. Таким образом, последовательности отдаленно родственных белков можно сопоставлять по остаткам, которые занимают геометрически сходные позиции в пространственной структуре.

Очень серьезные изменения связаны с удвоением гена, которое может привести к удвоению длины полипептидной цепи. Более того, известны случаи слияния различных структурных генов, которые указывают на перемещение одного или большего числа генов в другое положение в геноме.

Гомологичные белки возникают вследствие специализации или дифференциации. Сопоставление гомологичных белков позволяет выявить некоторые общие закономерности в строении белков [250], а, как уже отмечалось, исследование эволюции белков способствует решению многих общих биологических проблем. В связи с этим полезно установить различие в понятиях специализации и дифференциации белков [473]. Под специализацией подразумевается эволюция гомологичных белков, выполняющих одинаковую функцию в различных организмах. Напротив, дифференциация белков есть процесс, ведущий к функциональному разнообразию гомологичных белков часто внутри одного организма.

Исследование специализации белков позволяет устанавливать генеалогию организмов. Как правило, аминокислотная последовательность и характер свертывания цепи в процессе эволюции сохраняются настолько хорошо, что белки даже очень генетически далеких организмов имеют определенное сходство. Таким образом, путем исследования специализации белков можно устанавливать генеалогию, т. е. происхождение и родственные связи организмов. Поэтому филогенетический анализ, основанный на специализации белков, можно рассматривать как один из методов систематики.

Дифференциация белков отражает эволюцию биохимических путей. Исследования дифференциации белков преследуют иные цели. В этом случае структурное подобие обычно позволяет выявить неожиданные биологические связи, что можно затем использовать, например, при прослеживании эволюции метаболизма. Кроме того, дифференциация убедительно показывает, что биологическое разнообразие ограниченно и что в этом плане белки можно подраз-

делить на определенные группы [474]. Типичным примером класса образовавшихся в результате дифференциации белков является семейство сериновых протеаз (табл. 9.4), которые осуществляют контроль над самыми разнообразными жизненно важными процессами в организме человека.

Отсюда очевидно, какое огромное значение для (пато)физиологических проблем и для фармакологии будет иметь анализ структуры белков на молекулярном уровне [602].

9.1 Специализация белков

Частота допустимых мутаций

Биологическая специализация устанавливается в процессе случайных мутаций и последующего отбора. Наиболее важны мутации, возникающие в ДНК. В табл. 9.1 даны ориентировочные частоты мутаций ДНК высших организмов. В этих организмах частоты допустимых мутаций на уровне ДНК существенно не влияют на отбор белков, поскольку структурные зоны [78, 475], кодирующие белок, составляют менее 10% ДНК (разд. 4.1). Тем не менее, как было установлено, ДНК гомогенна в пределах одного вида [476, 477], что указывает на существование некоторого отбора и на уровне ДНК. Вопрос о том, определяется ли эта гомогенность малой популяцией предшественников и малым «генетическим дрейфом» (недарвиновская гипотеза, поддерживаемая «нейтралистами» [144, 478—480])

Таблица 9.1

Скорость эволюции макромолекул[a]

Макромолекула	Скорость эволюции, выраженная в замещениях на кодон за 10^{10} лет (т. е. в PAM/10^8 лет)
Гипервариабельные нуклеотидные последовательности ДНК [144, 483, 586]	500
Средняя ДНК млекопитающих [475, 476]	45—85
мРНК цепей гемоглобина [484, 485]	100
мРНК гистона IV в различных видах морского ежа [486]	35
Средние белки [145]	1—50
Гемоглобин, α- и β-цепи соответственно [145]	14
Гистон IV [145]	0,09
Цитохром *c* митохондриальный [794]	5
Гипервариабельные аминокислотные последовательности белков [488, 586]	85

[a] Все величины оценены очень приближенно и могут представлять нетипичные случаи. ДНК здесь обозначает неповторяющуюся ДНК.

или же она отражает отбор самой приспособленной ДНК данного вида (дарвиновская гипотеза, поддерживаемая «селекционерами»), по-прежнему остается предметом дискуссии (ср. с работами [481, 482]).

Генетическое расстояние измеряется в единицах РАМ. На уровне макромолекул частоту фиксации мутаций (называемую также частотой усвоения) обычно выражают числом усвоенных точечных мутаций на кодон за 10^{10} лет, что численно равно процентному содержанию фиксированных точечных мутаций за 10^8 лет. Процентное содержание фиксированных точечных мутаций (единица РАМ) является общепринятой мерой эволюционного расстояния. Величиной, обратной частоте фиксации, является «единичный эволюционный период», т. е. число лет, за которое происходит одно фиксируемое изменение кодона на 100 остатков.

Отбор нуклеотидных последовательностей осуществляется на уровне ДНК. Из-за вырожденности генетического кода только 75% всех мутаций структурной зоны ДНК влияют на последовательность белка [144, 483]. Остальные 25% образуют так называемые невыявленные мутации, при которых разные кодоны эквивалентны одной и той же аминокислоте (рис. 1.5, *б*). Поскольку используются все кодоны данной аминокислоты (хотя и с различной частотой [12, 13, 145]), ограничений на уровне белка не существует и частоты невыявленных мутаций должны составлять приблизительно 25% максимальных частот мутаций ДНК. Частоты мутаций информационной РНК (мРНК), кодирующей цепи гемоглобина [484, 485], составляют приблизительно четверть частот мутаций гипервариабельной нуклеотидной последовательности ДНК (табл. 9.1). Таким образом, по-видимому, она соответствует максимальным частотам невыявленных мутаций. Частота невыявленных мутаций мРНК гистона IV в три раза меньше, чем мРНК гемоглобина (табл. 9.1). Одна клетка морского ежа содержит 1200 копий активных генов гистона IV [487]; соответствующие 1200 последовательностей тРНК, по-видимому, гомологичны в пределах данного вида [486], хотя и различаются между видами. Этот факт служит еще одним свидетельством того, что отбор нуклеотидных последовательностей происходит на уровне ДНК.

Отбор на уровне белка обычно значительно более затруднителен, чем на уровне ДНК. Это следует из сравнения частот замен аминокислот гемоглобина и гистона IV с частотами невыявленных мутаций в соответствующих тРНК (табл. 9.1). Для различных белков частота замен сильно варьируется. Если сравнивать крайние случаи, гистоны IV и гипервариабельные сегменты белков, то частоты замен для них отличаются на три порядка (табл. 9.1). Такие различия свидетельствуют о том, что сохранение структуры данного белка важно для его функции в организме. Гистон IV закодирован в центральной части ДНК, и любое изменение его аминокислот может привести к

катастрофическим последствиям. Эволюция гистона IV строго следует дарвиновскому пути; мутации практически не фиксируются. Напротив, гипервариабельная область рибонуклеазы, по-видимому, не имеет определенной функции [488], и изменения в ней фиксируются значительно легче. В этом случае можно предположить недарвиновский путь эволюции [489].

Критерии фиксации аминокислотных замен в белках

Подробная биография белка необходима для оценки биологического значения данного остатка. Как было показано на предыдущем примере, фиксация замен аминокислот почти полностью зависит от биологической роли соответствующей аминокислоты. Однако оценить эту роль довольно трудно, поскольку она может быть связана не только с определенной функцией белка, например каталитическим действием фермента, а со *всеми* другими взаимодействиями белка в организме на протяжении его «жизненного пути» от активации соответствующего гена до разложения полипептида. И хотя такие подробные «биографии» белков пока еще не доступны, можно тем не менее сделать некоторые общие замечания*.

Остатки на поверхности белка заменяются чаще, чем внутренние. Как показывает рис. 7.1, *б*, частота замен аминокислотных остатков в данном белке сильно зависит от положения в полипептидной цепи и от расположения в трехмерной структуре. Как правило, остатки на поверхности заменяются чаще, чем внутренние. Поскольку наружные остатки имеют менее существенное значение для стабильности белка, чем внутренние, это правило отражает важность сохранения стабильности для функции белка. Исключение составляют остатки на поверхности, непосредственно участвующие в функционировании белка, как, например, остатки, образующие каталитически активные центры ферментов, остатки в местах связывания субстратов, кофакторов, простетических групп, аллостерических эффекторов, или центры связывания макромолекул. Так, у цитохрома *с*, который взаимодействует с другими макромолекулами и присоединяет простетическую группу, на поверхности практически отсутствуют неактивные участки. Это объясняет низкую частоту фиксированной мутации в молекуле цитохрома *с* (табл. 9.1).

Влияние замен аминокислот во внутренней части белка часто компенсируется другими заменами. Поскольку глобулярные белки

---

* Эти обобщения применимы не только к тем белкам, которые действуют в обычных условиях, но также и к тем, которые подвергаются экстремальным воздействиям тепла [491] или гидростатического давления [492], как, например, белки термофильных бактерий и абиссальных (глубоководных) рыб соответственно. По-видимому, для приспособления к этим экстремальным условиям достаточно лишь небольших изменений аминокислотных последовательностей.

упакованы так же плотно, как кристаллы органических молекул (разд. 1.6), замена внутренних остатков влечет за собой изменение расположения соседних остатков, что обычно приводит к понижению стабильности. Поэтому замена во внутренней части белка происходит сравнительно редко и во многих случаях сопровождается по крайней мере еще одной заменой. Примеры таких взаимно компенсирующихся замен были обнаружены в рибонуклеазах [67]. Случай очень тонкой внутренней компенсации был отмечен в двух отдаленно родственных сериновых протеазах: внутренний кластер химотрипсина, построенный из Trp-29, Ser-45, Val-53, Val-200, Leu-209, Val-210, Ile-212, в эластазе превращается в Ser-29, Thr-45, Met-53, His-200, Val-209, His-210 и Val-212 без каких-либо серьезных изменений остова полипептидной цепи [490].

Замены подобных остатков происходят чаще, чем другие замены. Требование сохранения функции налагает ограничения на частоту допустимых замен в данном положении полипептидной цепи. По-видимому, многие функции менее всего нарушаются при консервативных замещениях, т. е. заменах на подобные остатки. При этом имеют значение величина, форма, гибкость и заряд боковой цепи, а также ее способность к образованию водородных связей. Так, при замене Lys → Arg сохраняется подвижная боковая цепь, несущая положительный заряд, а при замене Ile → Leu — относительно объемная неполярная боковая цепь. Принцип консервативных замен использован в разд. 1.6 для выявления эмпирического подобия аминокислот по наблюдаемым частотам встречаемости (табл. 1.2).

Аномальные гемоглобины иллюстрируют возможные последствия случайных мутаций. Однако даже консервативные замены могут привести к серьезным последствиям, как было обнаружено на примере аномального гемоглобина Сиднея [493], в котором в положении 67 β-цепи вместо Val содержится Ala. Замещение двух метильных групп атомами водорода расшатывает группу гема и значительно снижает стабильность белка [494] и тем самым стабильность эритроцита.

Неконсервативные замены, как, например, введение отрицательного заряда в гидрофобную внутреннюю часть, несовместимы со стабильностью нативной конформации белка, поскольку при этом потери свободной энергии оказываются порядка 10 ккал/моль, что приближается к значению $\Delta G_{\text{общ}}$ перехода свернутая — развернутая цепь (разд. 8.1). Примером может быть аномальный гемоглобин Вены [494], в котором замена Tyr-130 → Asp в α-цепи приводит к большим конформационным перестройкам, в результате которых отрицательный заряд компенсируется положительным.

Оба описанных аномальных гемоглобина вызывают состояния тяжелой анемии, демонстрируя тем самым последствия случайных мутаций. Очевидно, что такие мутации никогда не станут фиксированными.

Таблица 9.2

Матрица вероятностей мутаций на эволюционном пути в 2 PAM [20][a]

	Gly	Pro	Asp	Glu	Ala	Asn	Gln	Ser	Thr	Lys	Arg	His	Val	Ile	Met	Cys	Leu	Phe	Tyr	Trp	Сумма
Gly	9870	17	13	22	40	22	11	42	8	5	0	1	7	0	0	3	2	0	0	0	10063
Pro	7	9850	1	13	23	9	13	11	5	3	0	0	4	3	0	0	0	0	0	0	9942
Asp	8	1	9757	96	13	45	27	26	2	8	0	6	4	0	1	0	2	0	0	0	9996
Glu	13	17	95	9726	21	9	40	15	12	13	0	4	7	4	1	0	4	0	0	0	9981
Ala	42	54	24	37	9730	31	34	99	45	18	0	5	32	3	19	5	5	5	0	0	10188
Asn	10	10	36	7	14	9701	20	51	17	19	7	24	4	4	1	0	2	0	0	0	9927
Gln	4	11	16	24	12	15	9736	13	10	9	14	14	5	4	11	0	2	0	0	0	9900
Ser	26	15	28	16	59	67	22	9598	69	14	2	17	7	4	23	27	3	6	0	0	10003
Thr	6	8	3	14	30	25	20	76	9759	10	0	8	20	24	11	8	5	3	0	0	10030
Lys	5	6	13	21	17	37	23	22	14	9845	65	14	13	9	11	0	6	0	4	0	10125
Arg	0	0	0	0	0	5	13	1	0	23	9881	17	0	0	18	0	0	2	0	0	9960
His	0	0	4	3	2	20	15	10	5	6	19	9865	1	4	0	0	3	3	4	11	9975
Val	6	8	5	10	27	7	12	9	25	12	0	3	9783	156	82	18	22	3	0	0	10188
Ile	0	2	0	3	1	3	4	3	14	4	0	4	70	9703	22	3	22	14	0	0	9872
Met	0	0	0	0	2	0	4	5	2	2	7	0	12	7	9672	5	14	5	0	0	9737
Cys	1	0	0	0	1	0	0	12	3	0	0	0	6	2	11	9928	0	0	0	0	9964
Leu	2	0	3	7	4	3	4	5	7	6	0	6	24	52	99	0	9899	19	0	0	10140
Phe	0	0	0	0	2	0	0	5	2	0	3	4	2	18	18	0	10	9879	74	30	10047
Tyr	0	0	0	0	0	0	0	0	0	2	0	4	0	0	0	0	0	51	9909	17	9981
Trp	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	8	7	9941	9960

[a] Аминокислоты расположены в том же порядке, что и в табл. 1.2. Все значения умножены на 10 000. Элемент этой матрицы $m(i, j)$ есть вероятность того, что аминокислота в колонке i будет замещена аминокислотой в строке i за эволюционный интервал в 2 PAM, т. е. 2 допущенные точечные мутации на 100 аминокислот. Таким образом, вероятность того, что Ala будет замещен на Ser, равна 0,0059, а Ser на Ala — 0,0099. Сумма величин в каждой колонке равна 1,0. Сумма величин в строке представляет «фактор роста за 2 PAM» соответствующего аминокислотного остатка; он варьирует от 0,9737 (для Met) до 1,0188 (для Ala и Val).

Вероятность мутаций аминокислотных остатков

Вероятность фиксированных мутаций зависит от типа остатка. Вероятности мутаций для различных остатков приведены в табл. 9.2. Значения, находящиеся на диагонали, представляют вероятность того, что данный остаток будет сохраняться. Ser, который часто находится на поверхности белка, имеет наименьший диагональный член, что означает наибольшую вероятность мутации. Наименьшая вероятность мутации соответствует Trp. Это вполне объяснимо, поскольку Trp обычно является внутренним остатком и его трудно заместить остатком с боковой цепью равной величины.

Сильный консерватизм в этом отношении обнаруживают также некоторые остатки Gly. На первый взгляд это до некоторой степени неожиданно, поскольку Gly не несет боковой цепи, которая могла бы выполнять определенную функцию. Однако анализ трехмерной структуры белков показал, что для таких инвариантных положений Gly (как, например, в положении контакта между спиралями B и E в семействе глобинов [145, 277] или в положении 7 рис. 7.8 между двумя α-спиралями цитохромов *c*-типа [495]) именно отсутствие боковой цепи имеет решающее значение. Только в этом случае спирали могут упаковываться достаточно плотно; введение боковой цепи привело бы к уменьшению плотности упаковки, а следовательно, и стабильности (разд. 3.6). Инвариантные остатки Gly были обнаружены также в третьем положении в шести реверсивных поворотах цепи типа II у белков семейства цитохрома *c* [495]. Как видно из рис. 5.7, *б*, боковая цепь в этом положении невозможна по стерическим причинам.

Эволюционный путь, начинающийся от случайного набора частот встречаемости аминокислот, приводит к установлению существующих частот встречаемости. Как показали Кинг и Джукс, данные которых иллюстрируются на рис. 9.1, *а*, существует строгая корреляция между наблюдаемыми частотами встречаемости аминокислот и теми, которые можно ожидать при применении существующего генетического кода (рис. 1.5, *б*) к случайным нуклеотидным последовательностям в ДНК [144]. Эта корреляция явно указывает на случайность замен аминокислот, что подтверждает гипотезу эволюции, выдвигаемую «нейтралистами» [496]. Можно ли согласовать эти данные с установленным фактом консервативности замен аминокислот, т. е. с тем, что эволюция белка в значительной степени определяется отбором? По-видимому, можно.

Роль консервативных изменений хорошо видна из матрицы вероятности мутаций, приведенной в табл. 9.2. Элементы этой матрицы весьма неоднородны, что согласуется с большими различиями в тенденциях аминокислот к заменам. Однако более подробный анализ этой матрицы показывает, что наблюдаемое число мутаций, приводящих к данному остатку (последняя колонка) коррелирует также

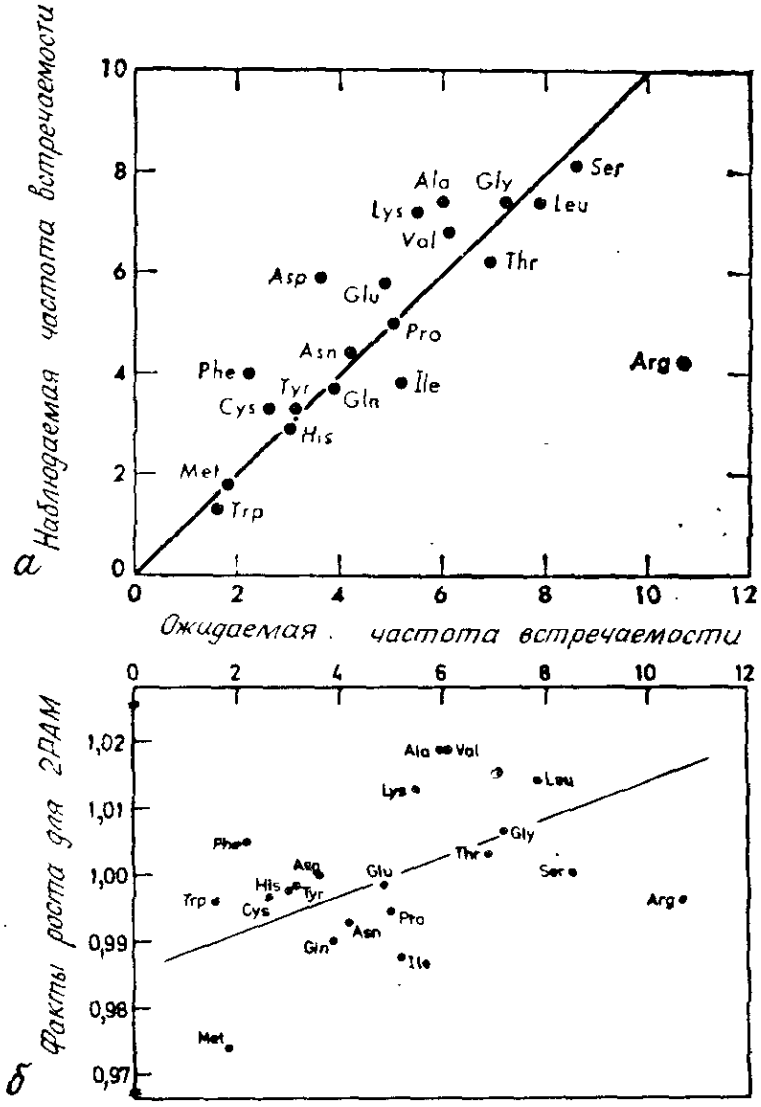


Рис. 9.1.

Частоты встречаемости аминокислот.

а — корреляция между наблюдаемой и ожидаемой частотами встречаемости аминокислот. Наблюдаемая частота встречаемости аминокислот получена по 53 полностью установленным аминокислотным последовательностям белков млекопитающих [144]. Ожидаемые частоты встречаемости рассчитаны по случайным последовательностям ДНК с помощью генетического кода [144]. Обе частоты встречаемости даны в процентах к общему аминокислотному составу. Ход прямой указывает на приблизительное равенство ожидаемых и наблюдаемых величин. *б* — корреляция между суммой значений, образующих строку табл. 9.2, и ожидаемой частотой встречаемости, определенной согласно рис. *а*. Эта сумма представляет фактор роста данной аминокислоты на эволюционном расстоянии в 2 PAM. Таким образом, если в данном наборе белков Ala имеет частоту встречаемости 1%, за 2 PAM она возрастет до 1,019%. Фактор роста приводит к достижению корреляции, показанной на рис. *а*, после значительных эволюционных изменений любого исходного соотношения аминокислот. Следовательно, корреляцию, приведенную на рис. *а*, можно объяснить с помощью табл. 9.2 вероятностей мутаций.

с частотой встречаемости этой аминокислоты, ожидаемой из случайных последовательностей ДНК (рис. 9.1, *б*). Таким образом, имеется соответствие между рис. 9.1, *а* и 9.1, *б*, т. е. тип остатка с высокой наблюдаемой частотой встречаемости имеет высокую вероятность появления в результате мутации. Поэтому любой эволюционный путь, начинающийся от данного (случайного) набора частот встре-

чаемости аминокислот, изменяется (т. е. модифицируется в соответствии с матрицей вероятности мутаций) таким образом, чтобы приблизиться к наблюдающемуся в настоящее время набору частот встречаемости аминокислот.

Итак, корреляцию на рис. 9.1, *а* можно объяснить с помощью матрицы вероятности мутаций. Корреляция выдерживается, хотя в каждом конкретном случае замены аминокислот находятся под сильным давлением отбора. Очевидно, что на основании корреляции между такими суммарными величинами, как частоты встречаемости аминокислот, невозможно заключить, что эволюция следовала нейтральному, т. е. недарвиновскому пути. Отметим, что корреляция, приведенная на рис. 9.1, *б*, пока еще никак не объяснена.

Влияние замен аминокислот на динамику свертывания

Замены, влияющие на процесс свертывания, исследуются на моделях — аналогах белков. В предыдущих разделах обсуждалось влияние замен аминокислот на функцию или на *стабильность* свернутых белков. Однако очевидно, что наиболее отрицательное воздействие мутация оказывает на динамику свертывания полипептидной цепи. Исследование этой проблемы на естественных мутантах затруднительно по двум причинам. Во-первых, если путь свертывания белка-мутанта полностью заблокирован, то полипептидную цепь невозможно идентифицировать и выделить обычными биохимическими методами (однако можно использовать иммунологические [94, 418] или комплементационные методики [446]). Кроме того, полипептиды, которые после их биосинтеза не свертываются совсем или свертываются слишком медленно, часто подвержены быстрому разрушению in vivo [154]. Эти трудности заставили искать модели для изучения влияния мутаций на свертывание белка среди полусинтетических аналогов белков [497—499] или белков с модифицированными боковыми цепями [445] (разд. 8.2).

Инициирующие спираль аминокислотные остатки были обнаружены путем сравнения последовательностей. Важность включения определенных групп для динамики свертывания была установлена при сравнительном исследовании молекул глобинов [500]. В этом случае все α-спирали несут на своих N-концевых участках или Pro, или остатки с короткими полярными боковыми цепями (Asn, Asp, His, Ser, Thr), которые могут образовывать водородные связи с основной цепью. Эти водородные связи, по-видимому, необходимы для инициации спирали, так же как в некоторых случаях и наличие пролина, фиксирующего необходимую для α-спирали величину двугранного угла. Эти факторы приводят к благоприятному соотношению между связывающей энергией и энтропией цепи; водородные связи увеличивают связывающую эиергию, а Pro понижает число возможных конформаций, а следовательно, и энтропию цепи. В соответ-

ствии с уравнением (3.2) значение $\Delta G_{общ}$ для этого локального участка цепи имеет большую отрицательную величину, в связи с чем такой участок проявляет сильную тенденцию к образованию небольшой нуклеации свертывания.

Высокий консерватизм укладки цепи связан с ограничениями, налагаемыми динамикой свертывания. Наблюдающаяся у белков тенденция к сохранению способа свертывания цепи в процессе эволюции, по-видимому, не может быть объяснена только на основе сохранения тонкого баланса энергии (или функции) данного белка. Оба эти условия в значительно большей степени относятся к боковым цепям аминокислот, чем к основной цепи. В связи с этим следует предположить, что укладка цепи сильно коррелирует с динамикой свертывания, которая и налагает основные ограничения на белок.

Замены аминокислот как эксперименты, проведенные природой

Белки-мутанты можно привлекать к интерпретации структурных принципов. Все фиксированные мутации белков можно рассматривать как эксперименты природы, которые указывают нам, какие вариации мало влияют на стабильность белка и на динамику свертывания. С другой стороны, случайные и, по-видимому, нефиксирующиеся мутации, как в аномальном гемоглобине, дают примеры вариаций, заметно понижающих стабильность белковой структуры. Оба типа мутаций можно использовать для совершенствования наших представлений о невалентных силах в белках. Для этой цели можно использовать процедуры минимизации энергии исходных и мутировавших полипептидных цепей на основе известных трехмерных структур [501]. Определенные таким образом разности энергий и геометрические отклонения можно сравнить с экспериментальными данными, полученными соответственно из термодинамических измерений [413, 417] и рентгеноструктурных исследований с высоким разрешением. Аналогичные сопоставления можно провести с помощью моделирования свертывания цепи (разд. 8.6), которое позволяет получить дополнительную информацию о некоторых аспектах процесса свертывания.

9.2 Филогенез на основе структур белков

Каждое положение остатка в цепи — это ее независимый признак с 20 различными характеристиками. Для определения филогенетической взаимосвязи необходимо выявить независимые признаки и затем провести количественное сопоставление соответствующих признаков. Поэтому при анализе специализации белков для установ-

ления филогенетических связей нужно вначале определить, какие признаки белковых структур могут считаться независимыми признаками. Наиболее подходящими для этого, по-видимому, являются последовательно расположенные аминокислотные остатки в первичной структуре белка. Именно они при сравнении любых белков дают наибольшее число признаков, т. е. число положений остатков. Каждое положение может быть в принципе занято любой из 20 стандартных аминокислот, т. е. каждый признак имеет 20 потенциальных отличительных характеристик. Хотя эти признаки явно зависят друг от друга в отношении отдельных мутаций (правило консервативных изменений, табл. 9.2), данные, приведенные на рис. 9.1, и многочисленные анализы аминокислотных последовательностей [145] свидетельствуют о том, что в первом приближении такие признаки можно считать независимыми, поскольку рассматривается большое число мутаций.

Белки, удобные для филогенетическнх исследований

Большинству требований филогенетических исследований соответствует цитохром *с*. В принципе для таких исследований можно использовать многие белки, однако наиболее удобным в этом аспекте оказался митохондриальный цитохром *с* (рис. 7.8), что было впервые продемонстрировано работами Смита, Марголиаша и Фича [502—506]. Критерии таких исследований и свойства цитохрома *с*, определившие выбор именно этого белка, перечислены ниже в пунктах *а — д*.

а) Однозначность пути гена в процессе эволюции. Построение филогенетического дерева основано на предположении, что рассматриваемый характер (признак) передается от предка к потомку и что другие пути отсутствуют. Перенос структурных генов от данного вида к неродственным видам, выполняемый, например, с помощью вирусов в качестве переносчиков, делает бесполезным сопоставление белков как средство установления филогенетической взаимосвязи. До сих пор явления межгенетического переноса гена для структурных генов в эукариотах не наблюдалось. Однако оно обычно для бактерий [507, 508], что означает невозможность установления филогенетического дерева для прокариотов на основе структурных сопоставлений белков.

б) Единственность гена, кодирующего белок. В гаплоидном геноме каждого эукариотического организма имеется только один ген цитохрома *с* (пищевые дрожжи — единственное известное исключение из этого правила) [509]. Присутствие в организме нескольких изофункциональных гомологичных белков усложнило бы выяснение взаимосвязи между генетически отдаленными организмами, поскольку в этом случае было бы неизвестно, какие белки нужно сравнивать. Изофункциональные белки часто встречаются на разных

онтогенетических стадиях организма, как, например, γ-, ε- и ζ-цепи гемоглобина приматов, в норме не развивающиеся в периоде внеутробной жизни. Для популяций или видов, которые используют эти типы цепей во «взрослой» жизни [510], было бы невозможно установить правильные связи по филогенетическому дереву, построенному по α- и β-цепям НьА, гемоглобина, который в норме образуется у взрослых особей [145]. Следует также учитывать принцип Гарстенга [511], согласно которому эволюция часто включает личиночные, а не взрослые формы; эволюционный путь киназ гуанидина можно проследить, только если следовать этому принципу [512]. Отметим, что правило Геккеля (онтогенез отражает филогенез) до сих пор не получило подтверждения на белковом уровне.

в) Одинаковая роль белка во всех сравниваемых организмах. Согласно имеющимся данным, все обнаруженные до сих пор цитохромы *с*-типа, выделенные из митохондрий, выполняют одну и ту же биологическую функцию.

г) Распространенность белка. Цитохром *с* найден во всех таксонах, для которых рассматриваются ответвления филогенетического дерева.

д) Простота анализа аминокислотной последовательности. Простота получения белка и длина цепи приблизительно в 110 аминокислотных остатков, содержащая достаточное число признаков, но еще сравнительно легко поддающаяся анализу аминокислотной последовательности — именно эти факторы в значительной мере определили выбор цитохрома *с*.

Филогенетическое дерево

Связи между очень отдаленными таксонами можно установить путем исследования специализации белка. Отправным пунктом любого эволюционного исследования, основанного на заменах аминокислот, является таблица различий между белками, относящимися к разным видам [145]. Эти данные представлены в виде упорядоченной матрицы различий; упорядоченность достигается путем помещения самых малых чисел, отражающих различия, вблизи диагонали. Основные этапы перехода от такой таблицы к установлению филогенетической связи описаны в разд. 9.3 на примере коротких пептидных гормонов. Для установления генеалогии белков, подобных рассмотренным митохондриальным цитохромам *с*, выделенным из более чем 70 организмов [509], были предложены детально разработанные схемы, например «метод родовой последовательности» [513] и «матричный подход» [504].

Установленная генеалогия цитохрома *с*, соответствующая филогенетическому дереву организмов, содержащих цитохром *с*, включает генетические расстояния (данные в единицах PAM) без указаний на времена эволюции. Однако можно произвести калибровку путем

сопоставления с филогенетическим деревом, основывающимся на макроскопических признаках и данных по ископаемым. Для датирования ископаемых используется геологическая временная шкала, устанавливаемая по радиоактивному распаду долгоживущих изотопов. Калибровочной константой является скорость эволюции, приведенная в табл. 9.1. Для случая митохондриальных цитохромов *c* эта скорость приблизительно постоянна для всех ветвей дерева, что приводит к линейному виду калибровочных кривых аналогично приведенным на рис. 9.2. Если предположить, что прямые для самых низких скоростей фиксации могут быть продолжены до времени, предшествующего эпохам, для которых известны ископаемые, то очевидно, что с помощью филогенеза, основанного на специализации белка, можно не только сопоставить такие отдаленные таксоны, как, например, растения и животные, но и указать дату для точки ветвления. Однако попытки таких обобщений принимаются весьма скептически. Действительно, если предположить, что окружающая среда налагала на ранние виды меньшее число ограничений, то скорость их эволюции могла быть значительно более высокой [483, 509].

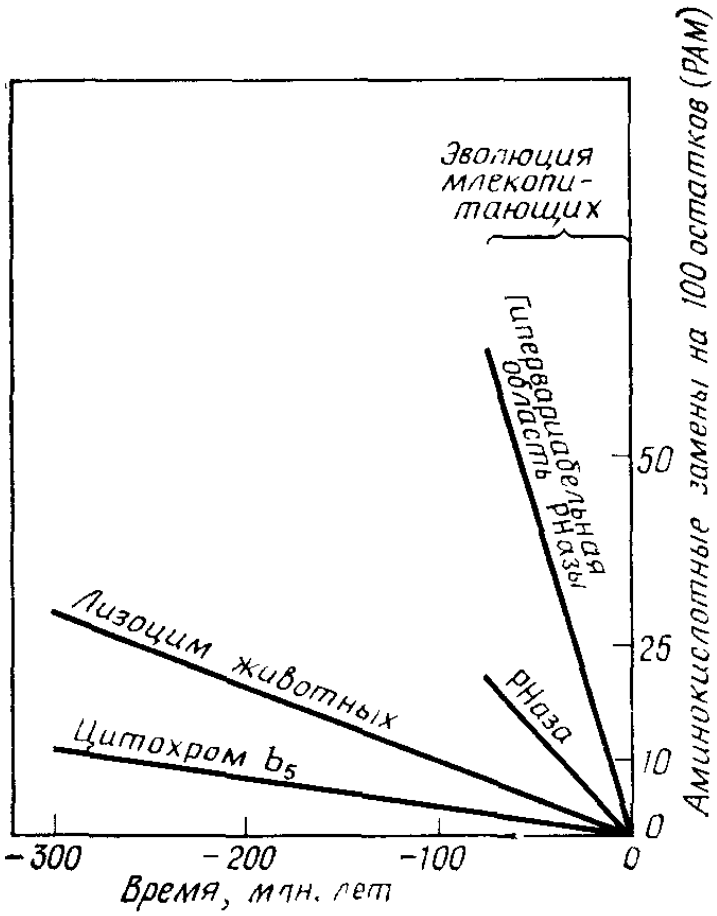


Рис. 9.2.

Графическое представление скоростей эволюции (даны в числе РАМ на 10^8 лет) в специализированных белках.

300 Ма (миллионов лет) назад произошла дивергенция линий, которая привела к пресмыкающимся и птицам, с одной стороны, и млекопитающим — с другой. Если вернуться на 600 Ма назад, эволюционные события, как указывает специализация белков, будут отвечать ископаемым данным.

Для цитохромов *c* характерна низкая скорость эволюции (табл. 9.1), поэтому их нельзя использовать для выяснения взаимосвязей в ответвлениях филогенетического дерева, например между коровой и другими парнокопытными. Для этой цели необходимо использовать белки, эволюционизирующие быстрее, например рибонуклеазу поджелудочной железы [145, 488] (рис. 9.2).

Эволюция поведения и морфологии в сравнении с эволюцией белка

Эволюция белка не является основой эволюции организмов. Тот факт, что специализацию белков можно использовать для изу-

чения эволюции, еще не означает, что эволюция организмов, заключающаяся в эволюции их морфологии и поведения, протекает параллельно эволюции белков. Это иллюстрируется следующими примерами, заимствованными из работ Уилсона и его сотр. [514—517].

Плацентарные млекопитающие претерпели более быструю эволюцию организмов по сравнению с низшими позвоночными, например с лягушкой. Однако белки млекопитающих специализировались, по-видимому, не быстрее, чем белки лягушки, и лягушки образуют более старшую группу (150 Ма*), чем плацентарные млекопитающие (75 Ма). Вследствие этого виды, настолько близкие по анатомии и способу существования, что их можно отнести к одному виду лягушки (например, *Rana*), могут отличаться последовательностью аминокислот данного белка так же сильно, как летучая мышь отличается от кита.

Другой пример относится к шимпанзе и человеку [514]. В среднем полипептид человека идентичен (более чем на 99%) соответствующему полипептиду шимпанзе; это согласуется с генетическим расстоянием между близко родственными видами разных млекопитающих. Однако анатомические различия между человеком и шимпанзе и различия в их поведении столь существенны, что их относят к разным семействам.

Специализация белков создает предпосылки для исследований на уровне организма. Согласно распространенной сейчас гипотезе в основе эволюции организма лежит, по-видимому, не эволюция структурных генов и соответствующих белков, а эволюция регуляторных генов [517]. Имеется много свидетельств тому, что адаптивная эволюция происходит прежде всего в результате изменений во взаимных отношениях структурных генов, а не в замещениях аминокислот в продуктах этих генов. В свете этой гипотезы увеличивается значение специализации белков для построения филогенетического дерева, поскольку предполагается, что положения аминокислот в полипептидной цепи являются признаками, не зависящими от эволюционных скачков. Таким образом, исследование специализации белка создает основу для более систематического изучения эволюции поведения и эволюции морфологии; это особенно важно для таксонов, которые бедны сведениями об ископаемых вариантах.

9.3 Дифференциация белков

Процесс, в результате которого гомологичные белки наделяются различной функцией, называют процессом дифференциации белков [473]. В ферментах мутация в одном или двух положениях аминокислот может изменять субстратную специфичность белка [508]. Соответственно изменение предрасположенности к субстрату S_x

* Ма = mega-anni = миллион лет.

в пользу субстрата S_y может привести к адаптации организма путем замещения фермента E_x на E_y. Экспериментально исследованы случаи такой адаптации в бактериях [519—524]. Хорошим примером является эволюция ацетамидазы в фенилацетамидазу [521] и рибитолдегидрогеназы в ксилитдегидрогеназу [508, 522]. В природе постоянно развиваются белки с новыми функциями. Это подтверждается эволюцией наследуемой сопротивляемости токсичным химическим веществам у насекомых и сопротивляемости антибиотикам у бактерий [519].

Обычно дифференциация белков начинается с дупликации соответствующего гена. Начиная с этого момента пути различных аминокислотных последовательностей расходятся в соответствии с различием функций. Классическим примером [522] дифференциации белков служит наличие нескольких цепей глобинов, например миоглобина и α-, β-, γ-, ε- и ζ-цепи гемоглобина человека.

Построение филогенетического дерева для вазотоцина, окситоцина и вазопрессина

При сравнении гомологичных белков, выполняющих разные функции, не предполагают и действительно часто не обнаруживают постоянства скорости фиксации аминокислотных замен. Это не создает больших помех при построении генеалогии белков, выполняющих различные функции, однако к датировке этапов эволюции на основе структурных сопоставлений таких белков нужно относиться с осторожностью.

В этом разделе рассматриваются генеалогические связи пептидных гормонов вазотоцина (ВТ),окситоцина (ОТ), бычьего вазопрессина (БВ) и свиного вазопрессина (СВ). Вазотоцин участвует в восстановлении и регуляции водно-солевого баланса многих позвоночных, окситоцин (млекопитающих) выполняет только первую, а вазопрессин — только вторую функцию. Эти гормоны, аминокислотная последовательность которых была определена в 1953 г. группой дю Виньё, представляют первый изученный пример дифференциации пептидов [145, 526]. Кроме того, их небольшая величина позволяет проследить основные этапы построения филогенетического дерева гомологичных (поли)пептидов.

Исходным пунктом является упорядоченная матрица различий. Первый этап состоит в построении для этих гормонов матрицы случайных и упорядоченных различий, показанной на рис. 9.3. При последующем уточнении эту матрицу превращают в матрицу «минимальных замен оснований» (МЗО)* в последовательностях, которую вновь упорядочивают. Введение шкалы МЗО приводит к увели-

* Возражения против использования шкалы МЗО при сравнении отдаленно родственных белков обсуждаются в работе [509].

чению составляющих матрицу элементов, так как более 33% наиболее часто встречающихся аминокислотных замен несовместимы с однократной заменой нуклеотида (см. рис. 9.11 из работы [203]). Это до некоторой степени повышает и точность, потому что в случае двойных и тройных нуклеотидных замен возможны более одной фиксированной аминокислотной замены в данном положении. Кроме того, многократные нуклеотидные замены могут соответствовать консервативным остаткам, замена которых может быть зафиксирована только в результате многократного нуклеотидного обмена. Таким образом, используемая шкала МЗО придает большие веса заменам в строго консервативных положениях.

Необходимо установить топологическую взаимосвязь между молекулярными видами. Перед построением филогенетического дерева в нашем примере нужно установить топологическую взаимосвязь между четырьмя видами пептидов. Это можно сделать графическим путем, как показано на рис. 9.3, *г*. Здесь однократные замены оснований отмечены крестиками, так что при переходе от одного вида молекул (например, СВ) к другому виду (например, ОТ) крестики, складываясь, дают элемент матрицы МЗО (например, 3). Если применить тот же подход, но воспользоваться матрицей аминокислотных замен, то для рассмотренного примера мы не получим решения, даже если расположить молекулы в виде топологического графа (рис. 9.3, *г*).

Более подробный анализ показывает, что проблема подгонки матрицы различий к топологическому графу соответствует системе шести линейных уравнений (шесть элементов матрицы, приведенной на рис. 9.3, *в*) и только пяти неизвестным (число крестиков в каждой ветви), которая имеет решения только в специальных случаях. В нашем случае мы получаем прямое решение матрицы МЗО для одной из трех возможных топологий. Можно применить и матрицу аминокислотных различий, если предположить, что мутация в ветви ОТ происходит в том же месте, что и мутация ветви СВ, т. е. при переходе от СВ к ОТ происходят три замены, хотя в матрице различий имеется только две.

Топологическую взаимосвязь можно превратить в филогенетическое дерево. Исходя из существующих молекулярных видов, нужно разместить предшественников в точках разветвлений. Точку разветвления выбирают таким образом, чтобы числа замен во всех ветвях до существующих видов были по возможности близкими. В нашем примере обе точки разветвления α и β в этом отношении эквивалентны (рис. 9.3), что приводит к выбору БВ и ВТ в качестве соответствующих предшественников. Если, однако, обратиться к организмам, то вопрос будет решен в пользу ВТ, поскольку он встречается в более древних таксонах [145]. С помощью дерева можно также установить, что аминокислотная последовательность бычье-

а

СВ	*Вазопрессин*	*(свиной)*	Cys-	Tyr-	Phe-	Gln-	Asn-	Cys-	Pro-	Lys-Gly
БВ	»	*(бычий)*	"	"	Phe	"	"	"	"	Arg "
ОТ	*Окситоцин*		"	"	Ile	"	"	"	"	Leu "
ВТ	*Вазотоцин*		"	"	Ile	"	"	"	"	Arg "

б

Обмен	*Кодоны*		МИО
Ile ⟷ Phe	AUU, AUC	⟷ UUU, UUC	1
	AUA		
Leu ⟷ Arg	CUU, CUC, CUA, CUG	⟷ CGU, CGC, CGA, CGG	1
	UUA, UUG	AGA, AGG	
Lys ⟷ Arg	AAA, AAG	⟷ AGA, AGG	1
		GCU, CGC, CGA, CGG	
Leu ⟷ Lys	CUU, CUC, CUA, CUG	AAA, AAG	2
	UUA, UUG		

в

	ОТ	БВ	ВТ	СВ
ОТ	0			
БВ	2	0		
ВТ	1	1	0	
СВ	2	1	2	0

→

	СВ	БВ	ОТ	ВТ
СВ	0			
БВ	1	0		
ОТ	2	2	0	
ВТ	2	1	1	0

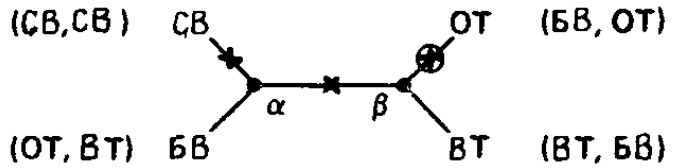
→

	СВ	БВ	ОТ	ВТ
СВ	0			
БВ	1	0		
ОТ	3	2	0	
ВТ	2	1	1	0

→

	СВ	БВ	ВТ	ОТ
СВ	0			
БВ	1	0		
ВТ	2	1	0	
ОТ	3	2	1	0

г



д

pv + *conn* + ot = 3 (2)
pv + *conn* + vt = 2
pv + bv = 1
bv + *conn* + ot = 2
bv + *conn* + vt = 1
ot + vt = 1

е

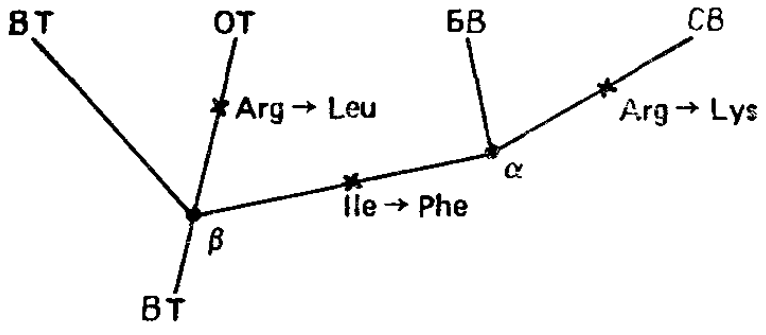


Рис. 9.3. Построение филогенетического дерева.

а — аминокислотные последовательности четырех пептидиых гормоиов. Замены происходят только в двух положениях. *б* — замены аминокислот, соответствующие кодоны и мнимальные изменения оснований (МИО). Кодоны, относящиеся к МИО-1 обведены. *в* — построение матрицы аминокнслотных различнй: отсутствие порядка → упорядочен-

го вазопрессина (БВ) имеет более древнее происхождение, чем последовательность свиного вазопрессина (СВ).

Отметим, что этот метод дает сведения об аминокислотной последовательности предшественников в точках разветвления. Именно этот этап является основным в упоминавшемся выше «методе предшествующих последовательностей». При сравнении более чем четырех (поли)пептидов в общем случае точного решения получить нельзя. В таких случаях, в частности, например, для цитохрома *c* (70 специализированных белков) и для сериновых протеаз (20 дифференцированных белков, см. ниже), метод предшествующих последовательностей [513] используется для достижения возможно лучшей подгонки при построении на основе имеющихся данных наилучшего филогенетического дерева. Очевидно, что в таком дереве будет наблюдаться тем больше неопределенностей в некой предшествующей последовательности, чем более удалена во времени точка разветвления.

α-Лактальбумин и лизоцим

Лактальбумин [517, 528] и лизоцим [518, 529—531] представляют классический пример двух белков с аналогичными последовательностями, но различными функциями и различными частотами фиксации мутаций. Предположение о структурном подобии обоих белков было впервые выдвинуто в 1958 г. и подтверждено спустя ~10 лет [523, 533] путем сравнения аминокислотных последовательностей. Некоторые важные для сопоставления свойства обоих белков приведены в табл. 9.3. Трехмерную структуру бычьего лактальбумина определили, основываясь на структуре лизоцима белка куриного яйца, путем построения модели [534] и последующей минимизации энергии [501, 535]. Эта процедура предполагает идентичность укладки обеих цепей, что представляется достаточно обоснованным, если учесть большое сходство аминокислотных последовательностей обоих белков (табл. 9.3). Этот пример показывает также, каким образом можно использовать данные по одному белку для структурного анализа отдаленно родственных гомологичных белков.

ность → переход от аминокислотных различий к МИО → другой тип упорядоченности. *г* — топологическая взаимосвязь между четырьмя пептидными гормонами. В скобках даны другие варианты топологии. Узловые точки α и β отвечают промежуточным, не обязательно существующим видам. Крестиками отмечены аминокислотные замены (или значения МИО) в соответствующих ответвлениях. Обведенный крестик представляет особый случай, когда происходит аминокислотная замена остатка, находящегося, как предполагается, в том же положении, где изменяется ветвь СВ—α. Следовательно, эта замена не означает перехода СВ—ОТ матрицы аминокислотных различий. *д* — система из шести линейных уравнений, отвечающая топологии, данной в *г*. *pv* означает число крестиков в ветви СВ—α, conn — число крестиков в ветви α—β и т. д. Система имеет решение, если первое уравнение приравнено 3, а не 2, т. е. если рассматривать аминокислотную замену в том же положении. При другой топологии решения нет. *е* — филогенетическое древо, построенное по топологии рис. *г*.

Таблица 9.3

Свойства α-лактальбумина [528] и лизоцима [518, 531]

	α-Лактальбумин	Лизоцим
Вероятное время дивергенции от общего предшественника [145]	$3 \cdot 10^8$ лет	
Судьба после дивергенции	Участвует в развитии новой биологической функции	Сохраняет функцию предшественника
Частота фиксации мутаций [145]	22 PAM/10^8 лет	10 PAM/10^8 лет
Различия аминокислот обоих белков человека	82 ≡ 62% ≡ 126 PAM	
Функция	Служит для инициации и прекращения синтеза лактозы путем взаимодействия со связанным с мембраной белком (аппарат Гольджи)	Гидролизует полисахариды оболочек бактерий
Местонахождение	В молоке (15% общего белкового состава молока)	Во многих тканях и секретах всех важных групп организмов; не ясно [241], все ли ферменты, проявляющие активность лизоцима, гомологичны

Трипсиноподобные сериновые протеазы

Трипсиноподобные сериновые протеазы [138, 536] образуют семейство расщепляющих белки ферментов, которые контролируют многие важнейшие физиологические процессы (табл. 9.4). Пищеварительный фермент трипсин, для которого и был впервые употреблен термин «энзим» (фермент), является наиболее изученным членом этого семейства. Он известен уже более ста лет, а его способность к расщеплению пептидных связей вблизи лизиновых и аргининовых остатков очень сходна со свойствами большей части других белков из этого семейства. Однако большинство родственных трипсину ферментов намного более специфичны, чем сам трипсин: каждый из них расщепляет в белке только одну или очень небольшое число пептидных связей. Структурная гомология сериновых протеаз была изучена и обобщена Хартли в 1970 г. [490]. Попарные сравнения трипсина, эластазы, химотрипсина и тромбина показывают, что около 40% их аминокислотных последовательностей идентичны (58 PAM). На сегодняшний день известны структуры первых трех из этих ферментов. Как и предсказывалось, все они имеют одинаковую укладку цепи [18, 243—245].

Таблица 9.4

Сериновые протеазы, участвующие в процессах, поддерживающих жизнедеятельность, и в процессах, угрожающих жизни человека

Протеаза	Местонахождение в активном состоянии	Управляемый (пато)физиологический процесс
Трипсин	Тонкий кишечник	Перевариванне белка
Тромбин	а) Кровь	Свертывание кровн
	б) Воспаленные ткани	Отторжение ткани
Плазмин	Раны и воспаленные ткани	Растворение временных сгустков фибрина
Акросомальная протеаза спермы	Яйцеклетка	Оплодотворение
Компоненты системы комплемента	Цитоплазматические мембраны бактерий, зараженные ткани	Разрушение чужеродных клеток, инициация процессов воспаления и т. д.
Калликреин	Кровь и ткани	Регуляция кровообращения и ощущение боли
Протеаза оболочки клетки	Воспаленные ткани	Разрушение патогенных и эластичных соединительных тканей
М-протеаза	Клетки саркомы	Рост опухоли
Д-протеаза	Клетки миокарда и другие клетки	Освобождение смертоносного дифтерийного фермента от токсина

Последовательности протеазы В и химотрипсина можно сопоставить на основании их трехмерной структуры. Трипсиноподобные ферменты были обнаружены также в микроорганизмах. *Streptomyces griseus*, например, содержит так называемую проте(ин)азу В, которая включает только 186 остатков (по сравнению с 234 в химотрипсине). Данные рентгеноструктурного анализа показывают, что укладка цепи протеазы В очень сходна с укладкой цепи химотрипсина [246]. Как можно заключить из трехмерных структур, делеции и вставки происходят только на поверхности молекулы. Это видно из рис. 9.4. Ранее уже говорилось о том, что сохранение внутренней части белка можно рассматривать как общее правило. В результате сопоставления аминокислотных последовательностей на основе геометрических данных выяснилось, что идентичны только 13% остатков, что отвечает генетическому расстоянию между химотрипсином и протеазой В в 432 PAM. При таком расстоянии невозможно установить гомологичность белков только по их аминокислотным последовательностям (разд. 9.6 и работа [803]).

Расположение остатков в каталитическом центре, своеобразном химическом ноже для расчленения пептидных связей, очень хорошо сохранилось во всех сериновых протеазах, тогда как специфичность

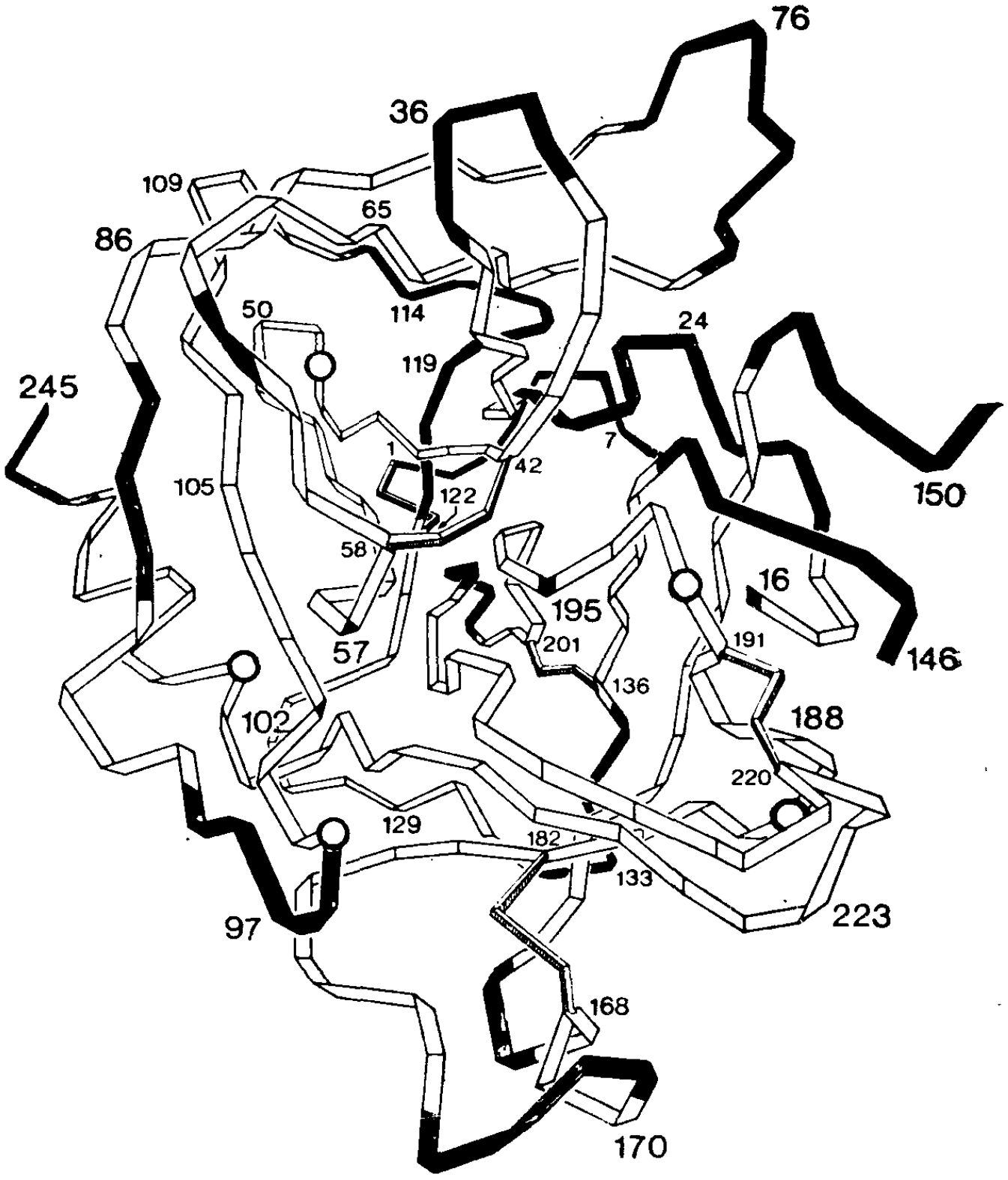


Рис. 9.4.

Полипептидные цепи α-химотрипсина и протеазы *B* из *Streptomyces griseus* [18]. Цепь химотрипсина показана в виде ленты, прогнутой при каждом атоме C_γ. Первый β-структурный цилиндр (рис. 5.17, *г*) находится в верхней левой части рисунка, а второй — в правой нижией части. Цепь протеазы *B* свернута так же. Но в ней отсутствуют те части ленты, которые зачериены. Кроме того, имеются вставки в местах, отмечеиных на ленте кружками. Отметим, что все вставки и делеции расположены иа поверхности молекулы.

к субстрату претерпела значительные изменения. Этот факт может означать, что каталитический механизм достиг физико-химического оптимума [537, 538]. Однако существуют и другие возможности для расщепления пептидных связей; в различных семействах протеиназ [145, 539] действуют разные каталитические механизмы.

Иммуноглобулины, антигены трансплантацин и пероксид-дисмутаза

Характерный участок свернутой цепи приблизительно из 100 остатков, присутствующий в совершенно различных белках, был назван «Ig-свертыванием» [540—543], поскольку впервые его обнаружили в иммуноглобулине [75, 292, 544, 545]. Ig-свертывание используется в качестве конструкционного элемента только в иммуноглобулинах, в цитозолиевой пероксид-дисмутазе [546] и, по-видимому, также в антигенах трансплантации (HL-A-белки) [86—88, 547]. Свойства этих белков представлены на рис. 4.2 и в табл. 9.5. Гомологичность последовательностей легкой цепи HL-A и С-домена у иммуноглобулинов настолько значительна (при генетическом расстоянии 217 PAM [541]), что представляется совершенно очевидным их общее эволюционное начало, а также подобие укладки цепи. Для последовательностей V- и С-доменов иммуноглобулинов гомологичность весьма незначительна, однако наблюдается очень большое подобие характера свертывания цепей. Среднеквадратичное отклонение расстояний между соответствующими C_α-атомами V_L- и $C_{H}1$-доменов составляет 1,2 Å по 56 эквивалентным остаткам [548], что, таким образом, указывает на еще один случай дифференциации (разд. 9.6). Для пероксид-дисмутазы и двух других типов белков не наблюдается гомологической связи и в аминокислотных последовательностях, однако укладка их цепей настолько сходна, что можно вполне предположить наличие отдаленного эволюционного родства [548]. Значение гомологичности в этом случае будет рассмотрено в разд. 9.6.

Обратимся теперь к функциям этих трех типов белков. Известно, что иммуноглобулины и HL-A-белки препятствуют вторжению в организм чужеродного белка. Пероксид-дисмутаза выполняет ту же функцию по отношению к небольшим радикалам O^-_2. Необходимость в такой функции должна была возникнуть с появлением кислорода в атмосфере Земли. Поскольку все три рассмотренных типа белков выполняют защитную функцию, они представляют такой случай дифференциации, при котором функции изменяются значительно, но не полностью.

«Свертывание типа глобинов»

Глобины — это мономерные или олигомерные гемсодержащие белки. Они встречаются в разнообразных организмах, в том числе в бобах, насекомых и человеке [145]. Представители семейства глобинов участвуют в транспорте O_2, в фиксации азота клубнями корней бобовых растений, регулировке наполнения плавательного пузыря некоторых видов рыб и в других биологических функциях [549, 550]. Были изучены трехмерные структуры большого числа

Таблица 9.5

Белки, содержащие фрагмент, свернутый по типу Ig

	Иммуноглобулины	Антигены трансплантации	Пероксид-дисмутаза
Свойства	Участвуют в процессах обнаружения и удаления вредных агентов, например токсинов, микроорганизмов или раковых клеток. Иммуноглобулин G содержит 12 g-доменов, по два V_L, C_L, V_H, C_H1, C_H2 и C_H3 (рис. 4. 2). Некоторые другие классы иммуноглобулинов содержат по несколько образований из 12 доменов [145, 542, 544]	Белки оболочки клетки, выполняющие различные функции во взаимодействиях клетка — клетка. Представляют главный барьер при пересадке тканей. Носят также название HL-A-белки (human leukocyte antigens) или H2-белки мышей. Каждая из двух идентичных тяжелых цепей содержит, по-видимому, три Ig-домена [86, 547]. Все легкие цепи идентичны; они состоят из отдельных Ig-доменов, которые были обнаружены в крови и моче и до выяснения их происхождения назывались «β_2-микроглобулинами» [88]	Димерный фермент, содержащий Cu^{2+} и Zn^{2+} [546]; расположен в цитозоле эукариотов; в митохондриях существует неродственный изофермент. Образует первичную защиту против радикалов O_2^-, важного кислородного токсичного агента, превращая их в O_2 и H_2O_2. Каждая субъединица содержит один Ig-домен
Последовательность аминокислот	Определены многие последовательности, см. обзор в [541]	Определена только последовательность легкой цепи [541, 603]	Последовательность известна [604—606]

Продолжение

	Иммуноглобулины	Антигены трансплантации	Пероксид-дисмутаза
Трехмерная структура	Исследовано несколько структур [42, 292, 294, 543, 545, 607, 608]		
Установленные гомологни последовательности	В *C*-доменах [541] н *V*-доменах [541]	В легкой цепн и *C*-доменах нммуноглобулинов [541]	Нет
Внутридоменные мостикн S—S в β-структурах, включающих 60—70 амнно-кислотных остатков	Присутствуют во всех доменах	Присутствуют во всех доменах [86]	Отсутствуют [286]
Положенне актнвного центра	Между V_L н V_H	Не обнаружено [547]	Две из трех петель между β-структурами, которые определяют места связывания антнгена в *V*-доменах нммуноглобулинов, соответствуют петлям, образующнм участки связывання Cu^{2+} и Zn^{2+} в активном центре [548]

этих белков [185, 188, 274—277] и выяснено, что для всех них характерен одинаковый способ свертывания цепи.

В историческом плане именно глобины определили появление идей о специализации и дифференциации белков. Еще в 1865 г. была выдвинута гипотеза, по которой глобины мышц [551] и глобины красных кровяных телец [552] представляют собой идентичные или родственные белки. Почти через 100 лет рентгеноструктурный анализ гемоглобина [553] и миоглобина [185] показал, что для обоих белков характерен особый способ свертывания цепи, так называемое свертывание типа глобинов (рис. 8.4). По данным анализа аминокислотной последовательности миоглобина [554] и гемоглобина [553], завершенного несколько позже, различие между ними составляет около 73% (192 РАМ).

Таким образом, уже на первых объектах анализа трехмерной структуры было установлено, что укладка цепи может сохраняться в высокой степени. Кроме того, стало очевидным, что белки, выполняющие различные функции, существующие в разных тканях и в разных организмах, могут иметь большое сходство и, следовательно, одинаковое происхождение. Эти факты явились важными стимулами к использованию структурного анализа белков, и в особенности анализа аминокислотной последовательности, для изучения эволюции на уровне макромолекул [145].

Морфологическое сходство глобинов и цитохромов *b*

Гомология цитохромов, находящихся в различных цепях электронного переноса, может указывать на гомологию цепей переноса. В настоящее время получены некоторые данные о том, что глобины родственно связаны с определенными цитохромами *b*-типа*; в этом отношении наиболее интересны цитохромы b_2, b_5 и b_{562}. Данные об этих цитохромах приведены в табл. 9.6. Как было предсказано по спектрам протонного магнитного резонанса [556] и установлено сравнением аминокислотных последовательностей [557], цитохром b_5-гсф (гем-связывающий фрагмент) и цитохром b_2-гсф представляют собой гомологичные белки с различием последовательностей около 72% (185 РАМ). Подобие оказывается еще более разительным, если сопоставить последовательность b_2-гсф с известной кристаллической структурой b_5-гсф [297, 557]. Гомологические связи цитохромов b_2-гсф и b_5-гсф показывают, что не только эти два белка, но и все цепи,

* Цитохромы классифицируют по их спектру поглощения в области длин волн между 400 и 600 нм. Эти спектры зависят от типа гема (цитохром *b*, например, содержит протогем IX) и от способа соединения гема с полипептидной цепью. Согласно этой номенклатуре, отнесение белков к цитохромам *b*-типа, не означает, что они имеют общее с ними свертывание цепи [509. 561].

Таблица 9.6

Свойства цитохромов b_2, b_5 и b_{562}

	Цитохром b_2 из митохондрий дрожжей [561], идентичный L-лактатдегидрогеназе (КФ 1.1.2.3) дрожжей	Цитохром b_5 из эндоплазматического ретикулума (микросом) печени теленка [561]	Цитохром b_{562} из *Escherichia coli* [561, 609]
Белок в целом	Растворимый тетрамерный белок; каждая субъединица (молекулярная масса 58 000) состоит из одного FMN-содержащего домена и одного гем-содержащего домена	Связанный с мембраной мономерный белок (молекулярная масса 16 700), способный к агрегированию in vivo. 40 из 50 С-концевых остатков находятся в мембране	Мономерный белок (молекулярная масса 12 000) из растворимой фракции экстрактов *E. coli*
Возможная функция всего белка	b_2 переносит электроны от L-лактата к цепи дыхания по следующему пути: L-лактат → b_2 (FMN-содержащий домен) → гем-содержащий домен b_2 → цитохром *c*	b_5 переносит электроны от NADH к некоторым функциональным звеньям метаболизма (например, к «десатуразе жирных кислот» [610]); путь электрона: NADH → FAD-содержащий белок → гем-содержащий домен b_5 → «цианид-чувствительный фактор»	Неизвестна
Гем-связывающий участок	b_2-гсф (ядро цитохрома b_2), гем-связывающий фрагмент, получен специфическим протеолизом b_2; содержит 95 аминокислотных остатков	b_5-гсф, гем-связывающий фрагмент, получен специфическим протеолизом b_5; содержит 93 аминокислотных остатка	Гемопротеид содержит 110 аминокислотных остатков
Известные структурные свойства гем-связывающего участка	Аминокислотная последовательность b_2-гсф [557] гомологична последовательности гем-связывающего фрагмента b_5 [611, 612] и может быть сопоставлена с известной третичной структурой b_5-гсф	Аминокислотная последовательность b_5-гсф [611, 612] гомологична последовательности гем-связывающего фрагмента b_2 [557]. Третичная структура b_5-гсф известна [297]	Аминокислотная последовательность b_{562} [613] может быть гомологична последовательности миоглобина кита [554]. Начат анализ третичной структуры b_{562} [563]

осуществляющие электронный транспорт (табл. 9.6), могут происходить от одного предшественника.

Структурное подобие цитохромов *b*-типа и глобинов. Аминокислотная последовательность цитохрома b_{562} отличается от последовательностей b_2-гсф и b_5-гсф, но, по-видимому, гомологична последовательности миоглобина [558]. Эти данные, а также результаты некоторых других исследований [559], показавших возможность существования непрерывного набора гомологичных структур в интервале от цитохромов *b*-типа до глобинов, дали толчок к сравнительным исследованиям [406, 560] известных* пространственных структур b_5-гсф и β-цепи гемоглобина. Это сопоставление показало [560], что из 85 остатков цитохрома b_5-гсф, пространственное расположение которых известно, 51 остаток имеет свой аналог в глобине. При совмещении обеих структур атомы железа гема оказываются на расстоянии, меньшем 1,4 Å, а угол между нормалями составляет около 10°. Это отвечает величинам, найденным при сравнении эритрокруорина (гемоглобин насекомого) с миоглобином [277] и при сравнении восстановленного цитохрома *c* с окисленным цитохромом *c* [273, 564]. Железокоординирующий гистидин и две соседние спирали гемоглобина соответствуют гистидину и соседним спиралям цитохрома b_5-гсф.

Хотя ориентации нормалей к гемам близки, сравнение цитохрома b_5-гсф и гемоглобина [560] выявило поворот кольца гема на 53° вокруг этой нормали. Авторы рассматривают это различие как энергетически минимальное и консервативное, поскольку боковые цепи гема непосредственно не вовлечены в функции какого-либо из белков: последнее положение нуждается в дополнительной проверке (ср. разд. 10.3). Таким образом, сопоставления структур гем-содержащих белков могут помочь в выяснении роли боковых цепей гема [565—568].

Итак, между глобином и цитохромом b_5-гсф существует подобие, распространяющееся на части спиралей *A*, *B*, *E*, *F* и *G* (рис. 8.4) и на расположение гема. Однако остается нерешенным вопрос о том, относится ли это подобие к гомологии, т. е. объясняется дивергентной эволюцией от общего предшественника, или к аналогии, т. е. возникло в результате конвергентной эволюции к физически предпочтительной пространственной структуре для невалентного связывания гема [569]. По-видимому, к решению этого вопроса может привести определение третичной структуры цитохрома b_{562} [563], которое восстановит недостающее в генеалогической связи звено (табл. 9.6). Вопросы выявления очень отдаленных взаимосвязей обсуждаются в разд. 9.6.

---

* Рентгеноструктурное исследование цитохрома b_2 [562] и b_{562} [563] в настоящее время продолжается.

Цитохромы типа *c*

Цитохромы типа *c* возникли в результате дифференциации белков. Как показано в разд. 9.2, митохондриальные цитохромы *c* образуют хорошо изученную группу специализированных белков. Другие цитохромы *c*-типа выполняют различные функции, большая часть которых, однако, состоит в присоединении и переносе электронов [509]. Все цитохромы *c* можно рассматривать как типичный случай дифференциации белков, однако поскольку сам термин «цитохром *c*» возник в результате спектральной, а не структурной классификации, какие-либо обобщения в настоящее время вряд ли возможны. Помимо митохондриального цитохрома *c* рентгеноструктурный анализ завершен для трех бактериальных цитохромов (табл. 9.7) и начат для еще трех прокариотных цитохромов, а именно cc_3, c_{555} и *c'* [509].

Цитохромы типа *c* подразделяются на малые и большие. Как показал анализ трехмерных структур и аминокислотных последовательностей, все известные цитохромы *c* имеют одинаковую укладку цепи. Однако на поверхности их молекул обнаружены большие вставки и делеции, аналогичные тем, которые отмечались для протеазы B и химотрипсина (рис. 9.4). Цитохромы *c* подразделяют на малые и большие структурные варианты в зависимости соответственно от отсутствия или наличия петли 40—55 (рис. 7.8).

Фотосинтетические и респираторные цитохромы *c* структурно столь же подобны, как химотрипсин и трипсин. Классификация цитохромов *c* с известной трехмерной структурой дана в табл. 9.7. Среди соединений большого варианта, первичная и вторичная структуры цитохрома *c* из митохондрий тунца, c_2 из *Rhodospirillum rubrum* и c_{550} из *Paracoccus denitrificans* (табл. 9.7), имеют различия, не бóльшие, чем отмечено для трипсина и химотрипсина; таким образом, цитохромы c_2, c_{550} и митохондриальный *c* можно рассматривать как родственные. Сравнения структур этих белков подтверждают гипотезу [570] о происхождении митохондрий от бактерий, родственных *Paracoccus*.

Попытки установить аминокислотную последовательность c_{551} (малый вариант) по последовательностям больших вариантов оказались неудачными. Только с помощью рентгеноструктурного анализа c_{551} удалось выявить родство обоих белков [571]. Благодаря этому рентгеноструктурному исследованию были ликвидированы неясности, связанные не только с установлением аминокислотной последовательности и не только у этого белка, но и у других цитохромов с известными последовательностями, например фотосинтетических цитохромов c_6 (называемых также *f*) прокариотических и эукариотических водорослей.

Таблица 9.7

Малый и большие структурные варианты цитохромов типа *c*[a]

	Малый вариант	Большие варианты		
Обозначение	c_{551}	c_2	c_{550}	*c*
Источник	*Pseudomonas aeruginosa*	*Rhodospirillum rubrum*	*Micrococcus denitrificans*	Митохондрии тунца
Молекулярная масса	8100	12 500	14 890	11 500
Имеющиеся структурные данные	Последовательность и пространственная структура	Последовательность и пространственная структура	Последовательность пространственная структура	Последовательность и пространственная структура
Структурная информация о родственных соединениях	>6 последовательностей	>5 последовательностей	>1 последовательности	> 67 последовательностей, 2 пространственные структуры
Аминокислоты, важные для присоединения гема	Cys-X-Y-Cys-His вблизи N-конца и Met ближе к С-концу полипептидной цепи. Оба остатка Cys присоединяются ковалентно к винильным боковым цепям группы гема. His и Met —аксиальные лиганды атома железа гема			
Характер ароматических боковых цепей	Только два или три положения в структуре непосредственно сопоставимы с большими вариантами	Идентичны в восьми положениях		
Включение со стороны «дна» молекулы (отвечающего положениям от 40 до 55 митохондриального *c*) приводит к делению на малый и большие варианты	Отсутствует		Имеется	
Роль в метаболизме	Дыхание (O_2)	Фотосинтез	Дыхание (O_2)	Дыхание (O_2)
Метаболическая роль ближайшего родственного соединения	Фотосинтез (c_6)	Фотосинтез и дыхание (O_2) (c_2 из *Rhodopseudomonas spheroides*)		Дыхание (O_2) (*c* из всех эукариот)

[a] Данные взяты из работы [509].

Среди больших вариантов имеются цитохромы, которые участвуют в переносе электронов в процессе дыхания, а также те, которые образуют части цепи переноса электронов при фотосинтезе; для малого варианта обнаружено то же различие функций. Это означает, что обе переносящие электроны цепи могут быть тесно связаны эволюционно. Эта гипотеза подтверждается тем фактом, что дыхание и фотосинтез объединены на молекулярном уровне в *Rhodopseudomonas spheroides*, где для выполнения обеих функций используется только один цитохром c_2 (большой вариант) [572]. До сих пор не известен организм, в котором обе функции выполнял бы малый вариант.

Цитохромы типа *c* могут способствовать выявлению эволюции метаболических путей. Цитохромы *c* вводят нас в обширную область прокариотов. В принципе структуры этих белков можно использовать для установления определенного порядка среди бактерий таким же образом, как митохондриальные цитохромы *c* были применены в таксономических целях к эукариотическим организмам. Первые попытки такой классификации бактерий уже сделаны [509, 571]. Однако, поскольку в бактериях может осуществляться переход межродового гена [507, 508], построение филогенетического дерева затрудняется генами, которые переходят из одной ветви в другую.

Однако, несмотря на это явление, предполагается, что цитохромы типа *c* могут помочь в прослеживании развития современного фотосинтеза и дыхания от более древних предшественников (например, фотосинтез на основе H_2S, дыхание, основанное на использовании сульфатов, и т. д.); они могут способствовать объяснению эволюции метаболических путей [571], что является одной из конечных целей исследований дифференциации белков [573, 574].

9.4 Слияние генов

Если замены аминокислот, вставки и делеции представляются сравнительно небольшими структурными модификациями в ходе эволюции, то процесс слияния генов вызывает очень существенные изменения: он приводит либо к объединению друг с другом разных полипептидных цепей, либо к дупликации данной цепи.

Процесс слияния генов

В прошлом слияние структурных зон ДНК, по-видимому, не было редким событием. Существуют надежные данные, указывающие на то, что слияния генов регулярно происходят в ходе онтогенеза у современных организмов, например на этапе, предшествующем синтезу иммуноглобулинов.

Слияние *V*- и *C*-генов иммуноглобулина происходит на уровне ДНК. Легкая (L) цепь молекулы иммуноглобулина состоит из вариабельного (*V*) и константного (*C*) участков (рис. 4.2, *в*). Струк-

турные зоны ДНК, кодирующие участки *V* и *C*, расположены раздельно, но на одной хромосоме. Несмотря на это разделение на уровне ДНК, было показано, что мРНК, отвечающая L-цепи, представляет собой отдельную непрерывную молекулу [576]. То же самое наблюдается и для тяжелых (H) цепей, которые состоят из одного *V*- и трех *C*-участков [477]. Более того, опыты с использованием меченых аминокислот выявили только одну точку роста при синтезе L-цепи, что указывает на отсутствие сращивания *V*- и *C*-цепей на белковом уровне*. В опытах по гибридизации с использованием L-цепи мРНК в качестве зонда наблюдалась перегруппировка ДНК в предшественниках клеток, продуцирующих иммуноглобулин [577]. Последний факт показывает, что слияние *V*- и *C*-участков происходит исключительно на уровне ДНК, а не на уровне РНК. Этот вопрос, однако, окончательно еще не решен [78].

Слияние разных цепей ДНК обычно также и у лизогенных бактериофагов, которые вводят свою ДНК в бактериальный геном, где она в течение некоторого времени остается недеятельной. Отметим также, что сплайсинг ДНК является основным этапом всех экспериментов в области генной инженерии [542, 575].

Слияние генов и эволюция

Как удалось установить на некоторых белках, слияние генов имело исключительно важное значение в процессе эволюции. Такого рода документация возможна потому, что составной ген в отличие от случая *V-C*-гена может оказаться и в линии клеток зародыша. Подобные случаи обнаружены при исследовании путей синтеза аминокислот и синтеза жирных кислот. Классическим примером одной полипептидной цепи, выполняющей две ферментативные функции, является аспартокиназа I — гомосериновая дегидрогеназа *E. coli* [578]. Сравнительное изучение ферментов, участвующих в синтезе Trp, выявило большое разнообразие в размещении по полипептидной цепи нескольких ферментативных центров [579], что указывает на возможность как слияния, так и расщепления генов.

Одним из наиболее удачных примеров является синтаза жирных кислот пищевых дрожжей [76, 580]. В гене имеются два локуса fas1 и fas2, в которых закодированы две полипептидные цепи, но восемь каталитических функций. Заметной корреляции между последовательностью биосинтетических реакций и порядком соответствующих функциональных доменов вдоль двух полипептидных цепей не существует. Например, порядок кодирования доменов в fas1 отвечает биосинтетическим стадиям 5, 6, 2, 8.

Слияние генов не приводит к резкому изменению свойств белков.

---

* Слияние на белковом уровне наблюдается в процессе биогенеза бактериофага λ [142].

Первым примером искусственного слияния двух ферментов на генетическом уровне было сращивание двух ферментов, участвующих в биосинтезе гистидина в *Salmonella typhimurium* [581]. Этот пример показывает, что само по себе слияние генов не приводит к резкому изменению свойств белков. Каждый из двух ферментов состоит из двух идентичных субъединиц, и эта структурная особенность сохраняется после слияния генов. Ферментативные и спектральные свойства индивидуальных белков также остаются неизменными.

Слившиеся белки имеют некоторые преимущества. Прежде всего функциональные домены синтезируются в стехиометрических количествах. Второе преимущество состоит в возможности появления выгодных домен-доменных взаимодействий, которые (в отличие от подобных взаимодействий в олигомерных белках) не зависят от концентраций белков [76]. И наконец, в специализированных биосинтетических путях образование лабильных промежуточных соединений, ведущих непосредственно к следующему ферменту, может оказаться весьма полезным.

Слиянию генов могла принадлежать важная роль в процессе эволюции основных метаболических путей. Энергетический путь метаболизма каждого из перечисленных ниже ферментов определился, вероятно, в результате объединения копии изначального (ди)нуклеотидсвязывающего домена с одним или большим числом других доменов, отличных от первого: фосфоглицераткиназа [235, 310, 311], дегидрогеназы, специфичные соответственно к глицеральдегид-3-фосфату, лактату, малату и алкоголю [91], и гликоген-фосфорилаза [236]. Как обсуждалось в разд. 5.4, (ди)нуклеотидсвязывающий домен представляет N-концевую часть первых четырех ферментов, тогда как в алкогольдегидрогеназе он расположен в С-концевой части, а в фосфорилазе — в середине цепи. Это указывает на то, что ограничения в пространственном расположении доменов не вызывали затруднений при их использовании в качестве составных блоков для построения самых сложных белков в процессе эволюции.

Мультипликация генов

До сих пор рассматривались вопросы слияния различных генов. Однако во многих случаях наблюдается слияние идентичных генов, что указывает на то, что перед слиянием генов происходит их дупликация.

Смежные мультиплетные гены кодируют генные продукты, например рибосомальную РНК и гистоны. Многократное повторение (без слияния) структурного гена, приводящее к специфическим новым генам, является известным эволюционным процессом. Смежные мультиплетные гены [582, 583] кодируют генные продукты, например рибосомальную РНК или гистоны [487]. В случае рибосомальной РНК преимущество по сравнению с одиночным геном оче-

видно: продукт транскрипции (если не учитывать реакции модификации) является конечным продуктом; многократного воспроизведения конечного продукта гена в рибосомальном механизме не предусмотрено. В случае гистонов было выдвинуто предположение, что транскрипция мультиплетного гена в гистон — мРНК дает большие количества этих белков, которые необходимы на ранних стадиях эмбриогенеза. Многочисленные копии гистон — мРНК, вероятно, вытесняют из рибосом другие виды мРНК [542].

По-видимому, мультипликация генов важна для дифференциации белков. Предполагается, что мультипликация генов играла важную роль в эволюции структуры и функции белков [523, 525, 582, 584, 585, 592]. После кратного воспроизведения генов одна копия выполняет первоначальную функцию, тогда как другая (или другие) может развиваться для выполнения близкой или новой функции. Наиболее хорошо известными примерами являются гены человека, кодирующие α-, β-, γ-, ε- и ζ-цепи гемоглобинов и миоглобин. Ведется полемика (работы [586] и [523]) по вопросу о том, сколько поколений избыточного гена могут существовать в геноме и являются ли недеятельные (спящие) формы этой экстракопии возможными промежуточными продуктами в процессе развития белка с новой функцией.

Повторяемость структуры в белке может также вызываться неодинаковым перекрестным соединением генов (кроссинговером). Мультипликация генов с последующим их слиянием приводит к генным продуктам с двумя или более идентичными субструктурами [587]. Однако, как показывает нижеследующий пример, к такому же результату могут привести и другие процессы. Случай частичной структурной дупликации обнаружен в редкой α_2-цепи гаптоглобина человека [145, 588]. Поскольку аминокислотные последовательности обеих частей идентичны, а также идентичны с большим участком обычной α_1-цепи, эта структурная дупликация должна была произойти совсем недавно. Скорее всего она вызвана хромосомной аберрацией (неэквивалентным кроссинговером) в предшествующей популяции (человека). Если бы это событие произошло намного раньше, так что гомология последовательностей оказалась бы стертой аминокислотными заменами, вставками и делециями, различить дупликацию и последующее слияние генов, с одной стороны, и хромосомную аберрацию — с другой, было бы невозможным. Поэтому все очень давно возникшие случаи структурных повторений обычно относят к «дупликациям генов», не пытаясь провести различия между разными механизмами.

Некоторые повторения можно обнаружить по повторяющемуся в структуре расположению дисульфидных связей. Структурные повторения часто могут быть идентифицированы по расположению связей S—S, как было показано на примере агглютинина зерна пшеницы, рис. 7.2, *а*. Четырехкратное повторение подтверждено нали-

чием соответствующего повторения в трехмерной структуре [316]. Сложная генеалогия сывороточного альбумина была постулирована на основе внутренних гомологий аминокислотной последовательности [82, 589], которые явствуют из расположения дисульфидных связей, рис. 7, 2, *б*. По-видимому, имели место следующие эволюционные этапы: на самой ранней стадии кодирование геном белка из приблизительно 80 аминокислот и образование S—S-связи → трипликация этой структуры → делеции приблизительно 30 остатков → дупликация полученной структуры, которая привела к белку, состоящему из 400 аминокислот, с последующей дупликацией половины этой структуры.

Изобилие структурных повторений свидетельствует о том, что новые формы укладки цепей были редкими, но очень важными нововведениями. Дупликация аминокислотной последовательности и укладки цепи обнаружены в ферредоксине (разд. 5.3). В парвальбумине повторение последовательности наблюдается в основном для остатков, находящихся в центре связывания Ca^{2+}, однако общее повторение укладки цепи явно проявляется в третичной структуре [59, 590]. Простое повторение типа свертывания цепи без какой-либо гомологии в аминокислотной последовательности встречается в роданезе (рис. 5.17, *а*), а также в трипсиноподобных белках, где эти участки имеют вид двух цилиндров (рис. 5.17, *г*). Все эти примеры показывают, что новый тип свертывания цепи возникал чрезвычайно редко, а новая укладка цепи повторялась и затем сохранялась в каждой копии.

Дупликация генов была постулирована для NAD-связывающих доменов четырех дегидрогеназ [91], в которых повторяется тип свертывания по Россману (рис. 5.12, *б*). Правда, эволюционная значимость этого обстоятельства (разд. 9.6) невелика, поскольку такое свертывание является энергетически предпочтительным элементом сверхвторичной структуры (разд. 5.2). Аналогичные случаи малой эволюционной значимости представляют структурные повторения внутри каждого из цилиндров сериновых протеаз (рис. 5.17, *г*), а также структурные повторения в триозофосфатизомеразе (рис. 5.17, *д*).

Гипотеза о происхождении всех белков от олигопептидов до сих пор не подтверждена. Наблюдающиеся структурные повторения привели к гипотезе [591], согласно которой все белки появились в результате генных дупликаций и слияний более коротких сегментов, состоящих приблизительно из 15 остатков. Однако эта гипотеза не была подтверждена при статистической обработке аминокислотных последовательностей 50 глобулярных белков [587]. Анализ трехмерных структур также мало способствует решению этого вопроса. В этом случае короткие повторения невозможно выявить, поскольку число возможных способов свертывания таких коротких участков цепи мало. В связи с этим любое структурное совпадение имеет до-

вольно низкую значимость (разд. 9.6). Число возможностей увеличивается, если сравнивать точные значения двугранных углов в основной цепи, а не только общий ход цепей. Однако эти углы не сохраняются точно неизменными в процессе эволюции.

В некоторых белках встречаются многократные повторения коротких последовательностей. Повторения коротких последовательностей обнаружены в так называемых периодических белках [145, 593], к которым относятся коллаген, кератин шерсти, гистоны, тропомиозин, липопротеин A1 человека и понижающий точку замерзания гликопротеин антарктической рыбы. Для последнего белка повторяющимся звеном во всей последовательности является Ala-Thr-Thr. В некоторых случаях периодичность может отражать специфические особенности соответствующей ДНК [593]; в других случаях структурные особенности (образование тройной спирали коллагена, показанной на рис. 5.6, характерное присоединение тропомиозина к нитевидному актину и гистона к двойной спирали ДНК) могли возникнуть под воздействием отбора.

9.5 Конвергентная эволюция белков

Число возможных аминокислотных последовательностей и способов свертывания цепи настолько велико, что случаи подобия нельзя рассматривать как случайно возникшие; очевидно, что они отражают историческое развитие, т. е. специализацию или дифференциацию белков. Однако громадное число вариантов строения не исключает возможности того, что некие структуры являются особенно предпочтительными и что определенные эволюционные процессы могут конвергировать к таким структурам. Наглядными примерами конвергенции в отношении белковой структуры являются α-спирали и β-складчатые листы; возможно, что это относится также к сверхвторичным структурам.

Эволюционная конвергенция различных структур к общей функции встречается довольно часто. Поскольку функция может осуществляться различными способами, разные белки могут приобрести способность к выполнению близких функций; их эволюционные пути конвергируют в отношении функции белка. Примером служит обратимое присоединение O_2 к гемоглобину, гемеритрину* и гемоцианину*. Молекула O_2 всегда присоединяется к железу или меди, ионы которых входят в состав совершенно различных белковых структур [594]. Аналогичные случаи конвергенции разных структур к общей функции отмечены для пероксид-дисмутаз, содержащих Mn^{2+} и пероксид-дисмутаз, содержащих Cu^{2+} — Zn^{2+} [546, 595].

Классическими примерами результатов конвергентной эволюции

---

* Эти названия произошли от греческого слова *haima* (кровь); они не означают, что белки обязательно содержат группу гема.

являются каталитические центры сериновых протеаз химотрипсина и субтилизина [18, 239, 240]. Для обоих ферментов при взаимодействии с субстратом характерно одинаковое расположение водородных связей, фиксирующих основную цепь субстрата, идентичное положение полости, принимающей боковые цепи, один и тот же механизм переключения заряда при расщеплении связей и один и тот же набор доноров водородных связей (NH-групп остова), фиксирующих карбонильные O^--группы каталитического промежуточного продукта [537]. Несмотря на все это, оба фермента совершенно не скоррелированы в отношении их аминокислотных последовательностей, укладки цепей и, например, нумерации остатков Asp, His, Ser [18, 239, 240] в цепи переключения заряда.

Таким образом, именно совершенствование каталитических поверхностей химотрипсина и субтилизина привело к принятию ими одинаковой функции. Как показано на рис. 11.1, механизм каталитического действия протеаз папаина и глицеральдегид-3-фосфатдегидрогеназы, фермента гликолитического пути, по-видимому, аналогичен механизму действия сериновых протеаз. Однако имеются и другие пути расщепления полипептидных цепей, как видно на примерах термолизина, катепсинов и кислых протеаз [539, 596].

Детали строения гомологичных белков могут быть следствием конвергентной эволюции. В процессе последующего развития белков от некоторого общего предшественника также можно выявить некоторые аспекты конвергентной эволюции в отношении общего построения этих белков [273, 597]. Например, в малом варианте цитохрома c_{551} глубоко лежащая пропионовая группа гема связана водородной связью с Trp-56, а в большом варианте митохондриального цитохрома *c* — с Trp-59* [509]. В этом случае важные для функции остатки Trp занимают неэквивалентные положения в гомологических полипептидных цепях. Это показывает, что моделирование аминокислотной последовательности с фиксацией положения функциональных остатков может привести к неверным выводам.

9.6 Обнаружение отдаленных родственных связей

Сравнение аминокислотных последовательностей

Априорная **значимость является верхней границей значимости родства.** Аналогия между двумя отдаленно родственными белками не всегда очевидна, и для ее нахождения нужны эффективные критерии. В качестве примера рассмотрим α-цепи гемоглобина человека

---

* Указанные остатки Trp занимают негомологичные положения в цепях, схемы нумерации в обоих цитохромах не были приведены в соответствие, как в случае схем на рис. 7.1, *а*. Рассматриваемый остаток Trp соответствует Trp-62 (рис. 7.8).

и миоглобин спермы кашалота, сопоставленные Дейхофом [20] при пренебрежении вставками и делециями. В этом случае длина совмещаемой цепи составляет 142 остатка, 37 из которых идентичны. «*Априорная* вероятность» нахождения общих остатков в любой из 37 позиций этих цепей очень мала:

$$\left(\frac{1}{20}\right)^{37} \cdot \left(\frac{19}{20}\right)^{105} \cdot \frac{142!}{37!\ 105!} = 5 \cdot 10^{-17}$$

И напротив, наблюдение такого подобия свидетельствует о том, что оно не может быть случайным и что имеет большую значимость. Соответственно «*априорную* значимость» определяют как величину, обратную «*априорной* вероятности». Высокая значимость подобия для биологических объектов указывает на эволюционное родство.

Биологические ограничения заметно снижают априорную значимость. *Априорная* вероятность является математическим понятием. Чтобы получить более реальные значения, следует учесть биологические ограничения. Так, не все аминокислоты встречаются с одинаковой частотой (табл. 1.1), в связи с чем вероятность появления данного остатка в данном положении не будет составлять $^1/_{20}$ для каждого из остатков. Кроме того, все остатки представляют собой составные элементы вполне определенной трехмерной структуры. Например, заряженный остаток, как правило, не может находиться внутри белка. Это повышает вероятность нахождения других остатков в этой области от $^1/_{20}$ до более высокого значения.

Биологические ограничения отражены в относительной частоте замен. Наилучшим образом эти ограничения можно учесть путем проведения анализа на основе экспериментальных данных по природным белкам. Для этой цели можно использовать матрицу частот относительных замен (табл. 1.2). Элементы этой матрицы представляют отношение частоты наблюдаемых аминокислотных замен к частоте, ожидаемой при случайных заменах для данного распределения частот встречаемости аминокислот (табл. 1.1). Поэтому эти элементы в среднем отражают природные ограничения, налагаемые на замены аминокислот.

При выявлении подобия двух последовательностей с помощью этой матрицы ответ для некоторого положения последовательности получают не в виде «да» (подходит) или «нет», а в более «сравнительной» форме: замена Ile ↔ Ile происходит с высоким уровнем значимости, а Ile ↔ Val (часто встречающаяся замена), Ile ↔ Thr и Ile ↔ ↔ Arg отвечают последовательно понижающимся абсолютным величинам элементов матрицы (табл. 1.2). Элементы, отвечающие всем положениям последовательности, можно сложить и сопоставить со случайными последовательностями. При этом необходимо предусмотреть, чтобы частота встречаемости аминокислот в случайных последовательностях была той же самой, что и в рассматриваемых белках. С помощью этой схемы можно обнаружить родство двух со-

вершенно различиых последовательиостей, содержащих большое число вероятных замеи. И иапротив, последовательиости с 15% идентичности и с большим числом маловероятиых замеи следует рассматривать как неродственные.

Эта схема была предложеиа Мак-Лахланом [598]. Он преобразовал матричные элементы $m(i, j)$ табл. 1.2 в целые числа от 0 до 7 путем пропорциоиальиого умеиьшения всех иедиагоиальиых элемеитов до уровня 0—5. Диагоиальные элементы оказались равными 6 и 7, что отражает иеиамиого большую зиачимость идентичности остатков по сравнению с часто встречающимися заменами. Для получеиия критерия подобия матричные элемеиты $m_r(i, j)$ для каждого положения r цепи, содержащей амииокислоту i в одной цепи и аминокислоту j в другой, суммируют. В иашем примере (миоглобин и α-цепь гемоглобина) этот критерий равен сумме:

$$M = \sum_{r=1}^{142} m_r(i,j)$$

Затем методами комбинаторики рассчитывают распределение всех сумм M, ожидаемое для случайных последовательностей с той же самой частотой встречаемостей аминокислот (см. работу [598]). Этот расчет заметно облегчается благодаря использованию в качестве элементов $m_r(i, j)$ небольших целых чисел.

Поскольку основной интерес заключается не в вероятности получеиия самих сумм M, а в вероятности того, что полученная сумма M будет означать родство, необходимо сравиивать совокупную вероятность (сумму всех вероятностей) всех сумм, больших или равиых M, с совокупной вероятностью всех сумм, меньших M. Последняя кумулятивная вероятность может быть положена равной 1,0, поскольку представляют интерес только высокие значеиия M, т. е. иизкие кумулятивиые вероятности всех сумм, больших или равных M, и поскольку сумма обоих типов кумулятивных вероятностей равиа 1,0. Распределеиие кумулятивных вероятностей всех сумм, больших или равных M, приведено на рис. 9.5. При сравнении миоглобииа кашалота с α-цепью гемоглобииа человека эта кумулятивная вероятиость достигает $2 \cdot 10^{-9}$ [598], что иа восемь порядков больше рассчитаниой выше *априорной* вероятности.

Метод Мак-Лахлана дает стандартную значимость для родства двух последовательностей. Поскольку кумулятивная вероятность учитывает биологические ограничения, налагающиеся на замену аминокислот (хотя и в весьма обобщенном виде), она намного более достоверна, чем *априорная* вероятность. Такую биологически согласованную вероятность называют «стандартной вероятностью», а обратную ей величину — «стандартной значимостью» [387]. Уровень, выше которого стандартная значимость означает эволюционное родство, не является достаточно четким. Ясно, что для величин, боль-

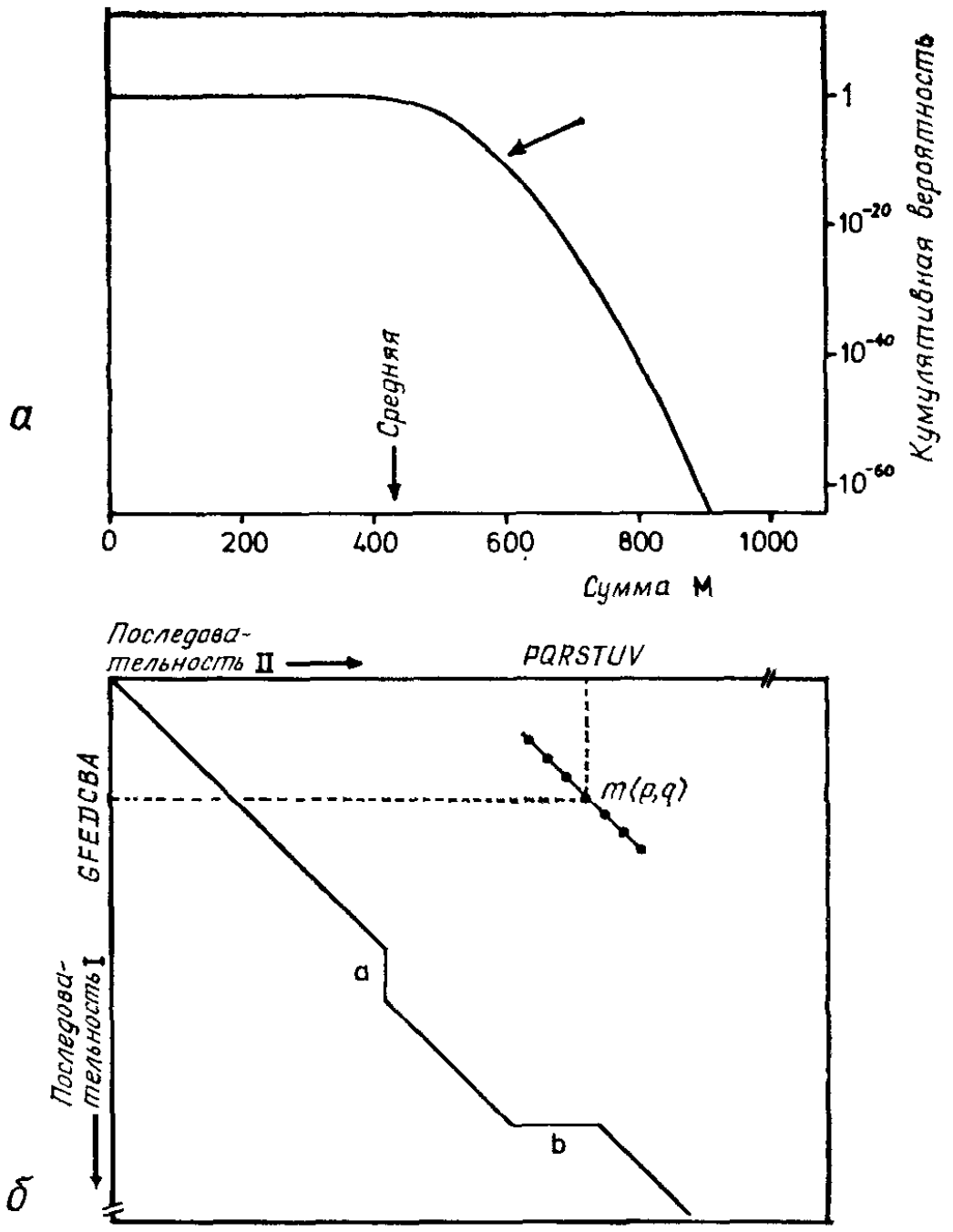


Рис. 9.5.

Сравнения аминокислотных последовательностей.

а — распределение кумулятивных вероятностей всех сумм, бо́льших или равных *M*, для случайных последовательностей в цепях, состоящих из 142 остатков [598]. О расчете *M* см. текст. Сумма и соответствующая кумулятивная вероятность для сравнения миоглобина кашалота и α-цепи гемоглобина человека указаны стрелкой. *б* — схема матрицы сравнения двух последовательностей I и II [598]. Оба отрезка *ABCDEFG* и *PQRSTUV* имеют длину 7 и при сравнении дают один элемент матрицы. Диагональ показывает корреляцию между последовательностями, которая выявляет вставки в последовательности I в положении *a* и в последовательности II в положении *b*.

ших 100, такое родство может рассматриваться как рабочая гипотеза. В принципе высокая стандартная значимость может исключить случай конвергентной эволюции, так как конвергенция определяется биологическими ограничениями, а они уже были учтены. В действительности, однако, никогда нельзя быть уверенным, что все ограничения известны.

Матрица сравнения может помочь в правильном расположении последовательностей. Рассчитанная выше стандартная значимость

предполагает точное совпадение без вставок и делеций, что мало вероятно в случае такого большого эволюционного расстояния, как между миоглобином и гемоглобином. Чтобы каким-то образом учесть вставки и делеции, необходимо произвести сдвиги определенных участков последовательности. Однако такие сдвиги увеличивают вероятность высоких значений суммы M, поскольку поиск хорошего соответствия осуществляется автоматически. Поэтому исходное распределение вероятностей (рис. 9.5, *а*) применять уже нельзя.

В этом случае прежде всего необходимо локализовать вставки и делеции. Для этой цели выбирают последовательность конечной длины (например, из семи остатков) и предполагают, что внутри ее не происходит сдвигов (т. е. вставок и делеций). Затем совмещают все возможные пары (p, q) нарушений, относящихся соответственно к положению p первой цепи и к положению q второй цепи. Каждая суперпозиция дает сумму $M(p, q)$, как показано на рис. 9.5, *б*. Эту сумму (или соответствующие обозначения) вносят в матрицу сравнения (рис. 9.5, *б*). В этой матрице линии высоких значений сумм означают наилучшее совпадение, нарушения этих линий указывают на вставки или делеции.

После установления совпадения по матрице, показанной на рис. 9.5, *б*, можно рассчитать с помощью процедуры, описанной выше для сравнения миоглобина с гемоглобином, кумулятивные вероятности, а следовательно, стандартную значимость родства.

Качественный критерий родства можно вывести из наблюдаемого распределения сумм *M*. Очень часто, однако, явного совпадения обнаружить не удается. В таких случаях наблюдаемое распределение *всех* значений $M(p, q)$ (20 000 при сравнении цепи из 100 остатков с цепью из 200 остатков) можно сравнить с распределением, ожидаемым при сравнении случайных последовательностей. Если высокие значения $M(p, q)$ встречаются намного чаще, чем можно было ожидать, то в последовательностях имеется много совпадающих участков и белки родственны друг другу. Однако степень этого родства трудно выразить количественно, поскольку стандартная значимость не может быть рассчитана.

Система суммирования исключает ложные признаки. Она может выявить структурные повторения. Поскольку этот метод очень высоко оценивает подобие аминокислот и не придает слишком большого веса их идентичности, он нечувствителен к неправильному выстраиванию, связанному с ложной маркировкой остатков, например Trp-59 и Trp-56 соответственно в цитохромах *c* и c_{551} (разд. 9.5).

Этот метод можно также применять и к сравнениям в пределах одной цепи для поиска повторений, которые могут указывать на дупликацию гена. Такие повторения могут проявляться в виде линий высокой значимости, параллельных диагонали.

В особых случаях, когда можно ожидать высокоупорядоченные повторения, как, например, в коллагене или в тропомиозине, такие

упорядоченности можно выявлять с помощью одномерного анализа Фурье [599]. В этом случае каждому типу остатков приписывается определенное число, например от 1 до 20 для аминокислот, расположенных в порядке табл. 1.2. Такое расположение есть одномерное представление вероятностей аминокислотных замен. Пользуясь этим распределением, данную последовательность переводят в непрерывную функцию, высота которой характеристична для свойств данной аминокислоты (очень высокие значения, например, отвечают ароматическим боковым цепям). Затем фурье-преобразование этой функции выявляет периодичности, т. е. повторения аминокислот с аналогичными свойствами.

Сравнения укладки цепи

Укладка цепи выявляет отдаленные эволюционные родственные связи. В предыдущих разделах упоминалось о том, что в процессе эволюции характер свертывания полипептидной цепи в пространстве сохраняется значительно лучше, чем ее аминокислотный состав. Поэтому тип свертывания цепи позволяет устанавливать такие отдаленные родственные связи, которые уже не проявляются в аминокислотных последовательностях. Для этого необходимо количественное описание способов свертывания цепей и сравнений укладки цепей.

Средние C_α-расстояния служат обычным критерием подобия укладки цепей. Тип свертывания цепи хорошо описывается координатами C_α-атомов. Поэтому укладку цепей можно сравнить путем наложения двух цепей с помощью минимизации расстояний между соответствующими C_α-атомами [600]. Среднеквадратичное или обычное среднее значение полученных расстояний между соответствующими C_α-атомами можно принять за показатель подобия при сравнении. Такой показатель (среднеквадратичное C_α-расстояние) в настоящее время является общепринятым. Его применяют, например, для проверки результатов моделирования свертывания цепи (разд. 8.6). Были также предложены и использованы для сравнения свертывания цепей [601] и более сложные показатели подобия, основывающиеся на расстояниях C_α, однако эти показатели не получили большого распространения.

Стандартную значимость для сравнения типов свертывания цепей пока определить нельзя. При сравнении свертывания цепей можно использовать среднеквадратичное C_α-расстояние таким же образом, как использовалась сумма M, введенная для сравнений аминокислотных последовательностей. Однако если кумулятивная вероятность для M могла быть рассчитана путем сравнения со случайными аминокислотными последовательностями (рис. 9.5, *а*), то в случае свертывания цепей ситуация более сложна, поскольку случайные укладки цепи неизвестны. Поэтому пока еще невозможно уста-

новить стандартную значимость такого сравнения, как это было сделано при сопоставлении последовательностей миоглобина и α-цепи гемоглобина, приведенном на рис. 9.5, *а*.

Генерация и сравнение случайных свертываний цепи для нахождения кумулятивных вероятностей среднеквадратичных C_α-расстояний требует больших объемов расчета [387]. Кроме того, при расчете таких распределений следует установить различия между структурным сходством, вызванным физико-химическими ограничениями (т. е. сходством вторичных и сверхвторичных структур) и сходством свободно варьируемых структур. Однако понимание структурных принципов в той мере, которая необходима для проведения таких различий, пока еще не достигнуто. Вместо случайным образом генерированных свернутых цепей можно использовать структурно изученные белки. Однако число известных типов свертывания цепей для этого недостаточно велико.

Матрица сравнения C_α-расстояний может выявить совпадение цепей. Структуры отдаленно родственных белков обычно сильно отличаются друг от друга за счет больших вставок и делеций, как это показано на рис. 9.4 для структур химотрипсина и протеазы В. При сравнении свертывания цепей, как и при сравнении аминокислотных последовательностей, требуется совпадение соответствующих остатков, хотя на среднеквадратичном C_α-расстоянии нарушения совпадения сказываются не так резко, как на сумме M. Такое совмещение можно производить по той же схеме, что и при сравнении последовательностей, рис. 9.5, *б*; $M(p, q)$ нужно заменить среднеквадратичными C_α-расстояниями между сегментами цепи данной длины, центрированными относительно p и q [802]. После установления совпадений можно определить общее среднеквадратичное расстояние между соответствующими остатками. Однако его опять-таки нельзя перевести в стандартную значимость подобия двух рассматриваемых белков, поскольку кумулятивные вероятности для среднеквадратичных C_α-расстояний пока не известны.

Матрица сравнения среднеквадратичных C_α-расстояний также может быть использована для выявлений структурного подобия между частями белков. Такое подобие обнаружено, например, между фрагментами лизоцима бактериофага Т4 и лизоцима белка куриного яйца [802]. Анализируемые длины сегментов составляли 40 и 80 остатков.

Сравнения укладки цепи должны учитывать вставки и делеции. При расчете среднеквадратичных C_α-расстояний между фрагментами цепи, которые сопоставляют с помощью матрицы сравнения, вставками и делециями полностью пренебрегают. Такое допущение, однако, слишком произвольно, поскольку информация о вставках и делециях необходима для определения показателя подобия. Действительно, если в сравниваемых белках их оказывается слишком много, соответствующие белки считают не проявляющими подобия.

При попарном сравнении свертываний цепи вставки и делеции можно учесть с соответствующими весами путем нахождения среднего C_α-расстояния по минимальной площади, заключенной между сравниваемыми цепями, как показано на рис. 9.6. Этот метод не требует совпадения. Однако минимизация площади, заключенной между дву-

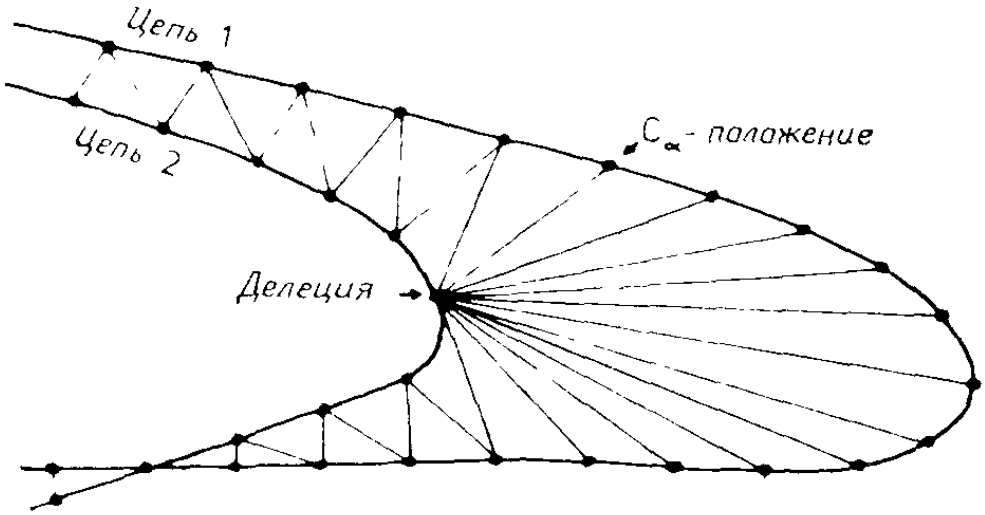


Рис. 9.6.

Расчет среднего C_α-расстояния между двумя цепями, учитывающий систематические вставки и делеции [387].

Среднее C_α-расстояние определяется как минимальная площадь, заключенная между двумя цепями (соответствующая площади эластичной пленки, натянутой между двумя проволоками) и отнесенная к средней длине цепи. Сумма площадей треугольников между C_α-атомами, которые показаны на рисунке, является хорошей аппроксимацией этой минимальной площади.

мя цепями, в вычислительном отношении более сложна, чем минимизация расстояний между двумя одинаковыми по величине наборами атомов, которые составляют базис при определении среднеквадратичного C_α-расстояния.

Сравнение топологии слоев

Сравнения топологии слоев подчеркивают важность для структур белков β-складчатых листов. Учитывая трудности перевода средних C_α-расстояний в стандартные значимости, следует упомянуть о менее строгом способе сравнения свертывания цепей, основанном на топологии β-структуры и, естественно, применимом только к белкам, содержащим β-складчатые листы. Подобную топологию определяют как ход цепи в β-структуре (субструктуре, считающейся жесткой) без учета точных координат [255] (рис. 7.7). Например, различное расположение листов вдоль цепи или сочленений между параллельными листами на разных сторонах β-структуры описывает различную топологию β-структур.

Топологию β-структур можно выразить количественно с помощью методов комбинаторики. Число различных топологий для

параллельной β-структуры из шести β-листов равно $6! \cdot 2^{6-2} = 11\,520$ [255]. Таким образом, *априорная вероятность* наличия идентичных топологий β-структуры в двух различных белках равна 1/11 520, а *априорная* значимость 11 520. Это сравнение ограничено только белками, включающими параллельные структуры из шести листов. Учет всех других белков увеличивает эту значимость приблизительно в 10 раз. Примером идентичной топологии β-структуры является NAD-связывающий домен дегидрогеназ [91]. В этом случае NAD-связывающие центры также идентичны по отношению к β-структуре (расположены над карбонильными концами β-листов, рис. 7.7), что приводит к увеличению априорной значимости [255] приблизительно в 10 раз, т. е. к общей величине значимости 1 152 000.

В двух случаях приближенные стандартные значимости родства определены по топологиям β-структуры. *Априорную* значимость можно снизить до приближенной стандартной значимости, равной 3500, если учесть наблюдаемую предпочтительность βαβ-звеньев (разд. 5.2) и наблюдаемую корреляцию по соседним остаткам в β-структурах (рис. 5.15, *а*); оба эти условия отражают пространственные ограничения, которые снижают число возможных топологий β-структур. Значимость, равная 3500, без сомнения, указывает на эволюционное родство дегидрогеназ.

Стандартная значимость была определена также для топологий β-структур иммуноглобулина и пероксид-дисмутазы [548]. Ее величина, равная 30 000 (по расчету авторов 3000, $\times 10$ в связи с учетом ограничения группы сравниваемых белков, как в случае дегидрогеназ), явно указывает на эволюционное родство. Хотя сравнения топологий β-структур возможны лишь в ограниченном числе структур, тем не менее они показывают, каким способом можно выполнять сопоставление и количественно оценивать по свертыванию цепи очень отдаленное родство белков.

Заключение

Огромное разнообразие белков есть следствие их эволюции. Эволюция явилась результатом многочисленных природных экспериментов (мутации и последующий отбор), которые могут быть использованы для изучения принципов строения белков. Основной мутационной ступенью эволюции белков является замена аминокислотного остатка; на следующих по значимости этапах происходят вставки и делеции одного или большего числа остатков; очень большие изменения являются результатом мультипликации и слияния генов.

Сопоставляя известные структуры белков, можно сделать следующие заключения: замены на поверхностях белков происходят значительно более легко, чем замены во внутренней части; функционально важные положения, например, активные центры практически инвариантны и т. д. Кроме того, стало очевидным, что очень близкие

трехмерные структуры могут быть образованы совершенно различными аминокислотными последовательностями. Необходимо отметить исключительно высокую сохранность свертывания цепи, которая показывает, что не только окончательная структура, но также и процесс свертывания находятся под мощным воздействием отбора. Дополнительную структурную информацию можно получить, исследуя мутации, еще не подвергшиеся действию отбора, например мутантный гемоглобин человека. В этом случае структурные принципы выявляются по отклонениям от нормальных свойств.

В процессе эволюции белков можно выделить тенденции к специализации и дифференциации. Специализированные белки выполняют одну и ту же функцию в разных организмах и могут использоваться для установления генеалогии организмов. Однако следует отметить, что специализация белков *не направляет* эволюцию организмов. Дифференциация белков — это процесс, ведущий к функциональному разнообразию гомологичных белков. Таким образом, исследование эволюции белков не только способствует проникновению в детали структурной организации белков, но также позволяет установить связи между белками, находящимися в совершенно различных частях метаболического пути. Таким образом, можно внести определенный порядок в огромный перечень существующих белков и вместе с тем выявить аспекты эволюции метаболических путей. Важным механизмом дифференциации белков является мультипликация и слияние генов.

Как и дивергентная эволюция (гомологичных белков), конвергентная эволюция (аналогичных белков) происходит на всех стадиях. Например, обычным явлением представляется конвергенция в отношении функции. Может также происходить конвергенция в отношении структуры, однако этот процесс трудно проследить на основании уровня современных знаний. В частности, при исследовании очень отдаленного эволюционного родства не всегда оказывается возможным провести различие между аналогичными и гомологичными чертами сходства.

Глава 10 Взаимодействия белок — лиганд

Белки проявляют селективность во взаимодействиях с компонентами клеток. В отличие от белков, встречающихся в природе, химически синтезированные полипептиды ведут себя как малые дети: они «задевают», связывают и разрушают многие низкомолекулярные метаболиты [614]. Природные полипептиды — белки, «приученные» эволюцией взаимодействовать только с небольшим набором молекул [586]. Это могло произойти только потому, что в отличие от синтетических полипептидов белки приобрели способность образовывать определенные компактные структуры. Специфическое связывание является индивидуальным свойством белков, и наиболее общим результатом организации белковых структур оказалась скорее тенденция к несвязываемости, чем к связываемости.

В живых системах роль кофакторов или метаболитов выполняют относительно немногие типы молекул. Кроме некоторых ионов [615, 616], наиболее типичными в этом отношении являются производные аденозина, глюкозы, порфирина и α-аминокислот. Большая часть белков специализируется во взаимодействии с одним или с большим числом таких стандартных строительных блоков [19].

10.1 Лигандсвязывающие центры иммуноглобулинов

Единственными белками, которые приобрели способность к образованию центров связывания с практически всеми видами молекул, являются иммуноглобулины (Ig). К наиболее хорошо изученным иммуноглобулинам относятся антитела плазмы крови, а также белки миеломы, которые образуются в раковых клетках, но в остальном ведут себя как обычные иммуноглобулины. Показано, что реакции, с помощью которых многие белки миеломы человека и мышей связывают лиганды*, очень близки к реакциям, характерным для системы лиганд — антитело (см. работы [540] и [617]).

---

* В биохимии «лигандом» называют малую молекулу, присоединенную к более крупной. В химии координационных соединений лиганды — это молекулы или ионы, группирующиеся вокруг некоторого центра, например иона железа. В этой главе термин «лиганд» используется и в том, и в другом смыслах.

Пространственное строение центров связывания было определено для двух различных белков миеломы: для иммуноглобулина New человека, который связывает производные витамина K [618], и для иммуноглобулина McPC 603 мыши [607], который связывает фосфорилхолин. На основании этих исследований были сделаны следующие выводы.

Случайное спаривание n_1 H-цепей и n_2 L-цепей* приводит к образованию $n_1 \cdot n_2$ различных центров связывания антигена. В обоих белках лигандсвязывающие центры расположены между доменом V_L легкой цепи и доменом V_H тяжелой цепи (рис. 4.2, *в*). Таким образом, центр связывания образуют две различные полипептидные цепи. Использование двух цепей может быть решением проблемы образования практически неограниченного числа различных центров связывания, исходя из ограниченного объема генетического материала. Если число различных L- и H-цепей равно соответственно n_1 и n_2, то имеются $n_1 \cdot n_2$ комбинаций V_L- и V_H-доменов, т. е. $n_1 \cdot n_2$ различных центров связывания [540]. Величина $n_1 \cdot n_2$ была оценена в 10^7 [542], поэтому числа n_1 и n_2 должны иметь порядок $10^3 — 10^4$.

Существует шесть гипервариабельных сегментов, каждый из которых содержит от 5 до 10 остатков на центр связывания. Укладка цепей во всех вариабельных доменах иммуноглобулина New и иммуноглобулина McPC 603, а также в других иммуноглобулинах очень сходна [540, 543, 544]. Структурные различия сосредоточены в основном в так называемых гипервариабельных петлях [619] — трех сегментах, содержащих от 5 до 10 остатков в обоих V_L- и V_H-доменах. Аминокислотные остатки, образующие место связывания, комплементарное данному лиганду, принадлежат этим шести гипервариабельным сегментам.

Картина вставок и делеций, характеристичная для гипервариабельных областей, определяет размеры центра связывания. В случае Ig New место присоединения представляет собой неглубокий (~6 Å) желоб с размерами 15 × 6 Å. Центр связывания фосфорилхолина в Ig McPC 603 расположен на глубине 12 Å.

Модель связывания в системе иммуноглобулин — лиганд

Иммуноглобулины с известной структурой используются для исследования иммуноглобулинов, обладающих другой специфичностью. Область иммуноглобулинов V_L/V_H можно рассматривать как состоящую из жесткого остова, к которому присоединены гипервариабельные цепи. Это создает основу для сравнительного ана-

* Буквами H и L обозначены соответственно тяжелая (heavy) и легкая (light) цепи.

лиза систем иммуноглобулин — лиганд путем построения модели с использованием координат остова (по областям известной структуры V_L/V_H) и аминокислотных последовательностей гипервариабельных областей анализируемого иммуноглобулина. Первым белком, исследованным с помощью такого подхода [291, 620], был белок IgA MOPC 315 миеломы мыши, который, как известно, присоединяет производные 2,4-динитрофенила. Полученные результаты были впоследствии уточнены Дьюком и сотр. [621], которые применили построение модели в сочетании с изучением ядерного магнитного резонанса, электронного парамагнитного резонанса и реакций химической модификации.

В принципе возможен систематический анализ любой системы иммуноглобулин — лиганд. Исследование моделей связывания может проводиться систематическим образом, поскольку у млекопитающих, например кроликов или коз, синтез антител может индуцироваться по отношению к любой специально подобранной молекуле (с низкой или высокой молекулярной массой). Используя аффинную хроматографию, получаемые антитела затем очищаются. Как правило, эта процедура приводит к вырожденному иммунному ответу, т. е. к синтезу нескольких видов антител несколькими клонами так называемых клеток плазмы [542]. Эти антитела отличаются константами сродства к лиганду, что проявляется также в химических, спектральных и иммунологических свойствах центров связывания. Последующие анализы аминокислотной последовательности показали, что различия вызваны аминокислотными заменами в гипервариабельных областях доменов V_L и V_H [622]. На первый взгляд кажется, что вырожденный ответ должен усложнить пространственную организацию центров связывания. В действительности, однако, в этом случае возможно объединить и взаимно контролировать данные по нескольким центрам связывания, что заметно увеличивает вероятность нахождения правильных моделей.

Системы иммуиоглобулии — лиганд как модели прочих взаимодействий белок — лиганд

Наши представления о взаимодействиях белок — лиганд до сих пор настолько ограниченны, что для получения надежной информации в каждом конкретном случае необходим рентгеноструктурный анализ [623]. Такой подход не может распространяться на все белки, хотя бы потому что большинство наиболее интересных белков, например мембранные белки, не поддаются кристаллизации. Поэтому важно по имеющимся уже данным вывести некие общие правила связывания лигандов.

Наиболее удобны для этой цели комплексы иммуноглобулин — лиганд, поскольку, как отмечалось выше, для них был развит достаточно систематический подход к проблеме.

Комплексы иммуноглобулин — лиганд имеют сходство с комплексами фермент — лиганд. Существуют явные параллели между связывающими свойствами иммуноглобулинов и ферментов [85, 624, 625]:

а) Как показывают значения констант скоростей (до 10^8 $M^{-1} \cdot c^{-1}$), стадия связывания часто представляет собой реакцию, контролируемую диффузией.

б) Связываемый лиганд имеет низкую вращательную свободу; он прочно и специфично удерживается невалентными взаимодействиями. Стандартные свободные энергии связывания варьируют от —6 до —15 ккал/моль, что отвечает константам диссоциации комплексов лиганд — фермент от $K = 10^{-4}$ M до $K = 10^{-10}$ M.

в) Связывание лиганда иммуноглобулинами, по-видимому, может быть описано моделью ключ — замок [44]. Это справедливо также и для многих взаимодействий фермент — лиганд, однако не относится к тем ферментам, которые при связывании лиганда претерпевают большие конформационные изменения (индуцированная комплементарность) [626].

Архитектура иммуноглобулина может служить основой для синтеза in vitro пептидов с заданными связывающими свойствами. Для теоретических и практических исследований может оказаться крайне полезным синтез in vitro полипептидной цепи с определенной специфичностью и сродством к данному соединению. Один из возможных путей может начаться с природной или синтетической области V_L/V_H без гипервариабельных петель в качестве остова. Путем включения подходящих последовательностей на место гипервариабельных сегментов можно затем сформировать специфичный центр связывания рассматриваемого лиганда без нарушения процесса свертывания и стабильности остова [498]. Пример Cu^{2+} —Zn^{2+}-содержащей пероксид-дисмутазы [286] можно рассматривать как природный прецедент этого метода пептидной инженерии. В этом случае геометрия координации атомов металла в активных центрах имеет очень много общего с соответствующими фрагментами кристаллических структур медь-имидазольных и цинк-имидазольных комплексов [661]. Таким образом, обе основные особенности этого фермента, структура иммуноглобулина и комплекс металла, могут быть воспроизведены химиками-органиками.

10.2 Центры связывания субстрата сериновых протеаз

Механизм связывания субстрата аналогичен у всех известных сериновых протеаз [537], включая субтилизин [627]. Поскольку субстратами сериновых протеаз являются белки, выяснение механизма образования комплексов этих ферментов с их субстратами также дает информацию о взаимодействии белок—белок.

Образование комплекса химотрипсина с субстратом

Связывание субстрата можно разделить на несколько стадий. В каждой стадии принимают участие взаимодействия с атомами основной цепи химотрипсина. Рассмотрим этапы, предшествующие расщеплению химотрипсином пептидной связи со стороны карбонильного конца аминокислотного остатка i [537, 538]. Этот остаток

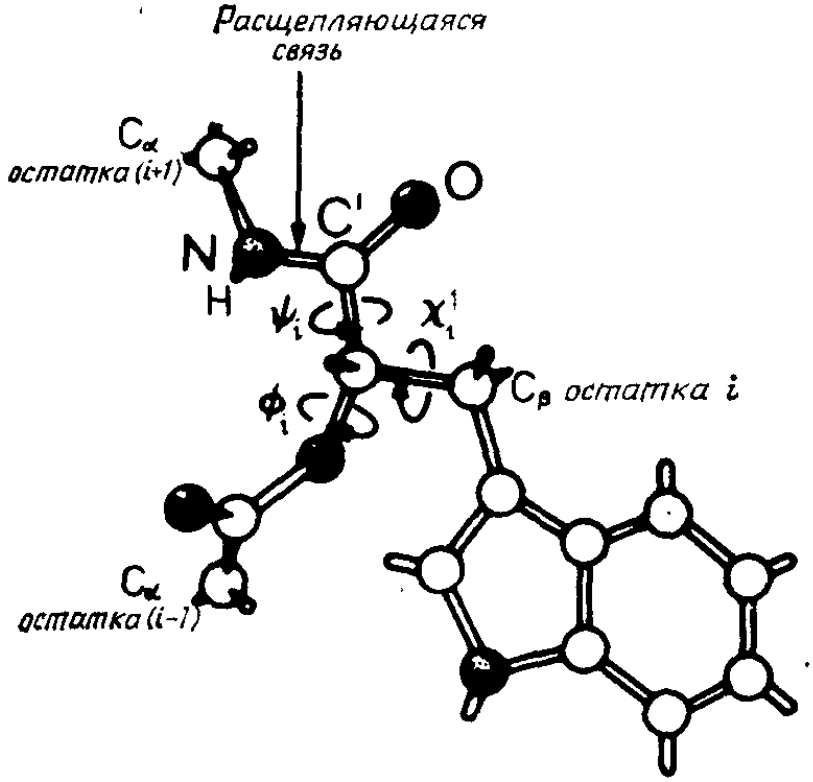


Рис. 10.1.

Присоединение субстрата к химотрипсину [537].

Показаны двугранные углы субстрата, которые должны принять определеиные значения перед тем, как пронзойдет расщепление пептидной связи между остатками i и $i+1$.

может быть либо ароматическим, либо нести объемную алифатическую боковую группу. При образовании комплекса расщепляющаяся связь субстрата ориентирована точно по направлению к атакующим группам [537]. Начиная с фиксации боковых цепей, этот процесс можно разделить на несколько стадий. Для этих стадий следует отметить важность взаимодействий, в которых участвуют неподвижные атомы основной цепи.

а) Боковая цепь i, включая C_β-атом (рис. 10.1), попадает в глубокий карман фермента, образованный тремя пептидными связями, и фиксируется гидрофобными взаимодействиями. Форма входа в карман оставляет мало позиционной свободы для C_α-атома этого остатка.

б) При фиксированной боковой цепи двугранный угол χ^1_i (рис. 2.2 и 10.1) определяет положение основной цепи субстрата. Допустимую вариацию χ^1_i определяет поверхность фермента. Значение χ^1_i фиксируется образованием водородной связи между NH-

группой остатка субстрата и карбонильной группой остатка 214 основной цепи, вследствие чего фиксируется также угол ϕ_i.

в) После установления этих взаимодействий необходима фиксация угла ψ_i. Выбор в этом случае также контролируется профилем поверхности фермента, который допускает для угла ψ_i только два интервала значений. Одна из этих возможностей позволяет карбонильной группе остатка i образовать водородные связи с NH-группами остатков 193 и 195, входящих в основную цепь. Это приводит к сближению амидной группы расщепляющейся связи (т. е. NH-группы остатка $i + 1$ субстрата) с центром химического управления фермента — линейной «релейно-зарядной системой» [628], состоящей из боковых цепей Asp, His и Ser. Теперь можно считать, что фермент-субстратный комплекс (комплекс Михаэлиса) образован и пептидная связь между остатками i и $i + 1$ готова к ферментативной атаке.

Непродуктивное связывание предотвращает гидролиз пептидов, состоящих из нежелательных D-аминокислот. Пептиды, состоящие из D-аминокислот, также могут прочно связываться химотрипсином. Однако в этом случае образуется сравнительно малореакционноспособный фермент-субстратный комплекс, поскольку расщепляющаяся связь не ориентирована должным образом относительно каталитического центра [629]; таким путем свободная энергия связывания расходуется на ингибирование реакции с аналогом субстрата, которая могла бы привести к нежелательным продуктам. Непродуктивное связывание, по-видимому, является общим механизмом, обеспечивающим специфичность фермента [630, 631].

Центр связывания субстрата в трипсине

Электростатические взаимодействия вносят вклад в специфичность трипсина к остаткам Lys и Arg. Трипсин [244, 245, 536] связывает свои субстраты существенно тем же способом, что и химотрипсин. Однако трипсин специфичен к положительно заряженным остаткам субстрата; боковая цепь Lys или Arg электростатически связывается с остатком аспарагиновой кислоты на дне связывающего кармана фермента. Кристаллографические исследования комплексов трипсина и белковых ингибиторов трипсина [269, 632] показали, что способ связывания очень сходен с образованием комплекса сериновая протеаза — субстрат. Очевидно, ингибитор точно воспроизводит субстрат. Механизм, ведущий к расщеплению субстрата трипсином и к стабилизации комплекса трипсин — ингибитор [269, 536], рассматривается в разд. 11.2.

10.3 Центры связывания гема

Протогем IX встречается в гемоглобине, в цитохромах типа *b* и *c*, а также в каталазах и пероксидазах. Мы остановимся на этих

семействах гем-содержащих белков в связи с тем, что трехмерные структуры этих белков известны. Такие данные получены для гемоглобинов [85, 550, 634], включая и мономерные виды, как, например, миоглобин млекопитающих [185, 635, 795] и эритрокруорин личинок насекомых [277], цитохромов *c* [509] и цитохромов *b* [297, 562]. Другими примерами гемсодержащих белков служат каталазы и пероксидазы. Из гемов, существующих в природе [637], наиболее хорошо изучен протогем IX*, который встречается исключительно в рассмотренных ниже примерах. Длина полипептидной цепи, образующей гемсвязывающий домен, составляет около 100 остатков в случае цитохромов и около 150 остатков в случае глобинов. Однако известен пример необычного способа присоединения гема. В цитохромах c_3 из видов *Desulfovibrio* четыре гемгруппы присоединены одной полипептидной цепью, состоящей приблизительно из 120 аминокислот. Это означает, что для связывания одной группы гема используются только от 20 до 30 аминокислотных остатков (см. обзор в [509]). Четыре гема функционируют более или менее независимо друг от друга [638].

Гемсодержащие белки являются переносчиками электронов или малых молекул, таких, как O_2. В гемоглобинах функция гема и окружающей его полипептидной цепи состоит в обеспечении связывания молекулярного кислорода железом и в защите координированного ферроиона от окисления [639]. В цитохроме *c* функция атома железа в геме заключается не в координации малой молекулы, а в переносе электронов; в ходе метаболизма энергии железо ферментативно восстанавливается ($Fe^{3+} \rightarrow Fe^{2+}$) и окисляется соответствующими белками — партнерами цитохрома *c* [509]. Цитохром b_5 — составная часть другой группы электронпереносящих белков, которые участвуют в расщеплении жирных кислот и других химических реакциях [297]. Было выдвинуто предположение [640], что цитохром b_5 может взаимодействовать in vivo с цитохромом *c*. Однако пока установлено, что восстановление цитохрома *c* цитохромом b_5 может происходить только in vitro. Недавно была предложена структурная модель этого взаимодействия [640].

Связи между гемом и апобелком

Шестьдесят атомов апобелка участвуют в вандерваальсовом контакте с порфириновым циклом. Главное сходство между гемоглобинами, цитохромами *c* и цитохромами *b* состоит в том, что во

---

* Протогем IX называют также ферропротопорфирином IX [633]. При нашем рассмотрении тетрапиррольный цикл этого соединения можно считать плоским. К этой циклической системе присоединены восемь боковых цепей — четыре метильные, две винильные и две пропионовые группы (рис. 10.2). «Прото» относится к происхождению соединения, а «IX» означает порядок чередования боковых групп.

всех этих молекулах группа гема находится в полости, образованной боковыми цепями неполярных аминокислот. В субъединице гемоглобина в вандерваальсовом контакте с атомами порфиринового цикла участвуют около 60 атомов [60].

Взаимодействие боковых цепей гема и апобелка может осуществляться всеми известными способами связывания. Боковые цепи гема в различных гемсодержащих белках выполняют разные роли. В цитохроме *c* обе винильные группы (в исключительных случаях только одна [509]) образуют ковалентные связи с Cys-14 и Cys-17 апобелка; одна заряженная пропионовая группа находится на поверхности молекулы, а вторая взаимодействует посредством водородной связи, по-видимому имеющей функциональное значение, с внутренней частью молекул. Одна или обе пропионовые группы, по-видимому, существенны для структуры или функции (или для обеих) пероксидазы из хрена [641] и цитохрома *b* из золотистой фасоли [642].

В случаях цитохрома b_5 и гемоглобина более ранние исследования показали, что боковые группы выполняют лишь небольшую и неспецифическую роль [560, 568, 637]. Однако значение отдельных боковых цепей гема было пересмотрено в свете недавних рентгеноструктурных данных [565, 566]. В гемсодержащем фрагменте цитохрома b_5, например, одна пропионовая группа может непосредственно участвовать в окислительно-восстановительном цикле, нейтрализуя избыточный положительный заряд атома железа в феррицитохроме b_5 и присоединяясь к другой катионной группе ферроцитохрома b_5 [556].

Атомы боковых цепей His и Met координируют атом железа гема. В цитохроме b_5 атом железа имеет симметричную координацию, которая образована шестью октаэдрическими позициями лигандов, занятыми четырьмя пиррольными атомами азота гема и двумя ε-атомами азота His-39 и His-63, находящимися в аксиальных положениях [297]. δ-Атомы азота His-39 и His-63 образуют водородные связи с атомами кислорода карбонильных групп основной цепи таким образом, что оба имидазольных цикла прочно закреплены благодаря жесткой структуре основной цепи.

Гемоглобины и цитохромы *c* имеют одинаковые лиганды, но их молекулы отличаются тем, что одно из аксиальных координационных мест занято соответственно молекулой O_2 и атомом серы метионина. Если молекула O_2 отсутствует, шестое координационное место («лигандный карман») гемоглобина остается незаполненным [60, 636, 643]. Предполагают, что в бактериальном цитохроме *c* не заняты оба аксиальные положения лигандов. По-видимому, в этом случае окружающая гем полипептидная цепь имеет конформацию, препятствующую связыванию потенциальных лигандов [644].

Влияние группы гема на структуру белков

Основное различие между структурами цитохрома *c*, с одной стороны, и гемоглобина и цитохрома b_5 — с другой, заключается в типе свертывания полипептидной цепи, а также в различной роли гема в пространственной организации белка.

Апоцитохром *c* полностью развернут. Молекулу цитохрома *c* можно представить в виде вытянутой полипептидной цепи, окружающей гем своеобразной оболочкой толщиной в одну полипептидную цепь (рис. 7.8). Апоцитохром *c* с отщепленной группой гема совершенно несвернут. Центральное положение гема, как составной части структуры, подчеркивается тем обстоятельством, что эволюционно наиболее хорошо сохраняются те аминокислотные боковые цепи (если сопоставлять цитохромы с известными последовательностями из 70 источников [509]), которые находятся вблизи атома железа и винильной группы, а также гидрофобные остатки, упакованные вокруг гема.

Апомиоглобин имеет богатую спиралями глобулярную структуру, подобную структуре миоглобина. В отличие от неустойчивой напоминающей кокон оболочки гема в цитохроме *c* в гемоглобине полость гема образована из жестких стабильных α-спиралей. Группа гема здесь не является основным элементом, определяющим укладку цепи: апомиоглобин, например, образует третичную структуру и в отсутствие простетических групп (гл. 8), однако группа гема увеличивает ее стабильность [415, 461, 645, 646]. Более мягкие структурные требования к взаимодействию гема и белка в гемоглобинах отражает также тот факт, что для членов семейства глобинов [277, 634] инвариантна только общая картина неполярных контактов гема. Структурно эквивалентные положения у них заняты различными аминокислотными остатками.

В апоцитохроме b_5 имеется такой же жесткий участок связывания гема, как и в гемоглобине. В случае цитохромов *b* базовый набор слишком мал для того, чтобы установить, насколько инвариантны неполярные остатки в специфических положениях, взаимодействующих с группой гема. Как и в глобинах, участок, связывающий гем, имеет достаточно жесткую структуру; стенки полости образованы двумя парами приблизительно антипараллельных спиральных сегментов, а дно представляет собой β-складчатый лист [297]. Структура гемсвязывающего центра может свидетельствовать об определенной свободе, с которой можно расположить образующие полость неполярные остатки. Предварительные данные, основанные на сопоставлении последовательности гемсвязывающего фрагмента b_2 с третичной структурой b_5 [557], показывают, что в семействе цитохрома *b*, как и в случае глобинов, сохраняется лишь общая картина неполярных контактов, но не отдельные боковые цепи.

Химические превращении атома железа гема в микросреде белка

В гемоглобине ион Fe^{2+} защищен от окисления неполярным окружением. Полипептидная цепь гемсодержащего белка создает для его простетических групп особую микросреду, которая способствует определенному химическому превращению и препятствует

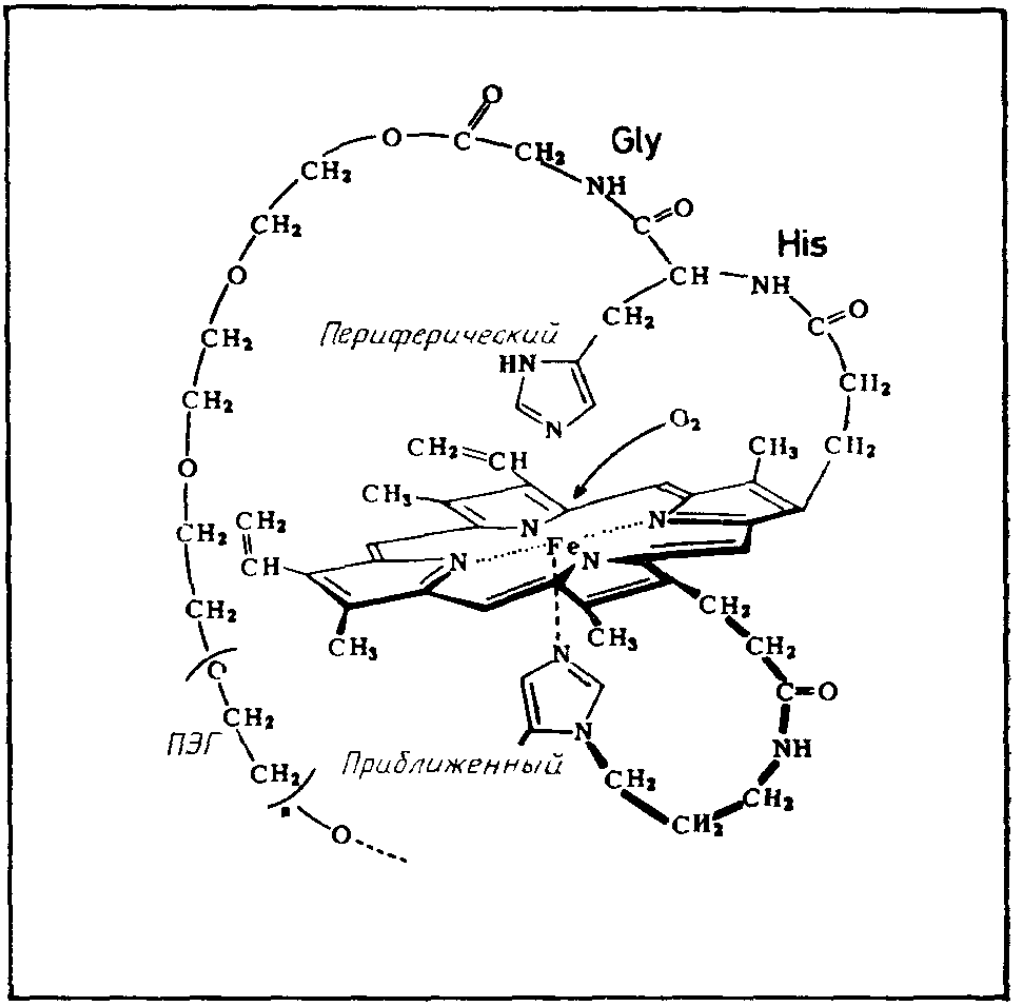


Рис. 10.2.

Химический аналог активного центра гемоглобина [727].

В центре схемы находится протогем IX. Одна пропионовая боковая цепь (вверху справа) образует пептидную связь с аналогом периферического His гемоглобина, который связан с полиэтиленгликолем (ПЭГ) через остаток Gly. Другая пропионовая группа (внизу справа) связана через разделительную цепь с имидазолом, который представляет аналог приближенного His. Химический переносчик O_2 качественно сходен с гемоглобином по характеру обратимого связывания молекулярного кислорода и по некоторым спектральным свойствам.

протеканию других реакций. Важную роль микросреды демонстрирует пример гемоглобинов. В воде феррогем связывает O_2, но при этом сразу же окисляется до ферригема, который не способен присоединять молекулярный кислород [637]. Однако гем может обратимо присоединять кислород и без окисления, если он заключен в неполярную среду, например бензол, где значительно труднее оттянуть электрон от ферроиона, чем в воде [639, 647]. Можно заключить, что основное назначение неполярного гемсвязывающего кармана — это предохранение ферросостояния гема от окисления путем его защиты от воды.

Химическая модель соединений, обладающих активностью гемоглобина. Синтез модели активного центра гемоглобинов основан на идее замены белка другим полимером [639, 727] (рис. 10.2) или присоединенными к гему гидрофобными боковыми цепями, образующими «частокол» [797]. Модельные системы такого типа могут быть

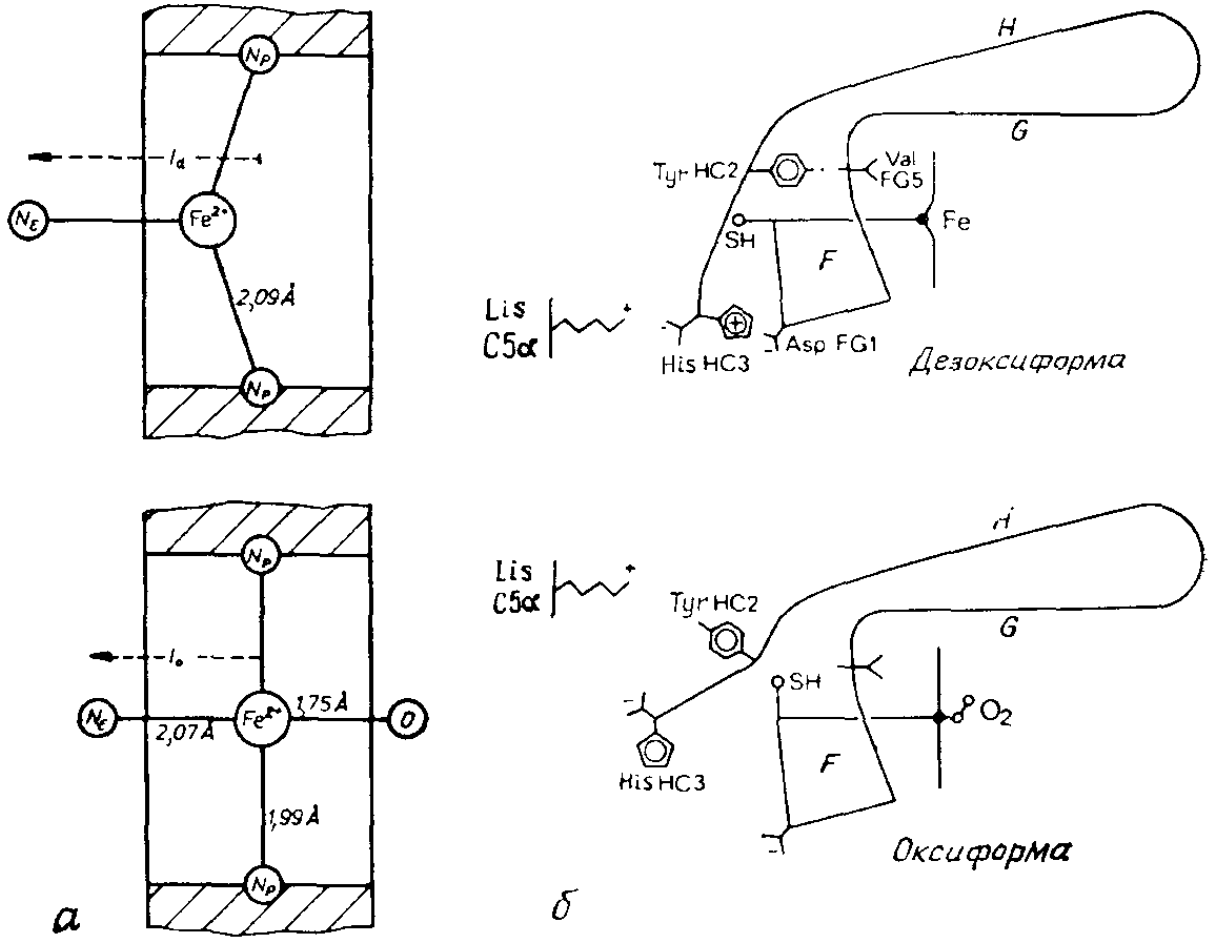


Рис. 10.3.

Изменение гемоглобина при окислении.

а — возможные положения иона железа относительно плоскости порфиринового цикла (заштриховаиа) в дезоксигемоглобине [804, 654] и в оксигемоглобине [550, 651]. Расстояиия центр — центр взяты нз работ, цитированных Филнпсом [690]. Следует отметить, что геометрия гема в оксигемоглобиие установлена по так называемым частокольным комплексам [651], т. е. по модельным соединениям. Указаны положения двух порфнриновых атомов азота (N_p) и атома N_ε приближенного His. При окислении атом N_ε смещается иа $I_d - I_0 = 0{,}6 \pm 0{,}1$ Å. *б* — коиформационные изменения в субъединице гемоглобина при окислеиии (Перутц, частное сообщение). Смещение спирали *F* к спирали *H* прнводит к вытеснению предпоследнего остатка β-субъединицы Tyr HC2 нз кармана между двумя спиралями. При этом радикально меняется положение карбоксильиой группы концевого остатка His HC3; Val FG5 — это Val-98.

полезными при выяснении взаимосвязи структура — функция природных макромолекул. Кроме того, они позволяют экстраполировать принципы структурной организации белков к условиям, которые несовместимы с функциональной целостностью самих белков (необычные значения температуры, pH и состава растворителя).

O_2 и CO присоединяются к гемоглобину как изогнутые лиганды. Этот способ присоединения, по-видимому, входит в механизм дискриминации CO. Сродство гемоглобинов к O_2 в два раза ниже, чем у простых железопорфириновых моделей, например гема, раст-

воренного в бензоле. По способности к связыванию CO обе системы различаются в 100 раз [648]. Следовательно, образованная белком микросреда дискриминирует CO, не изменяя заметно своей способности к связыванию O_2. Эта селективность основана отчасти на различии для CO и O_2 электронного влияния, связанного с *транс*-расположением второго аксиального лиганда атома железа [649] так называемого приближенного His (рис. 10.2).

Однако более важным представляется другой аспект. И в простой железопорфириновой системе [650, 651], и в гемоглобине [690] молекулярный кислород связан в виде изогнутого лиганда (рис. 10.3, *а*), а в такой форме он наилучшим образом может быть описан как пероксид O_2^- ([652, 653], ср., однако, [650]). Напротив, CO в простой системе связывается по линейному типу [648, 654], а во всех до сих пор изученных гемоглобинах — по изогнутому [655—657]. Таким образом, белковая среда, в которой находится гем, определяет способ связывания, который благоприятен для O_2 и неблагоприятен для CO. Необходимо упомянуть, что, несмотря на такую дискриминацию, сродство гемоглобинов к CO все-таки в 500 раз выше, чем к O_2 [634].

Эволюция связывания лиганда в направлении, неблагоприятном для CO, по-видимому, была обращена против эндогенного CO, который выделяется при разложении порфирина в пигменты желчи [85]. Если бы не такая предосторожность, образующийся CO мог бы занять примерно одну треть всех мест связывания гемоглобина [649]. Современная точка зрения на дискриминацию в отношении CO заключается в том, что появившиеся более чем 10^8 лет назад гемоглобины подверглись адаптации к условиям, которые в ином случае могли оказаться смертельными, — курению табака и загрязнению воздуха.

В цитохромах атом железа гема защищен от нежелательных лигандов. В цитохромах белковая микросреда выполняет две функции: защищает гем от нежелательных лигандов и поддерживает на нужном уровне окислительно-восстановительный потенциал атома железа. Было высказано предположение о защитной роли Met-80, который координирует атом железа гема в цитохроме *с* (рис. 7.8) [509]. В отсутствие такого контакта атом железа легко и неконтролируемо восстанавливается, например аскорбатом [658]. Можно утверждать, что полипептидная цепь гарантирует правильность транспорта электрона по его биологическому пути [509]. Необычна инертность к лиганду иона железа в цитохроме b_5 [297]; по-видимому, она связана с жесткостью аксиальных лигандов (разд. 10.3).

Окислительно-восстановительный потенциал атома железа гема в цитохроме *с* определяется аксиальными лигандами. Влияние апобелка на уровень окислительно-восстановительного потенциала исследовали Мур и Вильямс [659]. В этой работе были сопоставлены окислительно-восстановительные потенциалы соединений, моделирующих гем [660], а также использована структурная информация

о различных цитохромах и данные метода ядерного магнитного резонанса для отдаленно родственных цитохромов *c*. В результате авторы пришли к выводу, что окислительно-восстановительный потенциал атома железа гема определяют аксиальные лиганды. Так, две гистидиновые боковые цепи в качестве аксиальных лигандов, обнаруженные, например, в b_5, обеспечивают более низкий окислительно-восстановительный потенциал, чем пара гистидин — метионин, обычно присутствующая в цитохромах *c*. Если сравнивать цитохромы *c*, то обнаруживаются вариации длины связи Fe—S [659]. Укорочение этой связи на 0,1 Å связано с уменьшением оксилительно-восстановительного потенциала на 400 мэВ, что отражает увеличение электронодонорной способности атома серы метионина.

Лиганды гема как пусковые системы структурных изменений гемоглобина млекопитающих

Дезоксигемоглобин содержит высокоспиновый ион Fe^{2+}, а оксигемоглобин — низкоспиновый ион Fe^{2+}. До сих пор существуют разногласия по поводу конформационных изменений, происходящих в ходе окислительно-восстановительного цикла цитохрома *c* [459, 509, 564]. Поэтому мы ограничим обсуждение хорошо изученным случаем тетрамерного гемоглобина млекопитающих [60, 662, 666].

Кристаллографические исследования показали, что четвертичные структуры окси- и дезокси-форм гемоглобина существенно различаются. Эти различия указаны в табл. 10.1. В частности, при окислении четыре субъединицы смещаются относительно друг друга, в результате чего расстояние между атомами железа β-цепей уменьшается на 6,5 Å. Такое значительное изменение четвертичной структуры, вероятно, основано на чисто химических аспектах взаимодействий железопорфиринов [637, 663, 664]. В железопорфириновом комплексе возможны два типа распределения *d*-электронов. Высокоспиновая конфигурация, найденная в дезоксигемоглобине, характеризуется большим числом орбиталей с неспаренными электронами, чем низкоспиновая конфигурация оксигемоглобина. Одно из последствий состоит в том, что высокоспиновый атом железа имеет слишком большой размер для того, чтобы оставаться в плоскости порфиринового цикла, и выходит из нее на 0,5 Å [550] (рис. 10.3).

Механизм присоединения O_2 гемоглобином изучен на атомном уровне. Проследим теперь процесс насыщения дезоксигемоглобина кислородом [60, 550]. Молекулярный кислород подходит к высокоспиновому атому железа одной α-субъединицы (в β-субъединицах дезоксигемоглобина шестое координационное место блокировано остатком Val-67 и не доступно для O_2) и переводит его в низкоспиновую конфигурацию. В результате атом железа смещается на 0,5 Å и возвращается в плоскость порфиринового цикла. Имидазол (при-

Таблица 10.1

Различия между дезоксигемоглобином и оксигемоглобином [550, 666]

	Дезоксигемоглобин	Оксигемоглобин
Шестое координационное место Fe^{+}	Вакантно	Занято
Конфигурация *d*-электронов атома железа	Высокоспиновая	Низкоспиновая
Длина связи между Fe^{2+} и атомом азота пиррола	2,09 Å	2,01 Å
Смещение атома железа из плоскостн порфиринового цикла	0,75 Å	0,05 Å
Расстояние между ε-атомом азота координированного His и плоскостью порфирина	2,67±0,1 Å	2,07 Å
Положение предпоследней боковой группы тирозина	Между спиралью F и спиралью H	Удалены из этого положения (рис. 10.3,*б*)
Положение терминальных карбоксильных групп	Фиксировано электростатическим взаимодействием с другой субъединицей	Имеет свободу вращения
Расстояние между Fe^{2+} β_1-субъединицы и β_2-субъединицы	39,8 Å	33,4 Å
Положение дифосфоглицерата	Фиксирован между β_1- и β_2-субъединицами	Не связан
Карман лиганда в β-цепи	Блокирован боковой цепью Val-67; может быть заполнен кислородом только в том случае, если будет раскрыт за счет тепловой энергии	Содержит O_2

ближенного) гистидина, который координирует атом железа, передвигается на такое же расстояние или, возможно, даже на 0,7 Å, поскольку связь железо—аксиальный лиганд при переходе атома железа из высокоспинового в низкоспиновое ферросостояние укорачивается на величину от 0,1 до 0,2 Å.

Спираль F, к которой присоединяется приближенный His, так близко подходит к спирали H, что боковая цепь предпоследнего остатка Tyr уже не помещается между двумя спиралями и выталки-

вается, оттягивая С-концевой остаток, который удерживался солевым мостиком с другой α-цепью. Таким образом, нарушается главная связь, фиксирующая дезоксиформу (рис. 10.3, *б*). Присоединение O_2 к следующей субъединице оказывает такое же действие на связь между субъединицами и усугубляет изменение четвертичной структуры. Одно из последствий этого изменения заключается в оттягивании остатка Val-67 от лигандного кармана β-субъединиц, в результате чего их атомы железа также становятся доступны для O_2. Энергетические аспекты этого перехода обсуждались несколькими авторами [60, 268, 634].

Пример дезоксигемоглобина показывает, почему бывает трудно распознать химическое взаимодействие, которое дает толчок решающим изменениям четвертичной или третичной структур белка: «запускающее» событие должно быть структурно незаметным. В случае гемоглобина потребовались многие годы, прежде чем удалось установить, что пусковой механизм связан с атомом железа, который приводит в движение субъединицы гемоглобина [666].

В тетрамерном гемоглобине сродство гема к O_2 контролируется концентрациями O_2, CO_2, H^+ и 2,3-дифосфоглицерата. Что же было достигнуто в результате эволюции мономерного гемоглобина типа миоглобина в сложный гемоглобин млекопитающих? Основное преимущество заключается в возросшей физиологической приспособляемости тетрамерного белка, которая достигнута за счет постепенного перевода сродства к кислороду, свойственного центру связывания, под контроль внешних влияний [276, 549, 667] (рис. 10.4).

Одним из преимуществ является и улучшенный вариант эффекта Хальдана [668, 669], который заключается в освобождении протонов при присоединении молекулярного кислорода к гемоглобину, и наоборот. Этот эффект важен для транспорта молекулярного кислорода из легких к потребляющим кислород тканям, а также для транспорта протонов из этих тканей к легким. Кроме того, с помощью разных типов гемоглобина, которые имеют различную чувствительность к концентрации ионов водорода, некоторые типы рыб могут использовать O_2 для двух раздельно контролируемых функций: для наполнения воздушного пузыря и для перевода молекулярного кислорода в пути метаболизма [136]. Большое значение имеет также возможность использования некоторых специальных органических фосфорильных соединений, таких, как 2,3-дифосфоглицерат (ДФГ) или инозитолгексафосфат [670, 671], в качестве аллостерических эффекторов, благоприятствующих дезоксиконформации. Эти эффекторы играют решающую роль в переносе O_2 из материнской крови в кровь плода [274, 671], при адаптации к большой высоте [672] и при заболеваниях, при которых артериальная кровь лишь частично насыщается кислородом [85, 673] (рис. 10.4).

Кооперативность субъединиц облегчает сложную задачу переноса O_2. Еще одним важным фактором является так называемое

взаимодействие гем — гем, т. е. влияние связанного с O_2 гема на сродство к O_2 групп гема других субъединиц [634]. Этот эффект дает возможность гемоглобину захватывать O_2 в легких при концентрации свободного O_2 130 мкМ и освобождать приблизительно 30% O_2 в капиллярах тканей, потребляющих O_2 (рис. 10.4). Здесь концентра-

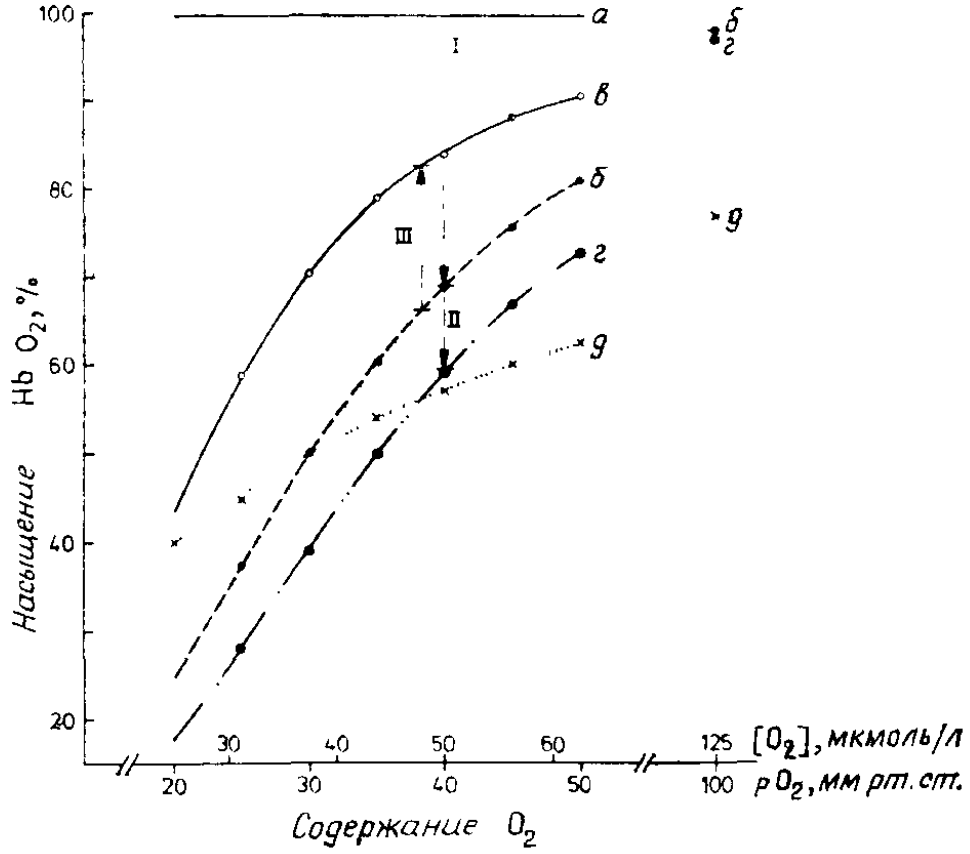


Рис. 10.4.

Кривые связывания O_2 гемоглобинами [85].

Дана зависимость насыщения гемоглобина (Hb) кислородом при pH 7,2 от концентрации свободного кислорода. Концентрации O_2 в капиллярах легких (125 мкМ) и в капиллярах тканей, потребляющих O_2 (50 мкМ), зафиксированы в узких пределах. Кривая *а*: в отсутствие дифосфоглицерата (ДФГ) гемоглобин насыщается O_2 в легких, но не может доставлять его к тканям. Кривая *б*: при физиологическом уровне ДФГ (4,5 мМ) приблизительно 30% O_2, поглощенного легкими, высвобождается в тканях (стрелка I). Кривые *б* и *в*: поскольку гемоглобин плода (кривая *в*) имеет более низкое сродство к ДФГ, чем материнский гемоглобин, освобожденный из материнской крови молекулярный кислород может захватываться гемоглобином плода (стрелка III). Кривая *г*: высокая концентрация ДФГ (8 мМ) приводит к повышенному снабжению тканей кислородом (стрелки I и II). Кривая *д*: при отсутствии кооперативности между субъединицами гемоглобина от легких к тканям транспортировалось бы меньше O_2. При построении гипотетической кривой связывания (*д*) для комплекса HbO_2 принята константа диссоциации 38 мкМ.

ция свободного O_2 должна оставаться на уровне 50 мкМ (чтобы обеспечить диффузию кислорода в клетки ткани). Доля освобождаемого кислорода (30%) не очень сильно зависит от уменьшения концентрации O_2 в легких и может быть увеличена за счет сокращения среднего диффузионного пути O_2 в тканях. Напротив, переносчик с независимыми центрами связывания O_2 освобождал бы при благоприятных условиях меньше 20% переносимого кислорода и не обладал бы возможностями к адаптации (рис. 10.4). Таким образом, кооперативность олигомера обеспечивает более высокую приспособленность к выполняемой функции.

Многие представления о действии и взаимодействии белков появились в ходе исследования гемоглобина. Многие представления и модели, относящиеся к взаимодействиям белок — лиганд и белок — белок, были развиты в процессе исследований гемоглобина; к ним относятся сигмоидное связывание [674—676], коэффициент Хилла [677], константы последовательного связывания лигандов в олигомерных белках [678], кооперативность, основанная на конформационных изменениях [679, 680], и аллостерический контроль белков [92, 681, 682]. Следует отметить, что многие из этих концепций были введены и математически формализованы до того, как стала известна структура какого-либо белка. Очевидно поэтому актуальное значение и полезность этих концепций должны подвергаться постоянной проверке. Пример дифосфоглицерата, влияние которого на действие и структуру гемоглобина игнорировалось десятилетиями, свидетельствует о потенциальной опасности жестких формулировок в биологии.

10.4 Центры связывания нуклеотидов

NAD-связывающие центры дегидрогеназ

Определены структуры четырех NAD-зависимых дегидрогеназ: лактатдегидрогеназы (LDH) [232], *s*-малатдегидрогеназы (MDH) 233], алькогольдегидрогеназы печени (ADH) и D-глицеральдегид-3-фосфатдегидрогеназы (GAPDH) [230, 231]. Эти ферменты катализируют перенос гидрид-иона из субстрата, по которому они названы, к атому C4 никотинамидного цикла NAD. Каждая субъединица фермента содержит один домен, связывающий NAD.

Все NAD-связывающие домены содержат параллельные β-структуры с идентичной топологией. Домены гомологичны [91], поскольку для них характерна одна и та же сложная топология β-складчатых листов (разд. 9.6). В частности, точно сохраняется расположение β-структур в области связывания NAD, как схематично показано на рис. 7.7. Это консервативное пространственное расположение выдержало много аминокислотных замен: не существует какой-либо заметной гомологии аминокислотных последовательностей двух любых из этих доменов.

NAD присоединяется в вытянутой конформации. Присоединение NAD к любой из четырех дегидрогеназ в этой характерной конформации можно рассматривать как отражение подобия NAD-связывающих доменов. Адениновый и никотинамидный циклы удалены друг от друга на 14 Å, а их плоскости почти перпендикулярны друг другу. Целесообразно рассматривать аденозиновую и никотинамидную части NAD отдельно. Одной из причин этого является возможность возникновения самого NAD-связывающего домена за счет дуплика-

ции меньшей структурной единицы — шпильки Россмана (разд. 9.4). Как схематически показано на рис. 7.7, каждый мононуклеотидный фрагмент NAD присоединяется к одной шпильке Россмана.

Карман, предназначенный для связывания аденина, не является высокоспецифичным. Аденозин NAD связывается в кармане, который не строго специфичен к этому лиганду, а к ароматическим остаткам вообще. Единственным инвариантным остатком аденозин-связывающих центров у всех четырех дегидрогеназ является последний остаток первого β-складчатого листа шпильки Россмана (рис. 5.12, *б*). Это — глицин; любой более крупный остаток помешал бы связыванию рибозильного фрагмента. Сохраняется также последний остаток второго β-складчатого листа. Это — аспарагиновая кислота, которая образует водородную связь с атомом O-2′ рибозы, имеющую решающее значение для способности ферментов различать NAD и NADP.

В дегидрогеназах LDH, ADH и MDH гидрид-ион связывает А-сторона никотинамида, а в GAPDH — В-сторона. Никотинамидный фрагмент NAD инвариантно связывается в полости, гидрофобной с одной стороны и гидрофильной с другой стороны, которая взаимодействует с субстратами, такими, как лактат, этанол, малат или глицеральдегидфосфат. Гидрид-ион стереоспецифически переносится к никотинамиду. В LDH, ADH и MDH так называемая А-сторона [85] никотинамидного кольца взаимодействует с субстратами и принимает гидрид-ион, тогда как в GAPDH никотинамидное кольцо повернуто на 180° вокруг гликозидной связи, связывающей рибозу и никотинамид, за счет чего В-сторона оказывается под воздействием субстрата и принимает гидрид-ион. Такой относительный поворот, очевидно, обусловлен в случае NAD пространственными требованиями центра связывания.

Нуклеотидсвязывающие центры других белков

Многие кофакторы содержат аденозиновые остатки. В NAD-зависимых дегидрогеназах субцентр связывания аденозинового остатка исключительно хорошо сохранился. Многие другие кофакторы белков, таких, как ATP, NADP, FAD, кофермент А, S-аденозилметионин, PAPS и дезоксиаденозилкобаламин, также содержат аденозиновые фрагменты. Поэтому важно определить, насколько обобщения, касающиеся связывания NAD дегидрогеназами, можно распространять на другие взаимодействия нуклеотид — белок.

Многие нуклеотидсвязывающие белки проявляют сходство с NAD-зависимыми дегидрогеназами. Все структурно изученные нуклеотидсвязывающие белки проявляют определенное сходство с дегидрогеназами. Фосфоглицераткиназа [235, 310, 311] содержит домен с такой же топологией β-структуры (разд. 9.6), что и NAD-связывающий домен, а аденозиновый фрагмент кофактора ATP при-

соединяется в положении, соотвечающем положению аденозинового остатка NAD в дегидрогеназах.

В аденилаткиназе центры связывания субстратов ATP и AMP [665] занимают положения, близкие к положению β-структуры NAD-дегидрогеназ; β-структуры имеют в этом случае также близкие топологии [255]. С точки зрения топологии положение ATP в аденилаткиназе в точности соответствует положению аденозинсодержащей части NAD в лактатдегидрогеназе. С другой стороны, расположение AMP в аденилаткиназе соответствует геометрически, но не топологически положению никотинамида в лактатдегидрогеназе. Топологическая неэквивалентность объясняется тем, что ATP и AMP образуют контакты под участком, соединяющим β-структуры A и B, а не над ним, как показано на рис. 7.7. Аденин ATP, по-видимому, связан водородной связью с фенольной OH-группой остатка Tyr-95 аденилаткиназы; изофункциональным остатком в LDH является Tyr-85. Очевидны аналогии между LDH и аденилаткиназой также в (пиро)фосфатсвязывающем субцентре. Кроме того, в обоих белках при присоединении нуклеотидов в соответствующих положениях наблюдаются большие конформационные изменения [665].

Во флаводоксине центр связывания простетической группы флавинмононуклеотида (FMN) эквивалентен центру никотинамидного фрагмента NAD в дегидрогеназах [229]. При совмещении FMN с никотинамидом укладка цепей флаводоксина и NAD-связывающего домена LDH почти полностью совпадает [91].

В дигидрофолатредуктазе и в глутатионредуктазе NADP присоединяется к доменам, содержащим β-структуры различной топологии. Некоторые другие белки, которые были сопоставлены с NAD-связывающим доменом, описаны в работах [91, 254, 683]. Среди них наибольший интерес представляют белки, специфичные к NADP. В той же степени, в какой NAD типичен для ферментов катаболизма, NADP характерен для ферментов анаболизма. Имеются два структурно изученных фермента с NADP в качестве кофактора: дигидрофолатредуктаза [308] и глютатионредуктаза [124].

В дигидрофолатредуктазе спирали и β-складчатые листы NADP-связывающего центра имеют некоторые общие черты с соответствующими элементами NAD-связывающего центра LDH. Однако топология и укладка цепей в этих двух ферментах совершенно различны. В глютатионредуктазе имеется NAD-связывающий домен с параллельной β-структурой в центре [124]. Эта β-структура содержит шпильку Россмана, связывающую аденозиновый остаток NADP в положении, отвечающем эквивалентному месту дегидрогеназ. В никотинамидном фрагменте NADP аналогичной шпильки Россмана нет. Глутатионредуктаза включает FAD в качестве простетической группы. Здесь также аденозиновый фрагмент FAD присоединен к структурному элементу, напоминающему складку Россмана.

Некоторые протеиназы разрушают апобелки нуклеотидсвязывающих ферментов. Задолго до того, как ученые выявили общие свойства нуклеотидсвязывающих белков [91], этими свойствами воспользовались другие существующие в природе агенты. К ним относятся, в частности, протеиназы, которые специфически разрушают апобелки нуклеотидсвязывающих ферментов [684]. Другой интересный пример представляет салицилат — соединение, которое встречается в некоторых растениях и которое образует действующее начало аспирина. Как в алгольдегидрогеназе [685], так и в аденилаткиназе [665] салицилат присоединяется в качестве (слабого) ингибитора к функциональным группам, которые обычно используются для связывания аденозиновых остатков соответственно NAD и ATP.

Индуцированное соответствие в аденилаткиназе

Индуцированное соответствие субстрата увеличивает специфичность. Многие ферменты, переносящие фосфорильные группы от ATP к акцепторным молекулам, используют механизм индуцированного соответствия для повышения своей специфичности, в частности для исключения H_2O как возможного акцептора фосфорильной группы. Типичным примером является аденилаткиназа (рис. 10.5), которая фосфорилирует H_2O в 10^5 раз медленнее, чем свой специфичный субстрат AMP.

Фермент, к которому применима модель индуцированного соответствия [686, 687], существует почти исключительно в неактивном состоянии E. Только небольшая доля молекул имеет активную конформацию E'. Согласно предположению Дженкса [631], отношение $[E]/[E']$ можно определить непосредственно по скоростям фосфорилирования H_2O и специфичного субстрата. Для аденилаткиназы величина $[E]/[E']$ в отсутствие субстрата должна быть равна 10^5. Присоединение специфичного субстрата вызывает изменение конформации активного центра, переводя таким образом фермент в активную форму E' (рис. 10.5).

Большая часть связывающей энергии расходуется на изменение конформации фермента. Если предположение Дженкса верно, то наблюдаемая константа связывания $K_{\text{набл}} = [E' - \text{AMP}]/[E] \times [\text{AMP}] = 10^4\ M^{-1}$ в 10^5 раз меньше, чем «общая» константа связывания $K_{\text{общ}} = [E' - \text{AMP}]/[E] \cdot [\text{AMP}]$. Пользуясь соотношением $\Delta G = -RT\ln K$, мы находим, что общая свободная энергия связывания, т. е. полученная в результате взаимодействия AMP с ферментом E' ($\Delta G_{\text{общ}} = -12{,}5$ ккал/моль), более чем в два раза превышает свободную энергию связывания, следующую из наблюдаемой величины константы диссоциации ($\Delta G_{\text{набл}} = -5{,}3$ ккал/моль). Таким образом, большая часть энергии связывания расходуется на обеспечение изменения конформации фермента от неактивной к активной форме, а оставшаяся часть проявляется в виде наблюдаемой энергии связывания. Общая энергия связывания была бы эквивалент-

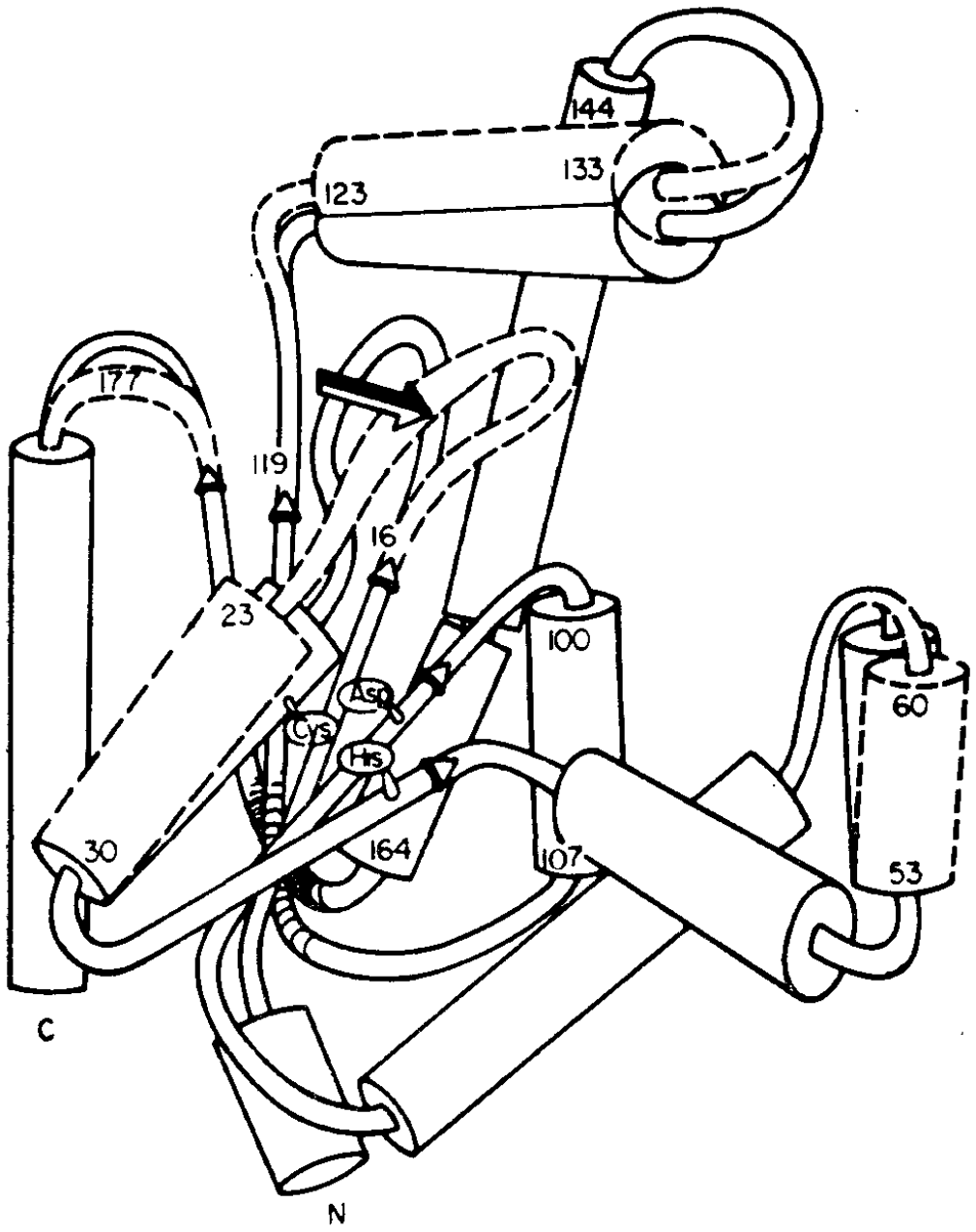


Рис. 10.5.

Смещение фосфорилсвязывающей петли (остатки 16-22) при переходе из конформации А (показана сплошными линиями) в конформацию В (пунктирные линии) кристаллической аденилаткиназы [665, 688].

В конформации В в центре присоединиения АТР имеются аденозинновый карман, широкая «расщелина» для принятия фосфорильной группы, а также открытый центр присоединения АМР. В конформацин А место присоединения также открыто, однако фосфорильная «расщелинна» узка, а аденозиновый карман центра присоединения АМР закрыт. Эти данные показывают, что конформацня В относится к свободному ферменту до индуцированного АМР соответствия, тогда как конформация А относится к ферменту Е′ после такого измеиения.

на наблюдаемой энергии связывания в том случае, если бы фермент был заморожен в активной форме E'. К сожалению, это можно осуществить только в кристаллической аденилаткиназе при условиях, в которых константы связывания определить нельзя [665, 688].

10.5 Центры связывания фосфорильных групп

В белках нет какого-либо единого и универсального фосфорилсвязывающего центра. Приблизительно 50% изученных белков

связываются или каким-либо другим образом взаимодействуют с соединениями, несущими фосфорильные группы. Контрастирующие примеры рибонуклеазы [689], гемоглобина человека (который присоединяет 2,3-дифосфоглицерат, ATP или инозитгексафосфат в положении между двумя β-цепями [672]) и стафилоккоковой нуклеазы [242] показывают, что в белках не существует единственного и универсального фосфорилсвязывающего центра.

Петля между первым листом β-структуры в шпильке Россмана и последующей α-спиралью является одним характеристичным фосфорилсвязывающим центром. Имеется группа белков, в которых фосфорильный или пирофосфорильный фрагмент присоединяется в характеристическом центре. Этим центром является петля полипептидной цепи, которая связывает первый* складчатый лист параллельной β-структуры и спираль, расположенную приблизительно антипараллельно этому листу. Обычно эта петля воспроизводит петлю между первым β-листом шпильки Россмана и соседней α-спиралью (рис. 5.12, *б*).

Фосфорильные группы фиксированы относительно остова петли водородными связями. В S-малатдегидрогеназе [691], D-глицеральдегид-3-фосфатдегидрогеназе [230] и лактатдегидрогеназе [230] пирофосфатный фрагмент NAD связан с основной цепью соответствующей петли водородными связями. В случае алкогольдегидрогеназы [692] исследование связывания на примере аналога NAD показало, что пирофосфат находится в том же положении (в пределах 3 Å) относительно этой петли. Во флаводоксине петля «обернута» вокруг фосфорильного фрагмента FMN [145, 237, 238]. Как следует из исследований связывания субстрата, в аденилаткиназе основная цепь петли обернута вокруг фосфата AMP, в кристаллической аденилаткиназе эта петля фиксирует сульфат из маточного раствора [665]. В аналогичном положении связывается ATP в фосфоглицераткиназе [310, 311]. В глутатионредуктазе [124] пирофосфорил NADP присоединяется к остову соответствующей петли шпильки Россмана. То же самое относится и к пирофосфату FAD. В триозофосфатизомеразе [305] фосфат также расположен у петель, связывающих β-листы и последующие α-спирали, но в С-концевой части полипептидной цепи (рис. 5.17, *д*). Во всех этих случаях используется петля между карбонильным концом β-листа и следующей α-спиралью. Это можно объяснить выгодным электростатическим взаимодействием между отрицательным зарядом фосфорильной группы и диполем α-спирали [792], который возникает при наложении диполей водородных связей.

Фосфорилсвязывающие петли могут быть гибкими и подвижными. Из рассмотренных примеров следуют и некоторые другие

* Следует напомнить, что листы β-структуры нумеруют по их расположению вдоль цепи линейного полипептида; остатки, принадлежащие первому листу, находятся ближе к N-концу.

обобщения. Во всех случаях фосфорильные группы образуют водородные связи с остовом полипептидной цепи. Связывание фосфорила осуществляется петлями, выступающими из жесткой вторичной или сверхвторичной структуры, например, шпильки Россмана. Поскольку эти петли имеют конформационно жесткие основания, сами они могут проявлять подвижность, не нарушая структуру белка в целом. В аденилаткиназе свиньи, например, фосфатсвязывающая петля имеет последовательность Gly -Pro-Gly-Ser-Gly-Lys-Gly [389], наличие глицильных остатков в каждом втором положении обеспечивает большой интервал двугранных углов при C_α-атомах (разд. 2.3). Следовательно, этот сегмент цепи может приобретать широкий спектр конформаций, что и приводит к гибкости и подвижности даже в невзаимодействующей молекуле фермента. Рентгеноструктурный анализ (рис. 10.5) показал, что при переходе аденилаткиназы из конформации A в конформацию B эта петля претерпевает смещение на 6 Å.

Сходство внутри рассмотренной группы фосфорилсвязывающих белков может и не иметь общей природы. Что касается дегидрогеназ, описанных в разд. 10.4, то способ связывания (пиро)фосфорильных групп, вероятно, относится к их консервативным особенностям, поскольку эти белки, по-видимому, гомологичны (разд. 9.6). Однако для других белков этой группы способ связывания фосфорила нельзя рассматривать как черту, указывающую на гомологию, поскольку он может быть просто следствием выгодного физико-химического взаимодействия.

Некоторые белки, модифицируемые ковалентным фосфорилированием, имеют типичные свойства, общие с нуклеотидсвязывающими белками. Имеется большое число белков, функции которых модифицируются или регулируются с помощью обратного фосфорилирования [175]. Аминокислотные последовательности в центрах фосфорилирования различных белков проявляют некоторое сходство (табл. 10.2). Обычно фосфорилируется боковая цепь остатков Ser, Thr или His. Часто перед таким остатком или сразу за ним располагается остаток Gly; в большинстве случаев остаток, находящийся в цепи на две позиции дальше, оказывается положительно заряженным Lys или Arg. Эти последовательности имеют черты некоторого сходства с фосфорилсвязывающей петлей аденилаткиназы (табл. 10.2). Гипотеза о том, что обратимое фосфорилирование данного остатка управляет конформационными изменениями фосфопротеида [175, 693], хорошо коррелирует с тем фактом, что в лактатдегидрогеназе [232] и в аденилаткиназе [688] петля является эпицентром больших конформационных изменений. Таким образом, исследования нуклеотидсвязывающих белков могут способствовать прояснению статических и динамических аспектов фосфорилирования белков.

Таблица 10.2

Центры присоединения фосфата в нуклеотидсвязывающих белках и в фосфопротеидах

Белок	Последовательность	Примечание
Миелиновый основной белок	G-Ser(P)-G-K-D	Фосфопротеиды; (P) означает обратимо фосфорилируемый остаток; данные из работы [693]
Гистон H2a	Ac-Ser(P)-G-R-G	
Гистон H4	Ac-Ser(P)-G-R-G	
Гистон HI	G-Ser(P)-F-K-L	
Гликогенсинтетаза	I-Ser (P)-V-R-X	
Тропонин I	I-Thr(P)-A-R-R	
HPr-фактор [720]	I-His(P)-A-R-P	HPr участвует в направленном фосфорилировании. Типичная топология петли, связывающей фосфат, предсказана по аминокислотной последовательности
Флаводоксин [145]	10 G-Thr-G-N-T	Эта петля охватывает фосфатный фрагмент FMN
Аденилаткиназа [665]	20 G-Ser-G-K-G	Петля охватывает P-фрагмент AMP: в ней происходят основные конформационные изменения
Глицеральдегид-3-фосфатдегидрогеназа [145[	R-Ile-G-R-L	Амиды Arg-10 и Ile-11 остова образуют водородные связи с пирофосфатным фрагментом NADH

10.6 Взаимодействия белков с другими макромолекулами

Белки могут специфично взаимодействовать с другими макромолекулами, например с нуклеиновыми кислотами и полисахаридами. К макромолекулам относят также липиды, поскольку они образуют в водных растворах крупные агрегаты. В нуклеопротеидах, гликопротеидах или липопротеидах белок может составлять менее 50%, и суммарные свойства комплексов часто определяются небелковыми фрагментами. Более того, и образование, и стабильность структуры белков могут зависеть от их партнеров по комплексам. Это наиболее очевидно для тех мембранных белков, которые соединяют различные углеводородные фрагменты липидного бислоя.

Липопротеиды

Белки проникают сквозь поверхность раздела водной и неполярной фаз. На взаимодействиях белков и липидов основаны две важ-

ные биологические функции: транспорт липидов в водных растворах и мембранная активность. Способность к транспорту в плазме крови обычно труднорастворимых липидов, например холестерола (2 мг растворимы в 1 л 150 мМ раствора NaCl, pH 7,4), или триацилглицеридов увеличивается в 1000 раз за счет образования специфических липопротеидов. Такие липидтранспортирующие белки могут играть роль в атеросклерозе и других заболеваниях, с чем и связано большое число посвященных им работ [145, 694]. Немалый интерес проявляется также к вторичной и третичной структурам липопротеидов. Предпринимались попытки предсказания их структуры на основе корреляции (гл. 6) между вторичной структурой и аминокислотной последовательностью обычных глобулярных белков [694]. Поскольку закристаллизовать липопротеиды достаточно трудно, по-видимому, более точных структурных данных в ближайшем времени получить не удастся.

В мембранах липиды образуют биологические барьеры и оболочки, тогда как специфические функции мембран, такие, как, например, транспорт, передача сигнала и преобразование энергии, выполняются белками [19, 695—697]. Информация об аминокислотных последовательностях во внутримембранных частях белков крайне ограниченна; известно, что в них имеются довольно протяженные участки неполярных остатков [698]. Наиболее подробная информация о третичной структуре получена для мембранного белка из *Halobacterium halobium* [699, 700]. Субъединица этого белка состоит в основном из семи параллельных или антипараллельных α-спиралей, вытянутых от одной поверхности мембраны до другой. Другая хорошо исследованная система обсуждается ниже.

Кальциевый насос — типичный хорошо исследованный мембранный белок. Саркоплазматический ретикулум из мышц [701, 702] представляет собой трубчатую систему с высокоспецифической мембраной, единственная функция которой состоит в освобождении и накоплении ионов кальция [703, 704]. Это отражается в том обстоятельстве, что один белок с молекулярной массой 100 000 так называемый Ca^{2+}-транспортирующая АТРаза или Ca^{2+}-насос, составляет более 50% массы мембраны и 80% общего содержания белков в мембране. Этот белок, представляющий собой цилиндр диаметром 50 Å и длиной 80 Å, проходит через весь слой мембраны толщиной от 60 до 90 Å. Мембрана включает 90 (фосфо)липидных молекул на одну молекулу Ca^{2+}-насоса.

Очевидно, мембрана и Ca^{2+}-насос образуют функциональную единицу, где мембрана служит барьером, препятствующим возврату накопленного кальция. Кроме того, мембрана является двумерным растворителем белка: как показано в опытах по связыванию липида с очищенным белком, поверхность транспортирующей Ca^{2+} АТРазы имеет большее сродство к неполярной среде, чем к воде.

Поскольку существует только один тип воды, но много типов молекул липидов, этот широкий спектр липидных молекул обеспечивает большое разнообразие специфических взаимодействий с мембранными белками. Роль индивидуальных классов липидов для данного мембранного процесса выясняют с помощью модельных экспериментов [705, 706]. В случае кальциевого насоса функционально активную мембранную систему можно воссоздать путем добавления к очищенному белку только фосфолипидов [701, 702].

Белки перемещаются в плоскости мембраны. Объемное антитело, присоединенное к наружной части Ca^{2+}-транспортирующей АТРазы, не препятствует перемещению ионов кальция, что указывает на то, что для транслокации Ca^{2+} не требуется поворота белка вокруг оси, параллельной поверхности мембраны [707]. По-видимому, это общее правило для мембранных белков. С другой стороны, белки способны поворачиваться [708] и перемещаться в боковых направлениях в пределах плоскости мембран; реальная степень подвижности зависит от физико-химических свойств мембраны и от направляющего действия белковых контактов с обеими поверхностями мембраны. Диффузия по горизонтали важна для взаимодействий между элементами многокомпонентной, связанной с мембраной системы, поскольку функционально связанные мембранные белки не всегда находятся в физическом контакте друг с другом [709, 710].

Гликопротеиды

Ковалентно связанная углеводная часть определяет величину, вязкость и местонахождение многих гликопротеидов. Наиболее распространенную группу биологически важных макромолекул представляют гликопротеиды, которые состоят из углеводных фрагментов, ковалентно связанных с полипептидами через боковые цепи остатков Thr, Ser или Asn [711, 712]. В единственном гликопротеиде с известной структурой белке IgG Kol миеломы [543] углеводный фрагмент имеет фиксированное расположение и выполняет, по-видимому, чисто структурную функцию. Вероятно, это относится к большинству гликопротеидов.

Углеводная часть может составлять от не менее одного до более восьмидесяти процентов общего веса гликопротеида. Поскольку углеводные фрагменты, присоединенные к определенному месту белка, могут варьировать, гликопротеид редко относится к гомогенному типу молекул. Первый пример из табл. 10.3 показывает, что вязкость и гидрофильные свойства углеводных фрагментов важны для функции многих гликопротеидов. В других случаях (типичными примерами являются рибонуклеазы поджелудочной железы различных млекопитающих [145]) наличие углеводного фрагмента не означает каких-либо преимуществ или недостатков, а связано лишь с тем фактом, что неспецифический присоединяющий сахар фермент распознает

Таблица 10.3

Влияние углеводных фрагментов на свойства и действие гликопротеидов[1]

Свойство или действие гликопротеида	Пример	Биологический аспект
Величиина	Белки плазмы крови [145, 721]	Предохранение от потери белков в почках
Вязкость	Муцины (гликопротеиды слизистых [722])	Муцины используются в качестве защнтных оболочек, смазкн и среды для транспорта
Антифризное действие	Антифризиые гликопротеиды антарктических видов рыб [723, 724]	Препятствуют росту, но не образованию кристаллов льда в жидких средах организма. Гликопротеиды понижают точку замерзания водных растворов в той же степени, что и эквивалентное количество NaCl
Ориентация в мембранах	Гликопротеиды клеточных мембран млекопитающих; углеводные фрагменты расположены на внешней поверхностн [709]	Поскольку существует высокоэнергетический барьер для прохождения углеводного фрагмента через гидрофобную часть мембраны, гликопротеиды помогают установлению и поддержанию асимметричного распределения мембранных белков
Взаимодействие со специфическими поверхностями	Протромбин [725] и другие белки плазмы крови млекопитающих [173]	Время жизни белков плазмы регулируется следующим образом. Потеря концевого невосстановленного сахара, связанного с белком углеводной боковой цепью, приводит к обнажению концевой D-галактозы. Этот остаток распознается рецептором на поверхности клеток печени, затем присоединяется весь гликопротеид и разлагается белками клетки
Действие как мембраисвязанного рецептора	Лектиновые рецепторы в клеточных мембранах, например рецептор конканавалина А [726]	У лектина (например, конканавалина А) имеется центр для присоединения специфического сахара (α-метилманиозида) рецептора (гликопротеида). Присоединение лектина является пусковым механизмом деления клетки. Лектины могут дать важные сведения для выяснения этого процесса. Их истинная функция пока еще не известна.

[1] Первые четыре примера представляют сбщие свойства углеводсв. Последние два примера иллюстрируют важность отдельных сахарных остатксв для специфическогс распозиования.

структурную особенность белка как субстрата [712]. Искусственные углевод-белковые образования более стабильны к нагреванию, денатурации и действию протеаз, чем соответствующие белки. По катаболизму и другим биологическим особенностям такие соединения отличаются от белков и могут рассматриваться как потенциальные новые терапевтические средства [796].

Специфичные сахарные остатки выполняют функции узнавания. Последние два примера табл. 10.3 показывают, что сахара выполняют важную роль в специфических взаимодействиях между поверхностями клеток и растворимыми макромолекулами. Межклеточное распознавание, например, при образовании тканей из различных типов клеток также основано на структурном разнообразии гликопротеидов [709, 713]. Сахара действительно являются подходящими элементами образования некоторых специфических структур [85]. Если из трех различных аминокислот можно составить только шесть различных пептидов (используя все перестановки), то из трех сахарных остатков можно образовать по меньшей мере в десять раз больше первичных структур; в связи с этим многие из возможных объединений моносахаридов используются in vivo. Однако механизмы узнавания с участием сахарных остатков часто основываются скорее на стохастических, чем на стехиометрических процессах, поскольку синтезу сложных углеводов недостает точности белкового синтеза.

Нуклеопротеиды

Этим термином мы обозначаем все белки, которые взаимодействуют не только с моно- или динуклеотидами, как, например, ферменты, рассмотренные в разделе 10.4, но и с целыми нуклеиновыми кислотами.

Рибосомы — комплексы белков и нуклеиновых кислот. Хорошо известным примером комплексов белков с нуклеиновыми кислотами является рибосома, катализирующая образование полипептидных цепей [714, 715]. Рибосома содержит несколько молекул РНК и различные белковые молекулы. Она состоит из двух субъединиц неодинаковой величины, между которыми заключена рибосомальная РНК. В *E. coli* более крупная субъединица содержит две молекулы РНК (названные «23S» и «5S» в соответствии с их константами седиментации) и 34 молекулы белков, названные от «L1» до «L34». Более мелкая субъединица содержит РНК «16S» и 21 белок (от «S1» до «S21»).

Многие из этих белков прочно связаны с нуклеиновыми кислотами. Некоторые белки настолько вытянуты [7], что число их внешних контактов с нуклеиновыми кислотами и другими рибосомными белками намного превосходит число внутримолекулярных контактов. Поэтому сама по себе полипептидная цепь не может принимать определенной структуры. Очень маловероятно поэтому, чтобы эти белки

можно было закристаллизовать и подвергнуть структурному исследованию. До сих пор удалось закристаллизовать только фрагменты белка L12 (и/или белка L17), который легко отделяется от рибосомы и, по-видимому, способен сам по себе образовывать определенную структуру. Ведется его рентгеноструктурное исследование [716].

Фактор элонгации Tu · GDP по форме напоминает головастика. Удалось также закристаллизовать фактор элонгации Tu · GDP, который образует с рибосомой и тРНК тройной комплекс [717]. Его структура определена с разрешением 6 Å [718]. Форма молекулы заметно отличается от формы обычных глобулярных белков. Это — «головастик», состоящий из глобулярного домена («головы») и удлиненного тонкого домена («хвоста»), разделенных расщелиной с меньшей плотностью. Голова содержит несколько α-спиралей и, по-видимому, имеет довольно жесткую структуру. Хвост, напротив, более лабилен и, по-видимому, представляет собой β-структуру. Интересно, что голова и хвост имеют и второе сочленение, так что в результате возникает некая циклическая структура. Вероятно, тРНК присоединяется вблизи отверстия этого цикла в большом желобе между доменами. Хорошо видно относительное смещение доменов во время стадии элонгации на рибосоме. Деление на подвижный и жесткий домены обнаружено также для нуклеопротеидов L7/L12 [716] и lac-репрессора [719]. В L7/L12 подвижный домен находится со стороны N-конца.

Аминоацил-тРНК-синтетазы также участвуют в синтезе белка. Однако эти ферменты взаимодействуют не с рибосомой, а только с тРНК. Известно свертывание цепи для одного из этих ферментов [221]. Оказалось, что фермент представляет собой симметричный димер и не отличается существенно от других глобулярных белков. Предполагалось, что центр связывания тРНК образован обеими субъединицами, что напоминает ситуацию, наблюдаемую в других ферментах, связывающих более мелкие субстраты [124, 313].

Большой интерес у молекулярных биологов вызывают белки, контролирующие экспрессию генов путем ассоциации с ДНК. Один из этих белков — lac-репрессор — удалось закристаллизовать; сейчас продолжается исследование его пространственной структуры [719].

Таким образом, наши представления о взаимодействиях между белками и нуклеиновыми кислотами пока еще очень ограниченны. Как было отмечено, важную роль в этих взаимодействиях могут играть β-структуры [206—208]. Примеры рибосомальных белков и фактора элонгации Tu · GDP показывают, что некоторые нуклеопротеиды имеют весьма своеобразную структуру.

Заключение

Нормальная деятельность живых организмов возможна только благодаря тому, что природный белок не проявляет присущей поли-

пептидам способности присоединять все виды малых молекул, но специализируется только на нескольких определенных видах. Лишь один тип белков, иммуноглобулины, имеет максимально возможное число центров присоединения. Поэтому систематический анализ систем иммуноглобулин — лиганд позволяет выяснить общие принципы взаимодействий белок — лиганд. Архитектура иммуноглобулинов может также служить основой для синтеза in vitro полипептидов с желаемыми свойствами.

Живые организмы включают огромное число разнообразных низкомолекулярных соединений, однако доминирующую роль имеет ограниченная группа метаболитов, включающая некоторые аминокислоты — производные аденозина и порфиринов; большинство белков специфически взаимодействует с одним или несколькими из таких метаболитов. Аденозиннуклеотиды и другие фосфорилсодержащие соединения, как правило, присоединяются к определенному месту складки Россмана.

Исследования гемсодержащих белков показывают, каким образом используются для биологических функций свойства полифункциональных химических соединений, таких, как железопорфириновые комплексы. Белковая часть молекулы обеспечивает протекание в каждом случае единственной специфичной реакции на атоме железа гема и модификацию присущих гему химических свойств под действием физиологически важных факторов.

Для биологических процессов наиболее важны белки, которые образуют комплексы с двойными липидными слоями, сахаридами и нуклеиновыми кислотами. Некоторую информацию об этих белках можно получить из уже исследованных взаимодействий белок — лиганд. Для расширения наших представлений в этой области необходимы данные о трехмерных структурах таких комплексов.

В заключение отметим, что для исследования взаимодействий белков с любыми другими соединениями наиболее подходящими объектами являются иммуноглобулины. Что же касается взаимодействий между белками и группой наиболее важных метаболитов, то здесь уже могут быть сделаны структурные предсказания, относящиеся к большинству белков.

Глава 11 Структурные основы механизма, действия и функции белков

11.1 Определения

Очень многие, а может быть, и все виды деятельности организма основаны на деятельности белков. Общая реакция всего организма на условия гомеостаза и на условия окружающей среды в конечном счете определяется изменениями на атомном или даже субатомном уровне [728]. В этой главе сделана попытка суммировать аспекты строения белков, которые можно отнести к категориям механизма, действия или активности и функции.

Механизм описывает последовательность во времени и пространстве основных процессов, составляющих определенное действие или реакцию. Для того чтобы сформулировать механизм действия белка, необходимо знать детали пространственного расположения атомов и направленность молекулярных орбиталей. В настоящее время это возможно только в исключительных случаях, классическими примерами которых могут служить гемоглобин (разд. 10.3) и химотрипсин (разд. 11.2). Вопрос, относящийся к механизму: *как работает белок?*

Понятие *действия*, или *активности*, относится к наиболее очевидным проявлениям свойств белка. Например, действие химотрипсина состоит в расщеплении пептида или полипептида. Вопрос, относящийся к действию: *что делает белок?*

Биологическая функция — это вклад некоего составляющего элемента в действие всей системы. Это означает, что функцию белка следует изучать в связи с более высокими уровнями функциональной иерархии [729]. В случае гемоглобина, например, такими более высокими уровнями являются циркуляция и метаболизм. Химотрипсин выполняет определенную функцию в системе пищеварения, перерабатывая белки пищи, а также функцию в системе защиты организма, инактивируя вредные полипептиды (некоторые гормоны, токсины и вирусные белки) до того, как они смогут атаковать эпителиальный барьер. Вопрос, относящийся к функции: *каково назначение белка?*

В последующих разделах на примере ферментов обсуждаются известные нам механизмы действия белков. В конце главы мы рассмотрим деятельность одного из органов — скелетной мышцы на уровне действия и функции белков.

11.2 Ферментативный катализ

Ферменты в миллиарды раз эффективнее искусственных катализаторов, хотя они работают в мягких биологических условиях. Необыкновенным свойством ферментов является эффективность их каталитического действия. Катализируемые ферментами реакции протекают со скоростями, в 10^8—10^{20} раз большими, чем соответствующие некатализируемые реакции. Ферменты традиционно сравнивают с искусственными катализаторами, которые, как правило, в 10^8—10^9 раз менее эффективны для ускорения данной реакции, чем соответствующий фермент [730—732]. Кроме того, в отличие от экстремальных условий, часто необходимых для ускорения химических реакций в органической лаборатории, каталитическое действие ферментов достигается в водном растворе при биологических значениях pH и умеренных температурах и давлениях.

Специфичность действия ферментов предохраняет организм от засорения побочными продуктами. Другая особенность ферментов, которой химические катализаторы обладают лишь в редких случаях, состоит в специфичности действия. Воздействию подвергается только одно или несколько соединений (субстратов), с которыми протекает только один тип реакций. Отсутствие побочных реакций или побочных продуктов—это отражение того факта, что неконтролируемое загрязнение живой клетки недопустимо. Отсюда следует вывод, что использования в промышленных целях заслуживает не только каталитическая эффективность, но также (а возможно, и в большей мере) специфичность ферментов.

Свободная энергия связывания субстрата расходуется на катализ и на обеспечение специфичности. Начиная с химотрипсина, возникла традиция раздельного рассмотрения специфических связывающих взаимодействий между субстратом и ферментом (разд. 10.2) и собственно каталитического процесса, который описывается в терминах химических «модельных» реакций (см. ниже).

Как было отмечено Дженксом [631], используя концепцию «утилизации», можно согласовать оба основных аспекта действия фермента. Невалентное взаимодействие специфичного субстрата с активным центром фермента количественно выражается свободной энергией связывания («связывающей энергией»). Так называемая собственная свободная энергия для большинства пар фермент — субстрат имеет порядок 10 ÷ 20 ккал/моль. Однако бóльшая часть этой энергии расходуется на каталитический процесс; катализ осуществляется за счет связывающих сил, которые могут быть удалены от каталитического центра. Наблюдаемая связывающая энергия пар фермент — субстрат представляет собой лишь ту часть, которая остается после такой утилизации; она составляет приблизительно от 5 до 7 ккал/моль. Связывающие силы могут также расходоваться на обес-

печение механизма индуцированного соотьетствия (разд. 10.4) или на непродуктивное связывание (разд. 10.2); оба этих механизма увеличивают специфичность действия фермента.

Каталитический механизм химотрипсина

Каталитический механизм химотрипсина — фермента, расщепляющего пептидные связи, изучен более подробно, чем механизм любого другого фермента [537]. Такие исследования упрощаются благодаря некоторым особенностям химотрипсина. Это мономерный фермент, не проявляющий аллостерических эффектов; структурные изменения, сопровождающие процесс расщепления пептидной связи, очень малы и, наконец, химотрипсин обладает способностью переносить ацильные группы самых разнообразных доноров, например пептидов и эфиров, к самым разнообразным акцепторам, например к воде, спиртам или аминам. Возможность сопоставлений процессов для различных доноров и для различных акцепторов значительно облегчают анализ отдельных каталитических стадий [733, 734].

На основе рентгеноструктурных, спектроскопических и химических данных (см. обзор в [537]) предложен возможный механизм действия химотрипсина (рис. 11.1). После того как фермент и субстрат образовали комплекс Михаэлиса (разд. 10.2), атом кислорода гидроксильной группы остатка Ser-195 атакует карбонильный атом углерода расщепляемой связи субстрата. Образуется неустойчивый тетраэдрический промежуточный продукт [735]. Эта реакция облегчается системой передачи заряда [628, 736], которая осуществляет передачу протона от гидроксильной группы остатка Ser-195, превращая его тем самым в сильный нуклеофил. Затем His-57 подает протон атому азота расщепляемой пептидной связи, в результате чего связь расщепляется. На этой стадии аминная часть образует водородную связь с остатком His-57, тогда как ацильная группа присоединяется к Ser-195 с образованием эфирной связи. На этом завершается стадия ацилирования гидролитической реакции.

Вторая половина процесса — это деацилирование. Аминная часть субстрата удаляется и ее место в активном центре занимает молекула воды. В принципе деацилирование есть процесс, обратный ацилированию, при котором H_2O замещает аминную компоненту [85]. Вначале зарядно-релейная система отрывает протон от воды, образующийся ион OH^- одновременно атакует карбонильный атом ацильной группы, присоединенной к Ser-195. Образуется неустойчивый тетраэдрический аддукт, в котором His-57 подает протон атому кислорода остатка Ser-195, что приводит к освобождению кислотного компонента субстрата, который удаляется диффузией. Теперь фермент готов для следующего каталитического цикла.

Химотрипсин, глицеральдегид-3-фосфатдегидрогеназа и папаин используют аналогичные механизмы реакции. Описанный выше механизм характерен не только для филогенетических родственни-

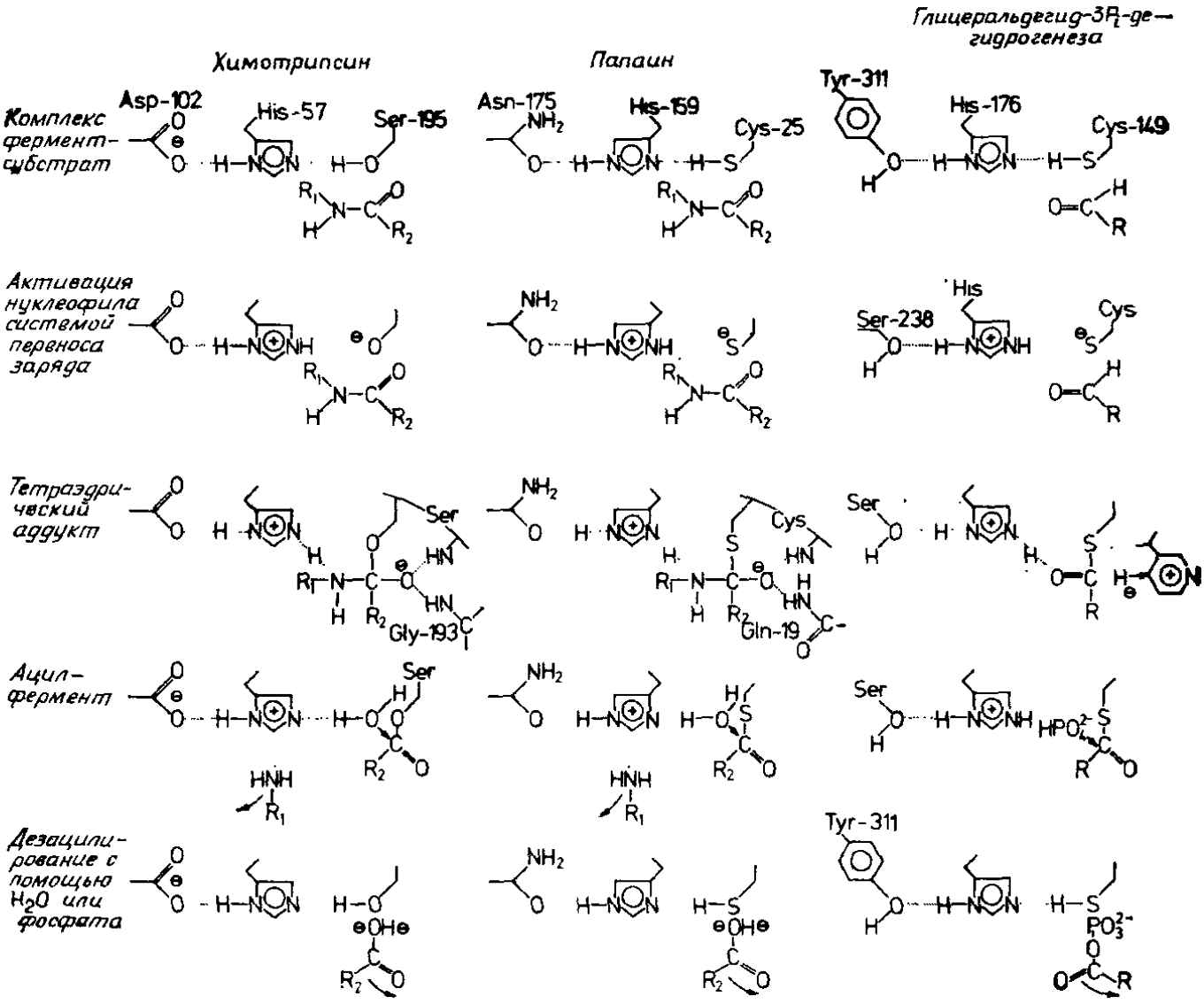


Рис. 11.1.

Аналогия механизмов катализа химотрипсина, папаина и глицеральдегид-3-фосфатдегидрогеназы [737].

В первом столбце указаны отдельные стадии, характерные для трех ферментов. Тетраэдрические промежуточные продукты (тетраэдрические аддукты), по-видимому, представляют переходные состояния. Приведена система переноса заряда химотрипсина Asp-His-Ser (слева). У субстрата показана только расщепляющаяся пептидная связь R_1—NH—CO—R_2. Пептидная цепь идет справа налево, как обычно на схемах механизмов протеаз. Все стадии реакции, катализируемой химотрипсином, обсуждаются в тексте; второй упомянутый там тетраэдрический промежуточный продукт показан на рисунке только в начальной фазе: H_2O атакует ацильный промежуточный продукт. Папаин — так называемая сульфгидрильная протеаза; нуклеофильной группой, аналогичной OH-группе остатка Ser-195 химотрипсина, в папаине служит группа SH. Фермент, приведенный в четвертой колонке, катализирует реакцию: глицеральдегид-3-фосфат + + NAD + неорганический фосфат $\rightleftarrows$ 1,3-дифосфоглицерат + NADH (в качестве субстратов показаны O=CHR, никотинамидный цикл и HPO_4^{2-}). В этом случае процесс, ведущий от тетраэдрического аддукта к ацилферменту, представляет собой реакцию элиминирования гидрид-ионов, т. е. стадию окисления. Ацильный промежуточный продукт атакуется не молекулой H_2O, как в обеих протеазах, а неорганическим фосфатом, так что конечным продуктом реакции является ангидрид кислоты.

ков химотрипсина, например трипсина или тромбина, но используется также и в других группах белков. Так, например, была отмечена постадийная аналогия реакций, катализируемых такими разными ферментами, как химотрипсин, папаин и глицеральдегид-3-фосфатдегидрогеназа [737] (рис. 11.1).

Теория переходного состояния

На каждой стадии химической реакции реагирующие вещества переходят из одного относительно стабильного состояния в другое через состояние с более высокой энергией (рис. 11.2). Переходное состояние — это состояние с более высокой свободной энергией, через которое реагирующие вещества должны пройти, чтобы перейти из одного стабильного состояния в другое [738—740]. Для достижения переходного состояния нужна дополнительная свободная энергия

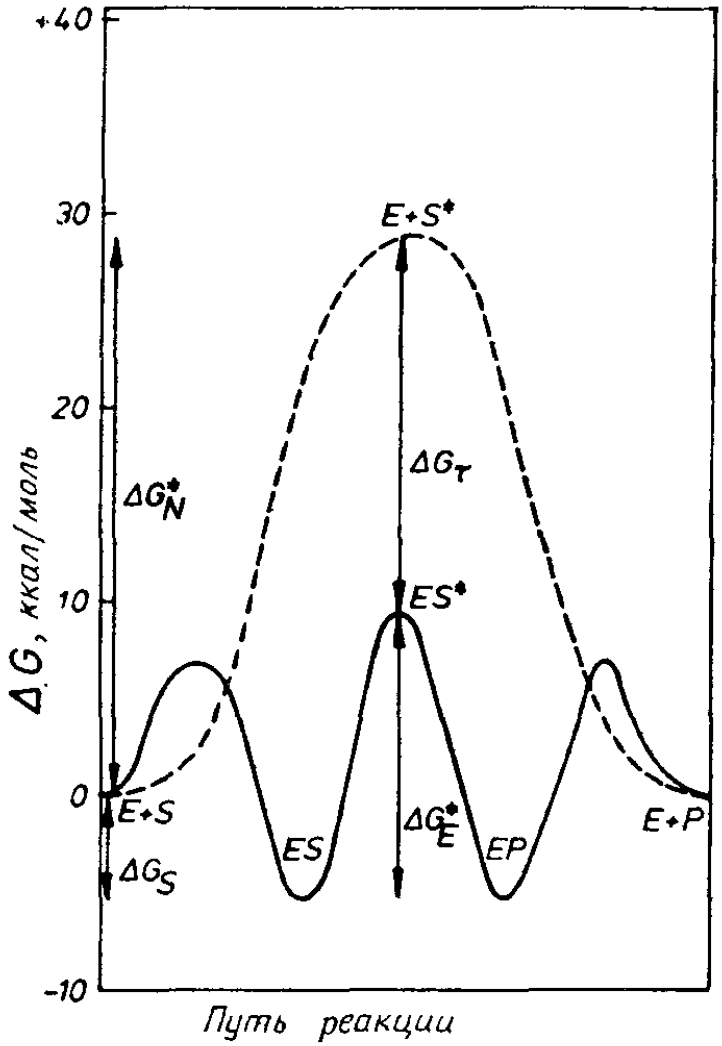


Рис. 11.2.

Теория переходного состояния и ферментативный катализ [740].

Показаны профили свободной энергии некатализируемой реакции (пунктирная линия) и соответствующей реакции, катализируемой ферментом (сплошная линия). Учитывается только одно соединение (субстрат); рассматриваемая реакция может быть, в частности, реакцией изомеризации. Предполагается, что и катализируемая, и некатализируемая реакции следуют одному и тому же механизму. Использованы следующие обозначения: E — фермент, S — субстрат, P — продукт, $S^{\neq}$ — переходное состояние субстрата. Изменения свободной энергии относятся к следующим реакциям: ΔG_S — образование ES из E и S; ΔG_τ — образование $ES^{\neq}$ из E и $S^{\neq}$; $\Delta G_N^{\neq}$ — образование $S^{\neq}$ из S; $\Delta G_E^{\neq}$ — образование $ES^{\neq}$ из ES. Величины, принятые для расчета изменений свободной энергии: K_S — константа ассоциации ES, 10^4 М$^{-1}$; K_P — константа ассоциации EP, 10^4 М$^{-1}$; k_d — константа скорости первого порядка для диссоциации ES в E и S и EP в E и P, 10^4 с$^{-1}$; k_E — константа скорости реакции первого порядка превращения ES в EP, 10^2 с$^{-1}$; k_N — константа скорости первого порядка соответствующей некатализируемой реакции, 10^{-8} с$^{-1}$; концентрации S и P приняты равными 1 М; константа равновесия при образовании P из S равна 1,0.

$\Delta G^{\neq}$ — свободная энергия активации. Согласно уравнению Аррениуса $k = s \cdot \exp(-\Delta G^{\neq}/RT) = s \cdot \exp(-\Delta H^{\neq}/RT + T\Delta S/RT)$ скорость химической реакции зависит от энергии активации [697].

Энергия активации имеет энтропийную составляющую, отражающую дополнительную упорядоченность, которую должна приобрести система, чтобы достигнуть переходного состояния, например, в в том случае, когда атомам для взаимодействия нужно существенно сблизиться. Другая, энтальпийная часть энергии активации представляет работу, которую следует затратить для сближения атомов, достаточного для образования ковалентной связи между ними. Ферменты могут понижать энтальпию активации (по сравнению с реакциями в растворе) в результате специфической фиксации субстратов, выбирая тем самым выгодный путь реакции из тех нескольких путей, по которым следует реакция в растворе.

Эффекты, способствующие высоким скоростям реакций

Сближенность, ориентация, направленность орбиталей и другие энтропийные эффекты

Ферменты могут ускорять химическую реакцию путем оптимальной для реакции ориентации реагентов в активном центре [741, 742]. Такая ориентация может включать и орбитальную направленность: точную взаимную ориентацию связывающих орбиталей атомов — партнеров по реакции [743, 744].

Связывание двух отдельных молекул в активном центре фермента превращает бимолекулярную реакцию в мономолекулярную, внутримолекулярную реакцию. Внутримолекулярные модельные реакции являются наиболее простым средством определения ускорения, которое может быть получено в результате сближения реагентов [631, 745]. Другими словами, энтропийный эффект фермента сказывается в увеличении эффективной концентрации субстрата. Поскольку скорости химических реакций пропорциональны концентрациям реагентов, то увеличения скорости в 10^8 раз можно ожидать на локальных участках с высокой концентрацией и упорядоченностью [631, 744].

Дестабилизация преобразуемых в субстратах атомных групп

Преобразуемая группа субстрата может быть дестабилизирована в тот момент, когда субстрат присоединяется к ферменту; энергия, затрачивающаяся на такую дестабилизацию [631], черпается из общей связывающей энергии. Механизмы дестабилизации могут включать замену растворителя, взаимодействия заряд — заряд и напряжения валентных связей и валентных углов. Дестабилизация может быть ослаблена в переходном состоянии, что означает снижение энергии активации, необходимой для достижения переходного состояния.

Помещая химические группы в неполярное окружение, можно увеличить скорость в 50 000 раз. Простой механизм, предложенный для пируватдекарбоксилазы, — тиаминпирофосфатсодержащего фермента основан на двух наблюдениях [746]. Одно состоит в том, что аддукт пирувата и аналог кофактора (рис. 11.3) декарбоксилируется в органических растворителях в $10^4 \div 10^5$ раз быстрее, чем в воде;

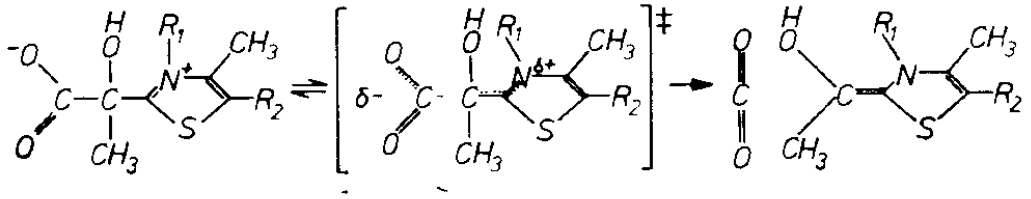


Рис. 11.3.

Катализ декарбоксилирования при замене растворителя.

Реакционноспособным фрагментом тиаминпирофосфата, кофактора ферментов декарбоксилирования, является тиазольная группа. В тиаминпирофосфате R_1 — пиримидиновый, а R_2 — пирофосфатный фрагменты; в обсуждаемом модельном соединении $R_1=CH_3$ и $R_2=H$. В органических растворителях модельное соединение теряет CO_2 в 50 000 раз быстрее, чем в воде [746]. Предполагается, что органические растворители благоприятствуют переходному состоянию, показанному на рисунке в скобках.

второе относится к гидрофобности пируватсвязывающего центра в пируватдекарбоксилазе. Увеличение скорости декарбоксилирования, ожидаемое при переходе аддукта пируват — кофермент в такое окружение, вносит существенный вклад в наблюдаемую скорость ферментативного действия. Связывающую энергию, необходимую для фиксации заряженного субстрата в таком невыгодном окружении и для обеспечения увеличения скорости реакции, могут предоставить пирофосфатная группа и пиримидиновый цикл кофермента.

Фрагменты ADP коферментов, таких, как ATP, FAD, NAD(P) или кофермент А, могут выполнять функцию, аналогичную функции пиримидиновой и пирофосфатной групп тиаминпирофосфата. Во всех изученных случаях (разд. 10.4) кофакторы присоединяются к своим коферментам в вытянутых конформациях, обеспечивая тем самым большое число взаимодействий с ферментами (рис. 11.4). Связывающая энергия, которая расходуется на специфические нужды других фрагментов кофермента, при этом всегда предоставляется фрагментом ADP.

Заряженные группы фермента становятся более активными, если сольватирующие молекулы воды заменить субстратом. Дестабилизация, основанная на десольватации, по-видимому, вносит вклад в каталитическое действие иона Zn^{2+}, связанного с карбоксипептидазой А [747]. Замещение сольватирующей молекулы воды на субстрат уменьшает диэлектрическую проницаемость окружения иона металла и увеличивает его активность в поляризации ацильной группы субстрата, необходимой для нуклеофильной атаки. Аналогичный эффект вносит вклад в процессы в D-субцентре лизоцима. Взаимодей-

ствие с субстратом десольватированной карбоксильной группы остатка Asp-42 ведет к локальной дестабилизации фермент-субстратного комплекса; эта дестабилизация ослабляется при образовании переходного состояния, напоминающего оксокарбониевый ион [531].

Длины связей и валентные углы в субстрате могут искажаться при его присоединении к ферменту. Фермент может действовать на химическую группу субстрата, придавая ей строение, близкое

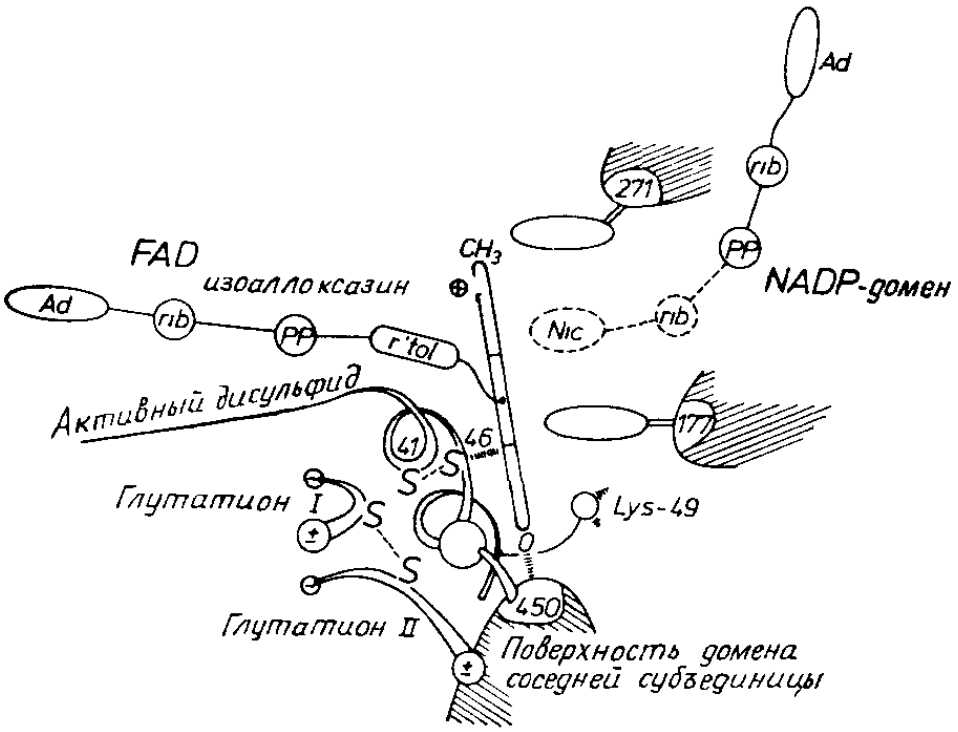


Рис. 11.4.

Актнвный центр глутатнонредуктазы [124].

В димерном ферменте имеются два идентичных активных центра. Восстанавлнваемые агенты проходят следующий путь (справа иалево): иикотииамидный фрагмент NADPH→изоаллоксазиновый цикл FAD (показан сбоку)→дисульфид с окислнтельно-восстановительной активностью→дисульфидиая связь глутатиона. Глутатионсвязывающий центр и каталнтический цеитр образованы остатками обеих субъединиц. Отметим, что NADPH и FAD присоедиияются к ферменту в вытянутых конформацнях.

к структуре переходного состояния; механизм такой структурной дестабилизации включает деформацию валентных углов, сближение реагирующих атомов до межатомных расстояний, меньших суммы вандерваальсовых радиусов, а также растяжение ковалентных связей до длины, превосходящей сумму ковалентных радиусов связанных атомов [631, 739, 740]. Наиболее хорошо изучен случай структурной дестабилизации, происходящей при присоединении N-ацетиламино-сахарного остатка, к субцентру D-лизоцима [748]. Тетраэдрический $(sp)^3$-атом углерода-1 искажен, и соответствующее напряжение ослабляется в переходном состоянии, где атом углерода-1 становится планарным (sp^2). Если учитывать десольватацию остатка Asp-42 при связывании субстрата, то энергия дестабилизации составит по крайней мере + 8,6 ккал/моль. Эта энергия восполняется общей связывающей энергией олигосахаридного субстрата,

Связывающие взаимодействия вдали от каталитического центра стабилизируют комплекс трипсин — ингибитор. Это следует из сравнения взаимодействия трипсин — субстрат с взаимодействием трипсин — ингибитор [269, 536]. Комплекс трипсина с псевдосубстратом — ингибитором характеризуется аномальным расстоянием 2,6 Å между атомом углерода карбонильной группы-15 псевдосубстрата и остатком Ser-195 фермента. Обычным субстратам химотрипсина и трипсина, в которых осуществляется несколько выгодных контактов, для достижения стадии ацилирования необходима энергия активации от + 5 до + 15 ккал/моль. Однако при образовании комплекса трипсин — ингибитор оптимизируются многочисленные другие взаимодействия, и величина ΔG оказывается равной —18 ккал/моль, несмотря на напряженность C—O^{γ} -аддукта (табл. 5.6). Таким образом, различие энергий стабилизации можно объяснить различием контактирующих областей в комплексах, которые трипсин образует с ингибитором и с обычными субстратами [749].

Процессы в каталитическом центре могут стабилизировать переходное состояние. До сих пор подчеркивался тот факт, что дальние взаимодействия поставляют свободную энергию активируемым группам в каталитическом центре фермент-субстратного комплекса. Однако взаимодействия и в самом каталитическом центре могут стабилизировать переходное состояние и тем самым вносить вклад в эффективность ферментативного катализа. В химотрипсине выигрыш энергии, обеспечивающийся образованием двух водородных связей между активированным субстратом и атомами азота остова, а также частичной компенсацией заряда скрытого внутри белка остатка Asp-102 (рис. 11.1), способствует компенсации энергии образования напряженной связи между ферментом и субстратом в тетраэдрическом комплексе [537].

Химические группы как части ферментов

Дестабилизирующие эффекты в фермент-субстратном комплексе оказывают влияние на состояние преобразуемых групп субстратов. Однако в ферменте предусмотрены также функциональные группы, которые более тонко воздействуют на преобразуемые группы. Общий кислотно-основной катализ довольно обычен в ферментах, и с его помощью скорость реакции может увеличиваться в 1000 раз. В химотрипсине эту функцию выполняет зарядно-релейная система, которая посредством водородных связей обеспечивает протонный транспорт в нескольких стадиях реакции (рис. 11.1). В других ферментах, например в глутатионредуктазе, белок обладает активными группами (FAD и цистеиновая пара с окислительно-восстановительной активностью) для транспорта электронов через молекулу фермента (рис. 11.4).

Раздельное рассмотрение эффектов, приводящих к увеличению скорости ферментативного катализа

Неоднократно предпринимались попытки количественно оценить вклады различных эффектов (например, сближенности, направленности орбиталей, дестабилизации, общего кислотно-основного катализа и др.) в увеличение скорости, к которому приводит действие данного фермента. Однако пример химотрипсина показывает, что эти «эффекты» представляют разные способы описания одного и того же действия в активном центре и в действительности их нельзя отделить друг от друга. С другой стороны, представления, развитые при анализе фермент-субстратных взаимодействий, расширили наши представления о химическом катализе и способствовали созданию аналогов ферментов на основе полимеров непептидной природы [745, 750].

11.3 Биологические и медицинские аспекты действия и функции белков

Белки как функциональные составляющие организма

Белки in situ должны обладать широким спектром свойств, которые позволяют им действовать в очень сложном окружении [586]. К этим свойствам относятся быстрое спонтанное свертывание, необходимые физико-химические свойства, например растворимость и изоэлектрическая точка, а также «хорошее» поведение (внутриклеточный белок не должен взаимодействовать с подавляющим большинством компонентов клетки) (разд. 10.1).

Многие белки in situ присутствуют в значительно более высоких концентрациях, чем обычно бывает необходимо для реализации их специфических функций, например катализа определенных химических реакций. Это указывает на то, что белки возможно используются в качестве формы хранения аминокислот и (или) что они выполняют некие еще не известные нам функции. Например, высокая концентрация креатинкиназы в мышце (20 мг на 1 г ткани [751]) оставалась непонятной до тех пор, пока не установили, что этот белок функционирует не только как фермент, но также в форме, составляющей М-диск, как структурная опора сократительного аппарата [752].

Одна из трудностей изучения биологической функции состоит в том, что эта функция может оставаться незамеченной до тех пор, пока она выполняется нормально. Поэтому часто выявление функции базируется на исследованиях заболеваний и мутаций, а также действия ядов, токсинов и лекарств. Недостаточную или нарушенную функцию, как правило, можно связать с неполной, недостаточной или нарушенной белковой структурой. Это показывает, что в

функционировании белка детали молекулярного строения наглядно и непосредственным образом отражаются на деятельности целого органа или организма [753].

Развитие новых функций на осиове старых белков

Структура белка может быть значительно старше, чем его функция. Лактальбумин, например, который появился как функциональная составляющая молочных желез около 100 миллионов лет назад, существовал, вероятно, в виде лизоцимподобного белка задолго до того, как возникла потребность в этой функции. По-видимому любая термодинамически возможная функция может развиться на основе существующих белковых структур (ср. раздел 9.3).

Некая новая функция также может быть развита на основе предшествующих белков в совершенно новом функциональном направлении [754]. Как видно из табл. 9.4, сериновая протеаза является прототипом функциональной единицы, которая неоднократно использовалась при развитии сложных физиологических систем. Другой распространенный пример — белки актин и миозин, которые широко распространены в подвижных клетках и их содержимом [755, 756]. У более высокоразвитых организмов актин-миозиновыми системами осуществляются такие различные функции, как сокращение мышц, освобождение соединений-переносчиков в нервной системе, амебовидное движение белых кровяных телец и закупорка поврежденных кровяных сосудов путем создания сгустка. Кроме того, в некоторых биологических процессах, когда должна стабилизироваться или изменяться форма клеток, используется свойство актина образовывать самые разнообразные структуры за счет обратимой полимеризации [757].

Тот факт, что большое разнообразие функций не сопровождается большим разнообразием структур [586], имеет позитивное значение для практической медицины, так как он показывает, что молекулярная физиология, по-видимому, не столь непостижима, как это считалось до недавнего времени. Однако этот позитивный аспект используется пока еще крайне мало, поскольку структурно-специфичные лекарства, например ингибиторы сериновых протеаз или ацетилсалицилат, могут вызывать очень разнообразные эффекты и поэтому действовать неизбирательно [84].

11.4 Скелетная мышца — система, в которой действие белка можно связать с общей деятельностью органа

Три свойства скелетной мышцы сделали этот орган удобным объектом структурного и функционального анализов: (а) относительная нечувствительность к манипуляциям; (б) большие количества типичных белков; клетка мышцы содержит 10% миозина, 3,5%

актина, 2% тропомиозина плюс тропонин и 2% переносящей Ca^{2+} АТРазы; (в) высокоупорядоченное геометрическое расположение сократительных белков. Применяя такие методы, как малоугловое рассеяние рентгеновских лучей и электронная микроскопия высокого разрешения в сочетании с биохимическими и физнологическими средствами анализа, можно связать детали молекулярного строения с общей деятельностью органа [755, 758].

Распознавание, ответ и регуляция — аспекты биологических функций белковых структур в клетке. Хотя клетка мышцы высоко специализирована, тем не менее она проявляет большинство черт, типичных для живых систем (табл. 11.1). Так, она обладает способностью к деятельности и к контролю своей деятельности [687]. Сигнал, попадающий в эту систему (нервный импульс), вызывает мощный ответ (движение или напряженность), который строго контролируется во времени, пространстве и по своей интенсивиости и который координируется с функционально родственными системами, например с процессами, поставляющими химическую энергию. В этом отношении функции этих белков подпадают под категории клеточной биологии: распознавание (с какими молекулами взаимодействует белок?), отклик (как белок реагирует на раздражение или сигнал?) и регуляция (как контролируется активность белка или какой процесс осуществляет этот контроль?). Однако все эти выражения описывают различные стороны структуры белка, и, следовательно, между ними нельзя провести четкой границы.

Структурная и функциональная организация сократительных белков

Главным структурным и функциональным звеном клетки мышцы является саркомер (рис. 11.5)—цилиндрическое образование диаметром 1,5 мкм и длиной 2 мкм, которое содержит около 2×2000 тонких и 1000 толстых белковых нитей. В левой части рис. 11.5 показано образование толстой нити из приблизительно 200 молекул миозина. Тонкая нить образуется путем ассоциации 2×175 мономеров глобулярного актина, 2×25 молекул тропомиозина и 2×25 звеньев трехкомпонентного белка тропонина. Нить актина представляет собой двойную спираль (разд. 5.1) с периодом идентичности от 360 до 370 Å [755, 758]. Саркомер — один из многих образующихся in vivo ансамблей макромолекул, имеющих пространственную группу симметрии.

Цикл образования мостиков — элементарный процесс сокращения мышц. Сокращение происходит в результате скольжения тонких нитей вдоль толстых нитей к центру саркомера [759, 760]. Сила между нитями возникает благодаря поперечным сшивкам, идущим от толстых нитей. Эти сшивки, называемые также миозиновыми головками, представляют собой биохимически активные участки мо-

Таблица 11.1

Клетка мышцы млекопитающего как система, в которой сигналы (нервные импульсы или присоединение гормона) приводят к контролируемому ответу[a]

Раздражение (сигнал), переданное мембране клетки	Нервный импульс	Присоединение гормона (адреналина)
Медиатор	↓ Освобождение ионов Ca^{2+} из саркоплазматической сети Ионы Ca^{2+}	↓ Синтез циклического AMP аденилатциклазой Ионы Ca^{2+}
Механизм усиления	↓ Нет	Циклический AMP ↓ Каскад ферментативного фосфорилирования
Эффектор	Аппарат сокращения; присоединение кальция к тропонину C приводит к кооперативным взаимодействиям, «включающим» аппарат сокращения	↓ Гликогенфосфорилаза; фосфорилирование этого фермента приводит к переходу из *b*-формы в более активную *a*-форму
Действие эффектора (ответ)	Образование поперечно-мостиковых циклов, основанное на взаимодействиях миозина, актина и Mg-ATP	Разрушение гликогена до глюкоза-1-фосфатных звеньев
Судьба медиатора после прекращения стимуляции	Поглощение ионов Ca^{2+} саркоплазматическим ретикулумом	Ферментативный гидролиз циклического AMP
Отставание во времени между окончанием стимуляции и окончанием действия эффектора	Миллисекунды	Секунды (каскад дефосфорилируется ферментативно)

[a] Основная функция клетки состоит в движении, основанном на сокращении и расслаблении, и в выполнении механической работы. Вспомогательные функции относятся к запасанию химической энергии, предназначенной для превращения в механическую энергию. В этом плане важной стадией является мобилизация гликогена, выполняющего функцию хранения энергии.

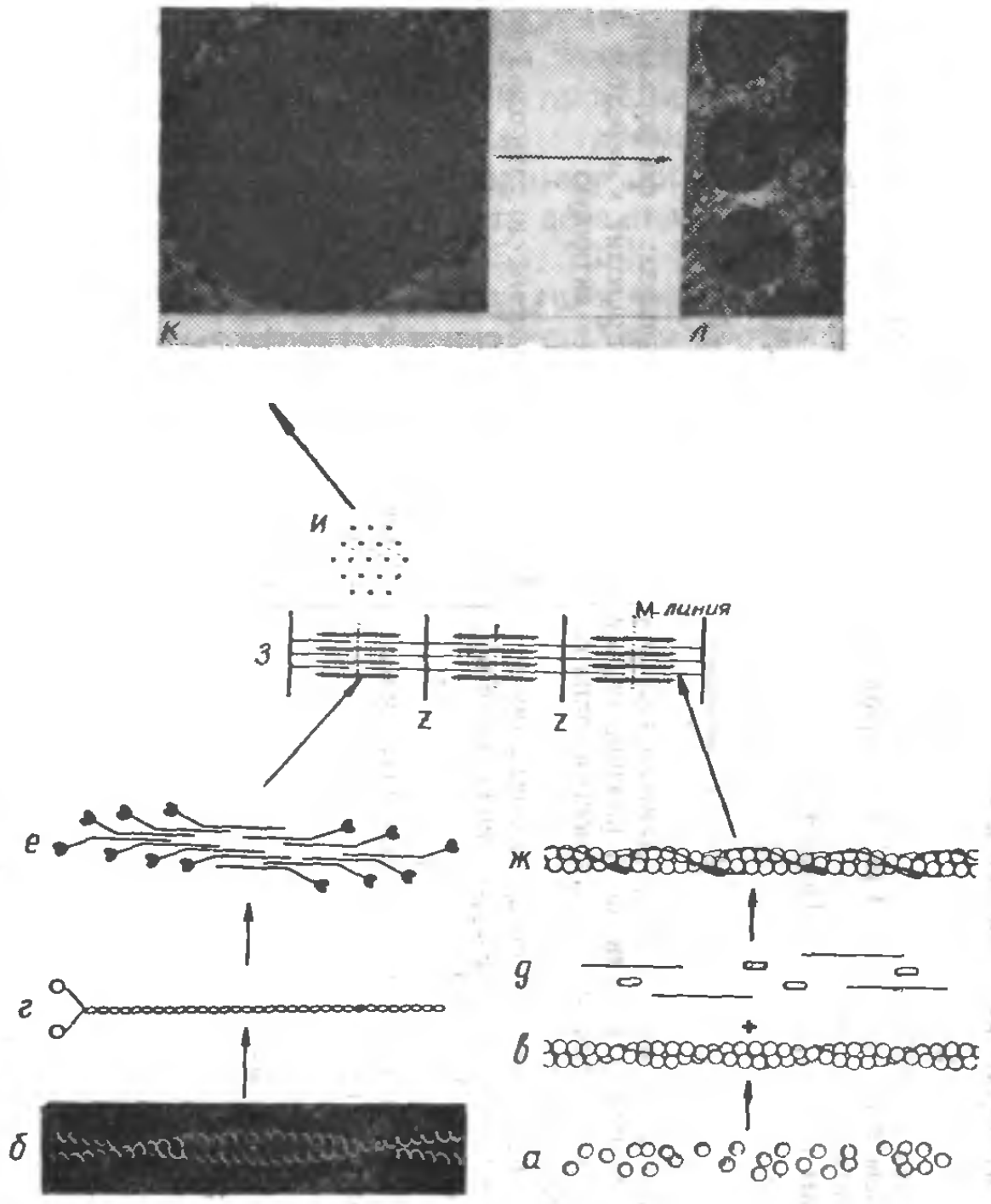


Рис. 11.5.

Структурная организацня аппарата сокращения скелетной мышцы [754].

Как схематически показано внизу слева, молекула миозина (*б* и *г*) состоит из длинной суперспирализованной α-спирали, «хвоста», и двух глобуляриых головок. «Хвост» образован тяжелыми цепями молекулы, которые вместе с четырьмя легкими цепями участвуют также в образовании обеих головок. Биполяриые агрегаты молекул миозиина, так называемые толстые филамеиты (*е*), размещены в саркомерах (*з*). Тонкий филамеит имеет длину 1200 нм и диаметр 10 им; его молекуляриая масса 10^8. Полимеризация актиина (*а*) приводит к длиным спиральиым филаментам (*F*-актин). Два других белка, тропонин и тропомиозин, присоединяются к *F*-актиину; в результате (пример самоагрегирования [754]) образуется тонкий филамеит (*ж*; длина 1000 им, диаметр 5 им, молекуляриая масса $2{,}5\times10^7$). Тоикие фнламенты прикрепляются к Z-дискам, отделяющим саркомеры друг от друга. Тоикие филаменты, выходящне из разных сторои Z-мембраны, имеют противоположиую полярность. В саркомере тонкие и толстые филаменты образуют гексагональную решетку (*и*). Как видио из электронных микрофотографий (предоставлены Хофмаиом), упорядоченность толстых и тонких филаментов близка к кристаллической (*к*). Миофибриллы (*л*), находящиеся на следующем уровие структурной иерархии, построеиы из саркомеров. Совокупность миофибрилл образует сократительный аппарат клетки мышцы.

лекулы миозина, которые взаимодействуют с Mg-ATP и с актином в так называемых поперечно-мостиковых циклах (рис. 11.6). Один мостик способен совершить от 10 до 100 циклов в секунду. Поскольку синхронизация циклов миозиновых головок (на саркомер их приходится целый миллион!) отсутствует, в результате возникает постоянно действующая сила и плавное движение.

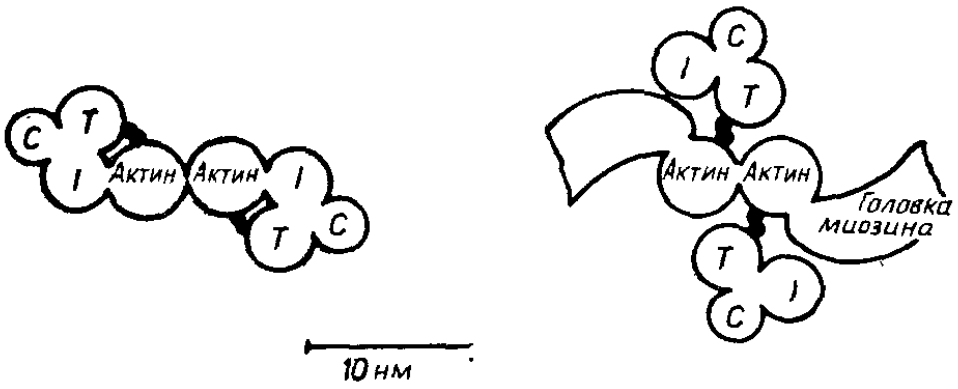


Рис. 11.6.

Взаимодействия белков в аппарате сокращения [615].

Слева показано сечение тонкого филамента в состоянии расслабления. C, I и T — белковые составляющие тропонинового комплекса. Тропомиозин блокирует связывающие центры головок миозина. При присоединении Ca^{2+} к тропонину C тонкий филамент переключается в активное состояние (справа). Тропомиозин перемещается к центру желоба спирали актина, и головки миозина (поперечные мостики) могут взаимодействовать с актином. Цикл начинается с присоединения головки миозина к актиновому филаменту [755]. Затем миозин отклоняется, подтягивая активный филамент к центру саркомера (на рисунке — из плоскости бумаги) приблизительно на 80 Å. После этого ATP отделяет тонкий филамент от головки миозина, которая может начать другой цикл образования мостика далее по актиновому филаменту.

Сокращение основано на превращении химической энергии в механическую. Если мышца активирована, расход ее химической энергии в форме ATP увеличивается на один-два порядка. Один из предложенных механизмов состоит в следующем. В расслабленном состоянии мышцы актин и миозин не взаимодействуют и головка миозина характеризуется слабой ATPазной активностью, которая лимитируется стадией освобождения продуктов, главным образом Mg-ADP. В работающей мышце актин усиливает ATPазную активность миозина примерно в 100 раз путем вытеснения Mg-ADP. До сих пор не ясно, является ли действие актина на ADP-связывающий центр миозина прямым или аллостерическим.

«Ответ» ATPазы миозина показывает, что одинаково важны и действия белка и контроль над этим действием. Миозин проявляет ATPазную активность только при сокращении мышцы; высокая активность ATPазы в отсутствие взаимодействия актин — миозин была бы бессмысленной затратой химической энергии.

Большие времена жизни Mg^{2+}-нуклеотидных комплексов могут объяснить их широкую распространенность в качестве субстратов. Для ответа на вопрос, почему физиологическим субстратом ATPазы миозина и многих других ферментов является Mg-ATP, а не Ca-ATP,

можно воспользоваться кинетическими данными. Константы стабильности Mg-нуклеотидных и Ca-нуклеотидных комплексов почти идентичны, но распад Ca^{2+}-комплексов происходит в 1000 раз быстрее, чем соответствующих Mg^{2+}-комплексов [762]. Не кальций, а магний, полураспад ATP- и ADP-комплексов которого имеет порядок миллисекунд, был избран для подавления АТРазной активности миозина в состоянии расслабления мышцы и для проведения относительно медленных конформационных изменений ($t_{1/2} > 1$ мсек), которые происходят на стадиях каталитического действия АТРазы актин-активированного миозина, равно как и некоторых других ферментов [758].

Такой пример показывает, что для оценки влияния комплексов на действие белков должны быть изучены равновесные и кинетические свойства этих комплексов, а также их структурных изомеров. Для этой цели могут применяться методики изучения кинетики быстрых (10^{-7} с) и очень быстрых (10^{-9} с) реакций [763—766], скорости которых близки к скоростям процессов, протекающих под диффузионным контролем (10^{-10} с).

Отклик на нервные импульсы и гормональное действие в клетке мышцы

Когда клетка мышцы находится в состоянии покоя, центры прикрепления поперечных мостиков в тонких нитях блокированы молекулами тропомиозина [767—769], прочными двуволокнистыми образованиями [214]. Тропомиозин — прототип суперспирализованной α-спирали (разд. 5.2) с длиной молекулы 400 Å, прикрывает семь глобулярных звеньев актина [768, 769]. Такое расположение тропомиозина, по-видимому, определяется тропонином, или, более точно, общей конформацией комплекса тропонина, свободного от Ca^{2+} (рис. 11.6).

Цепочка событий, приводящих к смещению тропомиозина, начинается на клеточной мембране. Когда нервные импульсы активируют клетку мышцы, имеющую объем 1 мкл, ионы Ca^{2+} выделяются из саркоплазматического ретикулума [770] в цитоплазму, где концентрация свободных ионов Ca^{2+} становится на два порядка выше 1 мкм (рис. 11.7). Это приводит к насыщению тропонина С — кальций-чувствительного компонента тонкой нити [771]; к молекулам тропонина С присоединяются 90% из общего количества 10^{14} ионов. Связывание Ca^{2+} вызывает конформационные изменения всего тропонинового комплекса [772]. При измененной структуре тропонина тропомиозин уже не может больше удерживаться в «выключенном» состоянии. Тропомиозиновая спираль соскальзывает в сторону к новому положению ближе к центру желоба. Таким образом, одна молекула тропомиозина освобождает семь мономеров актина, способных к взаимодействию с миозином [767, 769, 785].

Активация сократительного аппарата ионами Ca^{2+} — модель кооперативного процесса [772, 773]. Активация сократительного аппарата ионами Ca^{2+} (рис. 11.7) хорошо иллюстрирует кооперативное поведение [92, 678, 682, 774] и аллостерический контроль [92, 681, 775, 776] в белках. При кооперативных действиях компонент белка или надмолекулярных образований, например тонких нитей, осуществляется переход из одного стабильного состояния («выключенное») в другое («включенное»). Типичным «переключающим» процессом внутри структурных единиц, например в компонентах тропонина, является конформационный переход. Эта структурная изомеризация в свою очередь вызывается специфическим аллостерическим лигандом, в нашем случае ионом Ca^{2+}. Поведение тропомиозина показывает, что не все динамические взаимодействия белок — белок основаны на структурной изомеризации в пределах отдельных белков; тропомиозин проявляет свое действие, смещаясь как молекула в целом (рис. 11.6).

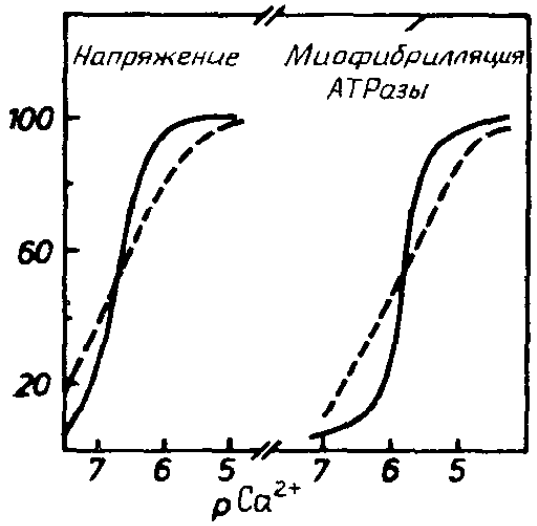


Рис. 11.7.

Влияние концентрации ионов Ca^{2+} на напряжение мышцы и АТРазную активность миозина [785].

pCa^{2+} является понятием, аналогичным pH. Наблюдаемая кооперативность (сплошные линии), по-видимому, отражает тот факт, что структурные и ферментативные переходы в сократительном аппарате «включаются» при присоединении четвертого иона Ca^{2+} к тропонину С [615, 773, 785]. Пунктирными кривыми показана гипотетическая активность сократительного аппарата при отсутствии структурной кооперативности.

Метаболические, нервные и гормональные раздражения, часто усиленные ферментативными каскадами, коррелируют расход и запас энергии. Какие механизмы обеспечивают химической энергией активную сократительную систему? Важную роль в непосредственной корреляции запаса энергии с энергетическими запросами играет ADP — основной продукт АТРазной деятельности миозина. ADP как метаболический сигнал приводит в действие выработку ATP активированными белками в ключевых положениях гликолитического пути и окислительного фосфорилирования [19]. Создание запаса энергии на продолжительный период регулируется ионами

Ca^{2+}, которые стимулируют разложение гликогена [777]. Однако в этом случае механизм активации кальцием включает несколько стадий в отличие от одностадийной активации этим же ионом сократительного аппарата (рис. 11.1).

На мобилизацию гликогена может влиять гормон эпинефрин, который действует независимо или параллельно со стимуляцией нервными импульсами. Каскад мембранных и цитоплазматических процессов, которые ведут от связывания эпинефрина к фосфорилированию глюкозных фрагментов гликогена, функционирует как высоко эффективный механизм кинетического усилия (табл. 11.1). Концентрация гормона во внеклеточном пространстве составляет приблизительно 10^{-11} М; циклический AMP, который образуется из ATP — первого эффектора гормонального действия, активирует при концентрациях уже от 10^{-8} до 10^{-7} М белок киназу. Затем происходит 10-кратное усиление при стимуляции фосфорилазы *b* киназы, и заключительное усиление в 20 ÷ 50 раз достигается при превращении фосфорилазы *b* в фосфорилазу *a*.

По аналогии с событиями, происходящими в мышце, можно полагать, что и в других физиологических процессах действие циклического AMP, второго передаточного звена для многих гормонов, связано с модуляцией системы реакций фосфорилирования — дефосфорилирования белков [777, 778].

Быстрая диссоциация комплексов кальций — белок служит механизмом переключения активности белка. Рассмотрим теперь внутриклеточные процессы, происходящие после того, как прекращается действие раздражения (нервных импульсов или присоединения гормонов). Метаболические процессы прекратятся с запаздыванием по времени порядка секунд после того, как содержание циклического AMP или ионов Ca^{2+} понизится ниже критической величины (табл. 11.1). В отличие от этого сократительный аппарат выключается в интервале 10—20 мс путем удаления ионов кальция из тонких нитей. Скорость этого процесса свидетельствует об эффективности саркоплазматического ретикулума как кальциевого насоса [701]. С другой стороны, наличие этой сложной системы, которая предназначена исключительно для поглощения и освобождения ионов Ca^{2+}, позволяет предположить, что ионам Ca^{2+} присущи уникальные свойства в отношении функций переключения. На это указывает также тот факт, что ионы Ca^{2+} принимают участие во многих других физиологических процессах в качестве посредников между поступающими раздражениями и клеточными реакциями [615], например, в светочувствительных клетках [779] или в яйцеклетке, где Ca^{2+}-зависимый механизм, по-видимому, обеспечивает оплодотворение только одной клеткой спермы [780]. Всем этим процессам свойственно быстрое включение и выключение активности белка. Это же относится и к сократительно-

му аппарату: быстрое сокращение и быстрое расслабление необходимы при многих движениях.

По-видимому, кальций был эволюционно отобран для контроля быстрых биологических процессов из-за свойств его координационных соединений [615], которые при биологических условиях способны распадаться с приемлемыми скоростями [781]. В пусковой системе, основанной на сигнальной молекуле S и молекуле-мишени T, ассоциация и диссоциация происходят в соответствии с уравнением $S + T \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} ST$, где k_{+1} и k_{-1} — соответственно константы скорости образования и распада комплекса. Константа стабильности (константа ассоциации) равна $K = k_{+1}/k_{-1} = [ST] / [S] \cdot [T]$. Оптимальное значение для системы запуска, действующей в физиологически значимом интервале концентраций, варьирует от 10^5 до 10^7 $м^{-1}$ [782]. Величина k_{+1} для замещения Ca^{2+} в воде составляет $10^8 M^{-1} \cdot с^{-1}$, а для Mg^{2+} $10^5 M^{-1} \cdot с^{-1}$ [762]. Отсюда величина $k_{-1} = k_{+1}/K$ имеет порядок 1000 в секунду в случае комплексов Ca^{2+} и 1 в секунду в случае комплексов Mg^{2+}. Кратчайшие периоды полураспада $(\ln \frac{2}{k_{-1}})$ кальциевых и магниевых комплексов составляют при этих условиях соответственно 0,7 мс и 0,7 с.

Таким образом, в отличие от ионов Mg^{2+} и других возможных биологических активаторов ионы Ca^{2+} могут наилучшим образом осуществлять функцию включение — выключение независимо от того, имеет ли время действия белка порядок секунд или миллисекунд.

И наконец, следует подчеркнуть, что, по-видимому, именно биохимические процессы в координационных соединениях кальция ограничивают скорость контроля в мышцах позвоночных. Быстродействующие наружные мышцы человеческого глаза совершают до 50 циклов сокращения — расслабления в секунду. С другой стороны, маховые мышцы крылатых насекомых могут колебаться с частотой (в звуковом диапазоне) свыше 1000 Гц. Столь быстрые колебания основаны на таких процессах в белках, в которых ионы Ca^{2+} участия не принимают, в частности на реакции АТРазы миозина на сжатие [783, 784].

Заключение

Анализ структуры белка является первым этапом в исследованиях механизма его действия и в конечном счете его биологической функции. В настоящее время наиболее хорошо изучены функции гемоглобина и химотрипсина. Высокая скорость и эффективность катализа химотрипсина (и других ферментов) можно приписать многим эффектам, ускоряющим химические реакции в модельных системах. К таким эффектам относятся ориентация субстрата(ов) в актив-

ном центре фермента в положение, оптимальное для реакции, стабилизация преобразуемых в субстрате групп относительно их переходного состояния и кислотно-основный катализ. Необходимая для увеличения скорости энергия, по-видимому, черпается из свободной энергии образования фермент-субстратных комплексов.

Некоторые свойства белков можно объяснить только в свете их функции, т. е. их вклада в более сложную деятельность. Одной из немногих систем, для которых удалось установить корреляцию между функцией белков и функцией органа, является скелетная мышца. Клетка мышцы активируется нервными импульсами (мембранно-направленными сигналами). В молекулярном аспекте мышечное сокращение основано на циклическом образовании поперечных мостиков за счет периодических взаимодействий между миозином, актином и Mg-ATP. Ионы Ca^{2+} и кальцийсвязывающие белки являются посредниками между нервными импульсами и эффекторами. Медиация ионами Ca^{2+} ограничивает скорость реакции на сигналы включение — выключение и предохраняет от сокращений без сигнала. Напротив, отдельные осцилляции маховых мышц крылатых насекомых контролируются не ионами Ca^{2+} или подобными низкомолекулярными соединениями, а самими сократительными белками. Это делает возможными очень быстрые периодические сокращения, которые, будучи инициированы (ионами Ca^{2+}), протекают сами по себе. В заключение отметим, что исследования мышцы показывают, что в функционировании белка отчетливо проявляется связь между деталями молекулярного строения и деятельностью всего организма.

Приложение Статистическая механика перехода спираль — клубок

A.1 Функция распределения

Функция распределения — наиболее строгое описание ансамбля молекул, находящихся в равновесии. Распределение конформационных состояний полипептидной цепи, построенной из N остатков, строго описывается функцией распределения.

$$Z = \sum_{\substack{\text{по всем}\\ \text{состо-}\\ \text{яниям } k}} \exp \frac{-E_k}{RT} \qquad \text{(A.1)}$$

где E_k — энергия всей системы (цепь плюс окружающий растворитель) в состоянии k. Член $\exp \frac{-E_k}{RT}$ называют статистическим весом состояния k. Если пренебречь вкладом растворителя и энергией напряжения длин связей и валентных углов, E_k сводится к функции двугранных углов (ϕ_i, ψ_i) остова и двугранных углов (χ_i^1, χ_i^2, ...) $\equiv \chi_i$ боковой цепи остатка i (рис. 2.2). Поскольку конформационное пространство непрерывно, сумма превращается в интеграл

$$Z = \text{const} \cdot \int \exp \frac{-E(\phi_1 \ldots \chi_N)}{RT}\, d\phi_1 d\psi_1 d\chi_1 d\phi_2 d\psi_2 \ldots d\chi_N \qquad \text{(A.2)}$$

A.2 Вероятность данной конформации остатка

Если пренебречь растворителем и всеми взаимодействиями остаток — остаток, функция распределения выражается произведением статистических весов отдельных остатков. Эти статистические веса можно разбить на веса для областей α_R, α_L и ε. Если пренебречь всеми взаимодействиями *между* остатками, то энергию можно разделить на вклады отдельных остатков и разложить интеграл на множители:

$$E(\phi_1 \ldots \chi_N) = E_1(\phi_1\psi_1\chi_1) + E_2(\phi_2\psi_2\chi_2) + \ldots + E_N(\phi_N\psi_N\chi_N)$$

$$Z = \text{const} \cdot \int \exp \frac{-E_1}{RT}\, d\phi_1 d\psi_1 d\chi_1 \ldots \int \exp \frac{-E_N}{RT}\, d\phi_N d\psi_N d\chi_N \qquad \text{(A.3)}$$

Полученные интегралы отдельных остатков можно аппроксимировать путем сведения пределов интегрирования к тем областям, в которых отсутствуют стерические напряжения, т. е. где значение E_i не слишком велико. Как видно из рис. 2.3, *б*, имеются области правой и левой α-спиралей, а также вытянутой цепи, обозначенные α_R, α_L и ε.

$$\int \exp \frac{-E_i}{RT}\, d\phi_i d\psi_i d\chi_i = \int_{\alpha_R} \exp \frac{-E_i}{RT}\, d\phi_i d\psi_i d\chi_i + \int_{\alpha_L} \ldots + \int_{\varepsilon} \ldots =$$

$$= z_i^{\alpha_R} + z_i^{\alpha_L} + z_i^{\varepsilon}$$

$$Z = \text{const} \prod_{i=1}^{N} \left(z_i^{\alpha_R} + z_i^{\alpha_L} + z_i^{\varepsilon} \right) \qquad \text{(A.4)}$$

Здесь $z_i^{\alpha R}$, $z_i^{\alpha L}$, z_i^{ε} — статистические веса соответственно областей α_R, α_L и ε. Они специфичны для данного типа остатка, т. е. $z_i^{\alpha R}$, $z_i^{\alpha L}$, z_i^{ε} представляют собой вероятности того, что фрагмент остова, отвечающий остатку i, примет какую-либо из конформаций соответственно в областях α_R, α_L и ε. Другими словами, величины $z_i^{\alpha R}$, $z_i^{\alpha L}$, z_i^{ε} отражают склонность остатка находиться в конформации α_R, α_L, ε. Эти вероятности зависят от энергии E_i в этой области и о размера области. Таким образом, при сравнении вероятностей α_R и ε более низкая энергия для α_R (лучшее связывание) вполне может быть компенсирована большей величиной области для ε (рис. 2.3).

Эмпирические статистические веса можно получить из наблюдаемых частот встречаемости. В принципе эти склонности можно рассчитать для каждого типа остатка по энергетическим картам, подобным приведенной на рис. 2.5. Однако эти карты следует расширить с тем, чтобы учесть зависимость от конформации χ_i боковой цепи. Кроме того, для получения разумных результатов нужно принять во внимание растворитель, что, однако, сделать достаточно сложно. С другой стороны, близкие значения с этими вероятностями $z_i^{\alpha R}$, $z_i^{\alpha L}$, z_i^{ε} можно получить путем приравнивания их частотам встречаемости состояний α_R, α_L, ε рассматриваемого типа остатка в глобулярных белках известной структуры. Эта процедура принята в большинстве методов предсказания.

Карта общей эмпирической энергии получена из частот встречаемости. Используя аппроксимацию уравнения (A.3), из наблюдаемых в глобулярном белке углов (ϕ, ψ) всех типов остатков можно получить карту общей эмпирической энергии $E(\phi, \psi)$. Для этой цели частоты встречаемости остатков (безотносительно к типу остатка), найденные в данной (ϕ, ψ)-области (рис. 2.4), приравни-

вают к статистическим весам $\exp\frac{-E(\phi,\psi)}{RT}$ в этой области. Отсюда статистические веса, а следовательно, и E можно выразить как функцию углов (ϕ, ψ). Эмпирическая энергия $E(\phi, \psi)$ достаточно хорошо согласуется с рассчитанной (рис. 2.5) [786].

А.3 Модель Айзинга

Линейный массив с взаимодействиями между ближайшими соседями впервые описан Айзингом. Упрощения функции распределения, помимо учтенных в уравнении (А.2), основаны на предположении об отсутствии взаимодействия между различными остатками. Это совершенно неверно в случае α-спиралей, поскольку в них существуют водородные связи между остатками i и $i + 3$ (рис. 5.4). Кроме того, кривые, описывающие переходы спираль — клубок в синтетических полипептидах [328, 787], имеют сигмоидальный характер, что указывает на кооперативность. Чтобы учесть этот факт, необходимо ввести иные аппроксимации функции распределения. Для подобного случая, а именно для линейного массива ферромагнетиков с взаимодействиями между ближайшими соседями, аппроксимация предложена Айзингом [788].

А.4 Модель Зимма — Брэгга для перехода спираль — клубок

Основная формула

Член, учитывающий взаимодействие ближайших соседей, вводится в упрощенную формулу. Зимм и Брэгг [789] применили модель Айзинга к переходу спираль — клубок гомополипептидных цепей. Для этой цели они разделили конформационное пространство на две области α_R или «α», но не α_R или «клубок». Кроме того, они использовали приближенное уравнение (А.4), не учитывающее взаимодействия остаток — остаток, а затем ввели член, учитывающий взаимодействие между ближайшими соседями. Для цепи, состоящей из N остатков данного типа, уравнение (А.4) принимает вид:

$$Z = \text{const}\,(z_1^a + z_1^c)\cdot(z_2^a + z_2^c)\cdots(z_N^a + z_N^c)$$

После перемножения получаем сумму 2^N слагаемых

$$Z = \text{const}\cdot \sum_{\substack{\text{по всем } 2^N \\ \text{конформа-} \\ \text{циям цепи}}} z_1^a \cdot z_2^a \cdot z_3^a \cdot z_4^c \cdot z_5^c \cdot z_6^a \cdot z_7^c \cdots z_N^a \qquad \text{(А.5)}$$

Для каждого возможного сочленения спираль — клубок и клубок — спираль вводится дополнительный множитель $\sqrt{\sigma}$, $\sigma \ll 1$

$$Z = \text{const} \sum_{\substack{\text{по всем } 2^N \\ \text{конформа-} \\ \text{циям цепи}}} z_1^{\gamma} \cdot z_2^{\gamma} \cdot z_3^{\alpha} \cdot \sqrt{\sigma} \cdot z_4^{c} \cdot z_5^{c} \cdot \sqrt{\sigma} \cdot z_6^{\alpha} \cdot \sqrt{\sigma} \cdot z_7^{c} \ldots z_N^{\alpha}$$

Это можно выразить в форме

$$Z = \text{const} \sum_{\substack{\text{по всем } 2^N \\ \text{конформа-} \\ \text{циям цепи}}} z_1^{\alpha} \cdot z_2^{\alpha} \cdot z_3^{\alpha} \cdot \sigma \cdot z_4^{c} \cdot z_5^{c} \cdot z_6^{\alpha} \cdot \sigma \cdot z_7^{c} \ldots z_N^{\alpha} \qquad \text{(A.6)}$$

путем попарного объединения множителей $\sqrt{\sigma}$ и использования их только для сочленений спираль — клубок. Это допустимо, поскольку α-сегментов столь же много, как и клубковых сегментов; в том случае, если один структурный тип превосходит по числу сегментов другой, различиями пренебрегают.

Множитель σ приводит к отрицательным результатам, если сочленений слишком много, и отдает предпочтение протяженным сегментам идентичной конформации. Это — альтернативное описание «кооперативности остатков идентичной конформации», или в данном случае «кооперативности остатков спиральной конформации». Можно также говорить о выгодности «нуклеаций клубка или спирали». Поэтому величину σ называют «параметром кооперативности или нуклеации».

Матричное представление

Функция распределения в матричной форме. Крамерс и Ванье [790] предложили записать уравнение (А.6) в компактной матричной форме, упрощающей расчет Z.

$$Z = \text{const} \cdot (1,\ 1) \cdot \begin{pmatrix} z_1^{c} & 0 \\ 0 & z_1^{\alpha} \end{pmatrix} \cdot \begin{pmatrix} z_2^{c} & z_2^{\gamma} \cdot \sigma \\ z_2^{c} & z_2^{\gamma} \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} z_N^{c} & z_N^{\alpha} \cdot \sigma \\ z_N^{c} & z_N^{\alpha} \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{(A.7)}$$

Проверим соответствие этого выражения уравнению (А.6) при малых N:

$$N = 1 \quad Z = \text{const} \cdot (z_1^{c} + z_1^{\alpha})$$

$$N = 2 \quad Z = \text{const} \cdot (z_1^{c} \cdot z_2^{c} + z_1^{c} \cdot z_2^{\alpha} + z_1^{\alpha} \cdot \sigma \cdot z_2^{c} + z_1^{\alpha} \cdot z_2^{\gamma})$$

$$N = 3 \quad Z = \text{const} \cdot (z_1^{c} \cdot z_2^{c} \cdot z_3^{c} + z_1^{c} \cdot z_2^{\gamma} \cdot \sigma \cdot z_3^{c} + z_1^{\alpha} \cdot \sigma \cdot z_2^{c} \cdot z_3^{c} + z_1^{\alpha} \cdot z_2^{\alpha} \cdot \sigma \cdot z_3^{c} + z_1^{c} \cdot z_2^{c} \cdot z_3^{\alpha} + z_1^{c} \cdot z_2^{\gamma} \cdot z_3^{\alpha} + z_1^{\alpha} \cdot \sigma \cdot z_2^{c} \cdot z_3^{\alpha} + z_1^{\alpha} \cdot z_2^{\alpha} \cdot z_3^{\alpha})$$

Эти соотношения можно распространить и на большие значения N.

Простой случай гомополимера удовлетворительно описывается матрицей N-го порядка в терминах собственных значений. Если применить уравнение (А.7) к *гомо*полимерам, то все величины z_i^a, как и величины z_i^c, будут одинаковы. Для дальнейшего упрощения относительные статистические веса s можно выразить как

$$s_i = z_i^a / z_i^c \equiv s \tag{A.8}$$

Будем считать $(z_i^c)^N$ за постоянную, которую опустим, поскольку интерес представляет только изменение Z при переходе спираль — клубок, а не ее абсолютное значение.

$$Z = (1,\ 1) \cdot \begin{pmatrix} 1 & 0 \\ 0 & s \end{pmatrix} \cdot \begin{pmatrix} 1 & s\sigma \\ 1 & s \end{pmatrix}^{N-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{A.9}$$

Путем диагонализации матрицы это можно свести к выражению:

$$Z = (1,\ 1) \begin{pmatrix} 1 & 0 \\ 0 & s \end{pmatrix} A\,(A^{-1}UA)^{N-1}\,A^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{A.10}$$

При $U = \begin{pmatrix} 1 & s\sigma \\ 1 & s \end{pmatrix}$

$$A^{-1}UA = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$\lambda_{1(2)} = \frac{1}{2}\left(1 + s \overset{+}{(-)} \sqrt{(1-s)^2 + 4s\sigma}\right)$$

$$A = \begin{pmatrix} 1-\lambda_2 & 1-\lambda_1 \\ 1 & 1 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 1 & \lambda_1 - 1 \\ -1 & 1-\lambda_2 \end{pmatrix}$$

$$(A^{-1}UA)^{N-1} = \begin{pmatrix} \lambda_1^{N-1} & 0 \\ 0 & \lambda_2^{N-1} \end{pmatrix}$$

то есть Z можно в явном виде выразить как функцию собственных значений.

$$Z = \lambda_1^N \frac{\lambda_1}{\lambda_1 - \lambda_2} - \lambda_2^N \frac{\lambda_2}{\lambda_1 - \lambda_2} \tag{A.11}$$

Поскольку $\lambda_1 > 1 > \lambda_2$ при всех s и σ, то для больших N статистическую сумму можно аппроксимировать как

$$Z = \lambda_1^N \tag{A.12}$$

Содержание спиральных остатков в гомополимере

Содержание спирали изменяется в соответствии с относительным статистическим весом и в зависимости от температуры. Значение Z можно использовать для определения среднего числа спиральных остатков $\langle n \rangle$ как функции s и σ. Согласно (А.7) и используя соотношение (А.8) для цепей, состоящих из ν спиральных сегментов, получаем

$$Z = \sum_{\text{по всем } 2^N \text{ конформациям цепи}} s^n \sigma^\nu \tag{A.13}$$

$$\langle n \rangle = \sum_{\text{по всем } 2^N \text{ конформациям цепи}} n \cdot \frac{s^n \cdot \sigma^\nu}{Z} \tag{A.14}$$

Таким образом, число спиральных остатков в каждой конформации цепи находится в соответствии с относительным статистическим весом этой конформации, который нормируется с помощью $1/Z$.

$$\langle n \rangle = \frac{s}{Z}\frac{\partial Z}{\partial s} = \frac{\partial \ln Z}{\partial \ln s} \approx \frac{\partial \ln \lambda_1^N}{\partial \ln s} = \frac{N \cdot s}{\lambda_1} \cdot \frac{\partial \lambda_1}{\partial s} = N\frac{\lambda_1 - 1}{\lambda_1 - \lambda_2} \tag{A.15}$$

$\langle n \rangle$ есть переменная величина, экспериментально наблюдаемая при переходе спираль — клубок. Теоретические кривые для $\langle n \rangle$ как функции s построены для нескольких значений σ на рис. А.1,*а*. Если $s = 1$, то при всех значениях σ половина остатков имеет спиральные конформации. Чем меньше σ, тем резче переход при $s = 1$, т. е. тем сильнее выражена кооперативность.

Число и длина спиральных сегментов

Наибольшее число спиральных сегментов отвечает точке перехода. В уравнении (А.13) величина ν означает число спиральных сегментов в данной конформации цепи. Среднее число спиральных сегментов $\langle \nu \rangle$.

$$\langle \nu \rangle = \sum_{\text{по всем } 2^N \text{ конформациям цепи}} \nu \frac{s^n \cdot \sigma^\nu}{Z} = \frac{\partial \ln Z}{\partial \ln \sigma} \approx \frac{\partial \ln \lambda_1^N}{\partial \ln \sigma} = N\frac{(\lambda_1 - 1)(1 - \lambda_2)}{\lambda_1 (\lambda_1 - \lambda_2)} \tag{A.16}$$

Эта функция, показанная на рис. А.1, *б*, имеет максимум при $s = 1$. Число сегментов уменьшается с уменьшением σ. Средняя длина спирального сегмента $\langle l \rangle$ равна

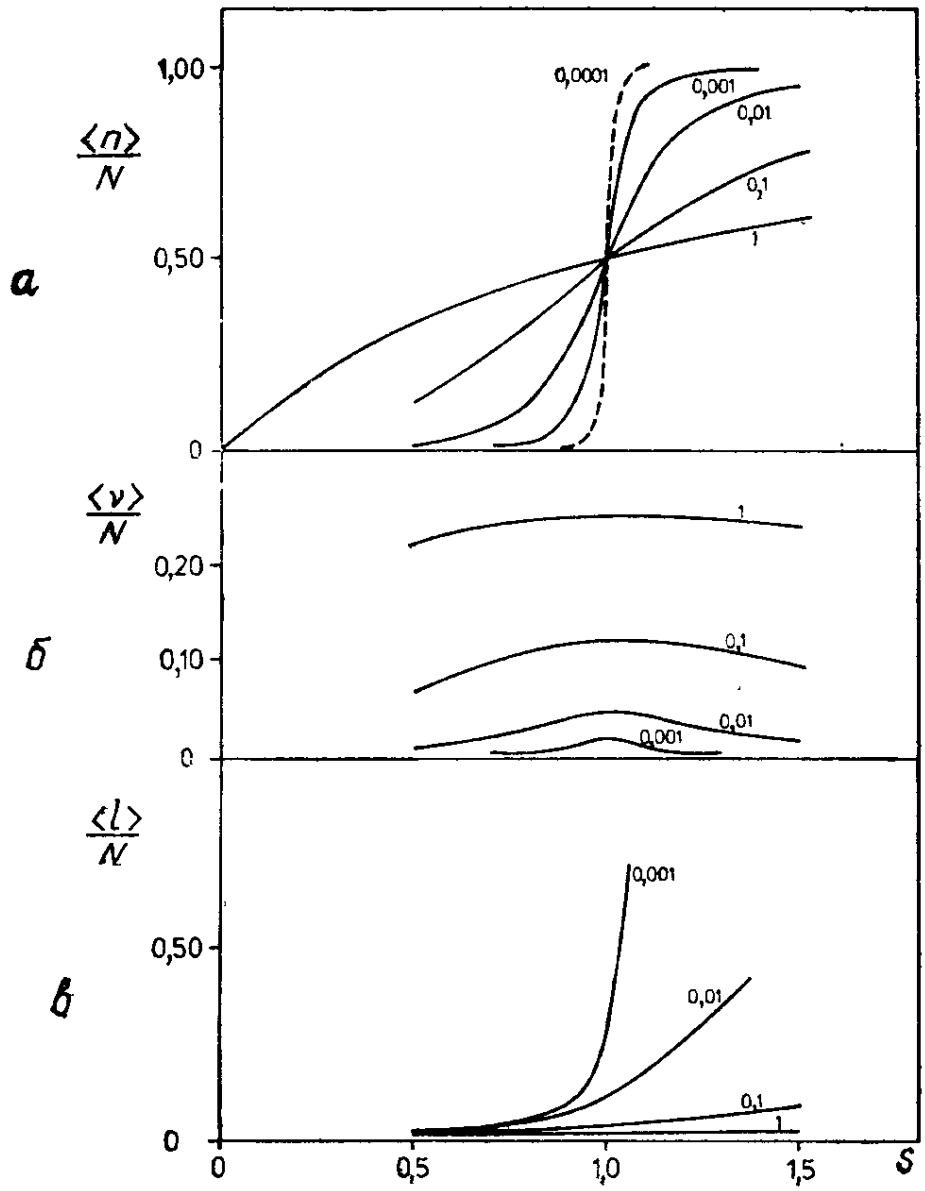


Рис. А.1.

Модель Зимма — Брэгга перехода спираль — клубок в полипептидной цепи длиной N. Рядом с кривыми указаны значения σ.

а — средняя доля спиралей как функция относительного статистического веса s. *б* — отношение среднего числа спиральных сегментов к N как функция s. *в* — отношение средней длины спиральных сегментов к N как функция s.

$$\langle l \rangle = \frac{\langle n \rangle}{\langle \nu \rangle} = \frac{\lambda_1}{1 - \lambda_2}$$

$$\langle l \rangle_{s=1} = \frac{1 + \sqrt{\sigma}}{\sqrt{\sigma}} \sim \frac{1}{\sqrt{\sigma}} \qquad \text{(A.17)}$$

Как видно из рис. А.1, *в*, длина спирали монотонно растет с увеличением s.

Необходимо отметить, что разделение цепи на несколько сегментов спиральной и клубковой конформаций согласуется с более общим рассмотрением Ландау и Лифшица [791]. Авторы утверждают, что в одномерном случае разделение на две отдельные фазы невозмож-

но, если только поверхностное натяжение небесконечно. Бесконечное поверхностное натяжение отвечает состоянию с $\sigma = 0$, когда к разделению на две фазы приводит также и модель Зимма — Брэгга.

Связь между s и температурой

Вблизи точки перехода относительный статистический вес линейно связан с температурой. Эта связь важна, так как переходы спираль — клубок часто наблюдаются при изменении температуры. Ее можно получить, если вновь обратиться к уравнению (А.4), но свести его к двум состояниям

$$\frac{1}{s}\frac{\partial s}{\partial T} = \frac{\partial \ln s}{\partial T} = \frac{\partial \ln z^{\alpha}}{\partial T} - \frac{\partial \ln z^{c}}{\partial T} = \frac{1}{z^{\alpha}}\frac{\partial z^{\alpha}}{\partial T} - \frac{1}{z^{c}}\frac{\partial z^{c}}{\partial T} =$$

$$= \frac{\int\limits_{\alpha} \frac{E}{RT^2}\exp\left(\frac{-E}{RT}\right)d\varphi d\psi d\chi}{\int\limits_{\alpha}\exp\left(\frac{-E}{RT}\right)d\varphi d\psi d\chi} - \frac{\int\limits_{\text{клубок}} \frac{E}{RT^2}\exp\left(\frac{-E}{RT}\right)d\varphi d\psi d\chi}{\int\limits_{\text{клубок}}\exp\left(\frac{-E}{RT}\right)d\varphi d\psi d\psi} =$$

$$= \frac{\langle E\rangle_{\alpha\text{ - область}}}{RT^2} - \frac{\langle E\rangle_{\text{клубок}}}{RT^2} \equiv \frac{\Delta H}{RT^2} \qquad \text{(А.18)}$$

Таким образом, в области перехода ($s \simeq 1$) относительное изменение s в зависимости от T приблизительно постоянно, т. е. изменение s приблизительно пропорционально изменению температуры (рис. А.2).

Температура перехода

Температура перехода приблизительно пропорциональна изменению энтальпии. Переход происходит при $s = 1$, то есть при $z^{\alpha} = z^{c}$ или

$$\int\limits_{\alpha}\exp\left(\frac{-E}{RT}\right)d\phi d\psi d\chi = \int\limits_{\text{клубок}}\exp\left(\frac{-E}{RT}\right)d\phi d\psi d\chi$$

Если принять, что $E = E_{\alpha} = \text{const}$ по всей α-области размера Ω_{α} и $E = E_{c} = \text{const}$ по всей области клубка размера Ω_{c}, то

$$\Omega_{\alpha}\exp\left(\frac{-E_{\alpha}}{RT}\right) = \Omega_{c}\exp\left(\frac{-E_{c}}{RT}\right)$$

$$T_{s=1} \equiv \theta = \frac{-(E_{\alpha} - E_{c})}{R\ln\frac{\Omega_{c}}{\Omega_{\alpha}}} = \frac{-\Delta H}{R\ln\frac{\Omega_{c}}{\Omega_{\alpha}}} \qquad \text{(А.19)}$$

Переходы происходят лишь в тех случаях, когда состояние с более низкой энергией занимает меньшую конформационную область. При $\Omega_c > \Omega_\alpha$ температура перехода становится отрицательной, если энергия α-конформации менее выгодна, чем энергия конформации клубка, $E_\alpha > E_c$. Кроме того, переходы совершаются

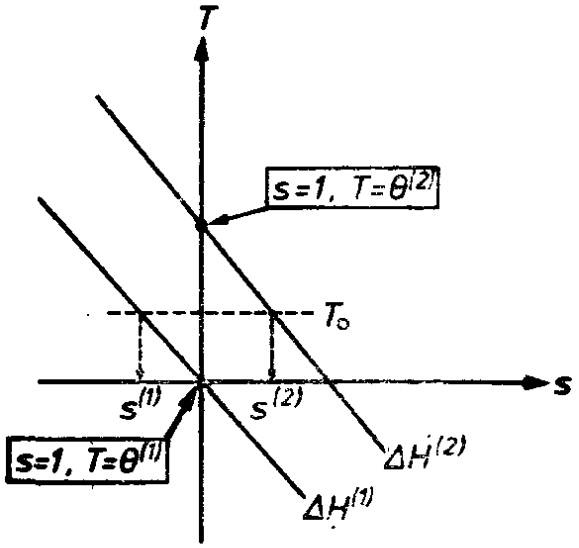


Рис. А.2.

Связь между относительным статистическим весом s, температурой T и разницей энтальпий ΔH вблизи точки перехода ($s = 1$, $T = \theta$) для двух различных полимеров.

Это — графическое представление уравнений (А.18) и (А.19). Обе величины $\Delta H^{(1)}$ и $\Delta H^{(2)}$ отрицательны и $\Delta H^{(2)} < \Delta H^{(1)}$. Различие температур перехода $\theta^{(2)} - \theta^{(1)}$ пропорционально разности $\Delta H^{(2)} - \Delta H^{(1)}$. При данной температуре T_0 оба полимера имеют различные относительные статистические веса $s^{(1)}$ и $s^{(2)}$. Эти значения s определяют склонности полимеров к образованию спиралн при данной температуре.

в том случае, когда состояние с меньшей энергией E (более прочное связывание) занимает меньшую конформационную область Ω. В нашем примере, когда $\Omega_c > \Omega_\alpha$ (рис. 2.3), спиральная конформация принимается, только если она энергетически выгодна. Направление перехода таково, что при температурах ниже θ превалирует состояние с меньшей энергией, а при температурах выше θ — состояние с большей конформационной областью (рис. А.2). Чем выше температура перехода θ, тем меньше различие конформационных областей и тем больше энергетически выгодных спиральных конформаций.

А.5 Сравнение с экспериментальными данными

Кривые перехода спираль — клубок и модель Зимма — Брэгга

Экспериментальные кривые перехода спираль — клубок описываются с помощью модели Зимма — Брэгга. Для апробации модели Зимма — Брэгга теоретические кривые перехода спираль — клубок (уравнение (А.15)), показанные на рис. А.1, *а*, следует сравнить

с экспериментальными. Если содержание спирали $\langle n \rangle$ измерять, как функцию температуры T, получаются сигмоидные кривые, аналогичные показанным на рис. А.1, *а*. Для сравнения температуру T нужно перевести в относительный статистический вес s. В области перехода T приблизительно пропорциональна s (разд. А.4 и рис. А.2). Если известна разность энтальпий ΔH перехода, T можно заменить на s, пользуясь уравнением (А.18). Затем полученную экспериментальную кривую перехода $\langle n \rangle$ как функцию s можно сопоставить с теоретической кривой с соответствующим параметром нуклеации σ. В большинстве случаев получается вполне удовлетворительное совпадение, что и подтверждает модель Зимма — Брэгга.

Приближеиные значения σ для синтетических полимеров

В случае перехода спираль — клубок, индуцируемого изменением температуры, параметр нуклеации можно оценить по наклону кривой в точке перегиба. Как показано выше, экспериментальную кривую индуцируемого изменением температуры перехода спираль — клубок можно сопоставить с теоретической кривой, если известна величина ΔH. Это сопоставление позволяет получить параметр нуклеации. Однако во многих случаях величина ΔH бывает неизвестна. Тем не менее оценить величину σ можно. Согласно уравнению (А.15) при учете (А.10) наклон кривой перехода в точке перегиба ($s = 1$) должен быть равен

$$\frac{1}{N}\frac{\partial \langle n \rangle}{\partial s}\bigg|_{s=1} = \frac{1}{4s\sigma} = \frac{1}{4\sigma} \qquad \text{(А.20)}$$

Величина ΔH обычно имеет порядок 1 ккал/моль. Пользуясь этим значением, обратим экспериментальную величину $\partial \langle n \rangle / \partial T$ в $\partial \langle n \rangle / \partial s$ (рис. А.2) и по уравнению (А.20) получим значения σ. Наблюдаемые величины σ для синтетических полимеров имеют порядок 10^{-4}. Это означает, что каждое сочленение между сегментами спирали и клубка понижает статистический вес соответствующей конформации цепи в 100 раз (уравнение (А.6)). При температуре перехода средняя длина спиральных сегментов составляет 100 остатков (уравнение (А.17)). Отметим, что эти значения получены в предположении очень больших величин N.

Температура перехода и склонность к спирали

Температуры переходов синтетических полимеров преобразуются в относительные статистические веса при данной температуре. Зависимость между s, T и ΔH вблизи точки перехода показана на рис. А.2. В первом приближении она линейна (разд А.4 с. 300). Сопоставим эти зависимости у двух различных гомополимеров с разностями энтальпий $\Delta H^{(1)}$ и $\Delta H^{(2)}$ и температурами перехода

$\theta^{(1)}$ и $\theta^{(2)}$. Если значения ΔH и θ достаточно близки, обе линейные зависимости можно объединить на одном графике (рис. А.2). График можно использовать для получения относительных статистических весов $s^{(1)}$ и $s^{(2)}$ для обоих типов гомополимеров при данной температуре T_0. Статистические веса отражают склонность полимеров к спиральной конформации (уравнения (А.18) и (А.4)) при данной температуре.

Обычно значение ΔH неизвестно. Однако его можно оценить по соответствующим значениям θ, используя уравнения (А.19). Таким образом, всякий раз, когда удается измерить θ для гомополимера данного остатка*, его значение можно использовать для приблизительной оценки склонности к спиральной конформации.

Вклад растворителя

Начиная с уравнения (А.1), влиянием растворителя во всех формулах ради простоты пренебрегли. При строгом решении задачи следовало бы учитывать трансляционные и вращательные параметры окружающих цепь молекул растворителя. Однако поскольку на практике это сделать невозможно, вклад растворителя всегда рассматривается отдельно.

Растворитель может индуцировать переходы спираль — клубок без температурных изменений. Кроме того, в зависимости от растворителей величина ΔH (уравнение (А.19)) может быть и положительной [789]; в результате $\Omega_c < \Omega_\alpha$, и спираль образуется и при высоких температурах. В этом случае величины Ω_c и Ω_α включают эффекты растворителя и их уже нельзя получить с помощью рис. 2.3.

* Гомополимеры многих остатков явно не образуют спирали. В таких случаях можно использовать смешаиные полимеры с остатками даиного типа в качестве «гостя» и с хорошим спиралеобразователем (например, поли-γ-бензил-L-глутаматом) в качестве «хозяина». Анализируя изменения θ, можно затем получить температуру перехода для остатка «гостя».

Литература

1. *Kühne W.*, Über das Verhalten verschiedener organisirter und sogenannter ungeformter Fermente. Über das Trypsin (Enzym des Pankreas). Verhandlungen des Heidelb. Naturhist.-Med. Vereins N. S. I, 2—5 (1976). Reproduced in FEBS Lett. **62**, E3 (1976).
2. *Hoppe-Seyler F.*, Über die chemischen und optischen Eigenschaften des Blutfarbstoffs. Virchows Arch. **29**, 233 (1864).
3. *Calvin M.*, Chemical Evolution, Clarendon Press, Oxford, 1969.
4. *Crick F. H. C.*, The origin of the genetic code. J. Mol. Biol. **38**, 367 (1968).
5. *Orgel L. E.*, Evolution of the genetic apparates. J. Mol. Biol. **38**, 381 (1968).
6. *Eigen M.*, Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften **58**, 465 (1971).
7. *Tischendorf G. W., Zeichardt H., Stöffer G.*, Architecture of the *Escherichia coli* ribosome as determined by immune electron microscopy. Proc. Natl. Acad. Sci. USA **72**, 4820 (1975).
8. *Miller S. L.*, Production of some organic compounds under possible primitive Earth conditions. J. Amer. Chem. Soc. **77**, 2351 (1954).
9. *Bonner W. A., van Dort M. A., Yearian M. R.*, Asymmetric degradation of DL-leucin with longitudinally polarised electrons. Nature **258**, 419 (1975).
10. *Tse-Fei Wong J.*, A co-evolution theory of the genetic code. Proc. Natl. Acad. Sci. USA **72**, 1909 (1975).
11. *Richards K. E., Guilley H., Jonard G., Hirth L.*, A specifically encapsidated fragment from RNA of tobacco mosaic virus: Sequence homology with the coat protein cistron. FEBS Lett. **43**, 31 (1974).
12. *Sanger F., Air G. M., Barrell B. G., Brown N. L., Coulson A. R., Fiddes J. C., Hutchison III C. A., Slocombe P. M.., Smith M.*, Nucleotide sequence of bacteriophage ϕX174 DNA. Nature **265**, 687 (1977).
13. *Fiers W., Contreras R., Duerinck F., Haegeman G., Inserentant D., Merregaert J., Min Jou W., Molemans F., Raeymaekers A., Van den Berghe A., Volckaert G., Ysebaert M.*, Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. Nature **260**, 500 (1976).
14. *Barrell B. G., Air G. M., Hitchison III C. A.*, Overlapping genes in bacteriophage ϕX174. Nature **264**, 34 (1976).
15. *Kauzmann W.*, Some foctors in the interpretation of protein denaturation. Adv. Prot. Chem. **14**, 1 (1959).
16. *Nozaki Y., Tanford C.*, The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. J. Biol. Chem. **246**, 2211 (1971).
17. *Chothia C.*, Hydrophobic bonding and accesible surface area in proteins. Nature **248**, 338 (1974).
18. *Birktoft J. J., Blow D. M.*, Structure of crystalline α-chymotrypsin. J. Mol. Biol. **68**, 187 (1972).
19. *Ленинджер А.* Биохимия. — М.: Мир, 1976

20. *Dayhoff M. O.*, Atlas of Protein Sequence and Structure, Natl. Biom. Res. Foundation, Washington D. C., 1972.
21. IUPAC — IUB Comission on Biochemical Nomenclature 1969, Abbreviations and symbols for the descriptions of the conformations of polypeptide chains. Biochemistry **9**, 3471 (1970).
22. *Уотсон Дж.* Молекулярная биология гена. — М.: Мир, 1978.
23. *Полинг Л.* Природа химической связи. — М.: Гос. изд. хим. лит., 1947 (перевод 1-го изд.).
24. *Lehninger A. L.*, Bioenergetics, W. A. Benjamin, Menlo Park, California, 1971, p. 173.
25. *Pauling L., Corey R. B., Branson H. R.*, The Structure of proteins: Two hydrogen-bonded helical configurations of the polipeptide chain. Proc. Acad. Sci. USA **37**, 205 (1951).
26. Tables of Interatomic Distances in Molecules and Ions, Suppl. Spec. Publ. No. 18, The Chemical Society, London, 1965.
27. *Edsall J. T., Flory P. J., Kendrew J. C., Liquori A. M., Némethy G., Remachandran G. N., Scheraga H. A.*, A proposal of standard conventions and nomenclature for the description-of polypeptide conformations. J. Mol. Biol. **15**, 399 (1966).
28. *Ramachandran G. N., Ramakrishnan C., Sasisekharan V.*, Stereochemistry of polypeptide chain configurations. J. Mol. Biol. **7**, 95 (1963).
29. *Ramachandran G. N., Sasisekharan V.*, Conformation of polypeptides and proteins. Adv. Prot. Chem. **23**, 283 (1968).
30. *Lavitt M.*, A simplified representation of protein conformations for rapid simulation of protein folding. J. Mol. Biol. **104**, 59 (1976).
31. *Burgess A. W., Ponnuswamy P. K., Scheraga H. A.*, Analysis of conformations of amino acid residues and prediction of backbone topography in protein. Israel J. Chem. **12**, 239 (1974).
32. *Brant D. A., Miller W. G., Flory P. J.*, Conformational energy estimates for statistically coiling polypeptide chains. J. Mol. Biol. **23**, 47 (1967).
33. *Németry G., Scheraga H. A.*, Protein folding. Duart. Rew. Biophys. **10**, 239 (1977).
34. *Ramachandran G. N., Venkatachalam C. M., Krimm S.*, Stereochemical criteria for polypeptide and protein chain conformations. Biophys. J. **6**, 849 (1966).
35. *Gieren A., Dederer B., Schanda F.*, Some new aspects concerning conformation of polypeptide chains in proteins. Z. Naturforsch. in press (1978).
36. *Winkler F. K., Dunitz J. D.*, The non-polar amide group. J. Mol. Biol. **59**, 169, (1971).
37. *Wieland T., Birr C.*, Homodetic cyclic peptides. International Review of Science, Ordanic Chemistry Series Two, Vol. 6, p. 183, Butterworth, London, 1976.
38. *Karle I. L., Karle J., Wieland T., Burgermeister W., Faulstich H., Witkop B.*, Conformations of the Li-antamanide complex and Na-(Phe^4, Val^6)antamanide complex in the crystalline state. Proc. Nat. Acad. Sci. USA **70**, 1836 (1973).
39. *Wyckoff H. W., Tsernoglou D., Hanson A. W., Knox J. B., Lee B., Richards F. M.*, The three-dimensional structure of ribonuclease-S. J. Biol. Chem. **245**, 305 (1970).
40. *Schubert-Wright C., Alden R. A., Kraut J.*, Structure of subtilisin BPN' at 2,5 Å resolution. Nature **221**, 235 (1969).
41. *Cahn R. S., Ingold. C. K., Prelog V.*, The specification of asymmetric configuration in organic chemistry. Experentia **12**, 81 (1956).
42. *Epp O., Lattman E. E., Schiffer M., Huber R., Palm W.*, The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2,0 Å resolution. Biochemistry **14**, 4943 (1975).

43. *Quiocho F. A., Lipscomb W. N.*, Carboxypeptidase A: A protein and an enzyme. Adv. Chem. **25**, 1 (1971).
44. *Fischer E.*, Einfluß der Configuration auf die Wirkung der Enzyme. Chem. Ber. **27**, 2985 (1894).
45. *London F.*, Über einige Eigenschaften und Anwendungen der Molekularkräfte. Z. Phys. Chem. Abt. B **11**, 222 (1930).
46. *Slater J. C., Kirkwood J. G.*, The van der Waals forces in gases. Phys. Rev. **37**, 682 (1931).
47. *Jones J. E.*, On the determination of molecular fields. Proc. Roy. Soc. London, Ser. **A106**, 441 (1924).
48. *Slater J. C.*, The normal state of helium. Phys. Rev. **32**, 349 (1928).
49. *Buckingham R. A.*, The classical equation of state of gaseous helium, neon and argon. Proc. Roy. Soc. London, Ser. **A 168**, 264 (1938).
50. Handbook of Chemistry and Physics, 41st ed., Chemical Rubber, Cleveland, Ohio, 1959, p. 2523.
51. *Brant D. A.*, Conformational analysis of biopolimers: Conformational energy calculation. Annu. Rev. Biophys. Bioeng. **1**, 369 (1972).
52. *Momany F. A., McGuire R. F., Burgess A. W., Scheraga H. A.*, Energy parameters in polipeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, and intrinsic torsional potentials for the naturally occurring amino acids. J. Phys. Chem. **79**, 2361 (1975).
53. *Haglar A. T., Lapiccirela A.*, Spatial electron distribution and population analysis of amides, carboxylic acids, and peptides, and their relation to empirical potential functions. Biopolymers **15**, 1167 (1976).
54. *Hagler A. T., Huler E., Lifson S.*, Energy functions for peptides and proteins. I. Derivation of a consistent force fields including the hydrogen bond from amide crystals. J. Am. Chem. Soc. **96**, 5319 (1974).
55. *Levitt M.*, Energy refinement of hen egg-white lysozyme. J. Mol. Biol. **82**, 393 (1974).
56. *Jeziorski B., van Hermet M.*, Variation — perturbation treatment of the hydrogen bond between water molecules. Mol. Phys. **31**, 713 (1976).
57. *Pauling L.*, The Nature of the Chemical Bond, Cornell University Press, Ithaca, New York, 1960, p. 461.
58. *Pimentel G. C., McClellan A. L.*, The Hydrogen Bond, Freeman Sc Co, London, 1960, pp. 242, 282.
59. *Moews P. C., Kretsinger R. H.*, Refinement of the structure of carp muscle calcium-binding parvalbumin by model building and difference Fourier analysis. J. Mol. Biol. **91**, 201 (1975).
60. *Perutz M. F.*, Stereochemistry of cooperative effects of haemoglobin. Nature **228**, 726 (1970).
61. *Kretsinger R. H., Nockolds C. E.*, Carp muscle calcium-binding protein. J. Biol. Chem. **248**, 3313 (1973).
62. *Bode W., Schwager P.*, The refined crystal structure of bovine β-trypsin at 1,8 Å resolution. J. Mol. Biol. **98**, 693 (1975).
63. *Richards F. M.*, The interpretation of protein structures: Totat volume, group volume distributions and packing density. J. Mol. Biol. **82**, 1 (1974).
64. *Richards F. M.*, Areas, volumes, packing and protein structure. Annu. Rev. Biophys. Bioeng. **6**, 151 (1977).
65. *Brandts J. F., Oliveira R. J., Westort C.*, Thermodinamics of protein denaturation. Effect of pressure on the denaturation of ribonuclease A. Biochemistry **9**, 1038 (1970).
66. *d'Ans-Lax*, Taschenbuch für Chemiker und Physiker, Vol. 1. Springer-Verlag, Berlin, 1967.
67. *Wyckoff H. W.*, Compensating nature of substitutions in pancreatic ribonucleases. Brookhaven Symp. Biol. **21**, 252 (1968).
68. *Finney J. L.*, Volume occupation, environment and accessibility in proteins. The problem of the protein surface. J. Mol. Biol. **96**, 721 (1975).

69. *Momany F. A., Carruthers L. M., McGuire R. F., Scheraga H. A.*, Intermolecular potentials from crystal data. III. Determination of empirical potentials and application to the packing configurations and lattice energies in crystals of hydrocarbons, carboxilic acids, amines and amides. J. Phys. Chem. **78**, 1595 (1974).
70. *Lifson S., Warchel A.*, Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkana and *n*-alkane molecules. J. Chem. Phys. **49**, 5116 (1968).
71. *Warshel A., Levitt M., Lifson S.*, Consistent force field for calculation of vibrational spectra and conformations of some amides and lactam rings. J. Mol. Spectrosc. **33**, 84 (1970).
72. *Adman E., Watenpaugh K. D., Jensen L. H.*, NH—S hydrogen bonds in *Peptococcus aerogenes* ferredoxin, *Chostridium pasteurianum* rubredoxin, and chromatium high potential iron protein. Proc. Nat. Acad. Sci. USA **72**, 4854 (1975).
73. *Bondi A.*, Wan der Waals volumes and radii. J. Phys. Chem. **68**, 441 (1964).
74. *Benzer S.* The Chemical Basis of Heredity, McElroy W. D., Glass B., Eds. John Hopkins University Press, Baltimore, 1957, pp. 70—93.
75. *Edelman G. M., Cunningham B. A., Gall W. E., Gottlieb P. D., Rutishauser U. Waxdal M. J.*, The covalent structure of an entire G immunoglobulin molecule. Proc Nat. Acad. Sci. USA **63**, 78 (1969).
76. *Kirschner K., Bisswanger H.*, Multifunctional proteins. Annu. Rev. Biochem. **45**, 143 (1976).
77. *Goldberg W. E.*, Tertiary structure of *Escherichia coli* β-D-galactosidase. J. Mol. Biol. **46**, 441 (1969).
78. *Gilbert W.*, Why genes in pieces? (editorial). Nature **271**, 501 (1978).
79. *Davidson E. H., Britten R. J.*, Organization, transcription and regulation in the animal genome. Quart. Rev. Biol. **48**, 565 (1973).
80. *Stammers D. K., Muirhead H.*, Three-dimensional structure of cat muscle pyruvate kinase at 3.1 Å resolution. J. Mol. Biol. **112**, 309 (1977).
81. *Klotz I. M., Darnall D. W., Langerman N. R.*, Quaternary structure of proteins. In: The Proteins (H. Neurath, Ed.), 3rd., Vol., I, Academic Press, New York, 1975, pp. 293—411.
82. *Brown J. R.*, Structural origins of mammalian albumin. Fed. Proc. 35, 2141 (1976).
83. *Noda L.*, Adenylate kinase. The Enzymes **8**, 279 (1973).
84. *von Zabern I., Wittmann-Liebold B., Untuch-Grau R., Schirmer R. H., Pai E.*, Primary and tertiary structure of the principal human adenylate kinase. Eur. J. Biochem. **68**, 281 (1976).
85. *Stryer L.*, Biochemistry. W. H. Freeman, San Francisco, 1975.
86. *Peterson P. A., Rask L., Sege K., Klareskog L., Anundi H., Ostberg L.*, Evolutionary relationship between immunoglobulins and transplantation antigens. Proc. Nat. Acad. Sci. USA **72**, 1612 (1975).
87. *Bach F. H.*, Genetics of transplantation: The major histocompatibility complex. Annu. Rev. Genet. **10**, 281 (1976).
88. *Peterson P. A., Cunningham B. A., Berggård I., Edelman G. M.*, β_2-*Microglobulin* — A free immunoglobulin domain. Proc. Nat. Acad. Sci. USA **69**, 1697 (1972).
89. *Frieden E.*, Protein — protein interaction and enzymatic activity. Annu. Rev. Biochem. **40**, 653 (1971).
90. *Chan W. H. C.*, Relationship between quaternary structure and enzyme activity. Trends Biol. Sci., **1**, 258 (1976).
91. *Rossmann M. G., Liljas A., Bränden C.-I., Banaszak L. J.*, Evolutionary and structural relationships among dehydrogenases. The Enzymes **11**, 61(1975).
92. *Monod J., Wyman J., Changeux J.-P.*, On the nature o allosteric transitions: A plausible model. J. Mol. Biol. **12**, 88 (1965).

93. *Koshland D. E., Jr.*, The evolution of function in enzymes. Fed. Proc. **35**, 2104 (1976).
94. *Anfinsen C. B., Scheraga H. A.*, Experimental and theoretical aspects of protein folding. Adv. Chem. **29**, 205 (1975).
95. *Sela M., White F. H., Anfinsen C. B.*, Reductive cleavage of disulfide bridges in ribonuclease. Science **125**, 691 (1957).
96. *White F. H., Jr.*, Regeneration of native secondary and tertiary structures by air oxidation of reduced ribonuclease. J. Biol. Chem. **236**, 1353 (1961).
97. *Anfinsen C. B., Haber E.*, Studies on the reduction and reformation of protein disulfide bonds. J. Biol. Chem. **236**, 1361 (1961).
98. *Olins D. E., Edelman G. M.* Reconstitution of 7S molecules from L and H polypeptide chains of antibodies and γ-globulins. J. Exp. Med. **119**, 789 (1964).
99. *Freedman M. H., Sela M.*, Recovery of antigenic activity upon reoxydation of completery reduced polyalanyl rabbit immunoglobulin G. J. Biol. Chem. **241**, 2383 (1966).
100. *Petersen J. G. L., Dorrington K. I.*, An *in vitro* system for studies the kinetics of interchain disulfide bond formation in immunoglobulin G. J. Biol. Chem. **249**, 5633 (1974).
101. *Givol D., DeLorenzo F., Goldberger R. F., Anfinsen C. B.*, Disulfide interchange and the three-dimensional structure of proteins. Proc. Nat. Acad. Sci. USA **53**, 676 (1965).
102. *Della Corte E., Parkhouse R. M. E.*, Biosynthesis of immunoglobulin (IgA) and immunoglobulin M (IgM). Requirement for J-chain and a disulfide-exchanging enzyme for polymerization. Biochem. J. **136**, 597 (1973).
103 *Ziegler D. M., Poulsen L.L.*, Protein disulfide bond synthesis, a possible intracellular mechanism. Thends Biol. Sci. **2**, 79 (1977).
104. *Venetianer P., Straub F. B.*, The enzymic reactivation of reduced ribonuclease. Biochim. Biophys. Acta **67**, 166 (1963).
105. *Goldberger R. F., Epstein C. J., Anfinsen C. B.*, Acceleration of reactivation of reduced bovine pancreatic ribonuclease by a microsomal system from rat liver. J. Biol. Chem. **238**, 628 (1963).
106. *DeLorenzo F., Goldberger R. F., Steers E., Givol D., Anfinsen C. B.*, Purification and properties of an enzyme from beef liver which catalyzes sulfhydryl-disulfide interchange in proteins. J. Biol. Chem. **241**, 1562 (1966).
107. *Fuchs S., DeLorenzo F., Anfinsen C. B.*, Studies on the mechanism of the enzymic catalysis of disulfide interchange in proteins. J. Biol. Chem. **242**, 398 (1967).
108. *Flory P. J.*, Principles of Polymer Chemistry. Cornell University Press, Ithaca, New York, 1953. Chap. 14.
109. *Straub F. B.*, SH-groups and SS-bridges in the structure of enzymes. Proc. 7th Int. Congress Biochem. (Tokyo), 42 (1967).
110. *Wall J. S.*, Disulfide bonds: Determination location and influence on molecular properties of proteins. J. Agr. Food Chem. **19**, 619 (1971).
111. *Williams C. H., Jr.*, Flavin-containing dehydrogenases. The Enzymes (3rd ed.) **13**, 89 (1976).
112. *Hatfield G. W., Burns R. O.*, Threonine deaminase from *Salmonella typhimurium*. J. Biol. Chem. **245**, 787 (1970).
113. *Pontremoli S., Horecker B. L.*, Fructose-1,6-diphosphatases. The Enzymes (3rd ed.) **4**, 611 (1971).
114. *Ernest M. J., Kim K. H.*, Regulation of rat liver glycogen synthetase D. J. Biol. Chem. **249**, 5011 (1974).
115. *Ernest M. J., Kim K. H.*, Regulation of the rat liver glycogen synthetase. J. Biol. Chem. **248**, 1550 (1973).
116. *DeLorenzo F., Fuchs S., Anfinsen C. B.*, Characterization of a peptide fragment containing the essential half-cystine residue of a microsomal disulfide interchange enzyme. Biochemistry **5**, 3961 (1966).

117. *Sakai H.*, A ribonucleoprotein which catalyzys thiol-disulfide exchange in the sea urchin egg. J. Biol. Chem. **242**, 1458 (1967).
118. *Cuatrecasas P.*, Proc. Symp. Insulin Action, Toronto, Canada, Academic Press, New York, 1972, p. 137.
119. *Czech M.*, Molecular basis of insulin action. Annu. Rev. Biochem. **46**, 359 (1977).
120. *Bhakdi S., Knüfermann H., Schmidt-Ullrich R., Fischer H., Wallach D. F.*, Interaction between erythrocyte membrane proteins and complement components. Biochim. Biophys. Acta **363**, 39 (1974).
121. *Morino Y., Snell E. E.*, The subunit structure of tryptophanase. J. Biol. Chem **242**, 5602 (1967).
122. *Hatfield G. W., Burns R. O.*, Specific binding of leucyl transfer RNA to an immature form of L-threonine deaminase: Its implications in repression. Proc. Nat. Acad. Sci. USA **66**, 1027 (1970).
123. *Jones E. T., Williams C. H., Jr.*, The sequence of amino acid residues around the oxidation — reduction active disulfide in yeast glutathione reductase. J. Biol. Chem. **250**, 3779 (1975).
124. *Schulz. G. E., Schirmer R. H., Sachsenheimer W., Pai E. F.*, The structure of the flavoenzyme glutathione reductase. Nature **273**, 120 (1978).
125. *Bornstein P.*, The biosynthesis of collagen. Annu. Rev. Biochem. **43**, 567 (1974).
126. *Tanzer M. L.*, The biological diversity of collagenous proteins. Trends Biochem. Sci. **3**, 15 (1978).
127. *Uitto J., Lichtenstein J. R.*, Defects in the biochemistry of collagen in diseases of connective tissue. J. Invest. Dermatol. **66**, 59 (1976).
128. *Byers P. H., Click E. M., Harper E., Bornstein P.*, Interchain disulfide bonds on procollagen are located in a large nontriple-helical COOH-terminal domain. Proc. Nat. Acad. Sci. USA **72**, 3009 (1975).
129. *Veis A., Brownell A. C.*, Collagen biosynthesis. CRC Crit. Rev. Biochem. 2, 417 (1974).
130. *Tanzer M. L., Church R. L., Yaeger J. A., Wampler D. E., Park E. D.*, Procollagen: Intermediate forms containing several types of peptide chains and non-collagen peptide extensions at NH_2 and COOH ends. Proc. Nat. Acad. USA **71**, 3009 (1974).
131. *Martin G. R., Byers P. H., Piez K. A.*, Procollagen. Adv. Enzymol. **42**, **167 (1975).**
132. *Narita K., Matsuo H., Nakajima T.*, End group determination. In: Protein Sequence Determination (S. B. Needleman, Ed.), 2nd ed., Springer-Verlag, Heidelberg, 1975, pp. 30—103.
133. *Jones G. H.*, Cell-free synthesis of amino-terminal L-pyroglutamic acid. Biochemistry **13**, 854 (1974).
134. *Burstein Y., Kantor F., Schechter I.*, Partial amino-acid sequence of the precursor of an immunoglobulin light chain containing NH_2-terminal pyroglutamic acid. Proc. Nat. Acad. Sci. USA **73**, 2604 (1976).
135. *Chen R., Brosius J., Wittmann-Liebold B.*, Occurrence of methylated amino acids as N-termini of proteins from *Escherichia coli*. ribosomes. J. Mol. Biol. **111**, 173 (1977).
136. *Gillen R. G., Riggs A.*, Structure and function of the isolated hemoglobins of the American eel, *anguilla rostrata*. J. Biol. Chem. **248**, 1961 (1973).
137. *Ottesen M.*, Induction of biological activity by limited proteolysis. Annu. Rev. Biochem. **36**, 55 (1967).
138. *Neurath H., Walsh K. A.*, Role of proteolytic enzymes in biological regulation (a review). Proc. Nat. Acad. Sci. USA **73**, 3825 (1976).
139. *Haenn de C., Neurath H., Teller D. C.*, The phylogeny of trypsin-related serine proteases and their zymogens. New methods for the investigation of distant evolutionary relationships. J. Mol. Biol. **92**, 225 (1975).

140. *Hershko A., Fry M.*, Post-translational cleavage of polypeptide chains: Role in assembly. Annu. Rev. Biochem. **44**, 775 (1975).
141. *Apte B. N., Zipser D.* Authors' statement on polypeptide splicing. Proc. Nat. Acad. Sci. USA **73**, 661 (1976).
142. *Hendrix R. W., Casjens S. R.*, Protein fusion: A novel reaction in bacteriphage λ head essembly. Proc. Nat. Acad. Sci. USA **71**, 1451 (1974).
143. *Maroux S., Baratti J., Desnuelle P.*, Purification and specificity of porcine enterokinase. J. Biol. Chem. **246**, 5031 (1971).
144. *King J. L., Jukes T. H.*, Non-Derwinian evolution. Science **164**, 788 (1969).
145. *Dayhoff M. O.*, Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 2 and earlier volumes. Nat. Biom. Res. Foundation, Washington, D. C., 1976.
146. *Steiner D. F.*, Proteolytic processing in the biosynthesis of insulin and other proteins. Fed. Proc. **33**, (10) 2105 (1974).
147. *Müller-Eberhard H. J.*, Complement. Annu. Rev. Biochem. **44**, 697 (1975).
148. *Davie E. K., Fujikawa K.*, Basic mechanisms in blood coagulation. Annu. Biochem. **44**, 799 (1975).
149. *Blobel G., Dobberstein B.*, Transfer of proteins across membranes. J. Cell. Biol. **67**, 835 (1975).
150. *Milstein C., Brownless G. G., Harrison T. M., Mathews M. B.*, A possible precusor of immunoglobulin light chains. Nature New Biol. **239**, 117 (1972).
151. *Schechter J., McKean D. J., Guyer R., Terry W.*, Partial amino acid sequence of the precursor of immunoglobulin light chain programmed by messenger RNA *in vitro*. Science **188**, 160 (1974).
152. *Kemper B., Habener J. F., Mulligan R. C., Potts J. T., Rich A.*, Pre-proparathyroid harmone: A direct translational product of parathyroid messenger RNA. Proc. Nat. Acad. Sci. USA **71**, 3731 (1974).
153. *Davillers-Thiery A., Kindt T., Scheele G., Blobel G.*, Homology in amino terminal sequence of precursors to pancreatic secretory proteins. Proc. Nat. Acad. USA **72**, 5016 (1975).
154. *Goldberg A. L., St. John A. C.*, Intracellular protein degradation in mammalian and bacterial cells. Annu. Rev. Biochem. **45**, 747 (1976).
155. *Katunuma N.*, New intracellular proteases and their role in intracellular enzyme degradation. Trends Biol. Sci. **2**, 122 (1977).
156. *Scornik O. A., Botbol V.*, Role of changes in protein degradation in the growth of regenerating livers. J. Biol. Chem. **251**, 2891 (1976).
157. *Schimke R. T., Doyle D.*, Control of enzyme levels in animal tissues. Annu. Rev. Biochem. **39**, 929 (1970).
158. *Williams-Ashman H. G., Notides A. C., Pabalan S. S., Lorand L.*, Transamidase reactions involved in the enzymic coagulation of semen: Isolation of γ-glutamyl-ε-lysine dipeptide from clotted secretion protein of quinea pig seminal vesicle. Proc. Nat. Acad. Sci. USA **69**, 2322 (1972).
159. *Lorand L., Goton T.*, Fibrinoligase — The fibrin stabilizing factor system. Meth. Enzymol. **19**, 770 (1970).
160. *Decker K.*, Structure and synthesis of a flovoprotein with covalently bound FAD. Trends Biochem. Sci. **1**, 184 (1976).
161. *Acharya A. S., Taniuchi H.*, A study of renaturation of reduced hen egg white lysozyme. Enzymatically active intermediate formed during oxydation of the reduced protein. J. Biol. Chem. **251**, 6934 (1976).
162. *Acharya A. S., Taniuchi H.*, Formation of the four isomers of hen egg white lysozyme containing three disulfide bonds and one open disulfide bond. Proc. Nat. Acad. Sci. USA **74**, 2362 (1977).
163. *Jacobson M. F., Baltimore D.*, Polypeptide cleavages in the formation of poliovirus proteins. Proc. Nat. Acad. Sci. USA **61**, 77 (1968).
164. *Baltimore D.*, Polio is not dead. Perspect Virol. **7**, 1 (1971).
165. *Pappenheimer A. M., Jr.*, Diphtheria toxin. Annu. Rev. Biochem. **46**, 69 (1977).

166. *Holzer H.*, Catabolite inactivation in yeast. Trends Biol. Sci. **1**, 178 (1976).
167. *Tager H. S., Steiner D. F.*, Peptide hormones. Annu. Rev. Biochem. **43**, 509 (1974).
168. *Stadtman E. R.*, Adenylyl transfer reactions. The Enzymes **8**, 2 (1973).
169. *Hagaishi O., Ueda K.*, Poly (ADP-ribose) and ADP-ribosylation of proteins. Annu. Rev. Biochem. **46**, 95 (1977).
170. *Stenflo J., Suttie J. W.*, Vitamin K-dependent formation of γ-carboxyglutamic acid. Annu. Rev. Biochem. **46**, 157 (1977).
171. *Price P. A., Poser J. W., Raman N.*, Primary structure of the γ-carboxyglutamic acid-containing protein from bovine bone. Proc. Nat. Acad. Sci. USA **73**, 3374 (1976).
172. *Lunney J., Ashwell G.*, A hepatic receptor of avian origin capable of binding specifically modified glycoproteins. Proc. Nat. Acad. Sci. USA **73**, 341 (1976).
173. *Ashwell G., Morell A.*, The role of surface sarbohydrates in the hepatic recognition and transport of circulating glycoproteins. Adv. Enzymol. **41**, 99 (1974).
174. *Paik W. K., Kim S.*, Protein methylation: Chemical, enzymological and biological significance. Adv. Enzymol. **42**, 227 (1975).
175. *Rubin C. S., Rosen O. M.*, Protein phosphorylation. Annu. Rev. Biochem. **44**, 831 (1975).
176. *Linderström-Lang K. U., Schnellman J. A.*, Protein structure and enzyme activity. The Enzymes (P. D. Boyer, ed.), Ademic Press, New York, Yol. 1, 2nd ed., 1959, p. 443.
177. *Anfinsen C. B.*, Principles that govern the folding of protein chains. Science **181**, 223 (1973).
178. *Moorhouse R., Winter W. T., Arnott S.*, Conformation and molecular organization in fibers of the capsular polysaccharide from *Escherichia coli* M41 mutant. J. Mol. Biol. **109**, 373 (1977).
179. *Arnott S.*, The geometry of nucleic acids. Progr. Biophys. Mol. Biol. **21**, 265 (1970).
180. *Stubbs G., Warren S., Holmes K.*, Structure of RNA and RNA binding site in tobacco mosaic virus from 4 Å map calculated from X-ray fibre diargams. Nature **267**, 216 (1977).
181. *Cohen C., Harrison S. O., Stephens R. E.*, X-ray diffraction from microtubules. J. Mol. Biol. **59**, 375 (1971).
182. *Moody M. F.*, Sheath of Bacteriophage T4. J. Mol. Biol. **80**, 613 (1973).
183. *Perutz M. F.*, New X-ray evidence on the configuration of polypeptide chains. Nature **167**, 1053 (1951).
184. *Dickerson R. E.*, X-ray analysis and protein structure. The Proteins (H. Neurath, ed.), 2nd ed., Vol. 2, Academic Press, New York, 1964, p. 603.
185. *Kendrew J. C., Dickerson R. E., Strandberg B. E., Hart R. G., Davies D. R., Phillips D. C., Shore V. C.*, Structure of myoglobin. Nature **185**, 422 (1960).
186. *Shulz G. E., Elzinga M., Marx F., Schirmer R. H.*, Threedimensional structure of adenylate kinase. Nature **250**, 120 (1974).
187. *Donohue J.*, Hydrogen-bonded helical configurations of the polypeptide chain. Proc. Nat. Acad. Sci. USA **39**, 470 (1953).
188. *Hendrickson W. A., Love W. E.*, Structure of lamprey haemoglobin. Nature New Biol. **232**, 197 (1971).
189. *Hendrickson W. A., Love W. E., Karle J.*, Crystal structure analusis of sea lamprey hemoglobin at 2 Å resolution. J. Mol. Biol. **74**, 331—361 (1973).
190. *Low B. W., Grenville-Wells H. J.*, Generalized mathematical relationships for polypeptide chain helices. The coordinates of the π-helix. Proc. Nat. Acad. Sci. USA **39**, 785 (1953).
191. *Ramachandran G. N., Kartha G.*, Structure of collagen. Nature **176**, 593 (1955).
192. *Rich A., Crick F. H. C.*, The molecular structure of collagen. J. Mol. Biol. **3**, 483 (1961).

193. *Yonath A., Traub W.*, Polymers of tripeptides as collagen models. J. Mol Biol. **43**, 461 (1969).
194. *Traub W., Piez K. A.*, The chemistry and structure of collagen. Adv. Prot. Chem. **25**, 243 (1971).
195. *Bolian G., Click E. M., Hermodsen M., Bornstein P.*, Structure of rat skin collagen α1-CB8. Amino acid sequence of the hydroxylamine-produced fragment HA2. Biochemistry **11**, 3798 (1972).
196. *Fietzek P. P., Rexrodt F., Hopper K., Kühn K.*, The covalent structrue of collagen. Eur. J. Biochem. **38**, 396 (1973).
197. *Hulmes D. J. S., Miller A., Parry D. A. D., Piez K. A., Woodhead-Galloway J.*, Analysis of the primary structure of collagen for the origins of molecular packing. J. Mol. Biol. **79**, 137 (1973).
198. *Reid K. B. M., Porter R. R.*, Subunit composition and structure of subcomponent C1q of the first component of human complement. Biochem. J. **155**, 19 (1976).
199. *Venkatachalam C. M.*, Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. Biopolimers **6**, 1425 (1968).
200. *Crawford J. L., Lipscomb W. N., Schellman C. G.*, The reverse turn as a polypeptide conformation in globular proteins. Proc. Nat. Acad. Sci. USA **70**, 538 (1973).
201. *Chou P. Y., Fasman G. D.*, Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. Biochemistry **13**, 211 (1974).
202. *Lewis P. N., Momany F. A., Scheraga H. A.*, Chain reversals in proteins. Biochim. Biophys. Acta **303**, 211 (1973).
203. *Kuntz I. D.*, Protein-folding. J. Amer. Chem. Soc. **94**, 4009 (1972).
204. *Pauling L., Corey R. B.*, Configurations of polypeptide chains with favored orientations around single bonds: Two new pleated sheets. Proc. Nat. Acad. Sci. USA **37**, 729 (1951).
205. *Chothia C.*, Conformation of twisted β-pleated sheets in proteins. J. Mol. Biol. **75**, 295 (1973).
206. *Blake C. C. F., Catley S. J.*, Protein—DNA and protein—hormone interactions in prealbumin: A model of the thyroid hormone nuclear receptor? Nature **268**, 115 (1977).
207. *Carter C., Kraut J.*, A proposed model for the interaction of polypeptides with RNA. Proc. Nat. Acad. Sci. USA **71**, 283 (1974).
208. *Church G. M., Sussman J. L., Kim S. H.*, Secondary structural complementarity between DNA and proteins. Proc. Nat. Acad. Sci. USA **74**, 1458 (1977).
209. *Richardson J. S., Getzoff E. D., Richardson D. C.*, The β-bulge: A common small unit of non-repetitive protein structure. Proc. Nat. Acad. Sci. USA **75**, 2574 (1978).
210. *Crick F. H. C.*, The packing of α-helices: Simple coiled coils. Acta Crystallogr. **6**, 689 (1953).
211. *Fraser R. D. B., MacRae T. P.*, Structure of α-keratin. Nature **233**, 138 (1971).
212. *Parry D. A. D., Crewther W. G., Fraser R. D. B, MacRae T. P.*, Structure of the α-keratin: Structural implication of the amino acid sequences of the type I and type II chain segments. J. Mol. Biol. **113**, 449 (1977).
213. *Caspar D. L. D., Cohen C., Longley W.*, Tropomyosin: Crystal structure, polymorphism and molecular interactions. J. Mol. Biol. **41**, 87 (1969).
214. *Cohen C., Holmes K. C.*, X-ray diffraction evidence for α-helical coiled-coils in native muscle. J. Mol. Biol. **6**, 423 (1963).
215. *Huxley H. E.*, The mechanism of muscular contraction. Science **164**, 1356 (1969).
216. *Klotz I. M., Klippenstein G. L., Hendrickson W. A.*, Hemerythrin: Alternative oxygen carrier. Science **192**, 335 (1976).

217. *Stenkamp R. E., Sieker L. C., Jensen L. H., Loehr J. S.*, Structure of methemerythrin at 5 Å resolution. J. Mol. Biol. **100**, 23 (1976).
218. *Champness J. N., Bloomer A. C., Bricogne C., Butler P. J. G., Klug A.*, The structure of the protein disk of tobacco mosaic virus to 5 Å resolution. Nature **259**, 20 (1976).
219. *Henderson R., Unwin P. N. T.*, Three-dimensional model of purple membrane obtained by electron microscopy. Nature 257, 28 (1975).
220. *Marvin D. A., Wachtel E. J.*, Structure and assembly of filamentous bacterial viruses. Phil. Trans. Roy. Soc. London **B 276**, 81 (1976).
221. *Irwin M. J., Nyborg J., Reid B. R., Blow D. M.*, The crystal structure of tyrosyl-transfer RNA synthetase at 2.7 Å resolution. J. Mol. Biol. **105**, 577 (1976).
222. *Stone D., Sodek J., Johnson P., Smillie L. B.*, Tropomyosin: Correlation of amino acid sequence and structure. Proc. 9th FEBS Meeting (Budapest), **31**, 125 (1975).
223. *Stewart M.*, Tropomyosin: Evidence for no stagger between chains. FEBS Lett. **53**, 5 (1975).
224. *Stewart M., McLachlan A. D.*, Structure of magnesium paracrystals of α-tropomyosin. J. Mol. Biol. 103, 251 (1976).
225. *Chothia C., Levitt M., Richardson D.*, Structure of proteins: Packing of α-helices and pleated sheets. Proc. Nat. Acad. Sci. USA **74**, 4130 (1977).
226. *Richardson J. S.*, Handedness of crosssover connections in β-sheets. Proc. Nat. Acad. Sci. USA 73, 2619 (1976).
227. *Sternberg M. J. E., Thornton J. M.*, On the conformation of proteins: The handedness of the β-strand — α-helix — β-strand unit. J. Mol. Biol. **105**, 367 (1976).
228. *Nagano K.*, Logical analysis of mechanism of protein folding. IV. Supersecondary structures. J. Mol. Biol. **109**, 235 (1977).
229. *Rao S. T., Rossmann M. G.*, Comparison of super-secondary structures in proteins. J. Mol. Biol. **76**, 241 (1973).
230. *Moras D., Olsen K. W., Sabesan M. N., Buehner M., Ford G. C., Rossmann M. G.*, Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase. J. Biol. Chem. **250**, 9137 (1975). -
231. *Biesecker G., Harris J. I., Thierry J. C., Walker J. E., Wonacott A. J.*, Sequence and structure of D-glyceraldehyde 3-phosphate dehydrogenase from *Bacillus stearothermophilus*. Nature **266**, 328 (1977).
232. *Rossmann M. G., Adams M. J., Buehner M., Ford G. C., Hackert M. L., Lentz P. J., Jr., McPherson A., Jr., Schevitz R. W, Smiley I. E.*, Structural constraints of possible mechanisms of lactate dehydrogenase as shown by high resolution studies of the apoenzyme and a variety of enzyme complexes. Cold Spring Harbor Symp. Quant. Biol. **36**, 179 (1971).
233. *Hill E., Tsernoglou D., Webb L., Banaszak L. J.*, Polypeptide conformation of cytoplasmic malate dehydrogenase from an electron density map at 3.0 Å resolution. J. Mol. Biol. **72**, 577 (1972).
234. *Eklund H., Nordström B., Zeppezauer E., Söderlund G., Ohlsson I., Boiwe T., Söderberg B. O., Tapia O., Bränden C. I.*, Three-dimensional structure of horse liver alcohol dehydrogenase at 2,4 Å resolution. J. Mol. Biol. **102**, 27 (1976).
235. *Blake C. C. F.*, X-ray studies of glycolytic enzymes. Essays Biochem. **11**, 37 (1975).
236. *Fletterick R. J., Sygusch J., Semple H., Madsen N. B.*, Structure of glycogen phosphorylase *a* at 3.0 Å resolution and its ligand binding sites at 6 Å. J. Biol. Chem. **251**, 6142 (1976).

237. *Burnett R. M., Darling G. D., Kendall D. S., LeQuesne M. E., Mayhew S. G., Smith W. W., Ludwig M. L.*, The structure of the oxidized form of clostridial flavodoxin at 1.9 Å resolution. J. Mol. Chem. **249**, 4383 (1974).
238. *Watenpaugh K. D., Sieker L. H., Jensen L. H., Legall J., Dubourdieu M.*, Structure of the oxidized form of a flavodoxin at 2.5 Å resolition: Resolution of the phase ambiguity by anomalous scattering. Proc. Nat. Acad Sci. USA **69**, 3185 (1972).
239. *Drehgth J., Hol W. G. J., Jansonius J. N., Koeckoek R.*, A comparison of the three-dimensional structures of subtilisin BPN' and subtilisin novo. Cold Spring Harbor Symp. Quant. Biol. **36**, 107 (1971).
240. *Wright C. S., Alden R. A., Kraut J.*, Structure of subtilisin BPN' at 2,5 Å resolution. Nature **221**, 235 (1969).
241. *Matthews B. W., Remington S. J.*, The three-dimensional structure of the lysozyme from bacteriophage T4. Proc. Nat. Acad. Sci. USA **71**, 4178 (1974).
242. *Arnone A., Bier C. J., Cotton F. A., Day V. W., Hazen E. E., Jr., Richardson D. C., Richardson J. S., Yonath A.*, A high resolution structure of an inhibitor complex of the extracellular nuclease of *Staphylococcus aureus*. J. Biol. Chem. **246**, 2302 (1971).
243. *Shotton D. M., Watson H. C.*, Three-dimensional structure of tosylelastase. Nature **225**, 811 (1970).
244. *Stroud R. M., Kay L. M., Dickerson R. E.*, The structure of bovine trypsin: Electron density maps of the inhibited enzyme at 5 Å and 2.7 Å resolution. J. Mol. Biol. **83**, 185 (1974).
245. *Fehlhammer H., Bode W., Huber R.*, Crystal structure of bovine trypsinogen at 1.8 Å resolution. J. Mol. Biol. **111**, 415 (1977).
246. *Dalbaere L. T. J., Hutcheon W. L. B., James M. N. G., Thiessen W. E.*, Tertiary structural differences between microbial serine proteases and pancreatic serine enzymes. Nature **257**, 758 (1975).
247. *Richardson J. S.*, β-sheet topology and the relatedness of proteins. Nature **268**, 495 (1977).
248. *Wetlaufer D. E.*, Nucleation, rapid folding and globular intrachain regions in proteins. Proc. Nat. Acad. Sci. USA **70**, 697 (1973).
249. *Levitt M., Chothia C.*, Structural patterns in globular proteins. Nature **261**, 552 (1976).
250. *Schulz G. E.*, Structural rules for globular proteins. Angew. Chem. Int. Edit. **16**, 23 (1977).
251. *Hamermesh M.*, Group. Theory, Addison-Wesley, London, 1962.
252. *Weyl H.*, Symmetry, Princeton University Press, Princeton, New Jersy, 1952.
253. *Wright C S.*, Multi-domain structure of the dimeric lectin wheat germ agglutinin. International Symposium of Biomolecular Structure, Conformation, Function and Evolution, Madras, 1978.
254. *Rossmann M. G., Moras D., Olsen K. W.*, Chemical and biological evolution of a nucleotide-binding protein. Nature **250**, 194 (1974).
255. *Schulz G. E., Schirmer R. H.*, Topological comparision of adenylate kinase with other proteins. Nature **250**, 142 (1974).
256. *Zappe H. A., Krohle-Ehrich G., Schulz G. E.*, Low resolution structure of human erythrocyte glutatione reductase. J. Mol. Biol. **113**, 141 (1977).
257. *Bergsma J., Hol W. G. J., Jansonius J. N., Kalk K. H., Ploegman J. H., Smith D. G.*, The double domain structure of rhodanese. J. Mol. Biol. **98**, 637 (1975).
258. *Clegg J. C. S., Kennedy S. I. T.*, Initiation of synthesis of the structural proteins of semliki forest virus. J. Mol. Biol. **97**, 401 (1975).
259. *Blundell T., Dodson G., Hodgkin D., Mercola D.*, Insulin: The structure in

the crystal and its reflection in chemistry and biology. Adv. Prot. Chem. **26**, 279 (1972).
260. *Moult J., Yonath A., Traub W., Smilansky A., Podjarny A., Rabinovich D., Saya A.*, The structure of triclinic lysozyme at 2.5 Å resolution. J. Mol. Biol. **100**, 179 (1976).
261. *Steitz T. A., Fletterick R. J., Anderson W. F., Anderson C. M.*, High resolution X-ray structure of yeast hexokinase, an allosteric protein exhibiting a non-symmetric arrangement of subunits. J. Mol. Biol. **104**, 197 (1976).
262. *Caspar D. L. D., Klug A.*, Physical principles in the construction of regular viruses. Cold Spring Harbor Symp. Quant. Biol. **27**, 1 (1962).
263. *Winkler F. K., Schutt C. E., Harrison S. C., Bricogne G.*, Tomato bushy stunt virus at 5.5 Å resolution. Nature **265**, 509 (1977).
264. *Liljas A., Rossmann M. G.*, X-ray studies of protein interactions. Annu. Rev. Biochem. **43**, 475 (1974).
265. *Rossmann M. G., Adams M. J., Buehner M., Ford G. C., Hackert M. L., Liljas A., Rao S. T., Banaszak L. J., Hill E., Tsernoglou D., Webb L.*, Molecular symmetry axes and subunit interfaces in certain dehydrogenases. J. Mol. Biol. **76**, 533 (1973).
266. *Chothia C., Janin J.*, Principles of protein-protein recognition. Nature **256**, 705 (1975).
267. *Janin J., Chothia C.*, Stability and specifity of protein-protein interactions: The case of the trypsin-trypsin inhibitor complexes. J. Mol. Biol. **100**, 197 (1976).
268. *Chothia C., Wodak S., Janin J.*, Role of subunit interfaces in the allosteric mechanism of hemoglobin. Proc. Nat. Acad. Sci. USA **73**, 3793 (1976).
269. *Huber R., Kukla D., Bode W., Schwager P., Bartels K., Deisenhofer J., Steigeman W.*, Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. J. Mol. Biol. **89**, 73 (1974).
270. *Levinthal C., Wodak S. J., Kahn P., Davivanian A. K.*, Hemoglobin interaction in sickle cell fibers. I: Theoretical approaches to the molecular contacts. Proc Nat. Acad. Sci. USA **72**, 1330 (1975).
271. *Timkovich R., Dickerson R. E.*, Recurrence of the cytochrome fold in a nitrate-respiring bacterium. J. Mol. Biol. **79**, 39 (1973).
272. *Salemme F. R., Freer S. T., Xuong N. G. H., Alden R. A., Kraut J.*, The structure of oxidized cytochrome c_2 of *Rhodospirillum rubrum*. J. Biol. Chem. **248**, 3910 (1973).
273. *Takano T., Kallai O. B., Swanson R., Dickerson R. E.*, The structure of ferrocytochrome *c* at 2.45 Å resolution. J. Biol. Chem. **248**, 5234 (1973).
274. *Frier J. A., Perutz M. F.*, Structure of human foetal deoxyhaemoglobin. J. Mol. Biol. **112**, 97 (1977).
275. *Vainshtein B. K., Harutyunuan E. H., Kuranova I. P., Borisov V. V., Sosfenov N. I., Pavlovsky A. G.. Grebenko A. I., Konareva N. V.*, Structure of leghaemoglobin from lupin root nodules at 5 Å resolution. Nature **254**, 163 (1975).
276. *Love W. E., Klock P. A., Lattmann E. E., Padlan E. A., Ward K. B., Jr., Hendrickson W. A.*, The structures of lamprey and bloodworm hemoglobins in relation to their evolution and function, Cold Spring Harbor Symp. Quant. Biol. **36**, 349 (1971).
277. *Huber R., Epp O., Steigemann W., Formaneck H.*, The atomic structure of erythrocruorin in the light of chemical sequence, and its comparison with myoglobine. Eur. J. Biochem. **19**, 42 (1971).
278. *Sasaki K., Dockerill S., Adamiak D. A., Tickle I. J., Blundell T.*, X-ray analysis of glucagon and its relationship to receptor binding. Nature **257**, 751 (1975).
279. *Drength J., Jansonius J. N., Koekoek R., Sluyterman L. A. A., Wolthers B. G.*,

The structure of the papain molecule. Phil. Trans. Roy. Soc. London **B 257**, 231 (1970).

280. *Colman P. M., Jansonius J. N., Matthews B. W.*, The structure of thermolysin: An electron density map at 2.3 Å resolution. J. Mol. Biol. **70**, 701 (1972).
281. *Hardman K. D., Ainsworth C. F.*, Structure of concanavalin A. at 2.4 Å resolution. Biochemystry **11**, 4910 (1972).
282. *Reeke G. N., Becker J. W., Edelman G. M.*, The covalent and three-dimensional structure of concanavalin A. Biol. Chem. **250**, 1525 (1975).
283. *Blake C. C. F., Geisow M. J., Swan I. D., Rerat C., Rerat B.*, Structure of human plasma prealbumin at 2.5 Å resolution. J. Mol. Biol. **88**, 1 (1974).
284. *Watenpaugh K. D., Sieker L. C., Herriott J. R., Jensen L. H.*, The structure of non-heme iron protein: Rubredoxin at 1,5 Å resolution. Cold Spring Harbor Symp. Quant. Biol. **36**, 359 (1971).
285. *Watenpaugh K. D., Sieker L. C., Herriott J. R., Jensen L. H.*, Refinement of the model of a protein: Rubredoxin at 1.5 Å resolution. Acta Crystallogr. **B 29**, 943 (1973).
286. *Richardson J. S., Thomas K. A., Rubin B. H., Richardson D. C.*, Crystal structure of bovine Cu, Zn superoxide dismutase at 3 Å resolution: Chain tracing and metal ligands. Proc. Nat. Acad. Sci. USA **72**, 1349 (1975).
287. *Low B. W., Preston H. S., Sato A., Rosen L. S., Searl J. S., Rudko A. D., Richardson J. S.*, Three-dimensional structure of erabutoxin *b* neurotoxic protein: Inhibitor of acetylcholine receptor. Proc. Nat. Acad. Sci. USA **73**, 2991 (1976).
288. *Subramanian E., Swan I. D. A., Liu M., Davies D. R., Jenkins J. A., Tickle I. J., Blundell T. L.*, Homology among acid proteases: Comparison of crystal structures at 3 Å resolution of acid proteases from *Rhizopus chinensis* and *Endothia parasitica*. Proc. Nat. Acad. Sci. USA **74**, 556 (1977).
289. *Hsu I. N., Delbaere L. T. J., James M. N. G., Hofmann T.*, Penicillopepsin from *Penicillum janthinellum* crystal structure at 2.8 Å and sequence homology with porcine pepcin. Nature **266**, 140 (1977).
290. *Tang J., James M. N. G., Hsu I. N., Jenkins J. A., Blundell T. L.*, Structural evidence for gene duplication in the evolution of acid proteases. Nature **271**, 618 (1978).
291. *Poljak R. J., Amzel L. M., Chen B. L., Phizackerley, R. P., Saul F.*, The three-dimensional structure of the Fab'fragment of a human myeloma immunoglobulin at 2.0 Å resolution. Proc. Nat. Acad. Sci. USA **71**, 3440 (1974).
292. *Schiffer M., Girling R. L., Ely K. R., Edmundson A. B.*, Structure of a λ-type Bence-Jones protein at 3.5 Å resoluton. Biochemistry **12**, 4620 (1973).
293. *Segal D. M., Padlan E. A., Cohen G. H., Rudikoff S., Potter M., Davies D. R.*, The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobuline Fab and the nature of the antigen binding site. Proc. Nat. Acad. Sci. USA **71**, 4298 (1974).
294. *Huber R., Deisenhofer J., Colman P. M., Matsuhima M., Palm W.*, Crystallographic structure studies of an IgG molecule and an F_c fragment. Nature **264**, 415 (1976).
295. *Liljas A., Kannan K. K., Bergstén P. C., Waara I., Fridborg K., Strandberg B., Carlbom U., Järup L., Lövgren S., Petef M.*, Crystal structure of human carbonic anhydrase C. Nature New Biol. **235**, 131 (1972).
296. *Kannan K. K., Notstran B., Fridborg K., Lövgren S., Ohlsson A., Petef M.*, Crystal structure of human erythrocyte carbonic anhydrase B. Three dimen-

sional structure at a nominal 2.2 Å resoulution. Proc. Nat. Acad. Sci. USA 72, 51 (1975).

297. *Mathews F. S., Argos P., Levine M.*, The structure of cytochrome b_5 at 2.0 Å resolution. Cold Spring Harbor Symp. **36**, 387 (1971).
298. *Mathews F. S., Levine M., Argos P.*, Three-dimensional Fourier synthesis of calf liver cytochrome b_5 at 2.8 Å resolution. J. Mol. Biol. **64**, 449 (1972).
299. *Carter C. W., Jr. Kraut J., Freer S. T., Xuong N. H., Alden R. A., Bartsch R. G.*, Two-Ångstrom crystal structure of oxidized *Chromatium* high potential iron protein. J. Biol. Chem. **249**, 4212 (1974).
300. *Blake C. C. F., Koenig D. F., Mair G. A., North A. C. T., Philips D. C., Sarma V. R.*, Structure of hen egg-white lysozyme. Nature **206**, 757 (1965).
301. *Kartha G., Bello J., Harker D.*, Tertiary structure of ribonuclease. Nature **213**, 862 (1967).
302. *Carlisle C. H., Palmer R. A., Mazumdar S. K., Gorinsky B. A., Yeates D. G. R.*, The structure of ribonuclease at 2.5 Å resolution. J. Mol. Biol. **85**, 1 (1974).
303. *Fenna R. E., Natthews B. W.*, Chlorophill arrangement in a bacteriochlorophill protein from *Chlorobium limicola*. Nature **258**, 573 (1975).
304. *Holmgren A., Söderberg B. O., Eklund H., Bränden C. I.*, Three-dimensional structure of *Escherichia coli* thioredoxin-S_2 to 2.8 Å resolution. Proc. Nat. Acad. Sci. USA 72, 2305 (1975).
305. *Banner D. W., Bloomer A. C., Petsko G. A., Phillips D. C., Pogson C. I., Wilson I. A., Corran P. H., Furth A. J., Milman J. D , Offord R. E., Priddle J. D., Waley S. G.*, Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. Nature, **255**, 609 (1975).
306. *Mavridis I. M., Tulinsky A.*, The folding and quaternary structure of trimeric 2-keto-3-deoxy-6-phosphogluconic aldolase at 3.5 Å resolution. Biochemistry **15**, 4410 (1976).
307. *Campbell J. W., Watson H. C., Hodgson G. L.*, Structure of yeast phosphoglycerate mutase. Nature **250**, 301 (1974).
308. *Matthews D. A., Alden R. A., Bolin J. T., Freer S. T., Hamlin R., Xuong M., Kraut J., Poe M., Williams M., Hoogsteen K.*, Dihydrofolate reductase: X-ray structure of the binary complex with methotrexate. Science **197**, 452 (1977).
309. *Quiocho F. A., Gilliland G. L., Phillips G. N.*, The 2.8 Å resolution structure of the L-arebinose binding protein from *E. coli*. J. Biol. Chem. **252**, 5142 (1977).
310. *Blake C. C. F., Evans P. R.*, Structure of horse muscle phosphoglycerate kinase. J. Mol. Biol. **84**, 585 (1974).
311. *Bryant T. N., Watson H. C., Wendell P. L.*, Structure of yeast phosphoglycerate kinase. Nature. 247, 14 (1974).
312. *Levine M., Muirhead H., Stammers D. K., Stuart D. I.*, Structure of pyruvate kinase and similarities with other enzymes: Possible implications for protein taxonomy and evolution. Nature **271**, 626 (1978).
313. *Shaw P. J., Muirhead H.*, Crystallographic structure analysis of glucose 6-phosphate isomerase at 3.5 Å resolution. J. Mol. Biol. **109**, 475 (1977).
314. *Adman E. T., Sieker L. C., Jensen L. H.*, The structure of a bacterial ferredoxin. J. Biol. Chem. 248, 3987 (1973).
315. *Drenth J., Enzing C. M., Kalk K. H., Vessies C. A.*, Structure of porcine pancreatic prephospholipase A_2. Nature **264**, 373 (1976).

316. *Wright C. S.*, The crystal structure of wheat germ agglutinine at 2.2 Å resolution. J. Mol. Biol. **111**, 439 (1977).
317. *Knox J. R., Wyckoff H. W.*, A crystallographic study of alkaline phosphatase at 7.7 Å resolution. J. Mol. Biol. **74**, 533 (1973).
318. *Adams M. J., Helliwell J. R., Bugg C. H.*, Structure of 6-phosphogluconate dehydrogenase from sheep liver at 6 Å resolution. J. Mol. Biol. **112**, 183 (1977).
319. *Warren S. G., Edwards B. F. P., Evans D. R., Wiley D. C., Lipscomb W. N.*, Aspartate transcarbamoylase from *Escherichia coli*: Electron density at 5.5 Å resolution. Proc. Nat. Acad. Sci. USA **70**, 1117 (1973).
320. *Hoare R. J., Harrison P. M., Hoy T. G.*, Structure of horse-spleen apoferritin at 6 Å resolution. Nature **255**, 653 (1975).
321. *Perutz M. F., Muirhead H., Cox J. M., Goamon L. C. G.*, Three-dimensional Fourier synthesis of horse oxyhaemoglobin at 2.8 Å resolution: The atomic model. Nature **219**, 131 (1968).
322. *Bolton W., Perutz M. F.*, Three dimensional Fourier synthesis of horse deoxyhaemoglobin at 2.8 Å resolution. Nature **288**, 551 (1970).
323. *Srinivasan R.*, Helical length distribution from protein crystallographic data. Ind. J. Biochem. Biophys. **13**, 192 (1976).
324. *Johnson C. N.*, Oak Ridge Nat. Lab Report-5138 ORTEP-II: A Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations, 1976.
325. *Trus B. L., Piez K. A.*, Molecular packing of collagen: Three-dimensional analysis of electrostatic interactions. J. Mol. Biol. **108**, 705 (1976).
326. *Lewis P. N., Momany F. A., Scheraga H. A.*, Folding of polypeptide chains in proteins: A proposed mechanism for folding. Proc. Nat. Acad. Sci. USA **68**, 2293 (1971).
327. *Sternberg M. J. E., Thornton J. M.*, On the conformation of proteins: An analysis of β-pleated sheets. J. Mol. Biol. **110**, 285 (1977).
328. *Blout E. R., de Lozé C., Bloom S. M., Fasman G. D.*, The dependence of the conformation of the synthetic polypeptides on amino acid composition. J. Amer. Chem. Soc. **82**, 3787 (1960).
329. *Davies D. R.*, A correlation between amino acid composition and protein structure. J. Mol. Biol. **9**, 605 (1964).
330. *Jirgensons B.*, Optical Rotary Dispersion of Proteins and Other Macromolecules. Springer-Verlag, Heidelberg, 1969.
331. *Guzzo A. V.*, The influence of amino acid sequence on protein structure. Biophys. J. **5**, 809 (1965).
332. *Prothero J. W.*, Correlation between the distribution of amino acids and alpha helices. Biophys. J. **6**, 367 (1966).
333. Havensteen, A study of the correlation between the amino acid composition and the helical content of proteins. J. Theor. Biol. **10**, 1 (1966).
334. *Cook D. A.*, The relation between amino acid sequence and protein conformation. J. Mol. Biol. **29**, 167 (1967).
335. *Periti P. R., Guagliarotti G., Liquori A. M.*, Recognition of α-helical segments in protein of known primary structure. J. Mol. Biol. **24**, 313 (1967).
336. *Dunnill P.*, The use of helical net-diagrams to represent protein structures. Biophys. J. **8**, 865 (1968).
337. *Low B. W., Lowell F. M., Rudko A. D.*, Prediction of α-helical regions in proteins of known sequence. Proc. Nat. Acad. Sci. USA **60**, 1519 (1968).
338. *Dirkx J.* Une méthode semi-empirique de prédiction des régions L-hélicoidales des chaînes polypeptidiques d'aprés leur structure primaire. Arch. Int. Physiol. Biochim. **80**, 185 (1972).
339. *Beghin F., Dirkx J.*, Une méthode statistique simple de prédiction des conformations proteiques. Arch. Int. Physiol. Biochim. **83**, 167 (1975).

340. *Chou P. Y., Fasman G. D.*, Prediction of protein conformation. Biochemistry **13**, 222 (1974).
341. *Птицын О. Б., Финкельштейн А. В.* Предсказание спиральных участков глобулярных белков по их первичной структуре. Докл. АН СССР **195**, 221 (1970).
342. *Птицын О. Б., Финкельштейн А. В.* Связь между вторичной и первичной структурами глобулярных белков. Биофизика **15**, 757 (1970).
343. *Finkelstein A. V., Ptitsyn O. B.*, Statistical analysis of the correlation among amino acid residues in helical, β-structural and non-regular regions of globular proteins. J. Mol. Biol. **62**, 613 (1971).
344. *Periti P. F.*, A Bayesian approach to the recognition of discrete patterns with an applications to a problem of protein molecular structure. Bull. Chim. Farm. **113**, 187 (1974).
345. *Pain R. H., Robson B.*, Analysis of the code relating sequence to secondary structure in proteins. Nature **227**, 62 (1970).
346. *Robson B., Pain R. H.*, Analysis of the code relating sequence to conformation in proteins: Possible implications for the mechanism of formation of helical regions. J. Mol. Biol. **58**, 237 (1971).
347. *Robson B.*, Analysis of the code relating sequence to conformation in globular proteins. Theory and application of expected information. Biochem. J. **141**, 853 (1974).
348. *Robson B., Pain R. H.* Analysis of the code relating sequence to conformation in globular proteins. Development of a stereochemical alphabet on the basis of intraresidue information. Biochem. J. **141**, 869 (1974).
349. *Robson B., Pain R. H.*, Analysis of the code relating sequence to conformation in globular proteins. An informational analysis of the role of the residue in determining the conformation of its neighbours in the primary sequence. Biochem. L. **141**, 883 (1974).
350. *Robson B., Pain R. H.*, Analysis of the code relating sequence to conformation in globular proteins. The distribution of residue pairs in turns and kinks in the backbone chain. Biochem. J. **141**, 899 (1974).
351. *Robson B., Suzuki E.*, Conformational properties of amino acid residues in globular proteins. J. Mol. Biol. **107**, 327 (1977).
352. *Suzuki E., Robson B.*, Relationship between helix-coil transition parameters for synthetic polypeptides and helix conformation parameters for globular proteins. A simple model. J. Molec. Biol. **107**, 357 (1977).
353. *Nagano K.*, Logical analysis of mechanism of protein folding. I. Predictions of helices, loops and β-structures from primary structure. J. Mol. Biol. **75** 401 (1973).
354. *Nagano K.*, Logical analysis of mechanism of protein folding. II. The nucleation process. J. Mol. Biol. **84**, 337 (1974).
355. *Nagano K., Hasegawa K.*, Logical analysis of mechanism of protein folding. III. Prediction of the strong long-range interactions. J. Mol. Biol. **94**, 257 (1975).
356. *Nagano K.*, Triplet information in helix prediction applied to the analysis of supersecondary structures. J. Mol. Biol. **109**, 251 (1977).
357. *Kabat E. A., Wu T. T.*, The influence of neighbor amino acids on the conformation of the middle amino acid in proteins. Comparision of predicted and experimental determination of β-sheets in concanavalin A. Proc. Nat. Acad. Sci. USA **70**, 1473 (1973).
358. *Kabat E. A., Wu T. T.*, The influence of nearest-neighboring amino acid residues on aspects of secondary structure of proteins. Attempts to locate α-helices and β-sheets. Biopolymers **12**, 751 (1973).
359. *Kabat E. A., Wu T. T.*, Further comparison of predicted and experimentally determined structure of adenylate kinase. Proc. Nat. Acad. Sci. USA **71**, 4217 (1974).
360. *Wu T. T., Kabat E. A.*, An attepmt to locate the non-helical and permissi-

vely helical sequences of proteins: Application to the variable regions of immunoglobulin light and heavy chains. Proc. Nat. Acad. Sci. USA **68**, 1501 (1971).

361. *Kabat E. A., Wu T. T.*, Constraction of a three-dimensional model of the polypeptide backbone of the variable region of kappa immunoglobulin light chains. Proc. Nat. Acad. Sci. USA **69**, 960 (1972).
362. *Wu T. T., Kabat E. A.*, An attempt to evaluate the influence of neighboring amino acids (n — 1) and (n + 1) on the backbone conformation of amino acid (n) in proteins. Use in predicting the threedimensional structure of the polypeptide backbone of other proteins. J. Mol. Biol. **75**, 13 (1973).
363. *Kotelchuck D., Scheraga H. A.*, The influence of shortrange interactions on protein conformation. I. Side chainbackbone interactions within a single peptide unit. Proc. Nat. Acad. Sci. USA **61**, 1163 (1968).
364. *Kotelchuck D., Scheraga H. A.*, The influence of shortrange interactions on protein conformation. II. A model for predicting the α-helical regions of proteins. Proc. Nat. Acad. Sci. USA **62**, 14 (1969).
365. *Leberman R.*, Secondary structure of tobacco mosaic virus protein. J. Mol. Biol. **55**, 23 (1971).
366. *Burgess A. W., Schegara H. A.*, Assessment of some problems associated with prediction of the three-dimentional structure of a protein from its amino-acid sequence. Proc. Nat. Acad. Sci. USA **72**, 1221 (1975).
367. *Ponnuswamy P. K., Warme P. K., Scheraga H. A.*, Role of medium-range interactions in proteins. Proc. Nat. Acad. Sci. USA, **70**, 830 (1973).
368. *Lewis P. N., Gō N., Gō M., Kotelchuck D., Scheraga H. A.*, Helix probability profiles of denatured proteins and their correlation with native structures. Proc. Nat. A cad. Sci. USA **65**, 810 (1970).
369. *Lewis P. N., Scheraga H. A.*, Predictions of structural homologies in cytochrome *c* proteins. Arch. Biophys. Biochem. **144**, 576 (1971).
370. *Ptitsyn O. B., Denesyuk A. I., Finkelstein A. V., Lim V. I.*, Prediction of the secondary structure of the L7, L12 proteins of the *E. coli* ribosome. FEBS Lett. **34**, 55 (1973).
371. *Finkelstein A. V., Ptitsyn O. B.*, A theory of protein molecule self-organization. IV. Helical and irregular local structures of unfolded protein chains. J. Mol. Biol. **103**, 15 (1976).
372. *Finkelstein A. V., Ptitsyn, O. B.*, Theory of protein molecule self-organisation. I. Thermodynamic parameters of local secondary structures it the infolded protein chain. Biopolymers **16**, 469 (1977).
373. *Finkelstein A. V., Ptitsyn O. B., Kozitsyn S. A.*, Theory of protein molecule self-organisation. II. A comparison of calculated thermodynamic parameters of local secondary structures with experiments. Biopolymers **16**, 497 (1977).
374. *Finkelstein A. V.*, Theory of protein molecule self-organisation. III. A calculating method for the probabilities of the secondary structure formation in an unfolded polypeptide chain. Biopolymers **16**, 525 (1977).
375. *Lifson S., Roig A.*, On the theory of helix-coil transitions in polypeptides. J. Chem. Phys. **34**, 1963 (1961).
376. *Schiffer M., Edmundson A. B.*, Use of helical wheels to represent the structures of protein and to identify segments with helical potential. Biophys. J. **7**, 121 (1967).
377. *Schiffer M., Edmundson A. B.*, Correlation of amino acid sequence and conformation in tobacco mosaic virus. Biophys. J. **8**, 29 (1968).
378. *Palau J., Puigdoménech P.*, The structural code for proteins: Zonal distribution of amino acid residues and stabilisation of helices by hydrophobic triplets. J. Mol. Biol. **88**, 457 (1974).
379. *Lim V. L.*, Structural principles of the globular organisation of protein chains. A stereochemical theory of globular protein secondary structure. J. Mol. Biol. **88**, 857 (1974).

380. *Lim V. I.*, Algorithms for prediction of α-helical and β-structural regions in globular proteins. J. Mol. Biol. **88**, 873 (1974).
381. *Argos P., Schwarz J., Schwarz J.*, An assessment of protein secondary structure prediction methods based on amino acid sequence. Biochem. Biophys. Acta. **439**, 261 (1976).
382. *Lenstra J. A.*, Evaluation of secondary structure prediction in proteins. Biochim. Biophys. Acta **491**, 333 (1977).
383. *Schulz G. E., Barry C. D., Friedman J., Chou P. Y., Fasman G. D., Finkelstein A. V., Lim V. I., Ptitsyn O. A., Kabat E. A., Wu T. T., Levitt M., Robson B., Nagano K.*, Comparison of predicted and experimentally determined secondary structure of adenylate kinase. Nuture **250**, 140 (1974).
384. *Matthews B. W.*, Comparison of predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta **405**, 442 (1975).
385. *Lenstra J. A., Hofsteenge J., Beintema J. J.*, Invariant features of the structure of pancreatic ribonuclease. J. Mol. Biol. **109**, 185 (1977).
386. *Wooton J. C.*, The coenzyme binding domains of glutamate dehydrogenases. Nature **252**, 542 (1974).
387. *Schulz G. E.*, Recognition of phylogenetic relationship from polypeptide chain fold similarities. J. Mol. Evol. **9**, 339 (1977).
388. *Argos P.*, Secondary structure predictions of calcium-binding proteins. Biochemistry **16**, 665 (1977).
389. *Heil A. F., Müller G., Noda L., Pinder T., Schirmer H., Schirmer I., von Zabern I.*, The amino-acid sequence of porcine adenilate kinase from skeletal muscle. Eur. J. Biochem. **43**, 131 (1974).
390. Protein Data Bank, Department of Chemistry, Brookhaven National Laboratory, Associated Universities, Inc., Upton, L. I., New York 1973.
391. Kendrew — Watson Models, Cambridge Repetition Engineers, Ltd., Green's Road, Cambridge, England.
392. *Richards F. M.*, The matching of physical models to three-dimensional electron-density maps: A simple optical device. J. Mol. Biol. **37**, 225 (1968).
393. Byron's Bender, Charles Supper Co., 15 Tech Circle, Natick, Massachusetts 01760.
394. Nicholson Models, Labquip Ltd. 18 Rosehill Park Estate, Caversham, Reading RG4 8XE, England.
395. CPK-models, The Ealing Corporation, South Natick, Massashusetts 01760.
396. *Lee B., Richards F. M.*, The interpretation of protein structures: Estimation of static accessibility. J. Mol. Biol. **55**, 379 (1971).
397. Evans and Sutherland, 580 Arpeen Drive, Salt Lake City, Utah 84108.
398. *Diamond R.*, Bilder: A computer graphics program for biopolymers. International Symposium of Biomolecular Structure, Conformation, Function and Evolution, Madras, 1978.
399. *Feldmann R. J., Porter T. K.*, Div. Computer Research and Technology National Institutes of Health, Bethesda, Maryland 20014 (1978).
400. Stereo viewer, Abrams Instr. Corp. 606 East Shiawassee Street, Lansing, Michigan 48901.
401. AMSOM, Tracor Jitco Inc., 1776 E. Jefferson St., Rockville, Maryland 20852.
402. *Balasubramanian R.*, New type of representation for mapping chain-folding in protein molecules. Nature **266**, 856 (1977).
403 *Srinivasan R., Balasubramanian R., Rajan S. S.*, Some new methods and general results of analysis of protein crystallographic structural data. J. Mol. Biol. **98**, 739 (1975).
404. *Levitt W., Warshell A.*, Computer simulation of protein folding. Nature **253**, 694 (1975).
405. *Balasubramanian R.*, A new type of representation of dipeptide conformation. Biochem. J. **157**, 769 (1976).

406. *Kuntz I. D.*, An approach to the tertiary structure of globular proteins. J. Amer. Chem. Soc. **97**, 4362 (1975).
407. *Rossmann M. G., Liljas A.*, Recognition of structural domains in globular proteins. J. Mol. Biol. **85**, 177 (1974).
408. *Kabat E. A., Wu T. T., Bilofsky H.* Variable Regions of Immunoglobulin Chains. Bolt Beranck and Newman, Cambridge, Massachusetts.
409. *Behrens P. G., Spiekerman A. M., Brown J. B.*, Structure of human serum albumin. Fed. Proc. **34**, 2106 (1974).
410. *Gutte B., Merrifield R. B.*, The total synthesis of an enzyme with-ribonuclease A activity. J. Amer. Chem. Soc. **91**, 501 (1969).
411. *Hirschmann R., Nutt R. F., Veber D. F., Vitali R. A., Varga S. L., Jacob T. A., Holly F. W., Denkewalter R. G.*, Studies on the total synthesis of an enzyme. V. The preparation of enzymically active material. J. Amer. Chem. Soc. **91**, 507 (1969).
412. *Anfinsen C. B., Haber E., Sela M., White F. H., Jr.*, The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Nat. Acad. Sci. USA **47**, 1309 (1961).
413. *Pase C. N.*, The stability of globular proteins. CRC Crit. Rev. Biochem. **3**, 1 (1975).
414. *Hvidt A., Nielsen S. O.*, Hydrogen exchange in proteins. Adv. Prot. Chem. **21**, 287 (1966).
415. *Ottesen M.*, Methods for measurement of hydrogen isotope exchange in globular proteins. Meth. Biochem. Anal. **20**, 135 (1971).
416. *Nakanishi M., Tsuboi M., Ikegami A.*, Fluctuation of the lysozyme structure. J. Mol. Biol. **70**, 351 (1972).
417. *Privalov P. L. Khechinashvili N. N.*, A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. J. Mol. Biol. **86**, 665 (1974).
418. *Sachs D. H., Schechter A. N., Eastlake A., Anfinsen C. B.*, An immunologic approach to conformational equilibria of polypeptides. Proc. Nat. Acad. Sci. USA **69**, 3790 (1972).
419. *Cooper A.*, Thermodynamic fluctuations in protein molecules. Proc. Nat. Acad. Sci. USA **73**, 2740 (1976).
420. *Callen H. B.*, Thermodynamics. John Wiley, New York, 1960.
421. *Privalov P. L.*, Thermal investigations of biopolymer solutions and scanning microcalorimetry. FEBS Lett. Suppl. **40**, 140 (1974).
422. *Suurkuusk J.*, Specific heat measurements on lysozyme, chymotrypsinogen, and ovalbumin in aqueous solution and in solid state. Acta Chem. Scand. **B28**, 409 (1974).
423. *Hetzel R., Wüthrich K., Deisenhofer J., Huber R.*, Dynamics of the aromatic amino acid residues in the globular conformation of the basic pancreatic trypsin inhibitor (BPTI). II. Semi-empirical energy calculations. Biophys. Struct. Mech. **2**, 159 (1976).
424. *Wagner G., DeMarco A., Wüthrich K*, Dynamics of the aromatic amino acid residues in the globular conformation of the basic pancreatic trypsin inhibitor (BPTI). I. ^{1}H-NMP studies. Biophys. Struct. Mech. **2**, 139 (1976).
425. *Campbell I. D., Dobson C. M., Williams R. J. P.*, Proton magnetic resonance studies of the tyrosine residues of hen lysozyme-assignment and detection of conformational mobility. Proc. R. Soc. London Ser. **B. 189**. 503 (1975).
426. *Allerhand A., Doddrell D., Glushko V., Cochran D., Wenkert E., Lawson P., Gurd F.*, Conformation and segmental motion of native and denatured ribonuclease A in solution. Application of natural-abundance carbon-13 partially relaxed Fourier transform nuclear magnetic resonance. J. Amer. Chem. Soc. **93**, 544 (1971).
427. *Oster O., Niereiter G. W., Clouse A. O., Gurd F. R. N.*, Deuterium nuclear magnetic resonance of deuterium-labeled diacetyldeuterohemin incorporated into sperm whale myoglobin. J. Biol. Chem. **250**, 7990 (1975).

428. *Lakowicz J. R., Weber G.*, Quenching of protein fluorescence by oxygen. Detection of structural fluctuations in proteins on the nanosecond time scale. Biochemistry **12**, 4171 (1973).
429. *Eftink M. R., Ghiron C. A.*, Dynamics of a protein matrix revealed by fluorescence quenching. Proc. Nat. Acad. Sci. USA **72**, 3290 (1975).
430. *Grinvald A., Steinberg I. Z.*, Fast relaxation processes in a protein revealed by the decay kinetics of tryptophan fluorescence. Biochemistry **13**, 5170 (1974).
431. *Saviotti M. L., Galley W. C.*, Room temperature phosphorescence and the dynamic aspects of protein structure. Proc. Nat. Sci. USA **71**, 4154 (1974).
432. *Careri G., Faselle P., Gratton P., Gratton E.*, Statistical time events in enzymes: A physical assessment. CRC Crit. Rev. Biochem. **3**, 141 (1975).
433. *Epstein C. J., Goldberger R. F., Anfinsen C. B.*, The genetic control of tertiary protein structure: Studies with model systems. Cold Spring Harbor Symp. Quant. Biol. **27**, 349 (1963).
434. *Wetlaufer D. B., Ristow S.*, Aquisition of three-dimensional structure of proteins. Annu. Rev. Biochem. **42**, 135 (1973).
435. *Haber E., Anfinsen C. B.*, Side-chain interactions governing the pairing of half-cystine residues in ribonuclease. J. Biol. Chem. **237**, 1839 (1962).
436. *Sela M., Lifson S.*, On the reformation of disulphide bridges in proteins. Biochim. Biophys. Acta **36**, 471 (1959).
437. *Kauzmann W.*, In: Sulfur in Proteins (R. Benesch, P. D. Boyer, I. M. Klotz, W. R. Middlebrook, A. Szent-Györgyi and Schwartz, eds.), p. 93. Academic press, New York, 1959.
438. *Epstein H. F., Schechter A. N., Chen R. F., Anfinsen C. B.*, Folding of staphylococcal nuclease: Kinetic studies of two processes in acid renaturation. J. Mol. Biol. **60**, 499 (1971).
439. *Shen L. L., Hermans J., Jr.*, Kinetics of conformation change of sperm-whale myoglobin. I. Folding and unfolding of metmyoglobin following pH jump. Biochemistry **11**, 1836 (1972).
440. *Hagerman P. J., Baldwin R. L.*, A quantitative treatment of the kinetics of the folding transition of ribonuclease A. Biochemistry **15**, 1462 (1976).
441. *Anson M. L.*, Protein denaturation and the properties of protein groups. Adv. Prot. Chem. **2**, 361 (1945).
442. *Steiner D. F., Oyer P.*, The biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma. Proc. Nat. Acad. Sci. USA **57**, 473 (1967).
443. *Steiner D. F., Clark J. L.* The spontaneous reoxidation of reduced beef and rat proinsulins. Proc. Nat. Acad. Sci. USA **60**, 622 (1968).
444. *Varandani P., Shroyer L. A.*, Insulin degradation. Biochim. Biophys. Acta **320**, 249 (1973).
445. *Friedmann M. E., Scheraga H. A., Goldberg R. F.*, Structural studies of ribonuclease. XXVI. The role of tyrosine 115 in the refolding of ribonuclease. Biochemistry **5**, 3770 (1966).
446. *Zabin I., Villarejo M. R.*, Protein complementation. Annu. Rev. Biochem. **44**, 295 (1975).
447. *Baldwin R. L.*, Intermediates in protein folding reactions and the mechanism of protein folding. Annu. Rev. Biochem. **44**, 453 (1975).
448. *Garel J. R., Baldwin R. L.*, Both the fast and slow refolding reactions of ribonuclease A yield native enzyme. Proc. Nat. Acad. Sci. USA **70**, 3347 (1973).
449. *Garel J. R., Nall B. T., Baldwin R. L.*, Guanidine-unfolded state of ribonuclease A contains both fast- and slaw-refolding species. Proc. Nat. Acad. USA **73**, 1853 (1976).
450. *Garel J. R., Baldwin R. L.*, The heat-unfolded state of ribonuclease A is an equilibrium mixture of fast and slow refolding species. J. Mol. Biol. **94**, 611 (1975).

451. *Creighton T. E.*, Energetics of folding and unfolding of pancreatic trypsin inhibitor. J. Mol. Biol. **113**, 295 (1977). (См. также более ранние работы в этом журнале.)
452. *Brandts J. F., Helvorson H. R., Brennan M.*, Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of protine residues. Biochemistry **14**, 4953 (1975).
453. *Furie B., Schechter A. N., Sachs D. H., Anfinsen C. B.*, An immunological approach to the conformational equilibrium of staphylococcal nuclease. J. Mol. Biol. **92**, 497 (1975).
454. *Curd J. G., Schechter A. N., Anfinsen C. B.*, Isolation of antibodies specific for a helical region of staphylococcal nuclease by affinity chromatography with synthetic polypeptides. Fed. Proc. **34**, 550 (1975).
455. *Flory P. J.*, Theory of elastic mechanisms in fibrous proteins. J. Amer. Chem. Soc., **78**, 5222 (1956).
456. *Nakajiama A., Scheraga H. A.*, Thermodinamic study of shrinkage and of phase equilibrium under stress in films made from ribonuclease. J. Amer. Chem. Soc. **83**, 1575 (1961).
457. *Creighton T. E.*, The single-disulphide intermediates in the refolding of reduced pancreatic trypsin inhibitor. J. Mol. Biol. **87**, 603 (1974).
458. *Creighton T. E.*, Kinetics of refolding of reduced ribonuclease. J. Mol. Biol. **113**, 329 (1977).
459. *Hantgan R. R., Hammes G. G., Scheraga H A.*, Pathways of folding of reduced bovine pancreatic ribonuclease. Biochemistry **13**, 3421 (1974).
460. *Anderson W. L., Wetlaufer D. B.*, The folding pathway of reduced lysozyme. J. Biol. Chem. **251**, 3147 (1976).
461. *Schechter A. N., Epstein C. J.*, Spectral studies on the denaturation of myoglobin. J. Mol. Biol. **35**, 567 (1968).
462. *Stellwagen E., Rysavy R., Babul G.*, The conformation of horse heart apocytochrome *c*. J. Biol. Chem. **247**, 8074 (1972).
463. *Fischer W., Taniuchi H., Anfinsen C. B.*, On the role of heme in the formation of the structure of cytochrome *c*. J. Biol. Chem. **248**, 3188 (1973).
464. *Light A., Taniuchi H., Chen R. F.*, A kinetic study of the complementation of fragments of staphylococcal nuclease. J. Biol. Chem. **249**, 2285 (1974).
465. *Taniuchi H., Bohnert J. L., Parker D. S.*, The strengthening of cooperative interactions in the structure of nuclease T′ by binding of ligands. Fed. Proc. **33**, 1589A (1974).
466. *Bornmann L., Hess B., Zimmermann-Telschow H.*, Mechanism of renaturation of pyruvate kinase of *S. carlsbergensis*: Activation by L-valine and magnesium and mangenese ions. Proc. Nat. Acad. Sci. USA **71**, 1525 (1974).
467. *Hess B., Bornmann L.*, Control of enzyme activity through gerulation of three-dimensional structure of proteins: Studies on pyruvate kinase. Adv. Enz. Reg. **13**, 235 (1975).
468. *Hatfield G. W., Burns S. O.*, Ligand-induced maturation of threonine deaminase. Science **167**, 75 (1970).
469. *Alpers J. B., Paulus H., Bazylawicz G. A.*, ATP-catalyzed preconditioning of phosphofructokinase. Proc. Nat. Acad. Sci. USA **68**, 2937 (1971).
470. *Ptitsyn O. B., Rashin A. A.*, A model of myoglobin self-organisation. Biophys. Chem. **3**, 1 (1975).
471. *Warshel A., Levitt M.*, Folding and stability of helical proteins: Carp myogen. J. Mol. Biol. **106**, 421 (1976).
472. *Paulus H., Alpers J. B.*, Preconditioning: An obligatory step in the biosynthesis of oligomeric enzymes and its promotion by allosteric ligands. Enzyme **12**, 385 (1971).
473. *Dickerson R. E.*, Energy and evolution in the folding of proteins. In: Molecular evolution and Polymorphism (M. Kimura, ed.), National Institute of Genetics, Mishima, Japan, 1977.

474. *Dayhoff M. O.*, The origin and evolution of protein superfamilies. Fed. Proc. **35**, 2132 (1976).
475. *Britten R. J., Davidson E. H.*, DNA sequence arrangement and preliminary evidence on its evolution. Fed. Proc. **35**, 2151 (1976).
476. *Kohne D. E.*, Evolution of higher-organism DNA. Quart. Rev. Biophys. **3**, 327 (1970).
477. *Cowan N. J., Secher D. S., Milstein C.*, Purification and sequence analysis of the mRNA coding for an Ig heavy chains. Eur. J. Biochem. **61**, 355 (1976).
478. *Kimura M.*, Evolutionary rate at the molecular level. Nature **217**, 624 (1968).
479. *Kimura M.*, The rate of molecular evolution considered from the standpoint of population genetics. Proc. Nat. Acad. Sci. USA **63**, 1181 (1969).
480. *Kimura M., Ohta T.*, On some principles governing molecular evolution. Proc. Nat. Acad. Sci. USA **71**, 2848 (1974).
481. *Harris H.*, Molecular evolution: The neutralist-selectionist controversy. Fed. Proc. **35**, 2079 (1976).
482. *Thoday J. M.*, Non-Darwinian evolution and biological progress, Nature **255**, 675 (1975).
483. *Fitch W. M., Langley C. H.*, Protein evolution and the molecular clock. Fed. Proc. **35**, 2092 (1976).
484. *Salser W., Bowen S., Browne D., F. El. Adli., Fedoroff N., Fry K., Heindell H., Paddock G., Poon R., Wallace B., Whitcome F.*, Investigation of the organisation of mammalian chromosomes at the DNA sequence level. Fed. Proc. **35**, 23—35 (1976).
485. *Marrotta C. A., Forget B. G., Weissmann S. M., Verma I. M., McCaffrey R. P., Baltimore D.*, Nucleotide sequences of human globin messenger RNA. Proc. Nat. Acad. Sci. USA **71**, 2300 (1974).
486. *Grunstein M., Schedl P., Kedes L.*, Sequence analysis and evolution of sea urchin (*Lytechinus pictus* and *Strongylocentrotus purpurratus*) histone H4 messenger RNA s. J. Mol. Biol. **104**, 351 (1976).
487. *Weinberg E. S., Birnstiel M. L., Purdom I. F., Williamson R.* Genes coding for polysomal 9S RNA of sea urchins: Conservation and divergence. Nature **240**, 225 (1972).
488. *Barnard E. A., Cohen M. S., Gold M. H., Jae-Kyoung K.*, Evolution of ribonuclease in relation to polypeptide folding mechanism. Nature **240**, 395 (1972).
489. *King J. L.*, Progress in the neutral mutation-random drift controversy. Fed. Proc. **35**, 2087 (1976).
490. *Hartley B. S.*, Homologies in serine proteases. Phil. Trans. Roy. Soc. London, Ser. **B 257**, 77 (1970).
491. *Levy P. L., Pangburn M. K., Burstein Y., Ericsson L. H., Neurath H., Walsh K. J.*, Evidence of homologous relationship between thermolysin and neutral protease A of *Bacillus subtilis*. Proc. Nat. Acad. Sci. USA **72**, 4341 (1975).
492. *Hochachka P. W.*, Why study proteins of abyssal organisms? Comp. Biochem. Physiol. **52**, B supplement (1975).
493. *Carrell R. W., Lehmann H., Lorkin P. A., Raik E., Hunter E.*, Haemoglobin Sydney: β67 valin → alanin: An emerging pattern of unstable haemoglobins. Nature **215**, 626 (1967).
494. *Perutz M. F., Lehmann H.*, Molecular pathology of human haemoglobin. Nature **219**, 902 (1968).
495. *Takano T., Trus B. L., Mandel N., Mandel G., Kallai T. B., Swanson R., Dickerson R. E.*, Tuna cytochrome *c* at 2.0 Å resolution. J. Biol. Chem. **252**, 776 (1977).
496. *Ohno S.* Ancient linkage groups and frozen accidents. Nature **244**, 259 (1973).
497. *Smyth D. G.*, Constructing semisynthetic polypeptides. (editorial) Nature **256**, 699 (1975).

498. *Rees A. R., Offord R. E.*, The semisynthesis of portions of hen's egg lysozyme by fragments condensation. Biochem. J. **159**, 487 (1976).
499. *Gutte B.*, Shortened ribonuclease A; Conformational and mechanistic aspects. Xth Int. Congr. Biochem. (Hamburg), Proceedings. Abstr. 4-6-376 (1976).
500. *Ptitsyn O. B.*, Invariant features of globin primary structure and coding of their secondary structure. J. Mol. Biol. **88**, 287 (1974).
501. *Warme P. K., Momany F. A., Rumball S. V., Tuttle R. W., Scheraga H. A.*, Computation of structures of homologous proteins. Alphalactalbumin from lysozyme. Biochemistry **13**, 768 (1974).
502. *Margoliash E.*, Primary structure and evolution of cytochrome *c*. Proc. Nat. Acad. Sci. USA **50**, 672 (1963).
503. *Smith E. L., Margoliash E.*, Evolution of cytochrome *c*. Fed. Proc. **23**, 1243 (1964).
504. *Fitch W. M., Margoliash E.*, Construction of phylogenetic trees. Science **155**, 279 (1967).
505. *Nolan C., Margoliash E.*, Comparative aspects of primary structures of proteins. Annu. Rev. Biochem. **37**, 727 (1968).
506. *Smith E. L.*, Evolution of enzymes. The Enzymes **1**, 267 (1970).
507. *Reiner A. M.*, Genes for ribitol and D-arabitol catabolism in *Escherichia coli*: Their loci in C strains and absence in K-12 and B strains. J. Bacteriol. **123**, 530 (1975).
508. *Hartley B. S., Altossar I., Dothie J. M. Neuberger M. S.*, Experimental evolution of a xylitol dehydrogenase. In: Structure Function Relatioships of Proteins (R. Markham, R. W. Horne, eds.), pp. 191, North Holland, Amsterdam, New York, 1976.
509. *Dickerson R. E., Timkevich R.*, Cytochrome *c*. The Enzymes **11**, 397 (1975).
510. *Hunt L. T., Dayhoff M. O.*, Globins. Atlas of Protein Sequences **5**, suppl., 191 (1976).
511. *Garstang W.*, Preliminary note on the new theory of the phylogeny of the chordata. Zool. Ahz. **17**, 122 (1894).
512. *Watts D. C.*, Evolution of phoshagen kinases. In: Molecular Evolution (E. Schoffeniels, ed.), Vol. 2, pp. 150—173, North Holland, Amsterdam, New York, Oxford, 1971.
513. *Dayhoff M. O.*, Computer analysis of protein evolution. Sci. Amer. Juli, 86 (1969).
514. *King M.-C., Wilson A. S.*, Evolution at two levels in humans and chimpanzees. Science **188**, 107 (1975).
515. *Wilson A. C., Maxson L. R., Sarich V. M.*, Two types of molecular evolution. Evidence from studies of interspecific hybridization. Proc. Nat. Acad. Sci USA **71**, 2843 (1974).
516. *Wilson A. C.*, Evolutionary importance of gene regulation. Studler Symp. 7, 117 (1975).
517. *Wilson A. C., Sarich V. M., Maxson L. R.*, The importance of gene rearrangement in evolution: Evidence from studies on rates of chromosomal, protein, and anatomical evolution. Proc. Nat. Acad. Sci. USA **71**, 3028 (1974).
518. *Blake C. C. F., Johnson L. N., Mair G. A., North A. C. T., Phillips D. C., Sarma V. R.*, Crystallographic studies of the activity of hen egg-white lysozyme. Proc. Roy. Soc. Ser. B **167**, 378 (1967).
519. *Clarke P. H.*, Genes and enzymes. FEBS Lett **62**, E37—E46 (1976).
520. *Wills C.*, Production of yeast alcohol dehydrogenase isoenzymes by selection. Nature **261**, 26 (1976).
521. *Betz J. L., Brown P. R., Smyth M. J., Clarke P. H.*, Evolution in action. Nature **247**, 261 (1974).
522. *Rigby P. W., Burleigh B. D., Jr., Hartley B. S.*, Gene duplication in experimental enzyme evolution. Nature **251**, 200 (1974).
523. *Harley B. S., Burleigh B. D., Midwinter G. C., Moore C. H., Morris H. R.,*

Rigby P. W. J., Smith M. J., Taylor S. S., Where do new enzymes come from? FEBS Proc. **29**, 151 (1966).
524. *Hall B. C.*, Experimental evolution of a new enzymatic function. Kinetic analysis of the ancestral ebg^0 and evolved ebg$^+$ enzymes. J. Mol. Biol. **107**, 71 (1976).
525. *Ingram V. M.*, Gene evolution and the haemoglobins. Nature **189**, 704 (1961).
526. *Acher R.*, The neurohypophyseal hormones: An example of molecular evolution. In: Molecular Evolution (E. Schoffeniels, ed.), Vol. 2, pp. 43, North Holland, Amsterdam, New York, Oxford, 1971.
527. *Wichmann A.*, Ueber die Krystallformen der Albumine (Krystallisation von Lactalbumin). Z. Phys. Chem. **27**, 575 (1899).
528. *Hill R. L., Brew K.*, Lactose synthetase. Adv. Enzymol. **43**, 411 (1975).
529. *Fleming A.*, On a remarkable bacteriolytic element found in tissues and secretions. Proc. Roy. Soc. **B 93**, 306 (1922).
530. *Abraham E. P., Robinson R.*, Crystallization of lysozyme. «Miss D. Crowfoot has kindly undertaken their crystallografic examination», Nature **140**, 24 (1937).
531. *Phillips D. C.*, The hen egg-white lysozyme molecule. Proc. Nat. Acad. Sci. USA **57**, 484 (1967).
532. *Brew K., Vanaman T. C., Hill R. L.*, Comparison of the amino-acid sequence of bovine α-lactalbumin and hen's egg-white lysozyme. J. Biol. Chem. **242**, 3747 (1967).
533. *Brew K., Castellino F. J., Vanaman T. C., Hill R. L.*, The complete amino acid sequence of bovine α-lactalbumin. J. Biol. Chem. **245**, 4570 (1970).
534. *Browne W. J., North A. C. T., Phillips D. S., Brew K., Vanaman T. C., Hill R. L.*, A possible three-dimensional structure of bovine α-lactalbumin based on that of hen's egg-white lysozyme. J. Mol. Biol. **42**, 65 (1969).
535. *Warme P. K., Scheraga H. A.*, Refinement of the X-ray structure of lysozyme by complete energy minimization. Biochemistry **13**, 757 (1974).
536. *Stroud R. M.*, A family of protein-cutting proteins. Sci. Amer. July, **74** (1974).
537. *Blow D. M.*, Structure and mechanism of chymotrypsin. Acc. Chem. Res. **9**, 145 (1976).
538. *Beppu Y., Yomosa S.*, Molecular orbital studies on the enzimatic reaction mechanism of serine proteases. J. Physic. Soc. Japan **42**, 1694 (1977).
539. *Phillips D. C., Blow D. M., Hartley B. S., Lowe G.* (eds.), A discussion on the structures and functions of proteolytic enzymes, Phyl. Trans. Roy. Soc. London, Ser. **B 257**, 65 (1970).
540. *Poljak R. J.*, Three-dimensional structure, function and genetic control of immunoglobulins. Nature **256**, 373 (1975).
541. *Barker W. C., Dayhoff M. O.*, Immunoglobulins and related proteins. Atlas of Protein Sequence and Structure **5**, suppl. 2, 165 (1976).
542. *Уотсон Дж.* Молекулярная биология гена. — М.: Мир, 1978.
543. *Huber R.*, Antibody structure. Trends Biochem. Sci. **1**, 174 (1976).
544. *Davies D. R., Padlan E. A., Segal D. M.*, Three-dimensional structure of immunoglobulins. Annu. Rev. Biochem. **44**, 639 (1975).
545. *Poljak R. J., Amzel L. M., Avey H. P., Chen B. L., Phizackerley R. P., Saul F.*, Three-dimensional structure of the Fab'fragment of a human immunoglobulin at 2.8 Å resolution. Proc. Nat. Acad. Sci. USA **70**, 3305 (1973).
546. *Fridovich I.*, Superoxide dismutases. Adv. Enzymol. **41**, 35 (1974).
547. *Stromingler J. L.*, Histocompatibility antigens. Annu. Rev. Biochem., in press.
548. *Richardson J. S., Richardson D. C., Thomas K. A., Silverton E. W., Davies D. R.*, Similarity of three-dimensional structure between the immunoglobulin domain and the copper, zinc superoxide dismutase subunit. J. Mol. Biol. **102**, 221 (1976).
549. *Riggs A.*, Factors in the evolution of hemoglobin function. Fed. Proc. **35**, 2115 (1976).

550. *Perutz M. F.*, Haemoglobin. Annu. Rev. Biochem., in press.
551. *Kühne W.*, Ueber den Farbstoff der Muskeln. Virchows Arch. **33**, 79 (1865).
552. *Hoppe-Seyler F.*, Ueber die chemischen und optischen Eingenchaften des Blutfarbstoffs. Virchows Arch. **29**, 233 (1864).
553. *Perutz M. F., Rossmann M. G., Cullis A. F., Muirhead H., Will G., North A. C. T.*, Structure of haemoglobin. Nature **185**, 416 (1960).
554. *Edmundson A. B.*, Amino-acid sequence of sperm whale myoglobin. Nature **205**, 883 (1965).
555. *Braunitzer G., Gehring-Müller R., Hilschmann N., Hilse K., Hobom G. Rudloff V., Wittmann-Liebold B.*, Die Konstitution des normalen adulten Humanhämoglobins. Z. Physiol. Chem. **325**, 283 (1961).
556. *Keller R., Groudinsky O., Wüthrich K.*, Proton magnetic resonances in cytochrome b_2 core. Structural similarities with cytochrome b_5. Biochim. Biophys. Acta **328**, 233 (1973).
557. *Guiard B., Groudinsky O., Lederer F.*, Homology between bakers' yeast cytochrome b_2 and liver microsomal cytochrome b_5. Proc. Nat. Acad. Sci. USA **71**, 2539 (1974).
558. *Warme P. K., Hager L. P.*, Mesoheme sulfuric anhydride as a heme protein structure probe. Reaction with cytochrome b_{562}. Biochemistry **9**, 4244 (1970).
559. *Ozolz J., Strittmatter P.*, The homology between cytochrome b_5, hemoglobin and myoglobin. Proc. Nat. Acad. Sci. USA **58**, 264 (1967).
560. *Rossmann M. G., Argos P.*, A comparison of the heme binding pocket in globins and cytochrome b_5. J. Biol. Chem. **250**, 7525 (1975).
561. *Hagihara B., Sato N., Yamanaka T.*, Type *b* cytochromes. The Enzymes **11**, 549 (1975).
562. *Mathews F. S., Lederer F.*, Crystallographic study of bakers' yeast cytochrome b_2. J. Mol. Biol. **102**, 853 (1976).
563. *Czerwinski E. W., Mathews F. S., Hollenberg P., Drickamer K., Hager L. P.*, Crystallographic study of cytochrome b_{562} from *E. coli*. J. Mol. Biol. **71**, 819 (1972).
564. *Takano T., Swanson R., Kallai O. B., Dickerson R. E.*, Conformational changes upon reduction of cytochrome *c*. Cold Spring Harbor Symp. **36**, 397 (1971).
565. *Seybert D. W., Moffat K.*, Structure of haemoglobin reconstituted with mesoheme. J. Mol. Biol. **113**, 419 (1977).
566. *Argos P., Mathews F. S.*, The structure of ferrocytochrome b_5 at 2.8 Å resolution. J. Biol. Chem. **250**, 747 (1975).
567. *Strittmatter P.*, The nature of the heme binding in microsomal cytochrome b_5. J. Biol. Chem. **235**, 2492 (1960).
568. *Ozols J., Strittmatter P.*, The interaction of porphyrins and metalloporphyrins with apocytochrome b_5. J. Biol. Chem. **239**, 1018 (1964).
569. *Rossmann M. G. Argos P.*, Exploring structural homology of proteins. J. Mol. Biol. **105**, 75 (1976).
570. *John P., Whatley F. R., Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. Nature **254**, 495 (1975).
571. *Dickerson R. F., Timkovich R., Almassy R. J.*, The cytochrome fold and the evolution of bacterial energy metabolism. J. Mol. Biol. **100**, 473 (1976).
572. *Connelly J. L., Jones O. T. G., Saunders V. A., Yates D. W.*, Kinetic and thermodynamic properties of membrane-bound cytochromes of aerobically and protosynthetically grown *Rhodopseudomonas spheroides*. Biochim. Biophys. Acta **292**, 644 (1973).
573. *Horowitz N. H.*, On the evolution of biochemical syntheses. Proc. Nat. Acad. Sci. USA **31**, 153 (1945).
574. *Astbury W. T.*, Adventures in molecular biology. Harvey Lect. **46**, 3 (1950).
575. *Vosberg H. P.*, Molecular cloning of DNA. Human Gen. **40**, 1 (1977).
576. *Brownlee G. G., Cartwright E. M., Cowan N. J., Jarvis M., Milstein C.*, Purification and sequence of messenger RNA for immunoglobulin light chains. Nature New Biol. **244**, 236 (1973).

577. *Tonegawa S., Hozumi N., Matthyssens G., Schuller R.*, Somatic changes in the content and context of immunoglobulin genes. Cold Spring Harbor Symp. **41**, 877 (1977).
578. *Truffa-Bachi P., Veron M., Cohen G. N.*, Structure, function, and possible origin of a bifunctional allosteric enzyme, *Escherichia coli* aspartokinase I-homoserine dehydrogenase I. Crit. Rev. Biochem. **2**, 379 (1974).
579. *Crawford I. P.*, Gene rearrangements in the evolution of the tryptophan synthetic pathway. Bacteriol. Rev. **39**, 87 (1975).
580. *Schweizer E., Knobling A., Manger G., Dietlein G.*, Gene fusion in the fatty acid synthetase system of *Saccharomyces cerevisiae*. FEBS Proc. **32**, 133 (1975).
581. *Yourno J., Kohno T., Roth J. R.*, Enzyme evolution: Generation of a bifunctional enzyme by fusion of adjacent genes. Nature **228**, 820 (1970).
582. *Tartof K. D.*, Redundant genes. Annu. Rev. Gen. **9**, 355 (1975).
583. *Hood L., Campbell J. H., Elgin S. C. R.*, The organisation, expression and evolution of antibodies and other multigene families. Annu. Rev. Genet. **9**, 305 (1975).
584. *Bridges C. B.*, The BAR «gene», a duplication. Science **83**, 210 (1936).
585. *Hood L.*, Antibody genes and other multigene families. Fed. Proc. **35**, 2158 (1976).
586. *Zuckerkandl E.*, The appearance of new structures and functions in proteins during evolution. J. Mol. Evol. **7**, 1 (1975).
587. *McLachlan A. D.*, Repeating sequence and gene duplications in proteins. J. Mol. Biol. **64**, 417 (1972).
588. *Black J. A., Dixon G. H.*, Amino-acid sequence of alpha chains of human haptoglobins. Nature **218**, 736 (1968).
589. *McLachlan A. D., Walker J. E.*, Evolution of serum albumin. J. Mol. Biol. **112**, 543 (1977).
590. *Collins J. H.*, Homology of myosin DTNB light chain with alkali light chains, troponin C and parvalbumin. Nature **259**, 699 (1976).
591. *Cantor C. R., Jukes T. H.*, The repetition of homologous sequences in the polypeptide chains of certain cytochromes and globins. Proc. Nat. Acad. Sci. USA **56**, 177 (1966).
592. *Russell R. L., Abelson J. N., Landy A., Gefter M. L., Brenner S., Smith J.D.*, Duplicate genes for tyrosine transfer RNA in *Escherichia coli*. J. Mol. Biol. **47**, 1 (1970).
593. *Ycas M.*, Origin of periodic proteins. Fed. Proc. **35**, 2139 (1976).
594. *Hendrickson W. A.*, The molecular architecture of oxygen-carrying proteins. Trends Biochem. Sci. **2**, 108 (1977).
595. *Bridgen J., Harris J. I., Northron F.*, Evolutionary relationships in superoxide dismutase· FEBS Lett. **49**, 392 (1975).
596. *Subramanian E.*, Molecular structure of acid-proteases. Trends Biochem.Sci. **3**, 1 (1978).
597. *Edmundson A. B., Ely K. R., Abola E. E., Schiffer M., Panagiotopoulos N., Deutsch H. F.*, Conformational isomerism, rotational allomerism, and divergent evolution in immunoglobulin light chains. Fed. Proc. **35**, 2119 (1976).
598. *McLachlan A. D.*, Tests for comparing related amino acid sequences. Cytochrome *c* and cytochrome c_{551}. J. Mol. Biol. **61**, 409 (1971).
599. *McLachlan A. D., Stewart M.* The 14-fold periodicity in α-tropomyosin and the interaction with actin. J. Mol. Biol. **103**, 271 (1976).
600. *McLachlan A. D.*, Gene duplication in carp muscle calcium binding protein. Nature New Biol. **240**, 83 (1972).
601. *Rossmann M. G., Argos P.*, The taxonomy of protein structure. J. Mol. Biol. **109**, 99 (1977).
602. *Schirmer R. H., Schirmer I.*, The first enzyme — Kühne's trypsin — a hundred years old favourite. Pharma Int. **5/6**, 10 (1976).
603. *Cunningham B. A., Wang J. L., Berggård I., Peterson P. A.*, The complete

amino acid sequence of β_2-microglobulin. Biochemistry **12**, 4811 (1973).
604. *Evans H. J., Steinmann H. M., Hill R. L.*, Bovine erythrocyte superoxide dismutase II. J. Biol. Chem. **249**, 7315 (1974).
605. *Steinmann H. M., Naik V. R., Abernethy J. L., Hill R. L.*, Bovine erytrocyte superoxide dismutase I. J. Biol. Chem. **249**, 7326 (1974).
606. *Abernethy J. L., Steinmann H. M., Hill R. L.*, Bovine erythrocyte superoxide dismutase III. J. Biol. Chem. **249**, 7339 (1974).
607. *Segal D. M., Padlan E. A., Cohen G. H., Rudikoff S., Potter M., Davies D. R.*, The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. Proc. Nat. Acad. Sci. USA **71**, 4298 (1974).
608. *Colman P. M., Deisenhofer J., Huber R., Palm W.*, Structure of the human antibody molecule Kol (Immunoglobulin G1). Am electron density map at 5 Å resolution. J. Mol. Biol. **100**, 257 (1976).
609. *Itagaki E., Hager L. P.*, Studies on cytochrome b_{562} of *Escherichia coli*. J. Biol. Chem. **241**, 3687 (1966).
610. *Oshino N., Imai Y., Sato R.*, A function of cytochrome b_5 in fatty acid desaturation by rat liver microsomes. J. Biochem. (Tokyo) **69**, 155 (1971).
611. *Ozols J., Strittmatter P.*, Correction of the amino acid sequence of calf liver microsomal cytochrome b_5. J. Biol. Chem. **244**, 6617 (1969).
612. *Tsugita A., Kobyashi M., Tani S., Kyo S., Rashid M. A., Yoshida Y., Kajihara T., Hagihara B.*, Comparative study of the prymary structures of cytochrome b_5 from four speries. Proc. Nat. Acad. Sci. USA **67**, 442 (1970).
613. *Itagaki E., Hager L. P.*, The amino acid sequence of cytochrome b_{562} of *Escherichia coli*. Biochem. Biophys. Res. Commun. **32**, 1013 (1968).
614. *Cavadore J. G.*, Polycondensations d'α-amino acides en milieu aqueux. Thèse doctorat ès sciences physiques, Université des Sciences et Techniques du Languedoc, Montpellier, 1971.
615. *Kretsinger R. H.*, Calcium-binding proteins. Annu. Rev. Biochem. **45**, 239 (1976).
616. *Nozaki M.*, Structure and function of metalloenzymes (review). J. Syn. Org. Chem. **34** (NM), 805 (1976).
617. *Potter M., Rudikoff S., Vrana M., Rao D. N., Mushinksi E. B.*, Primary structural differences in myeloma proteins that bind the same haptens. Cold Spring Harbor Symp. **41**, 661 (1976).
618. *Amzel L. M., Poljak R. J, Saul F., Varga B. M., Richards F. F.*, The three-dimensional structure of a combining region-ligand complex of immunoglobulin NEW at 3.5 Å resolution. Proc. Nat. Acad. Sci. USA **71**, 1427 (1974).
619. *Wu T. T., Kabat E. A.*, An analysis of the sequences of the variable regions of Bence-Jones Proteins and myeloma light chains and their implications for antibody complementarity. J. Exp. Med. **132**, 211 (1970).
620. *Padlan E. A., Davies D. R., Pecht I., Givol D., Wright C.*, Model-building studies of antigen-binding sites: The hapten-binding site of MOPC-315. Cold Spring Harbor Symp. **41**, 627 (1976).
621. *Dwek R. A., Wain-Hobson S., Dover S., Gettins P., Sutton B., Perkins S. J., Givol D.*, Structure of an antibody combining site by magnetic resonance. Nature **266**, 31 (1977).
622. *Hunkapiller M. W., Goetze A. M., Richards J. H., Hood L. E.*, Structure-function correlates for DNP and PC binding myeloma proteins. Immunogenetics **4**, 424 (1977).
623. *Klotz J., Hunstron D. L.*, Protein-interaction with small molecules. J. Biol. Chem. **250**, 3001 (1975).
624. *Jolley M. E., Rudikoff S., Potter M., Glaudemans C. P.*, Spectral changes on binding of oligosaccharides to murine IgA myeloma immunoglobulins. Biochemistry **12**, 3039 (1973).
625. *Kabat E. A.*, In: Experimental Immunochemistry (E. A. Kabat, M. M. May-

er, eds.), 2nd ed., Charles C. Thomas, Springfield, Illinois (1961).
626. *Citri N.*, Conformational adaptability in enzymes. Adv. Enzymol. **37**, 397 (1973).
627. *Robertus J. D., Alden R. A., Birktoft J. J., Kraut J., Powers J. C., Wilcox P. E.*, An X-ray crystallographic study of the binding of peptide chloromenthyl ketone inhibitors to subtilisin BPN'. Biochemistry **11**, 2439 (1972).
628. *Blow D. M., Birktoft J. J., Hartley B. S.*, Role of a buried acid group in the mechanism of action of chymotrypsin. Nature **221**, 337 (1969).
629. *Ingles D. W., Knowles J. R.*, Specificity and stereo-specificity of α-chymotrypsin. Biochem. J. **104**, 369 (1967).
630. *Haldane J. B. S.*, Enzymes. Longmans, Green, London (1930).
631. *Jencks W. P.*, Binding energy, specificity and enzymic catalysis: The Circe effect. Adv. Enzymol. **43**, 220 (1975).
632. *Sweet R. M., Wright H. T., Janin J., Chothia C. H., Blow D. M.*, Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6 Å resolution. Biochemistry **13**, 4212 (1974).
633. *Fischer H., Zeile K.*, Synthese des Hämatoporphyrins, Protoporphyrins und Hämins. Ann. Chem. **468**, 98 (1929).
634. *Baldwin J. M.*, Structure and function of haemoglobin. Progr. Biophys. Mol. Biol. 29, 225 (1975).
635. *Watson H. C., Nobbs C. L.*, The Structure of Oxygenated and Deoxygenated Myoglobin, Collog. Ges. Biol. Chem. (Mosbach/Baden). Springer-Verlag, Heidelberg, 1968, pp. 37—48.
636. *Watson H. C.*, The stereochemistry of the protein myoglobin. Progr. Stereochem. **4**, 299 (1969).
637. *Falk J. E.*, Porphyrins and Mettalloporphyrins, BBA Library 2, Elsevier, Amsterdam — London — New York, 1964.
638. *Le Gall J., Postgate J. R.*, Bacterial cytochromes *c*. Adv. Microbiol. Physiol. **10**, 81 (1973).
639. *Wang J. H.*, Synthetic biochemical models. Accounts Chem. Res. **3**, 90 (1970).
640. *Salemme F. R.*, A Hypothetical structure for an intermolecular electron transfer complex of cytochromes *c* and b_5, J. Mol. Biol. **102**, 563 (1976).
641. *Maehley A. C.*, Haemin-protein binding in peroxidase and methaemalbumin. Nature, 192, 630 (1961).
642. *Schichi H., Hackett D. P., Funatsu C.*, Studies on the *b*-type cytochromes from mung bean seedlings. J. Biol. Chem. **238**, 1156 (1963).
643. *Nobbs C. L., Watson H. C., Kendrew J. C.*, Structure of deoxymyoglobin: A crystallographic study. Nature. **209**, 339 (1966).
644. *Maltempo M. M., Moss T. H., Cusanovich M. A.*, Magnetic studies on the changes in the environment in *chromatium* ferricytochrome c^1. Biochim. Biophys. Acta, **342**, 290 (1974).
645. *Harrison S. C., Blout E. R.*, Reversible conformational changes of myoglobin and apomyoglobin. J. Biol. Chem. **240**, 299 (1965).
646. *Lim V. I., Efimov A. V.*, The folding pathway for globins. FEBS. Lett. 78, 279 (1977).
647. *Corwin A. H., Erdman J. G.*, A synthetic ferroporphyrin complex that is passive to oxygen. J. Amer. Chem. Soc. **68**, 2473 (1946).
648. *Collmann J. P., Brauman J. J., Halbert T. R., Suslick K. S.*, Nature of O_2 and CO binding to mettalloporphyrins and haem proteins. Proc. Nat. Acad. Sci. USA **73**, 3333 (1976).
649. *Rougee M., Brault D.*, Influence of trans weak or strong field ligands upon the affinity of deuteroheme for carbon monoxide. Monoimidazoleheme as a reference for unconstrained five-coordinate hemoproteins. Biochemistry **14**, 4100 (1975).
650. *Pauling L.*, Nature of the iron-oxygen bond in oxyhaemoglobin. Nature **203**, 182 (1964).

651. *Collmann J. P., Gagne R. R., Reed C. A., Robinson W. T., Rodley G. A.*, Structure of an iron II dioxygen complex; a model for oxygen-carrying hemeproteins. Proc. Nat. Acad. Sci. USA **71**, 1326 (1974).
652. *Weiss J. J.*, Nature of the iron-oxygen bond in oxyhaemoglobin. Nature **202**, 83 (1964).
653. *Thomson A. J.*, Ligand binding properties of the haem group. Nature **265**, 15 (1977).
654. *Hoard J. L.*, In: Porphyrins and Mettalloporphyrins (K. M. Smith, ed.), Elsevier, New York, 1975, p. 356.
655. *Huber R., Epp O., Formanek H.*, Structures of deoxy-and carbonmonoxyerythrocruorin. J. Mol. Biol. **52**, 349 (1970).
656. *Padlan E. A., Love W. E.*, Three-dimensional structure of hemoglobin from the polychaete annelid, *Glycera dibranchiata*, at 2.5 Å resolution. J. Biol. Chem. **249**, 4067 (1975).
657. *Norvell J. G., Nunes A. C., Schoenborn B. P.*, Neutron diffraction analysis of myoglobin: Structure of the carbon monoxide derivative. Science **190**, 568 (1975).
658. *Kaminsky L. S., Miller V. J.*, The ascorbate reduction of denatured ferricytochrome *c*. Biochem. Biophys. Res. Commun. **49**, 252 (1972).
659. *Moore G. R., Williams R. J. P.*, Structural basis for the variation in redox potential of cytochromes. FEBS Lett. **79**, 229 (1977).
660. *Kassner R. J.*, Effects of nonpolar environments on the redox potentials of heme complexes. Proc. Nat. Acad. Sci. USA **69**, 2263 (1972).
661. *Freeman H. C.*, Crystal structures of metal-peptide complexes. Adv. Prot. Chem. **22**, 257 (1967).
662. *Perutz M. F.*, The haemoglobin molecule. Proc. Roy. Soc. (B) **173**, 113 (1969).
663. *Pauling L., Coryell C. D.*, The magnetic properties and structure of the hemochromogens and related substances. Proc. Nat. Acad. Sci. USA **22**, 159 (1936).
664. *Hoard J. L.*, In: Hemes and Hemoproteins (B. Chance, R. W. Estabrook, T. Yonetani, eds.), Academic Press, New York and London, 1966.
665. *Pai E. F., Sachsenheimer W., Schirmer R. H., Schulz G. E.*, Substrate positions and induced-fit in crystalline adenylate kinase. J. Mol. Biol. **114**, 37 (1977).
666. *Perutz M. F.*, Structure and mechanism of haemoglobin. Br. Med. Bull. **32**, 195 (1976).
667. *Edsall J. T.*, Blood hemoglobin. J. Hist Biol. **5**, 205 (1972).
668. *Kilmartin J. V., Rossi-Bernardi L.*, Interaction of haemoglobin with hydrogen ions, Carbon dioxide, and organic phosphates. Physiol. Rev. **53**, 836 (1973).
669. *Baggot J.*, The magnitudes of the Bohr and Haldane effects. Trends Biochem. Sci. **3**, N30 —N32, (1978).
670. *Benesch R., Benesch R. E.*, Interacellular organic phosphates as regulators of oxygen release by haemoglobin. Nature **221**, 618 (1969).
671. *Tyuma I., Shimizu K.*, Effect of organic phosphates on the difference in oxygen affinity between fetal and human haemoglobin. Fed. Proc. **29**, 1112 (1970).
672. *Arnone A.*, X-ray diffraction study of binding of 2,3-diphosphoglycerate to human deoxyhaemoglobin. Nature **237**, 146 (1972).
673. *Lenfant C., Torrance J. D., Woodson R. D., Jacobs P., Finch C. A.*, Role of organic phosphates in the adaptation of man to hypoxia. Fed. Proc. **29**, 1115 (1970).
674. *Bohr C.*, Theoretische Behandlung der quantitativen Verhältnisse bei der Sauerstoffaufnahmen des Häemoglobins. Zentr. Physiol. **17**, 682 (1903).
675. *Bohr C., Hasselbalch K., Krogh A.*, Über einen in biologischer Beziehung

wichtigen Einfluss, den die Kohlensäurespannung des Blutes auf dessen Sauerstoffbindung übt. Skand. Arch. Physiol. **16**, 402 (1904).
676. *Christiansen J., Douglas C. G., Haldane J. S.*, The absorption and dissociation of carbon dioxide by human blood. J. Physiol. **48**, 245 (1914).
677. *Hill A. V.*, The possible effects o the aggregation of the molecules of haemoglobin on its dissociation curves. J. Physiol. **40**, IV (1910).
678. *Adair G. S.*, The haemoglobin system. J. Biol. Chem. **63**, 493 (1925).
679. *Wyman J., Allen D. W.*, The problem of the heme interactions in haemoglobin and the basis of the Bohr effect. J. Polymer Sci. **7**, 499 (1951).
680. *Wyman J.*, Heme proteins. Adv. Prot. Chem. **4**, 407 (1948).
681. *Monod J., Changeux J. P., Jacob F.*, Allosteric proteins and sellular control systems. J. Mol. Biol. **6**, 306 (1963).
682. *Koshland D. E., Jr., Nemethy G., Filmer D.*, Comparison of experimental binding data and theoretical models in proteins containing subunit. Biochemistry **5**, 365 (1966).
683. *Eventoff W., Rossmann M. G.*, The evolution of dehydrogenases and kinases. CRC Crit. Rev. Biochem. **3**, 112 (1975).
684. *Katunuma N., Kito K., Kominami E.*, A new enzyme that specifically inactivates apo-protein of NAD-dependent dehydrogenases. Biochem. Biophys. Res. Commun. **45**, 76 (1971).
685. *Einarsson R., Eklund H., Zeppezauer E., Boiwe T., Brändén C.-I.*, Binding of salicylate in the adenosine-binding pocket of dehydrogenases. Eur. J. Biochem. **49**, 41 (1974).
686. *Koshland D. E., Jr.*, Application of a theory of enzyme specifity to protein synthesis. Proc. Nat. Acad. Sci. USA **44**, 98 (1958).
687. *Koshland D. E.*, Protein shape and biological control. Sci. Amer. October, 52 (1973).
688. *Sachsenheimer W., Schulz G. E.*, Two conformations of cryctalline adenylate kinase. J. Mol. Biol. **114**, 23 (1977).
689. *Richards F. M., Wyckoff H. W.*, Bovine pancreatic ribonuclease. The Enzymes **4**, 647 (1971).
690. *Phillips S. E. V.*, Structure of oxymyoglobin. Nature **273**, 247 (1978).
691. *Webb L. E., Hill E. J., Banaszak L. J.*, Conformation of nicotinamide adenide dinucleotide bound to cytoplasmic malate dehydrogenase. Biochemistry **12**, 5101 (1973).
692. *Ohlsson I., Nordström B., Brändén C.-L.*, Structural and functional similarities within the coenzyme binding domains of dehydrogenases. J. Mol. Biol. **89**, 339 (1974).
693. *Williams R. E.*, Phosphorylated sites in substrates of intracellular protein kinases: A common feature in amino acid sequences. Science **192**, 473 (1976).
694. *Scanu A. M.*, Plasma lipoprotein structure (editorial). Nature **270**, 209 (1977).
695. *Fox C. F.*, The structure of cell membranes. Sci. Amer. **226**, 30 (1972).
696. *Singer S. J., Nicolson G. L.*, The fluid mosaic model of the structure of cell membranes. Science **175**, 720 (1972).
697. *Metzler D. E.*, Biochemistry. Academic Press, New York, 1977.
698. *Ozols J., Gerard C.*, Covalent structure of the membranous segment of horse cytochrome p_5. J. Biol. Chem. **252**, 8549 (1977).
699. *Henderson R., Unwin P. N. T.*, Three-dimensional model of purple membrane obtained by electron microscopy. Nature **257**, 28 (1975).
700. *Blaurock A. E.*, Bacteriorhodopsin: a trans-membrane pump containing α-helix. J. Mol. Biol. **93**, 139 (1975).
701. *Hasselbach W.*, The sarcoplasmic calcium pump— a most efficient ion translocating system. Biophys. Struct. Mech. **3**, 43 (1977).
702. *MacLennan D. H., Holland P. C.*, Calcium transport in sarcoplasmic reticulum. Annu. Rev Biophys. Bioeng. **4**, 377 (1975).
703. *Hasselbach W., Makinose M.*, Die Calciumpumpe der Erschlaffungsgrana

des Muskels und ihre Abhängigkeit von der ATP-Spaltung. Biochem. Z. 333, 518 (1961).
704. *Makinose M., Hasselbach W.*, ATP synthesis by the reverse of the sarcoplasmic calcium pump. FEBS Lett. **12**, 271 (1971).
705. *Rothfield L., Romeo D., Hinckley A.*, Reassembly of purifield bacterial membrane components. Fed. Proc. **31**, 12 (1972).
706. *Levey G. S.*, Restoration of norepinephrine responsiveness of solubilised myocardial adenylate cyclase by phosphatidylinositol. J. Biol. Chem. **246**, 7405 (1971).
707. *Martonosi A. N.*, In: Calcium Transport in Secretion and Contraction (E. Carafoli, F, Clementi, W. Drabikowski A. Margreth, eds.), North Holland, Amsterdam, 1975, pp. 77—86.
708. *Brown P. K.*, Rhodopsin rotates in the visual receptor membrane. Nuture New Biol. **236**, 35 (1972).
709. *Nicolson G. L.*, Transmembrane control of the receptors on normal and tumor cells. Biochem. Biophys. Acta **457**, 57 (1976).
710. *Singer S. J.*, The molecular organisation of membranes. Annu. Rev. Biochem. **43**, 805 (1974).
711. *Kornfeld R., Kornfeld S.*, Comparative aspects of glycoprotein structure. Annu. Rev. Biochem. **45**, 217 (1976).
712. *Marshall R. D.* Glycoproteins. Annu. Rev. Biochem. **41**, 673 (1972).
713. *Edelman G. M.*, Summary: Understanding selective molecular recognition. Cold Spring Harbor Symp. **39**, 891 (1977).
714. *Wittmann H. G.*, Structure, function and evolution of ribosomes. Eur. J. Biochem. **61**, 1 (1976).
715. *Kurland C. G.*, Structure and function of the bacterial ribosome. Annu. Rev. Biochem. **46**, 173 (1977).
716. *Liljas A., Eriksson S., Donner D., Kurland C. G.*, Isolation and crystallization of stable domains of the protein L7/L12 from *Escherichia coli* ribosomes. FEBS. Lett. **88**, 300 (1978).
717. *Leberman R., Wittinghofer A., Schulz G. E.*, Polymorphism in crystalline elongation factor Tu·GDP from *E. coli*. J. Mol. Biol. **106**, 951 (1976).
718. *Kabsch W., Gast W. H., Schulz G. E., Leberman R.*, Low resolution structure of partially trypsin-degraded polypeptide elongation factor, EF-Tu, from *E. coli*. J. Mol. Biol. **117**, 999 (1977).
719. *Steitz T. A., Stenkamp R. E., Geisler N., Weber K., Finch J*, X-ray and electron microscopic studies of crystals of native and proteolytically cleaved lac repressor protein. Int. Symp. Biomol. Struct. Conform., Funct. and Evol. Madras (1978).
720. *Beyreuther K., Raufuss H., Schrecker O., Hengstenberg W.*, The phosphoenolpyruvate-dependent phosphotransferase system of *Staphylococcus aureus*. Eur. J. Biochem. 75, 275 (1977).
721. *Winzler R. J.*, Glycoproteins. In: Handbook of Biochemistry, Publ. Chemical Rubber Co., Cleveland, Ohio, 1970, p. C42—C49.
722. *Gallagher J. T., Corfield A. P.*, Mucin-type glycoproteins — new perspectives on their structure and synthesis. Trends Biochem. Sci. 3, 38 (1978).
723. *DeVries A. L.*, Glycoproteins as biological antifreeze agents in Antarctic fish. Science **172**, 1152 (1971).
724. *Feeney R. E., Osaga D. T.*, Polar fish proteins. Trends Biochem. Sci. 2, 269 (1977).
725. *Nolsestuen G. L., Suttie L. W.*, Properties of asialo-and aglycoprothrombin. Biochem. Biophys. Res. Commun. **45**, 198 (1971).
726. *Edelman G. M., Yahara J., Wang J. L.*, Receptor mobility and receptor cytoplasmic interactions in lymphocytes. Proc. Nat. Acad. Sci. USA **70**, 1442 (1973).
727. *Bayer E., Holzbach G.*, Synthetische Hämolopolymers zur reversiblen Anlagerung von molekularem Sauerstoff. Angew. Chem. **89**, 120 (1977).

728. *Schmidt F. O., Schneider D. M., Crothers D. M.*, (eds.), Functional Linkage in Biomolecular Systems. Raven Press, New York, 1975.
729. *v. Bertalanffy L.*, General System Theory. George Braziller, New York, 1968.
730. *Koshland D. E.*, The comparison of non-enzymic and enzymic reaction velocities. J. Theoret. Biol. **1**, 75 (1962).
731. *Jencks W. P.*, Catalysis in Chemistry and Enzymology. McGraw-Hill, New York, 1969 chap. 5.
732. *Hammes G.G.*, Mechanism of enzyme catalysis. Nature **204**, 342 (1964).
733. *Hamilton C. L., Nieman C., Hammond G. S.*, A quantitative analysis of the binding of N-acyl derivatives of α-aminoamides by α-chymotrypsin. Proc. Nat. Acad. Sci. USA **55**, 664 (1966).
734. *Bender M. L., Kezdy F. J.*, The current status of the α-chymotrypsin mechanism. J. Amer. Chem. Soc. **86**, 3704 (1964).
735. *Bürgi H. B., Shefter E., Dunitz J. D.*, Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group. J. Amer. Chem. Soc. **95**, 5065 (1973).
736. *Hunkapiller M., Smallcombe S. H., Whitaker D. R., Richards J. H.*, Ionization behaviour of the histidine residue in the catalytic triad of serine proteases. J Biol. Chem. **248**, 8306 (1973).
737. *Garavito R. M., Rossmann M. G., Argos P., Eventoff W.*, Convergence of active center geometries. Biochemistry **16**, 5065 (1977).
738. *Pauling L.*, Molecular architecture and biological reactions. Chem. Engineer. News **24**, 1375 (1946).
739. *Wolfenden R.*, Analog approaches to the structure of the transition state in enzyme reactions. Acc. Chem. Res. **5**, 10 (1972).
740. *Lienhard G. E.*, Enzymatic catalysis and transition state theory. Science **180**, 149 (1973).
741. *Page M. I., Jencks W. P.*, Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. Proc. Nat. Acad. Sci. USA **68**, 1678 (1971).
742. *Bruice T. C. B.*, Proximity effects and enzyme catalysis. The Enzymes **2**, 217 (1970).
743. *Storm D. R., Koshland D. E.*, A source for the special catalytic power of enzymes: Orbital steering. Proc. Nat. Acad. Sci. USA **66**, 445 (1970).
744. *Storm D. R., Koshland D. E., Jr.*, An indication of the magnitude of orientation factors in esterification, J. Amer. Chem. Soc. **94**, 5805 (1972); Effect of small changes in orientation on reaction rate. J. Amer. Chem. Soc. **94**, 5815 (1972).
745. *Roberts G. C. K.*, What's special about enzymes? (editorial). Nature **271**, 409 (1978).
746. *Crosby J., Stone R., Lienhard G. E.*, Mechanisms of thiamine-catalyzed reactions. Decarboxylation of 2-(1-carboxyl-1-hydroxyethyl)-3,4-dimethylthiazolium chlotide, J. Amer. Chem. Soc. **92**, 2891 (1970).
747. *Lipscomb W. N., Hartsuck J. A., Quioche F. A., Reeke G. N., Jr.*, The structure of carboxypeptidase A. Proc. Nat. Acad. Sci. USA **64**, 28 (1969).
748. Ford *L. O., Jehnson L. N., Machin P. A., Philips D. C., Tjian R.*, Crystal structure of a lysozyme-tetrasaccharide lactone complex. J. Mol. Biol. **88**, 349 (1974).
749. *Chothia C. H., Janin J.*, Stability and specifity of protein-protein interactions: The case of trypsin-trypsin inhibitor complexes. J. Mol. Biol., **100**, 197 (1976).
750. *Wulff G.*, Synthetische Polymere mit chiralen Hohlräumen. Nachr. Chem. Techn. Lab. **25**, 239 (1977).
751. *Watts D. C.*, Creatine kinase. The Enzymes **8**, 383 (1973).
752. *Turner D. C., Wallimann T., Eppenberger H. M.*, A protein that binds spe-

cifically to the M-line of skeletal muscle is identified as the muscle form of creatine kinase. Proc. Nat. Acad. Sci. USA **70**, 702 (1973).
753. *Atkinson D. E.*, Functional linkage as basis principle of biology, In: Functional Linkage in Biomolecular Systems (F. O. Schmitt, Schneider D. M., Crothers D. M., eds). Raven Press, New York, 1975, p. 43.
754. *Holmes K. C.*, Selbstorganisation biologischer Strukturen. Klin. Wochenschr. **53**, 997 (1975).
755. *Huxley H. E.*, Muscular contraction and cell motility. Nature **243**, 445 (1973).
756. *Clarke M., Spudich J. A.*, Nonmuscle contractile proteins: The role of actin and myosin in cell motility and shape determination. Annu. Rev. Biochem. **46**, 797 (1977).
757. *DeRosier D., Mandelkow E., Silliman A., Tilney L., Kane R.*, Structure of actin-containing filaments from two types of nonmyscle cells. J. Mol. Biol. **113**, 679 (1977).
758. *Mannherz H. G., Goody R. S.*, Proteins of contractile systems. Annu. Rev. Biochem. **45**, 427 (1976).
759. *Huxley A. F., Niedergerke R.*, Structural changes in muscle during contraction. Nature **173**, 971 (1954).
760. *Huxley H. E., Hanson J.*, Changes in the cross-striations of muscle during contraction and strech and their structural interpretation. Nature **173**, 973 (1954).
761. *Koshland D. E., Jr., Neet K. E.*, The catalitic and regulatory properties of enzymes. Annu. Rev. Biochem. **37**, 359 (1968).
762. *Diebler H., Eigen M., Hammes G. G.*, Relaxations-spektrometrische Untersuchungen schneller Reaktionen von ATP in wässriger Lösung. Z. Naturforsch. **15b** 554 (1960).
763. *Chance B.*, Catalysis in biochemical reactions. Z. Elektrochem. **64**, 7 (1960).
764. *Eigen M., De Maeyer L.*, Relaxation methods. In: Investigazion of Rates And Mechanisms of Reaction (Friess S. L., Lewis E. S., Weissberger A., eds.), Techique of Ogganic Chemistry, Vol. 8, Part II, 1963, p. 895.
765. *Einstein A.*, Schallausbreitung in teilweise dissoziierten Gasen. Sitz. ber. Preuss. Akad. Wiss. Phys. Math. Kl. Berlin, 1920, p. 382.
766. *Hartridge H., Reughton F. J. W.*, A method of measuring the velocity of very rapid chemical reactions. Proc. Roy. Soc. Ser. A**104**, 376 (1923).
767. *Huxley H. E.*, Structural changes in the actin- and myosin-containing filaning filaments during contraction. Cold Spring Harbor Symp. Quant. Biol. **37**, 361 (1972).
768. *Spudich J. A., Huxley H. E., Finch J. T.*, Regulation of skeletal muscle contraction. II. Structural studies of the interaction of the tropomyosinetroponin complex with actin. J. Mol. Biol. **72**, 619 (1972).
769. *Parry D. A.D., Squire J. M.*, Structural role of tropomyosin in muscle regulation: Analyse of the X-ray diffraction patterns from relaxed and contracting muscles. J. Mol. Biol. **75**, 33 (1973).
770. *Hasselbach V.*, The sarcoplasmic calcium transport, a basic mechanism of the regulation of muscle activity, Verh. Dtsch. Zool. Ges., **1976**, 100 (1976).
771. *Ebashi S., Endo M.*, Calcium ion and muscle contraction. Progr. Biophys. Mol. Biol. **18**, 125 (1968).
772. *Cohen C.*, The protein switch of muscle contraction. Sci. Amer. November, 36 (1975).
773. *Weber A.*, Regulation of muscle contraction through cooperative interaction in the actin filament. In: Functional Linkage in Biomolecular Systems (Schmidt F. O., Schneider D. M., Crothers D. M., eds.), Raven Press, New York, 1975, p. 312.
774. *Kirschner K.*, Kinetic analysis of allosteric enzymes. Curr. Top. Cell. Reg. **4**, 167 (1971).
775. *Dische Z.*, Interdependence of various enzymes of the glycolytic system and the automatic regulation of their activity within the cells. I. Inhibition of

the phosphorylation of glucose in red corpuscles by monophosphoglyceric and diphosphoglyceric acids; state of the diphosphorylation of glucose. Bull. Soc. Chim. Biol. (Paris) **23**, 1140 (1971).
776. *Umbarger H. E.*, Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. Science **123**, 848 (1956).
777. *Walsh D. A., Ashby C. D.*, Protein kinases: Aspects of their regulation and diversity. Recent Progr. Horm. Res. **29**, 329 (1973).
778. *Robinson G. A., Butcher F. R. W., Sutherland E. W.*, Cyclic AMP, Academic Press, New York, 1971.
779. *Hagins W. A., Yoshikami S.*, A role for Ca^{2+} in excitation of retinal rods and cones. Exp. Eye Res. **18**, 299 (1974).
780. *O'Dell D.*, Egg activation (editorial). Nature **270**, 558 (1977).
781. *Eigen M., Hammes G.*, Elementary steps in enzyme reactions (as studied by relaxation spectrometry). Adv. Enzymol. **25**, 1 (1963).
782. *Eigen M.*, Construction of control and self-organizing systems. In: Functional linkage in Biomolecular Systems (Schmidt F. O., Schneider D. M., Crothers D. M., eds.), Raven Press, New York, 1975, p. 38.
783. *Rüegg J. C., Tregear R. T.*, Mechanical factors affecting the ATPase activity of glycerol-extracted insect fibrillar flight muscle. Proc. Roy. Soc. **B165**, 497 (1966).
784. *Pringle J. W. S.*, The contractile mechanism of insect fibrillar muscle. Progr, Biophys. Mol. Biol. **17**, 3 (1967).
785. *Weber A., Murray J. M.*, Molecular control mechanism in muscle contaction. Physiol. Rev. **53**, 612 (1973).
786. *Pohl F. M.*, Empirical protein energy maps. Nature New Biol. **234**, 277 (1971).
787. *Doty P., Holtzer A. M., Bradbury J. H., Blout E. R.*, Polypeptides. II. The configuration of polymers of γ-benzyl-L-glutamate in solution. J. Amer. Soc. **76**, 4493 (1954).
788. *Ising E.*, Beitrag zur Theorie des Ferromagnetismus, Zeitschr. Phys. **31**, 253 (1925).
789. *Zimm B. H., Bragg J. K.*, Theory of the phase transition between helix and random coil in polypeptide chains. J. Chem. Phys. **31**, 526 (1959).
790. *Kramers H. A., Wannier G. H.*, Statistics of the two-dimensional ferromagnet. Phys. Rev. **60**, 252 (1941).
791. *Ландау Л. Д., Лифшиц Е. М.* Теоретическая физика. Т. 5. Статистическая физика. — М.: Наука, 1976.
792. *Hol W. G. J., van Duijnen P. T., Berendsen H. J. C.*, The α-helix dipole and the properties of proteins. Nature **273**, 443 (1978).
793. *Creighton T. E.*, Experimental studies of protein folding and unfolding. Progr. Biophys. Mol. Biol. **33**, 231 (1978).
794. *Dickerson R. E., Geis I.*, The Structure and Action of Proteins. Harper & Row, New York, Evanston, London, 1969.
795. *Takano T.*, Structure of myoglobin refined at 2 Å resolution. J. Mol. Biol. **110**, 537 (1977).
796. *Marshall J. J.*, Manipulation of the properties of enzymes by covalent attachment of carbohydrate. Trends Biochem. Sci. **3**, 79 (1978).
797. *Collman J. P., Brauman J. I., Rose E., Suslick K. S.*, Cooperativity in O_2 binding to iron porphyrins. Proc. Nat. Acad. Sci. USA **75**, 1052 (1978).
798. *Mitsui Y., Satow Y., Sakamaki T., Iitaka Y.*, Crystal structure of a protein proteinase inhibitor, Streptomyces subtilisin inhibitor, at 2.3 Å resolution. J. Biochem. (Tokyo) **82**, 295 (1977).
799. *Ploegman J. H., Drent G., Kalk K. H., Hol W. G. J., Heinrikson R. L., Keim P., Weng L., Russell J.*, The covalent and tertiary structure of bovine liver rhodanese. Nature **273**, 124 (1978).
800. *Richardson J. S., Richardson D. C., Thomas K. A., Silverton E. W., Davies*

D. R., Similarity of three-dimensional structure between the immunoglobulin domain and the cooper, zinc superoxide dismutase subunit. J. Mol. Biol. **102**, 221 (1976).

801. *Hagler A. T., Honig B.*, On the formation of protein tertiary structure on a computer. Proc. Nat. Acad. Sci. USA **75**, 554 (1978).
802. *Remington S. L., Matthews B. W.*, A general method to assess similarity of protein Structures, with applications to T4 bacteriophage lysozyme. Proc. Nat. Acad. Sci. USA **75**, 2180 (1978).
803. *Needleman S. N., Wungch C. D.*, A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**, 443 (1970).
804. *Fermi G.*, Three-dimensional Fourier synthesis of hyman deoxyhaemoglobin at 2,5 Å resolution: Refinement of the atomic model. J. Mol. Biol. **97**, 237 (1975).
805. *Gutfreund H., Trentham D. R.*, Energy changes during the formation and interconversion of enzyme-substrate complexes. Ciba Foundation Symposium **31**, 69 (1975).

Предметный указатель

Агглютинии из зерна пшеницы 110, 119
Агрегаты глобулярных белков 116
Аденилаткиназа 71, 87, 109, 262, 266
  изоферменты 64
  корреляция по соседним остаткам 104, 107
  модели 162, 169, 170
  тест предсказаний 142, 144, 150
Аденилатциклаза 285
Аденинсвязывающий центр 260
Аденозин 243
Адениозинсвязывающий центр 260
Айзинга модель 295
«Айсберга» структура 51, 53
Активный центр
  конвергентная эволюция 233, 276
  плотность упаковки 57
  примеры 60, 62, 169, 221, 280
  расположение между доменами 60, 117
  см. также Каталитический центр
Аланин 10, 11, 20
Алкогольдегидрогеназа 109, 110, 112, 120, 259; *см.* также Дегидрогеназы NAD-зависимые
Аллостерические лиганды 191, 201
Аллостерический контроль 259, 289
Альдолаза 119
Аминоацил-тРНК-синтетаза 271
α-Амилаза 67
Аминокислотные изменения (вставки, замены) 197, 199
  вероятности 171
  влияние на динамику свертывания 206
  как эксперименты природы 207
  компенсация 57, 202
  скорость 202
  частота 24, 234
Аминокислотные остатки 9
  взаимозамещающиеся 25
  — признаки в филогенетических исследованиях 207
  конформационное пространство 32
  молекулярные массы 11
  обозначения 10
  относительное содержание 11, 24
  парциальные заряды 41
  подобие 24, 202
  частоты встречаемости 204, 234
Аминокислотные последовательности 129
  анализ 18, 160, 208
  коллинеарность с нуклеотидными последовательностями 12, 18
  повторения 231, 232
  представление 157, 158
  предшественники 211
  сопоставления 217, 227, 235
  сравнения 198, 233
  фосфопротеидов 266
Аминокислоты 9
  биохимические обозначения 10
  гомополимеры 89, 129, 295
  происхождение 14
  свойства 10
β-Аминокислоты 15, 179
D-Аминокислоты 15, 248
Аналогичные белки 190
Антигенсвязывающие центры 62, 244
Антитела *см.* Иммуноглобулины
Антифризные гликопротеиды 269
Аполипопротеиды 267
Апоферритин 119
Апоцитохром 251
Априорная вероятность 234, 241
Априорная значимость 234, 241
L-Арабинозасвязывающий белок 110, 112
Аргинин 10, 11, 20, 248
Аррениуса уравнение 277
Аспарагин 10, 11, 20, 268
Аспартат 10, 11, 21, 260, 276, 281
Аспартат-траискарбамилаза 119
Аспартокиназа I — гомосериновая дегидрогеназа 228
Аспирин 262
Ассоциации энергия 123
Атеросклероз 267
C_α-Атом 15

Cβ -Атом 13, 15, 28, 30, 31
АТР 26, 260, 283, 287,
АТР-связывающие цеитры 283, 287

Бактериородопсин 98
Бактериофаг fd 98
Бактериофаг ФХ 174 19
Бактериофаг λ 228
Бактериохлорофилла белок 109, 112
«Банк» данных о белках 160
Белка внутренняя часть 52
  водородиые связи 54
  диэлектрическая проиицаемость 38, 54
  сравиеиие с каплей масла 49, 57
  — — кристаллом 57
  сохраиение 201, 217
Белка свертываиие *см.* Свертывание белка
Белка структура 108
  вторичиая 82, 126
  документация 157
  модели 157
  первичная 82
  сверхвторичиая 97, 116, 127
  четвертичиая 82, 255, 257
HL-A-Белки 63, 219, 220
Белки глубоководных рыб 201
Белок
  ассоциация 123
  биографии 201
  действие 273
  денатурация 50, 177
  диффереицииация 198, 211
  ииактивация 77, 184
  классы 107, 151, 154
  комплементация 184
  кристаллы 111, 123, 181
  механизм 273
  мономер 61, 63
  наращивание 7
  олигомер 61
  потребиость 78
  сжимаемость 57
  разрушение 77, 183, 206
  функция 273
  эволюция 197
Белок — белок, взаимодействие 117, 125, 246
Белок — лигаид, взаимодействие 243, 259
Бенс-Джонса белки 35
Биография белков 71, 201
Биологическая фуикция 273, 282
Биологические ограничеиия 234
Боковая цепь 9, 11, 20
  в спиралях 86, 87, 88
  гидрофобность 22
  двугранные углы 28
  искажеиия 98
  модификации 78
  объем 57
  парциальные заряды 41
  перенос свободной энергии 10, 53
  размер 25
  стерические трудиости 29
  p*K* 10
Боли, ощущение 217
Больцмана распределение 137
Бора эффект *см. Хальдана* эффект
Букингема потенциал 39
Быстро свертывающиеся цепи 184, 188
Бычий панкреатический трипсииовый иигибитор (БПТИ) 109, 181
  дисульфидные мостики 68, 187
  комплекс с трипсином 122, 124, 126, 248, 281
  моделироваиие свертываиия 195
  путь свертывания 184

Вазопрессин 212
Вазотоцин 212
Валин 10, 11, 20, 191
Вандерваальсово притяжеиие 87, 90, 100
Вандерваальсовы контакты 29, 30, 43, 125, 250
Вероятностиые методы 131
Виртуальные связи С—С 173
Вирус MS2 19
Вирус ФХ174 19, 174
Вирус кустистого карликового томата 119
Вирус лесов семлики 117
Вирус табачной мозаики 85, 118
Витамин B_6 192
Витамин К 244
Внеэпителиальиое пространство 63, 190
Внутреиние остатки 57, 198, 201
Внутриклеточные белки 64
Водородные связи 38, 45, 206
  внутреиией части белка 21
  вторичных структур 86
  коллагена 90
  поверхностей раздела 125
  представление в моделях 86, 171
Воспалительный процесс 217
Вставки и делеции 197
  антигеисвязывающих цеитров 244
  карты расстояний между Cα-атомами 174

расположение на поверхности 217
стрктурные сравнения 217, 236
эволюция белка
ВТМ *см.* Вирус табачной мозаики
ВТМ, белок оболочки 98, 100, 108
блокироваиный N-конец 71
последовательность РНК 18
самообъединение 65
Вторичная передача (Ca^{2+}, циклический AMP) 285, 290
Вторичные модификации *см.* также Посттрансляционная модификация
в изображеииях свертывания 170
как основа для классификации 107
кооперативные эффекты 137, 139
предсказание 129
содержание в белках 147
Вырождеиность геиетического кода 17, 200
Вырожденный иммунный отклик 245

Гаптоглобии 230
Гарстенга принцип 209
Гексокиназа 110, 118
Гельмгольца свободная энергия 48
Гем — гем взаимодействие 255
Гемсодержащие белки 248
Гема
атом железа 244, 248
боковые цепи 224, 250
связывающий домен 223
— цеитр 248
Гемеритрии 97, 100, 108, 112, 116, 119, 232
Гемоглобин 108, 248
аномальный 202, 207
Вены 202
в оитогенезе 209
дифференциация 211, 230
история изучения 87, 184, 222, 259
механизм присоединения O_2 255
морской миноги 88
реиатурация 184
рыб 72, 257
свертывание типа глобинов 219
серповидноклеточиый 126
Сиднея 202
симметрия 118, 119, 122
солевые мостики 54
субъединиц контакты 122
эволюция 200, 219, 233
Гемоцианин 232

Ген
дуплнкация 229
мультипликация 229
перекрывание 19
перенос межгенетический 192, 208, 227
расщепление 228
регуляторный 211
слияние 198, 228
Генеалогия белков 197, 209
Генетическая взаимосвязь 213, 233
Генетическое расстояние 209, 215
Генная инженерия 228
Гиббса свободная энергия 48
Гидрофобиое ядро 22, 51, 55, 102, 141
Гидрофобность 22
Гидрофобные взаимодействия 38, 51
Гипервариабельные последовательности 199, 212, 244
Гистидин 10, 11
в активных центрах 21
— «релейно-зарядных» системах 248, 275
как аксиальный лиганд 226, 250, 253
Гистоны 199, 229
Главная цепь 26, 53, 166, 264
Главной цепи атомы
в ферментативном катализе 246, 276, 282
водородные связи 46
на поверхности и внутри белка 53
парциальиые заряды 38, 41
роль в зарождении спирали 206
— связывании лигандов 246, 250, 264, 265
стерические трудиости 29
Гликогенфосфорилаза 79, 112, 117, 119, 168, 229, 290
Гликолитический путь 117, 289
Гликопротеид, понижающий температуру замерзания 232
Гликопротеиды 268
Глицин 10, 11, 53, 92, 204
в коллагене 90
в подвижиых цепях 263, 265
коиформационное пространство 20, 32, 33
сохранение 20, 204
Глицеральдегидтриозофосфатдегидрогеназа (GAFDH) 110, 112; *см.* также Дегидрогеназы NAD-зависимые
Глобальный эиергетический минимум 181

Глобина свертывание 219
Глобулярные белки 108, 116
Глутаминовая кислота 10, **11**, 22, 79
Глутатион 60, 67, 69, 280
Глутатионредуктаза 60, 119
  домены 60, 109, 261
  группы с окислительно-восстановительной активностью 68, 69, 261, 280, 281
Глюкагон 74, 108, 119
Глюкогенсинтетаза 266
Глюкогенфосфорилаза 79, 112, 117, 119, 229, 290
Глюкозофосфатизомераза **91, 119**
Глютенин 68
Гомологичные белки 151, 197, 233
  контакты между субъединицами 121
  нумерация последовательности 157
  обнаружение 233
  определение 224
  построение моделей 215, 224
  предсказание свертывания 135
Гомополимеры аминокислот 129, 295
Гомосериновая дегидрогеназа 228
Гормоны 68, 72, 212, 288, 290
  действие 285, 288, 290
  разрушение 183
  эволюция 213

Дарвина гипотеза 26, 200
Двугранные углы аминокислотных остатков 20, 28, 161, 172
Дегидрогеназы NAD-зависимые 114, 117, 120, 231, 241, 259
Дезоксигемоглобин 122, 124, 253
Делеции и вставки *см.* Вставки и делеции
Денатурация белков 50, 177
Десмозин 78
Диаграмма контактов 174
Дивергентная эволюция 197, 224, 233
Дигидрофолатредуктаза 110, 261
Диполь-дипольные взаимодействия 41, 48
Дипольный момент α-спирали 43, 264
Дисперсионные силы 38, 44
Дисульфидизомераза 69, 183
Дисульфидная связь 66, 184
  в свернутых петлях 69, 280
  вклад в стабильность белка 190
  индикатор свертывания 186
  — структурных повторений 159, 220, 230
  каталитическая группа 280
  сохранение 155
Дисульфидный мостик *см.* Дисульфидная связь
Дисульфидсодержащие белки 69
  процесс свертывания 182, 186
2,3-Дифосфоглицерат (DPG) 256
Дифтерийный токсин 73, 217
Диффузия по горизонтали 268
Диэлектрическая проницаемость 42
ДНК 9, 19, 59
  взаимодействие с β-слоями 95, 271
  слияние гена 227
  скорость эволюции 200
Додецилсульфат Na 50
Домены 59, 94, 103, 116, 229
  примеры 60, 61, 108
  структурные 60, 106
  функциональные 59, 106
Дыхание 223, 226

Единица эволюционного пути, PAM 199
Единичный эволюционный период 200

Железосодержащий белок с высоким окислительно-восстановительным потенциалом 109

Заболевания как эксперименты природы 282
Засорение 274
βαβ-Звено 100, 112, 241
βξβ-Звено 100, 101, 170
βαβαβ-Звено 100, 116, 174
β-Зигзаг 101, 102, 112
Зимма — Брэгга модель 138, 295
Зимоген 74

Иерархия структурных уровней 82, 126, 192
Изоаллоксазиновый цикл 69
Изолейцин 10, 11, 20
цис-транс-Изомеры пептидной связи 35, 184
Изофункциональные гомологичные белки *см.* Специализированные белки
Имидазольный цикл 11, 21
Иммуноглобулины 62, 77, 109, 135, 219, 220
  антигенсвязывающие центры 68, 126, 243

доменная структура 63, 103, 109, 116, 120
C- и V-домены 219, 244
как гликопротеиды 268
образования из ковалентно-связанных цепей 63, 64, 67
симметрия 63, 112, 116, 119
слияние C- и V-доменов 227
топологические сравнения 241
H- и L-цепи 63, 227, 244
Индольный цикл 21
Индуцированное соответствие 262
Инициации параметр 139, 295
Инозитолгексафосфат 227
Инсулин 61, 66, 108, 183
дисульфидная связь 68, 69
кристаллы in situ 118
образование 72, 73, 183
олигомерная структура 119
прогормон 74, 183
разрушение 183
Информационная РНК *см.* мРНК
Искажение химических связей 280

Кадмием отравление 64
Калликреин 217
Кальцийсвязывающие центры 79, 112, 116, 155, 231
Кальцийтранспортирующая АТРаза 267, 284, 289
Кальция ионы 79, 190, 288
Кана — Ингольда — Прелога правило 13
Каноинческие аминокислоты 10, 11, 14
«Капля масла» (метафора) 50
Карбоангидраза 109
Карбоксипептидаза А 35, 110, 279
профермент 74
Карбоксипептидаза В 76
Фф-Карта 28, 31, 96, 172, 178
α-Картины рентгенограмм 87
β-Картины рентгенограмм 87
Каталаза 249
Катализ 274
Каталитический центр
примеры 62, 164, 275, 280, 281
сохранение 201, 217; *см.* также Активный центр
Катепсин 233
Квазисимметрия
Кератин 71, 87, 97, 232
Киназа 290
Кинга — Джукса корреляция 204
Кислород молекулярный 249
FeS-Кластер 46, 114, 116
Клатратная структура 50, 53
Клубка конформация 301
область 295
«Ключ — замок», модель 38, 120, 246
Колебательная подвижность 181
Коллаген 29, 70, 85, 88
биогенез 70, 79
периодический белок 232, 237
повторное свертывание 183
поперечные сшивки 78
цепи 70
Коллагеноподобные структуры глобулярных белков 91
Коллинеарность 12, 18
Комплемент
белок 68, 72, 73, 75, 217
белок C1q 91
белок C3 76
центр связывания 62
Комплементация 184, 206
Компьютерная графика 164, 167
Конканавалин А 108, 119, 122, 269
Конвергентная эволюция 69, 232, 242, 276
Консервативные замещения 201
Конфигурация C_α-атома 13, 15
Конформационные изменения
кинетика 287
при свертывании 186
сократительных белков 284
третичной структуры 202, 253, 262, 265
четвертичной структуры 124, 255
эпицентр 262, 265
ε-Конформация 31, 136, 137
N-Концевая аминогруппа 71
Кооперативность
в олигомерных белках 117, 121, 257
— процессах свертывания (развертывания) 137, 140, 188, 295
определение 257, 289
сократительного аппарата 289
Корреляция между близко расположенными остатками 102, 103, 185, 194
Кофакторы 170, 201, 243, 260, 279
Коферменты *см.* Кофакторы
Коэруплазмин 79
КПК-модели 166
Креатинкиназа 282
Крови
плазма 63, 243
свертывание 72, 217
Кулона закон 41, 42
Кэндрью — Уотсона проволочные модели 161

Лактальбумин 215, 216, 283
Лактатдегидрогеназа (LDH) 102, 109, 171, 174, 259; *см.* также Дегидрогеназы NAD-зависимые
Лед 46, 47
Лекарства 212, 222, 283
Лектиновый рецептор 269
Леннард-Джонса потенциал 39, 40
Лиганды 243
  микросреда 252
  определение 243
  стабильность белка 190
  формирование ферментов 191
Лизин 10, 11, 20, 22, 78, 248
Лизогенный бактериофаг 228
Т4-Лизоцим 102, 108, 109, 146, 239
Лизоцим позвоночных 63, 109, 216
  гомология с лактальбумином 215
  дисульфидные мостики 68
  каталитический механизм 279
  конформационная подвижность 181
  сравнение с Т4-лизоцимом 239
Линейные группы 83, 89, 111, 118
Липоамидная дегидрогеназа 69
Липидтранспортирующие белки 267
Липиды 51, 266
Липопротеиды 232, 266
Лондона дисперсионные силы 39, 43

Магния комплексы 287, 291
Мак-Лахлана схема 235
Матричное представление 286
Маховые мышцы крылатых насекомых 291
Медленно свертывающиеся цепи 184, 188
Межгенетический переход гена 208, 227
Мембрана 51, 66, 266
Мембранные белки 77, 245, 266
Миозина легкая цепь 97, 100
Метаболическое раздражение 289
Металл как лиганд 62, 170, 190
Метионин 10, 11, 21
  водородные связи 23, 46
  как аксиальный лиганд 250, 252
  фактор роста за 2 РАМ 203
Механизм
  индуцированного соответствия 262
  определение 273
  переноса восстанавливающих агентов 260, 280
  присоединения O_2 258
  ферментативного катализа 274
Миеломы белок 243, 268
α_2-Микроглобулин 63, 64, 65, 220
Микротубулы 68, 85
Микрофишей атлас 169
Минимальные изменения оснований (МИО) 212, 215
Миоглобин
  апобелок 190, 193
  быстрота свертывания 182
  дифференциация 211, 230, 257
  количественное сравнение последовательностей 233
  свертывание глобина 220
  структура 108, 219, 249
  структурные флуктуации 180
Миозин 97, 100, 283
Митохондриальный цитохром *c* 208
Митохондрий, происхождение 225
Михаэлиса комплекс 248, 275
Мицелла 51
Модели молекулярные 161
  построение 161, 215, 224, 244
Модельные системы 253, 278
Молекулярная физиология 283
Молчащие мутации 200
Мономерные белки, определение 61, 63
мРНК 17, 26, 59, 199, 200, 228
Морского ежа яйцо 68, 199
Мостиковая область 34
Мостики S—S *см.* Дисульфидная связь
Мостики в нитях мышц 284, 285
Мультиплетные гены 229
Мутации 197
  матрица вероятностей 203
  молчащие 200
  случайные 202
  частота 199
Муцин 269
Мышцы сокращение 283

Напряжения, влияющие на химические связи 280
Нативная конформация цепи 49
Невалентные взаимодействия 38, 207
Недарвиновская гипотеза 199
Недопредсказание 147
«Нейтралистов» гипотеза 199
Непродуктивное связывание 248
Нервный импульс 285, 288
Николсона модель 165, 166
Нуклеиновые кислоты 9, 12, 95, 199, 270
Нуклеопротеиды 270

Нуклеотидсвязывающий центр 259
Нуклеотидная последовательность, анализ 18
Нуклеофильная атака 276
NAD 117, 259, 276
NAD-связывающие домены 114, 115, 249
NADP 60, 260, 261, 280
NADP-связывающие домены 60, 261

α-Область 131, 295
β-Область 131
Овальбумин 59
Ограниченный протеолиз 60, 72, 217
Окислительно-восстановительные потенциалы цитохромов 254
Окислительное фосфорилирование 289
Оксигемоглобин 124, 253, 256
Оксилизин 71, 79
Оксипролин 71, 79, 90
Окситоцин 212
Олигопептиды 231
Олигомерные белки 61
  контактные поверхности 121
  симметрия 118
  тепловые флуктуации 180
  формирование 191
  эволюция 121
Оплодотворение 217
Опознание 38, 125, 284
  специфичными аминокислотными последовательностями 75
  специфичными сахарными остатками 270
Оптический компаратор 165
Орбитальная направленность 278, 282
Отбор 199
Ответ как биологическая функция 284, 285, 291
ORTEP 89, 169

Папаин 108, 109, 233, 276
Парамиозин 87, 97
Парвальбумин 46, 54, 108, 231
  блокированный N-конец 71
  моделирование свертывания 192, 194
Парциальные заряды 41
Пепсин 69
Пептидная связь 27, 246, 275
Пептидные гормоны 68, 72, 73, 212
Первичная структура, определение 82
Перекрывающиеся гены 19
Пересадка тканей 220
Переходного состояния теория 277
Периодические белки 232
Пероксидаза 248, 250
Пероксид-дисмутаза 62, 108, 219
  агрегация субъединиц 119
  конвергентная эволюция 232
  топология слоя 241
Пиридоксальфосфат 68, 191
Пиримидонкарбонильная группа 71
Пиримидиновый цикл 11, 32
Пируватдекарбоксилаза 279
Пируваткиназа 61, 110, 112, 119, 191
Пищеварительные ферменты 50, 58, 62, 217
Плавательный пузырь 72
Плазмин 217
Плазмы
  белки 63, 269
  клетки 245
Плотность упаковки
  критерий предсказания 57
  на поверхности 125
  обеспечение специфичности 243
  сохранение в процессе эволюции 57
Поверхности остатки
  в реверсивных поворотах 91, 140
  представление 164
  процесс свертывания 52
  растворимость белка 22, 25, 52
  суперспирализованной α-спирали 98
  эволюция 216
Повторное свертывание, эксперименты 183
Поли-L-*цис*-аланин 31, 36
Полиглицин 31
Полинга — Кори параметры 27, 33
Полиовируса белок 73, 117
Полипептидная цепь, определение 26
Полипептиды синтетические 243
Полипролин 31, 89
TIM-Полость 61, 115
Полиэтиленгликоль 252
Полусинтетические белки 206
Порфирин 243, 250, 256
Посттрансляционная модификация 71
Потенциальной энергии распределение 33
Потенциалы
  вандерваальсовы 43
  вторичных структур 131
Почки 63, 65, 220, 269
Преальбумин 108, 119
Предсказания методы 57, 129, 192
  критерии оценки 146
  опробование критериев 152
  применение 143, 154

Прекоидиционированиие 192
Принцип спонтанности 177
Пробелки 77, 75
Проколлагеи 71
Пролин 10, 11, 89, 185, 206
Пропаратероидальиый гормон 77
Пропердиновый путь 76
Простетическая группа 80, 191, 201, 252, 261
Протеаза (протеиназа), определение 26
D-Протеаза 217
M-Протеаза 217
Протеаза B 109, 112, 217, 225, 239
Протромбии 79
Протогем IX 248, 252
Псевдосимметрия 118

Равновесия константа 50, 178
Различий таблица 209, 212
Радиусы вандерваальсовы 29, 30, 43, 44, 55, 56, 89
Разрешение номинальное и эффективное 160
Рамаландрана карта *см.* Φψ-Карта
β-Распад 15
C_α -Расстояний карта 174
C_α -Расстояния 238
Растворителя эффекты 38, 49, 51, 303
  агрегация 124
  катализ 278
  свертывание 179
Реакции скорость 277, 278
Реверсивные повороты 91, 92, 204
  как точки изгиба 102, 153
  обилие в белках 160
  предсказание 129
  частота встречаемости 131, 132, 144
Релейно-зарядная система 248, 275, 276
Ренатурация 177, 182
Рибонуклеаза поджелудочной железы 35, 39, 89, 136, 181
  комплементация 184
  предсказание структуры 151
  свертывание 177, 182
  эволюция 201, 202, 210
Рибонуклеиновая кислота (РНК) 9, 19, 59, 95
Рибосома 14, 118, 270
Рибосомальная РНК 12, 229, 270
Рибосомные белки 14, 72, 270
Ричардса оптический компаратор 165
РНК *см.* Рибонуклеиновая кислота
Россмана шпильки 101, 111, 115, 231, 260
Рубредоксин 46, 108

Салицилат 262
Самообъединение 65, 71, 286
Самопроизвольное свертывание 13, 14, 36, 38, 177
Саркомер 284
Саркоплазматическнй ретикулум 267, 285, 288
Сближенность 278, 282
Сверхвторичиые структуры 97, 116, 126
Сверхпредсказание 147
Свертывания белка
  единицы 106, 114
  моделирование 57, 192, 207, 238
  процесс 22, 106, 123, 154, 177, 206
  центры 154, 185, 206
Свободиой эиергии изменение 48
при ассоциации 123
— переносе 22, 38, 53, 123
— присоединенни лигандов 246, 274
— свертывании 48, 194
Ig-Свертывание 62, 220
O_2-Связывание 256
Связывание, центры
  быстрый распад 77
  кофакторов 60, 171, 259, 278
  между субъединицами 256, 264, 271, 280
  представление с помощью моделей 163, 164, 171
  примеры 62, 114, 116, 243
  сохранение в ходе эволюции 201
Связывающая энергия (свободная энергия связываиия) 262, 274
Связывающая энергия (связывающая энтальпия) 38, 48, 178
Серин 10, 11, 20, 79
  в сериновых протеазах 248, 274, 275
Сериновые протеазы 116, 216, 246
  каталитический механизм 275
  β-структуры 102, 109, 114, 115, 231
  функциональное разиообразие 198, 216, 283
Сжимаемость белков 57
Сигмоидальное присоединение 257, 289
Сигнальный пептид, гипотеза 73
Симметрия 111
  линейные группы 83, 111, 118
  пространственные группы 118, 284

точечные группы 111, 113, 118
Синтетаза жирных кислот 155, 228
Синтетические гомополимеры 129, 295
Синтетические полипептиды 243, 246
Скелетная мышца 283
Слейтера потенциал 39
Слияние на уровне гена 227
β-Слой 93
в нуклеопротеидах 271
листы 95, 142, 264
область на Φψ-карте 31
определение 102
параметры 85
полости 112, 114, 116, **218**, 230
предсказание 129
представление 170
примеры 114, 259, 260
склонности 132
сравнение 240, 261
топология 240
у поверхности раздела белок — белок 121
ширина 95
Случайная конформация цепи 49, 177, 184
Солевой мостик 38, 54, 125
в активных центрах 248
— гемоглобине 202, 256
Сократительный аппарат 282, 286
Сократительный белок 68, 284
Сохранение в процессе эволюции внутренней части белка 201, 217
генетического кода 17
глицильных остатков 20, 204
каталитических центров 201
процесса свертывания 207
свертывания 197, 207, 217, 222
Специализированные белки 198
Специфичность 72, 125, 243, 248, 269, 274
Спираль (α, α_L, 3_{10}, π) 83
Спираль 31, 32, 85
зарождение 137, 143
инициация 206
как диполь 43, 264
методы предсказания 129, 192
область на Φψ-карте 31, 131, 295
остатки, образующие 129, 130
— нарушающие 129, 130, 135
роль в качестве зародыша свертывания 185, 192
склонность 130, 131
содержание в белках 129, 150
Спираль — клубок, переход 139, 293
Стандартная вероятность 235
Стандартная значимость 235
Статистический вес 96, 100, 136, **293**
относительный 138, 300, 302
Стереоизображения 62, 114, 164, **168**
Структурно-специфичные лекарства 283
Структурные зоны ДНК 59, 200
Структурные макеты 169
Структурные повторения 111, **159**, 230, 237
Структурные флуктуации 180
Структурный домен 59, 103
Структурный класс 107
Структурный мотив 111, 113
Субстрат 170, 201, 211, 246, 274
Субтилизин 93, 110, 233, 246
Субтилизина ингибитор 109
Субъединицы 63, 190
диспропорционирование 66
контакты 120
определение 61
симметрия агрегатов 120
Суперспирализация α-спиралей 97, 287
Суперспираль 89, 97
Сывороточный альбумин 64, 139, 231

Φψ-Таблица 135
Таксономия 212, 227
Температура перехода 139, 300
Тенденции к встраиванию (во вторичную структуру) 130, 302
Тепловые флуктуации 180
Теплостойкие белки 201
Термодинамическая гипотеза свертывания 180
Термолизин 108, 109, 233
Тетраэдрический интермедиат
Тиаминпирофосфат 279
Тимидин-3,5-дифосфат 190
Тиолдисульфидный обмен 68
Тиоредоксин 69, 110
Тирозил-тРНК-синтетаза 97, 108, 110
Тирозин 10, 11, 20, 46, 181
Тканевой совместимости антигены 63, 219, 220
Токсины 68, 282
Топологическая взаимосвязь 213
Транскрипция 9
Трансляция 9, 15
Трансплантации антиген *см.* HL-A-Белки
Треонин 10, 11, 20, 79, 268
Треонилдезаминаза 68, 191
Тригональная диаграмма 120
Триозофосфатизомераза (TIM)
агрегация субъединиц 119
связывание фосфата 264
структура 61, 110, 112, 115, 116,

Трипсин 107, 112, 114, 232
дисульфидные связи 68
комплекс с БПТИ 122, 124, 126, 248
проверка компактности свертывания 183
родственные связи 216
специфический протеолиз 74, 75
Трипсиноген 74
пробелок 75
Триптофан 10, 11, 181, 237
Триптофаназа 68
тРНК 12
тРНК-связывающий центр 271
Тромбин 79, 217
Тропомиозин 97
блокированный N-конец 71
как периодический белок 232, 238
сокращение мышц 284

Узел в цепи 105
Уотсона — Крика двойная спираль 9, 12
Усиление сигнала 73, 285, 288
Утилизации концепция 276

Фактор HPr 266
Фенилаланин 10, 11, 20
Фенилацетамидаза 212
Ферменты
аллостерический контроль 289
аналогичные механизмы 276
в ранней эволюции 12
индуцированного соответствия гипотеза 262
«ключ—замок», модель 120, 246
ковалентная модификация 79, 226
магнийзависимые 287
Михаэлиса комплекс 248, 275
модельные реакции 274
основные остатки 232
предшественники 72
специфичность 211, 243, 274
субстрата связывание 246
образование 191
активные центры *см.* Активные центры
Ферментативная адаптация 212
Ферментативная кинетика 277
Ферментативный каскад 74, 285, 289
Ферментативный катализ 274
Фибриллярные белки 97
Физиологические виды белков 63
Филогенетический анализ 207, 212, 227
Флаводоксин 110, 261, 264, 266
Фосфоглицераткиназа 110, 112, 229, 260, 268
Фосфоглицератмутаза 110, 119
Фосфолипаза 74, 110
Фосфорилсвязывающие центры 263
Фосфопротеины 266
Фосфорилаза *см.* Гликогенфосфорилаза
Фосфорилирование и дефосфорилирование 263, 285, 290.
Фосфорилхолин 244
Фотосинтез 226
Флори формула 185
FAD 60, 69, 80, 223, 279, 281
FAD-связывающий центр 60, 260, 279
FMN 223, 261

Хальдана эффект 257
«Хвост» фага T4 85
Химические модели действия белков 252, 279
Химотрипсин 109, 172, 273
контакты в димере 122
корреляция по соседним остаткам 103, 104, 106
механизм катализа 247, 275
полости между β-слоями 112, 115
сравнение с гомологами 216, 239
α-Химотрипсин *см.* Химотрипсин
Химотрипсиноген 74

Цепи
ансамбли 59
прослеживание на карте электронной плотности 160
свертывание 173, 231
корреляция с процессом свертывания 206
повторения 114, 231
сохранение 197, 207, 216, 222
сравнение 174, 238
Циклические пептиды 35
Циклический AMP 285
Цилиндрическая развертка 83, 86, 98, 142
Цистамин 66
Цистеин 10, 11, 20, 66, 226, 281
Цитохромы 222, 248
Цитохром b_2 222, 223
Цитохром b_5 109, 222, 260, 254
апобелок 261
Цитохром b_{562} 222
Цитохром c_2 171, 226, 227
Цитохром c_3 248

Цитохром c_6 225
Цитохром c_{550} 225, 226
Цитохром c_{551} 225, 226, 237
Цитохром c_{555} 225
Цитохром cc_3 225
Цитохром c' 225
Цитохром f 225
Цитохромы b 222, 250
Цитохромы c 71, 108, 135, 171, 225
апобелок 251
варианты большие 225
— малые 225
окислительно-восстановительный потенциал 254
поверхность 164, 201
последовательности сопоставление 158, 236
частота замен 158, 204
филогенез 208, 215
функциональная дифференциация 225

Четвертичная структура 82, 255, 257
Чоу — Фасмана метод 140, 143

Шелка фиброин 69, 87, 93
Шенфлиса обозначения 119

Щелочная фосфатаза 119

Эволюционное расстояние (генетическое расстояние) 168, 209, 213
Эволюционные события, датирование 197, 210
Эволюционный путь 206
Эволюция 197
аналогичных белков 224
анатомии и поведения 210
в действии 211
генетического кода 17
гомологичных белков 197, 224, 233
дивергентная 197, 224, 233
ДНК 199
конвергентная 69, 225, 232, 242, 276
метаболизма 227, 229
регуляторных генов 211
Эластаза 109, 112, 202, 217
Эластин 78
Электронная плотность 56
карта 159, 161
Электронного транспорта цепи 222
Электронных оболочек расталкивание 39, 45
Электростатические взаимодействия 38, 41, 202, 279
Элонгации фактор 79
Эмбриогенез 230
Энергетические расчеты свертывания цепи 192, 207
Энергия
превращение в мышце 285
процедуры минимизации 42, 215
Энтальпия 48, 300
Энтерокиназа (энтеропептидаза) 74
Энтропия 48, 178
барьер 106, 185, 278
силы 51
Эпинефрин 285, 290
Эпителиальные поверхности 63, 74
Эрабутоксин 108
Эритрокруорин 249
Эритроциты 64, 222

Содержание

Предисловие редактора перевода 5

Предисловие редактора серии «Springer Advanced Texts in Chemistry» 7

От авторов 8

Глава 1 Аминокислоты 9

1.1 20 каноинческих амииокислот 9
1.2 Почему имеиио эти амииокислоты? 9
1.3 Коллиинеариая взаимосвязь между нуклеиновыми кислотами и полипептидами 18
1.4 Свойства боковых цепей 20
1.5 Эмпирические аналогии между аминокислотными остатками 24
Заключение 25

Глава 2 Структуриая роль пептидиой связи 26

2.1 Синтез на рибосомах 26
2.2 Параметры пептидной связи 27
2.3 Стерические затруднения 29
2.4. Коиформациоииая энергия 32
2.5 *Цис-* и *транс*-конфигурации 35
Заключение 37

Глава 3 Взаимодействия, определяющие структуру белка 38

3.1 Дисперсионные силы и отталкивание электрониых оболочек 39
3.2 Электростатические взаимодействия 41
3.3 Потенциалы Ван-дер-Ваальса 43
3.4 Водородиые связи 45
3.5 Эитропия 48
Концепция свободиой энергии 48
Смесь масла и воды 50
Полипептидная цепь в воде 51
Солевые мостики 54
3.6 Молекулярная упаковка 54
Заключение 57

Глава 4 Ковалентная структура белков 59

4.1 Ансамбли цепей 59
Функциоиальные и структуриые домеиы 59

Влияние окружения 61
Три основных физиологических вида белковых молекул 63
Иммуноглобулины и белки HL-A 64
Олигомерные белки 65
4.2 Дисульфидные связи 66
Роль мостиков S—S во внеклеточных и внутриклеточных белках 68
Дисульфидные мостики как составные части структуры белков 69
4.3 Модификации главной цепи, контролируемые ферментами 71
N- и C-концы 71
Специфическое расщепление полипептидной цепи 72
Активация трипсиногена 74
Действие комплемента — результат защитной системы пробелков 75
Сигнальные последовательности в пробелках . . . 77
Расщепление белка 77
4.4 Контролируемые ферментами модификации боковых цепей 78
Поперечные связи на основе модифицированных остатков Lys 78
Другие разновидности реакций модификации 80
Заключение 80

Глава 5 Способы свертывания и ассоциации полипептидных цепей 82

5.1 Вторичная структура 83
Линейные группы 83
3_{10}-, α- и π-Спирали 85
Спираль коллагена 88
Реверсивные повороты пептидной цепи 91
β-Складчатый лист 93
5.2 Сверхвторичные структуры 97
Суперспирализация α-спиралей 97
ββ-Структурное образование 100
β-Зигзаги и другие особенности β-структур 102
5.3 Структурные домены 103
Корреляция между близко расположенным в последовательности остатками 103
Структурные классы 107
Симметрия 111
5.4 Глобулярные белки 116
5.5. Агрегаты глобулярных белков 117
Симметрия агрегатов 118
Контакты между субъединицами в дегидрогеназах . . . 120
Поверхности контакта 121
Свободная энергия ассоциации 123
Специфичность взаимодействия белок — белок 125
5.6 Иерархия уровней 126
Заключение 127

Глава 6 Предсказание вторичной структуры по аминокислотной последовательности 129

6.1 Вероятностные методы 131
Склонность остатка встраиваться во вторичную структуру 131
Склонность двух остатков одновременно встраиваться во вторичную структуру 133

Склоииость трех остатков встраиваться во вторичиую структуру 135
6.2 Физико-химические методы 136
Методы, осиоваиные иа статистической механике . . . 136
Методы, осиоваииые иа стереохимических даииых . . . 141
6.3 Применеиие предсказательиых методов 143
6.4 Оцеика предсказательиых методов 145
Произвольиость выбора параметров 145
Критерии оценки методов 146
6.5 Замечаиия общего характера 153
Качество предсказаиия и локальиые взаимодействия 153
Успех предсказаиия и цеитры свертывания 154
Предсказаиие как аиалитический метод 155
Заключеиие 155

Глава 7 Модели, изображение и докумеитация белковых структур 157

7.1 Даииые о строеиии белка 157
Ковалеитиое строеиие 157
Пространственное строение 159
Полиые структуры 160
7.2 Представление полиой структуры 161
Пространственные модели 161
Двумериые представлеиия 168
7.3 Абстрактиые представлеиия о свертываии цепи 172
Графики для углов осиовиой цепи 172
Карты расстояний между C_α -атомами 174
Заключеиие 175

Глава 8 Термодинамика и кинетика свертываиия полипептидной цепи 177

8.1 Термодииамические аспекты 177
Переход между двумя термодииамическими состояииями цепи 177
Балаис эиергетических вкладов в глобулярном белке 179
Тепловые флуктуации в белковых структурах 180
Нативиое состояиие: глобальный или локальиый эиергетические миниимумы? 181
8.2 Быстрота, безошибочиость и ограиичеиия свертываиия in vitro 182
8.3 Структуриые элементы в иесвериутых цепях 184
8.4 Пути свертываиия 186
8.5 Влияиие лигаидов 190
8.6 Моделироваиие процесса свертываиия 192
Заключеиие 195

Глава 9 Эволюция белков 197

9.1 Специализация белков 199
Частота допустимых мутаций 199
Критерии фиксации амииокислотиых замеи в белках 201
Вероятиость мутаций аминокислотиых остатков 204
Влияиие замеи амииокислот иа динамику свертывания 206
Замеиы амииокислот как эксперимеиты, проведеиные природой 207
9.2 Филогеиез иа осиове структур белков 207
Белки, удобные для филогенетических исследоваиий . . 208
Филогеиетическое дерево 209

Эволюция поведения и морфологии в сравнении с эволюцией белка 210
9.3 Дифференциация белков 211
Построение филогенетического дерева для вазотоцина, окситоцина и вазопрессина 212
α-Лактальбумин и лизоцим 215
Трипсиноподобные сериновые протеазы 216
Иммуноглобулины, антигены трансплантации и пероксид-дисмутаза 219
«Свертывание типа глобинов» 219
Морфологическое сходство глобинов и цитохромов *b* . . . 222
Цитохромы типа *c* 225
9.4 Слияние генов 227
Процесс слияния генов 227
Слияние генов и эволюция 228
Мультипликация генов 229
9.5 Конвергентная эволюция белков 232
9.6 Обнаружение отдаленных родственных связей 233
Сравнение аминокислотных последовательностей 233
Сравнения укладки цепи 238
Сравнение топологии слоев 240
Заключение 241

Глава 10 Взаимодействия белок — лиганд 243

10.1 Лигандсвязывающие центры иммуноглобулинов 243
Модель связывания в системе иммуноглобулин — лиганд 244
Системы иммуноглобулин — лиганд как модели прочих взаимодействий белок — лиганд 245
10.2 Центры связывания субстрата сериновых протеаз 246
Образование комплекса химотрипсина с субстратом . . . 247
Центр связывания субстрата в трипсине 248
10.3 Центры связывания гема 248
Связи между гемом и апобелком 249
Влияние группы гема на структуру белков 251
Химические превращения атома железа гема в микросреде белка 252
Лиганды гема как пусковые системы структурных изменений гемоглобина млекопитающих 255
10.4 Центры связывания нуклеотидов 259
NAD-связывающие центры дегидрогеназ 259
Нуклеотидсвязывающие центры других белков 260
Индуцированное соответствие в аденилаткиназе . . . 262
10.5 Центры связывания фосфорильных групп 263
10.6 Взаимодействия белков с другими макромолекулами . . . 266
Липопротеиды 266
Гликопротеиды 268
Нуклеопротеиды 270
Заключение 271

Глава 11 Структурные основы механизма, действия и функции белков 273

11.1 Определения 273
11.2 Ферментативный катализ 274
Каталитический механизм химотрипсина 275
Теория переходного состояния 277
Эффекты, способствующие высоким скоростям реакций 278

11.3 Биологические и медицинские аспекты действия и функции белков 282
Белки как функциональные составляющие организма 282
Развитие новых функций на основе старых белков . . . 283
11.4 Скелетная мышца — система, в которой действие белка можно связать с общей деятельностью органа 283
Структурная и функциональная организация сократительных белков 284
Отклик на нервные импульсы и гормональное действие в клетке мышцы 288
Заключение 291

Приложение Статистическая механика перехода спираль — клубок 293

A.1 Функция распределения 293
A.2 Вероятность данной конформации остатка 293
A.3 Модель Айзинга 295
A.4 Модель Зимма — Брэгга для перехода спираль — клубок . . 295
Основная формула 295
Матричное представление 296
Содержание спиральных остатков в гомополимере . . . 298
Число и длина спиральных сегментов 298
Связь между s и температурой 300
Температура перехода 300
A.5 Сравнение с экспериментальными данными 301
Кривые перехода спираль — клубок и модель Зимма — Брэгга 301
Приближенные значения σ для синтетических полимеров 302
Температура перехода и склонность к спирали 302
Вклад растворителя 303

Литература 304
Предметный указатель 339

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, издательство «Мир».

Георг Е. Шульц, Р. Хайнер Ширмер

ПРИНЦИПЫ СТРУКТУРНОЙ ОРГАНИЗАЦИИ БЕЛКОВ

Научный редактор Т. Т. Орловская
Мл. научный редактор И. С. Ермилова
Художник Э. А. Смирнов
Художественный редактор М. Н. Кузьмина
Технический редактор З. И. Резник
Корректор Н. В. Серегин

ИБ № 2957

Сдано в набор 02.07.81. Подписано к печати 26.02.82. Формат 60×90/16. Бумага типографская № 1. Гарнитура литературная. Печать высокая. 11,25 бум. л. Усл. печ. л. 22,50. Усл. кр.-отт. 22,50. Уч.-изд. л. 23,93. Изд. № 3/1169. Тираж 4100 экз. Зак. 462. Цена 3 р. 90 к.

ИЗДАТЕЛЬСТВО «МИР»
Москва, 1-й Рижский пер., 2.

Ярославский полиграфкомбинат Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 150014. Ярославль, ул. Свободы, 97.

Отсканировал Семенюченко Владимир
chem_vova@mail.univ.kiev.ua; vova2002@mail.ru

ИЗДАТЕЛЬСТВО «МИР»

выпускает в свет в 1982 году
книги по химии

СТЬЮПЕР А., БРЮГГЕР В., ДЖУРС П. **Машинный анализ связи химической структуры и биологической активности:** *Пер. с англ., 17 л., 2 р. 90 к.*

Актуальность темы — связь между химической структурой и биологической активностью, перспективность метода — распознавание образов, а также средства его реализации — ЭВМ — все это характеризует данное издание как крайне актуальное и интересное. В нем изложены принципы организации распознавания образцов. На примере психотропных и снотворных агентов, одорантов и раздражителей носовой полости убедительно продемонстрирована эффективность метода при классификации и прогнозировании лекарственных препаратов.

Для научных сотрудников, работающих в области химической кибернетики, биохимии.

Заблаговременно оформляйте заказы на интересующие Вас книги. Заказы принимают магазины, распространяющие научно-техническую литературу. Своевременно сданный заказ гарантирует приобретение нужных Вам книг.

ИЗДАТЕЛЬСТВО «МИР»

выпускает в свет в 1982 году
книги по химии

Пептиды. Основные методы образования пептидных связей: *Пер. с англ./Под ред.* Э. ГРОССА, И. МАЙЕНХОФЕРА, *35 л., 5 р. 60 к.*

Фундаментальное пособие по химии и биологии пептидов — сложных соединений, регулирующих практически все важнейшие физиологические и биохимические процессы живой клетки и организма в целом. Пептидные препараты пользуются большим спросом как для научных целей, так и для нужд практической медицины, ветеринарии, животноводства.

Для химиков и биологов — научных сотрудников, преподавателей, аспирантов, студентов химических и биологических специальностей.

Заблаговременно оформляйте заказы на интересующие Вас книги. Заказы принимают магазины, распространяющие научно-техническую литературу. Своевременно сданный заказ гарантирует приобретение нужных Вам книг.

ИЗДАТЕЛЬСТВО «МИР»

выпускает в свет в 1982 году
КНИГИ ПО ХИМИИ

ТИНСЛИ И. **Поведение химических загрязнений в окружающей среде:** *Пер. с англ., 17 л., 2 р. 50 к.*

Посвящена серьезнейшей проблеме современности — охране окружающей среды от вредных промышленных выбросов. На основе представлений и идей, развитых к настоящему времени в рамках различных областей химии, предлагается решение важного вопроса — как то или иное вещество будет распределяться в окружающей природной среде и каким образом будет изменяться при этом его способность к превращениям. Ценность книги обусловлена возможностью «ранней диагностики» вредных последствий применения того или иного решения определенной производственно-технологической задачи.

Для широкого круга специалистов разного профиля во всех отраслях промышленности.

Заблаговременно оформляйте заказы на интересующие Вас книги. Заказы принимают магазины, распространяющие научно-техническую литературу. Своевременно сданный заказ гарантирует приобретение нужных Вам книг.

ИЗДАТЕЛЬСТВО «МИР»

выпускает в свет в 1982 году
книги по химии

Биохимическая термодинамика: *Пер. с англ./Под ред.* М. ДЖОУНСА, *30 л., 4 р. 20 к.*

Коллективная монография, посвященная приложению термодинамики к исследованию биологических систем в теоретическом и экспериментальном аспектах. Содержит четкое и полное изложение термодинамических основ темновых биологических процессов (взаимодействие ионов металлов, аминокислот, нуклеотидов, белково-липидные взаимодействия, процессы в клетках бактерий и млекопитающих и т. п.).

Для специалистов в области биохимии, биофизики, физической химии. Может быть использована как учебное пособие по соответствующим дисциплинам.

Заблаговременно оформляйте заказы на интересующие Вас книги. Заказы принимают магазины, распространяющие научно-техническую литературу. Своевременно сданный заказ гарантирует приобретение нужных Вам книг.

Шуль

Г. Шульц, Р. Ширмер

ПРИНЦИПЫ СТРУКТУРНОЙ ОРГАНИЗАЦИИ БЕЛКОВ

