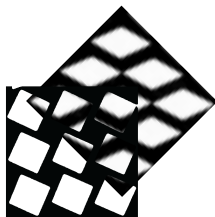


## Конформационная энергия и пространственная организация биополимеров



В этой главе рассматривается пространственная структура биополимеров, приведены расчеты конформационных энергий и принципы построения энергетических карт простых полипептидов.

### § 1. Конформационная энергия полипептидной цепи

Полипептиды составляют структурную основу белков и обладают уникальным сочетанием структурно-молекулярных особенностей.

**Природа пептидной связи.** Общий характер строения полипептидной цепи, состоящей из последовательно соединенных аминокислотных остатков, представлен

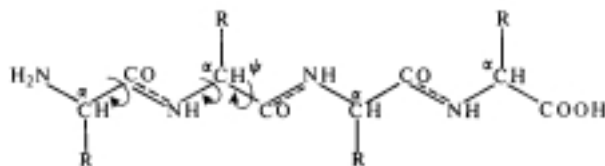
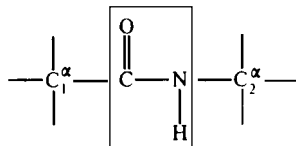


Рис. IX.1

Фрагмент полипептидной цепи

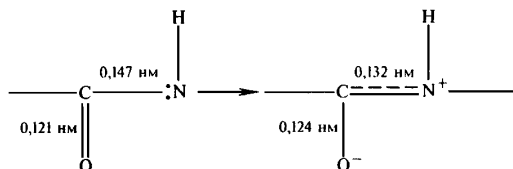
на рис. IX.1. В основе ее лежит пептидная связь  $-\text{CO}-\text{NH}-$ , образованная  $-\text{COOH}$ -группой одной и  $-\text{NH}_2$ -группой другой аминокислоты



где  $\text{C}_1^\alpha$ ,  $\text{C}_2^\alpha$  —  $\alpha$ -углеродные атомы, принадлежащие этим аминокислотам.

Важнейшей особенностью пептидной связи является ее частично двойной характер. Рентгенографические исследования простых пептидов показали, что длина пептидной связи  $\text{N}-\text{C}$  существенно меньше величины 0,147 нм, характерной для одиночной связи, и составляет 0,132 нм, что несколько превышает значение 0,125 нм,

характерное для длины обычной двойной связи  $C=N$ . Причина этого состоит в том, что неподеленная пара  $2s^2$ -электронов атома азота обобществляется между углеродами и азотом, давая примесь двойной связи между ними. При этом электрон углерода «выталкивается» из двойной  $\pi$ -связи  $C=O$  и локализуется на атоме кислорода, частично превращая связь  $C=O$  в одиночную и увеличивая ее длину от 0,121 до 0,124 нм:



Таким образом, между атомами N, C, O происходит делокализация электронов, сопровождающаяся сдвигом электронной плотности в направлениях  $N \rightarrow C \rightarrow O$ :

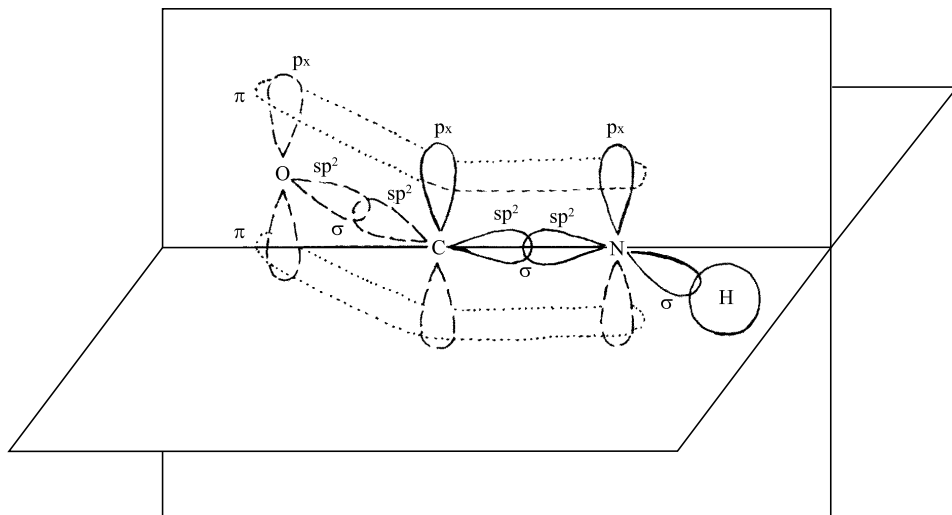
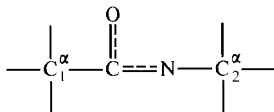


Рис. IX.2

Электронная конфигурация полипептидной цепи

Поскольку все три атома N, C, O участвуют в формировании двойной связи, их  $s$ - и  $p$ -орбитали гибридизуются с образованием  $sp$ -гибридных орбиталей. Одиночные  $\sigma$ -связи возникают между атомами за счет орбиталей  $O(sp^2)$  и  $C(sp^2)$ , а также  $C(sp^2)$  и  $N(sp^2)$ . Три  $p_x$ -орбитали атомов N, C, O, перпендикулярные плоскости  $\sigma$ -связей, перекрываются и образуют делокализованную систему  $\pi$ -электронов

(рис. IX.2). Энергия делокализации электронов, стабилизирующих структуру пептидной группы, зависит от степени перекрывания электронных орбиталей. Стабилизация оказывается максимальной, когда атомы пептидной группы, включая  $\alpha$ -углеродные атомы соседних аминокислот, расположены в одной плоскости. В соответствии с изомерной теорией пространственное строение полипептидной цепи определяется вращением вокруг единичных связей между соседними атомами (см. рис. IX.1). В силу двойного характера пептидной связи  $C=N$  вращение вокруг нее (угол  $\omega$ ) затруднено, и поэтому необходимо учитывать только вращение вокруг связей  $N-C^\alpha$  (угол  $\varphi$ ) и  $C^\alpha-C$  (угол  $\psi$ ) (рис. IX.3).

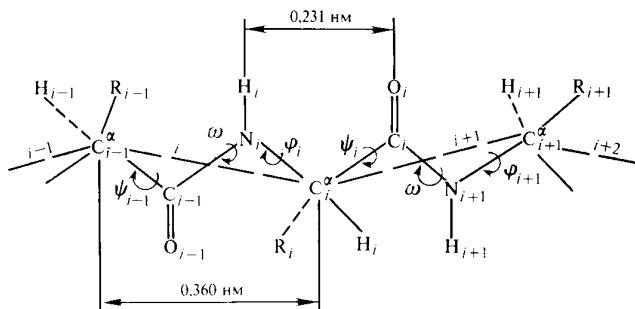


Рис. IX.3

Пространственная конфигурация полипептидной цепи

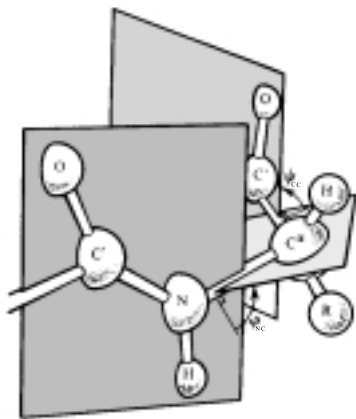


Рис. IX.4

Две пептидные группировки, связанные с помощью углеродных атомов (по A. G. Walton, J. Blackwell, 1973)

Два угла  $\varphi$  и  $\psi$  определяют вращение вокруг одиночной связи  $N-C^\alpha$  и  $C^\alpha-C'$  соответственно;  $\varphi$  и  $\psi = 0$ , когда две плоскости, определяющие угол, совпадают; значение  $\varphi$  увеличивается,  $\psi$  — уменьшается, когда плоскости вращаются, как обозначено стрелками. Рисунок представляет сегмент правовращающей L-аминокислоты

**Энергия внутреннего вращения полипептидной цепи.** Значения барьеров внутреннего вращения  $E_0$  для углов  $\varphi$  и  $\psi$  принимаются равными 0,84–2,52 кДж/моль, что намного меньше значения  $E_0 \sim 84$  кДж/моль для пептидной связи  $C-N$ . Пептидная группа  $CO-NH$  вращается вокруг этих связей как единое целое, без нарушения расположения атомов в единой плоскости (рис. IX.4). Изменение одного из углов

пары  $(\varphi_i, \psi_i)$  приводит соответственно к такому изменению стерических взаимодействий, которое зависит от значения другого угла. Иными словами, имеет место попарное кооперативное взаимодействие при вращении вокруг единичных связей, принадлежащих одному и тому же  $\alpha$ -углеродному атому. Однако взаимосвязанные вращения в пределах одной пары углов  $(\varphi_i, \psi_i)$  совершенно не зависят от значений соседней пары углов  $(\varphi_{i-1}, \psi_{i-1})$  или  $(\varphi_{i+1}, \psi_{i+1})$ .

Таким образом, полипептидная цепь является цепью с «изолированными» парами углов внутреннего вращения, где каждая пара углов может рассматриваться независимо. Причина этого состоит в том, что при изменении пар  $\varphi_i, \psi_i$  и  $\varphi_{i+1}, \psi_{i+1}$

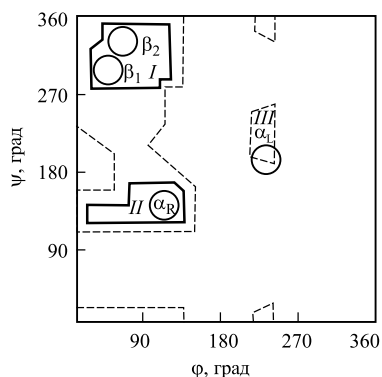


Рис. IX.5

Сферическая стерическая карта поли-L-аланина (по A. G. Walton, J. Blackwell, 1973)

Пунктиром обведены контактные положения, в которых сплошной линией показаны области равновесных положений (остальные объяснения см. в тексте)

или  $\varphi_i, \psi_i$  и  $\varphi_{i-1}, \psi_{i-1}$  отсутствуют стерические перекрытия атомов  $i$ -й пептидной единицы с  $(i+2)$ -й или  $(i-2)$ -й единицами. Данное важное свойство позволяет свести взаимодействия ближнего порядка к взаимодействиям соседних пептидных единиц. Принимая это во внимание и исходя из взаимодействий Ван-дер-Ваальса, Г. Рамачандран построил стерические диаграммы, позволяющие определять допустимые значения углов  $\varphi$  и  $\psi$ , при которых данный остаток не испытывает сильных стерических перекрытий с другими атомными группами. На рис. IX.5 представлена стерическая диаграмма для поли-L-аланина. В областях, ограниченных сплошными линиями, расстояния между парами валентно не связанных атомов соответствуют верхнему пределу значений расстояний между атомами или превышают его. Для областей, ограниченных пунктирными линиями, эти расстояния меньше и соответствуют нижним предельно допустимым значениям (см. § 1 гл. VIII). На рис. IX.5 имеются две полностью разрешенные области I и II, ограниченные сплошными линиями; область I включает параллельную и антипараллельную складчатые  $\beta$ -структуры поли-L-аланина. Области II принадлежит левая  $\alpha$ -спиральная структура. Допустимая область составляет лишь часть от всего интервала углов, который включает также левую  $\alpha$ -спираль (область III).

Границы полностью разрешенных областей могут несколько расширяться на  $5-10^\circ$  при изменении валентных углов вследствие их небольшой тепловой деформации.

**Особенности структуры белка.** Прежде чем перейти к расчетам конформационной энергии полипептидов, необходимо кратко вспомнить основные черты

пространственной структуры белковых молекул, установленные в настоящее время. Важную роль в формировании конформации полипептидов играют ван-дер-ваальсовы взаимодействия, водородные связи, электростатические эффекты и гидрофобные взаимодействия, причем решающее значение имеют первые два фактора. Природа гидрофобных взаимодействий будет рассмотрена ниже в связи с проблемой состояния воды в биоструктурах.



В классических работах Л. Полинга и Р. Б. Кори были сформулированы принципы образования структуры регулярной полипептидной цепи. 1. Длины связей и валентные углы всех пептидных групп одинаковы. 2. Полипептидная цепь должна быть максимально насыщена водородными связями. 3. Конформационные состояния всех звеньев полипептидной цепи являются эквивалентными. □

Построенная, на основе этих предположений  $\alpha$ -спиральная форма характеризуется следующими основными параметрами: высота витка  $L = 0,54$  нм, число остатков в одном витке 3,6, проекция остатка на ось 0,15 нм. Отдельные витки спирали связаны внутримолекулярными водородными связями вдоль оси спирали.

$\beta$ -конформация представляет собой складчатую структуру, включающую параллельные цепи, связанные между собой поперечными межцепочечными водородными связями. Основная роль в стабилизации  $\alpha$ - и  $\beta$ -структур принадлежит водородным связям, хотя определенные значения имеют взаимодействия боковых цепей и взаимодействия участков основных цепей с растворителем (рис. IX.6 и IX.7).



Согласно классификации белковых структур, насыщенные водородными связями регулярные участки  $\alpha$ -спирали и  $\beta$ -структуры принадлежат ко вторичной структуре белка. Первичная структура представляет собой аминокислотную последовательность белка, а третичная — соответствует набору взаимодействующих между собой регулярных конформаций. □

Некоторые природные аминокислоты стабилизируют  $\alpha$ -спираль (*ала, гли, мет* и др.), другие, наоборот, дестабилизируют (*вал, сер, цис* и др.). Разрушение спирализации происходит за счет невыгодных стерических контактов между спиральной цепью и боковыми цепями, имеющими разветвление при атоме  $C^\beta$  (*вал*). Другая причина связана с образованием между амидной группой основной цепи и  $\gamma$ -гетероатомом в боковой цепи сильной водородной связи, которая может конкурировать с внутрицепочечной водородной связью между пептидными группами.

В качестве примеров пространственных белковых структур можно привести миоглобин где 80% аминокислотных остатков составляют восемь правых  $\alpha$ -спиралей а также гемоглобин, состоящий из четырех глобул: двух  $\beta$ - и двух  $\alpha$ -субъединиц. Наряду с  $\alpha$ - и  $\beta$ -структурами доля неупорядоченных областей в молекуле белка составляет в ряде случаев до 50–60%. Это указывает на то, что регулярные участки в молекуле белка далеко не всегда доминируют, а третичная структура не всегда полностью насыщена водородными связями.

**Определение конформационной энергии.** Определение конформационной энергии больших участков белковых структур весьма сложно, оно связано с учетом всех взаимодействий и будет обсуждаться в последующих разделах. Здесь будут рассмотрены лишь отдельные участки полипептидной цепи. Важно отметить что потенциалы заторможенного внутреннего вращения  $E_0$  вокруг единичных связей, примыкающих к атому  $C^\alpha$  аминокислотного остатка, весьма малы ( $\sim 42$  кДж/моль)

и сравнимы с величиной  $k_B T$ . Следовательно, минимумы отдельных дискретных состояний, возникающих при изменении углов  $\varphi$  и  $\psi$ , разделены невысокими энергетическими барьерами, которые могут быть компенсированы другими вкладами

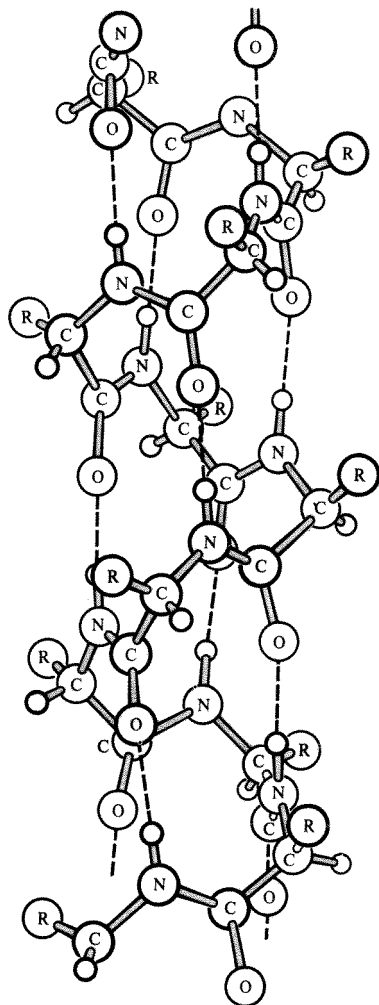
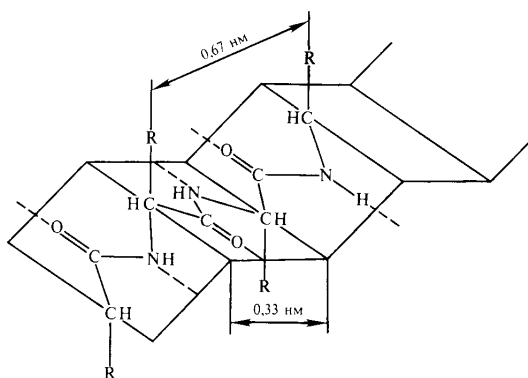


Рис. IX.6

Схема  $\alpha$ -спирали белковой молекулы (Полинг, Кори)

Рис. IX.7

Общая структура полипептидной цепи в  $\beta$ -форме (Полинг, Кори)



в конформационную энергию. Поэтому нельзя использовать поворотно-изомерное дискретное приближение и расчет конформационной энергии, определяемой вращением связей  $i$ -го остатка полипептидной цепи, необходимо вести по формуле (VIII.3.3).

С учетом ван-дер-ваальсова взаимодействия (VIII.1.16) потенциала внутреннего вращения (VIII.3.3) и электростатических эффектов (VIII.2.2) конформационная

энергия имеет вид

$$E(\varphi, \psi) = \sum_{k,l} E_{kl}(\varphi, \psi) + \{E_{\varphi}^0(1 - \cos 3\varphi)/2 + E_{\psi}^0(1 - \cos 3\psi)/2 + E_{\text{эл.-стат}}\}, \quad (\text{IX.1.1})$$

где члены  $E_{kl}(\varphi, \psi)$  определяются потенциалом Букингема (VIII.1.15) с расстояниями между атомами  $r_{kl}$ , зависящими от  $\varphi$  и  $\psi$ .

Конформационная энергия молекулы в (IX.1.1) представляется суммой отдельных видов взаимодействий в попарно-аддитивном приближении, когда каждая пара атомов взаимодействует независимо от их внутримолекулярного окружения. Оптимум потенциальной энергии и соответствующее ему конформационное состояние находят путем минимизации функции  $E(\varphi, \psi)$  при вариации значений двугранных углов вращения вокруг единичных связей. Естественно, учитывая большой объем вычислений, эта задача может быть решена только с помощью ЭВМ. Определение конформационных состояний отдельных аминокислотных остатков по формуле (IX.1.1) представляет собой необходимый и важный этап на пути построения конформационных карт, пространственной структуры белка. Конформационное состояние отдельно взятого аминокислотного остатка определяется взаимодействием атомов боковой цепи данного остатка с прилегающими пептидными группами. Вычисления для аминокислот с длинными боковыми цепями связаны с необходимостью учитывать как вращение вокруг их единичных связей (связи  $C_{\alpha}-C_{\beta}$  и  $C_{\beta}-C_{\gamma}$ ), так и взаимодействие соседних боковых цепей. Барьеры вращения в боковых цепях (углы  $\gamma$ ) зависят от природы атомных групп. Для алифатических групп  $E_0 \sim 12,6$  кДж/моль, а для фенильной группы в боковой цепи  $E_0 \sim 2,52$  кДж/моль (Г. Шерага, 1968).

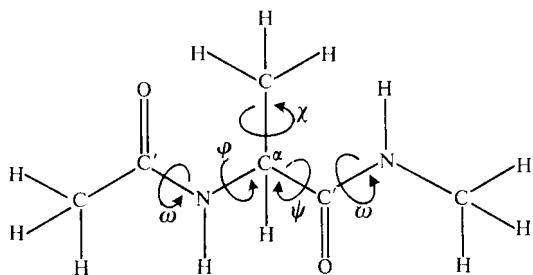


Рис. IX.8

Молекула метиламида-N-ацетил-L-аланина в плоской конформации (объяснение см. в тексте)

Удобными объектами для изучения этих взаимодействий являются метиламиды N-ацетил- $\alpha$ -аминокислот ( $\text{CH}_3\text{-CO-NH-CHR-CO-NH-CH}_3$ ), где R — радикал, соответствующий боковой цепи аминокислотного остатка. На рис. IX.8 приведена молекула метиламида N-ацетил-L-аланина ( $\text{R} = \text{CH}_3$ ) в плоской конформации с нулевыми значениями двугранных углов  $\varphi(\text{N}-\text{C}^\alpha)$ ,  $\chi(\text{C}^\alpha-\text{C}^\beta)$ ,  $\psi(\text{C}^\alpha-\text{C}^\beta)$  и углом  $\omega = 180^\circ$  ( $\text{C}' \cdots \text{N}$ ).

На рис. IX.9 представлена диаграмма поверхности конформационной энергии в функции углов  $\varphi$  и  $\psi$  (Рамачандран). Видно наличие четырех областей низкой энергии, минимумы которых лежат в областях, окаймленных контурами 4–8 кДж/моль. Аналогичная картина была получена и для других молекул с одним аминокислотным

остатком при оптимальных состояниях боковых цепей. Рассчитанные теоретические конформационные состояния отдельных метиламидов N-ацетил- $\alpha$ -аминокислот подтверждены экспериментально. Это доказывает справедливость изложенных выше (IX.1.1) методов учета взаимодействий атомов в пределах одного аминокислотного остатка. Таким образом, конформация полностью определяется при постоянных длинах связей и валентных углах значениями двугранных углов ( $\varphi$ ,  $\psi$ ,  $\chi$ ,  $\omega$ ).

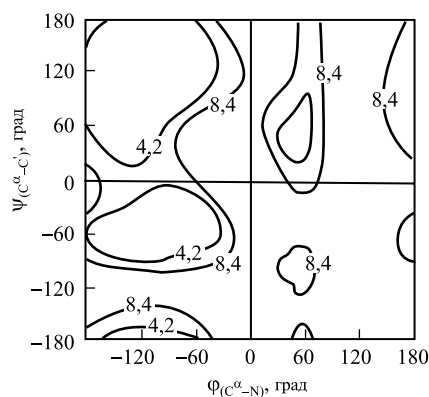


Рис. IX.9

Диаграмма потенциальной поверхности молекулы метиламида-N-ацетил-L-аланина (цифрами указана энергия конформации, кДж/моль) (по Е. М. Попову, 1981)

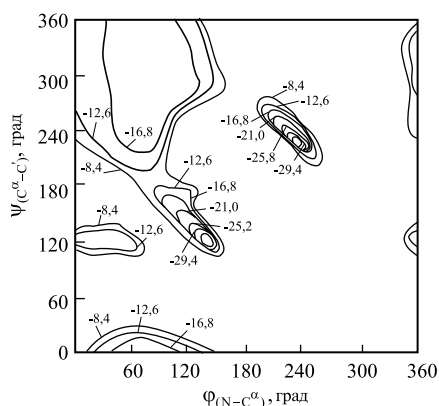


Рис. IX.10

Карта конформационной энергии поли-L-аланина (энергии конформации указаны так же, как на рис. IX.9) (по А. G. Walton, J. Blackwell, 1973)

Подобные расчеты были проведены также для более сложных систем и, в частности, полипептидов. На рис. IX.10 приведена диаграмма конформационной энергии для поли-L-аланина с учетом внутримолекулярных водородных связей. Области минимумов ( $\varphi = 240^\circ$ ,  $\psi = 240^\circ$  и  $\varphi = 120^\circ$ ,  $\psi = 120^\circ$ ) принадлежат левым и правым  $\alpha$ -спиралям, причем энергетический минимум последней на  $\sim 0,84$  кДж/пептид более устойчив. В верхнем левом углу диаграммы расположена область складчатых параллельных ( $\varphi = 61^\circ$ ,  $\psi = 239^\circ$ ) и антипараллельных ( $\varphi = 380^\circ$ ,  $\psi = 325^\circ$ )  $\beta$ -структур, стабилизированных за счет межмолекулярных водородных связей. Карта конформационной энергии поли-L-аланина (рис. IX.10) в общих чертах сходна с таковыми для других полиаминокислот, где нет замещения по  $C^\beta$ -атому (поли-L-лейцин).

В целом можно заключить, что существующие в настоящее время методы расчета позволяют с достаточной степенью точности учесть наиболее существенные межатомные невалентные взаимодействия и тем самым определить энергетически оптимальные конформации сравнительно небольших по размеру участков полипептидной цепи. Однако, как будет показано ниже, в отношении пространственной структуры больших белковых молекул положение является гораздо более сложным.



## § 2. Предсказания пространственной организации белков

Основной особенностью пространственной структуры белка является способность полипептидной цепи образовывать детерминированную структуру, обладающую необходимыми динамическими свойствами для осуществления биологических функций. □

Трехмерная структура белка в обычных физиологических температурных условиях строго центрирована, обладает большой средней плотностью  $1,33 - 1,42 \text{ г/см}^3$  что несколько больше плотности твердого сухого белка и фактически приближается к плотности кристаллов низкомолекулярных органических соединений. Значения удельного объема белковых молекул в растворе приближаются к таковым для сухого белка и немного меньше, чем в случае синтетических полимеров.

В области строения белка молекулярная биофизика сталкивается с важной проблемой: в какой степени уникальность белковой структуры (вторичной и третичной) определяется первичной аминокислотной последовательностью. Эта проблема включает следующие вопросы: 1) определяет ли однозначно данная первичная последовательность структуру белковой глобулы при сворачивании цепи, т. е. насколько детерминирован сам процесс укладки белковой цепи в нативную форму; 2) возможна ли укладка случайной последовательности остатков в структуру, подобную нативной структуре белка; 3) каким образом возникли и были закреплены в ходе эволюции существующие природные первичные последовательности аминокислотных остатков.

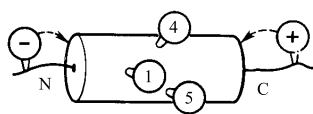
Важность этой проблемы можно иллюстрировать простым расчетом. Известно что число различных белков в природе составляет  $\sim 10^5$ . Каждый белок содержит около 200 аминокислотных остатков. В этом случае полное число различных последовательностей с 20 различными аминокислотными остатками составит  $20^{200}$ , что неизмеримо больше, чем число белков отобранных природой ( $20^{200} \gg 10^5$ ). Уже в первых работах (Р. Ламри, Г. Эйринг, 1954) было указано на связь между характером аминокислотной последовательности и нативной конформацией белка, обладающей минимальной свободной энергией. Об этом же свидетельствовали и опыты по обратимости перехода белков (рибонуклеаза) из денатурированного состояния в нативную конформацию с восстановлением всех дисульфидных связей.

В современных исследованиях делают попытки предсказать на основе конкретной аминокислотной последовательности вторичную и третичную структуры белка. Существуют в основном два подхода к решению этой проблемы.

Эмпирический метод составляет сущность подхода, который хорошо развит в настоящее время. Он основан на выявлении корреляции между вторичной структурой участка белка и его аминокислотным составом и последовательностью. Исходя из этих корреляций сформулированы эвристические принципы организации белковых молекул, пространственного строения белка и полипептидов, а также эмпирические правила построения белка и сворачивания пептидной цепи с образованием вторичной и третичной структур. Экспериментальной базой этих правил стали результаты рентгеноструктурного анализа глобулярных белков в кристаллическом состоянии, близком к нативному. □

Как правило, кристаллизация существенно не влияет на конформацию белка, тем более что в белковых кристаллах могут осуществляться многие ферментативные реакции. В глобулярных белках присутствуют полярные неполярные группы в боковых цепях. Общая топография белковой глобулы определяется тем, что на ее поверхности в основном расположены полярные группы, связанные с водой электростатическими и водородными связями, а неполярные группы находятся внутри глобулы и образуют ее гидрофобное ядро (см. § 4 этой главы). Внутримолекулярные водородные связи между пептидными группами основной цепи являются важнейшим фактором, стабилизирующим глобулу. Гидрофобные группы образуют целые гидрофобные области на поверхности основных сегментов вторичной структуры:  $\alpha$ -спиралей и  $\beta$ -структур. Тем самым длина каждого структурного сегмента определяется размером гидрофобной поверхности или числом гидрофобных групп, включенных в его состав.

I



II

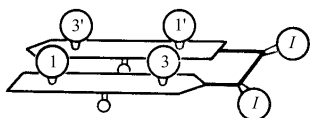


Рис. IX.11

Типичное распределение аминокислотных остатков в  $\alpha$ -спирали (I) и  $\beta$ -шпильке (II) (по О. Б. Птицыну, А. В. Финкельштейну, 1979)

Кружки с номерами означают гидрофобные группы, причем номера отражают их взаимное расположение в цепи. Кружками с «+» и «-» обозначены заряженные группы. На концах спирали стрелками указаны водородные связи, которые могут возникать между полярными боковыми группами и открытыми NH-группами N-конца спирали и CO-группами C-конца спирали

На рис. IX.11 приведена схема распределения аминокислотных остатков в  $\alpha$ - и  $\beta$ -участках, обогащенных гидрофобными группами, в то время как соединяющие их изгибы и петли преимущественно содержат короткие полярные группы, заряженные группы, остатки глицина и пролина (О. Б. Птицын, В. И. Лим, А. В. Финкельштейн). В средней части спирали во внутренних ветвях находятся преимущественно гидрофобные остатки (*ала, лей*). Остатки отрицательными и положительно заряженными боковыми цепями локализуются соответственно на N- и C-концах.



Различные типы белковых структур, образованных из структурных сегментов, составляют структурную иерархию, изображенную на рис. IX.12. По-видимому, эта иерархия белковых структур отражает также и последовательность стадий при блочном механизме сворачивания белка. Это значит, что предсказание сегмента вторичной структуры из первичной последовательности равносильно расшифровке начальной стадии сворачивания белковой глобулы из полипептидной цепи. □

Эмпирический стереохимический метод предсказания вторичной структуры основан на распределении по поверхности участка вторичной структуры аминокислотных остатков трех типов: гидрофобных боковых групп, длинных полярных боковых цепей и коротких полярных боковых групп. Постулируют, что реально осуществляется только такое распределение этих остатков, при котором на поверх-

ности  $\alpha$ -спиралей и  $\beta$ -участков образуются сплошные гидрофобные области. Это справедливо для гидрофобных остатков, занимающих в боковой цепи положения 1–5, 1–2–5, 1–4–5, а также 1–4 или 1–2 на поверхности  $\alpha$ -спирали и положение 1–3 для  $\beta$ -структур.

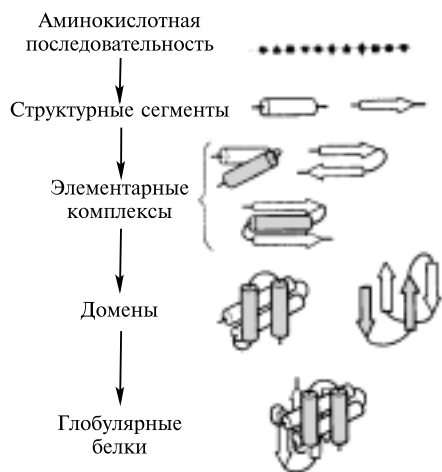


Рис. IX.12

Иерархия структуры белка и последовательность его сворачивания (по О. Б. Птицыну, А. В. Финкельштейну, 1979)

На основе простых стереохимических соображений можно видеть, что гидрофобные группы данной  $\alpha$ -спирали или  $\beta$ -участка включают в промежутках между собой гидрофобные группы других участков, создавая тем самым плотное гидрофобное ядро белка. Последнее экранируется от воды и стабилизируется также длинными гидрофобными группами, расположенными на краях гидрофобной области. Короткие полярные группы, расположенные там же, наоборот, дестабилизируют данный структурный сегмент.

Дедуктивный метод предсказания вторичной структуры белка дает правильные результаты примерно в половине случаев. 20 рассмотренных таким методом белков (3100 остатков) в  $\alpha$ -спиралях содержат 1145 остатков, из них правильно предсказаны 590, неправильно — 555, в  $\beta$ -структурах находится 425 остатков, 124 из которых определены правильно.

Статистические корреляции между первичной и вторичной структурами используют для установления связи между частотами появления тех или иных аминокислот и степенью стабилизации  $\alpha$ -спиральных участков. Обширное статистическое исследование структуры белков позволило определить частоту появления каждого остатка в  $\alpha$ -спиралях, внутренних витках  $\alpha$ -спиралей,  $\beta$ -структурах и клубках без учета взаимодействий между остатками.

В других исследованиях особое внимание обращается не только на природу данного остатка, но и на характер и частоту появления соседних с ним в цепи остатков. Тем самым учитывается стабилизирующая роль взаимодействий остатков внутри участка вторичной структуры. Действительно, в третичной структуре миоглобина, гемоглобина, лизоцима, рибонуклеазы и химотрипсина содержатся определенные пентапептидные и тетрапептидные последовательности, которые присутствуют в

спиральных областях. В ряде случаев при анализе  $\alpha$ -структур оказалось необходимым учитывать взаимодействия в нонапептидном, а для  $\beta$ -участков — в пентапептидном фрагментах. Вместе с тем на основе статистического анализа структуры девяти белков был сделан вывод, что взаимодействия боковых групп с непосредственно примыкающими участками основной цепи могут играть также большую роль, подчас более важную, чем взаимодействия между отдельными боковыми группами.

Изложенный эмпирический метод основан на экспериментально установленных корреляциях между первичной последовательностью и структурой белка и на представлениях о блочном характере строения белков. Строго говоря, здесь не учитываются в явном виде энергия и механизмы взаимодействия между соседними остатками в основной цепи. Это, в частности, относится и к конформациям боковых цепей, составляющих по числу атомов  $2/3$  всего белка, для которых недостаточно эмпирических корреляций их структурного положения. В эмпирическом методе ставится на первом этапе задача предсказания в общих чертах модели всего белка с последующим ее уточнением путем минимизации энергии при вариации степеней свободы. Очевидно, дальнейший прогресс на этом пути может быть достигнут при накоплении большого экспериментального материала и соответствующих статистических корреляций между аминокислотной последовательностью и особенностями белковой структуры. Не меньшее значение имеют разработка методов обобщенного топологического описания структур белковых молекул.

▽ Метод физического моделирования структуры белка основан на поэтапной оценке взаимодействий валентно не связанных атомов между собой и с растворителем. Предполагают, что нативная конформация белка отвечает минимуму свободной энергии и характеризуется согласованностью между всеми видами внутримолекулярных взаимодействий. Последние превалируют над межмолекулярными взаимодействиями. Непосредственное предсказание наиболее стабильной конформации белковой молекулы путем минимизации суммарной энергии всех видов взаимодействий невозможно вследствие непреодолимых математических трудностей. Они состоят в исключительно большом числе возможных конформаций белка и, следовательно, столь же большом числе локальных минимумов. Это делает поиск глобального минимума свободной энергии сразу для всей белковой молекулы практически невыполнимым. □

Так, если считать, что в среднем каждый аминокислотный остаток из  $n$  их общего числа в белковой цепи обладает 10 низкоэнергетическими формами, то общее число структурных вариантов составляет  $10^n$ . Такие расчеты можно выполнить с помощью ЭВМ только для 5–6-членного пептидного фрагмента. Очевидно, единственным правильным путем здесь является поэтапное рассмотрение отдельных видов взаимодействий, определяющих конформацию каждого остатка с ближайшими, соседними и удаленными по цепи остатками.

Учет ближних взаимодействий с ближайшими пептидными группами был выполнен на молекулах метиламидов ацетиламинокислот (см. рис. IX.8). Ближние взаимодействия задают набор низкоэнергетических конформаций свободного монопептида, а средние и дальние — определяют уже выбор тех конформаций из этого набора, которые включаются в нативную структуру белка. Вследствие согласованного

характера взаимодействий в белке средние и дальние взаимодействия не «портят» выгодной низкоэнергетической конформации монопептида. Поэтому низкоэнергетические состояния свободных остатков автоматически должны войти в состав низкоэнергетических состояний фрагментов. В свою очередь, это предполагает наличие структурных особенностей у свободных фрагментов, удовлетворяющих средним и дальним взаимодействиям в глобуле.

Иными словами, детерминированность механизма и кооперативный характер самоорганизации белковой глобулы должны приводить к энергетической дифференциации и появлению низкоэнергетических конформационных состояний на ранних стадиях самосборки. Отсюда следует, что на первом этапе расчета необходимо учитывать весь набор низкоэнергетических состояний свободных остатков, из которых затем происходит отбор для включения в нативную структуру.  $\square$

Конформационное состояние остатка характеризуется определенными значениями двугранных углов вращения, задающих соответственно низкоэнергетические области на стерических картах. На следующем этапе учитывают небольшое число низкоэнергетических состояний олигопептидов (трипептид  $n - 1, n, n + 1$ ) на основе известных конформационных состояний составляющих его фрагментов дипептида  $n - 1, n, n + 1$ . Расчеты показывают, что взаимодействия остатков в дипептиде приводят к заметной энергетической дифференциации оптимальных форм каждого из них. В то же время наиболее предпочтительные конформации трипептида представляют собой комбинации низкоэнергетических форм смежных дипептидов, что свидетельствует о согласованности три- и дипептидных взаимодействий.

Конформационный анализ более сложных олигопептидов проводят методом последовательного увеличения цепи на один остаток. Рассчитывают все комбинации, составленные из наиболее выгодных состояний вновь образованной пептидной пары  $(n + 1, n + 2)$  и предпочтительных состояний уже рассмотренного фрагмента  $(n - 1, n, n + 1)$ . Новые взаимодействия, возникающие при удлинении цепи, стабилизируют фрагмент и не нарушают уже сложившихся взаимодействий, что является прямым следствием общей согласованности всех типов взаимодействий. Учет конформационных возможностей свободных фрагментов ведет к идентификации в первичной последовательности жестких структур-нуклеаций, которые стабилизируются дальними взаимодействиями (появление таких нуклеаций предполагают и в эмпирическом подходе). Одновременно с этим под действием средних и дальних взаимодействий образуются и лабильные участки белковой последовательности с определенным набором выгодных конформационных состояний.

Таким образом, в предложенном методе расчета нахождение наиболее стабильной конформации белковой молекулы сводится к решению «блочной» задачи — поиску оптимальных условий упаковки жестких блоков при вариации низкоэнергетических состояний лабильных частей цепи, известных из анализа ближних и средних взаимодействий. Как видно, здесь используют ту же идею о детерминированном и последовательном характере самоорганизации белковой глобулы, что и в эмпирическом методе. Однако вместо статистических корреляций между первичной последовательностью и структурой глобулы в предложенном методе используют непосредственный расчет внутримолекулярных взаимодействий и принцип минимума энергии как критерий оптимальности упаковки.  $\square$

Конкретные расчеты основаны на анализе атом-атомных взаимодействий, включающих невалентные, электростатические, торсионные взаимодействия и водородные связи. В качестве примера рассмотрим данные по изучению пространственного строения и конформационных возможностей фрагмента *арг-1* — *цис-38* молекулы бычьего панкреатического трипсинового ингибитора (БПТИ), включающей 58 остатков с известной аминокислотной последовательностью (Е. М. Попов).

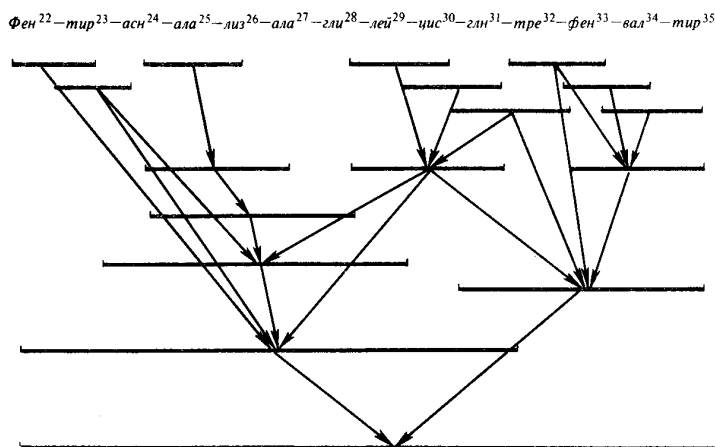


Рис. IX.13

Последовательность учета межпептидных взаимодействий при расчете фрагмента *фен-22* — *тир-35* БПТИ (по Е. М. Попову, 1981)

На первом этапе рассчитывали структуры трипептидов *ала-25* — *ала-27*, *лей-29* — *гли-31* и *фен-33* — *тир-35*, тетрапептида *ала-25* — *гли-28*. Полученные результаты позволили рассчитать конформации фрагментов *асн-24* — *лей-29*, *гли-31* — *тир-35* и *фен-22* — *гли-31* (рис. IX.13). Были выделены наиболее выгодные по энергии состояния, которые использовали при последующем рассмотрении конформационных возможностей более крупных фрагментов.

Эффективность подобного подхода демонстрирует тот факт, что для гексапептида *асн-24* — *лей-29* из 768 вариантов удалось выделить всего 17 предпочтительных конформаций с относительной энергией в интервале от 0 до 21 кДж/моль; для фрагмента *фен-21* — *гли-31* из 408 вариантов, полученных после процедуры минимизации в интервал 0–21 кДж/моль, попал лишь один структурный тип, а в интервал до 33,6 кДж/моль — 10 типов. Рассмотрение данных структурных типов совместно с фрагментом *гли-31* — *тир-35* позволило еще более сократить число возможных конформационных состояний для всего фрагмента *фен-22* — *тир-35*, состоящего из 14 остатков (рис. IX.14). Эти данные использовали в дальнейшем для выбора предпочтительных конформаций при совместном рассмотрении с семью возможными конформациями, полученными ранее при расчете структуры другого фрагмента *арг-1* — *тир-21*.

Совместное рассмотрение структуры двух указанных фрагментов *фен-22 — тир-35* и *арг-1 — тир-21* позволило выделить участок антипараллельной  $\beta$ -структуры, разные направления которой принадлежат разным фрагментам. Дальнейший анализ включал выбор возможных конформационных состояний фрагмента при присоединении трипептидного участка *тир-35 — гли-37*.

Сложная задача возникла при согласовании двух рассчитанных фрагментов, поскольку при этом обнаружены стерические препятствия. Еще более сложные си-

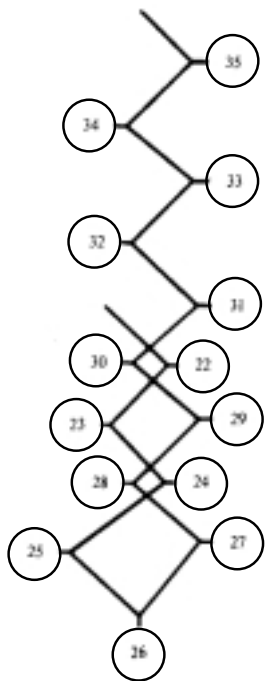


Рис. IX.14

Структурный план основной цепи в глобальной конформации фрагмента *фен-22 — тир-35* (цифрами в кружках обозначены номера фрагментов)

туации могут возникнуть при переходе к глобулярным белкам, где число граничащих блоков должно резко возрасти. Возможно, именно стерические препятствия служат одним из факторов, определяющих появление у белков гетерогенной по плотности структуры. Для устранения стерических препятствий были использованы вариации двугранных углов у различных остатков. На основе минимизации энергии при вариации двугранных углов удалось устранить все неблагоприятные стерические контакты, оставаясь в пределах низкоэнергетических конформационных карт. Полученная пространственная структура с точностью до 0,1 нм совпала с экспериментально определенной геометрией этого участка. Хорошее совпадение было получено и для 215 двугранных углов.

Эти данные показывают, что метод полуэмпирического конформационного анализа может быть успешно использован для определения пространственной структуры достаточно сложных полипептидов, включающих по крайней мере четыре десятка аминокислотных остатков, оставаясь в рамках первоначально отобранных наиболее выгодных конформационных состояний свободных пептидов. Однако при расчете достаточно больших биополимеров приходится накладывать дополнительные ограничения, делая априорные предположения о возможной структуре белка.

Ограниченность изложенных эмпирических методов предсказания связана с тем, что необходимая информация касается характера только непосредственного окружения аминокислотного остатка и не затрагивает дальних взаимодействий. Можно повысить точность предсказания вторичной структуры до 70%, если проводить сравнительный анализ близких по вторичной структуре последовательностей.

Полезно включить многие белки (гормональные рецепторы, транспортные белки), локализованные в мембранах за счет гидрофобных взаимодействий. В мембранах находятся в основном их  $\alpha$ -спирали, расположенные перпендикулярно плоскостям, хотя в ряде случаев возможно и внутримембранное расположение  $\beta$ -частей. С помощью определения участков гидрофобных последовательностей можно предска-

зять «внутримембранные» сегменты белка. Показателем служит степень гидрофобности соответствующих сегментов. Что касается петель между пронизывающими мембрану тяжами  $\alpha$ -спиралей, то они преимущественно несут положительные заряды при расположении во внутренней обращенной к цитозолю части белка.

Возможность предсказания третичной структуры белка фактически равносильна знанию характера сворачивания заданной первичной последовательности. В общем виде эта задача еще далека от своего решения (см. § 6), однако на основе экспериментальных закономерностей здесь достигнуты определенные успехи. Существует ограниченное количество известных типов укладки, каждому из которых соответствуют разные группы генетически родственных белков. Около 40% всех известных белковых последовательностей можно разделить на 1000 групп, которым соответствуют менее 7000 типов укладки. Компьютерная обработка массива этих данных позволяет сопоставить конкретные последовательности с наиболее вероятными третичными структурами из имеющегося набора.

Критерием отбора нужной нативной структуры из общего числа близких упаковок данной группы служит требование минимума энергии. Этот подход дополняется статистическими методами поиска места боковых цепей при фиксированном положении основного скелета белка.

В целом в этой области наблюдается значительный прогресс, который в ближайшем будущем позволит с помощью компьютеров проводить определение третичной структуры с высокой точностью.

**Самоорганизация белка.** Определенный характер распределения гидрофобных групп в развернутой цепи способствует образованию в ней  $\alpha$ - или  $\beta$ -структурных участков в основном за счет локальных взаимодействий внутри данного участка. Одновременно за счет дальних взаимодействий структурного сегмента с другими участками белковой цепи обеспечиваются ее стабилизация и встраивание в белковую глобулу. Это приводит к тому, что образованные в развернутой цепи наиболее вероятные структурные сегменты вторичной структуры предопределяют и структуру целой компактной глобулы, возникающую при сворачивании первичной последовательности. Вместе с тем для  $\alpha$ -спиральных белков  $\alpha$ -спиральная структура уже доминирует в развернутой цепи, а  $\beta$ -структура доминирует в развернутой цепи  $\beta$ -структурных белков.

Так, для кальций-связывающего белка мышц карпа с доминирующей  $\alpha$ -структурой в сформированной глобуле локальные максимумы вероятности образования спиральных участков в развернутой последовательности соответствуют гидрофобным участкам цепи со «спиральным» распределением гидрофобных, иницирующих и терминальных полярных групп. Уже на первых этапах самоорганизации в этом белке вторичной структуры  $\alpha$ -спирали подавляют  $\beta$ -структуру практически по всей длине цепи (рис. IX.15). Аналогичная картина наблюдается и у  $\beta$ -структурных белков (рибонуклеаза А), где  $\beta$ -структура доминирует с самого начала самоорганизации. Конечная структура белковой глобулы является результатом направленного процесса сворачивания цепи с предсуществующими структурными сегментами.

Описаны различные, пока еще далекие от завершенности, молекулярные модели образования третичной структуры белка, в которых учитываются стереохимические особенности  $\alpha$ - и  $\beta$ -структур. Например,  $\alpha$ -спирали должны упаковываться так, чтобы боковые группы одной спирали («хребты» в положениях 1–5 и 1–4) входили



в промежутки («ущелья» в положении 1–2) между боковыми группами другой спирали. Структурные сегменты соединены, как правило, гидрофобными участками цепи, где отсутствуют гидрофобные связи между сегментами. Именно в «слабых» точках комплекс может «переломиться» с образованием компактной структуры, состоящей уже из коротких нативных структурных сегментов.

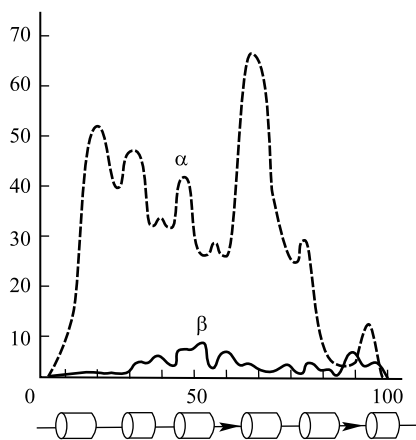


Рис. IX.15

Вторичная структура развернутой цепи кальций-связывающего белка мышц карпа на 1-й стадии самоорганизации (по О. Б. Птицыну, А. В. Финкельштейну, 1979)

Ось абсцисс — номера остатков в белковой цепи, оси ординат — вероятность  $\alpha$ -спирального (пунктир) и  $\beta$ -структурного (сплошная линия) состояния для каждого остатка цепи. Внизу показана экспериментальная локализация  $\alpha$ -спиралей (цилиндры) и  $\beta$ -участков (стрелки)

Сравнение кинетики сворачивания родственных блоков, полученных методом генной инженерии и отличающихся между собой аминокислотным составом боковых цепей, позволяет подойти к определению структуры промежуточных состояний при сворачивании. Незначительные изменения в боковых цепях белка мутантов не нарушают общей картины сворачивания, но выявляют роль определенных взаимодействий в этом процессе. Анализ более 120 мутантов, включающих изменение более половины боковых цепей барназы (РНК-аза белок со 110-ю остатками), подтвердил определяющую роль гидрофобных взаимодействий в стабилизации зародышей вторичной и третичной структуры (Фершт, 1993). По-видимому сворачивание в барназе начинается на N-конце, образующем  $\alpha$ -спирали, и центральной  $\beta$ -шпильке, которые затем притягиваются и стабилизируются гидрофобным взаимодействием. При сохранении общего направления сворачивания основной последовательности возможны небольшие различия в структуре интермедиатов, отражающие роль боковых цепей. В целом становится ясным, что в белках вторичная структура служит блоком при формировании третичной структуры и должна поэтому быстро образовываться и быть относительно стабильной, чтобы «прожить» достаточно долго. Начало формирования третичной структуры связано с образованием инициирующей шпильки из соседних по цепи сегментов. Это требует фиксации структуры двух структурных сегментов и большой затраты энергии и поэтому невыгодно с энергетической точки зрения.

Дальнейший рост третичной структуры осуществляется путем присоединения соседних участков вторичной структуры и фиксацией структуры только одного сегмента. При этом не происходит большой потери свободной энергии и поэтому в конечном итоге более выгодные энергетически стадии роста могут скомпен-

сировать невыгодные стадии инициации при образовании компактной третичной структуры.

Таким образом, можно считать, что процесс самоорганизации белка состоит по крайней мере из трех стадий (О.Б.Птицын). Формирование флуктуирующих зародышей вторичной структуры, стабилизированных главным образом водородными связями в главной цепи, происходит на первой стадии. На второй стадии из этих зародышей формируется промежуточная компактная структура белковой молекулы, где главную стабилизирующую роль играют гидрофобные взаимодействия неполярных боковых групп. Наконец, на третьей стадии формируется уникальная третичная структура белка, стабилизированная всеми основными видами объемных взаимодействий, включая стерическую комплементарность атомных групп. □

**Проблема уникальности структуры белка.** Первичная структура — последовательность аминокислотных остатков образует глобулярную структуру, в которой при большой плотности звеньев отсутствуют, тем не менее, перекрытия пар отдельных единиц. Общим условием является также определенное топографическое расположение полярных и гидрофобных боковых цепей на поверхности и внутри глобулы.

Удовлетворение этих достаточно жестких требований в сочетании с плотной нерегулярной конфигурацией и при отсутствии большого числа остатков в конформациях с высокой энергией обуславливает уникальный характер белковой глобулы. Определенная конформация данной полипептидной цепи в глобулярном белке является в этом смысле действительно исключительным явлением, а не просто одной из многих возможных конформаций, лишь слегка различающихся по энергии. Согласованность ближних и дальних взаимодействий обеспечивает направленный характер самосборки глобулы при зарождении элементов будущей вторичной структуры уже на ранних стадиях сворачивания. Для обеспечения этого необходимо, чтобы в цепи возникали регулярные ( $\alpha$ - и  $\beta$ -) участки, имеющие хотя бы одну непрерывную гидрофобную поверхность. Это условие является «грубым». Более «тонкие» свойства структуры функциональных белков связаны с определенным расположением остатков на поверхности для топографической совместимости белков вторичной структуры, с наличием гидрофобных и гидрофильных поверхностей на разных сторонах некоторых  $\alpha$ -спиральных участков.

При синтезе случайных последовательностей достаточно длинные неполярные кластеры также возникают в цепи с некоторой вероятностью, хотя в природных цепях они, конечно, определяются генетическим кодом. В обоих случаях возникшие неполярные сгущения затем слипаются своими непрерывными неполярными поверхностями. Распределение  $\alpha$ - и  $\beta$ -участков по длинам и по  $\alpha$ - и  $\beta$ -конформациям в случайных последовательностях близки к соответствующим распределениям в реальных глобулярных белках. Это было показано, в частности, на статистических сополимерах глутаминовой кислоты с лейцином. Кроме того, оказалось, что все наиболее распространенные типы укладки и топологии  $\alpha$ - и  $\beta$ -участков в реальных белках (см. рис. IX.12) могут быть получены в химически однородных полипептидных цепях с неполярными остатками.

Однако, реальные белки содержат как неполярные, так и полярные остатки и являются гетерополимерами, на которые вряд ли можно распространять выводы, полученные в экспериментах на неполярных гомополимерах.

Ниже мы увидим как отличаются между собой природные и произвольные последовательности аминокислотных остатков в отношении их способности сворачиваться в плотную белковую глобулу, а сейчас рассмотрим вопрос о том, насколько вероятен случайный синтез белковой структуры.

Получение количественной оценки возможно, если обратиться к информационному аспекту проблемы. Необходимо оценить количество ценной информации, требуемой для обеспечения заданной функции белка, и на основе этого определить вероятность его случайного синтеза.

Максимальное количество информации, содержащееся в цепи из 200 остатков, составляет (см. (VI.3.2))

$$I_B^{\max} = \log_2 W, \quad I_B^{\max} = \log_2 20^{200} = 860 \text{ бит.}$$

Вероятность случайного возникновения такой структуры исчезающе мала:

$$p = 1/W, \quad P_{\min} = 2^{-I_B} \simeq 10^{-240}.$$

Величина 860 бит, является максимальной, так как не все аминокислоты на самом деле строго фиксированы и незаменимы для выполнения белком своих функций. Можно дать предельную нижнюю оценку содержащейся в белке ценной информации, считая, что существенной является только фиксация аминокислот в активном центре. Тогда  $I_B^{\min} = \log_2 20^v = 4,3v$ , где  $v$  — число аминокислот в активном центре. При  $v = 5 \div 6$   $I_B^{\min} \simeq 20 \div 30$  бит, т. е.  $I_B^{\min} = I_B^{\max}$ . Вероятность случайного синтеза при этом максимальна:  $p_{\max} = 2^{-20} \sim 10^{-6}$ .

Эта величина уже не так мала ( $p_{\max} \gg p_{\min}$ ). Это означает, что синтез данного активного центра вполне мог произойти случайно во время эволюции. Разница между  $I_B^{\max}$  и  $I_B^{\min}$  и соответственно  $p_{\min}$  и  $p_{\max}$  очень велика. Очевидно, реальные оценки вероятности случайного возникновения функциональных белков лежат где-то в середине этого колоссального интервала. Они должны учитывать информацию, содержащуюся не только в активном центре, но и в участках и элементах структуры, характерных для данного белка.

Была предпринята попытка (Д. С. Чернавский) оценить количество информации для бактериородопсина, который содержит семь спиральных участков и неспирализованных мостиков между ними (см. гл. XXIX). При этом учитывали наличие сплошных гидрофобных боковых поверхностей, обращенных к липидной мембране, необходимое расположение гидрофильных остатков в центре молекулы и наличие фиксированных остатков в активном центре. Оказалось, что реальная оценка составляет  $I_B^{\text{реал}}$  бит, а соответственно вероятность синтеза  $p_{\text{реал}} \simeq 10^{-30}$ . Можно думать, что и другие белки содержат информацию того же порядка:  $\sim 100$  бит.

Таким образом, необходимость синтеза не только самого активного центра, но и необходимых для его функции элементов структуры с заданными особенностями приводит к уменьшению вероятности случайного возникновения функционального белка в  $p_{\max}/p_{\text{реал}} = 10^{-6}/10^{-30} \simeq 10^{24}$  раз. Все это свидетельствует о том, что, хотя и в произвольных последовательностях возможно образование отдельных элементов структуры белка, случайный синтез всего функционального белка не мог произойти сразу, «в один акт». Более вероятны представляется путь постепенного накопления информации и постепенного повышения ее ценности. Это означает, что хотя сами


по себе отдельные структурные изменения могли возникать в белке случайно, но биологическое значение определилось уже последующими процессами отбора и воспроизведения. Именно в результате этого утвердилась биологическая уникальность структуры белка, что является иллюстрацией принципа биологического запоминания случайного выбора как условия создания ценной биологической информации (см. §3 гл. IV). При этом за счет отсева альтернативных вариантов повышается ценность информации в структурах, оставшихся после «браковки».

Как мы увидим ниже, одним из главных критериев отбора последовательностей могла быть их способность полностью сворачиваться в нативную глобулу.

Уникальная способность к самоорганизации в сложные надмолекулярные системы была обнаружена недавно для химических многоатомных макроциклов (Lehn, 1997). Наблюдалось спонтанное образование в растворах супрамолекулярных систем, обладающих разными типами симметрии и архитектоники, объединение их в новые молекулярные ансамбли. Это говорит о том, что способность к структурной самоорганизации возникает на добиологическом уровне эволюции. Вслед за абиогенным синтезом элементарных единиц (аминокислот, нуклеотидов, сахаров) это могло быть основой для появления сложных структурных элементов живой материи и дальнейшего развития процессов жизнедеятельности.

### § 3. Особенности пространственной организации нуклеиновых кислот

Основная биологическая функция молекулы ДНК состоит, как известно, в хранении и передаче генетической информации, записанной в виде определенной последовательности нуклеотидов в двойной спирали (рис. IX.16). Поэтому связанное с этим основное требование к структуре ДНК — стабильность и сохранность генов — должно вполне определенным образом сочетаться с конкретными изменениями ее структуры, в частности в процессах взаимодействия с белками. Обычные тепловые флуктуации вызывают структурные изменения ДНК, не приводящие, однако, ни к разрыву водородных связей, ни к изменению межплоскостных расстояний между основаниями. В то же время спираль ДНК сохраняется при тепловых флуктуациях, а также и при более сильных внешних воздействиях. Для оценки потенциальных структурных возможностей ДНК большое значение имеют теоретические методы и модели. Они справедливы и когда необходимо оценить изменение структуры ДНК при ее участии в молекулярно-генетических процессах рекомбинации, репликации и транскрипции. Существует два ряда теоретических методов изучения структуры ДНК: методы построения упрощенных физических моделей, основанные на экспериментальных данных и отражающие совокупность свойств целостной молекулы ДНК, и методы конформационного анализа и квантовой химии.

 **Физические модели ДНК.** В первой группе методов наибольшее значение имеют модель гибкого стержня и модель спираль — клубок. □

В обычных полимерных цепях гибкость определяется поворотной изомерией (см. §3 гл. VIII). Однако в модели, где ДНК представляется в виде упругого стержня, ее жесткость характеризуется двумя параметрами: жесткостью на изгиб оси и жесткостью кручения. Жесткость на изгиб оси двойной спирали определяет способность ДНК укладываться в вирусных частицах, хромосомах и т.д. Энергия

изгиба цепочки, состоящей из  $N + 1$  звеньев или  $N + 1$  пар оснований, с углами  $\Theta_1, \Theta_2, \dots, \Theta_N$ , между звеньями имеет вид

$$E_b = (g_b/2) \sum_{i=1}^N \Theta_i^2, \quad (\text{IX.3.1})$$

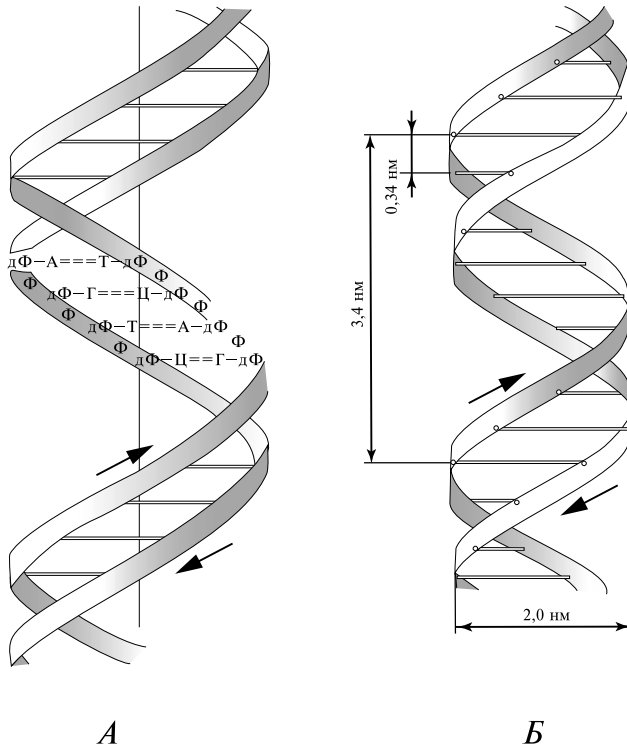
где  $g_b$  — константа жесткости на изгиб оси двойной спирали.

Как было показано (см. §1 гл. VII), жесткость полимерной цепи связана с корреляцией в направлении звеньев и характеризуется персистентной длиной  $a$  (см. (VII.1.11), (VII.1.12)). Для нахождения связи между  $a$  и  $g_b$  зададимся средней энергией колебаний равной  $k_B T$  и приравняем

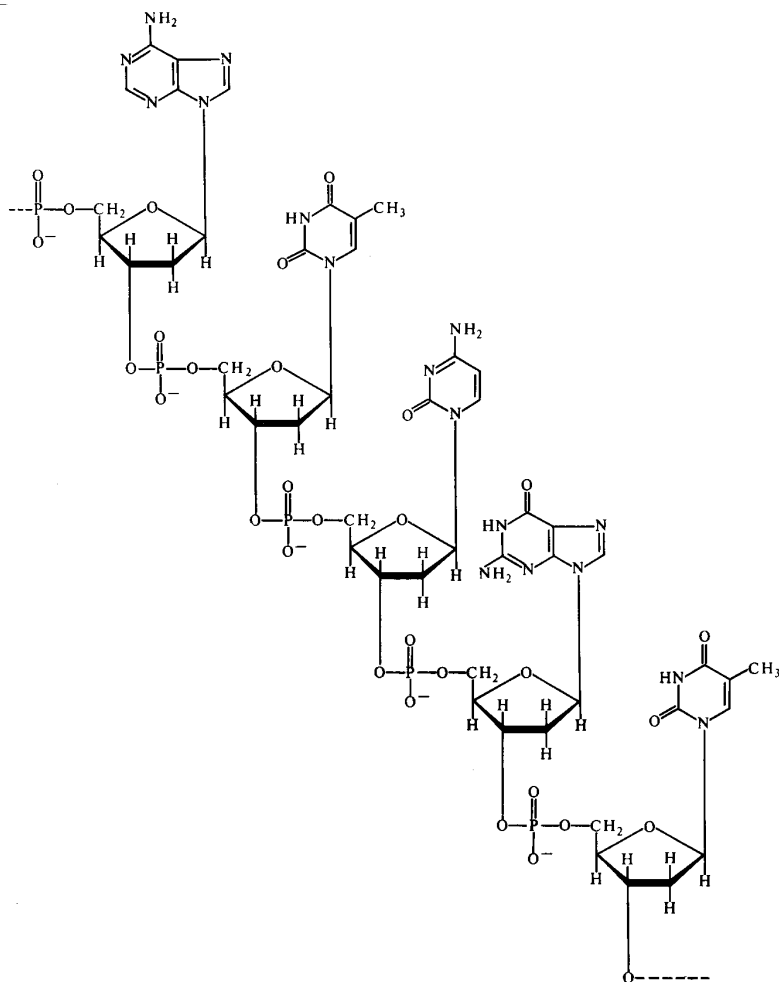
$$g_b = ak_B T/l, \quad (\text{IX.3.2})$$

где  $l$  — длина звена. Отсюда, зная персистентную длину  $a = 50,0$  нм, можно найти  $g_b \simeq 10^{-18}$  Дж.

Жесткость на кручение, или торсионная жесткость, ДНК определяется как изменение энергии при отклонении угла между соседними парами оснований от его равновесного значения. Согласно расчетам, числовое значение торсионной жесткости  $g_t$  для кольцевой полимерной цепи составляет  $g_t = 0,036 k_B T$ . При такой жесткости особое значение торсионная гибкость приобретает в случае замкнутых кольцевых молекул ДНК, состоящих из двух взаимно зацепленных одонитевых колец.



В обычных условиях угол ( $36^\circ$ ) между осями пар оснований, соседних по длине двойной спирали ДНК, испытывает отклонение  $\pm 5^\circ$  вследствие тепловых колебаний. В то же время для перехода между разными формами ДНК (*B*- и *C*-формы) достаточно изменения этого угла всего на  $2,5^\circ$ . Отсюда следует, что в пределах семейства одного типа форм (*B*-формы) происходят непрерывные переходы из одной конформации в другую. По существу, разные участки одной и той же молекулы ДНК могут находиться одновременно в разных конформациях.




*B*

Рис. IX.16

Строение ДНК. *A*, *B* — двойная спираль, *B* — участок цепи ДНК

В модели гибкого стержня изменения конформации ДНК связываются с небольшим изгибом и кручением двойной спирали в каждой паре оснований, которые накапливаются на большой длине. В результате происходит непрерывное изгибание молекулы ДНК с образованием кольца диаметром около 10,0 нм. Согласно другой точке зрения, укладка двойной спирали при компактизации ДНК может происходить за счет редких изломов на большой угол вследствие нарушения межплоскостных (стекинг) взаимодействий между парами оснований (см. §1 гл. VIII). Наряду с флуктуационными изгибами оси спирали и поворотами соседних пар оснований в ДНК может происходить и раскрытие отдельных пар оснований. Хотя этот процесс идет с очень малой вероятностью, тем не менее он приводит к сильным изменениям конформации и играет определяющую роль в реакциях ДНК с химическими агентами (например, с формальдегидом), которые могут реагировать лишь с раскрытыми нуклеотидами. Эти маловероятные, но значительные изменения структуры учитываются в моделях спираль–клубок, которые будут рассмотрены ниже. Здесь лишь необходимо отметить, что согласно модели спираль–клубок тепловые флуктуации могут приводить к раскрытию участков внутри спиральной области и переходу пары оснований (А–Т или Г–Ц) из закрытого (спиральное) в открытое (клубкообразное) состояние. При обычных физиологических температурах в ДНК происходят в основном флуктуационные раскрытия отдельных пар оснований с вероятностями не более  $10^{-5}$ . Вероятности одновременного раскрытия двух и большего числа пар оснований соответственно на один и два порядка меньше.

 **Конформационный анализ ДНК.** В отличие от хорошо изученных переходов спираль–клубок переходы ДНК в пределах двухтяжевого состояния без разрыва комплементарных пар стали широко изучаться сравнительно недавно. Целью этих работ является выявление возможных форм ДНК и физических причин конформационных переходов внутри двухтяжевого состояния. □

Важность этих задач очевидна с биологической точки зрения, поскольку в процессе функционирования ДНК может менять свою конформацию вследствие взаимодействия с другими компонентами клетки, прежде всего с белками. Двойная спираль нуклеиновой кислоты, обладая общей жесткостью по длине спирали, представляет собой вместе с тем своего рода шарнирное устройство с ограниченным числом вращательных степеней свободы вокруг единичных химических связей (рис. IX.17, IX.18). Общая геометрия спиральной молекулы описывается пятью параметрами (рис. IX.19), из которых наиболее важные: спиральное вращение  $\tau$  — угол поворота между соседними парами и  $D$  — расстояние пары от оси спирали. Отличие А- и В-форм состоит в различии значений углов  $\chi$  и  $\sigma$  (на  $60^\circ$ ) и большими положительными значениями  $D$  и меньшими значениями  $\tau$  у А-формы.

Основная причина всех этих различий связана с альтернативной геометрией сахарного кольца у форм В- и А-семейств. Фуранозное кольцо рибозы и дезоксирибозы неплоское. Обычно либо атом С2, либо атом С3 выходят из плоскости других четырех атомов кольца. Если данный атом выдвинут в сторону 5'-углерода и расположен выше кольца, т. е. находится от него со стороны основания, то кольцо обладает эндоконформацией. Если атом лежит ниже кольца, то получается экзoконформация (рис. IX.20). При изменении геометрии сахара с С2'-эндо (С3'-экзо) к С3'-эндо резко меняется расположение в пространстве выходящих из кольца свя-

зей, что и соответствует переходу от *B*- к *A*-формам. Наличие столь большого числа внутренних степеней свободы позволяет рассматривать двойную спираль как структуру, способную к непрерывному изменению своей геометрии.

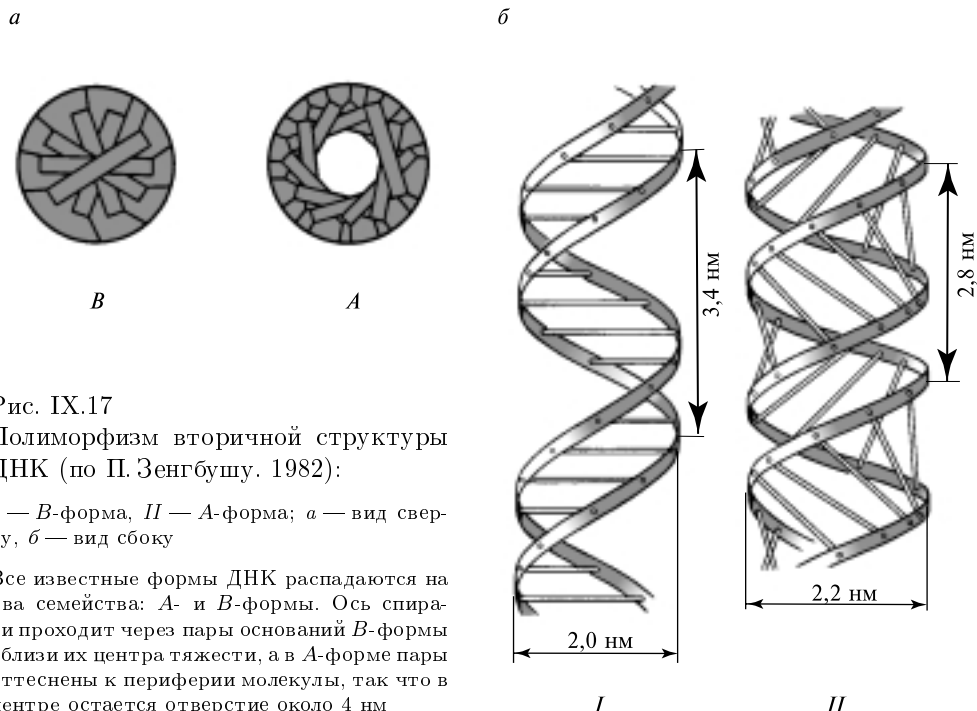


Рис. IX.17

Полиморфизм вторичной структуры ДНК (по П. Зенгбушу, 1982):

*I* — *B*-форма, *II* — *A*-форма; *а* — вид сверху, *б* — вид сбоку

Все известные формы ДНК распадаются на два семейства: *A*- и *B*-формы. Ось спирали проходит через пары оснований *B*-формы вблизи их центра тяжести, а в *A*-форме пары отеснены к периферии молекулы, так что в центре остается отверстие около 4 нм

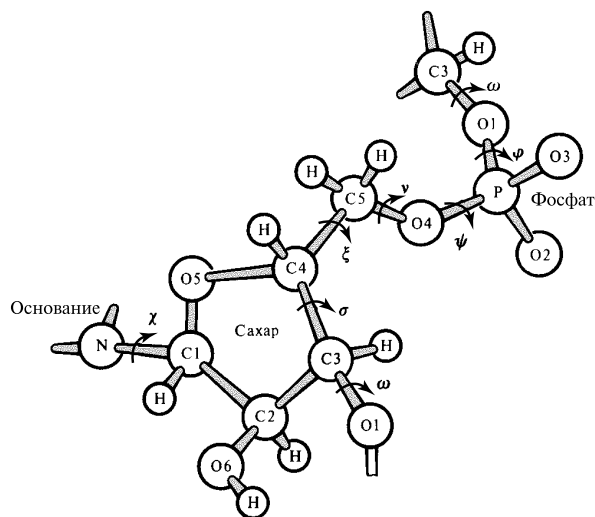


Рис. IX.18

Обозначения атомов и двугранных углов повторяющейся единицы сахарофосфатной цепи

Конформация мономера нуклеиновой кислоты задается конформацией сахарного кольца, пятью углами вращения вокруг единичных связей в сахарофосфатной цепи  $\zeta, \nu, \varphi, \omega, \theta$ . одним углом  $\chi$  (N—C1, гликозидная связь), определяющим ориентацию основания относительно сахарного кольца



Основная трудность в решении задачи полного описания множества энергетически разрешенных конформаций двойных спиралей состоит в чрезвычайно большом переборе всех структурных вариантов в пространстве двугранных углов или

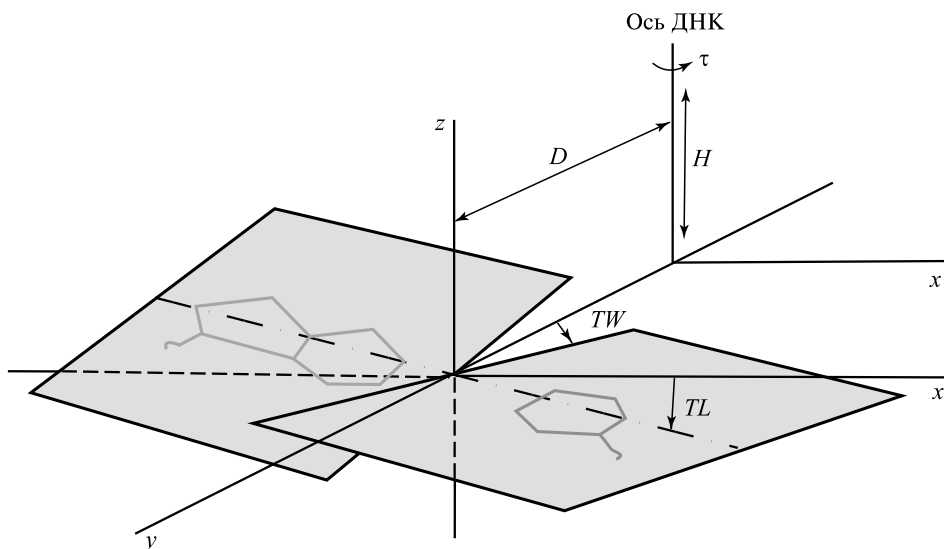


Рис. IX.19

Пять параметров Арнотта, задающих расположение оснований в регулярной двойной спирали: поворот  $\tau$ , расстояние от оси спирали  $D$ , наклон  $TL$ , «пропеллер» оснований в паре  $TW$ , проекция повторяющейся единицы на ось спирали  $H$  по В. И. Иванову, 1980)

параметров двойной спирали (см. рис. IX.19; IX.20). Область перебора сужается, если исходно задавать лишь «разумные» взаимоположения пар оснований, рассчитывать для них конформации сахарофосфатного остова. Оказывается, что для любого набора параметров оснований в случае регулярной спирали существует только конечное число конформаций сахарофосфатного остова. При заданном расположении оснований соединяющая их сахарофосфатная цепь является жесткой, однако при движении пар фиксированных оснований становятся возможными непрерывные переходы в пределах различных семейств форм. Непосредственно энергию спирали рассчитывают в виде суммы ван-дер-ваальсовых и электростатических взаимодействий с использованием потенциала Леннарда–Джонса (см. (VIII.1.16)) и значений  $\epsilon \sim 4$ .

Конформационная карта комплементарных спиралей приведена на рис. IX.21 в координатах  $(D, \tau)$ . Она представляет собой проекцию 5-мерного пространства переменных параметров на плоскость  $(D, \tau)$ , причем остальные параметры имеют такие значения, при которых энергия спирали является минимальной.

На карте изображены канонические семейства  $B$  и  $A$ , содержащие все известные модели форм двойной спирали. Энергетические «овраги», ограниченные

линиями равной энергии, соответствуют областям существования энергетически выгодных структур. Видно, что профиль энергии для  $B$ -«оврага» характеризуется довольно плоским дном в области  $36^\circ \leq \tau \leq 42^\circ$ .

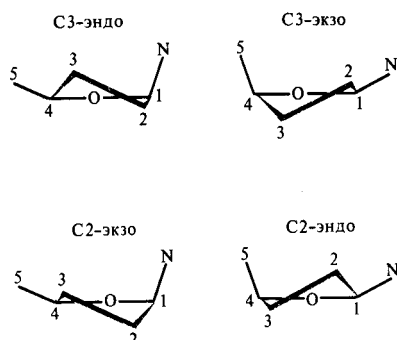


Рис. IX.20

Четыре наиболее устойчивые конформации сахарного кольца в нуклеотиде (по В. И. Иванову, 1980)

В  $B$ -формах присутствует  $C2'$ -эндо- ( $C3'$ -экзо) сахар; в  $A$ -формах —  $C3'$ -эндо. При изменении геометрии сахара от  $C2'$ -эндо ( $S$ -тип) к  $C3'$ -эндо ( $N$ -тип) связи  $C4-C5$  и  $N-C1$ , занимающие вначале почти экваториальное и меридиональное положения соответственно меняют свою ориентацию на противоположную.

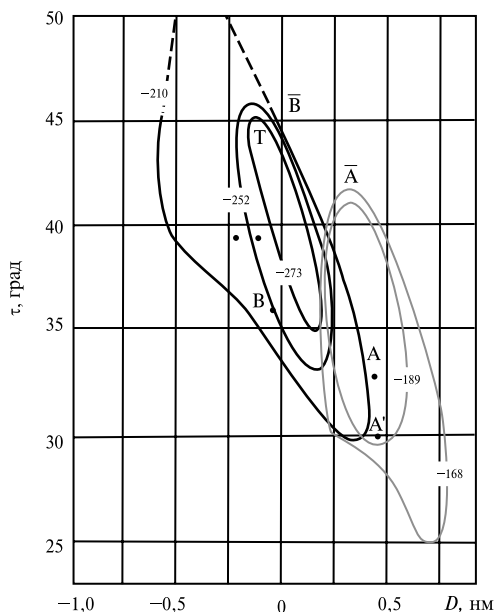


Рис. IX.21

Карта  $(D, \tau)$  для канонических семейств спиралей нуклеиновой кислоты (по В. И. Иванову, 1980)


Черными линиями обозначены семейства форм с  $C2'$ -эндо-сахарами ( $B$ -форма), цветными линиями — с  $C3'$ -эндо-сахарами ( $A$ -форма). Линии уровней маркированы величинами энергии в кДж/моль. Верхние части некоторых «оврагов» изображены точками, чтобы выделить области, в пределах которых спирали имеют очень узкий сахарный желоб (тот, в который входят гликозидные связи)

Видимо, в реальных условиях электростатическое отталкивание фосфатов противоположных цепей стабилизирует формы с большей шириной сахарного желоба, что приводит к опытным значениям  $\tau \leq 36^\circ$ , т. е. меньшей закрученности спирали. Энергия  $A$ -форм в целом выше, а «яма» уже, чем у  $B$ -форм.

Это объясняется энергетически невыгодным вкладом кулоновых взаимодействий в  $A$ -формах, где расстояние между фосфатами одной и той же цепи примерно на 0,1 нм короче, чем в  $B$ -формах.

Конформационные черты энергетически оптимальных полимеров и тримеров  $B$ -области в отличие от  $A$ -области (см. рис. IX.21) не зависят от состава и последо-

вательности оснований. Области значений углов сахарофосфатной цепи совпадают с областями наиболее вероятных значений двугранных углов в кристаллах фрагментов нуклеиновых кислот. В целом при полном учете всех внутримолекулярных взаимодействий *B*-форма оказывается единственной устойчивой формой ДНК, параметры спирали которой отвечают рентгенографическим данным. Перестройки спирали при конформационном переходе  $A \rightarrow B$  могут происходить и без заметных пространственных напряжений с энергетическим барьером перехода меньше 12,6 кДж/моль на нуклеотидную пару.

 Главное структурное отличие *A*-формы от *B*-формы — это большое отверстие в центре спирали у *A*-формы (рис. IX.17), которое занимает почти половину диаметра всей спирали ( $d \sim 20 \text{ \AA}$ , у *B*-формы  $d \sim 9 \text{ \AA}$ ). Молекулярным «переключателем» для перехода двойной спирали ДНК между семействами форм *B*- и *A*- служит сахарное кольцо пентозы. □

Биологическая роль *A*-формы является предметом дискуссий. Известно, например, что двуспиральные участки и РНК находятся в *A*-конформации. Это происходит потому, что 2-ОН-группа рибозы в РНК, в отличие от пентозы ДНК, не может вписаться в тесное пространство сахарно-фосфатного остова *B*-формы. В образовательной *A*-форме РНК этот гидроксил направлен наружу и не испытывает никаких стерических ограничений. В процессах транскрипции должен образовываться гибрид РНК–ДНК. Очевидно, для образования такого гибридно рибозо-дезоксирибозного РНК–ДНК дуплекса предпочтительно наличие *A*-формы ДНК. Промоторы, или места связывания РНК-полимеров с ДНК в процессе транскрипции, занимают малую долю природной ДНК, находящейся в *B*-форме. Оказалось, что РНК-полимеразы могут, взаимодействуя с ДНК, переводить ее наибольшую часть в *A*-конформацию (Иванов В. А., 1972; Жакобо–Молина, 1993).

**Необычные структуры ДНК.** В настоящее время помимо *A*-формы большое внимание уделяется изучению других необычных структур ДНК. К ним относятся крестообразные структуры, левозакрученная *Z*-форма, различного рода тройные и четвертные спирали, структура со сдвинутыми петлями SLS (Slipped Loop Structure).

При физиологических условиях все необычные структуры, как и *A*-форма, обладают более высокой по сравнению с *B*-формой энергией.

Мы остановимся кратко на свойствах некоторых из них.

Важное значение имеет проблема функционирования в клетке сверхзакрученных кольцевых замкнутых ДНК. При закрученном нормальном двуспиральном комплекса общее число оборотов  $\alpha$  одной нити относительно другой равно числу витков  $\beta$  в ненапряженной двойной спирали плюс число супервитков  $\tau$ , т. е.  $\alpha = \beta + \tau$ .

Величина  $\beta$  всегда положительна,  $\tau$  может быть и отрицательным — это левоспиральные супервитки, противоположные по знаку двойной правой спирали.

На рис. IX.22 изображена сверхспирализованная молекула ДНК, у которой двойная спираль правая, а сверхспирализация отрицательная (сверхспираль левая).

Различные характеристики ДНК зависят от плотности супервитков (степени суперспиральности)  $\sigma$ , равной числу этих супервитков  $\tau$  на 10 пар оснований в состоянии, когда все пары в двойной спирали закрыты. Для природных кольцевых молекул ДНК вирусов  $\sigma$  чаще всего отрицательна. Наиболее типичное значение

$-\sigma = 0,05$  ( $\sim 5$  супервитков на 1000 пар оснований). С ростом плотности сверхспирали  $-\sigma$  вероятность раскрытия пары возрастает, причем наиболее резко от  $\sim 9 \cdot 10^{-5}$  до  $10^{-3}$  в области изменения плотности от 0,06 до 0,12. Здесь начинают раскрываться не только отдельные пары, но и более протяженные области (рис. IX.23), что приводит к появлению крестообразных структур. Плотность сверхспирализации  $-\sigma$  не может превысить некоторого предельного значения ( $-\sigma = 0,085$ ), поскольку дальнейшее увеличение  $-\sigma$  будет компенсироваться соответствующим раскрытием пар оснований в двойной спирали (М. Д. Франк-Каменецкий).

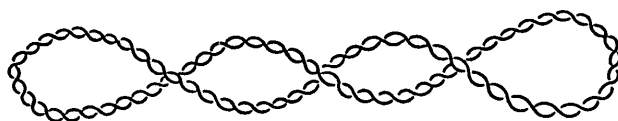


Рис. IX.22

Сверхспиральная ДНК. Сверхспирализация отрицательна (по М. Д. Франк-Каменецкому, 1988)

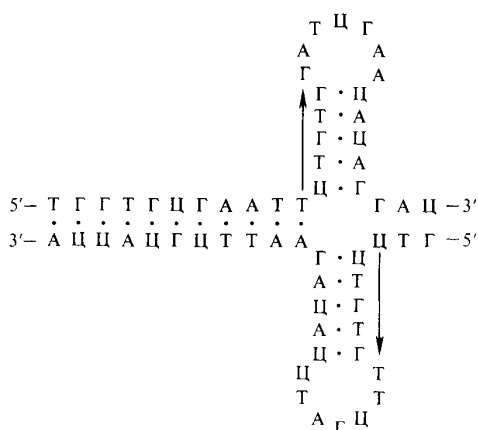


Рис. IX.23

Две конформации участка гена, детерминирующего дрожжевую фенилаланиновую тРНК (по Д. Мецлеру, 1980)

В области значений  $-\sigma = 0,06 \div 0,08$  появляются крестообразные структуры, конкурирующие с расплавленными структурами и образующиеся при раскрытии целых участков ДНК

Состояние сверхспирализации наблюдается не только у концевых замкнутых ДНК, но также на отдельных участках линейной ДНК. Сверхспирализация отдельных доменов происходит под влиянием специальных белков, активирующих фермент ДНК-перазу, который превращает «расслабленную» несверхспирализованную ДНК в сверхспираль. Имеются данные о том, что транскрипция в клетке идет на ДНК, находящейся в напряженном сверхспирализованном состоянии.

Обратим внимание, что образующие кресты последовательности ДНК, которые находятся в разных цепях, являются зеркальными превращениями, или палиндромами, т. е. одинаково читаются справа налево и слева направо. На рис. IX.23 они обозначены стрелками, которые указывают направление считывания.

До недавнего времени рентгеноструктурный анализ ДНК ограничивали ориентированными волокнами. Разрешение на таком образце не достигало атомного уровня. Поэтому интерпретация данных о структуре ДНК требовала модельных представлений. Впервые на монокристаллах синтетического олигонуклеотида  $\frac{ЦГЦГЦГ}{ГЦГЦГЦ}$  рентгеноструктурный анализ показал, что этот дуплекс образует спираль с антипараллельными нитями и уотсоновскими парами оснований, закрученную, однако, влево и содержащую повторяющуюся единицу не одного, а двух нуклеотидов (Рич, 1978). Из-за этого сахарофосфатный остов имеет зигзагообразный вид. Структура получила название *Z*-формы. Анализ распространенной структуры *Z*-ДНК привел к заключению, что для ее образования важно регулярное чередование пуриновых и пиримидиновых нуклеотидов в каждой из комплиментарных нитей и отрицательная сверхспирализация всей молекулы ДНК. □

В изучении свойств необычных структур ДНК существенно определить однонитевые участки, которые являются элементами этих конструкций. В этом отношении наиболее продуктивным оказался ферментативный метод, который позволяет локализовать место атаки одноститовой нуклеазы. Таким образом, например, были обнаружены одноститовые участки крестовидной ДНК, а также показано, что при образовании *Z*-формы атакуются границы между ними и *B*-формой. Оказалось, что одноститовые участки, чувствительные к одноститовой эндонуклеазе, находятся в важных регуляторных участках генома.

Они представляют собой последовательности, которые, как правило, в одной цепи содержат только пурины (А, Г), а в другой только пиримидины (Ц, Т), т. е. гомопурины — гомопиримидиновые последовательности типа  $(Г) \cdot (Ц)$  или  $(ГА) \cdot (ТЦ)$ . Изучение этого вопроса привело к открытию еще одной необычной формы ДНК, так называемой *H*-формы (Франк-Каменецкий).

$\nabla$  *H*-форма связана с существованием тройного комплекса, состоящего из двух полипиримидиновых (Ц-Т) цепей и только одной полипуриновой (А-Г) цепи. □

В основе комплекса лежит структурный элемент — триада, образованная тремя основаниями из трех одноститовых цепей: ТАТ или ЦГЦ<sup>+</sup>. Как видно, триады включают в себя обычные уотсон-криковские пары ТА и ЦГ, к которым присоединяется третье основание Т или Ц из полипиримидиновой (Ц-Т) цепи. В случае цитозина образование ЦГЦ<sup>+</sup> происходит путем захвата из среды одного протона. На рис. IX.24 приведена структура *H*-формы для участка цепи  $(ГА)_{16} \cdot (ТЦ)_{16}$ , встроенного в плазмидную ДНК. Основной элемент структуры — тройная спираль, состоящая из триад ЦГЦ<sup>+</sup>. Кружками обозначено уотсон-криковское спаривание оснований ДНК и крестиками — присоединение третьего цитозина с дополнительным протоном. Как видно, полипиримидиновая (Ц-Т) последовательность, входящая в тройной тяж *H*-формы, является перевертышем и читается одинаково справа налево и слева направо по одной и той же цепи.

Заметим, что все регуляторные гомопуринов-гомопиримидиновые последовательности являются зеркальными перевертышами и легко образуют *H*-формы.

$\nabla$  Другая неканоническая структура со сдвинутыми петлями (SLS-форма) была обнаружена для ДНК и РНК (Траверс А., Иванов В. И.). SLS-форма образуется, если в цепях ДНК содержатся короткие прямые повторы. Она получается

при проскальзывании относительно друг друга комплементарных нуклеотидных цепей. При этом возникают две одонитевые петли, выступающие из каждой полинуклеотидной цепи и разделенные длиной повтора. На рис. IX.25 показана схема образования SLS-структуры. Как видно, проскальзывание цепей приводит к тому,

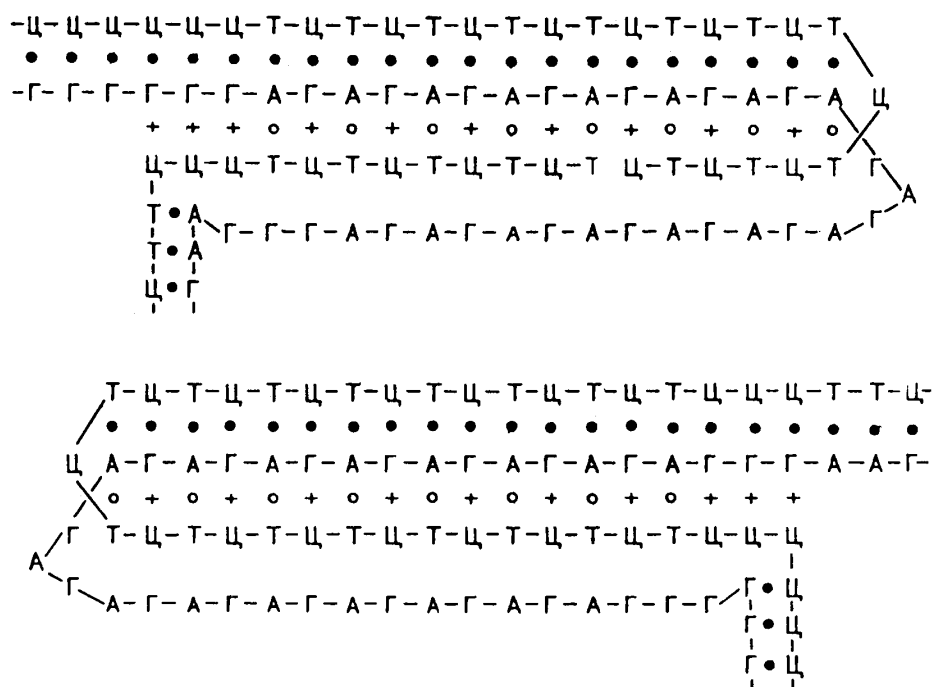


Рис. IX.24

Предполагаемое строение *H*-формы ДНК для участка  $(GA)_{16} \times (OC)_{16}$ . Показаны два возможных «изомерных» варианта структуры. (по Франк-Каменецкому, 1988)

что вместо исходного спаривания повторов  $A_1 - A_2$  и  $B_1 - B_2$  теперь происходит спаривание повторов  $B_1 - A_2$  в параллельных цепях. Кроме того, между выдвинутыми из цепей повторами  $A_1$  и  $B_2$  происходит также Уотсон-Криковское комплементарное спаривание с образованием миниспирали, или своего рода «паруса». На рис. IX.25 это обозначено наклонными штриховыми линиями.  $\square$

На рис. IX.26 приведен пример структуры ДНК, дающей РНК-транскрипт, который свертывается в SLS-форму.

Описанные выше необычные формы ДНК и РНК показывают большие потенциальные структурообразующие возможности этих молекул, многие из которых реализуются в клетке и играют важную биологическую роль в генной регуляции.

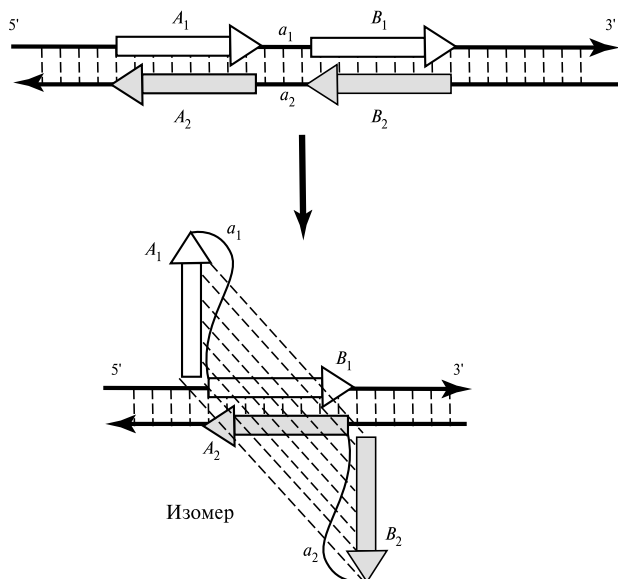


Рис. IX.25  
Схема образования SLS-структуры ДНК (по Иванову В. И.)

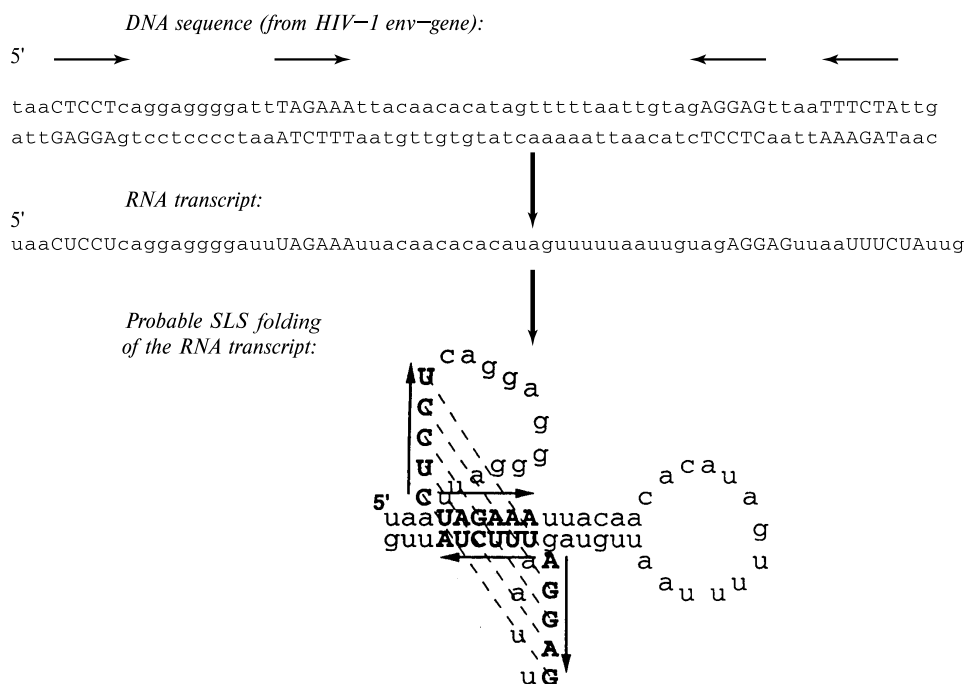


Рис. IX.26

Первичная структура ДНК, дающая РНК-транскрипт с образованием SLS-формы  
(по Иванову В. И.)

#### § 4. Состояние воды и гидрофобные взаимодействия в биоструктурах

Современные данные показывают, что вода принимает непосредственное участие в формировании структуры важнейших биополимеров, а также в процессах самосборки сложных надмолекулярных биоструктур. □

**Особенности строения воды.** Как известно, свойства воды обусловлены способностью ее молекул образовывать друг с другом водородные связи между атомами кислорода и водорода  $O-H \cdots O$  (см. § 2 гл. VIII). Молекула воды обладает такой электронной структурой, что может образовывать с соседними молекулами воды сразу четыре водородных связи, в которых участвуют по два протона и по две неподеленные пары электронов кислорода от каждой из них (рис. IX.27 и IX.28).

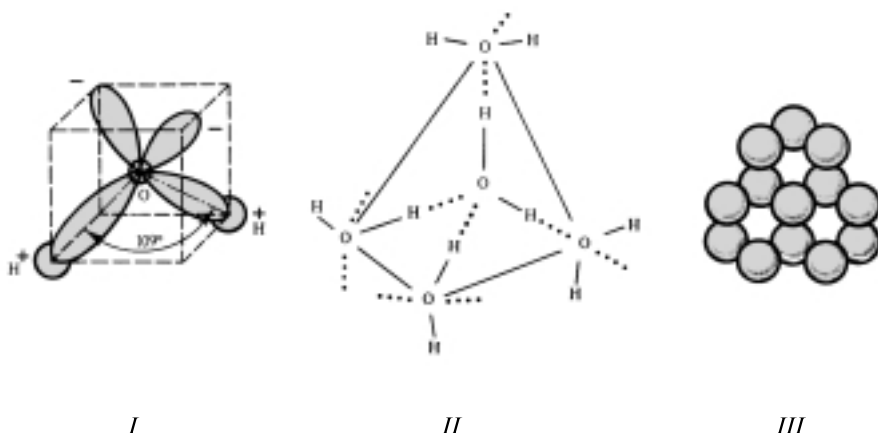


Рис. IX.27

Строение молекулы воды (I) и ближайшее ее окружение в структуре льда (II, III) (по Скрышевскому, 1971)

Жидкая вода представляет собой систему с сильно развитыми водородными связями, свойства которой хорошо описываются непрерывной, или континуальной, моделью, где практически все молекулы воды в среднем образуют водородные связи. Однако водородные связи в воде характеризуются широким набором углов и длин. Это означает, что потенциальная энергия водородной связи является непрерывной функцией угла межмолекулярной водородной связи и геометрических характеристик молекул воды. Максимум функции распределения энергии водородной связи соответствует расстоянию между атомами кислорода  $R_{O-O} = 0,286$  нм. В упорядоченной структуре водородные связи линейны и  $R_{O-O}$  сокращается до 0,275 нм.

Обычно лед (лед I) обладает тетраэдрической решеткой, где каждый атом кислорода воды расположен в центре тетраэдра, а в вершинах последнего находятся соседние атомы О. Такая ячеистая структура с четырьмя молекулами в каждой



ячейке определяет рыхлость и меньшую плотность льда. Такая же структура, хотя и с большим количеством нарушений, характерна и для жидкой воды. Наличие каркаса из молекул воды, связанных довольно прочно друг с другом четырьмя водородными связями, обуславливает относительно высокую температуру замерзания жидкой воды, высокую теплоту плавления льда, высокую теплоемкость и теплоту парообразования жидкой воды.

Д. Айзенберг и В. Каутцман ввели понятия о трех типах состояния жидкой и твердой воды (рис. IX.29). Ясно, что термодинамические параметры характеризуют в основном *D*-структуру. Обсуждается (Г. Шерага) кластерная модель состояния воды, согласно которой вода является смесью мономерных молекул и водородно-связанных кластеров. Кластеры распадаются и вновь образуются, что создает усредненное постоянное окружение для каждой молекулы воды. Согласно модели Самойлова, мономерные молекулы воды заполняют пустоты ячеистой решетки из водородных связей. При повышении температуры льда происходит разрушение его упорядоченной структуры, уплотнение молекул с уменьшением удельного объема. Однако одновременно идет и расширение жидкости вследствие увеличения амплитуд межмолекулярных колебаний. В результате

взаимодействия разных факторов минимум удельного объема достигается при  $4^\circ\text{C}$ . Уменьшение объема в модели Самойлова связано с активационными процессами отрыва молекул воды и заполнением ими полостей решетки.

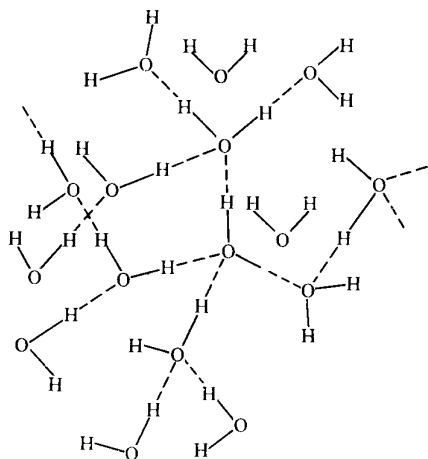


Рис. IX.28

Структура жидкой воды

С термодинамической точки зрения, образование упорядоченной структуры воды вызывает уменьшение энтропии ( $\Delta S < 0$ ), что невыгодно, так как приводит к увеличению полного термодинамического потенциала (свободная энергия)  $\Delta G = \Delta H - T\Delta S$  за счет положительного члена  $(-T\Delta S) > 0$ . Однако это увеличение с избытком компенсируется образованием четырех водородных связей, приводящим к снижению энергии (энтальпия) молекулы воды  $\Delta H < 0$ . Существенно, что при этом  $|\Delta H| > |T\Delta S|$ .  $\square$

Таким образом, в целом изменение свободной энергии при самопроизвольном образовании упорядоченной структуры воды отрицательно:  $\Delta F < 0$ . Это и определяет энергетическую выгодность этого процесса  $\Delta F \simeq \Delta G$  при  $\Delta V \simeq 0$  (см. (VI.1.10)). Молекулы воды испытывают колебательные движения около положения равновесия с характерным временем колебаний  $\tau \sim 10^{-13}$  с. Кроме того, за счет флуктуационных разрывов водородных связей (с энергией активации разрыва  $\sim 20$  кДж/моль при 298 К) может происходить перемещение отдельных молекул воды в пустоты в соседних ячейках. В жидком состоянии характерное среднее время такой самодиффузии составляет  $\tau_D \sim 10^{-10} \div 10^{-11}$  с, а во льду при  $0^\circ\text{C}$

оно значительно увеличивается — до  $\tau_D \geq 10^{-5}$  с (для сравнения: время валентных колебаний связи О–Н  $\tau_B \sim 10^{-14}$  с).

Именно за счет эффективной самодиффузии в жидком состоянии, когда молекулы воды занимают не только узлы решетки, но и некоторое число междоузлий, плотность воды превышает плотность льда. Важнейшие физикохимические свойства

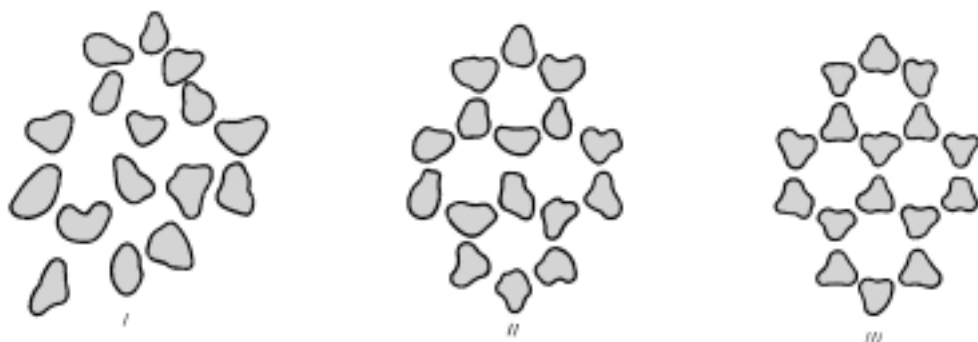


Рис. IX.29

Мгновенная (I), колебательно усредненная (II) и диффузионно усредненная (III) структуры воды (по М. В. Волькенштейну, 1975)

Молекулы воды осциллируют около своих средних положений с периодом колебаний  $\tau_k \sim 2 \cdot 10^{-13}$  с. При временах  $t \ll \tau_k$  возникает мгновенная I-структура, где положения молекул не усредняются (I) и могут отличаться от равновесия не более чем на 15%. При временах больших  $\tau_k$ , но меньших характерных времен  $\tau_n$  вращательных и трансляционных перемещений ( $\tau_n \sim 10^{-5}$  с для льда,  $\tau_n \sim 10^{-11}$  с для воды) получается колебательно усредненная K-структура ( $\tau_k < t < \tau_n$ ) (II). При  $t \gg \tau_n$  происходит усреднение диффузионных перемещений с образованием D-структуры воды, которая имеет упорядоченный характер (III)

воды, включая модификации льда, связаны, очевидно, с существованием большого числа переходных структурных элементов, взаимопревращения между которыми еще не вполне ясны.

В формировании структуры биополимеров важную роль играют эффекты изменения структуры самой воды вблизи растворенных веществ или в пограничном слое с другой поверхностью. Взаимодействие воды с полярными группами сводится к гидратации последних, причем энергии гидростатических взаимодействий в сольватной оболочке между водой и ионом могут достигать величин порядка энергии химических связей с увеличением времени «оседлой» жизни воды в несколько раз (для  $\text{Na}^+$  в два раза). Однако возможны и такие явления, когда вокруг иона  $\text{K}^+$  молекулы воды из-за сильного притяжения и большой плотности упаковки не образуют равномерного слоя. Это приводит к выталкиванию «лишних» молекул воды наружу и некоторому уменьшению вследствие этого среднего времени жизни в слое (для  $\text{K}^+$  на 10%). □

**Гидрофобные взаимодействия.** Наиболее интересны с точки зрения стабилизации структуры макромолекул взаимодействия воды с гидрофобными соединениями. Изучение растворимости неполярных углеводов при переносе их из неполярных

растворителей в воду показало, что процесс растворения этих соединений в воде носит экзотермический характер. Он идет с повышением температуры и выделением тепловой энергии, что свидетельствует о понижении энтальпии системы при смешивании углеводов с водой ( $\Delta H < 0$ ). Однако это понижение энтальпии компенсируется уменьшением энтропии ( $\Delta S < 0$ ) и, как следствие этого, возрастанием полного термодинамического потенциала ( $\Delta G > 0$ ). В целом уменьшение энтальпии ( $\Delta H < 0$ ) по абсолютной величине меньше, чем уменьшение энтропии, и, следовательно, растворение в воде неполярных соединений термодинамически невыгодный процесс:

$$\Delta G = \Delta H - T\Delta S > 0.$$

Так, при переносе из бензола и растворении этана в воде при 298 К энтропия понижается на  $\Delta S = -84$  Дж/(моль · К), а энтальпия — на  $\Delta H = -9240$  Дж/моль. Общее возрастание термодинамического потенциала системы этан–вода по сравнению с системой этан–бензол составляет

$$\begin{aligned}\Delta G &= \Delta H - T\Delta S = -9240 + 298 \cdot 84 \simeq \\ &\simeq 25500 - 9240 \simeq 16000 \text{ Дж/моль} = 16 \text{ кДж/моль}.\end{aligned}$$

Следствием этого является отталкивание водой молекул неполярных веществ, которым термодинамически выгоднее взаимодействовать друг с другом, чем с водой. Эффект отталкивания растворителя и представляет собой так называемые гидрофобные взаимодействия, которые, как видно, не связаны с существованием каких-либо особых сил взаимного притяжения между элементами системы, а имеют чисто термодинамическую причину.

Молекулярная картина, отражающая изменение структуры воды и взаимную ориентацию молекул растворенного вещества и растворителя, достаточно сложна. Так, в данном случае в отличие от процесса образования ассоциатов молекул в чистой воде уменьшение энтропии не компенсируется снижением энтальпии системы.

Прежде всего возникает вопрос о том, какие конкретные изменения в системе углеводород–вода приводят к уменьшению ее энтропии. Очевидно, растворение углеводорода в воде связано с внедрением его молекул в структурированные области воды и их разрушением. Небольшие молекулы, способные встраиваться в льдоподобный каркас водных кластеров, лучше растворяются, поскольку они относительно мало разрушают водородные связи. Однако, как правило, достаточно большие молекулы гидрофобных соединений разрушают водородные связи в воде. Так как энергия возникающих ван-дер-ваальсовых взаимодействий между молекулами растворенного вещества и водой меньше энергии разрушенных водородных связей чистого растворителя, то в целом это должно было бы привести к увеличению энергии системы. Следовательно, данное обстоятельство и является причиной разделения системы на две фазы молекулы чистого растворителя и гидрофобного соединения. Вместе с тем общее уменьшение энтропии системы углеводород–вода при растворении неполярных веществ в воде интерпретируется как образование «льдоподобной оболочки» вокруг неполярной молекулы или «структурирования воды неэлектролитами».

Эти понятия неточно отражают природу явления, поскольку в действительности никаких «айсбергов», структурно подобных льду, вокруг неполярных молекул

не образуется. При помещении молекул неэлектролита в воду молекулы последней вынуждены ориентироваться на его поверхности для обеспечения максимального числа водородных связей в условиях нарушенной регулярной ячеистой структуры и исчезновения свободных пустот в пограничной области. Это приводит к уплотнению молекул воды, прилежащих к гидрофобной поверхности, увеличивает среднее время «оседлой» жизни до  $\tau \sim 10^{-7}$  с и ведет к уменьшению энтропии.

Уплотненный упорядоченный каркас из молекул воды образуется вокруг каждой растворенной в воде молекулы гидрофобного соединения. Образование уплотненного слоя находит свое отражение, в частности, в сдвиге инфракрасных полос поглощения воды ( $5180$  и  $2130 \text{ см}^{-1}$ ), свидетельствующих об упрочении водородных связей между молекулами воды. Дальнейшее повышение концентрации неэлектролита в воде приводит уже к гидрофобным ван-дер-ваальсовым взаимодействиям между самими растворенными молекулами и образованию их ассоциатов с уплотненным упорядоченным слоем воды на гидрофобной поверхности. Наряду с уменьшением подвижности молекул воды вклад в понижение энтропии системы дает ограничение подвижности углеводородных цепей растворенных молекул неэлектролитов.

**Стабилизация биоструктур.** Гидрофобные взаимодействия играют существенную роль в формировании биоструктур, представляя собой один из основных факторов их стабилизации. В самом деле, эффект взаимодействия полярных групп белка с полярными молекулами воды связан с преобладанием полярных аминокислотных остатков на поверхности белковой глобулы. Однако наряду с этим возможно и взаимодействие посредством водородных связей полярных пептидных связей ( $\text{NH} \cdots \text{OC}$ ), принадлежащих разным участкам цепи внутри глобулы. Так как энергия водородных связей между пептидными связями в белке и между ними и водой примерно одинакова, это должно было бы приводить к рыхлой структуре макромолекулы в водном растворе. Однако реально существующая структура упорядочена и компактна и, как можно заключить, в основном определяется именно гидрофобными взаимодействиями. Отдельные аминокислотные остатки различаются по своим гидрофобным свойствам и могут вести себя как полярные или неполярные соединения. Термодинамическую оценку степени гидрофобности делают по величине изменения  $\Delta G$ , приходящегося на боковую группу аминокислоты при ее переносе из этанола в воду (К. Танфорд).

Так, к гидрофобным аминокислотам относятся такие, как *три* ( $\Delta G = +12570$  Дж/моль), *иле* ( $+12440$ ), *фен* ( $+11100$ ), *гис* ( $+5900$ ), *мет* ( $+5500$ ), а гидрофильные представлены *ала* ( $+3000$ ), *глу* ( $+2300$ ), *сер* ( $+170$ ), *гли* ( $0$ ), *асн* ( $-40$ ), *гли* ( $-420$ ). Распределение полярных и неполярных свойств не всегда совпадает с гидрофобностью аминокислот, которые могут проявлять те или иные свойства в зависимости от своего положения и окружения в белке. Взаимодействие неполярных групп с водой приводит к их преимущественному «заталкиванию» внутрь белковой глобулы и соответственно выходу наружу полярных групп. В предельном случае общая топографическая модель белковой глобулы предполагает существование ядра глобулы, заполненного гидрофобными аминокислотами и защищенного слоем обращенных в воду полярных групп. На основании этих представлений Фишер приближенно оценил форму глобулы, разбив все аминокислоты на две группы: полярные и гидрофильные (*арг*, *асп*, *гис*, *глу*, *лиз*, *сер*, *тир*, *тре*) и гидрофобные (остальные 12). Если все остатки занимают примерно

одинаковые объемы то в этом случае отношение  $b_s$  числа гидрофильных остатков к гидрофобным равно

$$b_s = \frac{r^3}{(r-d)^3} - 1,$$

где  $r$  — радиус глобулы;  $d$  — толщина внешнего гидрофильного слоя, равная  $\sim 0,4$  нм.

Когда число гидрофильных остатков достаточно для покрытия всей глобулы, последняя принимает сферическую форму ( $b = b_s$ ).

В реальных условиях, однако, наблюдаются отклонения величины  $b$  измеренной экспериментально, от теоретических значений  $b_s$  для сферической глобулы. При  $b > b_s$  число гидрофобных остатков превышает минимальное их число, необходимое для защиты от воды внутреннего гидрофобного ядра, и глобула становится эллипсоидальной. Наоборот, при  $b < b_s$  незащищенные гидрофобные участки на поверхности глобул приводят к их слипанию. Эта картина, конечно, дает лишь грубое представление о реальной действительности.

Детальные расчеты (Д. Ф. Брандтс), выполненные с учетом статистического характера распределения гидрофобных, гидрофильных и нейтральных групп в пределах макромолекулы и значений свободной энергии переноса группы на поверхность или внутрь глобулы, показали, что около 10–25% гидрофобных групп может располагаться на поверхности глобулы. Эксперименты выявили, что в ряде случаев число неполярных групп на поверхности даже превышает число их внутри белка (в молекуле лизоцима их числа соответственно 32 и 16). Очевидно, гидрофобные силы имеют решающее значение в определении общего характера топографии белковой глобулы, но истинная картина взаимодействий в пределах макромолекулы значительно сложнее.

☐ Оценивая роль различных взаимодействий в стабилизации глобулярных белков, следует считать, что характер нативной конформации определяется не каким-либо одним эффектом, а представляет собой результат совместного тонко сбалансированного действия целого ряда энергетических и энтропийных факторов. Водородные связи, образованные между полярными группами и водой и внутри глобулы, — главный фактор в обеспечении стабильности отдельных областей молекулы белка. Они ограничивают локальные конформационные изменения внутри белка, определяя жесткость конструкции и общий характер потенциальных барьеров для внутренних движений частей нативной структуры. В то же время гидрофобные взаимодействия между боковыми группами на отдельных участках основной цепи играют решающую роль в процессах сворачивания глобулы из первичной аминокислотной последовательности и в определении ее общей формы. В обоих случаях вода как растворитель имеет огромное значение, облегчая полярные взаимодействия за счет образования водородных связей как на поверхности, так и внутри макромолекулы белка. ☐

☐ **Неоднородность распределения молекул воды.** Гетерогенность организации макромолекулярной структуры влияет на распределение молекул воды. Очевидно, вода может существовать в виде объемной фазы свободного растворителя и связанной воды, состояние которой зависит от природы и мест локализации белковых групп, с которыми она взаимодействует. ☐

Каковы данные по состоянию воды в гидратной оболочке белка? Основной вклад в энергию гидратации дают водородные связи между водой и полярными группами молекулы белка. Для образования гидратной оболочки глобулярных белков имеет значение пространственная доступность протон-донорных и протон-акцепторных центров для взаимодействия с молекулами воды. Оказалось, что гетероатомы нерегулярно расположены на поверхности глобулы, которая не может служить матрицей для кристаллизации воды. Так как число и размеры гидрофобных участков на поверхности также невелики, то «шуба» из уплотненных молекул воды вокруг глобулы не образуется. количество гидратационной воды, определенное различными методами, составляет 0,3–0,4 г  $\text{H}_2\text{O}$ /г сухого белка, а общее содержание воды в кристаллах глобулярных белков не превышает, как правило, 0,45–0,60 г  $\text{H}_2\text{O}$ /г сухого белка. Следовательно, количество свободной воды в белке невелико. Она, в частности, может заполнять внутренние «полости», свободные от белкового вещества, содержание воды в этих полостях также невелико (в лизоциме — 2, трипсине — 12 молекул). Она может обмениваться с поверхностными водными слоями вследствие флуктуационных «открытий» внутренних полостей.

Реальная поверхность белковой глобулы составляет несколько сотен квадратных нанометров. Так, для инсулина  $S = 34 \text{ нм}^2$ , лизоцима —  $65 \text{ нм}^2$ . Количества связанной воды может хватать для покрытия не более 60–70% поверхности белковой глобулы и явно недостаточно для образования на поверхности глобулы сплошного монослоя воды.

Измерение изотерм сорбции и десорбции воды позволило оценить число молекул воды в более прочно удерживаемом первом слое на поверхности белка. Оно составляет от нескольких десятков (лизоцим 45–64) до сотен ( $\alpha$ -химотрипсин 72–103) молекул. Основными центрами первичной гидратации белка являются гетероатомы боковых полярных групп аминокислот и пептидные группы. Амидные группы и ионные пары не принимают такого активного участия в удержании гидратационной воды и представляют собой более «слабые» центры связывания. Среднее число молекул воды составляет в среднем четыре на один центр.

Общая картина гидратации поверхности белка приведена на рис. IX.30. Отдельные центры окружены молекулами воды, которые образуют, кроме того, и водные мостики между гидратационными областями, принадлежащими отдельным центрам. В результате образуется сплошная гидратационная «кружевоподобная» оболочка, соединяющая отдельные полярные группы. Гидратная оболочка на поверхности глобулы включает молекулы воды, принадлежащие локальным центрам гидратации. Время жизни слабосвязанных молекул воды в такой гидратной оболочке невелико ( $\tau \sim 10^{-11} \div 10^{-10} \text{ с}$ ), но около центра оно намного больше ( $\tau \sim 10^{-6} \text{ с}$ ).

Молекулы воды в первичном прочном слое образуют друг с другом в среднем не менее двух-трех водородных связей. В целом около поверхности белка может удерживаться до двух-трех слоев воды, многослойной «шубы» кристаллизационной воды при этом не образуется. При сорбции воды отдельные гидрофобные участки на поверхности «заполняются» водой лишь после ее энергетически более выгодного связывания с соседними полярными группами. Прочно связанная вода составляет около 10% от массы биополимеров и характеризуется большими временами жизни  $10^{-3} - 10^{-7} \text{ с}$ .

Для понимания свойств прочносвязанной воды необходимо учесть, что ее фракция неоднородна в отношении участков и групп биополимера, с которыми взаимодействуют молекулы воды. Вода связывается прежде всего на граничащей с ней поверхности биополимера. Следовательно, движение прочно связывающих ее соответствующих боковых групп должно определять и времена жизни этой фракции молекул связанной воды. Это будет, конечно, справедливо, если время жизни молекул

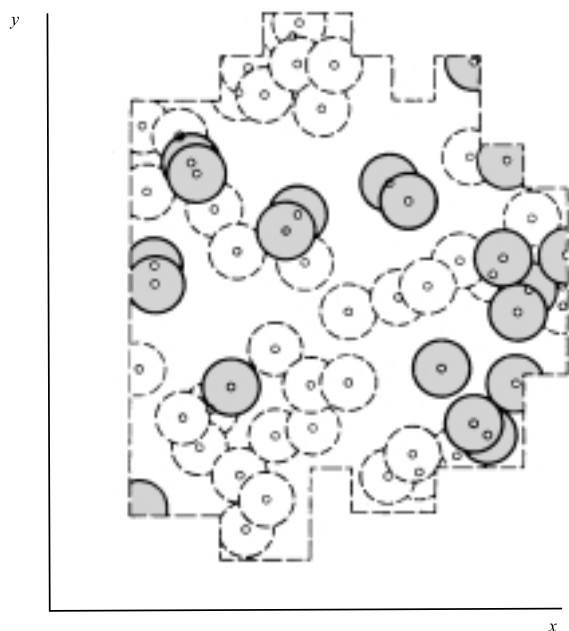


Рис. IX.30

Карта гидратации поверхности молекулы *p*-химотрипсина (по Ю. И. Хургину, 1980)

Сплошными линиями обозначена область локализации молекул воды, связанных гетероатомами поверхностных боковых полярных групп; пунктирными окружностями — область локализации молекул воды на пептидных группах

воды в связанном состоянии заметно больше времени, определяемого движением биополимера в целом или его отдельных участков, что в общем выполняется для прочносвязанной воды ( $\tau \sim 10^{-3} \div 10^{-7}$  с). Кроме того, время жизни молекул воды в растворе ( $\tau \sim 10^{-11}$  с) биополимеров может увеличиваться за счет обмена с водой, прочно связанной в подповерхностных медленно движущихся ( $\tau \sim 10^{-5}$  с) слоях белка. Очевидно, на время жизни воды будут влиять и изменения подвижности макромолекул, например при изменении температуры, образовании межмолекулярных связей и т. д.

Таким образом, физико-химическая природа связанной воды во многом обусловлена ее взаимодействиями с малоподвижными структурами. Этим определяется существование в растворах биополимеров нескольких фракций связанной воды, времена жизни которой занимают промежуточное положение между значениями, характерными для свободной ( $10^{-11}$  с) и сильносвязанной воды ( $10^{-3} - 10^{-5}$  с) (С. И. Аксенов). □

Динамические свойства внутриклеточной воды в значительной степени отражают состояние клеточных структур. Существует также ряд данных, указывающих на непосредственное участие небольших количеств воды в изменении конформации глобулярных белков. В следующей главе будут описаны подробно характеристики и модели динамической подвижности биомакромолекул. Сейчас лишь необходимо отметить, что функционирование белков тесно связано не только с характером их конформации, но, главное, с их конформационной подвижностью, зависящей от присутствия воды. Так, при низкой степени гидратации препаратов  $\alpha$ -химотрипсина возникающие дополнительные контакты между поверхностными дегидратированными полярными группами приводят к увеличению жесткости глобулы  $\alpha$ -химотрипсина и потере им ферментативной активности в диметилсульфоксиде. В сильно высушенных препаратах, вплоть до некоторого критического значения гидратации, вообще не наблюдается никакой активности. Восстановление последней при увеличении степени гидратации образца происходит резко в узком диапазоне увеличения числа молекул  $H_2O$  от 170 до 180 на одну молекулу белка. Очевидно, в этой области происходит растормаживание определенных степеней свободы, функционально важных для ферментативного акта. Существенно, что необходимое для этого процесса количество воды намного меньше, чем было бы нужно для завершения образования гидратной оболочки (Ю. И. Хургин).

Появление некоторых изменений в структуре фибриллярных и глобулярных белков, полипептидов при их гидратации отмечалось и по результатам ИК-спектроскопии. Наряду с данными о наличии внутри белковых глобул отдельных молекул воды это свидетельствует о неотъемлемой роли воды в формировании молекулы белка. Очевидно, внедрение некоторого количества воды в качестве структурного элемента в белковую глобулу может облегчить необходимое для обеспечения функциональной активности изменение ее конформации за счет эффекта «разрыхления» и «смазки». Таким образом, роль воды в формировании биоструктур связана, с одной стороны, со стабилизирующими гидрофобными взаимодействиями, а с другой — состоит в облегчении конформационной подвижности макромолекул.

**Энергетические эффекты гидратации.** Анализ эффектов гидратации при расчетах конформационной энергии на примере метиламидов *N*-ацетил-аминокислотных остатков (Г. Шерага) показал, что конформационная энергия белковой глобулы может не зависеть существенно от присутствия воды. Гидратация включает специфические взаимодействия воды с полярными атомами белковых групп за счет образования с ними водородных связей и неспецифические взаимодействия, представляющие собой попарные ван-дер-ваальсовы взаимодействия между всеми атомами воды в гидратационной оболочке и атомами белковой группы (см. § 1, 2, гл. VIII). Преимущественные места гидратации (метиламидов аминокислот) представляют собой, как правило, амидные, гидроксильные и карбонильные группы. Водородные связи образуются между водой и кислородом карбонильной группы и атомами пептидной связи.

Подробные вычисления и построенные на их основе энергетические карты (см. § 1 гл. IX; рис. IX.9 и IX.10) показали, что при гидратации существуют области значений углов  $\varphi$ ,  $\psi$ , где происходит стабилизация и дестабилизация с соответству-



ющими изменениями глубин энергетических минимумов. Стабилизация в области изменения углов  $\varphi$  ( $-40^\circ, -100^\circ$ ),  $\psi$  ( $0, 80^\circ$ ) достигается за счет снятия при гидратации энергетически невыгодных электростатических взаимодействий между диполями пептидной группы. Образование водородных связей с водой при гидратации ослабляет выгодные электростатические и ван-дер-ваальсовы взаимодействия между расположенными друг против друга CO- и NH-группами. Это ведет к дестабилизации структуры в областях изменения  $\varphi$  ( $-40^\circ, -100^\circ$ ),  $\psi$  ( $50^\circ, 110^\circ$ ). Для аланина значения пары углов  $\varphi, \psi$ , соответствующих области минимума энергии, где наблюдается дестабилизация, составляют  $-86^\circ, 82^\circ$ .

Таким образом, влияние воды на конформационную энергию пептидов существенно не изменяет мест локализации энергетических минимумов на конформационной карте. Однако при этом может сильно изменяться относительная стабильность отдельных минимумов конформационных участков, что создает, в свою очередь, предпосылки для характерных изменений конформации белка в различных функциональных процессах.  $\square$

Кооперативный характер системы биополимер — вода ярко проявляется в процессах внутримолекулярного плавления биологических макромолекул. Температурные зависимости изменения теплоемкости, времен подвижности протонов и вязкости растворов биополимеров характеризуются одновременным и резким изменением этих параметров в одном и том же температурном диапазоне (рис. IX.31).

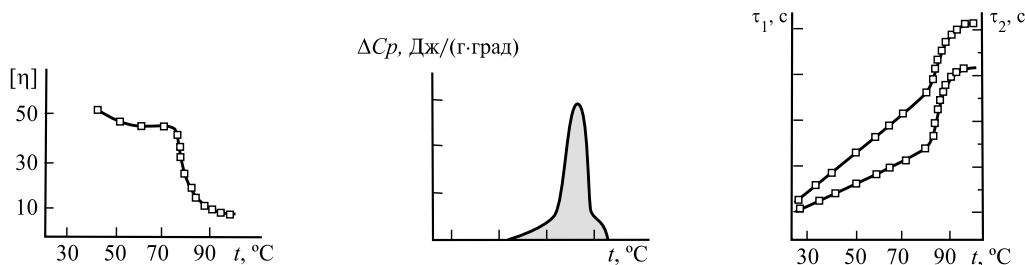


Рис. IX.31

Изменение вязкости  $\eta$ , теплоемкости  $\Delta C_p$  и времени спин-решеточной ( $\tau_1$ ) и спин-спиновой ( $\tau_2$ ) релаксаций в растворе ДНК тимуса телят (по Г. М. Мревришвили, 1981)

Небольшие изменения в количестве и состоянии относительно небольшой фракции молекул воды, образующих гидратный слой макромолекулы, приводят к резким изменениям термодинамических и релаксационных параметров всего раствора в целом. Быстрый обмен между протонами молекул воды вовлекает все фракции воды, включая молекулы, находящиеся во внутренних слоях белка. Все это делает систему биополимер — вода единой кооперативной системой, в которой любые изменения в состоянии как растворителя, так и макромолекулы носят взаимосвязанный и взаимообусловленный характер.

### § 5. Динамика фазовых переходов в макромолекулах

Регулярный характер конформации биополимеров зависит от характера объемных взаимодействий, которые изменяются при разных внешних условиях (температура, рН, ионная сила раствора). Это соответственно нарушает регулярность в строении макромолекулы и способствует превращению ее структуры в статистический клубок. Так, изменение рН раствора синтетического полипептида поли- $\gamma$ -бензил-L-глутамата (ПБГ) и полиглутаминовой кислоты (ПГК) приводит к потере  $\alpha$ -спиральности (рис. IX.32), что характеризуется изменением значений характеристической вязкости  $[\eta]$ , удельной вращающей способности  $[\alpha]$ , степени ионизации. Видно, что при рН = 6 происходит резкий переход спираль — клубок, связанный с исчезновением упорядоченной структуры. Сходные результаты были получены для полилизина, сополимеров лизина и глутаминовой кислоты.

Наблюдаемые изменения происходят в отдельной молекуле, а не в кристалле и по своему резкому S-образному характеру свидетельствуют о кооперативном характере перехода. Это внутримолекулярное плавление, аналогичное внешне фазовому переходу. Физическая основа переходов спираль — клубок состоит в том, что состояние макромолекулы, у которой все мономерные звенья участвуют в образовании водородных связей, более выгодно энергетически, а состояние свободной макромолекулы более выгодно энтропийно из-за ее гибкости. Поэтому энергии этих двух состояний по-разному меняются при изменении температуры, состава растворителя, рН. Температура перехода соответствует равенству энергий двух состояний.

**Кооперативный характер перехода.** Изменение энергии молекулы при образовании водородной связи одним из звеньев существенно зависит от наличия или отсутствия водородных связей у соседних с ним звеньев. Эта кооперативность носит энтропийный характер.

Рассмотрим причины кооперативного характера превращения, исходя из того, что плавление приводит к разрушению упорядоченной вторичной структуры и разрыву водородных связей между звеньями полимера. Если разность энергий состояний спираль — клубок  $\Delta G = \Delta H - T\Delta S$  известна как функция температуры, то доля спирализованных звеньев  $\Theta$  может быть найдена по формуле

$$\Theta = \frac{\exp(\Delta G/RT)}{1 + \exp(\Delta G/RT)}. \quad (\text{IX.5.1})$$

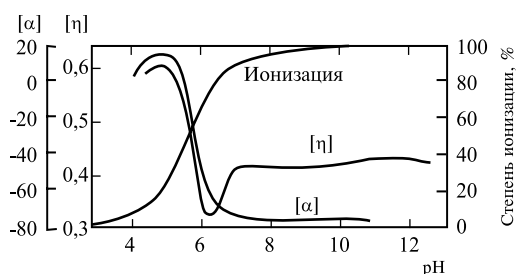


Рис. IX.32

Зависимость степени ионизации, характеристической вязкости ( $[\eta]$ ) и удельной вращательной способности ( $[\alpha]$ ) от рН среды в области перехода спираль — клубок (по М. В. Волькенштейну, 1975)

а соответственно доля расплавленного материала составит

$$1 - \Theta = \frac{1}{1 + \exp(\Delta G/RT)}. \quad (\text{IX.5.2})$$

Резкость перехода при нагревании можно вычислить, дифференцируя выражение (IX.5.2) по  $T$  и учитывая, что при  $T = T_{\text{пл}}$ ,  $\Delta G = 0$

$$\left| \frac{d\Theta}{dT} \right|_{T_{\text{пл}}} = \left| \frac{\Delta H}{4RT_{\text{пл}}^2} \right|. \quad (\text{IX.5.3})$$

где  $H$  — энтальпия перехода спираль–клубок. Величина  $\Delta H$  должна зависеть от числа  $n$  освобожденных звеньев, т. е.  $\Delta H \simeq n\Delta H_0$ , где  $H_0$  — энтальпия разрыва связей одного звена. Кооперативный характер процесса плавления означает, что в области перехода одновременно меняется состояние  $n$  звеньев цепи, а резкость перехода пропорциональна  $n$ .

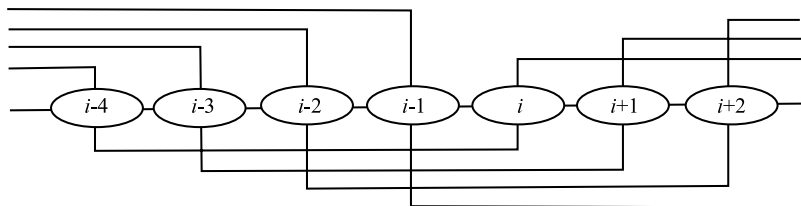


Рис. IX.33

Модель  $\alpha$ -спирали как одномерного гомополимера (объяснение см. в тексте) (по М. В. Волькенштейну, 1965)

**Количественная теория кооперативных переходов спираль–клубок.** Эта теория должна давать выражения для зависимости числа разорванных водородных связей от температуры. В  $\alpha$ -спирали каждое звено соединено водородными связями с двумя другими, отстоящими от него на четыре звена в цепи (рис. IX.33). Для освобождения данного звена надо одновременно разорвать не две, а сразу несколько водородных связей. Теория переходов спираль–клубок основана на одномерной модели кооперативных явлений. Энергия всей цепи зависит от набора свободных и связанных звеньев, поскольку их энергии отличаются друг от друга. Для нахождения условий перехода спираль–клубок необходимо вычислить статистическую сумму цепи, учитывая всевозможные распределения водородных связей между звеньями (см. § 3 гл. VIII). Статистическая сумма цепи равна

$$Z = \sum_{\{\mu_i\}} \exp(-G\{\mu_i\}/RT). \quad (\text{IX.5.4})$$

где  $G$  — собственная энергия  $i$ -го звена, включающая и энергию его водородных связей, и, принимает значение, равное нулю для каждого звена в свободном состоянии и равное единице в связанном состоянии. Подробные расчеты (Б. Г. Зимм,

Дж. К. Брегг) показывают, что в произведении (IX.5.4) наибольшее значение имеют два типа множителей, которые определяют энергию состояния звена, следующего за связанным или за свободным звеньями.

Множитель для каждого звена в связанном состоянии ( $\mu_i = 1$ ) равен

$$S = \exp(-\Delta G/RT) = \exp[-(\Delta H - T\Delta S)/RT]. \quad (\text{IX.5.5})$$

Выражение (IX.5.5) имеет смысл константы равновесия для образования водородной связи в звене, следующем за связанными звеньями, т. е. константы равновесия процесса увеличения серии связанных звеньев на одно звено за счет соседнего участка свободных звеньев. Величина  $\Delta H$  определяет суммарное изменение энергии связей при фиксации звена, а  $\Delta S$  включает уменьшение энтропии при ограничении его подвижности. Величина  $S$  определяет, очевидно, долю спирализованных звеньев в цепи. Другой множитель в (IX.5.4) для каждого звена в связанном состоянии ( $\mu_i = 1$ ), следующего за тремя (и более) свободными звеньями ( $\mu_{i-1} = \mu_{i-2} = \mu_{i-3} = 0$ ), равен

$$\sigma = \exp(-G_{\text{иниц}}/RT),$$

где  $G_{\text{иниц}}$  — энергия инициирования спирального участка. Эта энергия затрачивается для закрепления сразу четырех звеньев, что необходимо для фиксации одного звена, следующего за тремя свободными. Фактически в этом случае при фиксации одного звена организуется сразу целый спиральный участок с понижением энтропии, что и является следствием кооперативности системы.

Если принять, что при вращении свободного звена около каждой связи возможны три поворотных изомера, то снижение энтропии при фиксации одного звена

$$G_{\text{иниц}} \simeq G_i(0, 0, 0, 1) = -RT \ln(1/3),$$

а при образовании спирального участка из четырех звеньев

$$G_{\text{иниц}} \simeq -4RT \ln(1/3) \simeq 10,5 \text{ кДж/моль.}$$

Отсюда следует, что множитель  $\sigma$ :

$$\sigma \simeq 10^{-2}. \quad (\text{IX.5.6})$$

Степень кооперативности процесса тем больше, чем меньше множитель  $\sigma$ , зависящий от понижения энтропии при фиксации одного звена. При  $\sigma = 1$  кооперативность отсутствует ( $G_{\text{иниц}} = 0$ ). Статистическая сумма в (IX.5.4), найденная матричными методами (ср. (VIII.3.10 и VIII.3.11)), оказывается равной

$$Z = (1 + S)^N, \quad (\text{IX.5.7})$$

где  $N$  — число звеньев в цепи.

Доля звеньев, связанных водородными связями,

$$\Theta = \frac{1}{N} \frac{\partial \ln Z}{\partial \ln S} = \frac{S}{1 + S}. \quad (\text{IX.5.8})$$

плавно зависит от температуры и не дает резкого перехода.

При  $\sigma = 0$  (полная кооперативность,  $G_{\text{иниц}} \rightarrow \infty$ )

$$Z = 1 + S^N,$$

а доля связанных звеньев

$$\Theta = \frac{1}{N} \frac{\partial \ln Z}{\partial \ln S} = \frac{S^N}{1 + S^N}. \quad (\text{IX.5.9})$$

В этом случае цепь может находиться только в двух состояниях: полная спиральность ( $S > 1$ ) и клубок ( $S < 1$ ). При  $S > 1$ ,  $S^N \gg 1$ , т. е. цепь является полностью спиралью, а при  $S < 1$ ,  $S^N \ll 1$ ,  $Z_{\text{клуб}} = 1$ , т. е. цепь переходит в состояние клубка. Таким образом, при  $S = 1$  происходит резкий кооперативный переход спи-

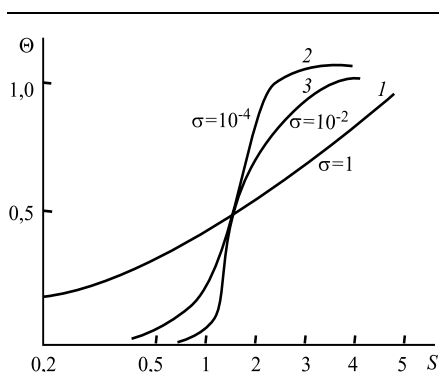


Рис. IX.34

Переход спираль–клубок (по М. В. Волькенштейну, 1965):

1 — некооперативный, 2, 3 — кооперативный

раль–клубок (рис. IX.34). Этот результат следует и из условия  $\sigma = \exp(-G_{\text{иниц}}/RT)$ , так как при бесконечно большой энергии инициирования спирального участка в цепи невозможно сосуществование в одной молекуле спиральных и клубкообразных областей. Естественно, что при конечных, но малых значениях  $\sigma$  переход осуществляется в некоторой области изменения  $S$  около точки  $S = 1$ . Это означает, что в реальных системах имеется некоторый небольшой, но конечный интервал изменения температуры, ширина которого определяет область сосуществования спиральных и клубкообразных участков цепи.

Изменения  $\Delta H$  и  $\Delta S$  в процессе перехода сильно зависят от взаимодействия мономерных звеньев с растворителем, молекулы которого могут конкурировать с группами цепи

за водородные связи. Так, в ряде случаев увеличение энтропии самой макромолекулы при переходе спираль–клубок компенсируется уменьшением энтропии при образовании водородных связей полимер–растворитель и перестройкой структуры самого растворителя (см. § 4 этой главы). Поэтому приведенные выше параметры переходов спираль–клубок следует интерпретировать с учетом изменения состояния растворителя.

**Роль выделенных степеней свободы в фазовых переходах.** До сих пор в спиральном и клубкообразном состояниях не рассматривались какие-либо определенные степени свободы, по которым совершаются фазовые переходы. Разрыв и образование связей таким образом могли происходить в любой области макромолекулы при переходе спираль–клубок.

В следующих главах (X, XI) будут подробно рассмотрены конформационные превращения биологических полимеров, сопровождающие их функционирование; они совершаются по выделенным степеням свободы и в основном носят не статистический, а детерминистический характер. Особенность фазовых переходов в

макромолекулярных системах также связана с тем, что они происходят в гетерогенных молекулярных конструкциях, где заранее запрограммировано расположение жестких ( $\alpha$ -спирали в белке) и мягких участков, типичное для данного класса белковых макромолекул. В силу гетерогенности конструкции разные вклады в свободную энергию концентрируются в разных местах макромолекулы в зависимости от вида энергии. Так, энергия электростатического взаимодействия концентрируется вблизи заряженных групп, а энтропийная часть свободной энергии — в размыкающихся и замыкающихся связях-фиксаторах, роль которых выполняют водородные связи. В отличие от этого в термодинамически равновесных однородных системах все части свободной энергии распределены по телу равномерно. Значение теплоты перехода ( $\Delta Q \sim 42$  кДж/моль) в пересчете на одну молекулу сопоставимо с обычными значениями теплоты переходов первого рода (см. § 2 гл. VII). Однако в расчете на единицу массы или на моль углерода величина  $\Delta Q \sim 4,2 \cdot 10^{-3}$  кДж/г, т. е. в десятки тысяч раз меньше. Столь же малыми оказываются удельные изменения энтропии и теплоемкости. По сравнению с общим числом межатомных связей в макромолекуле число разрываемых связей  $n$  очень мало. Это означает, что разрыв всего лишь тысячной доли имеющихся связей приводит уже к фазовому (конформационному) переходу в макромолекуле. Именно это характерно не для равновесных тел, а для конструкций, где изменение небольшого числа связей-фиксаторов влечет перестройку системы в целом (Д. С. Чернавский). Тепловые эффекты, сопровождающие эти перестройки, весьма малы (см. § 2, гл. VII).

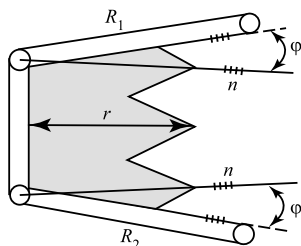


Рис. IX.35

Эквивалентная механическая схема сегментальных движений макромолекул (по Чернавскому):

$\tau$  — резервуар энергии;  $R_1, R_2$  — рычаги (порядка размеров макромолекул  $\sim 200$  нм);  $n$  — водородные связи, выполняющие роль фиксаторов;  $\phi$  — углы вращения

Рассмотрим фазовый переход при изменении температуры в макромолекуле, где он сопровождается разрывом определенной связи по выделенной степени свободы (рис. IX.35). Необходимо учитывать изменение энтропии как выделенной степени свободы, так и других микроскопических степеней свободы. Возможны два состояния: 1) пружина напряжена и фиксирующие связи замкнуты — это конструкция типа фермы ( $f$ ); 2) пружина расслаблена и фиксаторы разомкнуты — это машина ( $m$ ). При подсчете свободных энергий надо принимать во внимание, что группы, участвующие в фиксирующих связях, обладают различной энтропией в состояниях  $m$  и  $f$ .

В том случае, когда величина  $\Delta G$  зависит от взаимодействия с водой, ее можно оценить по числу центров связывания молекул воды. Например, в рассматриваемой модели связи, расположенные в полости, могут фиксировать воду в состоянии  $m$  (полость открыта), но не в состоянии  $f$ . Таким образом, при переходе  $m \rightarrow f$  вода вытесняется из полости и переходит на поверхность глобулы. Можно оценить чи-

сло групп, участвующих в фиксации состояния  $f$ . В обычных условиях величина  $n$  порядка 10.

Кооперативные фазовые переходы в макромолекулярной конструкции обусловлены разрывом и фиксацией небольшого числа  $n$  связей. Это определяется наличием конструктивных элементов в системе и силами упругости. В обычных макроскопических системах фазовые превращения связаны с процессами на границе фаз, а силы поверхностного натяжения определяют минимальный размер «зародыша» новой фазы и соответствующий ему барьер перехода.

Роль конструктивных элементов и взаимодействий между ними проявляется в наличии особого промежуточного состояния, возникающего в процессе денатурации белков (Е. М. Шахнович, А. В. Финкельштейн, О. Б. Птицын). Промежуточное состояние (расплавленная глобула) получается из нативного путем кооперативного температурного плавления в узком интервале температур. Оно энергетически менее выгодно, а энтропийно значительно более выгодно, чем нативное. Это связано с резким ослаблением в нем внутримолекулярных взаимодействий и уменьшением внутримолекулярной упорядоченности. Небольшое набухание в промежуточном состоянии резко ослабляет короткодействующие силы притяжения Ван-дер-Ваальса по сравнению с гидрофобными взаимодействиями. Одновременно происходит резкое увеличение микроконформационной подвижности и флуктуаций структуры белка. Таким образом, по своим свойствам термодинамически стабильное промежуточное состояние близко к нативной вторичной структуре, но обладает флуктуирующей пространственной структурой (подробнее см. гл. X). С этой точки зрения фазовые переходы в белке могут быть обусловлены не разворачиванием белковой глобулы, а разрушением ее уникальной пространственной структуры. Фазовый переход совершается между более плотным (нативное) состоянием с сильным ван-дер-ваальсовым притяжением, но заторможенными боковыми группами и менее плотным (промежуточное) состоянием, где боковые группы «разморожены», а ван-дер-ваальсовые контакты разрушены. Так как боковые группы прикреплены к жесткому структурному каркасу глобулы, состоящему из  $\alpha$ - и  $\beta$ -участков, нарушение плотной упаковки в одном месте белка или белкового домена может произойти лишь при смещении или разрушении этого каркаса по всему объему глобулы. Таким образом, двум состояниям белка соответствуют и два различных объема компактной глобулы, а промежуточные между ними объемы термодинамически неустойчивы. Поэтому локальные нарушения плотной упаковки оказываются невозможными и разрушение нативной структуры белка является фазовым переходом первого рода (см. § 2 гл. VII).

## § 6. Сворачиваемость белка

В нормальных условиях любая природная аминокислотная последовательность обладает уникальной структурой. Она характеризуется минимумом конформационной энергии, который является достаточно «размытым» по пространственным координатам и относительно небольшим по глубине (IX, § 1). Он соответствует нескольким близким конформационным подсостоянием разделенными барьерами  $kT$ , между которыми белок совершает спонтанные флуктуационные переходы, оставаясь однако все время в пределах нативной конформации.

Еще в работах Анфинсена (1973) было показано, что пространственная структура белка способна к спонтанной самоорганизации. Денатурированная белковая глобула после возвращения к нормальным температурным условиям и удаления денатуранта из раствора снова за несколько минут восстанавливает исходную первоначальную нативную структуру, в которой каждый атом занимает свое определенное место.

При огромном количестве конформационных подсостояний в нативном белке ( $\sim 10^{100}$  для цепи из 100 остатков, если каждый из них имеет  $\sim 10$  конформаций) реальный процесс сворачивания должен носить в известной степени направленный характер. Он не может осуществляться путем последовательного перебора всех вариантов для нахождения минимума энергии, необходимого для стабилизации структуры.

В самом деле, при сворачивании белковой цепи происходит падение энтропии из-за роста упорядоченности и одновременное падение энергии вследствие образования контактов между сближающимися звеньями. Если свернутая структура стабильна, то в ней проигрыш за счет уменьшения энтропии ( $T\Delta S < 0$ ) должен быть скомпенсирован выигрышем ( $\Delta U < 0$ ) по энергии, так что в сумме

$$\Delta F = \Delta U - T\Delta S < 0 \quad (|\Delta U| > |T\Delta S|).$$

Допустим, что поиск этого энергетического минимума происходит путем перебора разных типов укладки. Тогда цепь в процессе сворачивания должна в начале случайным образом приблизиться к нативной конформации, потеряв почти всю энтропию, и только затем в результате образования соответствующих контактов произойдет выигрыш в энергии.

Незначительная величина энергетического барьера на этом пути  $\Delta G^*$  обусловлена тем, что уменьшение энтропии при фиксации одного звена компенсируется выигрышем энергии при установлении «нужного» контакта, который в дальнейшем не нарушается. Это резко снижает общее время последовательного сворачивания цепи.

Такой путь сворачивания белка обеспечивается в итоге тем, что глобулярная часть растущих зародышей белковой структуры не перестраивается по ходу сворачивания. Это предположение соответствует принципу согласования ближних и дальних взаимодействий в белке, лежащем в основе его кооперативных свойств (IX, § 2).

При последовательном сворачивании энтропия падает примерно пропорционально числу звеньев:  $n$ , фиксированных в финальной конформации с общим числом звеньев  $N$ , а энергия — пропорционально числу контактов между звеньями. Расчеты показывают, что максимальная разность между падением энтропии и выигрышем по энергии при фиксации  $n$  звеньев

$$\Delta S(n)/\Delta S(N) - \Delta E(n)/\Delta E(N)$$

оставляет всего 0,19, а время сворачивания белка при  $N = 58$  равно  $10^{58 \cdot 0,19}$  нс  $\sim 10^{11}$  нс  $\sim 1$  мин (Финкельштейн, 1996), что примерно соответствует опыту.

Быстрое сворачивание белка без задержки в пути в промежуточных квазистабильных ловушках возможно, если нативная структура более стабильна, чем клубок, а все остальные конкурирующие (рис. IX.36) с ней структуры менее стабильны.



Конечно, рассмотренные выше соображения носят феноменологический характер. Они не снимают необходимости изучения реальных путей быстрого сворачивания белка и природы соответствующих взаимодействий между аминокислотными остатками. Следует сразу сказать, что теоретические исследования, которые ведутся в настоящее время в этой области, еще не позволяют учесть роль конкретных

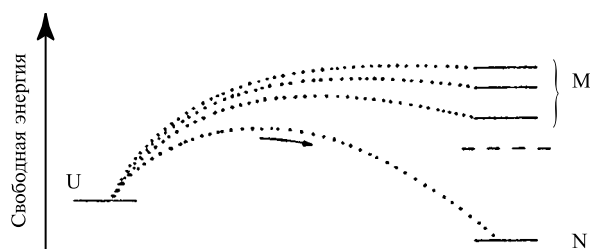


Рис. IX.36

Энергетические уровни различных состояний на пути сворачивания белка (объяснения в тексте) (по Финкельштейну, 1998)

остатков, характера последовательности в быстром свертывании цепи и стабилизации нативной структуры. В основном, внимание сосредоточено на следующих проблемах:

- общие кинетические и термодинамические закономерности сворачивания;
- различие в способности к сворачиванию у разных случайных последовательностей.

**Моделирование сворачиваемости.** В последние годы процесс сворачивания аминокислотных последовательностей интенсивно изучается методами компьютерного моделирования. Сворачиваемая цепь должна удовлетворять двум различным требованиям.

В термодинамическом отношении нативная конформация должна быть уникальной и обладать явно выраженным энергетическим минимумом (стабильностью) по сравнению с другими близкими конформациями.

В кинетическом отношении последовательность должна сворачиваться достаточно быстро и быть поэтому стабильной при не слишком низких температурах.

Иными словами, необходимы температуры достаточно высокие для обеспечения высокой скорости самого сворачивания и одновременно достаточно низкие, при которых достигнутое нативное состояние было бы стабилизировано. В этом случае процесс сворачивания будет очень медленным, так как на первом этапе потеря энтропии ( $-T\Delta S$ ) сопряжена с соответствующим ростом свободной энергии переходного состояния, образующегося по ходу сворачивания. От величины  $\Delta G^*$  в свою очередь экспоненциально зависит и время протекания процесса  $\tau \sim \exp(\Delta G^*/RT)$ . Именно это обстоятельство (падение энтропии до начала выигрыша энергии) и лежит в основе «парадокса Левинталя», согласно которому белковая цепь не может найти свою самую стабильную структуру за разумное время.

Время перехода одного звена из клубка в растущую глобулу  $\tau \sim 1$  нс, а падение энтропии на один аминокислотный остаток имеющий 10 конформаций, порядка  $R \ln 10$ . Тогда сворачивание в целом займет  $10^N$  нс. По оценке Левинталя время достижения стабильной структуры для белка с  $N = 58$  остатками составит  $\sim 10^N$  нс  $\sim 10^{58}$  нс  $\sim 10^{41}$  лет.

Между тем природные аминокислотные последовательности за время  $10^{-2} - 10^2$  с спонтанно сворачиваются в нативную глобулу, которая лишь ненамного ( $\sim 0,1$  ккал/моль в расчете на звено цепи) превосходит по стабильности развернутую форму (неглубокий энергетический минимум, гл. IX, § 1).

С точки зрения термодинамики для сворачивания необходима компенсация падения энтропии падением энергии на каждом этапе сворачивания (Финкельштейн, 1996).

Последовательный процесс сворачивания белковой структуры начинается с фиксации одного звена сразу же в его финальной конформации, т. е. начинается с формирования зародышей вторичной структуры. Затем в финальной конформации фиксируются и соседние по цепи звенья, в которой устанавливаются стабилизирующие структуру взаимодействия.

Большое разнообразие природных последовательностей белков ( $10^5$ ) и разброс значений параметров, характеризующих взаимодействие отдельных остатков, говорит о том, что статистические свойства глобулярных белков по существу совпадают со статистическими свойствами случайных последовательностей. Это позволяет использовать стохастический подход к моделированию сворачиваемости белка, а случайные последовательности могут служить объектом модельных исследований.

В гл. VII § 2 мы видели, что при определенных температурах в полимерах возможны переходы между состояниями клубок  $\rightleftharpoons$  глобула, зависящими от величины взаимодействия звеньев цепи. Исследование моделей полипептидных цепей показало, что гомополимерная последовательность может переходить в глобулярное компактное состояние с большей по сравнению с клубком плотностью звеньев. Однако только гетерогенные цепи, включающие разные по своей природе звенья, способны не только переходить в более плотное глобулярное состояние, но и принимать в нем определенную свернутую конформацию, соответствующую нативному состоянию белка.

Можно выделить три фазовые состояния полимерной цепи: клубок, глобула «расплавленная» и глобула «замерзшая» с небольшим числом конформаций. Между этими фазовыми состояниями возможны переходы при определенных температурах

$$\text{клубок} \xrightarrow{T_\theta} \text{глобула} \xrightarrow{T_c} \text{свернутая конформация}.$$

Температура  $T_\theta$  определяет границу уплотнения клубка до глобулярного состояния, которому однако соответствует много конформаций.  $\square$

При  $T = T_c = T_\theta$  происходит «замерзание» расплавленной глобулы, что называют «стеклованием» по аналогии с переохлажденной жидкостью, которая фиксируется в одном из многих неравновесных состояний, не достигнув еще своей равновесной кристаллической структуры.

На этом этапе уже может образоваться нативная структура, которая имеет только одну определенную конформацию. Но именно здесь возникает противоречие между кинетическими и термодинамическими требованиями к процессу сворачивания. Если истинная температура сворачивания  $T_f$  ниже точки замерзания  $T_f < T_c$ , то стабильность конечной конформации обеспечена, однако, время сворачивания будет очень большим. С другой стороны, при  $T_f > T_c$  сворачивание будет быстрым, однако оно приведет к структуре не стабильной при этих температурах.

Выход из этого противоречия, по-видимому, обеспечивается характером энергетического профиля сворачивания, пример которого приведен на рис. IX.37. Здесь характерен крутой наклон энергетического профиля и более глубокое расположение уровня нативной конформации по отношению к ближайшим к ней состояниям. Именно эти обстоятельства определяют не только большую скорость сворачивания, но и стабильность нативной структуры по отношению к тепловым флуктуациям при  $T_f > T_c$ .

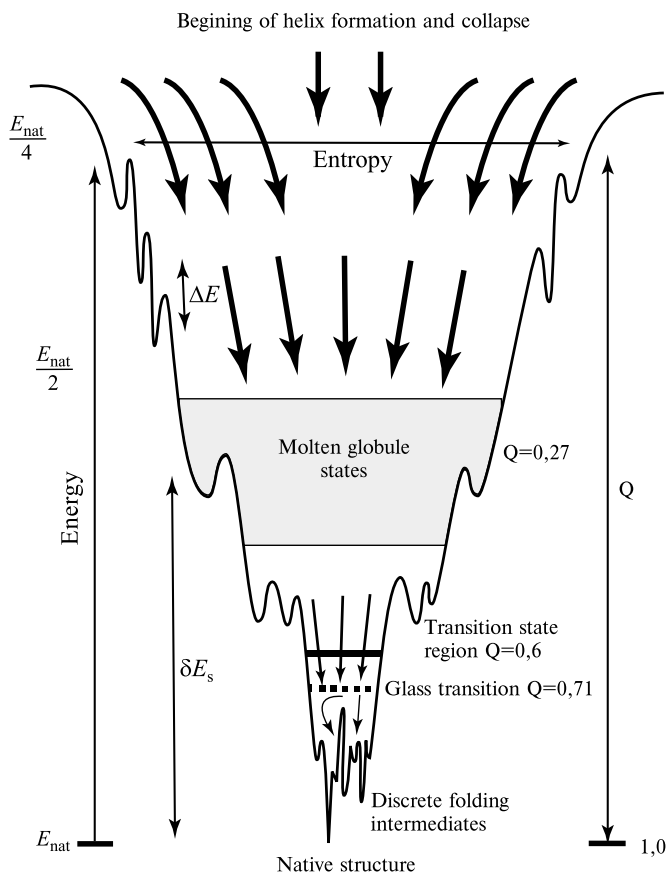


Рис. IX.37

Схема «воронки свертывания» для быстрого сворачивания спирального белка из 60 аминокислотных остатков (по Wolynes, Onuchic, Thirumalai)

Ширина воронки пропорциональна энтропии, а глубина — энергии. Доля независимых контактов точно определяется,  $Q$ , и устанавливается для каждого из состояний

Развиваемые теоретические представления посвящены анализу энергетического профиля процесса сворачивания (Онучик, Тирумала, 1990–1996) и моделированию сворачивания стохастическими методами (Карплюс, Шахнович, 1990–1996).

Рассмотрим результаты по моделированию динамики сворачивания случайных последовательностей.

В работах Карплюса, Шахновича и соавторов использовалась модель решетки, в узлах которой располагались взаимодействующие друг с другом бусины — аминокислотные остатки (рис. IX.38). В модели изучалась кубическая решетка  $(3 \times 3 \times 3)$  с 27-ю элементами — фрагментами цепи. Эта модель может имитировать состояние расплавленной глобулы, поскольку в ней отсутствуют боковые цепи при сохранении

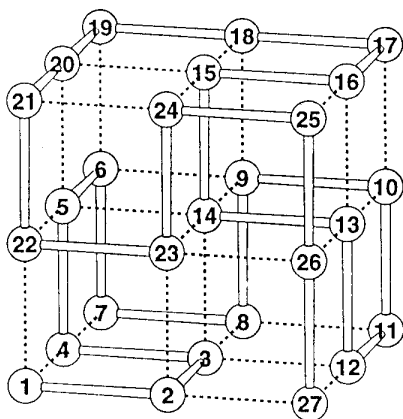


Рис. IX.38

Решетчатая модель скручивания белка: пример компактной структуры цепи из 27 мономеров (пронумерованные кружки) с 28 контактами (пунктирные линии) (по Sali, Shakhnovich, Karplus)

компактного «скелета». Предполагается, что энергия такой полимерной цепи зависит только от взаимодействия ближайших фрагментов в соседних узлах решетки. Общая величина энергии взаимодействия выразится как

$$E = \sum_{i < j} B_{ij} \Delta(r_i - r_j), \quad (\text{IX.6.1})$$

где  $B_{ij}$  — энергия взаимодействия между мономерами  $i$  и  $j$ . Если они находятся в соседних узлах решетки, то  $\Delta(r_i - r_j) = 1$ , в противном случае  $\Delta(r_i - r_j) = 0$ .

Можно ввести некую среднюю величину взаимодействия  $B_0$  между мономерами. При  $B_0 < 0$  (гидрофобное сжатие) происходит компактизация клубка и переход в глобулу. В рассматриваемых моделях энергия взаимодействий отдельных мономеров  $B_{ij}$  отклоняется от  $B_0$  случайным образом. Стохастический характер взаимодействий в цепи случайных последовательностей описывается гауссовым распределением

$$P(B_{ij}) = \frac{1}{2\pi\sigma_B} \exp \left[ -\frac{1}{2} \frac{B_{ij} - B_0}{\sigma_B} \right], \quad (\text{IX.6.2})$$

где  $\sigma_B$  — стандартное отклонение величин  $B_{ij}$  от среднего значения  $B_0$ . Это характеризует степень гетерогенности цепи. В гомополимере, где все энергии взаимодействий одинаковы  $\sigma_B = 0$ .

В гетерополимерной цепи отдельные члены  $B_{ij}$  в (IX.6.1) могут быть разных знаков. Поэтому даже при  $B_0 > 0$ , в гетерополимерной цепи возможно образование глобулярной структуры за счет вклада отрицательных членов  $B_{ij} < 0$ . Критическая температура  $T_c$  перехода в глобулярное состояние определяется средним числом контактов  $n$  одного мономера и количеством конформаций на один мономер  $\gamma$ :

$$T_c = \frac{\sigma_B}{2R_B} \frac{\sqrt{n}}{\sqrt{\ln \gamma}}.$$

В такой системе, вводя общее условие компактности (гидрофобности) глобулы ( $B_0 < 0$ ) возможно определить минимальную энергию, соответствующую нативному состоянию. Затем генерируется случайный набор значений энергий взаимодействий  $B_{ij}$  из гауссовского распределения (IX.6.2) с помощью генератора случайных чисел. Тем самым получается набор различных случайных последовательностей. При дальнейшем моделировании происходит сравнение способностей этих случайных последовательностей сворачиваться в низкоэнергетическую нативную конформацию. Необходимо выбрать из них те, которые могут это делать относительно быстро (за ограниченное число шагов), а затем сравнить характер энергетических спектров различных последовательностей вблизи нативного состояния.  $\square$

Методом Монте-Карло моделируется последовательное сворачивание цепи путем перебора случайных шагов, которые фиксируют звенья на тех или иных узлах решетки. Выбираются те шаги, которые уменьшают общую энергию системы. Процесс продолжается до достижения заданного энергетического минимума конечной нативной конформации. Наибольшее число возможных конформаций в модели  $\sim 10^{16}$ , однако, как оказалось, некоторые избранные последовательности могут найти нативную конформацию за значительно меньшее число шагов  $\sim 10^7$ . Это говорит о преодолении в этих случаях парадокса Левинталя в результате «движения» по «выделенному» пути сворачивания.

Максимальное количество шагов в расчетах по методу Монте-Карло было равно  $50 \cdot 10^6$ , поскольку обычно сворачивание достигалось за  $10 \cdot 10^6$  шагов. Если исходная последовательность не успевала достичь минимума по энергии за это число шагов, то она считалась не способной к сворачиванию. На рис. IX.39 показано, как на траектории сворачивания меняется энергия  $\epsilon$  сворачиваемой последовательности, число образуемых контактов  $N_c$  между мономерами, доля контактов  $Q_0$  из общего числа  $N_c$ , которые совпадают с нужными контактами в нативной конформации. Как видно в области  $10^4$  шагов Монте-Карло наблюдается резкий переход к компактной глобуле. На этой стадии из  $\sim 10^{16}$  конформаций клубка остается только  $\sim 10^{10}$ . Этот переход соответствует компактизации клубка  $T < T_\theta$  и переходу к расплавленной глобуле. Затем (при  $T < T_c$ ) идет «стеклование» и поиск одного из  $10^3$  переходных состояний. Переход к нативной конформации быстро происходит из любого переходного состояния и требует  $\sim 10^5$  шагов.

На рис. IX.40 приведены энергетические спектры 10 случайных последовательностей, из которых 4 сворачиваются. Спектры включают энергетические уровни 400 низких по энергии конформаций. У сворачиваемых последовательностей уровень энергии нативной конформации занимает самое низкое положение. Принципиально важно, что он отделен от ближайшего переходного состояния

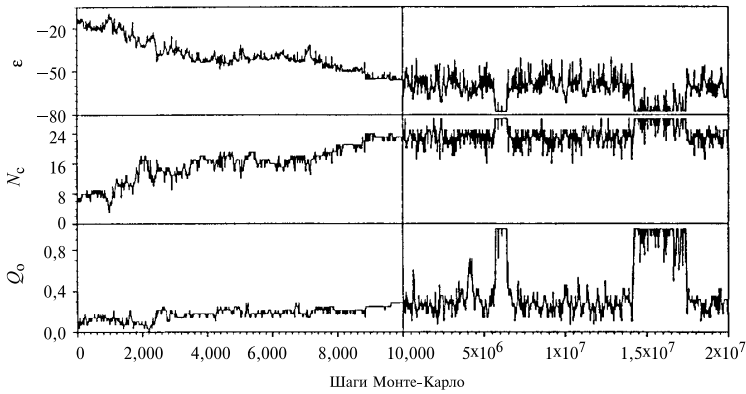


Рис. IX.39  
Типичная траектория для скручивания случайной последовательности (по Sali, Shakhnovich, Karplus)

$\epsilon$  — энергия (в единицах  $k_B T$ , где  $k_B$  — константа Больцмана),  $N_c$  — число контактов,  $Q_0$  — доля нативных контактов в общем числе  $N_c$ . Мгновенные значения этих величин регистрировались каждые 10 шагов Монте-Карло.

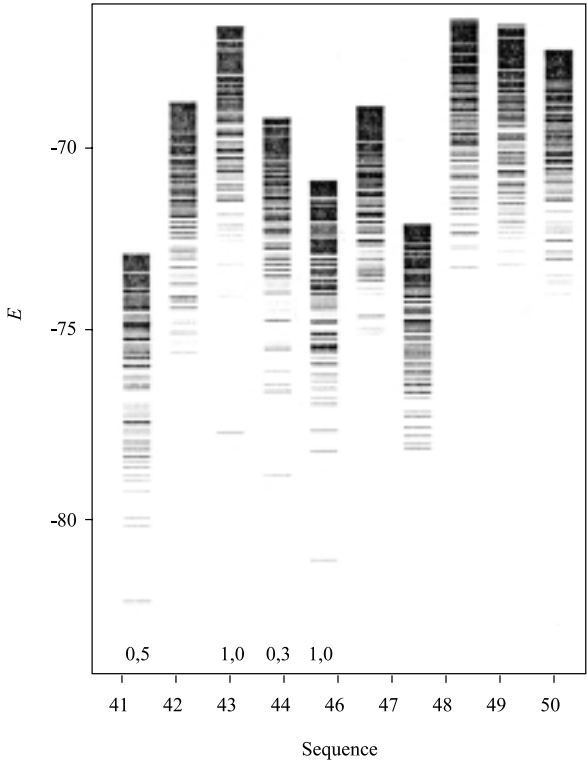


Рис. IX.40  
Энергетический спектр для 10 сворачивающихся и не сворачивающихся последовательностей (по Sali, Shakhnovich, Karplus)

Нативное состояние соответствует нижней полоске. Числа под спектрами показывают «тенденцию» к сворачиванию, соответствующих последовательностей. Если число отсутствует, «тенденция» равна 0. Последовательность сворачивается, если ее нативная конформация находится через  $50 \times 10^6$  шагов Монте-Карло.

энергетическим зазором  $\Delta E$ , который отсутствует в спектрах несворачиваемых последовательностей.

Мы уже говорили, что температурные условия стабильности нативного состояния и быстроты его достижения противоречат друг другу.

Как видно из рис. IX.40, быстро образующаяся при относительно высоких температурах нативная конформация будет оставаться при этом тем не менее устойчивой из-за энергетического зазора  $\Delta E$ , отделяющего ее от ближайших состояний (ср. рис. IX.37). Таким образом, сворачиваемые последовательности обладают таким энергетическим спектром, который позволяет им сворачиваться достаточно быстро при  $T_f > T_c$  с сохранением устойчивости нативного состояния в этих условиях. Все случайные последовательности стабильные лишь при  $T < T_c$  не относятся к сворачиваемым. Если сворачивание происходит при  $T_f \simeq T_c$ , то около 1% всех последовательностей имеет величину зазора в несколько  $k_B T$ .

Стохастический подход к выяснению механизмов самоорганизации белковой глобулы позволяет прояснить некоторые закономерности этого процесса. Однако для выяснения реальных путей сворачивания конкретных белков необходимо изучение их динамического поведения и строения гиперповерхности потенциальной энергии макромолекулы. Именно особенности строения гиперповерхности потенциальной энергии в конфигурационном пространстве определяют направленность движения системы в процессе сворачивания к нативной конформации.

На пути решения этой проблемы лежат большие вычислительные трудности построения гиперповерхностей уровней конформационной энергии макромолекулы. Результаты, получаемые при построении участков гиперповерхностей уровней вблизи локальных минимумов энергии даже для простых макромолекул, показывают наличие анизотропных областей, характеризующихся сильно вытянутыми в определенных направлениях энергетическими доменами.

▽ В последнее время появились новые результаты модельных экспериментов, которые указывают на роль структурных факторов в быстром сворачивании. Было показано, что в основе механизма быстрого сворачивания лежит образование уже на ранних этапах «ядра сворачивания» с участием определенных аминокислот. Такое ядро сворачивания образуется в быстро сворачивающихся и отсутствует в медленно сворачивающихся последовательностях. Ускорение сворачивания сопровождается стабилизацией ядра за счет внутренних контактов (Shakhnovich E., Abkevich V., Ptysyn O., 1996). Оказалось, что аминокислоты, принадлежащие ядру сворачивания, сохраняются в семействе родственных быстро сворачиваемых последовательностей (Shakhnovich E., 1998). Этот вывод был подтвержден на примере семейства структурноподобных белков, которые сворачиваются со связыванием монофосфата (Schulz G., 1992). Так в белке CheY наиболее «консервативны» четыре аминокислоты (*асп* 12, *асп* 57, *мет* 17, *ала* 88), которые сохраняют свое положение в ядре сворачивания в семействе родственных быстро сворачиваемых последовательностей. Интересно, что остатки, входящие в состав ядра сворачивания белка CheY, одновременно выполняют и важную роль в его функционировании при связывании фосфата. Очевидно, что в данном случае ядро сворачивания является также местом локализации активного центра полностью сформированного белка. □

В образовании ядра сворачивания важную роль играют гидрофобные взаимо-

действия. В модельных экспериментах было показано, что ядро обладает плотно упакованной и геометрически регулярной структурой (Wingreen N., Tong C., 1998).

Таким образом в белке существует небольшое число определенных положений аминокислотных остатков, которые ответственны за быстрое сворачивание. В процессе эволюции быстрота сворачивания могла быть следствием включения аминокислот в эти позиции, стабилизирующие ядро.

Итак, мы рассмотрели своеобразные физические факторы кооперативного процесса быстрого сворачивания, приводящего к достижению стабильной нативной конформации.

Существует точка зрения, что природные белки и соответствующие им аминокислотные последовательности обладают особыми свойствами, которые и обеспечивают уникальный характер их сворачивания. Это значит, что отобранные в эволюции природные белки, число которых ( $\sim 10^5$ ) намного меньше их общего возможного числа, не являются результатом «запоминания случайного выбора». Скорее, они действительно обладают некоторыми фундаментальными свойствами, которые их отличают от других неприродных последовательностей.

Сложность этой проблемы иллюстрируется данными о том, как происходит сворачивание белка в живой клетке. Именно здесь видно, что реализации рассмотренных выше физических закономерностей сворачивания происходит способом, отличным от такового *in vitro*. В самом деле, в клетке микроокружение полипептидной цепи включает рибосомальные структуры, ферменты, белки шапероны и другие факторы, отсутствующие в растворе. Векторный характер синтеза пептида от N- к C-концу приводит к тому, что сворачивание начинается уже на рибосоме в процессе трансляции немедленно вслед за появлением последовательности аминокислот. Связь C-конца с рибосомой обеспечивает возможность формирования  $\alpha$ -спиралей и влияет на скорость образования третичной структуры (Spirin A., 1986). Формирование нативной структуры белка в клетке происходит намного быстрее, чем ренатурация белков в растворе. Все это приводит к выводу о том, что сворачивание последовательности в живой клетке происходит не из состояния стохастического клубка, как при ренатурации в растворе, а осуществляется еще на рибосоме без выхода цепочки в окружающую рибосому среду, т. е. котрансляционным способом.

В работах Спирина А. С. были получены экспериментальные доказательства этого положения. Основная идея состоит в том, что активность белка проявляется лишь тогда, когда сформируются участки в белковой структуре, ответственные за функционирование белка.

Поэтому способность белка выполнять биологические функции является лучшим критерием образования его нативной структуры.

Этот подход использовался для изучения сворачивания люциферазы — белка светлячков и глобина. Было показано, что активность люциферазы по окислению люциферина кислородом в присутствии АТФ и  $Mg^{2+}$  проявляется практически сразу же после ухода молекулы люциферазы с рибосомы.

В противоположность этому ренатурация денатурированной люциферазы с восстановлением ферментативной активности занимала намного большее время ( $\sim 15$  минут). Более того, с помощью специально разработанных методов удалось показать, что вновь синтезированный белок люциферазы становится ферментативно активным уже на рибосоме, когда на его C-конце появляется не менее 26–27



аминокислотных остатков (Spirin A., 1995). Таким образом, сворачивание люциферазы в активную конформацию происходит уже на рибосоме, т. е. котрансляционным образом без ухода полипептида с рибосомы.

Аналогичный вывод был сделан и в отношении глобина, который способен связывать гем на рибосоме, для чего в молекуле глобина должна была сформироваться определенная структура. Было показано с помощью радиоактивных меток ( $^{35}\text{S}$ ,  $^3\text{H}$ ), что гем образует комплекс с глобином, находящимся на рибосоме, если *N*-конец глобина содержит не менее 86 аминокислотных остатков. Отсюда ясно, что цепь глобина уже на рибосоме свертывается в активную конформацию, обеспечивающую ее взаимодействие с группами гема (Spirin A., 1997).

Таким образом, вновь синтезированная последовательность начинает сворачиваться в нативную конформацию, не покидая рибосому, т. е. котрансляционно.

По-видимому, в клетке существуют специальные механизмы (котрансляционное сворачивание, шапероны), которые в полной мере обеспечивают проявление рассмотренных выше физических принципов направленного кооперативного сворачивания белковой последовательности в стабильную нативную конформацию.

Возвращаясь к принципу Анфинсена, следует признать, что он справедлив в отношении того, что первичная структура в ряде белков определяет характер стабильной вторичной и третичной конформации белка. Однако, этот принцип не может означать тождественности конкретных путей сворачивания белка в клетке и его ренатурации в растворе.