



УДК 577.156.3.02 : 541.63

ПОИСК ПЛОТНО УПАКОВАННЫХ СТРУКТУР ПАР α -СПИРАЛЕЙII. ШПИЛЬКИ $A - E$ И $G - H$ МИОГЛОБИНА КАШАЛОТА*Рашид А. А.***Институт белка Академии наук СССР,
г. Пушкино, Московская область*

Предложенным в предыдущем сообщении методом проведен поиск плотно упакованных структур α -спиральных шпилек $A - E$ и $G - H$ молекулы миоглобина кашалота. Полученные структуры использовались в качестве результата первого приближения для дальнейшего их улучшения и исследования с помощью пространственных моделей. Показано, что конформации, очень близкие к нативным шпилькам $A - E$ и $G - H$, находятся среди небольшого числа найденных наиболее выгодных структур. Для рассмотренных шпилек выяснен механизм локализации N-концов α -спиралей за счет встраивания ближайших к N-концам массивных гидрофобных групп в гидрофобные карманы шпилек.

В предыдущей статье этой серии [1] был предложен графический метод поиска окрестностей всех плотно упакованных конформаций для пар α -спиралей. В настоящей статье этот метод с последующим стереохимическим рассмотрением полученных конформаций при помощи пространственных моделей используется для исследования упаковок предсказываемых теорией локальных взаимодействий [2] пар спиралей (от пролина до пролина) или их частей, участвующих в образовании шпилек $A - E$ и $G - H$ в миоглобине кашалота, и проверки некоторых предположений о процессе самоорганизации миоглобина кашалота, выдвинутых ранее [3, 4]. Рассматривались только упаковки спиралей антипараллельного типа.

На первом этапе проводился полный перебор окрестностей всех плотно упакованных конформаций методом разверток [1] и отбрасывались упаковки с внутренними полостями, способными вместить молекулу воды, которые можно было идентифицировать на этой стадии. Проверка плотности упаковок и степени дегидратации оставшихся конформаций осуществлялась путем сборки пространственных моделей в соответствии с поворотными изомерами боковых групп и их взаимным расположением, отобраным на развертках. При необходимости поворотные изомеры боковых групп α -спиралей [1] деформировались в пределах стерически разрешенных областей так, чтобы сделать упаковку максимально плотной. Торсионные потенциалы не учитывались.

Для полученных конформаций проводилась грубая оценка степени дегидратации массивных гидрофобных боковых групп спиралей. Учитывались также неспиральные массивные гидрофобные группы, близкие

* Настоящий адрес автора: Институт биологической физики АН СССР, г. Пушкино, Московская область.

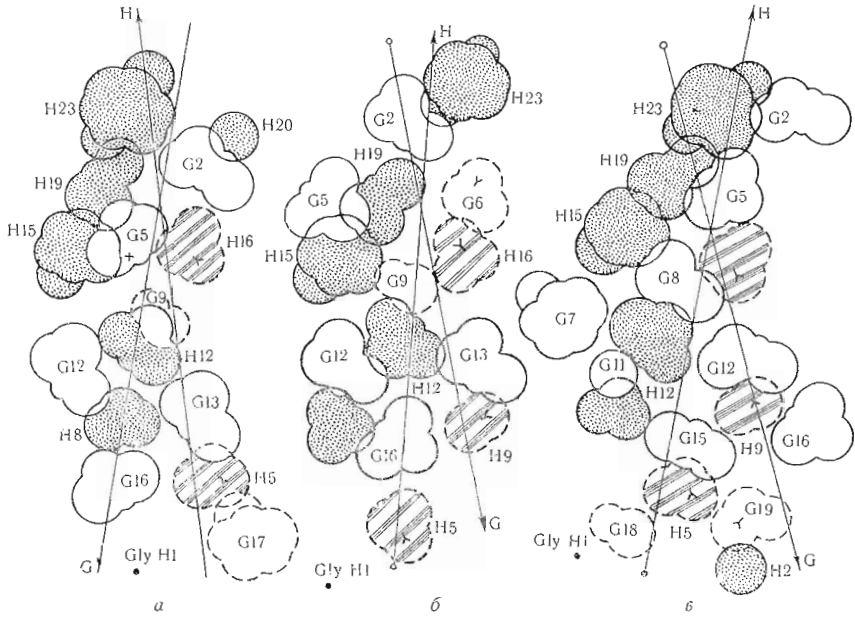


Рис. 1. Развертки трех наиболее выгодных плотно упакованных конформаций шпильки $G-H$. Сплошной линией обведены гидрофобные группы, пунктирной — гидрофильные. Группы на одной из спиралей заштрихованы

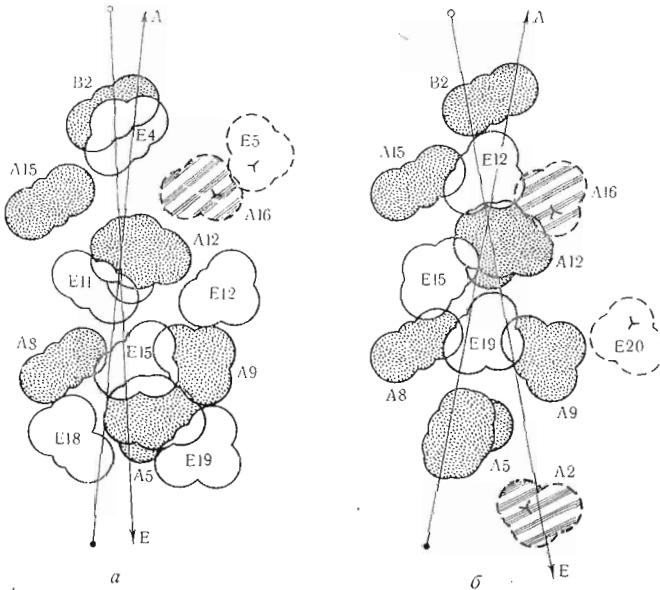


Рис. 2. Развертки двух наиболее выгодных плотно упакованных конформаций шпильки $A-E$

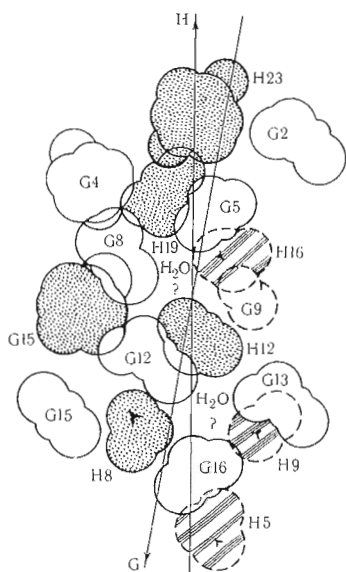


Рис. 3. Пример упаковки шпильки $G-H$ с внутренними полостями, способными вместить молекулу воды

вание показало, что учет указанных взаимодействий существенно не влияет на результат.

На рис. 1 и 2 в виде разверток изображены окрестности трех плотно упакованных структур шпильки $G-H$ и двух структур шпильки $A-E$, наиболее выгодных по степени дегидратации массивных гидрофобных групп. Каждая из конформаций является представителем некоторого набора конформаций, различающихся относительными смещениями спиралей до 2 \AA и поворотными изомерами отдельных боковых групп. Других окрестностей плотно упакованных конформаций, сравнимых по степени дегидратации с приведенными, найдено не было. Пример отбракованной (отброшенной как негодная) конформации шпильки $G-H$ приведен на рис. 3. Всего для шпильки $G-H$ было обследовано ~ 40 таких наборов различных конформаций, а для шпильки $A-E$ — около 15. Изображения боковых групп на развертках соответствуют углам χ идеальных поворотных изомеров, использованных в работе [1].

При улучшении плотности упаковки на пространственных моделях углы χ иногда довольно сильно отклонялись от их идеальных значений. Так, например, остаток Leu G5 (на рис. 1, *a* помечен крестиком), находящийся в конформации с $\chi_1 = 170^\circ$ и $\chi_2 = 60^\circ$ *, после улучшения упаковки на моделях имеет $\chi_2 \approx -10^\circ$. Естественно также, что конечные конформации «деформированных» при улучшении упаковки поворотных изомеров сильно зависят от исходной идеальной конформации, как при уточнении координат атомов [7] в рентгеноструктурном анализе или при локальной минимизации [8]. Так, если в упаковке типа GHI (рис. 1, *a*) исходная конформация остатка Leu G5 — $\chi_1 = 290^\circ$, $\chi_2 = 180^\circ$, то конечная находится в области $\chi_1 = 230^\circ$, $\chi_2 = 160^\circ$. Возможно, учет торсионных потенциалов привел бы к иному компромиссу между значениями χ_1 и χ_2 . В сходном положении находится в конформации GHI остаток Leu G16.

Таким образом, набор поворотных изомеров (рис. 1, 2) следует рассматривать как один из исходных для улучшения при помощи пространственных моделей или локальной минимизации, с использованием методов кон-

к концам спиралей. Сравнение выгодности различных структур осуществлялось по суммарной степени дегидратации массивных гидрофобных групп, числу водородных связей в основной цепи и между основной цепью и боковыми цепями остатков Ser, Thr, Asp, Asn, Tyr, His. При одинаковом числе водородных связей структуры, отличавшиеся от наиболее дегидратированной более чем на одну группу, отбрасывались как невыгодные. Кроме того, отбрасывалась любая конформация с недоступным для воды свободным донором или акцептором водородной связи (ср. [5]), а также обнаруженные на этом этапе конформации с внутренними полостями, способными вместить молекулу воды. При сравнении выгодности различных структур не учитывались дегидратация неполярной части полярных боковых групп и водородные связи между боковыми цепями или между длинными боковыми цепями (остатки Gln, Glu, Arg, Lys) и основной цепью. Считалось, что эти взаимодействия компенсируют энтропийные потери при фиксации боковой цепи полярного радикала. Дополнительное исследование указанных взаимодействий существенно не

* Отсчет углов проводится в соответствии с [6].

формационных расчетов [8]. Значения исходных углов χ_1 и χ_2 в конформациях GH_I , GH_{II} , GH_{III} , AE_I , AE_{II} (рис. 1, 2), углов χ для этих же групп в нативной структуре миоглобина кашалота, степени дегидратации массивных гидрофобных групп в каждой из улучшенных конформаций, а также суммарная степень дегидратации в каждой из них (т. е. число массивных гидрофобных групп, полная дегидратация которых эквивалентна дегидратации в соответствующей конформации) приведены в табл. 1 и 2. Никаких различий между группами не проводится.

Более детальные сведения о конформациях GH_I , GH_{II} , GH_{III} , улучшенных при помощи пространственных моделей, содержатся в матрицах контактов [9], показанных на рис. 4—6. Отмечались только контакты с расстоянием до 5 Å между неводородными атомами боковых цепей, а также те водородные связи между боковыми цепями, которые приводили к дегидратации массивных гидрофобных групп.

Игнорирование массивных гидрофобных групп, не входящих в экспериментально локализованные спирали [10, 11], но находящихся вблизи их концов, в ряде случаев серьезно влияет на результаты. Например, неучет остатка Leu NA_2 приводит к обрасыванию конформации AE_{II} (табл. 2), неучет остатка Phe GH_5 — к отбрасыванию структуры GH_I , способной вместить несколько молекул воды в глубокой узкой впадине между С-концом G -спирали и N-концом H -спирали. Встраивание остатка Phe GH_5 в эту полость вызывает его полную дегидратацию, небольшое выталкивание из полости остатка Leu G_{16} и одновременно локализацию N-конца H -спирали: первую водородную связь с кислородом основной цепи завязывает либо остаток Gly H_1 , либо за счет небольшого вытягивания остатка Phe из полости (кармана) остаток Phe GH_5 .

Для проверки достаточности энергетического выигрыша за счет встраивания остатка Phe GH_5 в карман для локализации N-конца H -спирали к конформациям GH_I и GH_{II} пристраивался участок GH и в полной структуре каждой пилочки завязывалось максимальное число водородных связей при частично или полностью дегидратированном остатке Phe GH_5 и при остатке Phe GH_5 , полностью экспонированном на воду (рис. 7).

В конформации GH_I в остаток Phe GH_5 полностью экспонирован на воду, дегидратация остальных групп практически не меняется, а карман лишь слегка уменьшается за счет сближения боковых цепей остатков Leu G_{16} и Gln H_5 . В конформациях GH_{IIb} и GH_{IIc} остаток Phe GH_5 также полностью экспонирован на воду и дополнительно экспонируется $1/4$ Met H_8 . Таким образом, полная потеря в дегидратации равна $3/4$. В GH_I потери в дегидратации и плотной упаковке компенсируются $\sim 1,5$ водородными связями (вторая, вилочковая связь, завязываемая связанным донором или акцептором водородной связи, считалась по энергии эквивалентной половине первой водородной связи), а в GH_{II} — одной дополнительной водородной связью. При продлении системы водородных связей на N-конце A -спирали потери дегидратации от 1 до 1,5 гидрофобных групп (см. табл. 2) компенсируются выигрышем лишь 1,5 дополнительных водородных связей (место кислорода основной цепи остатка NA_2 в системе водородных связей конформации AE_{II} занимает кислород боковой группы Ser A_1). Ситуация, близкая к подобной, наблюдается на N-конце G -спирали, где потеря дегидратации примерно одной массивной группы из-за разрыва гидрофобных контактов остатком Ile FG_5 может быть скомпенсирована выигрышем не более чем 1,5 водородных связей из-за наличия остатка Pro в положении G_I . В этом случае, как и для AE_{II} , в конформациях GH_I , GH_{II} , GH_{III} кислород основной цепи остатка FG_5 , способный участвовать в продолжении системы водородных связей G -спирали, завязывает водородную связь с атомами боковых групп остатков Tyr H_{23} или Lys H_{22} .

В конформации GH_I 30 из 35 отмеченных нативных контактов совпадают с теоретически полученными, в то время как в конформациях GH_{II} и

Таблица 2

Сравнение нативной и наиболее выгодных конформаций шпильки А — Е миоглобина кашалота.

Конформации	Боковой радикал													Суммарная деградация					
	NA2 Leu	A2 Glu	A5 Trp	A8 Val	A9 Leu	A11 Val	A12 Trp	A15 Val	A15 Glu	B2 Val	E1 Leu	E5 Lys	E11 Val		E12 Leu	E15 Leu	E18 Ile	E19 Leu	E20 Lys
ИИ			170, 270	170	290, 180	170	170, 270	170	170, 180	170	170	170, 60	170	290, 180	290, 180	290, 180	290, 180		
АЕИ	1/4		3/4	1/2	1/4		3/4	1/4		1/4		1/2	1/2	1/4	3/4	3/4	1/4		6
	(0)		(1/2)													(1/2)			(5/4)
Native	1/2		3/4	1/2	1/4		3/4	1/2		1/4		1/2	1/2	1/2	1/2	(1/4)	3/4		5 1/2
	(0)	251, 8	283, 117	174 (1/4)	284, 164	218	179, 271	210	225, 147	173	230, 103	236, 176	177	264, 231	185, 65	287, 169	280, 193	314, 177	(4)

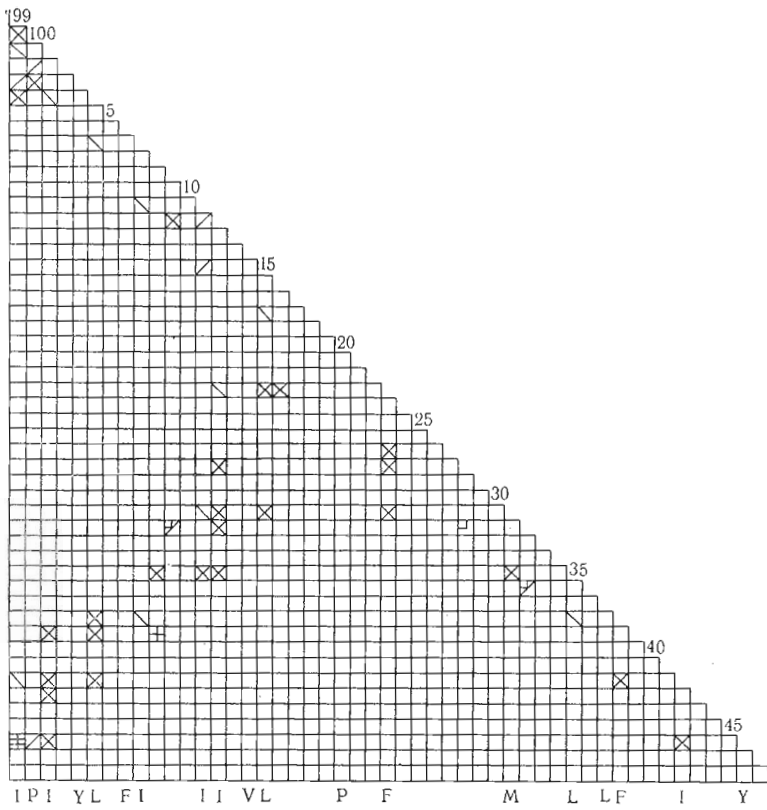


Рис. 4

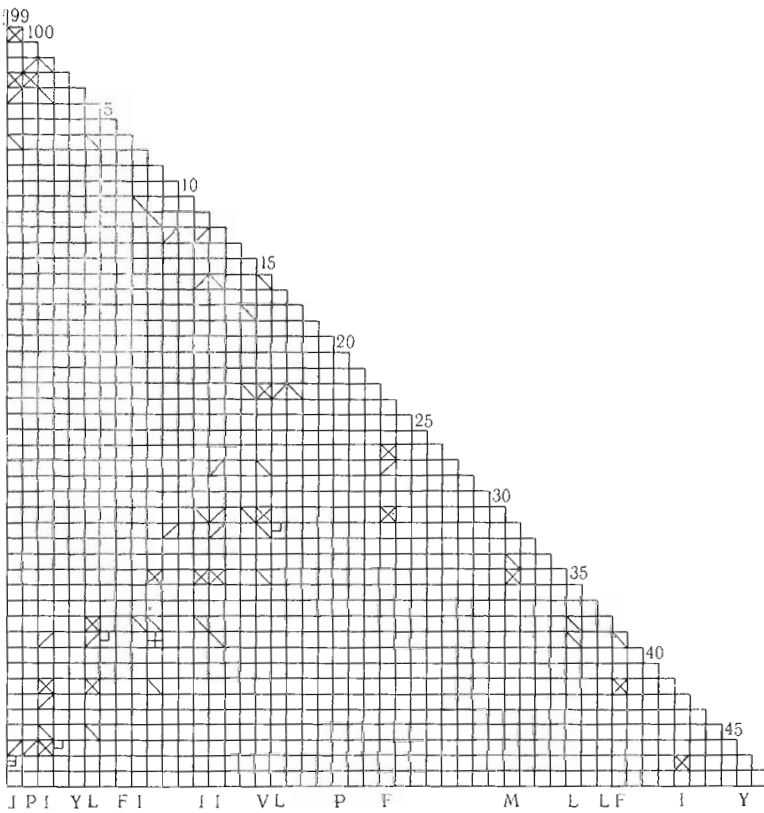


Рис. 5

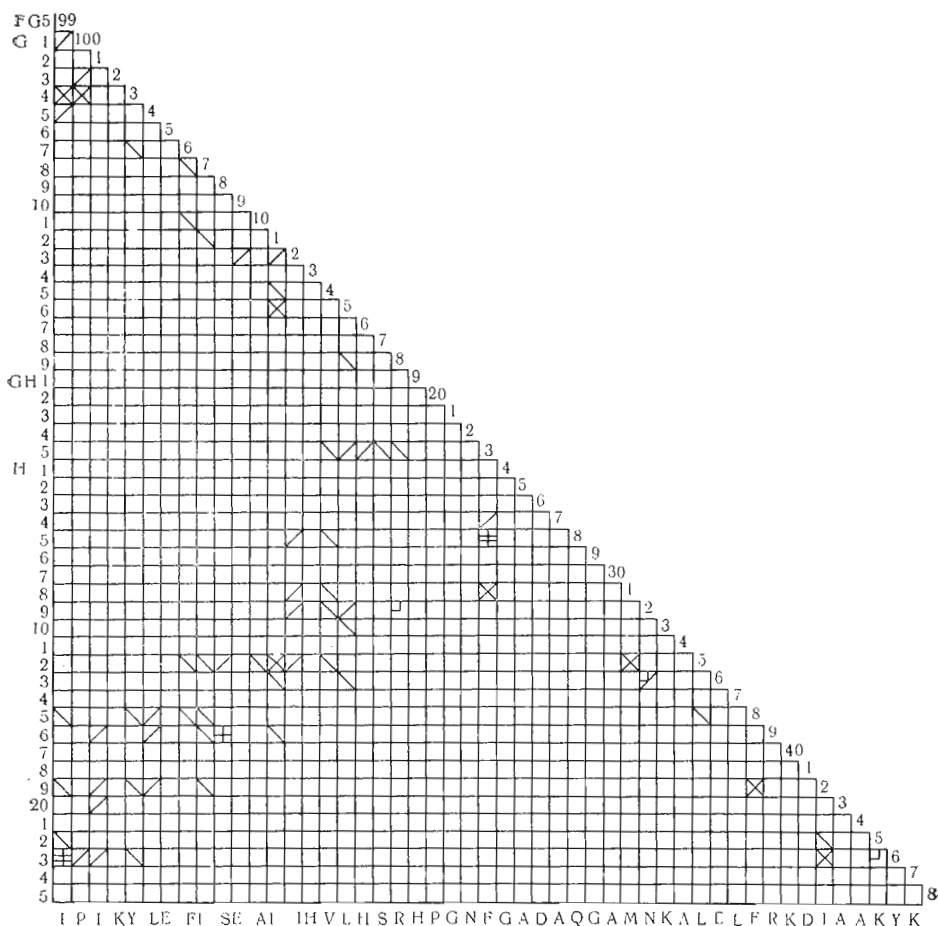


Рис. 6

Рис. 4. Треугольная матрица контактов, дегидратирующих массивные гидрофобные группы в конформации *GH*. \ — наличие контакта в собранной модели; / — наличие контакта в нативной структуре; | — водородная связь между боковыми цепями в модели; — — водородная связь «основная цепь — боковая цепь» в модели; ⊕ — та же связь в нативной структуре; X — водородная связь между боковыми группами в модели, но только контакт между полярными группами в нативной структуре

Рис. 5. Матрица контактов конформации *GHII*

Рис. 6. Матрица контактов конформации *GHIII*

GHIII теоретически их получено лишь 18 и 12 из 34 (рис. 4—6). Наличие избыточного числа контактов в собранной модели (рис. 8) может объясняться чувствительностью матрицы нативных контактов к изменению расстояния между атомами, при котором контакт засчитывается, недостаточной точностью построения моделей ($\sim 2 \text{ \AA}$) и ошибками в визуальном определении контактов в модели. В области дальних взаимодействий в плотной упаковке *GHI* только 5 из 27 заполненных клеток матрицы указывают на несовпадение теоретических и нативных контактов. Во всяком случае ясно, что конформация *GHI* имеет все основные черты нативной шпильки *G — H* или по крайней мере намного ближе к ней, чем *GHII* и *GHIII*. Аналогичные рассуждения свидетельствуют о том, что близкой к нативной является конформация *AEII*.

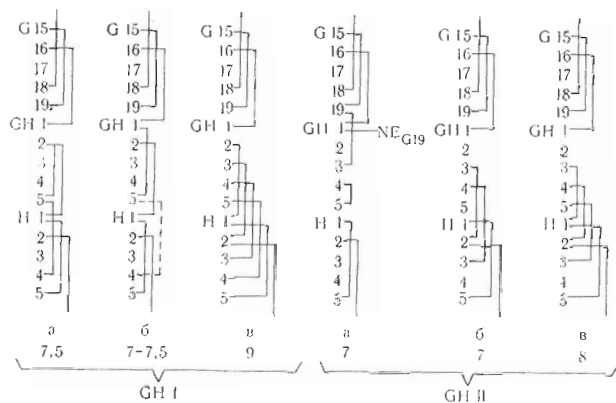


Рис. 7. Максимальные системы водородных связей в петле *GH*. *GH1a*, *GH1б*, *GH1в* — водородные связи в петле с дегидратированным Phe *GH5*; *GH1г*, *GH1д*, *GH1е* — с Phe *GH5*, полностью экспонированным на воду. Указано число водородных связей в приведенном участке для каждой из конформаций

Сравнение углов χ в конформациях *GH1* и *AEII* и соответствующих нативных структурах показывает очень хорошее согласие между экспериментальными и теоретическими значениями для *A — E* и худшее — для *G — H*. Последнее объясняется, видимо, выбором исходного варианта конформации *GH1*. Анализ привел к заключению, что конформация остатка Met *H8* легко может быть приближена к экспериментальной после встраивания остатка Phe *GH5* в карман. Конформации остатков Leu *G5* и Leu *G16* обсуждались выше. Расхождения могут происходить и вследствие ошибок рентгеноструктурного анализа, так как углы χ соответствуют *R*-фактору для миоглобина $\sim 0,4$ [12]. Возможно также, что те или иные варианты конформации одной и той же группы отбираются на более поздних стадиях самосборки, не различаясь в пределах шпильки по выгодности.

Ранее в работах [3, 4] было высказано предположение, что сборка миоглобина начинается с формирования «центров кристаллизации» из соседних по цепи участков, причем растут только наиболее выгодные из них. Следовательно, выросшая нативная структура таких центров также должна быть одной из наиболее выгодных структур соответствующей части белка. Результаты, полученные для шпильки *G — H*, представляющей один из таких центров, подтверждают это предположение в рамках использованных в настоящей работе методов. Результаты для шпильки *A — E* нельзя истолковывать таким же образом, так как соответствующие спирали не являются соседними по цепи. Обсуждение этого результата будет проведено в работе о самоорганизации всей молекулы. Тем не менее можно предположить, что формирование шпильки *A — E* играет важную роль в процессе самоорганизации миоглобина.

Интересно, что процесс самоорганизации оказывается, по-видимому, частично направленным, так как на уровне шпильки *G — H* (и, как можно показать, и на уровне комплекса *F — G — H*), т. е. центра кристаллизации [3, 4], есть несколько одинаково выгодных структур, что ведет к некоторому перебору возможностей на следующих стадиях самоорганизации. Число наиболее выгодных структур, к которым принадлежит нативная, оказывается небольшим. Возможно, более точные численные расчеты могли бы выявить различия между альтернативными структурами, но в рамках методов настоящей работы указаний на это нет.

Еще одним интересным результатом является выяснение механизма локализации N-концов спиралей дальними и средними взаимодействиями

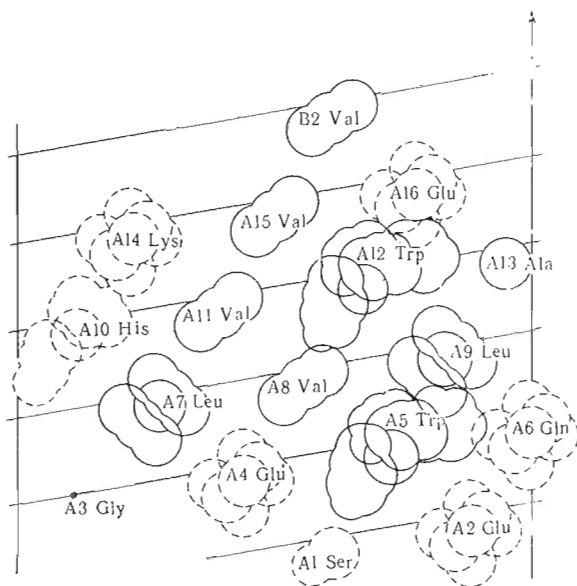
за счет встраивания ближайших к N-концам неспиральных массивных гидрофобных групп в гидрофобные карманы шпилек. Возможно, способ локализации, обнаруженный в настоящей работе для частного случая, имеет общий характер. Для конформаций, близких к нативным, этот механизм оказывается более специфичным, т. е. выигрыш в дегидратации и плотности упаковки за счет встраивания гидрофобных групп в гидрофобные карманы оказывается для этих конформаций большим, чем для явно ненативных.

Анализ конкуренции между дегидратацией и водородными связями (если принять свободную энергию дегидратации одной массивной гидрофобной группы равной $-2-3$ ккал/моль [13] и энергию завязывания водородной связи в фиксированной структуре ~ -2 ккал/моль [14]) показывает, что состояния с максимально увеличенным числом водородных связей имеют примерно ту же (или лишь в несколько раз большую) вероятность, что и состояния с плотно упакованными дегидратированными группами у концов спиралей, уже на уровне шпилек. Дальнейший рост структуры, видимо, еще больше закрепляет плотно упакованные конформации. В одном из трех рассмотренных случаев локализации концов нативная конформация отличается от предсказываемой этим механизмом на уровне шпильки. Это конформация *He FG5*, отличающаяся поворотом вокруг одной из связей в боковой группе примерно на 100° , что и приводит к одному из несовпадений теоретических и нативных контактов (рис. 4). Но в то же время эта боковая группа в нативной структуре участвует во взаимодействиях с гемом, и, возможно, изменения возникают при его присоединении.

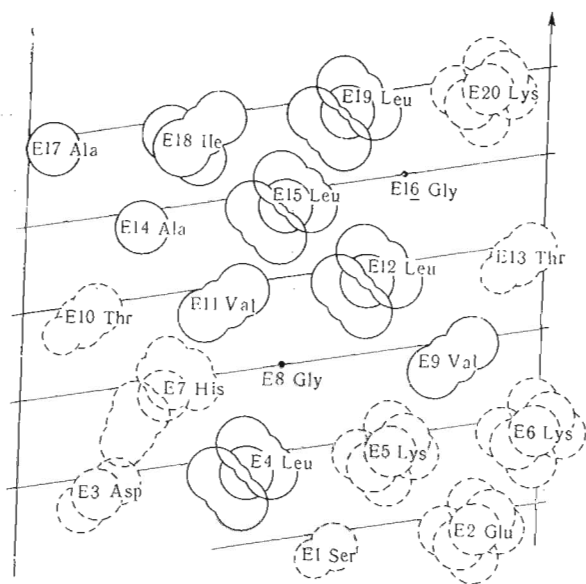
Результаты данной работы, по мнению автора, указывают на возможность теоретического получения наиболее выгодных конформаций фрагментов высокоспиральных белков типа глобинов с точностью $2-3 \text{ \AA}$. Дальнейший прогресс в этой области может быть связан с комбинированием данного метода с конформационными расчетами типа, описанного в работах [7, 15]. Полученные этим методом конформации могут использоваться как исходные для уточнения структур при помощи минимизации энергии взаимодействий.

Методическая часть. Для спиралей *A*, *E*, *G* и *H* миоглобина кашалота [10] (номенклатуру для *H*-спирали см. в работе [11]) строились развертки на плоскость (рис. 8) методом, описанным в работе [1]. Гидрофобные области на каждой развертке произвольно разбивались на несколько узких полос с шириной порядка угловой ширины массивной гидрофобной группы на развертке для искусственного внесения дискретности в перебор конформаций. Такое разбиение приведено на рис. 8, *в*, *г* для спиралей *G* и *H*, на примере которых ниже поясняется метод перебора окрестностей всех плотно упакованных конформаций пары спиралей.

На развертке спирали *H* оказывается одна полоса шириной 60° , на развертке спирали *G* — три полосы. При этом возможны следующие взаимные расположения полос, попарно входящих в область плотной упаковки и почти полностью исчерпывающих эту область: *HGI*, *HGII*, *HGIII*, *GIH*, *GIIH*, *GIIIH* (запись *HGI* означает, что при расположении спиралей, аналогичном представленному на рис. 2, гидрофобная полоса *H*-спирали, входящая в упаковку, находится левее первой гидрофобной полосы *G*-спирали). Кроме разбиения на полосы развертки сдвигались друг относительно друга на $\pm \Delta = 3-5 \text{ \AA}$ от положения, при котором C-конец *G*-спирали находится на уровне N-конца *H*-спирали. В случае шпильки *G* — *H* большие сдвиги невозможны из-за малой длины неспирального участка *GH*. Для спиралей, взаимные сдвиги которых, параллельные их осям, ограничены менее жестко, можно рассматривать любые сдвиги, кратные Δ , образуя таким образом дискретную сетку взаимных расположений разверток. Вблизи каждого из взаимных расположений разверток проводился поиск окрестностей плотно упакованных структур

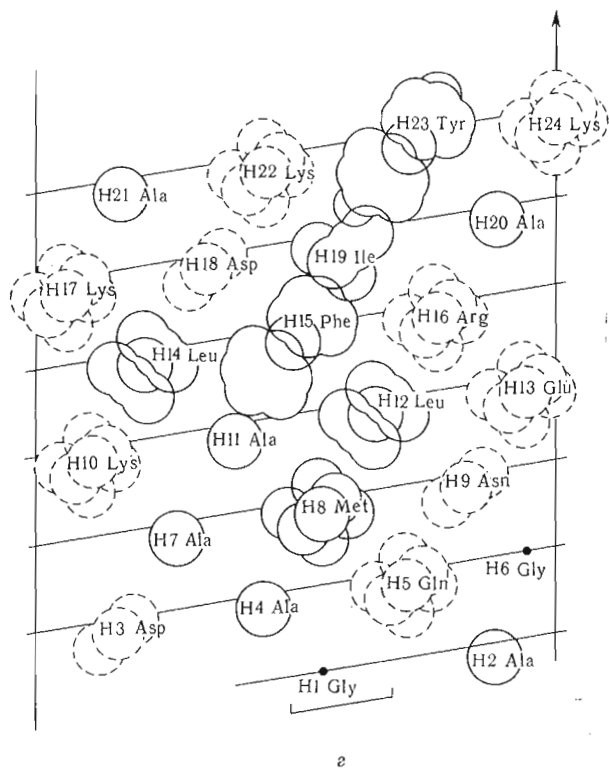
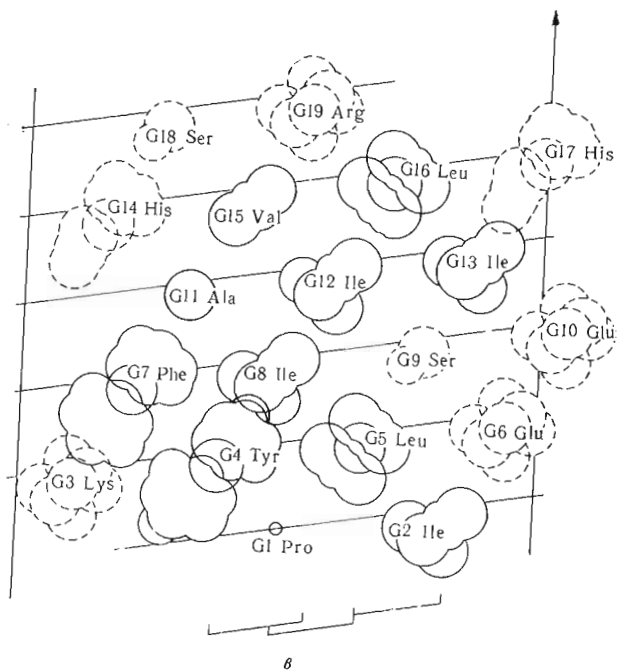


a



b

Рис. 8. Развертки на плоскость спиралей А (а), Е (б). Приведены номер каждого остатка в спирали и его трехбуквенный код



Развертки на плоскость спиралей G (θ), H (z)

методом, описанным в сообщении [1], при положительном и отрицательном значении угла $\alpha - 180^\circ$ (α — угол между положительными направлениями осей спиралей). В настоящем сообщении рассматривались только упаковки антипараллельного типа; при рассмотрении упаковок параллельного типа необходимо проделать те же операции при положительном и отрицательном значениях угла α . Ввиду возможных искажений самих α -спиралей и углов χ поворотных изомеров допускались небольшие налезания изображений боковых групп на развертках различных спиралей в окрестности плотной упаковки. Каждое взаимное расположение разверток без явно различимых внутренних полостей, больших размеров молекулы воды, фиксировалось на накладываемом листе кальки (изображения поворотных изомеров, соответствующих плотной упаковке, обводились фломастером и группы на одной из спиралей заштриховывались).

Отобранные окрестности плотно упакованных конформаций собирались из пространственных моделей СРЖ [16] или модифицированных моделей Courtauld фирмы Griffin & George.

Грубая оценка степени дегидратации массивных гидрофобных групп (остатки Cys, Met, Val, Leu, Ile, Phe, Tyr, Trp) проводилась при помощи визуального определения части поверхности бокового радикала, недоступной для контакта с молекулой воды. Шкала оценок степени дегидратации бралась дискретной: 0, $1/4$, $1/2$, $3/4$, 1. Величина среднеквадратичной ошибки в определении суммарной степени дегидратации равнялась $1/4 \times \sqrt{n}$ (n — число участвующих в плотной упаковке групп) и не превышала единицы для рассмотренных в настоящей работе случаев. (Один из способов оценки степени дегидратации состоял в подсчете относительного числа водородных атомов бокового радикала, недоступных для контакта с водой.) Для большей уверенности измерения проводились по нескольку раз при повторных сборках соответствующей плотной упаковки.

Массивные гидрофобные группы, близкие к концам спиралей, но не входящие в них, пристраивались к плотной упаковке спиралей в конформации, позволяющей заполнить впадины на гидрофобной поверхности пары спиралей и увеличить степень дегидратации ее гидрофобных групп.

Автор признателен О. Б. Птицыну и С. А. Козицыну за полезные обсуждения, а В. Н. Швыркову за помощь в работе.

ЛИТЕРАТУРА

1. Рашин А. А. (1976) Биоорг. химия, 2, 655—665.
2. Денесюк А. И., Птицын О. Б., Финкельштейн А. В. (1974) Биофизика, 19, 549—561.
3. Птицын О. Б., Рашин А. А. (1973) Докл. АН СССР, 213, 473—475.
4. Ptitsyn O. B., Rashin A. A. (1975) Biophys. Chem., 3, 1—20.
5. Финкельштейн А. В. (1974) Препринт, ОНТИ Научного центра биологических исследований АН СССР в Пушкине.
6. IUPAC — IUB Commission (1971) Biochem. J., 121, 577—585.
7. Levitt M. (1974) J. Mol. Biol., 82, 393—420.
8. Scheraga H. A. (1968) Adv. Phys. Org. Chem., 6, 103—184.
9. Nishikawa K., Ooi T., Isogai Y., Saito N. (1972) J. Phys. Soc. Jap., 32, 1331—1337.
10. Watson H. C. (1969) In Progress in Stereochemistry (Harris M. M., Aylett B. J., eds.), vol. 4, pp. 299—333, Butterworth & Co. (Publishers) Ltd, London.
11. Perutz M. F. (1965) J. Mol. Biol., 13, 646—668.
12. Jensen L. H. (1974) Ann. Rev. Biophys. and Bioeng., 3, 81—93.
13. Tanford Ch. (1962) J. Amer. Chem. Soc., 84, 4240—4253.
14. Брандтс Дж. Ф. (1973) в сб. Структура и стабильность биологических молекул (под ред. Волькенштейна М. В.), с. 174, «Мир», М.
15. Warne P. K., Scheraga H. A. (1974) Biochemistry, 13, 757—767.
16. Koltun W. L. (1965) Biopolymers, 3, 665—679.

Поступила в редакцию
15.IX.1975

SEARCH OF TIGHTLY PACKED STRUCTURES OF α -HELIX PAIRS.
II. *A-E* AND *G-H* HAIRPINS OF SPERM WHALE MYOGLOBIN

RASHIN A. A.

*Institute of Protein Research, Academy of Sciences
of the USSR, Pushchino*

Using the method proposed in the previous paper, a search is carried out of the tightly packed structures of α -helical hairpins *A - E* and *G - H* of a sperm whale myoglobin molecule. The structures obtained were used as the result of the first approximation for their further improvement with the help of three-dimensional models. It is shown that conformations very close to native hairpins *A - E* and *C - H* are among the few most favourable structures obtained. For the considered hairpins the mechanism of localization of N-termini of α -helices at the expense of building-in of bulky hydrophobic groups closest to the N-termini into the hydrophobic pockets of hairpins is clarified.
