



УДК 547.962:541.63

АПРИОРНЫЙ РАСЧЕТ ТРЕХМЕРНОЙ СТРУКТУРЫ МОЛЕКУЛЫ
БЫЧЬЕГО ПАНКРЕАТИЧЕСКОГО ТРИПСИНОВОГО ИНГИБИТОРА*Попов Е. М.**Институт биоорганической химии им. М. М. Шемякина
Академии наук СССР, Москва**Годжаев Н. М., Измаилова Л. И., Мусаев Ш. М.,
Алиев Р. Э., Ахмедов Н. А., Максумов И. С.**Азербайджанский государственный университет им. С. М. Кирова, Баку*

На основе количественной структурной теории белков выполнен априорный конформационный анализ молекулы бычьего панкреатического трипсинового ингибитора. Конформация белка, рассчитанная при использовании аминокислотной последовательности и стандартной валентной схемы, сопоставлена с трехмерной структурой белка, найденной с помощью рентгеноструктурного анализа. Среднеквадратичные отклонения теоретической и кристаллографической структур составили по двугранным углам ϕ и ψ 13° , ω -9° , χ -18° , а по расстояниям между всеми парами C^α -атомов основной цепи $-1,4$ Å и длинам пептидных водородных связей $-0,5$ Å. Обсуждаются конформационные аспекты образования дисульфидных связей, функции ближних, средних и дальних взаимодействий в пространственной организации белковых молекул, механизм самосборки полипептидной цепи в нативную глобулярную конформацию белка.

Введение

Настоящее сообщение посвящено обсуждению результатов априорного расчета нативной конформации молекулы бычьего панкреатического трипсинового ингибитора (БПТИ), включающей 58 аминокислотных остатков, последовательность которых была установлена Касселем и Ласковским [1]. Трехмерная структура БПТИ с разрешением $1,9$ Å получена Хубером и соавт. [2] и уточнена до $1,5$ Å Дайзепхофером и Стайгеманном [3]. Белок включает шесть остатков Cys, которые образуют три дисульфидные связи Cys^5-Cys^{55} , $Cys^{14}-Cys^{38}$ и $Cys^{30}-Cys^{51}$. Конформационный анализ БПТИ производился для линейной последовательности белка, так как априорный расчет должен автоматически привести к близости соответствующих остатков Cys в наиболее предпочтительной структуре.

Для определения структуры БПТИ использовался физический подход, предложенный одним из авторов настоящей работы [4]. Априорный расчет конформации БПТИ выполнен на основании разработанной им же конформационной теории белка и отвечающего ей метода расчета [5, 6] в предположении жесткой валентной схемы молекулы. Длины связей основной цепи приняты равными параметрам Поллица и Кори [7]. Валентные углы амидной группы, а также длины связей и валентные углы боковых цепей соответствовали значениям, приведенным в работе [8]. Пропиловое кольцо предполагалось плоским, его геометрия дана в сообщении [9].

Расчет белка выполнен с учетом невалентных и электростатических взаимодействий, водородных связей и торсионных вкладов. Невалентные взаимодействия оценивались по потенциалу Леннарда — Джонса с параметрами Скотта и Шераги [10]. Электростатическая энергия рассчитывалась в монополярном приближении по закону Кулона с использованием зарядов на атомах, предложенных Шерагой [8]. Конформационные состояния фрагментов и всей молекулы БПТИ рассчитаны применительно к условиям водной среды, в связи с чем величина диэлектриче-

ской проницаемости принята равной 10 [11]. Водородные связи, оцениваемые по потенциалу типа Морзе [12], предполагались ослабленными (энергия связи в воде при расстоянии $\text{NH}\dots\text{OC}$ 1,8 Å принята равной -1,1 ккал/моль [13-15]). Торсионные потенциалы и величины барьеров вращения вокруг связей $\text{C}-\text{N}$ (φ), $\text{C}^\alpha-\text{C}'$ (ψ), $\text{C}'-\text{N}$ (ω) $\text{C}^\alpha-\text{C}^\beta$ и других связей боковых цепей (χ) взяты такими же, как в работе [8].

Рамки журнальной статьи не позволяют привести и обсудить все детали независимого расчета трехмерной структуры молекулы БПТИ. Выбирая материал для публикации, мы стремились тем не менее представить хотя бы в очень сжатой форме данные о всех этапах конформационного анализа и на этой основе описать процесс самоорганизации белковой глобулы.

Результаты и обсуждение

Расчету трехмерной структуры молекулы БПТИ предшествовал детальный анализ конформационных возможностей большого числа фрагментов, начиная с N-конца последовательности. При фрагментарном исследовании использован метод конформационного анализа олигопептидов, предложенный в работе [4] и апробированный на многих примерах в работах [16-26]. Он заключается в предсказании сравнительно ограниченного набора низкоэнергетических конформаций олигопептида на основе известных конформационных состояний составляющих его фрагментов. В качестве структурной единицы выбран дипептид. В основе расчета дипептидов лежат низкоэнергетические конформации монопептидов, полученные из расчета молекул метиламидов N-ацетил- α -аминокислот. Таким образом, выбор исходных конформационных состояний аминокислотных остатков никак не связан с их геометрией в известной структуре БПТИ. Использованные в расчете наборы низкоэнергетических вариантов структур монопептидов являются универсальными при рассмотрении любых последовательностей. Количество конформационных состояний, учитываемых в начале расчета для каждой аминокислоты, значительно превышает число конформаций, обнаруживаемых в различных белках известного строения.

Фрагмент $\text{Arg}^1-\text{Pro}^9$. Начнем рассмотрение с результатов анализа N-концевого полипептидного участка молекулы. Ему предшествовал расчет всех перекрывающихся дипептидов. На основе их низкоэнергетических конформаций далее были исследованы структуры трипептидов $\text{Arg}^1-\text{Asp}^3$, $\text{Pro}^2-\text{Phe}^4$, $\text{Phe}^4-\text{Leu}^6$ и тетрапептида $\text{Leu}^6-\text{Pro}^9$. Данные о конформационных состояниях первых двух трипептидов послужили исходными для расчета тетрапептида $\text{Arg}^1-\text{Phe}^4$, который вместе с трипептидом $\text{Phe}^4-\text{Leu}^6$ стал базовым для анализа гексапептида $\text{Arg}^1-\text{Leu}^6$. Знание конформационных возможностей гексапептида $\text{Arg}^1-\text{Leu}^6$ и тетрапептида $\text{Leu}^6-\text{Pro}^9$ сделало реальным структурное исследование фрагмента $\text{Arg}^1-\text{Pro}^9$. На основе полученных данных был составлен набор структур $\text{Arg}^1-\text{Pro}^9$ для минимизации энергии при варьировании 43 двугранных углов $\varphi, \psi, \omega, \chi$, включающий 299 конформационных вариантов, которые распадаются на 24 шейпа. Данные табл. 1 свидетельствуют о чрезвычайно резкой энергетической дифференциации конформаций $\text{Arg}^1-\text{Pro}^9$. В интервал 0-5,0 ккал/моль попадают лишь три конформации шейпа $e_2f_4e_2$, а в интервал 0-10,0 ккал/моль - 12 конформаций, причем пять из них принадлежат одному шейпу. В чем же заключается причина столь четко проявившейся при состыковке двух лабильных фрагментов детерминации структуры пептида?

Рассмотрим прежде всего особенности глобальной конформации фрагмента $B_{1221}^1-B_{31}^2-R_{11}^3-R_{21}^5-R_{321}^6-B_{121}^7-B^8-B^9$, шейп которой, $e_2f_4e_2$, представлен на рис. 1. Основная цель этой конформации на начальном трипептидном участке развернута, затем образует примерно два витка неправильной спирали и удаляется в сторону, опять приобретая развернутый характер. При такой упаковке цепи имеет место сближенность почти

Энергетическое распределение конформаций фрагмента
 Arg¹-Pro²-Asp³-Phe⁴-Cys⁵-Leu⁶-Glu⁷-Pro⁸-Pro⁹

Шейп основной цепи	Интервал $E_{отн}$, ккал/моль						
	>10	10—5	5—4	4—3	3—2	2—1	1—0
e_8	10	—	—	—	—	—	—
e_3fe_2	12	—	—	—	—	—	—
e_4fe_3	6	—	—	—	—	—	—
$e_4fe_2e_2$	7	—	—	—	—	—	—
e_3fe_4	11	—	—	—	—	—	—
$e_3fe_2e_2$	12	—	—	—	—	—	—
$e_3fe_2e_3$	17	—	—	—	—	—	—
$e_3fe_2e_2$	12	—	—	—	—	—	—
e_2fe_5	5	—	—	—	—	—	—
$e_2fe_2fe_2$	13	—	—	—	—	—	—
$e_2fe_2fe_3$	6	—	—	—	—	—	—
$e_2fe_2fe_2e_2$	6	—	—	—	—	—	—
$e_2fe_2e_4$	19	3	—	—	—	—	—
$e_2fe_2fe_2e_2$	24	—	—	—	—	—	—
$e_2fe_2e_3$	28	—	—	—	—	—	—
$e_2fe_2e_2$	24	2	1	—	—	1	1
e_2fe_6	10	—	—	—	—	—	—
$e_2fe_3fe_2$	9	1	—	—	—	—	—
$e_2fe_2fe_3$	12	—	—	—	—	—	—
$e_2fe_2fe_2e_2$	7	—	—	—	—	—	—
$e_2fe_2e_4$	6	1	—	—	—	—	—
$e_2fe_2fe_2e_2$	10	1	—	—	—	—	—
$e_2fe_2e_3$	9	1	—	—	—	—	—
$e_2fe_2e_2$	12	—	—	—	—	—	—

всех аминокислотных остатков фрагмента и реализация тетра-, пента- и т. д. вплоть до октапептидных взаимодействий. Возникновение у глобальной конформации целого ряда новых стабилизирующих контактов не нарушает взаимодействий, присущих соответствующим состояниям участков Arg¹—Leu⁶ и Leu⁶—Pro⁹, которые отвечают наиболее низкоэнергетическим конформациям свободных гекса- ($E_{отн}$ 0,2 ккал/моль) и тетрапептида ($E_{отн}$ 1,7 ккал/моль). Общий стабилизирующий эффект взаимодействий между удаленными по цепи остатками (тетра-октапептидные взаимодействия) приблизительно такой же (−19,3 ккал/моль), как и эффекты взаимодействий на всех перекрывающихся ди- (−16,3 ккал/моль) и трипептидных (−21,5 ккал/моль) участках. Значительная стабилизирующая роль принадлежит остаткам Leu⁶ и Glu⁷. В компактной глобальной конформации нонапептида состояния основных и боковых цепей остатков в условиях весьма ограниченной свободы отвечают своим наиболее предпочтительным состояниям у соответствующих коротких фрагментов, где они обладают значительно большей конформационной свободой.

Рассмотрим теперь конформацию $B_{2212}^1-B_2^3-B_{11}^3-B_{21}^4-B_2^5-B_{321}^6-B_{221}^7-B^8-B^9$ ($E_{отн}$ 17,3 ккал/моль), являющуюся наиболее выгодной среди структур с полностью развернутым шейпом основной цепи e_8 (рис. 1, табл. 1). Она не отличается от глобальной структуры ($E_{отн}$ 0,0 ккал/моль) по энергии ди- (−16,2 ккал/моль) и трипептидных (−20,9 ккал/моль) взаимодействий. Большая относительная энергия этой конформации обусловлена не ее напряженностью, а незначительностью взаимодействий между удаленными остатками; суммарный эффект тетра-, пента- и т. д. пептидных взаимодействий составляет всего −1,4 ккал/моль, т. е. на 17,9 ккал/моль меньше, чем у глобальной структуры. Приблизительно на столько же отличаются у них величины $E_{отн}$. Следовательно, можно заключить, что основной особенностью глобальной конформации является реализация значительных стабилизирующих взаимодействий между всеми остатка-

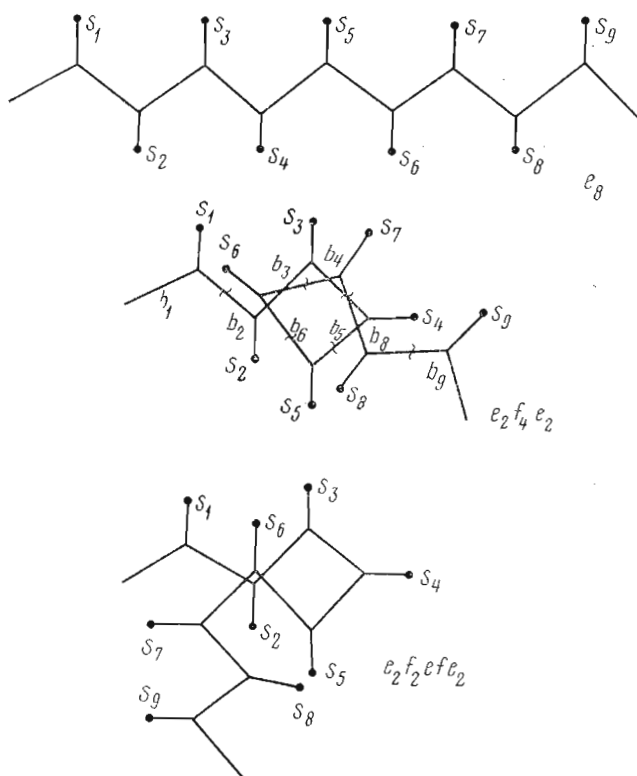


Рис. 1. Шейцы основной цепи фрагмента Arg¹ – Pro⁹

ми нонапептида Arg¹ – Pro⁹. В данном случае мы сравнили эту структуру с полностью развернутой структурой, где заведомо отсутствуют контакты между удаленными остатками. Поэтому этот пример иллюстрирует лишь наличие у глобальной конформации согласованности между всеми взаимодействиями.

Рассмотрим свернутую конформацию $B_{1221}^4-B^2-R_{31}^3-R_{11}^4-B_2^5-R_{321}^6-B_{331}^7-B^8-B^9$ ($E_{отн}$ 13,9 ккал/моль), имеющую шейку $e_2f_2efe_2$ (рис. 1). Эффективными здесь являются взаимодействия Leu⁶ с Arg¹ и Asp³ (–4,2 ккал/моль) и Glu⁷ с Arg¹ (–6,8 ккал/моль). Общий энергетический эффект тетра-, пента- и т.д. пептидных взаимодействий равен –10,6 ккал/моль, т.е. на 8,7 ккал/моль меньше, чем у конформации с $E_{отн}$ 0,0 ккал/моль. Сближенные в шейке $e_2f_2efe_2$ остатки Cys⁵ и Pro⁸ стабилизирующих контактов между собой не образуют. Энергия ди- и трипептидных взаимодействий здесь на 3,6 ккал/моль выше, чем у глобальной структуры. Следовательно, достаточно эффективные взаимодействия между удаленными остатками у лучшей конформации шейки $e_2f_2efe_2$ приводят к ослаблению ди- и трипептидных контактов по сравнению с соответствующими состояниями свободных участков Arg¹ – Leu⁶ и Leu⁶ – Pro⁹. В той или иной степени аналогичная ситуация наблюдается у конформаций всех свернутых шейков, за исключением $e_2f_4e_2$. В двух случаях ($efe_2f_2e_2$ и $efef_3e_2$) образование стабильных конформаций нонапептида невозможно из-за стерических напряжений.

На рис. 2 показан путь расчета конформационных возможностей нонапептида Arg¹ – Pro⁹. Из рис. 2 видно, что конформационные состояния отдельных участков в глобальной структуре отвечают довольно низкокэнергетическим конформациям свободных фрагментов. Максимальный разброс составляет 0–3,8 ккал/моль, причем большие величины $E_{отн}$ имеют короткие фрагменты, включающие Arg¹, в пределах которых не может быть в полной мере реализована потенция этого остатка к взаимодействиям. Однако среди конформаций, имеющих ту же форму основ-

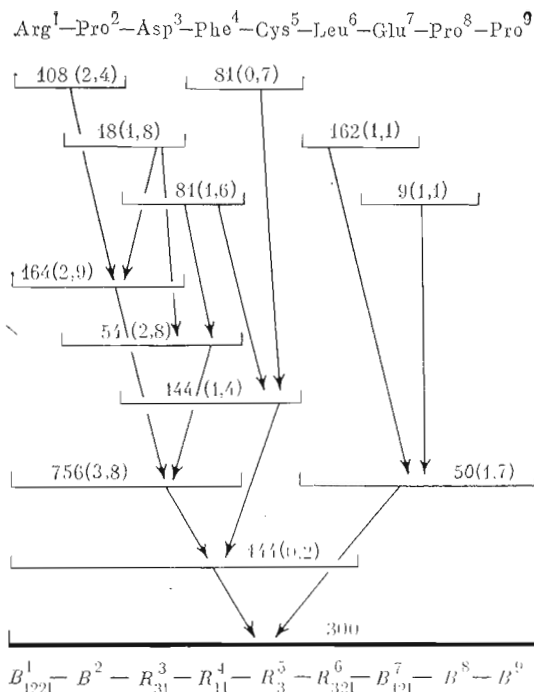


Рис. 2. Схема расчета конформаций фрагмента Arg¹ - Pro⁹. Приведены количество рассчитанных вариантов и в скобках относительная энергия его конформационного состояния (ккал/моль), вошедшего в глобальную структуру нонапептида

ной цепи, они наиболее предпочтительны. Таким образом, величина $E_{отн}$ какой-либо конформации короткого участка белковой цепи не является вполне удовлетворительным критерием для оценки вероятности реализации той или иной конформации в белке. Более показательным представляется отношение энергии данной конформации к энергии других вариантов, имеющих сходную форму основной цепи или шейки. Поэтому при фрагментарном конформационном анализе белка на более поздних стадиях расчета следует учитывать возможно большее число различных конформеров, отвечающих коротким участкам основной цепи. Необходимость этого обусловлена неполной реализацией средних взаимодействий на коротких участках. В нашем случае величины $E_{отн}$ конформационных состояний фрагментов, составивших глобальную структуру, попадают в начало или в середину энергетических интервалов, которые были приняты для отбора перспективных вариантов.

В табл. 2 приведены теоретические значения двугранных углов ϕ , ψ , ω и χ в глобальной конформации свободного нонапептида Arg¹ - Pro⁹ и экспериментальные значения этих углов, полученные с помощью рентгеноструктурного анализа [2, 3]. Сравнение теоретической и кристаллической структур указывает на их полное совпадение по шейке и форме основной цепи. Более того, имеется вполне удовлетворительное количественное согласие между теоретическими и опытными значениями подавляющего большинства конформационных параметров; расхождение, как правило, не выходит за общепринятые границы экспериментальных ошибок при рентгеноструктурном анализе белков. Трудно предположить, что совпадение случайно. Из более 10⁷ возможных конформационных вариантов нонапептида (произведение исходных низкоэнергетических конформаций всех монопептидов) на завершающем этапе рассчитано лишь 299 (а всего около 1850, рис. 2). Таким образом, была рассмотрена чрезвычайно малая часть, казалось бы, равновероятных ситуаций. Структура нонапептидного фрагмента, реализующаясь в белке, не только вошла в состав рассчитанных конформаций, но и оказалась сре-

Экспериментальные (верхняя строчка) и теоретические значения конформационных параметров (град) фрагмента Arg¹ - Pro⁹ молекулы БПТИ

Остаток	Основная цепь			Боковая цепь				
	φ	ψ	ω	χ^1	χ^2	χ^3	χ^4	χ^5
Arg ¹	—	144	179	82	-179	56	82	1
	60	146	170	61	175	180	74	0
Pro ²	-58	152	177	-174	177	-170		
	-60	143	170	180	180	180		
Asp ³	-61	-30	-177	-86	5			
	-94	-40	-173	-61	30			
Phe ⁴	-74	-12	169	66	71			
	-68	-29	173	63	83			
Cys ⁵	-65	-17	176	-56				
	-61	-36	-172	-59				
Leu ⁶	-92	-5	175	-54	-174	—	—	
	-88	-67	-173	-52	177	66	62	
Glu ⁷	-75	147	175	17	177	73		
	-129	149	173	30	178	90		
Pro ⁸	-76	157	178	-178	-165	162		
	-60	159	176	180	180	180		
Pro ⁹	-64	145	-173	178	-175	157		
	-60	149	-179	180	180	180		

Таблица 3

Низкоэнергетические конформации фрагмента Arg¹-Pro²-Asp³-Phe⁴-Cys⁵-Leu⁶-Glu⁷-Pro⁸-Pro⁹

Шейп	Конформация	Энергия E _{отн} , ккал/моль
e ₂ /e ₂	$B_{122}^1 - B^2 - R_{31}^3 - R_{11}^4 - R_3^5 - R_{32}^6 - B_{121}^7 - B^8 - B^9 (R^9)$	0
	$B_{221}^1 - B^2 - R_{31}^3 - R_{11}^4 - R_3^5 - R_{32}^6 - B_{121}^7 - B^8 - B^9 (R^9)$	2,0
e ₁ 'e ₂ e ₃	$B_{2212}^1 - R^2 - B_{11}^3 - R_{31}^4 - R_3^5 - B_{32}^6 - B_{221}^7 - B^8 - B^9 (R^9)$	5,4
e ₂ 'e ₂ e ₄	$B_{1221}^1 - B^2 - R_{31}^3 - R_{11}^4 - B_3^5 - B_{31}^6 - B_{331}^7 - B^8 - B^9 (R^9)$	6,1
e ₁ e ₂ e ₃ e ₂	$B_{2222}^1 - R^2 - B_{11}^3 - R_{31}^4 - B_3^5 - R_{32}^6 - B_{311}^7 - B^8 - B^9 (R^9)$	7,5

ди них термодинамически самой устойчивой. Отметим, что этот факт является первым серьезным доводом в пользу конформационной теории белка [4—6], на основе которой выполняется расчет БПТИ.

Оставленные для последующего конформационного анализа пять низкоэнергетических структур с *B*- и *R*-состояниями Pro⁹ приведены в табл. 3.

Фрагмент Arg¹-Cys¹⁴. Конформационный анализ тетрадекапептидного участка БПТИ включает предварительный расчет пентапептида Tyr¹⁰-Cys¹⁴ и октапептида Gly⁷-Cys¹⁴, а затем, с использованием данных по нонапептиду Arg¹-Pro⁹, — последовательности Arg¹-Cys¹⁴ (рис. 3).

Расчет фрагмента Tyr¹⁰-Cys¹⁴ выполнен исходя из предпочтительных конформаций перекрывающихся дипептидных участков, которые были получены из анализа всех возможных комбинаций конформационных состояний соответствующих молекул метиламидов *N*-ацетил- α -аминокислот [16—17]. Сочетания отобразивших низкоэнергетических дипептидных конформаций составили 288 начальных приближений участка Tyr¹⁰-Cys¹⁴. Они распадаются на 16 шейпов, каждый из которых включает 18 конформаций с двумя различными формами основной цепи. Рас-

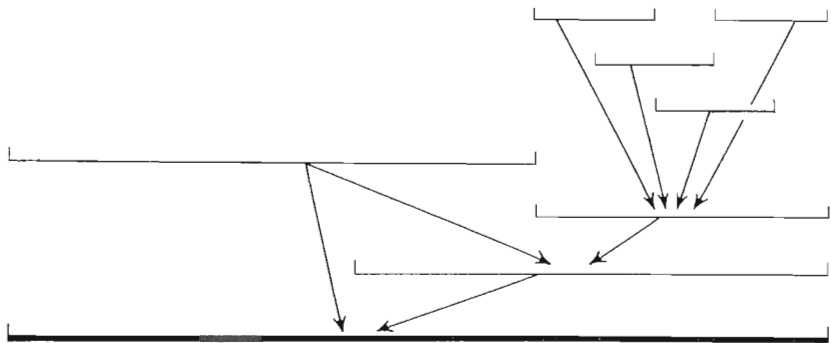
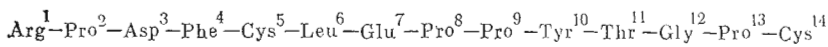


Рис. 3. Схема расчета конформаций фрагмента Arg¹ — Cys¹⁴

пределение рассмотренных конформаций пентапептида по энергии, шейпам и формам приведено в табл. 4. Среди большого числа возможных вариантов выделяется весьма ограниченный набор низкоэнергетических состояний фрагмента. Всего восемь конформаций шейпов e_2f_3 , f_2f_2 и f_4 имеют относительную энергию меньше 5 ккал/моль.

Конформации шейпа e_2f_2 характеризуются двумя поворотами пептидной цепи, что приводит к изменению ее направления на противоположное, сближая тем самым концы фрагмента, а при соответствующих конформационных состояниях остатков — к образованию эффективных взаимодействий. В лучшей конформации этого шейпа, $B_{112}^{10} - R_{12}^{11} - P^{12} - R^{13} - B^{14}$, дополнительная стабилизация за счет тетра- и пентапептидных взаимодействий составляет $-5,9$ ккал/моль.

Различие между конформациями пептидной цепи шейпов f_2f_2 и e_2f_2 касается только положения остатка Tyr¹⁰ относительно свернутого одинаковым образом участка Thr¹¹ — Cys¹⁴. Лучшая конформация шейпа f_2f_2 $R_{112}^{10} - R_{12}^{11} - P^{12} - R^{13} - B^{14}$ изоэнергетична конформации $B_{112}^{10} - R_{12}^{11} - P^{12} - R^{13} - B^{14}$ шейпа e_2f_2 .

Полностью свернутые конформации шейпа f_4 образуются при двух формах основной цепи: $R^{10} - R^{11} - P^{12} - R^{14} - B^{14}$ и $R^{10} - B^{11} - P^{12} - R^{13} - B^{14}$. Все варианты с α -спиральной формой высокоэнергетичны из-за отсутствия значительных стабилизирующих взаимодействий, а не вследствие напряжений, вызываемых Pro¹³. Для дальнейшего анализа отобраны 13 конформаций шейпов e_3f , e_2f_2 , f_2f_2 , f_2ef и f_4 , энергия которых не превышает 6,0 ккал/моль.

Исходные приближения для минимизации энергии октапептида Glu⁷ — Cys¹⁴ сформированы из самого предпочтительного конформационного состояния $B_{121}^7 - B^8 - B^9$ (R^9) участка Glu⁷ — Pro⁸ — Pro⁹ и отмеченных выше 13 состояний участка Tyr¹⁰ — Cys¹⁴. Полученные после минимизации при варьировании углов φ , ψ , ω и χ величины относительной энергии различных конформаций октапептида приведены в табл. 5.

У конформации шейпа e_6f существенными могут оказаться взаимодействия между остатками в первом и третьем положениях. Энергия всех конформаций этого типа равна сумме энергии изолированных фрагментов Glu⁷ — Pro⁹ и Tyr¹⁰ — Cys¹⁴ и энергии ди- и трипептидных взаимодействий на участке Pro⁸ — Pro⁹ — Tyr¹⁰ — Thr¹¹. Поскольку вклад последних весьма невелик (в лучшем варианте шейпа e_6f $-1,3$ ккал/моль), все конформации октапептида этого типа из-за высокой энергии конформеров e_3f пентапептидного участка Tyr¹⁰ — Cys¹⁴ ($E_{\text{отн}}$ 5,5–9,5 ккал/моль, табл. 4) малоперспективны и исключены из дальнейшего анализа.

В конформации шейпа $e_2f_2e_3f$ состояния фрагментов Glu⁷ — Pro⁹ и Tyr¹⁰ — Cys¹⁴ сходны с их состояниями в только что рассмотренном случае (табл. 5). Изменение угла φ_9 остатка Pro⁹ приблизительно на 180°

Относительная энергия конформаций фрагмента
Glu⁷-Pro⁸-Pro⁹-Tyr¹⁰-Thr¹¹-Gly¹²-Pro¹³-Cys¹⁴

Шейп	Конформация	Отн, ккал/моль
e_6f	${}^7_{121}B^8-B^9-B_{212}^{10}-R_{32}^{11}-L^{12}-R^{13}-B_3^{14}$	6,4
e_2fe_3f	$B_{121}-B-B-B_{212}-R_{32}-L-R-B_3$	3,3*
e_5f_2	$B_{121}-B-B-B_{112}-R_{12}-P-R-B_3$	1,7*
	$B_{321}-B-B-B_{212}-R_{12}-P-R-B_3$	1,3*
	$B_{121}-B-B-B_{212}-R_{22}-P-R-B_3$	4,2
	$B_{121}-B-B-B_{212}-R_{32}-P-R-B_3$	2,7*
	$B_{121}-B-B-B_{312}-R_{12}-P-R-B_3$	1,5*
	$B_{121}-B-B-B_{212}-B_{12}-R-R-B_3$	6,7
$e_2fe_2f_2$	$B_{121}-B-R-B_{112}-R_{12}-P-R-B_3$	2,3*
	$B_{121}-B-R-B_{212}-R_{12}-P-R-B_3$	0*
	$B_{121}-B-R-B_{212}-R_{22}-P-R-B_3$	4,7
	$B_{121}-B-R-B_{212}-R_{32}-P-R-B_3$	0,2*
	$B_{121}-B-R-B_{312}-R_{12}-P-R-B_3$	3,4*
	$B_{121}-B-R-B_{212}-B_{12}-R-R-B_3$	4,6
$e_3fe_2f_2$	$B_{121}-B-B-R_{112}-R_{12}-P-R-B_3$	3,4*
	$B_{121}-B-B-R_{112}-R_{32}-P-R-B_3$	4,2
	$B_{121}-B-B-R_{212}-R_{12}-P-R-B_3$	3,3*
	$B_{121}-B-B-R_{312}-R_{12}-P-R-B_3$	2,7*
e_3f_2ef	$B_{121}-B-B-R_{212}-R_{12}-B-R-B_3$	6,5
e_3f_4	$B_{121}-B-B-R_{212}-B_{32}-P-R-B_3$	3,0*
e_2f_5	$B_{121}-B-R-R_{212}-B_{32}-P-R-B_3$	0,3*

* Конформации, отобранные для последующего анализа.

меняет их взаимное расположение и сближает остатки Pro⁸ и Thr¹¹. Это сопровождается появлением дополнительных ди-, три- и тетрапептидных взаимодействий, понижающих энергию на 3,0–4,0 ккал/моль.

Конформации шейпа e_5f_2 имеют две формы основной цепи, одна из которых представлена пятью вариантами, а другая — одним (табл. 5). В конформациях обеих форм дополнительные взаимодействия осуществляются в районе стыковки фрагментов Glu⁷ — Pro⁹ и Tyr¹⁰ — Cys¹⁴ на уровне ди- и трипептидов.

Близкая ситуация наблюдается у конформаций типа $e_2fe_2f_2$, которые отличаются от e_5f_2 изгибом на участке Pro⁹ — Tyr¹⁰. Благодаря возникновению незначительных тетра- и пентапептидных контактов энергия некоторых конформаций этого шейпа на 1,0–2,0 ккал/моль ниже соответствующих конформаций e_5f_2 . В ряде случаев образование новых стабилизирующих взаимодействий сопровождается ослаблением взаимодействий внутри фрагментов Glu⁷ — Pro⁹ и Tyr¹⁰ — Cys¹⁴.

Некоторые конформации пентапептида с формой основной цепи $R-R-P-R-B$ (fe_2f_2) представляют собой компактные структуры и имеют низкую энергию (табл. 4). При их включении в конформации октапептида шейпа $e_3fe_2f_2$ возникают дополнительные контакты между остатками Pro⁹ и Glu¹² (~–2,0 ккал/моль). Однако это приводит к дестабилизации фрагмента Tyr¹⁰ — Cys¹⁴ и увеличению энергии приблизительно на 5,0 ккал/моль, что указывает на некомплементарность состояний типа e_2 у трипептида и fe_2f_2 у пентапептида, соединенных между собой цепью B^3-R^{10} (e). При состыковке этих фрагментов, находящихся в тех же конформационных состояниях, с помощью свернутой цепи R^9-R^{10} (f), т. е. в случае конформации октапептида шейпа $e_3f_2ef_2$, возникают сильные стерические препятствия и эти конформации становятся нереальными.

Шейп e_3f_2ef представлен одной конформацией, в которой сближены почти все остатки; суммарная энергия дополнительных тетра-, пента- и гексапептидных взаимодействий составляет –5,0 ккал/моль. Тем не менее эти контакты неэффективны, так как их образование сопровождается

ухудшением взаимодействий во фрагментах $\text{Glu}^7 - \text{Pro}^9$ и $\text{Tyr}^{10} - \text{Cys}^{14}$. Конформация e_3f_2ef уступает глобальной структуре 6,5 ккал/моль. Конформация близкого шейпа e_2f_3ef вообще нестабильна из-за перекрывания ван-дер-ваальсовых радиусов атомов.

Среди возможных конформеров пентапептида $\text{Tyr}^{10} - \text{Cys}^{14}$ с полностью свернутой формой шейпа f_4 лишь один вариант с основной цепью, $R^{10}-B^{11}-P^{12}-R^{13}-B^{14}$, имеет низкую энергию (табл. 4). В конформации октапептида шейпа e_3f_4 фрагмент $\text{Glu}^7 - \text{Pro}^9$ с развернутой формой цепи и фрагмент $\text{Tyr}^{10} - \text{Cys}^{14}$ со свернутой формой остаются в значительной степени автономными, возникает только одно новое взаимодействие между Pro^9 и Pro^{13} (-1,2 ккал/моль). Ситуация заметно меняется, когда фрагмент $\text{Pro}^9 - \text{Tyr}^{10}$ становится свернутым, т. е. при переходе к конформации e_2f_5 . Фрагменты сближаются, и появляются стабилизирующие контакты между Pro^8 и Thr^{11} (-2,3 ккал/моль), Pro^9 и Pro^{13} (-0,8 ккал/моль) и Glu^7 и Pro^{13} (-2,1 ккал/моль). Однако и в этом случае между новыми и ранее существовавшими у свободных три- и пентапептидных фрагментов взаимодействиями нет удовлетворительной согласованности — появление одних сопровождается ослаблением других. Тем не менее энергия единственной конформации шейпа e_2f_5 остается одной из самых минимальных.

Таким образом, проведенный анализ октапептида $\text{Glu}^7 - \text{Cys}^{14}$ не обнаружил конформаций, содержащих такие контакты между фрагментами $\text{Glu}^7 - \text{Pro}^9$ и $\text{Tyr}^{10} - \text{Cys}^{14}$, которые привели бы к значительному понижению общей энергии. Лучшими вариантами у $\text{Glu}^7 - \text{Cys}^{14}$ оказались те, в которых при минимальных контактах между $\text{Glu}^7 - \text{Pro}^9$ и $\text{Tyr}^{10} - \text{Cys}^{14}$ фрагменты имеют структуры, отвечающие наиболее предпочтительным структурам при свободном состоянии этих фрагментов. К ним прежде всего относятся конформации шейпов e_5f_2 и $e_2fe_2f_2$ с формами основной цепи $B^7-B^8-B^9-B^{10}-R^{11}-P^{12}-B^{13}-B^{14}$ и $B^7-B^8-R^9-B^{10}-R^{11}-P^{12}-R^{13}-B^{14}$.

Начальные конформационные варианты тетрадекапептида $\text{Arg}^1 - \text{Cys}^{14}$ были составлены из предпочтительных состояний фрагментов $\text{Arg}^1 - \text{Pro}^9$ (табл. 3) и $\text{Glu}^7 - \text{Cys}^{14}$ (табл. 5). Задача свелась к рассмотрению 70 конформаций 24 шейпов, причем каждый шейп представлен одной формой основной цепи. Энергетическое распределение оптимальных конформаций тетрадекапептида $\text{Arg}^1 - \text{Cys}^{14}$ (табл. 6) свидетельствует, что только варианты шейпов $e_2f_4e_3f_2$ и $e_2f_4e_2fe_2f_2$ имеют низкую энергию и перспективны для дальнейшего рассмотрения. К ним также можно отнести одну конформацию шейпа $e_2fe_4e_2f_5$. У первых двух шейпов наиболее предпочтительны те варианты, в которых остаток Arg^1 находится в состоянии B_{1221}^1 . Следовательно, у всех низкоэнергетических конформаций обоих шейпов N-концевой фрагмент ($\text{Arg}^1 - \text{Pro}^8$) тетрадекапептида находится в одном

и том же конформационном состоянии $B_{1221}^1-B^2-R_{31}^3-R_{11}^4-R_3^5-R_{32}^6-B_{121}^7-B^8$, а фрагмент $\text{Tyr}^{10} - \text{Cys}^{14}$ — в состоянии $B^{10}-R^{11}-P^{12}-R^{13}-B^{14}$, причем в двух случаях наборы конформационных состояний боковых цепей Tyr^{10} и Thr^{11} одинаковы.

Восемь низкоэнергетических конформаций шейпов $e_2f_4e_3f_2$ и $e_2f_4e_2fe_2f_2$ представляют собой комбинации самых выгодных состояний свободных нонапептида $\text{Arg}^1 - \text{Pro}^9$ (табл. 3) и пентапептида $\text{Tyr}^{10} - \text{Cys}^{14}$ (табл. 4); к наиболее предпочтительным относятся также состояния октапептида $\text{Glu}^7 - \text{Cys}^{14}$, приведенные в табл. 5. В конформациях шейпа $e_2f_4e_3f_2$ противоположные свернутые концы пептидной цепи удалены друг от друга (рис. 4). В более компактных конформациях $e_2f_4e_2fe_2f_2$ сближенными оказываются остатки Pro^2 , Cys^5 и Pro^8 с остатками Tyr^{10} и Cys^{14} . Однако из-за специфики боковых цепей этих остатков и их взаимного расположения эффективных невалентных взаимодействий здесь не возникает. Поэтому энергия конформаций этого типа также складывается практически аддитивно из вкладов свободных фрагментов $\text{Arg}^1 - \text{Pro}^9$ и $\text{Tyr}^{10} - \text{Cys}^{14}$. Конформации обоих шейпов различаются только состоянием (B или R) остатка Pro^9 . Из-за отсутствия дальних взаимодействий изменение одного угла ψ приблизительно на 180° не сопряжено со значительными энергетическими затратами.

Энергетическое распределение конформаций фрагмента Arg¹-Pro²-Asp³-Phe⁴-Cys⁵-
Leu⁶-Glu⁷-Pro⁸-Pro⁹-Tyr¹⁰-Thr¹¹-Gly¹²-Pro¹³-Cys¹⁴

Шейп	Интервал энергии, ккал/моль								
	>8	8—7	7—6	6—5	5—4	4—3	3—2	2—1	1—0
$e_2f_4e_5f_2$	4	—	—	1	—	2	1	—	—
$e_2f_4e_2fe_2f_2$	—	2	1	—	1	—	2	1	1
$e_2f_4e_3fe_2$	3	—	1	2	—	—	—	—	—
$e_2f_4e_3f_4$	1	1	—	—	—	—	—	—	—
$e_2f_4e_2f_5$	—	—	1	—	1	—	—	—	—
$e_2f_4e_2fe_3f$	2	—	—	—	—	—	—	—	—
$e_2f_2e_7f_2$	4	—	—	—	—	—	—	—	—
$e_2f_2e_4fe_2f_2$	2	2	—	—	—	—	—	—	—
$e_2f_2e_5fe_2f_2$	3	—	—	—	—	—	—	—	—
$e_2f_2e_5f_4$	1	—	—	—	—	—	—	—	—
$e_2f_2e_4f_5$	—	1	—	—	—	—	—	—	—
$e_2f_2e_4fe_3f$	1	—	—	—	—	—	—	—	—
$e_2f_2e_4fe_5f_2$	4	—	—	—	—	—	—	—	—
$e_2f_2e_2fe_2fe_2f_2$	4	—	—	—	—	—	—	—	—
$e_2f_2e_2fe_3fe_2f_2$	3	—	—	—	—	—	—	—	—
$e_2f_2e_2fe_3f_4$	1	—	—	—	—	—	—	—	—
$e_2f_2e_2fe_2f_5$	1	—	—	—	—	—	—	—	—
$e_2f_2e_2fe_2fe_3f$	1	—	—	—	—	—	—	—	—
$e_2f_2e_6f_2$	4	—	—	—	—	—	—	—	—
$e_2f_2e_3fe_2f_2$	3	1	—	—	—	—	—	—	—
$e_2f_2e_4fe_2f_2$	3	—	—	—	—	—	—	—	—
$e_2f_2e_4f_5$	1	—	—	—	—	—	—	—	—
$e_2f_2e_3f_5$	—	1	—	—	—	—	—	—	—
$e_2f_2e_3fe_3f$	1	—	—	—	—	—	—	—	—

Найденные низкоэнергетические структуры двух шейпов тетрадекапептида представляют интерес потому, что в случае шейпа $e_2f_4e_5f_2$ форма основной цепи полностью совпадает с конформацией фрагмента Arg¹ — Cys¹⁴ в кристаллографической структуре БПТИ. Более того, у самой выгодной конформации этого типа рассчитанные значения двугранных углов φ , ψ , ω и χ совпадают с экспериментальными значениями.

Вторая группа низкоэнергетических конформаций шейпа $e_2f_4e_2fe_2f_2$ может реализоваться при свертывании белковой цепи в условиях *in vitro* в процессе ренатурации восстановленной молекулы БПТИ. Крейтон [27, 28] при исследовании промежуточных состояний, образующихся при свертывании денатурированной белковой цепи, обнаружил продукт с дисульфидной связью между Cys⁵ и Cys¹⁴, которая отсутствует в нативной структуре БПТИ. В свете полученных нами результатов образование такой связи весьма вероятно. Расчет показал, что в самых предпочтительных по энергии (табл. 6) конформациях $e_2f_4e_2fe_2f_2$ свободного тетрадекапептида Arg¹ — Cys¹⁴ остатки Cys⁵ и Cys¹⁴ оказываются сближенными (рис. 4). Таким образом, из найденных теоретически низкоэнергетических конформаций двух форм основной цепи одна совпадает со структурой в нативном белке, а другая отвечает промежуточному метастабильному состоянию при свертывании денатурированной цепи. Обе конформации представляют собой два варианта упаковки одних и тех же структур; на ранней стадии ренатурации они легко переходят друг в друга.

Самые предпочтительные конформации Arg¹ — Cys¹⁴ приведены в табл. 7. Обращает на себя внимание то обстоятельство, что наиболее низкоэнергетические варианты предпочтительны не только по величине общей энергии, но и по энергии каждого отдельного вида взаимодействий. Наличие такой согласованности между невалентными, электростатическими и торсионными взаимодействиями и водородными связями делает весьма устойчивым энергетическое распределение конформаций в отношении параметризации потенциальных функций.

Для более детального исследования силового поля конформации фрагмента Arg¹ — Cys¹⁴, совпадающей с экспериментом ($E_{отн}$ 2,8 ккал/моль,

Рис. 4. Шейны наиболее предпочтительных конформаций фрагмента Arg¹ — Cys¹⁴

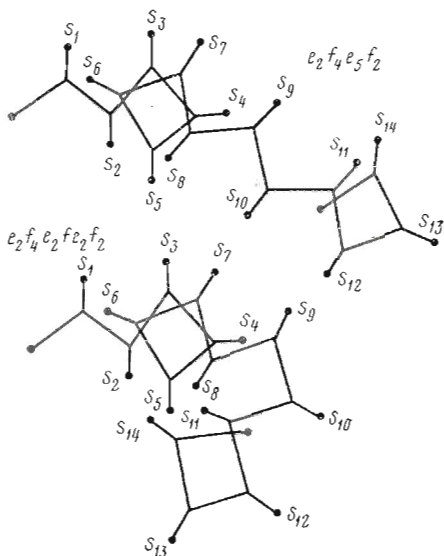


Рис. 4

Рис. 5. Совмещенные конформационные карты ϕ — ψ метиламида N-ацетилаланина и глицина (пунктир) и остатков Asp³, Phe⁴, Cys⁵ и Leu⁶ (а), Glu⁷ (б) и Gly¹² (в) в конформации фрагмента Arg¹ — Cys¹⁴ с $E_{отн}$ 2,8 ккал/моль (табл. 6). Замкнутые эквипотенциальные линии соответствуют относительной энергии 1,0 и 2,0 ккал/моль. Знаком (х) отмечены оптимальные значения углов ϕ , ψ

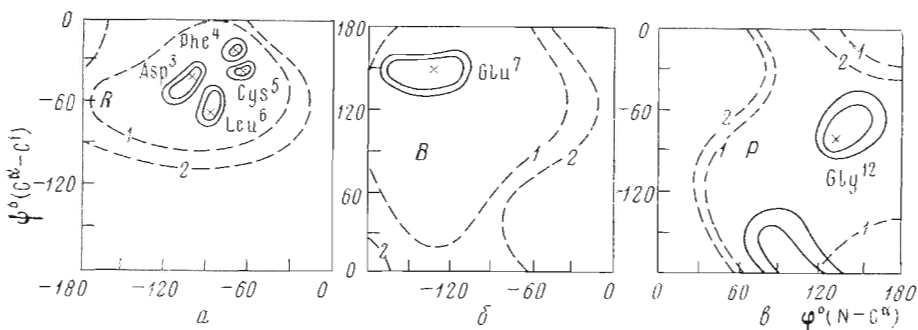


Рис. 5

табл. 7, рис. 4), нами получен целый ряд сечений ϕ — ψ потенциальной поверхности. Для их построения определялась энергия во всех точках той низкоэнергетической области конформационной карты ϕ — ψ монопептида, в которую попадают значения двугранных углов соответствующего остатка в исследуемой конформации Arg¹ — Cys¹⁴. Положение боковой цепи этого остатка и конформационные состояния других остатков фрагмента при этом остаются неизменными. Сечения были построены для Asp³, Phe⁴, Cys⁵, Leu⁶, Glu⁷, Pro⁸, Pro⁹, Gly¹² и Pro¹³.

На конформационных картах ϕ — ψ остатков Asp³, Phe⁴, Cys⁵ и Leu⁶ с *R*-формой основной цепи обнаруживается лишь один энергетический минимум. На рис. 5а потенциальные минимумы остатков Asp³, Phe⁴, Cys⁵ и Leu⁶ совмещены с областью *R* конформационной карты метиламида N-ацетилаланина. На карте ϕ — ψ свободного монопептида они занимают сравнительно небольшую площадь в средней части области *R*, имеющей самую низкую энергию. Найденные для оптимальной конформации фрагмента Arg¹ — Cys¹⁴ ($E_{отн}$ 2,8 ккал/моль) значения ϕ , ψ точно отвечают самым низкоэнергетическим точкам соответствующих минимумов. Таким образом, остатки Asp³, Phe⁴, Cys⁵ и Leu⁶ в кристаллической структуре фрагмента имеют весьма ограниченную конформационную свободу и падают в наиболее низкоэнергетических состояниях. Поскольку в каждом случае имеется только одна область низкой энергии, конформационное состояние каждого остатка пептидной цепи фрагмента Arg¹ — Cys¹⁴ взаимнообусловлено. Кооперативность столь велика, что делает невозможным изменение конформации основной цепи одного остатка без одновременного изменения конформаций других остатков, т. е. без разрушения

Низкоэнергетические конформации фрагмента Arg¹-Pro²-Asp³-Phe⁴-Cys⁵-Leu⁶-Glu⁷-Pro⁸-Pro⁹-Tyr¹⁰-Thr¹¹-Gly¹²-Pro¹³-Cys¹⁴

Ко фор мла	E _{отп} , ккал/моль	Энергия взаимодействий			
		невалент.	торс.	вод. св.	электрост.
$e_2^4 e_5^2$ $B_{1221}^1 - B_{31}^2 - R_{31}^3 - R_{11}^4 - R_3^5 - R_6^6 - B_{121}^7 - B_{121}^8 - B_{121}^{10} - R_{12}^{11} - P_{12}^{12} - R_{13}^{13} - B_3^{14} \quad (R_3^{14})$ » $B_{312} - R_{12} - P - R - B_3 \quad (R_3)$ » $B_{112} - R_{12} - P - R - B_3 \quad (R_3)$ » $B_{212} - R_{32} - P - R - B_3 \quad (R_3)$	2,8	-60,8	6,4	-6,0	-0,7
	3,7	-60,2	6,5	-6,0	-0,8
	3,8	-60,8	7,2	-6,0	-0,8
	5,4	-58,4	6,3	-6,0	-0,7
$e_2^4 e_2^4 e_2^2$ $B_{1221}^1 - B_{31}^2 - R_{31}^3 - R_{11}^4 - R_3^5 - R_6^6 - B_{121}^7 - B_{121}^8 - R_{121}^{10} - R_{12}^{11} - P_{12}^{12} - R_{13}^{13} - B_3^{14} \quad (R_3^{14})$ » $B_{212} - R_{32} - P - R - B_3 \quad (R_3)$ » $B_{312} - R_{12} - P - R - B_3 \quad (R_3)$ » $B_{112} - R_{12} - P - R - B_3 \quad (R_3)$	0	-65,0	7,9	-6,0	-1,1
	1,3	-64,4	8,3	-5,0	-1,1
	2,4	-62,6	8,4	-6,0	-1,3
	2,5	-62,5	8,3	-6,0	-1,5
$e_2^4 e_2^5$ $B_{1221}^1 - B_{31}^2 - R_{31}^3 - R_{11}^4 - R_3^5 - R_6^6 - B_{121}^7 - B_{121}^8 - R_{121}^{10} - B_{11}^{11} - P_{12}^{12} - R_{13}^{13} - B_3^{14} \quad (R_3^{14})$	4,5	-62,9	10,7	-6,4	-1,4

всей системы средних взаимодействий в данном конформационном состоянии. В силу этого обстоятельства, а тем более при наличии согласованности ближних, средних и дальних взаимодействий маловероятно не только изменение структуры основной цепи этого участка, но и большие отклонения углов φ , ψ в рамках той же структуры при генерации аминокислотной последовательности и ее укладке в нативную структуру БПТИ. Именно в этом одна из причин удовлетворительного совпадения расчетных для свободного фрагмента конформационных параметров с параметрами в сложившейся структуре белка. Простой профиль потенциальных сечений объясняет эффективность процедуры минимизации.

Сказанное выше относится и к остатку Glu⁷ с *B*-формой основной цепи. Его сечение φ — ψ (рис. 5б) также содержит одну область низкой энергии, вытянутую вдоль оси φ ; ее минимум отвечает расчетным углам φ , ψ остатка Glu⁷ в рассмотренной конформации Arg¹—Cys¹⁴. Сечение потенциальной поверхности φ — ψ остатка Gly¹² (рис. 5в) имеет две области низкой энергии, которые занимают большую по сравнению с рассмотренными случаями площадь. Несмотря на вполне удовлетворительное совпадение уже полученных результатов расчета с опытными данными, необходимо выяснить причину наличия в одной низкоэнергетической области нескольких потенциальных ям, так как в ином случае последующий конформационный анализ белка может пойти по неправильному руслу. В самом деле, расчет Arg¹—Cys¹⁴ привел нас к набору низкоэнергетических конформаций, на основе которых будет проводиться дальнейшее исследование БПТИ. В этом наборе все варианты содержат остаток Gly¹² в более предпочтительном для свободного фрагмента состоянии с φ 126° и ψ -84°. С помощью используемой процедуры минимизации невозможно прийти к иному состоянию, отвечающему второй потенциальной яме. Для этого нужно задать соответствующее нулевое приближение, что требует знания потенциальной поверхности. Пример с Gly¹² показывает, что исследование у наиболее предпочтительных конформаций сечений потенциальной поверхности должно явиться составной частью априорного расчета трехмерной структуры белка.

Таким образом, структурное исследование Arg¹—Cys¹⁴-участка БПТИ выявило девять низкоэнергетических и перспективных для последующего анализа конформаций, принадлежащих шейнам $e_2f_1e_3f_2$, $e_2f_4e_3fe_2f_2$ и $e_2f_4e_2f_5$ (табл. 6).

Фрагмент Arg¹—Tyr²¹. Схема расчета фрагмента Arg¹—Tyr²¹ приведена на рис. 6. Сложность задачи потребовала разделить анализ этого фрагмента на большое число ступеней. Кратко остановимся на одном из промежуточных этапов расчета гептадекапептида Arg¹—Arg¹⁷. Его начальные приближения были составлены из предпочтительных конформаций перекрывающихся по четырем остаткам тетрадекапептида Arg¹—Cys¹⁴ и гептапептида Thr¹¹—Arg¹⁷. Они могут быть разбиты на три группы. К первой относятся 24 конформации Arg¹—Arg¹⁷, в которых отобранные четыре варианта фрагмента Arg¹—Cys¹⁴ имеют шейп $e_2f_4e_3f_2$, а фрагмента Cys¹⁴—Arg¹⁷—шейпы e_3 , e_2f , efe , fe_2 , ef_2 и fef , отвечающие состояниям этого тетрапептида в низкоэнергетических конформациях фрагмента Thr¹¹—Arg¹⁷. Варианты второй группы различаются лишь конформацией Pro⁹ и на участке Arg¹—Cys¹⁴ относятся к шейпу $e_2f_4e_2fe_2f_2$. Третью группу образуют структуры, имеющие одинаковое конформационное состояние фрагмента Arg¹—Cys¹⁴ шейпа $e_2f_4e_2f_5$ и восемь различных состояний участка Cys¹⁴—Arg¹⁷ шейпов e_3 , e_2f , efe , fe_2 , ef_2 и fef .

У первой группы имеются четыре низкоэнергетические конформации шейпов $e_2f_4e_3f_2e_3$ и $e_2f_4e_3f_2e_2f$ (табл. 8). Среди 24 рассчитанных конформаций второй группы лишь две имеют относительно низкую энергию (4,4 и 5,9 ккал/моль). Они принадлежат шейпу $e_2f_4e_2fe_2f_2e_2f$. Из восьми рассмотренных структурных вариантов третьей группы шесть имеют энергию от 0 до 5,5 ккал/моль.

Таким образом, минимизация 56 исходных вариантов гептадекапептида свидетельствует о резкой энергетической дифференциации его конформаций. Запрещенной из-за многочисленных неблагоприятных контактов

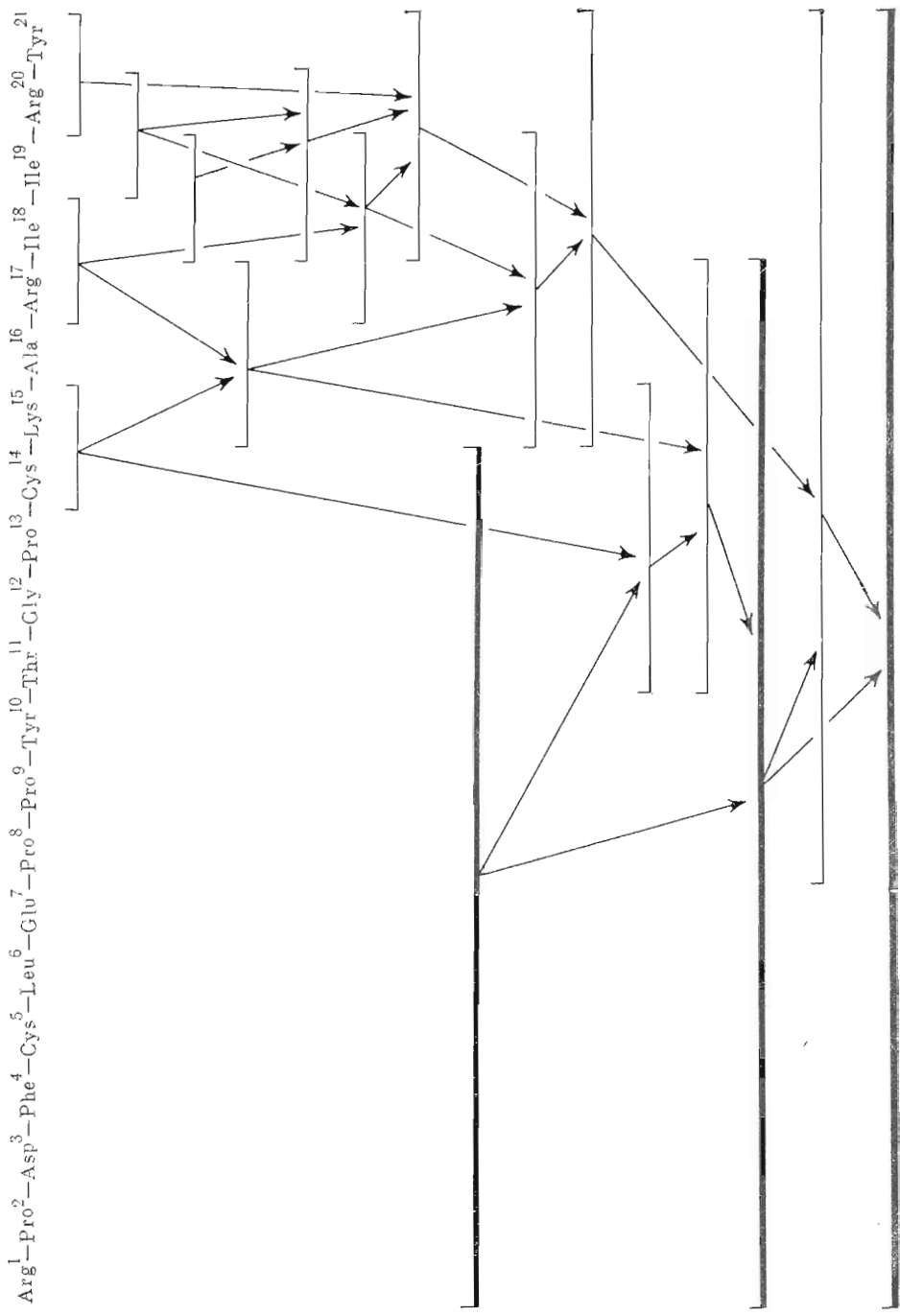


Рис. 6. Схема расчета конформаций фрагмента Arg¹ - Tyr²¹

между С-концевым и центральным участками фрагмента оказывается 41 конформация. Конформации Arg¹—Arg¹⁷, приведенные в табл. 8, являются исходными для анализа более сложного фрагмента молекулы БПТИ.

Следующий за гептадекапептидом Arg¹—Arg¹⁷ участок Ile¹⁸-Ile¹⁹-Arg²⁰-Tyr²¹ содержит объемные и конформационно лабильные остатки. Рассмотрение конформационных возможностей свободных фрагментов Ile¹⁸—Tyr²¹, Lys¹⁵—Ile¹⁹, Lys¹⁵—Tyr²¹, Pro⁸—Tyr²¹ с учетом данных для Arg¹—Arg¹⁷ позволило оценить взаимоотношение между всеми частями участка Arg¹—Tyr²¹ БПТИ. В результате задача была сведена к нахождению взаимодействий между остатками на противоположных концах фрагмента Arg¹—Tyr²¹. Результаты минимизации фрагмента Arg¹—Tyr²¹ приведены в табл. 9. Все найденные низкоэнергетические варианты представляют исключительный интерес.

Одна из низкоэнергетических конформаций фрагмента Arg¹—Tyr²¹ ($E_{\text{отн}}$ 2,1 ккал/моль) имеет шейп $e_2f_1e_5f_2e_7$. Начальный участок этой конформации (Arg¹—Pro⁹) представляет собой конформационно жесткое ядро. Далее, благодаря следующим друг за другом остаткам пролина цепь отходит от него, на участке Pro⁸—Thr¹¹ имеет развернутую форму, а на участке, включающем остатки Thr¹¹—Lys¹⁵, делает поворот на 180°. Поворот цепи в этом месте имеется во всех низкоэнергетических конформациях фрагмента Arg¹—Tyr²¹. После поворота на участке Lys¹⁵—Tyr²¹ наиболее выгодной опять становится развернутая форма основной цепи. Ее предпочтительность объясняется двумя причинами. Во-первых, как было выяснено при анализе целого ряда фрагментов разной длины, такое состояние основной цепи с точки зрения энтальпии и энтропии является одним из самых предпочтительных как для свободного фрагмента Lys¹⁵—Tyr²¹, так и для составляющих его участков. Во-вторых, при развернутой цепи создаются наиболее благоприятные условия для образования стабилизирующих контактов фрагмента Lys¹⁵—Tyr²¹ с подавляющим большинством остатков предшествующего участка.

Знаменательным фактом является совпадение описанной конформации с геометрией этого участка в трехмерной структуре молекулы БПТИ, найденной с помощью рентгеноструктурного анализа. Совпадение теоретической и кристаллической структур Arg¹—Tyr²¹ касается не только шейпа и формы основной цепи, но и численных значений конформационных параметров ϕ , ψ , ω и χ подавляющего большинства остатков.

Много общего с рассмотренной выше конформацией имеет структура с $E_{\text{отн}}$ 6,2 ккал/моль (табл. 9). На участке Arg¹—Pro⁹ она отличается лишь значением угла ψ_9 , который отвечает R-форме остатка Pro⁹. Далее, на участке Pro⁹—Ala¹⁶ цепь образует около двух витков спирали (не α -спирали) и затем вновь, как и в конформации с $E_{\text{отн}}$ 2,1 ккал/моль, приобретает развернутый характер. Таким образом, низкоэнергетические структуры Arg¹—Tyr²¹ состоят из конформационно близких, но по-разному упакованных микроглобул, которые различаются состояниями боковых цепей.

Конформации с величинами $E_{\text{отн}}$ 0, 6,2 ккал/моль и некоторые другие (табл. 9) представляют интерес в связи с результатами, полученными Крейтоном [27, 28] при исследовании процесса укладки денатурированной белковой цепи и локализации у метастабильных промежуточных продуктов дисульфидных связей. На разных стадиях окисления восстановленного белка Крейтон обнаружил продукты с S—S-мостиками между Cys⁵ и Cys¹⁴, Cys⁵ и Cys³⁰, Cys⁵ и Cys³⁸. В конформации с энергией 6,2 ккал/моль остатки Cys⁵ и Cys¹⁴ оказываются сближенными. Соответствующая конформация у фрагмента Arg¹—Cys¹⁴ была глобальной (табл. 5), а структура, близкая к экспериментальной, прониравала ей 2,8 ккал/моль. У свободного фрагмента Arg¹—Arg¹⁷ последняя оказалась уже на 3,1 ккал/моль более предпочтительной (табл. 8), а у фрагмента Arg¹—Tyr²¹ — на 4,4 ккал/моль. Поэтому можно полагать, что метастабильное конформационное состояние молекулы БПТИ с дисульфидным мостиком Cys⁵—Cys¹⁴ характерно для ранней стадии ренатурации белка. Глобальная и близкие ей низкоэнергетические структуры могут при удли-

Предпочтительные конформации фрагмента Arg¹—Arg¹⁷

Шейп	Конформации	Е _{отн.} , ккал/моль
$e_2 f_4 e_5 f_2 e_3 f$	$B_{1221}^1 - B_2 - R_{81}^3 - R_{11}^4 - R_3^5 - R_3^6 - B_7 - B_8 - B_9 - B_{10}^{10} - R_{12}^{11} - R_{12}^{11} - P_{12}^{12} - R_{13}^{13} - B_3^{14} - B_{3333}^{15} - R^{16} - B_{3222}^{17}$	1,0
$e_2 f_4 e_5 f_2 e_3$	» $B_{212} - R_{32} - P - R - B_3 - B_{3122} - B - B_{3222}$	1,3
$e_2 f_4 e_5 f_2 e_3$	» $B_{213} - R_{12} - P - R - B_3 - B_{3333} - B - B_{3222}$	4,8
$e_2 f_4 e_5 f_2 e_3$	» $B_{112} - R_{12} - P - R - B_3 - B_{3221} - B - B_{3222}$	5,2
$e_2 f_4 e_2 f_2 e_2 f$	$B_{1221} - B - R_{91} - R_{11} - R_3 - R_{32} - B_{121} - B - R - B_{212} - R_{32} - P - R - B_3 - B_{3333} - R - B_{3222}$	4,4
$e_2 f_4 e_2 f_5 e_3$	» $R_{212} - B_{32} - P - R - B_3 - B_{3333} - B - B_{3322}$	1,4
$e_2 f_4 e_2 f_5 e_2 f$	» $R_{212} - B_{32} - P - R - B_3 - B_{3333} - R - B_{3222}$	4,5
$e_2 f_4 e_2 f_6 e_2$	» $R_{212} - B_{32} - P - R - R_3 - B_{3333} - B - R_{1222}$	0
$e_2 f_4 e_2 f_6 e f$	» $R_{212} - B_{32} - P - R - R_3 - B_{3333} - R - B_{3322}$	4,4

Относительная энергия конформаций фрагмента Arg¹...Tyr¹

Шт-фп	Конформация	Еотп, ккал/моль
$e_2^1 e_4^6 e_5^2 e_6^7$	$B_{1221}^1 - B_{31}^2 - R_{31}^3 - R_{11}^4 - R_3^5 - R_{32}^6 - B_7^{121} - B_8^{131} - B_{10}^{213} - R_{11}^{32} - R_{12}^{33} - R_{13}^{34} - B_{14}^{35} - B_{15}^{3123} - B_{16}^{3123} - B_{17}^{3123} - B_{18}^{3123} - B_{19}^{3123} - B_{21}^{3123}$	2,1*
$e_2^1 e_4^5 e_5^2 e_6^7$	»	(·)
$e_2^1 e_5^2 e_6^7$	»	(·)
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4$	$B_{1221}^1 - B - R_{31} - R_{11} - R_3 - R_{32} - B_{121} - B - R - B_{212} - R_{32} - P - R - B_3 - B_{333} - R - B_{2222} - B_{32} - B_{32} - B_{322} - B_{312}$	8,5
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4 e_3^1$	»	7,5
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4 e_3^2$	»	7,9
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4 e_3^3$	»	6,2†
$e_2^1 e_4^2 e_5^2 e_6^7$	$B_{1221}^1 - B - R_{31} - R_{11} - R_3 - R_{32} - B_{121} - B - R - R_{12} - B_{32} - B_{32} - B_{32} - B_{312}$	0*
$e_2^1 e_4^2 e_6^5 e_7$	»	8,7
$e_2^1 e_4^2 e_6^7 e_3^4$	»	4,8*
$e_2^1 e_4^2 e_6^7 e_3^4 e_3^1$	»	4,3*
$e_2^1 e_4^2 e_6^7 e_3^4 e_3^2$	$B_{1221}^1 - B - R_{31} - R_{11} - R_3 - B_{121} - B - R - R_{12} - B_{32} - P - R - B_3 - B_{333} - B - B_{4222} - B_{32} - B_{12} - B_{3322} - B_{312}$	7,0*
$e_2^1 e_4^2 e_6^7 e_3^4 e_3^3$	»	10,0
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4$	»	4,2*
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4 e_3^1$	»	8,1
$e_2^1 e_4^2 e_5^2 e_6^7 e_3^4 e_3^2$	»	(·)

* Конформации, отобранные для дальнейшего анализа.

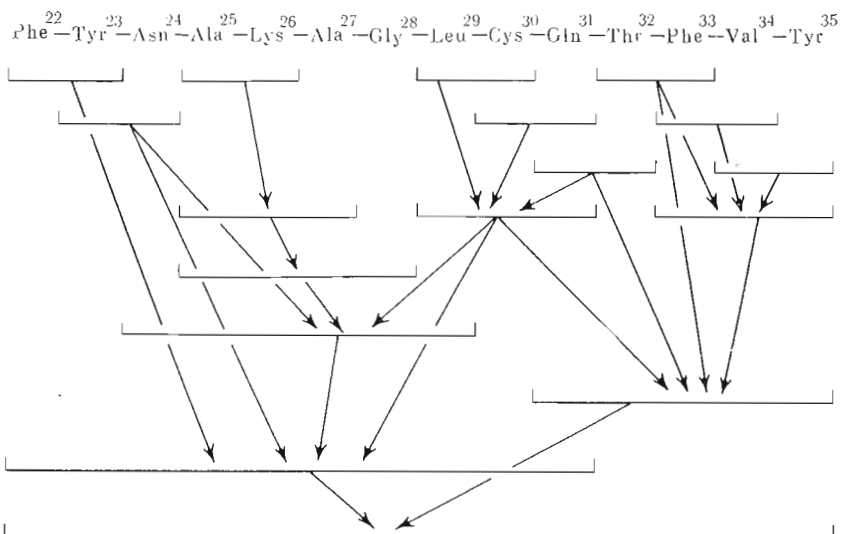


Рис. 7. Схема расчета конформаций фрагмента Phe²² - Tyr³⁵

нении цепи привести к сближенности остатков Cys⁵ и Cys³⁰, Cys⁵ и Cys⁵⁸. В связи с этим обстоятельством низкоэнергетические структуры разных типов, отмеченные в табл. 9 звездочками, оставлены для дальнейшего анализа.

Фрагмент Phe²² - Tyr³⁵ (схема расчета приведена на рис. 7). Мы не будем останавливаться на анализе конформационных состояний простых участков фрагмента, а начнем рассмотрение сразу с результатов расчета участка Phe²² - Gln³¹. Основу его нулевых приближений составили низкоэнергетические конформации гексапептида Asn²⁴ - Leu²⁹. Для N-концевого фрагмента Phe²² - Tyr²³ выбраны шесть конформационных состояний, имеющих четыре формы основной цепи шейпов *e* и *f*, полученные из анализа дипептидных фрагментов Phe²² - Tyr²³ и Tyr²³ - Asn²⁴. Нулевые приближения C-концевого участка декапептида представляют четыре самых низкоэнергетических состояния участка Cys³⁰ - Gln³¹, полученных при расчете трипептида Leu²⁹-Cys³⁰-Gln³¹. Всего для декапептида были сформированы 408 исходных для минимизации приближений. Они представляют при выбранных 17 конформациях гексапептида Asn²⁴ - Leu²⁹ все возможные 112 шейпов и 128 форм основной цепи Phe²² - Gln³¹.

Минимизация энергии выявила чрезвычайно резкую дифференциацию конформаций декапептидного фрагмента. Из большого числа рассчитанных вариантов с самыми разнообразными формами основной цепи и ориентациями боковых цепей в интервал 0-5,0 ккал/моль попадает лишь одна конформация, а в интервал 0-8,0 ккал/моль - 10 конформаций (табл. 10). В чем же причина столь однозначного конформационного детерминирования декапептида Phe²² - Gln³¹? В глобальной конформации основная цепь на участке Phe²² - Ala²⁵ имеет развернутый характер, затем она делает поворот приблизительно на 180° и на участке Ala²⁷ - Gln³¹ вновь становится развернутой. На рис. 8 приведены шейп *ef₂e₁* глобальной конформации декапептида и для сравнения шейп *ef₁ef* конформации с $E_{отн}$ 5,5 ккал/моль. Основная цепь во второй конформации почти полностью свернута и на первый взгляд представляется более компактной. Однако для последовательности Phe²² - Gln³¹ α -спираль и свернутые формы других типов оказываются неудовлетворительными, так как они не в состоянии обеспечить образование целого ряда благоприятных контактов. Из рис. 8 видно, что в случае шейпа *ef₁ef* остатки Tyr²³, Asn²⁴ и Leu²⁹, Gln³¹, отделенные друг от друга остатками с малой боковой цепью, находятся по разные стороны спирали и ни при каких положениях боковых цепей не могут взаимодействовать между собой. При конформации основной цепи с $E_{отн}$ 0 ккал/моль, напротив, имеет место сближенность всех тех остатков, которые особенно важны для межостаточных взаимодействий.

Низкоэнергетические конформации фрагмента Phe²²-Tyr²³-Asn²⁴-Ala²⁵-Lys²⁶-Ala²⁷-
Gly²⁸-Leu²⁹-Cys³⁰-Gln³¹

Шейп	Конформация	$E_{отн}$, ккал/моль
$e_3f_2e_3$	$B_{11}^{22}-B_{23}^{23}-B_{23}^{24}-R^{25}-R_{3322}^{26}-R^{27}-L^{28}-B_{11}^{29}-B_3^{30}-B_{321}^{31}$	0
$e_1e_1f_5e$	$B_{11}-R_{31}-B_{23}-R -R_{2222}-R -R -R_{32}-B_3 -B_{333}$	5,4
$f_2e_1f_5e$	$R_{23}-R_{31}-B_{23}-R -R_{2222}-R -R -R_{32}-B_3 -B_{333}$	5,3
$e_1e_1e_1$	$B_{21}-R_{31}-R_{11}-R -R_{2222}-R -R -B_{23}-R_3 -B_{213}$	5,5
$e_3f_2e_3f$	$B_{21}-B_{31}-B_{23}-R -R_{1222}-B -B -B_{32}-R_3 -B_{333}$	5,7
$e_1e_1f_4e_2$	$B_{21}-R_{31}-R_{13}-R -R_{2222}-R -R -B_{21}-B_3 -B_{333}$	5,8
$f_2e_1f_5e$	$R_{21}-B_{31}-B_{23}-R -R_{2222}-R -R -R_{32}-B_3 -B_{233}$	6,0
$e_1e_1e_2$	$B_{21}-R_{31}-R_{11}-R -R_{3322}-R -R -B_{21}-B_3 -B_{333}$	6,4
$e_1e_1f_4e_1$	$B_{21}-R_{31}-B_{13}-R -R_{2222}-R -R -B_{21}-R_3 -B_{333}$	7,4
e_3f_5e	$L_{11}-B_{31}-B_{23}-R -R_{2222}-R -R -R_{32}-B_3 -B_{333}$	7,9

Конформационный анализ свободных декапептида Phe²² — Gln³¹ и пентапептида Gln³¹ — Tyr³⁵ показал, что у первого фрагмента реальна одна конформация $B_{11}^{22}-B_{23}^{23}-B_{23}^{24}-R^{25}-R_{3322}^{26}-R^{27}-L^{28}-B_{11}^{29}-B_3^{30}-B_{321}^{31}$ шейпа $e_3f_2e_3$ (табл. 10), а у второго — несколько конформаций одной и той же формы основной цепи $B^{31}-B^{32}-B^{33}-B^{34}-B^{35}$ шейпа e_4 . Эти результаты однозначно определяют для тетрадекапептидного участка БПТИ Phe²² — Tyr³⁵ шейп пептидного скелета, форму основной цепи и конформационные состояния большинства остатков. Таким образом, задача анализа фрагмента Phe²² — Tyr³⁵ сводится к установлению наиболее предпочтительных ориентаций боковых цепей Gln³¹, Thr³², Phe³³. Расчет привел к единственной структуре фрагмента: $B_{11}^{22}-B_{23}^{23}-B_{23}^{24}-R^{25}-R_{3322}^{26}-R^{27}-L^{28}-B_{11}^{29}-B_3^{30}-B_{321}^{31}-B^{32}-B_{11}^{33}-B_2^{34}-B_{31}^{35}$ шейпа $e_3f_2e_3$.

Самым интересным фактом является совпадение найденной априорно для свободного участка Phe²² — Tyr³⁵ конформации с геометрией этого участка в структуре БПТИ, полученной с помощью рентгеноструктурного анализа. Расхождения между опытными и теоретическими значениями углов φ , ψ , ω и χ не выходят за общепринятые границы экспериментальных ошибок рентгеноструктурного анализа белков хорошего разрешения.

Фрагмент Arg¹ — Cys³⁸. На завершающем этапе расчета для фрагмента Arg¹ — Tyr²¹ были получены семь низкоэнергетических структурных вариантов, оставленных нами для последующего анализа. В двух самых выгодных конформациях с $E_{отн}$ 0 и 2,4 ккал/моль (табл. 9) С-концевой участок Lys¹⁵ — Tyr²¹ имеет полностью развернутую форму цепи. Расчет фрагмента Phe²² — Tyr³⁵ показал, что для него реальна лишь одна конформация, которая является частью антипараллельной β -структуры. Ее дальнейшее формирование представляется тем более вероятным, что в самых выгодных конформациях предшествующего фрагмента Arg¹ — Tyr²¹ С-концевой участок имеет комплементарную β -структуре развернутую форму основной цепи. В связи с этим перед расчетом фрагмента Arg¹ — Cys³⁸ были рассмотрены конформационные возможности фрагментов Cys¹⁴ — Tyr³⁵, Tyr²⁵ — Gly³⁷ и Cys¹⁴ — Cys³⁸.

Исходные варианты фрагмента Cys¹⁴ — Tyr³⁵ составлены из конформационных состояний октапептида Cys¹⁴ — Tyr²¹, которые он принимает в семи оставленных для дальнейшего анализа низкоэнергетических структурах Arg¹ — Tyr²¹, и в одной конформации тетрадекапептида Phe²² — Tyr³⁵. Всего было рассчитано 14 вариантов, так как при включении Cys¹⁴ — Tyr²¹ в последовательность Cys¹⁴ — Tyr³⁵ снимается энергетическое возмещение R- и B-форм остатка Tyr²¹ и число вариантов увеличивается. Результаты расчета полностью подтвердили сделанное выше предположение о предпочтительности встраивания остатков Cys¹⁴ — Tyr²¹ в β -структуру. В результате минимизации энергии при варьировании двугранных углов

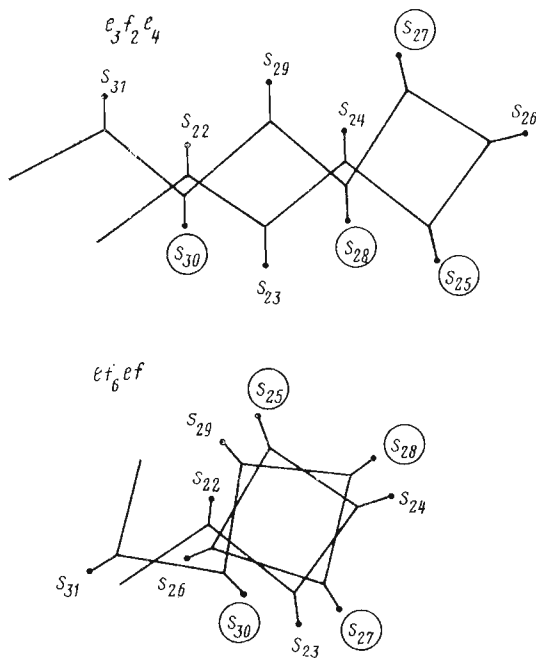


Рис. 8. Шейцы основной цепи низкоэнергетических конформаций фрагмента Phe²² — Gln³¹ (кружками выделены остатки с малыми боковыми цепями)

основных и боковых цепей фрагментов Cys¹⁴ — Phe²² и Ala²⁵ — Tyr³⁵ выяснилось, что для фрагмента Cys¹⁴ — Tyr³⁵ реальны только два конформационных состояния — именно те, которые на участке Cys¹⁴ — Tyr²¹ имеют развернутую форму и входят в самые предпочтительные структуры фрагмента Arg¹ — Tyr²¹ ($E_{\text{отн}}$ 0 и 2,1 ккал/моль; табл. 9). Другие структурные варианты по крайней мере на 12 ккал/моль менее стабильны.

Рассмотрим теперь результаты анализа фрагмента Tyr³⁵—Gly³⁶—Gly³⁷, завершающего построение β -структуры. Его нулевые приближения составлены из *B*- и *R*-форм остатка Tyr³⁵ с $\chi^1 \sim 60, 180$ и -60° и *B*-, *R*-, *L*- и *P*-форм остатков Gly³⁶ и Gly³⁷. Они имеют 32 формы основной цепи и разбиваются на четыре шейпа (по 24 варианта в каждом). Энергетическое распределение конформаций свидетельствует о большой лабильности фрагмента Tyr³⁵ — Gly³⁷. Почти все его конформационные состояния попадают в узкий интервал энергии 0—3,0 ккал/моль и, следовательно, являются вероятными. Поэтому определить конформационные состояния фрагмента Arg¹ — Cys³⁸ обычным путем нецелесообразно. Выбранный нами путь состоял в последовательном присоединении к Cys¹⁴ — Tyr³⁵ остатков лабильного участка Gly³⁶ — Cys³⁸ и в выяснении их наилучших конформационных состояний. С этой целью вначале строились сечения потенциальной поверхности $\psi_{35} - \psi_{36}$ при фиксированных значениях остальных углов фрагмента Cys¹⁴ — Tyr³⁵, а затем $\psi_{36} - \psi_{37}$ и $\psi_{37} - \psi_{38}$ при фиксированных параметрах фрагментов Cys¹⁴ — Gly³⁶ и Cys¹⁴ — Gly³⁷ соответственно. Каждый раз в области самой низкой энергии выбирались углы, которые фиксировались при следующем шаге. На каждом этапе и после окончания всей процедуры проводили минимизацию энергии при варьировании двугранных углов вначале у вновь присоединяемых остатков, а потом всего фрагмента.

В глобальной структуре Cys¹⁴ — Tyr³⁵ самой предпочтительной для Tyr³⁵ — Cys³⁸ оказалась конформация $B_{21}^{35} - R^{36} - P^{37} - B^{38} (R^{38})$ с развернутой формой основной цепи. Это не удивительно, поскольку такое состояние стерически наиболее комплементарно участку Cys¹⁴ — Ile¹⁸, входящему в противостоящий участок β -структуры.

При состыковке фрагмента $\text{Tyr}^{35} - \text{Cys}^{38}$ с фрагментом $\text{Cys}^{14} - \text{Tyr}^{35}$, имеющим конформацию с $E_{\text{отн}}$ 1,8 ккал/моль, наиболее выгодной также оказалась развернутая форма основной цепи. Однако в этом случае взаимодействия $\text{Tyr}^{35} - \text{Cys}^{38}$ с предшествующими остатками несколько менее эффективны и поэтому разница в энергии двух возможных конформационных вариантов фрагмента $\text{Cys}^{14} - \text{Cys}^{38}$ увеличилась до 3,0 ккал/моль. Кроме того, структура $\text{Cys}^{14} - \text{Tyr}^{35}$ с $E_{\text{отн}}$ 1,8 ккал/моль не отвечает важному геометрическому критерию, о котором мы будем говорить ниже.

Молекула трипсинового ингибитора содержит шесть остатков цистеина, которые образуют в нативной структуре три дисульфидных мостика: $\text{Cys}^5 - \text{Cys}^{55}$, $\text{Cys}^{14} - \text{Cys}^{38}$ и $\text{Cys}^{30} - \text{Cys}^{51}$. Исследование пространственного строения БПТИ проводится нами для линейной последовательности без каких-либо предположений о локализации S—S-связей. В соответствии с термодинамической концепцией Анфинсена [29] мы полагаем, что сближенность соответствующих остатков Cys является следствием строго детерминированного процесса укладки пептидной цепи в наиболее предпочтительную по энергии межостаточных взаимодействий структуру. Поэтому рассматриваемые расчетные данные о конформационных состояниях линейного и циклического фрагментов $\text{Cys}^{14} - \text{Cys}^{38}$ представляют интерес не только в изучении трехмерной структуры БПТИ, но и в значительно более общем плане.

Рассмотрев пространственное строение линейного фрагмента $\text{Cys}^{14} - \text{Cys}^{38}$, мы подошли к одному из наиболее интересных этапов исследования - выяснению стерической предрасположенности найденной конформации к образованию дисульфидной связи между Cys^{14} и Cys^{38} . Для образования дисульфидной связи между этими остатками в низкоэнергетических конформациях $\text{Cys}^{14} - \text{Cys}^{38}$ ($E_{\text{отн}}$ 0 и 3,0 ккал/моль) была выбрана процедура, предложенная в работе [8]. Найденные геометрические параметры дисульфидного мостика в глобальной конформации фрагмента приведены

на рис. 9. Связи $\text{C}_{(14)}^{\beta} - \text{S}_{(14)}$ и $\text{C}_{(38)}^{\beta} - \text{S}_{(38)}$ сближены и имеют *goss*-ориен-

тацию, а угол вращения вокруг связи $\text{S}_{(14)} - \text{S}_{(38)}$ (χ^{SS}) равен 100° , что отвечает вкладу торсионной составляющей менее 1,0 ккал/моль [30—32]. Конформационная энергия фрагмента $\text{Cys}^{14} - \text{Cys}^{38}$ с S—S-мостиком практически совпадает с энергией линейного фрагмента. Следовательно, образование дисульфидной связи не нарушило стабилизирующих контактов, сложившихся в линейной конформации с величиной $E_{\text{отн}}$ 0 ккал/моль.

У другой низкоэнергетической конформации линейной последовательности $\text{Cys}^{14} - \text{Cys}^{38}$ с $E_{\text{отн}}$ 3,0 ккал/моль остаток Cys^{14} имеет *R*-форму основной цепи. В этом случае боковая цепь Cys^{14} отвернута от Cys^{38} и расстояние между атомами $\text{S}_{(14)}$ и $\text{S}_{(38)}$ составляет 5,5—7,0 Å, в зависимости от значения угла χ^1 . При наложении стягивающего потенциала образование дисульфидной связи $\text{S}_{(14)} - \text{S}_{(38)}$ происходит за счет перехода *R*-формы основной цепи Cys^{14} в *B*-форму, т. е. путем инверсии конформации $\text{Cys}^{14} - \text{Cys}^{38}$ с $E_{\text{отн}}$ 3,0 ккал/моль в конформацию с $E_{\text{отн}}$ 0 ккал/моль. Таким образом, только глобальная структура линейной последовательности $\text{Cys}^{14} - \text{Cys}^{38}$ обнаруживает предрасположенность к созданию дисульфидной связи между остатками Cys^{14} и Cys^{38} .

Приход к единственной циклической структуре $\text{Cys}^{14} - \text{Cys}^{38}$ автоматически исключает из дальнейшего рассмотрения все конформационные состояния фрагмента $\text{Arg}^1 - \text{Tyr}^{21}$, кроме одного с энергией $E_{\text{отн}}$ 2,1 ккал/моль (табл. 9). Следовательно, расчет структуры участка $\text{Arg}^1 - \text{Cys}^{38}$ сводится к поиску наилучшей упаковки друг относительно друга всего двух конформаций фрагментов $\text{Arg}^1 - \text{Cys}^{14}$ и $\text{Cys}^{14} - \text{Cys}^{38}$. Объединение их в структуру приводит, однако, к резкому повышению энергии, вызванному образованием многочисленных неблагоприятных контактов между найденными конформациями участков $\text{Phe}^4 - \text{Gly}^{12}$ и $\text{Phe}^{33} - \text{Gly}^{28}$. Причина стерических затруднений заключается в том, что расчетная конформация $\text{Cys}^{14} - \text{Cys}^{38}$ не является классической β -структурой. Нисходящая и восходящая ветви в ней располагаются не строго антипараллельно, а, обла-

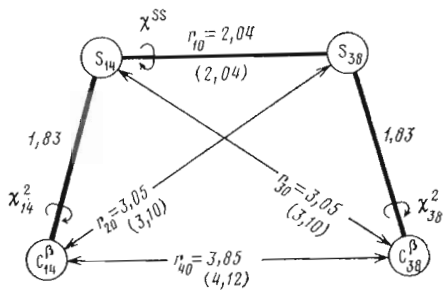


Рис. 9

Рис. 9. Геометрия дисульфидного мостика в *gauche*-конфигурации. В скобках приведены расчетные значения параметров мостика $\text{Cys}^{14}-\text{Cys}^{38}$

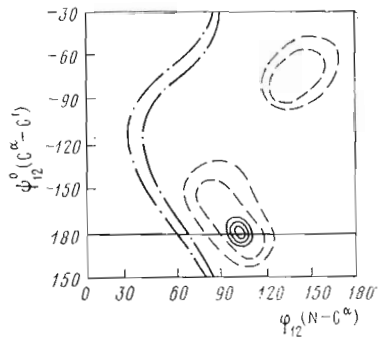


Рис. 10

Рис. 10. Сечение φ - ψ остатка Gly^{12} (---), фрагмента $\text{Arg}^1-\text{Cys}^{14}$ (- - -) и фрагмента $\text{Arg}^1-\text{Cys}^{38}$ (-)

дая несколько изогнутой формой, огибают друг друга и образуют сегмент неправильной двойной спирали с большим шагом. В результате сближенным с участком $\text{Phe}^4-\text{Gly}^{12}$ на расстояния, значительно меньше сумм ван-дер-ваальсовых радиусов, оказывается наиболее удаленный по цепи участок $\text{Phe}^{33}-\text{Gly}^{37}$.

Возникшие стерические осложнения нельзя снять без разрушения циклической β -структуры. Поэтому нами предпринята попытка ликвидировать стерические препятствия путем изменения геометрии по возможности минимального линейного фрагмента $\text{Arg}^1-\text{Cys}^{14}$. С этой целью для входящих в него остатков в поле участка $\text{Arg}^1-\text{Cys}^{38}$ были получены семейства сечений потенциальной поверхности φ - ψ для тех областей двугранных углов, которые соответствуют глобальной конформации основной цепи $\text{Arg}^1-\text{Cys}^{14}$. Подобная процедура была выполнена ранее при исследовании конформационных возможностей тех же остатков и в той же структуре свободного фрагмента $\text{Arg}^1-\text{Cys}^{14}$ (рис. 5). Оказалось, что варьирование в указанных пределах двугранных углов у большинства остатков не приводит к понижению конформационной энергии фрагмента $\text{Arg}^1-\text{Cys}^{38}$, т. е. не снимает стерические напряжения. Исключения составили углы φ , ψ остатка Gly^{12} , угол ψ остатка Leu^6 и угол φ остатка Glu^7 . Последовательные изменения их значений, направляемые и контролируемые в каждом случае сечениями потенциальной поверхности $\varphi_{12}-\psi_{12}$ и $\psi_6-\psi_7$, и минимизация энергии привели в конечном счете к снятию всех неблагоприятных контактов между $\text{Arg}^1-\text{Cys}^{14}$ и $\text{Cys}^{14}-\text{Cys}^{38}$. Форма основной цепи глобальной конформации $\text{Arg}^1-\text{Cys}^{14}$ осталась при этом прежней, и все изменения двугранных углов, коснувшиеся главным образом остатков Gly^{12} , Leu^6 и Glu^7 , произошли в пределах низкоэнергетических областей конформационных карт φ - ψ , полученных для свободного фрагмента $\text{Arg}^1-\text{Cys}^{14}$. Иными словами, при снятии стерических осложнений удалось практически полностью сохранить у тетрадекапептида согласованность всех межостаточных взаимодействий. На рис. 10 приведены совмещенные карты φ - ψ свободного и встроенного в $\text{Arg}^1-\text{Cys}^{14}$ и $\text{Arg}^1-\text{Cys}^{38}$ остатка Gly^{12} . Как видно из рисунка, состояния остатка в обоих фрагментах соответствуют области низкой энергии монопептида. У $\text{Arg}^1-\text{Cys}^{14}$ конформационная свобода Gly^{12} уменьшается, и он может принимать значения φ , ψ только в двух областях — вблизи $140^\circ, -80^\circ$ и $100^\circ, 180^\circ$, причем первый минимум предпочтительнее второго на 1–2 ккал/моль. На участке $\text{Arg}^1-\text{Tyr}^{21}$ различие в энергии двух потенциальных ям возрастает до ~ 5 ккал/моль, так как при значениях φ , $\psi \sim 140^\circ, -80^\circ$ у Gly^{12} имеет место сближенность многих остатков фрагментов $\text{Arg}^1-\text{Cys}^{14}$ и $\text{Cys}^{14}-\text{Tyr}^{21}$ и образование между ними более эффективных, чем при $\sim 100^\circ, 180^\circ$, стабилизирующих взаимодействий. Однако эти

взаимодействия, как было отмечено выше, невозможны у фрагмента

Arg¹ — Cys³⁸ при наличии циклической структуры Cys¹⁴ — Cys³⁸. Поэтому минимум в области $\varphi, \psi \sim 140, -80^\circ$ становится запрещенным. При других значениях φ, ψ , соответствующих второму минимуму ($\sim 100, 180^\circ$), частично сохраняются старые контакты между Arg¹ — Cys¹⁴ и Cys¹⁴ — Tyr²¹ и, кроме того, образуется целый ряд новых, значительно более эффективных контактов этого участка с β -структурой. Таким образом, несмотря на возникшие затруднения мы опять-таки приходим к согласованной в отношении всех средних и дальних взаимодействий конформации фрагмента Arg¹ — Cys³⁸.

О согласованности взаимодействий в молекуле БПТИ можно судить по данным, приведенным на рис. 11. То обстоятельство, что относительная энергия фрагментов во всех случаях находится в низкоэнергетической половине указанных интервалов или даже равняется нулю, свидетельствует, во-первых, о весьма совершенной согласованности в белке всех взаимодействий и, во-вторых, о том, что расчет молекулы трипсинного ингибитора проводится с достаточно большим «запасом прочности».

Расчитанная для Arg¹ — Cys³⁸ структура совпадает с геометрией этого участка в молекуле БПТИ, полученной с помощью рентгеноструктурного анализа [2, 3].

Фрагмент Arg¹ — Ser⁴⁷. Вначале были последовательно рассмотрены конформационные возможности ряда свободных ди-, три-, тетра-, пента- и октапептидов участка Arg³⁹ — Ser⁴⁷. Результаты анализа фрагментов Arg³⁹ — Lys⁴¹ и Ala⁴⁰ — Ser⁴⁷, а также найденная для Arg¹ — Cys³⁸ конформация с дисульфидной связью Cys¹⁴ — Cys³⁸ были затем использованы для предсказания структуры Arg¹ — Ser⁴⁷. Нонапептидная последовательность Arg³⁹ — Ser⁴⁷ включает объемные и локальные остатки Arg³⁹, Lys⁴¹, Arg⁴², Lys⁴⁶, несущие целочисленные положительные заряды и остатки Asn⁴³, Asp⁴⁴, Phe⁴⁵, также обладающие достаточно объемными и подвижными боковыми цепями. Большая конформационная свобода остатков фрагмента Arg³⁹ — Ser⁴⁷ и их способность образовывать разные по своей природе, энергетически весьма эффективные взаимодействия потребовали расчета многочисленных фрагментов постепенно увеличивающейся длины, каждый из которых имел представительный набор структурных вариантов различных форм основной цепи и шейпов. Расчет показал, что фрагмент Arg³⁹ — Ser⁴⁷ в отношении средних взаимодействий является конформационно лабильным участком БПТИ. Хотя выполненный анализ не выявил для этого участка какой-либо одной формы основной цепи, обладающей бесспорной предпочтительностью по взаимодействиям между входящими в его состав остатками, тем не менее он позволил сократить количество возможных вариантов для последующей оценки дальних взаимодействий, что сделало реальным решение более сложной задачи. В частности, расчет показал, что для Arg³⁹ — Ser⁴⁷ более предпочтителен развернутый характер основной цепи. Чрезвычайно мала вероятность образования на этом участке α -спиральных структур, в которых оказываются сближенными все одноименно заряженные остатки аргинина и лизина.

Нами также выполнен конформационный анализ свободных пентапептида Gly³⁶ — Ala⁴⁰ и гептапептида Tyr³⁵ — Lys⁴¹, у которых были рассмотрены состояния всех возможных форм основной цепи соответственно 16 и 64 шейпов. Оказалось, что взаимодействия между остатками в пределах этих фрагментов также не приводят к образованию жестких структур, изоэнергетичными являются практически все структурные варианты с *B*- и *R*-состояниями остатков.

Результаты расчета позволяют предположить, что детерминация конформационного состояния последовательности БПТИ на участке Cys³⁸ — Ser⁴⁷ осуществляется главным образом за счет дальних взаимодействий, обеспечивающих при сборке нативной структуры белка плотную упаковку этого фрагмента с ранее сложившейся конформацией предшествующего участка Arg¹ — Cys³⁸. Вследствие очень большого количества низкоэнергетических по средним взаимодействиям конформационных состояний

Cys³⁸ — Ser⁴⁷ дальнейший поиск наиболее предпочтительной структуры участка Arg¹ — Ser⁴⁷ заключался в последовательной генерации цепи, начиная с Cys³⁸, и количественной оценке взаимодействий присоединяемых остатков с уже найденной для Arg¹ — Cys³⁸ структурой. При наращивании белковой цепи сначала до Lys⁴¹, затем до Arg⁴², Asn⁴⁴, Lys⁴⁶ и, наконец, до Ser⁴⁷ учитывались конформационные состояния остатков фрагмента Cys³⁸ — Ser⁴⁷, отвечающих низкоэнергетическим вариантам свободных фрагментов Arg³⁹ — Lys⁴¹, Ala⁴⁰ — Ser⁴⁷ и Tyr³⁵ — Lys⁴¹.

При рассмотрении участка Arg¹ — Lys⁴¹ оценены взаимодействия большого набора конформационных состояний Cys³⁸ — Lys⁴¹ с жесткой структурой Arg¹ — Cys³⁸. У присоединяемого фрагмента учитывались *B*- и *R*-формы остатка Cys³⁸ и все комбинации *B*-, *R*- и *L*-форм основной цепи остатков Arg³⁹, Ala⁴⁰ и Lys⁴¹, составляющих максимально возможные для гетрапептида восемь шейпов пептидного скелета, а также полученные из предшествующего анализа конформационные состояния боковых цепей Arg³⁹ и Lys⁴¹. Найденные величины относительной энергии конформаций участка Arg¹ — Lys⁴¹, попадающие в интервал 0—20,0 ккал/моль, приведены в табл. 11.

Расчет свидетельствует о высокой чувствительности конформационной энергии Arg¹ — Lys⁴¹ к структуре фрагмента Cys³⁸ — Lys⁴¹. Из многих сотен рассмотренных вариантов в указанный широкий энергетический интервал попали лишь восемь конформационных состояний Arg¹ — Lys⁴¹. Низкоэнергетическими являются только две конформации с $E_{\text{отн}}$ 0 и 2,2 ккал/моль. Они имеют одинаковые (*B*³⁸—*L*³⁹) состояния дипептида Cys³⁸—Arg³⁹, которые обеспечивают поворот основной цепи, сближенность фрагмента Arg³⁹ — Lys⁴¹ с предшествующими остатками и образование с ними многочисленных стабилизирующих контактов.

В расчете фрагмента Arg¹ — Arg⁴² приняты во внимание все конформации Arg¹ — Lys⁴¹, приведенные в табл. 11. Для Arg⁴² варьировали допустимые по средним взаимодействиям положения его боковой цепи при *B*- и *R*-формах остатка. В интервал 0—20,0 ккал/моль попали 28 структурных вариантов. Самыми предпочтительными оказались структуры Arg¹ — Arg⁴²

с состоящими *S*-концевого пентапептида $B^{38}\text{-}L_{3322}\text{-}B^{40}\text{-}B_{3222}^{41}\text{-}R_{3222}^{42}$

($E_{\text{отн}}$ 0 ккал/моль) шейпа fe_3 и $B^{38}\text{-}L_{3322}^{39}\text{-}B^{40}\text{-}R_{3222}^{41}\text{-}B_{3222}^{42}$ ($E_{\text{отн}}$ 0,1 ккал/моль)

шейпа fe_{2f} .

Исходные структурные варианты фрагмента Arg¹ — Asn⁴⁴ сформированы из 28 конформаций фрагмента Arg¹ — Arg⁴² с относительной энергией от 0 до 20,0 ккал/моль и *B*-*R*- и *R*-*R*-форм дипептидного участка Asn⁴³ — Asn⁴⁴ с ориентациями боковых цепей, выгодными по средним взаимодействиям. Результаты минимизации энергии при варьировании двугранных углов основной и боковой цепей лабильного участка Cys³⁸ — Asn⁴⁴ для низкоэнергетических вариантов (0—12,0 ккал/моль) представлены в табл. 11. Расчет выявил у фрагмента Arg¹ — Asn⁴⁴ лишь одну предпочтительную конформацию, которая отвечает лучшим состояниям фрагментов Arg¹ — Arg⁴² и Arg¹ — Lys⁴¹, а также низкоэнергетическому по межостаточным взаимодействиям состоянию свободного участка Cys³⁸ — Asn⁴⁴. Однако предпочтительность глобальной структуры обусловлена прежде всего наилучшей комплементарностью между конформацией Cys³⁸ — Asn⁴⁴ шейпа fe_{3fe} и микроглобулой, образованной фрагментом Arg¹ — Cys³⁸. Только в этом случае все остатки первого фрагмента, не нарушая средних взаимодействий, образуют многочисленные стабилизирующие контакты с удаленными остатками, суммарный энергетический эффект этих взаимодействий составляет —27,1 ккал/моль.

Распределение конформаций фрагмента Arg¹ — Lys⁴⁶ по энергии не претерпевает серьезных изменений при удлинении белковой цепи на два остатка (Phe⁴⁵—Lys⁴⁶). Самая выгодная структура Arg¹ — Lys⁴⁶ (fe_{3fe} ; $E_{\text{отн}}$ 0 ккал/моль; табл. 11) представляет собой комбинацию глобальной конформации Arg¹ — Asn⁴⁴ и $B_{31}\text{-}R_{3222}$ -состояния Phe⁴⁵—Lys⁴⁶. В этом случае при-

Относительная энергия (ккал/моль) конформаций фрагментов Arg¹-Lys⁴¹, Arg¹-Asn⁴⁴, Arg¹-Lys⁴⁴, Arg¹-Lys⁴⁶

Arg ¹ -Lys ⁴¹			Arg ¹ -Asn ⁴⁴			Arg ¹ -Lys ⁴⁶		
Шейп	Форма	Е _{отн}	Шейп	Форма	Е _{отн}	Шейп	Форма	Е _{отн}
<i>fef</i>	<i>B</i> ³⁸ - <i>L</i> ³⁹ - <i>R</i> ⁴⁰ - <i>B</i> ⁴¹	0	<i>fe₃fe</i>	<i>B</i> ³⁸ - <i>L</i> ³⁹ - <i>B</i> ⁴⁰ - <i>B</i> ⁴¹ - <i>R</i> ⁴² - <i>B</i> ⁴³ - <i>B</i> ⁴⁴	0	<i>fe₃fe₃</i>	<i>B</i> ³⁸ - <i>L</i> ³⁹ - <i>B</i> ⁴⁰ - <i>B</i> ⁴¹ - <i>R</i> ⁴² - <i>B</i> ⁴³ - <i>B</i> ⁴⁴ - <i>B</i> ⁴⁵ - <i>R</i> ⁴⁶	0
<i>fe₂</i>	<i>B</i> - <i>L</i> - <i>B</i> - <i>B</i>	2,2	<i>fe₃e₂</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>R</i> - <i>B</i> - <i>B</i> - <i>B</i>	8,0	<i>fe₂fe₂f</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>B</i> - <i>B</i> - <i>R</i> - <i>R</i>	9,0
<i>e₂f</i>	<i>B</i> - <i>B</i> - <i>R</i> - <i>B</i>	12,4	<i>fef₄</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>R</i> - <i>R</i> - <i>R</i> - <i>B</i>	8,3	<i>fe₂fe₃</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>B</i> - <i>B</i> - <i>R</i>	9,4
<i>e₁f₂</i>	<i>B</i> - <i>R</i> - <i>R</i> - <i>B</i>	12,8	<i>fe₂ef</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>B</i>	9,6	<i>fefef₄</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>R</i> - <i>R</i> - <i>R</i>	9,8
<i>efe</i>	<i>B</i> - <i>B</i> - <i>L</i> - <i>B</i>	15,5	<i>fe₂ef₂f</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>B</i> - <i>R</i> - <i>B</i>	10,2	<i>fe₂fe₃se</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>R</i> - <i>B</i> - <i>R</i>	10,2
	<i>B</i> - <i>R</i> - <i>R</i> - <i>B</i>	16,7	<i>fe₂e₃</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>B</i> - <i>B</i> - <i>B</i>		10,6	<i>fe₂e₃f</i>	
<i>e₃</i>	<i>B</i> - <i>B</i> - <i>B</i> - <i>B</i>	16,7	<i>fe₂efe</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>B</i> - <i>B</i>	11,2	<i>fe₂efef</i>	<i>B</i> - <i>L</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>B</i> - <i>R</i> - <i>R</i>	10,8
	<i>B</i> - <i>R</i> - <i>L</i> - <i>B</i>	19,2	<i>e₁fe</i>	<i>B</i> - <i>B</i> - <i>B</i> - <i>B</i> - <i>B</i> - <i>B</i>		11,3	<i>fe₂e₃f₃</i>	

Примечания: 1) во всех конформациях фрагментов Arg¹-Lys⁴¹, Arg¹-Asn⁴⁴ и Arg¹-Lys⁴⁶ участок Arg¹-Cys⁸⁸ имеет одинаковую структуру; 2) различия в величинах при *B*- и *R*-формах *C*-концевых остатков фрагментов не превышают 1,0 ккал/моль.

соединенный участок оказывается сближенным с тетрапептидным фрагментом Arg²⁰—Tyr²³. Стабилизирующие контакты с удаленными в последовательности остатками образует также Lys⁴⁶.

Таким образом, постепенная генерация белковой последовательности и анализ на каждом этапе большого числа конформационных состояний фрагментов Arg¹—Lys⁴¹, Arg¹—Arg⁴², Arg¹—Asn⁴⁴ и Arg¹—Lys⁴⁶ привели в конечном счете к выделению из многочисленных, практически изоэнергетичных структурных вариантов свободного участка Cys³⁸—Ser⁴⁷ единственной реальной трехмерной структуры фрагмента Arg¹—Lys⁴⁶. Ее однозначная детерминация определяется предпочтительностью участка Cys³⁸—Lys⁴⁶ образовывать стабилизирующие контакты с удаленными по цепи остатками и согласованностью между всеми средними и дальними взаимодействиями. Присоединение остатка Ser⁴⁷, обладающего небольшой боковой цепью, естественно, не может сколько-нибудь существенно повлиять на энергетическое распределение конформаций фрагмента Arg¹—Ser⁴⁷ по сравнению с фрагментом Arg¹—Lys⁴⁶. Учитывая это обстоятельство и больший разрыв в энергии глобальной и следующих за ней структур, для дальнейшего анализа пространственного строения молекулы БПТИ нами оставлена лишь конформация Arg¹—Lys⁴⁶ типа je_3je_3 ($E_{отн}$ 0 ккал/моль, табл. 11).

Фрагмент Ala⁴⁸—Ala⁵⁸. Конформационный анализ ундекапептидного участка БПТИ Ala⁴⁸—Ala⁵⁸ выполнен по схеме, согласно которой вначале рассчитывались шесть перекрывающихся дипептидных фрагментов и фрагменты Ala⁴⁸—Asp⁵⁰, Asp⁵⁰—Met⁵² и Cys⁵⁵—Ala⁵⁸. Полученные для них данные явились исходными для расчета пентапептида Ala⁴⁸—Met⁵², который, в свою очередь, вместе с дипептидами Met⁵²—Arg⁵³ и Arg⁵³—Thr⁵⁴ стал базовым для анализа гептапептида Ala⁴⁸—Thr⁵⁴. Результаты его расчета и данные по фрагментам Thr⁵⁴—Cys⁵⁵ и Cys⁵⁵—Ala⁵⁸ позволили оценить конформационные возможности ундекапептидного участка Ala⁴⁸—Ala⁵⁸.

Структурные варианты фрагмента Ala⁴⁸—Thr⁵⁴ составили 96 конформаций с 42 различными формами основной цепи, принадлежащими 28 шейпам. Результаты минимизации их энергии свидетельствуют о том, что увеличение длины цепи от пента- до гептапептида сопровождается резкой детерминацией α -спиральной формы основной цепи. В табл. 12 приведено энергетическое распределение конформаций в соответствии с их формой и шейпом пептидного скелета; не включенные в таблицу структурные варианты 34 форм 24 шейпов имеют энергию $>8,0$ ккал/моль. В широкий энергетический интервал 0—5,0 ккал/моль попадают 25 конформаций с одной и той же формой основной цепи, $R^{48}R^{49}R^{50}R^{51}R^{52}R^{53}R^{54}$ (f_6), и лишь одна конформация родственного типа f_6e с величиной $E_{отн}$ 4,4 ккал/моль. Таким образом, расчет показывает, что единственной реальной формой основной цепи гептапептида Ala⁴⁸—Thr⁵⁴ является α -спираль. В чем же причина столь очевидной предпочтительности этой структуры у данной аминокислотной последовательности по сравнению с большим числом рассмотренных для нее структурных вариантов? α -Спираль — структура с наиболее плотно упакованными элементами основной цепи, и в этом отношении она обладает предпочтительностью по сравнению со всеми другими, особенно с полностью развернутыми формами пептидного остова. Однако этот фактор не может быть доминирующим в любой аминокислотной последовательности. В конечном счете форму основной цепи определяют природа остатков и порядок их расположения. Для белков это чаще всего не α -спираль. Выгодность α -спирали у пентапептида Ala⁴⁸—Met⁵² обусловлена согласованностью всех межкостаточных взаимодействий. Наиболее весомый вклад в стабилизацию этой структуры вносят взаимодействия между элементами основной цепи. В случае Ala⁴⁸—Met⁵² именно они сыграли определяющую роль в энергетической дифференциации различных конформаций, поскольку ни один из альтернативных структурных вариантов не обладает перед α -спиралью предпочтительностью в отношении взаимодействий боковых цепей. Присоединение к пентапептиду остатков Arg⁵³—Thr⁵⁴ увеличивает энергетический и энтропийный разрыв между α -спиральной и другими возможными формами основной цепи. Это вызва-

Энергетическое распределение предпочтительных конформаций фрагмента
Ala⁴⁸-Glu⁴⁹-Asp⁵⁰-Cys⁵¹-Met⁵²-Arg⁵³-Thr⁵⁴

Шейп	форма	интервал энергии, ккал/моль							
		>8	8—6	6—5	5—4	4—3	3—2	2—1	1—0
f_6	<i>R-R-R-R-R-R</i>	11	11	7	5	7	4	4	5
f_{3e}	<i>R-R-R-R-B-R</i>	1	2	1	1	—	—	—	—
	<i>R-R-R-B-L-R</i>	5	—	—	—	—	—	—	—
f_{4ef}	<i>R-R-R-B-R-R</i>	3	1	1	—	—	—	—	—
	<i>R-R-B-L-R-R</i>	5	—	—	—	—	—	—	—
f_{4e2}	<i>R-R-R-B-B-R</i>	3	1	1	—	—	—	—	—
	<i>R-R-B-L-B-R</i>	4	1	—	—	—	—	—	—
	<i>R-R-R-R-L-R</i>	5	—	—	—	—	—	—	—

но не только более эффективными контактами между элементами пептидной цепи, но и взаимодействиями боковых цепей, которые в α -спиральной форме основной цепи гептапептида Ala⁴⁸-Thr⁵⁴ становятся более выгодными, чем при других формах. В полностью свернутой структуре типа f_6 ($E_{отн}$ 0 ккал/моль) положительно заряженная боковая цепь Arg⁵³ пространственно сближена с боковыми цепями Glu⁴⁹ и Asp⁵⁰, несущими отрицательные заряды, и образует с ними исключительно эффективные контакты (−16,2 ккал/моль), взаимодействуя одновременно и с основными цепями этих остатков (−3,9 ккал/моль). Рассчитанные конформационные параметры Ala⁴⁸-Thr⁵⁴ совпадают с опытными значениями практически с точностью рентгеноструктурного анализа белков с хорошим разрешением. Совпадение имеет место в отношении всех двугранных углов основной цепи (ϕ , ψ , ω) и углов χ^1 , χ^2 боковых цепей. Энергия расчетной конформации, полностью отвечающая геометрии участка Ala⁴⁸-Thr⁵⁴ в кристалле, лишь немного уступает глобальной структуре свободного гептапептида. Таким образом, конформационный анализ Ala⁴⁸-Thr⁵⁴ и сопоставление результатов расчета с экспериментальными данными приводит к заключению об образовании у гептапептидного фрагмента за счет взаимодействий между его остатками жесткой структуры, которая без существенных конформационных перестроек входит в нативную трехмерную структуру белка.

Для дальнейшего расчета оставлены 12 конформаций фрагмента Ala⁴⁸-Thr⁵⁴ с формой основной цепи $R^{48}-R^{49}-R^{50}-R^{51}-R^{52}-R^{53}-R^{54}$ (f_6) и по одной конформации с формами $R^{48}-R^{49}-R^{50}-R^{51}-R^{52}-B^{53}-R^{54}$ (f_{3e}), $R^{48}-R^{49}-R^{50}-R^{51}-B^{52}-R^{53}-R^{54}$ (f_{4ef}) и $R^{48}-R^{49}-R^{50}-R^{51}-B^{52}-B^{53}-R^{54}$ (f_{4e2}). Нулевые приближения ундекапептида Ala⁴⁸-Ala⁵⁸ получены стыковкой указанных конформаций гептапептида Ala⁴⁸-Thr⁵⁴ с 37 конформациями тетрапептида Cys⁵⁵-Ala⁵⁸. Поскольку в расчете первого фрагмента для Thr⁵⁴ учитывалось только одно состояние *R* (*B*-состояние для *C*-концевого остатка изоэнергетично *R*, а *L* отвечает высокоэнергетическим конформациям дипептидов Arg⁵³-Thr⁵⁴ и Thr⁵⁴-Cys⁵⁵), в расчете ундекапептида принимались во внимание состояния *R* и *B* этого остатка. Всего для Ala⁴⁸-Ala⁵⁸ рассчитано 129 структурных вариантов, принадлежащих 22 шейпам пептидного скелета. Относительной энергией менее 8,0 ккал/моль обладают структуры трех шейпов. Лучшие из них имеют конформационные состояния гептапептидного участка Ala⁴⁸-Thr⁵⁴, отвечающие глобальной конформации свободного фрагмента шейпа f_6 ($E_{отн}$ 0 ккал/моль), а у тетрапептидного участка Cys⁵⁵-Ala⁵⁸ — конформациям свободного фрагмента соответственно $R^{55}-R^{56}-P^{57}-R^{58}$, $R^{55}-R^{56}-L^{57}-R^{58}$ и $R^{55}-R^{56}-R^{57}-R^{58}$ шейпов f_{ef} , f_{e2} и f_3 . Следовательно, образование низкоэнергетической конформации ундекапептида происходит путем подстройки лабильного тетрапептида к конформационно жесткой микроглобуле, образованной гептапептидом, причем остаток Cys⁵⁵ и Gly⁵⁶ во всех низкоэнергетических состояниях фрагмента Ala⁴⁸-Ala⁵⁸ входят в α -спираль. Комплементарными к спиральной форме оказываются не более четырех состояний дипептида Gly⁵⁷-Ala⁵⁸ (*P-R*, *L-R*,

R-R и *B-R*). В этом заключается причина резкой энергетической дифференциации конформаций фрагмента $\text{Ala}^{48}-\text{Ala}^{58}$.

Фрагмент $\text{Ala}^{48}-\text{Ala}^{58}$ содержит два остатка Cys, которые располагаются друг против друга в соседних витках α -спирали. Сближение атомов $\text{S}_{(51)}$ и $\text{S}_{(55)}$ на валентное расстояние приводит к искажению структуры и к повышению конформационной энергии более чем на 15 ккал/моль вследствие несогласованности взаимодействий. Поэтому образование дисульфидной связи $\text{Cys}^{51}-\text{Cys}^{55}$ представляется чрезвычайно маловероятным.

Молекула БПТИ. Полученные данные о конформационных возможностях фрагментов БПТИ позволили перейти непосредственно к расчету трехмерной структуры всей молекулы белка. Для этого необходимо было определить конформационное состояние лабильного участка $\text{Lys}^{46}-\text{Ser}^{47}-\text{Ala}^{48}$, обеспечивающее одновременно как энергетически наиболее предпочтительное взаимное расположение конформационно жестких фрагментов $\text{Arg}^1-\text{Lys}^{46}$ и $\text{Ala}^{48}-\text{Ala}^{58}$, так и образование двух дисульфидных связей между остатками $\text{Cys}^{51}-\text{Cys}^{30}$ и $\text{Cys}^5-\text{Cys}^{51}$.

На первом этапе нами выполнен расчет фрагмента $\text{Arg}^1-\text{Ala}^{48}$ путем перебора всех возможных низкоэнергетических состояний трипептида $\text{Lys}^{46}-\text{Ala}^{48}$ в потенциальном поле жесткой структуры $\text{Arg}^1-\text{Lys}^{46}$ (табл. 11). Далее были рассмотрены конформации фрагмента $\text{Arg}^1-\text{Cys}^{51}$, начальные приближения которых образованы из низкоэнергетических структурных вариантов $\text{Arg}^1-\text{Ala}^{48}$ и α -спирального сегмента $\text{Ala}^{48}-\text{Cys}^{51}$. У рассчитанных конформаций $\text{Arg}^1-\text{Cys}^{51}$ оценивалась потенция к образованию S-S-мостиков между Cys^{51} , с одной стороны, и Cys^5 и Cys^{30} — с другой. На заключительном этапе исследовалась трехмерная структура всей молекулы и определялась локализация дисульфидной связи с участием Cys^{55} .

При рассмотрении конформационных возможностей участка $\text{Arg}^1-\text{Cys}^{51}$ приняты во внимание 14 структурных вариантов $\text{Arg}^1-\text{Ala}^{48}$ с величинами $E_{\text{отн}}$ 0—4,6 ккал/моль. Минимизация энергии фрагмента $\text{Arg}^1-\text{Cys}^{51}$ при варьировании двугранных углов основной и боковых цепей $\text{Cys}^{38}-\text{Cys}^{51}$ выявила резкую дифференциацию конформаций. Как видно из табл. 13, структурные варианты, ближайšie к глобальной конформации, имеют на 14,0—20,0 ккал/моль более высокую энергию. Важно, что столь значительное различие в конформационной энергии фрагмента $\text{Arg}^1-\text{Cys}^{51}$ связано с невалентными взаимодействиями его сравнительно небольшого участка $\text{Lys}^{46}-\text{Cys}^{51}$ с $\text{Arg}^1-\text{Lys}^{46}$. В конформации $\text{Lys}^{46}-\text{Cys}^{51}$ с формой основной цепи $R^{46}-B^{47}-R^{48}-R^{49}-R^{50}-R^{51}$ ($E_{\text{отн}}$ 0 ккал/моль) α -спиральный сегмент $\text{Ala}^{48}-\text{Glu}^{49}-\text{Asp}^{50}-\text{Cys}^{51}$ оказывается сближенным со многими остатками и осуществляет с ними эффективные стабилизирующие взаимодействия как дисперсионной, так и электростатической природы.

Следующий этап теоретического конформационного анализа трипептидного ингибитора является одним из наиболее ответственных, узловых моментов данного исследования. Он контролирует результаты всех предшествующих расчетов и тем самым правильность положенных в их основу принципов пространственной организации белка [4—6]. Речь идет о независимой от эксперимента идентификации системы дисульфидных связей в молекуле БПТИ и выяснении стерических условий их образования у рассчитанных конформаций белкового участка $\text{Arg}^1-\text{Cys}^5$ (табл. 13). При анализе фрагмента $\text{Arg}^1-\text{Cys}^{38}$ было обнаружено, что только в самой низкоэнергетической структуре линейной последовательности остатки Cys^{14} и Cys^{38} оказываются сближенными. Создание между ними S-S-связи не приводит к существенному изменению геометрии фрагмента и не нарушает сложившихся стабилизирующих межостаточных взаимодействий. В последующих расчетах предполагалось наличие валентного мостика

$\text{Cys}^{14}-\text{Cys}^{38}$. Было показано также отсутствие стерических условий для образования дисульфидной связи между Cys^5 и Cys^{30} . Расстояние между S-атомами этих остатков в низкоэнергетической структуре $\text{Arg}^1-\text{Cys}^{38}$, имеющей мостик $\text{Cys}^{14}-\text{Cys}^{38}$, составляет ~ 10 Å. Их сближение сопровож-

Относительная энергия (ккал/моль) и расстояния между атомами C^β-остатков Cys⁵, Cys³⁰ и Cys⁵¹ в конформациях фрагмента Arg¹—Cys⁵¹

Конформация	Расстояние, Å		E _{отн}
	C ^β ₍₅₎ ...C ^β ₍₅₁₎	C ^β ₍₃₀₎ ...C ^β ₍₅₁₎	
R ⁴⁶ ₃₂₂₂ -B ⁴⁷ ₁₂ -R ⁴⁸ -R ⁴⁹ ₃₁₁ -R ⁵⁰ ₃₁ -R ⁵¹ ₂	8,6	4,5	0
B ³²²² -B ²² -R -R ³¹¹ -R ³¹ -R ₂	13,7	9,3	14,3
R ³²²² -R ₁₂ -B -R ³¹¹ -R ³¹ -R ₂	11,9	8,6	14,6
B ³²²² -R ₁₂ -R -R ³¹¹ -R ³¹ -R ₂	10,9	4,9	14,9
B ³²²² -R ₂₂ -B -R ³¹¹ -R ³¹ -R ₂	18,9	13,2	15,1
R ³²²¹ -R ₂₂ -R -R ³¹¹ -R ³¹ -R ₂	19,8	16,5	16,3
B ³²²² -B ₂₂ -B -R ³¹¹ -R ³¹ -R ₂	20,9	13,1	19,7
B ³²²² -B ₂₂ -L -R ³¹¹ -R ³¹ -R ₂	20,6	14,7	19,9

Примечание. Во всех конформациях фрагмента Arg¹—Cys⁵¹ участок Arg¹—Phe⁴⁵ имеет одинаковую структуру.

дается рассогласованностью стабилизирующих взаимодействий и возникновением напряженности. При конформационном анализе фрагмента Ala⁴⁸—Ala⁵⁸ выяснилась невозможность образования дисульфидной связи между противостоящими друг другу в соседних витках α-спирали остатками Cys⁵¹ и Cys⁵⁵. Итак, для решения вопроса о системе дисульфидных связей в молекуле БПТИ на основе данных теоретического анализа нам осталось рассмотреть стерические условия их образования между боковыми цепями остатков Cys⁵ и Cys³⁰, с одной стороны, и Cys⁵¹ и Cys⁵⁵ — с другой. На участке Arg¹—Cys⁵¹ возможно оценить конформационные предпосылки создания S—S-мостика между Cys⁵¹ и Cys⁵ или Cys³⁰.

В табл. 13 приведены расстояния C^β₍₅₎...C^β₍₅₁₎ и C^β₍₃₀₎...C^β₍₅₁₎ в конформациях Arg¹—Cys⁵¹ с величинами E_{отн} 0—20,0 ккал/моль. Напомним, что при их расчете не было сделано какого-либо предположения о способе замыкания S—S-связей. Наименьшие расстояния между валентно не связанными остатками цистеина, особенно в положениях 30 и 51, имеет глобальная конформация фрагмента Arg¹—Cys⁵¹ и конформация с E_{отн} 14,9 ккал/моль.

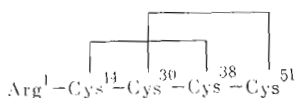
Все попытки сблизить у конформаций с величинами E_{отн} 14,0—20,0 ккал/моль остаток Cys⁵¹ с остатком Cys⁵ или Cys³⁰ путем наложения стягивающего потенциала между соответствующими атомами серы при предоставлении свободы двугранным углам φ, ψ на участке Cys³⁸—Cys⁵¹ фрагмента Arg¹—Cys⁵¹ неизменно приводили к резкому увеличению энергии. При фиксировании отмеченных структурных вариантов фрагмента Arg¹—Cys⁵¹ конформации трипептидного участка Lys⁴⁶—Ala⁴⁸ (что, однако, не исключает возможности значительных изменений его углов φ, ψ в пределах областей B и R) образование дисульфидных связей S₍₅₁₎—S₍₅₎ и S₍₅₁₎—S₍₃₀₎ может осуществляться только при значительной дестабилизации средних и дальних взаимодействий и нарушении согласованности между ними, т. е. при энергетически крайне невыгодных ситуациях.

В этой группе конформаций мостик Cys₍₅₁₎—Cys₍₃₀₎ относительно легко образуется только у структуры с величиной E_{отн} 14,9 ккал/моль. Но и в этом случае конформационная энергия остается значительно выше (>15 ккал/моль) энергии глобальной конформации.

Не привели к положительным результатам и попытки создать дисульфидную связь Cys₍₅₁₎—Cys₍₅₎ в самой низкоэнергетической структуре фрагмента Arg¹—Cys⁵¹. Сближение атомов C^β₍₅₁₎ и C^β₍₅₎ неизменно сопровождается повышением энергии, которое при расстоянии 5,3 Å достигает

15 ккал/моль; стремление установить между остатками Cys⁵¹ и Cys⁵ еще более тесный контакт вызывает прогрессирующий рост энергии из-за многочисленных стерических препятствий.

Совершенно иное положение складывается при образовании у глобальной конформации Arg¹-Cys⁵¹ дисульфидного мостика $\overline{\text{Cys}_{(51)}-\text{Cys}_{(30)}}$. Уменьшение расстояния между атомами серы вплоть до длины валентной связи не встречает здесь серьезных стерических осложнений и не вызывает заметного ослабления сложившихся до образования цикла невалентных межостаточных взаимодействий. Энергия этой конформации в бициклическом варианте



выше энергии соответствующего моноциклического состояния $\overline{\text{Arg}^1-\text{Cys}^{14}-\text{Cys}^{30}-\text{Cys}^{38}-\text{Cys}^{51}}$ всего на 2,0 ккал/моль, а геометрические параметры обоих дисульфидных мостиков (табл. 14) практически совпадают с экспериментальными значениями ($r_{\text{S-S}}$ 2,04, $r_{\text{S-C}\beta}$ 3,05 Å, τ (C^βSS) 104° [8, 30-31]).

Уникальную у глобальной структуры Arg¹-Cys⁵¹ стерическую компонентарность соответствующих участков белковой молекулы, их склонность к образованию такой структуры, которая неизбежно (т. е. в самом выгодном линейном структурном варианте) приводит к сближенности остатки Cys⁵¹ и Cys³⁰ и обеспечивает беспрепятственное образование еще более тесного контакта между ними, подтверждают и попытки создать дисульфидную связь $\overline{\text{Cys}^{51}-\text{Cys}^{30}}$ исходя из высокоэнергетических конформаций фрагмента.

Таким образом, рассмотрение конформационного состояния участка Arg¹-Cys⁵¹ привело к заключению о возможности реализации у этого фрагмента только одного низкоэнергетического бициклического структурного варианта с дисульфидными связями $\overline{\text{Cys}^{14}-\text{Cys}^{38}}$ и $\overline{\text{Cys}^{30}-\text{Cys}^{51}}$. Его параметры и α-спиральный участок Cys⁵¹-Gly⁵⁶ использованы в качестве начального приближения для выяснения структуры фрагмента Arg¹-Gly⁵⁶ и возможности образования S-S-мостика между остатками Cys⁵ и Cys⁵⁵. Стерическая предрасположенность составленной таким образом конформации Arg¹-Gly⁵⁶ проявляется в сближенности боковых цепей Cys⁵⁵ и Cys⁵ (расстояние между атомами C₍₅₅₎^β и C₍₅₎^β не превышает 5,2 Å) и в отсутствии неблагоприятных контактов. Сближение атомов S₍₅₅₎ и S₍₅₎ на валентное расстояние при одновременной минимизации энергии с варьированием двугранных углов участка Cys⁵¹-Gly⁵⁶ не вызывает, как и в случаях образования $\overline{\text{Cys}^{14}-\text{Cys}^{38}}$ и $\overline{\text{Cys}^{51}-\text{Cys}^{30}}$ существенного изменения конформационных состояний остатков и рассогласованности сложившихся между ними ранее взаимодействий. Найденные геометрические параметры мостика $\overline{\text{Cys}^{55}-\text{Cys}^5}$ совпадают с экспериментальными значениями (табл. 14).

На последнем этапе анализа молекулы БПТИ рассмотрены конформационные возможности C-концевого дипептида Gly⁵⁷-Ala⁵⁸ в потенциальном поле рассчитанной структуры Arg¹-Gly⁵⁶. Полученные результаты свидетельствуют о подвижности этого участка, который может паходиться в четырех, практически изоэнергетичных состояниях: *P-R*, *B-R*, *L-R* и *R-R*.

Итак, теоретический конформационный анализ трехмерной структуры БПТИ завершен. Рассчитанные значения 298 двугранных углов φ, ψ, ω и χ, определяющих конформационные состояния основных и боковых цепей всех 58 аминокислотных остатков молекулы, приведены в табл. 15. Там же даны экспериментальные значения этих углов, полученные рент-

Геометрические параметры дисульфидных связей в молекуле БПТИ

Дисульфидная связь	Расстояние между атомами, Å			
	$S_i - S_j$	$C_i \cdots C_j^{\beta}$	$C_i^{\beta} \cdots S_j$	$C_j \cdots S_i$
$\overbrace{\text{Cys}^{14} - \text{Cys}^{38}}$	2,04	4,12	3,1	3,1
$\overbrace{\text{Cys}^{30} - \text{Cys}^{51}}$	2,04	4,25	3,0	3,0
$\overbrace{\text{Cys}^5 - \text{Cys}^{55}}$	1,98	4,15	3,1	3,1

геноструктурным анализом белка с разрешением 1,5 Å [2, 3]. Из сопоставления данных видно, что найденная априорно конформация БПТИ полностью соответствует экспериментальной кристаллографической структуре. Подавляющее большинство конформационных параметров количественно совпадает между собой. Заметные различия, выходящие за условно принимаемые границы погрешностей рентгеноструктурного анализа в определении двугранных углов белковой цепи ($\pm 10-20^\circ$), касаются лишь боковых цепей единичных остатков, например Arg²⁰ (χ^4), Asn⁴³ (χ^2) и Asn⁴⁴ (χ^1). Однако и здесь экспериментальные значения углов χ также вполне приемлемы для найденной нами структуры; при сохранении других расчетных параметров они приводят к повышению конформационной энергии молекулы лишь на несколько килокалорий. Среднеквадратичные отклонения опытных и теоретических углов основной цепи составляют 13° для ϕ и ψ и 9° для ω . В пределах допустимых экспериментальных ошибок находится также среднеквадратичное отклонение углов χ^1 (18°).

При оценке значимости совпадения вычисленных и опытных значений двугранных углов нужно иметь в виду то обстоятельство, что по ряду причин они не являются вполне удовлетворительными количественными характеристиками пространственного строения белка, найденного теоретическим и опытным путем. Их величины зависят от длины связей и валентных углов молекулы, которые в двух случаях были не идентичны. Так, полученные Дайзенхофером и Стайгеманом [3] при уточнении кристаллографической структуры БПТИ по методу Даймона [33] величины валентных углов $\tau(\text{NC}^{\alpha}\text{C}')$ обнаруживают существенный разброс ($95-124^\circ$), который не может отвечать реальной ситуации. Приведение углов $\tau(\text{NC}^{\alpha}\text{C}')$ к действительно наблюдаемым у пептидов значениям (интервал $106-114^\circ$) неминуемо повлечет к изменению найденных в работе [2] двугранных углов ϕ , ψ , ω основной цепи и χ^1 боковых цепей. В нашем расчете была выбрана иная валентная геометрия белка, основанная на параметрах Полинга и Кори [7]. Другая причина неполной корректности сопоставления структур по двугранным углам связана с точностью их расчета. Сравнительно небольшие ошибки в значениях отдельных двугранных углов, особенно основной цепи, могут привести к весьма значительному изменению всей структуры. Поскольку в расчете они неизбежны, на первый взгляд представляется даже бесперспективным теоретический конформационный анализ белков. На самом деле такое опасение оказалось излишним. Вследствие высокой конформационной чувствительности потенциальной энергии, уникальности трехмерной структуры белка и большой гибкости пептидной цепи на ряде участков аминокислотной последовательности двугранные углы не являются независимыми друг от друга и отклонения расчетных значений одних углов от их истинных величин в той или иной степени компенсируются отклонениями других. Поэтому допускаемые в определении углов погрешности радикальным образом не сказываются на окончательном результате. Однако при сопоставлении их опытных и теоретических значений трудно оценить, насколько серьезно наблюдаемое численное расхождение между ними. Несмотря на сказанное выше, имеющее место совпадение двугранных углов основной и боковых цепей

Экспериментальные (верхняя строка) и теоретические значения конформационных параметров (град) молекулы БПТИ

Остаток	Идентификатор	Основная цепь			Боковая цепь				
		φ	ψ	ω	χ^1	χ^2	χ^3	χ^4	χ^5
Arg ¹	B_{1214}	-	144	179	82	-179	56	82	1
	B_{1214}	56	149	175	62	176	62	77	0
Pro ²	B	-58	152	177	-174	177	-170		
	B	-60	140	-177	180	180	180		
Asp ³	R_{31}	-61	-30	-177	-86	5			
	R_{31}	-97	-48	-172	-64	31			
Phe ⁴	R_{11}	-74	-12	169	66	71			
	R_{11}	-64	-24	-172	30	83			
Cys ⁵	R_{333}	-65	-17	176	-56	-88	-87		
	R_3	-77	-22	-179	-69				
Leu ⁶	R_{32}	-92	-5	175	-54	-174	-	-	
	R_{3214}	-111	-14	-176	-42	178	66	62	
Glu ⁷	B_{121}	-75	147	175	17	177	73		
	B_{121}	-55	137	180	78	169	90		
Pro ⁸	B	-67	157	178	-178	-165	162		
	B	-60	157	-179	180	180	180		
Pro ⁹	B	-64	145	-173	178	-175	157		
	B	-60	144	178	180	180	180		
Tyr ¹⁰	B_{21}	-126	106	-179	178	77	-		
	B_{212}	-132	110	180	-177	86	180		
Thr ¹¹	R_3	-70	-41	173	-77	-	-		
	R_{312}	-85	-61	180	-58	62	179		
Gly ¹²	L	94	179	-176					
	L	93	170	179					
Pro ¹³	R	-86	-7	178	178	-170	148		
	R	-60	-29	179	180	180	180		
Cys ¹⁴	B_{311}	-89	165	-170	-73	96	101		
	B_3	-88	140	-178	-70				
Lys ¹⁵	B_{31}	-127	32	163	-111	92	-	-	-
	B_{31221}	-125	63	-169	-75	83	151	-168	60
Ala ¹⁶	B	-77	178	-171	-				
	B_1	-86	173	174	60				
Arg ¹⁷	B_{32223}	-134	88	-163	-62	162	-137	113	-12
	B_{32223}	-120	70	180	-62	-179	180	180	0
Ile ¹⁸	B_{32}	-125	118	-176	-65	-173	-	-	
	B_{3211}	-116	152	-170	-56	176	56	61	
Ile ¹⁹	B_{33}	-84	118	178	-57	-55	-	-	
	B_{3311}	-92	105	171	-45	-56	60	61	
Arg ²⁰	B_{33213}	-126	172	175	-64	-77	177	89	-3
	B_{33223}	-130	167	-178	-49	-68	179	-154	0
Tyr ²¹	B_{31}	-140	146	170	-72	77	-		
	B_{311}	-135	151	180	-30	114	0		
Phe ²²	B_{11}	-129	151	167	76	100			
	B_{11}	-130	136	-170	81	103			
Tyr ²³	B_{23}	-80	130	-174	-171	-108	-		
	B_{232}	-98	131	176	179	-83	180		
Asn ²⁴	B_{23}	-110	104	-175	173	-11			
	B_{23}	-100	100	-177	-174	-24			
Ala ²⁵	R	-65	-28	171	-				
	R_1	-68	-35	-173	60				
Lys ²⁶	R_{321}	-67	-34	-177	-88	-121	105	-	-
	R_{33221}	-62	-48	176	-87	-81	180	180	60

Таблица 15 (продолжение)

Остаток	Идентификатор	Основная цепь			Боковая цепь				
		φ	ψ	ω	χ^1	χ^2	χ^3	χ^4	χ^5
Ala ²⁷	<i>R</i>	-93	-20	-179	-				
	<i>R</i> ₁	-79	-31	180	60				
Gly ²⁸	<i>L</i>	81	15	-175					
	<i>L</i>	79	34	177					
Leu ²⁹	<i>B</i> ₁₁	-159	175	-172	44	72	-	-	
	<i>B</i> ₁₁₁₁	-131	163	177	59	84	59	58	
Cys ³⁰	<i>B</i> ₃₂₃	-98	146	-177	-68	-122	-90		
	<i>B</i> ₃	-89	140	169	-74				
Gln ³¹	<i>B</i> ₃₂₂	-133	161	173	-67	178	145		
	<i>B</i> ₃₂₁	-125	160	176	-62	-167	104		
Thr ³²	<i>B</i> ₁	-83	150	170	53	-	-		
	<i>B</i> ₁₁₃	-106	152	180	70	30	-2		
Phe ³³	<i>B</i> ₁₁	-146	166	179	74	114			
	<i>B</i> ₁₁	-150	171	-176	104	106			
Val ³⁴	<i>B</i> ₂	-96	120	-175	172	-	-		
	<i>B</i> ₂₁₁	-104	128	-176	165	61	86		
Tyr ³⁵	<i>B</i> ₂₁	-100	134	178	161	53	-		
	<i>B</i> ₂₁₁	-92	140	-178	-165	49	0		
Gly ³⁶	<i>R</i>	-73	-11	170					
	<i>R</i>	-60	-20	174					
Gly ³⁷	<i>P</i>	105	0	174					
	<i>P</i>	90	-40	178					
Cys ³⁸	<i>B</i> ₁₃₁	-146	151	178	74	-118	103		
	<i>B</i> ₁	-154	146	180	45				
Arg ³⁹	<i>L</i> ₃₃₂₂₁	61	40	-178	-46	-52	-178	-174	2
	<i>L</i> ₃₃₂₂₁	61	36	-171	-50	-64	-179	180	0
Ala ⁴⁰	<i>B</i>	-61	152	177	-				
	<i>B</i> ₁	-101	180	-176	60				
Lys ⁴¹	<i>B</i> ₃₂₂	-103	180	-172	-88	-165	-175	-	-
	<i>B</i> ₂₂₂₂₁	-84	180	180	-128	-156	165	180	60
Arg ⁴²	<i>R</i> ₃₂₁₂₃	-85	-20	167	-83	-160	110	179	-8
	<i>R</i> ₂₂₂₂₃	-82	-17	-177	-60	180	180	180	0
Asn ⁴³	<i>B</i> ₂₁	-77	74	-166	-161	7			
	<i>B</i> ₂₁	-75	81	180	-178	82			
Asn ⁴⁴	<i>B</i> ₂₃	-166	107	-175	176	-30			
	<i>B</i> ₁₃	-170	131	-176	73	-70			
Phe ⁴⁵	<i>B</i> ₃₁	-129	158	-177	-56	91			
	<i>B</i> ₃₂	-118	160	-177	-64	138			
Lys ⁴⁶	<i>R</i> ₃₂₂₁	-88	-7	174	-80	160	-142	91	-
	<i>R</i> ₃₂₂₁₁	-98	-28	179	-58	180	-179	60	60
Ser ⁴⁷	<i>B</i> ₁	-154	162	179	76	-			
	<i>B</i> ₁₂	-151	170	180	54	177			
Ala ⁴⁸	<i>R</i>	-67	-34	-175	-				
	<i>R</i> ₁	-61	-23	180	60				
Glu ⁴⁹	<i>R</i> ₃₁₂	-69	-42	173	-73	86	173		
	<i>R</i> ₃₁₁	-52	-43	180	-78	73	88		
Asp ⁵⁰	<i>R</i> ₃₁	-65	-41	175	-97	52			
	<i>R</i> ₃₁	-71	-35	180	-53	98			
Cys ⁵¹	<i>R</i> ₂₃₃	-59	-50	-179	175	-94	-87		
	<i>R</i> ₂	-60	-45	-176	180				
Met ⁵²	<i>R</i> ₃₃₃	-71	-32	176	-70	-62	-71	-	
	<i>R</i> ₃₃₃₁	-84	-37	-175	-64	-52	-74	60	
Arg ⁵³	<i>R</i> ₂₂₁₂₁	-63	-42	175	178	156	37	-125	7
	<i>R</i> ₂₂₁₂₁	-63	-40	-171	171	166	59	-175	0

Остаток	Идентификатор	Основная цепь			Боковая цепь				
		φ	ψ	ω	χ^1	χ^2	χ^3	χ^4	χ^5
Thr ⁵⁴	R_3	-76	-46	-173	-60	-	-	-	-
	R_{312}	-71	-66	-173	-62	60	175	-	-
Cys ⁵⁵	R_{333}	-106	-4	-171	-69	-64	-91	-	-
	R_3	-114	-20	-169	-65	-	-	-	-
Gly ⁵⁶	R	-78	-14	174	-	-	-	-	-
	R	-60	-47	-179	-	-	-	-	-
Gly ⁵⁷ *	-	-	-	-	-	-	-	-	-
	P	159	-61	-179	-	-	-	-	-
Ala ⁵⁸ *	-	-	-	-	-	-	-	-	-
	R	-160	-55	-	60	-	-	-	-

* Численные значения этих параметров в работах [9, 10] не приведены.

или, иными словами, близость конформационных состояний всех аминокислотных остатков в опытной и теоретической структурах белка несомненно является объективным, очень веским доводом в пользу правильности результатов расчета. Однако лучшее представление о точности априорного расчета трехмерной структуры белка может быть получено из сравнения межатомных расстояний в двух структурах и прежде всего расстояний между C^α -атомами основной цепи. С этой целью нами найдено среднеквадратичное отклонение всех парных расстояний $C_i^\alpha \dots C_j^\alpha$, число которых составило 1540. При сравнении не учтены лишь остатки Gly⁵⁷ и Ala⁵⁸, так как в работах [2, 3] отсутствуют данные об их конформационных состояниях. По-видимому, это связано с подвижностью С-концевого фрагмента Gly⁵⁷-Ala⁵⁸, о чем свидетельствуют и результаты расчета. Наидежная величина среднеквадратичного отклонения расстояний между атомами C^α составила 1,4 Å; отклонения на участках Arg¹-Cys¹⁴, Cys¹⁴-Cys³⁸ и Cys³⁸-Gly⁵⁶ равны соответственно 1,0; 1,2 и 1,4 Å. Таким образом, среднеквадратичные отклонения, касающиеся всей структуры молекулы и ее отдельных фрагментов, мало различаются между собой и невелики. Это обстоятельство свидетельствует о том, что априорный расчет весьма удовлетворительно и приблизительно с одинаковой точностью предсказывает конформационные состояния отдельных частей молекулы и их взаимное расположение в нативной трехмерной структуре белка. На это указывает также хорошее совпадение габаритных размеров теоретической и экспериментальной конформаций БПТИ. Приведем несколько примеров. Так, расчетное и наблюдаемое расстояния $C_{(1)}^\alpha \dots C_{(15)}^\alpha$ равны 29,1 и 29,8; $C_{(2)}^\alpha \dots C_{(56)}^\alpha - 10, 2$ и 9,5; $C_{(8)}^\alpha \dots C_{(42)}^\alpha - 11,4$ и 10,0; $C_{(14)}^\alpha \dots C_{(5)}^\alpha - 21,1$ и 22,0; $C_{(14)}^\alpha \dots C_{(30)}^\alpha - 23,3$ и 22,3; $C_{(14)}^\alpha \dots C_{(51)}^\alpha - 22,3$ и 22,1; $C_{(14)}^\alpha \dots C_{(55)}^\alpha - 26,3$ и 25,3; $C_{(28)}^\alpha \dots C_{(38)}^\alpha - 27,5$ и 26,0 Å. Правильность полученных нами результатов подтверждается также сравнением длин водородных связей, обнаруженных в кристаллографической и расчетной структурах БПТИ. Данные об их геометрии количественно характеризуют взаимное расположение соответствующих атомных групп, которые принадлежат далеко отстоящим в линейной последовательности остаткам. Среднеквадратичное отклонение длин водородных связей основной цепи равно 0,5 Å. Лишь в двух случаях из 18 различие в расстояниях O...N достигает 1,0 Å.

Таким образом, результаты проведенного сопоставления геометрии двух структур как по двугранным углам вращения, так и по расстояниям между атомами C^α основной цепи и по длинам водородных связей позволяют утверждать о хорошем совпадении всех элементов априорно рассчитанной конформации БПТИ и кристаллографической трехмерной структуры белка.

Заключение

В заключение остановимся на краткой характеристике найденной нами трехмерной структуры молекулы БПТИ и в связи с этим рассмотрим некоторые общие вопросы пространственной организации белков.

Результаты теоретического конформационного анализа БПТИ прежде всего свидетельствуют о справедливости положенной в основу расчета количественной структурной теории белка [4—6], т. е. о правильном понимании принципов его пространственной организации и о возможности количественного предсказания нативной конформации молекулы, исходя только из аминокислотной последовательности. Расчет подтвердил одно из основных положений теории о наличии в плотно упакованной трехмерной структуре белка согласованности между всеми близкими, средними и дальними взаимодействиями. Именно это свойство, присущее только эволюционно отобранной аминокислотной последовательности, ответственно за резкую энергетическую дифференциацию и выделение из множества возможных структурных вариантов стабильной и уникальной в каждом случае глобулярной конформации.

В найденной нами структуре БПТИ геометрия всех 58 остатков молекулы отвечает наиболее предпочтительным конформациям соответствующих свободных монопептидов ($CH_3-CONH-C^\alpha RH-CONH-CH_3$). Ближние взаимодействия определяют конформационные возможности каждого аминокислотного остатка, т. е. набор дозволенных для него конформационных состояний. Средние взаимодействия — взаимодействия остатка с четырьмя предшествующими и четырьмя последующими в цепи остатками, только из этих наборов выбирает конформации, сочетание которых оказывается наиболее выгодным. Суммарный эффект средних взаимодействий в расчетной конформации молекулы БПТИ составил около -180 ккал/моль (в среднем $-3,1$ ккал/моль на остаток). На некоторых участках последовательности белка выбор конформационного состояния остатков только за счет средних взаимодействий практически однозначен (особенно в отношении формы основной цепи). В этих случаях соответствующие фрагменты образуют конформационно стабильные микроглобулы. В БПТИ к таким участкам можно отнести Arg^1-Pro^9 , $Phe^{22}-Gln^{31}$ и $Ala^{48}-Gly^{56}$. В других местах последовательности средние взаимодействия в той или иной мере сокращают число возможных конформационных состояний остатков и определяют для соответствующих фрагментов ограниченные наборы приблизительно равновероятных структурных вариантов.

Функция дальних взаимодействий по отношению к конформационно жестким по средним взаимодействиям фрагментам сводится к их дополнительной стабилизации, а также к еще большему ограничению свободы боковых цепей и, возможно, к незначительной корректировке формы основной цепи. У микроглобулы фрагмента Arg^1-Pro^9 энергия средних взаимодействий составляет около -35 ккал/моль, а дальних — около -39 ккал/моль; у $Phe^{22}-Gln^{31}$ соответствующие величины равны ~ -33 и -34 ккал/моль, а у $Ala^{48}-Gly^{56}$ ~ -36 и -30 ккал/моль. Таким образом, согласованность имеет место между сопоставимыми по своей эффективности взаимодействиями.

По отношению к конформационно лабильным участкам БПТИ функция дальних взаимодействий заключается в выборе из равновероятных по средним взаимодействиям состояний единичных конформаций, обеспечивающих образование между ними и микроглобулами эффективных контактов, а также близость элементарных друг другу глобул. Так, расчет показал, что значительно большей подвижностью по сравнению с

глобулами обладают, например, свободные фрагменты Cys¹⁴—Tyr²¹ и Thr³²—Cys³⁸. Но когда они присоединяются к жесткому по средним взаимодействиям участку Phe²²—Gln³¹, их состояние полностью детерминировано при встраивании в антипараллельную β-структуру. При этом между остатками фрагментов Cys¹⁴—Tyr²¹ и Thr³²—Cys³⁸ образуются эффективные стабилизирующие контакты. Это происходит только в том случае, если каждый из них принимает развернутую форму основной цепи, которая у того и другого фрагмента является одной из наиболее выгодных по средним взаимодействиям. Предпочтительность этой формы у Cys¹⁴—Tyr²¹ и Thr³²—Cys³⁸ на белковом участке Cys¹⁴—Cys³⁸ по сравнению с рядом других, изоэнергетичных у свободных фрагментов форм обусловлена дальними взаимодействиями. В β-структуре энергия межостаточных взаимодействий Cys¹⁴—Tyr²¹ с Thr³²—Cys³⁸ составляет около −21 ккал/моль при их собственной энергии в развернутых конформациях соответственно −17 и −15 ккал/моль. Кроме того, они образуют стабилизирующие контакты с предшествующим участком белка Arg¹—Pro¹³ (∼−15 ккал/моль), промежуточным Phe²²—Gln³¹ (∼−5 ккал/моль) и последующим Arg³⁹—Ala⁵⁸ (∼−12 ккал/моль).

Весьма лабилен по средним взаимодействиям также фрагмент Arg³⁹—Ser⁴⁷. Но лишь при одном из большого числа низкоэнергетического конформационного состояния он обеспечивает сближенность и образование многочисленных стабилизирующих контактов между жесткими β-структурным (Cys¹⁴—Cys³⁸) и α-спиральным (Ala⁴⁸—Ala⁵⁸) участками белка; суммарный энергетический эффект таких контактов составляет около −22 ккал/моль. Кроме того, при таком конформационном состоянии (энергия его средних взаимодействий равна ∼−25 ккал/моль) сам фрагмент Arg³⁹—Ser⁴⁷ осуществляет одновременно дальние взаимодействия с Arg¹—Pro⁹ (−20 ккал/моль), Cys¹⁴—Cys³⁸ (−30,0 ккал/моль) и Ala⁴⁸—Ala⁵⁸ (−25 ккал/моль). Общая энергия дальних взаимодействий в найденной нативной конформации БПТИ равна ∼−120 ккал/моль (в среднем −2,4 ккал/моль на остаток).

Конформационный анализ молекулы БПТИ показал, что согласованные между собой ближние, средние и дальние взаимодействия выполняют на разных стадиях самоорганизации белковой структуры вполне определенные функции. Поскольку роль этих взаимодействий различна и в равной мере необходима для осуществления быстрой и безошибочной сборки белка, нам представляется беспредметным так часто обсуждаемый в литературе вопрос о том, какие взаимодействия (ближние, средние или дальние) имеют определяющее значение в этом процессе.

О высокой степени согласованности всех межостаточных взаимодействий свидетельствует тот факт, что у рассмотренных фрагментов БПТИ независимо от их длины конформационные состояния, входящие в нативную структуру белка, оказывались или глобальными, или одними из самых предпочтительных, т. е. их энергия всегда была меньше энергии большинства конформационных вариантов, отобранных для последующего анализа.

В расчете БПТИ учитывались внутримолекулярные невалентные, электростатические и торсионные взаимодействия, а также водородные связи. Каковы же вклады этих взаимодействий в энергию нативной конформации белка?

Аминокислотная последовательность БПТИ включает 18 остатков, несущих 12 положительных и 13 отрицательных целочисленных зарядов. Стабилизирующий вклад электростатических взаимодействий заряженных остатков составляет около −50 ккал/моль, а дестабилизирующий — +35 ккал/моль. Следовательно, хотя электростатические взаимодействия и понижают конформационную энергию приблизительно на −15 ккал/моль, их интегральный вклад в стабилизацию структуры БПТИ сравнительно невелик. Суммарный стабилизирующий эффект водородных связей составил около −14 ккал/моль. Обнаруженные в расчете водородные связи обусловлены сложным комплексом многих межостаточных взаимодействий.

вии, среди которых собственно водородные связи играют незначительную роль. Общая энергия торсионных взаимодействий в найденной конформации БПТИ равна около +58 ккал/моль, т. е. ~1,0 ккал/моль на остаток. Из этого следует, что взаимные расположения практически всех атомных групп основных и боковых цепей аминокислотных остатков отвечают минимумам торсионных потенциалов. Таким образом, доминирующий вклад (75%) в конформационную энергию межкостаточных взаимодействий белка вносят невалентные, точнее, дисперсионные контакты.

Согласно теории [4—6], расчет трехмерной структуры белка можно проводить поэтапно, путем анализа аминокислотной последовательности постепенно увеличивающейся длины. На каждом этапе, начиная с отдельных аминокислотных остатков и дипептидных фрагментов, выполняется минимизация энергии при варьировании двугранных углов вращения вокруг связей основных и боковых цепей. В результате для всех рассматриваемых фрагментов выполняются оптимальные условия реализации межкостаточных взаимодействий при низкоэнергетических конформационных состояниях остатков и всех возможных формах и шейпах пептидного скелета. Такой путь поиска трехмерной структуры белка, по-видимому, наиболее естествен, так как он в той или иной степени моделирует реальный процесс укладки белковой цепи в глобулу. Действительно, формирование нативной конформации из неструктурированной флуктуирующей формы неизбежно начинается с выяснения взаимоотношений между соседними по цепи остатками. Конформационные возможности остатков определяются близкими взаимодействиями, т. е. их природой, а конформационные возможности фрагмента — средними взаимодействиями, т. е. конкретной аминокислотной последовательностью. Дальние взаимодействия возникают только на определенной стадии структурирования цепи, уже после того, как средние взаимодействия в основном выполнили свою функцию. Таким образом, самоорганизация белковой глобулы в естественных условиях — это также процесс последовательной «минимизации» энергии путем выбора конформационных состояний, наиболее предпочтительных сначала для простых, а затем для все более и более усложняющихся участков белка, вплоть до всей молекулы. Ограниченное число возможных для остатков конформационных состояний, образование микроглобул и их ассоциация без существенных изменений конформации при наличии согласованности ближних, средних и дальних взаимодействий значительно упрощают механизм сборки белковой структуры, делают его строго детерминированным, быстрым и безошибочным.

Опишем процесс самоорганизации пространственного строения молекулы БПТИ от полностью развернутой формы до нативной структуры, опираясь при этом на полученные нами данные о конформационных возможностях конкретной последовательности и ее отдельных фрагментов. Процесс свертывания белка начинается одновременно и независимо на трех участках последовательности с образования за счет средних взаимодействий микроглобул Arg¹—Pro⁹, Phe²²—Gln³¹ и Ala⁴⁸—Gly⁵⁶. Они являются доминирующими по энергии свободных фрагментов и впоследствии входят в нативную конформацию белка без существенных изменений. Фрагменты Cys¹⁴—Tyr²¹ и Thr³²—Cys³⁸, будучи более подвижными, по обладая предрасположенностью к развернутым формам цепи, приводят вместе с Phe²²—Gln³¹ к созданию конформационно жесткой антипараллельной β-структуры со сближенными остатками Cys¹⁴ и Cys³⁸. Образование между последними дисульфидного мостика не сопровождается заметным изменением геометрии и, следовательно, не приводит к рассогласованию сложившихся у линейной последовательности стабилизирующих межкостаточных взаимодействий. Необходимое для валентного связывания уменьшение расстояния между атомами серы с ~3,5 до ~2,0 Å происходит за счет конформационно подвижного концевой участка β-структуры Gly³⁶—Gly³⁷—Cys³⁸.

Формирование микроглобул Arg¹—Pro⁹ и ее взаимодействия с последующими остатками Tyr¹⁰—Cys¹⁴ и образовавшейся циклической β-структурой Cys¹⁴—Cys³⁸ однозначно детерминирует конформационное состояние

фрагмента Arg¹—Cys¹⁴. Аналогичное воздействие β-структура оказывает и на лабильный по средним взаимодействиям фрагмент Arg³⁹—Lys⁴⁶. Из его многочисленных состояний лишь одно стерически комплементарно конформации Cys¹⁴—Cys³⁸. Укладка фрагмента Arg³⁹—Lys⁴⁶ относительно β-структуры создает условие для образования эффективных стабилизирующих контактов с микроглобулой Ala⁴⁸—Gly⁵⁶. Ее взаимодействия с Cys¹⁴—Cys³⁸ и Arg¹—Pro⁹ приводят к сближенности остатков Cys⁵¹ и Cys⁵⁵ соответственно с остатками Cys³⁰ и Cys⁵, между которыми, также не вызывая стерических осложнений, образуются две дисульфидные связи. Непизбежное при сближении S-атомов на валентное расстояние незначительное изменение геометрии пептидной цепи делает низкоэнергетичным лабильный участок Lys⁴⁶—Ser⁴⁷—Ala⁴⁸.

Конформационный анализ молекулы БПТИ выполнялся нами для линейной последовательности с учетом всех возможных способов образования трех S—S-мостиков. Результаты расчета однозначно свидетельствуют в пользу только одной системы дисульфидных связей Cys⁵—Cys⁵⁵, Cys¹⁴—Cys³⁸ и Cys³⁰—Cys⁵¹, отвечающей нативной конформации белка. Создание S—S-связей осуществлялось по одному и тому же механизму, в основе которого лежат конформационные аспекты. Во всех случаях сближенность соответствующих остатков Cys происходила автоматически в самых выгодных конформациях липейных участков.

Как уже отмечалось, Крайтон [27, 28] при исследовании промежуточных состояний БПТИ, образующихся в процессе свертывания денатурированной белковой цепи, обнаружил продукты с дисульфидными связями, которые отсутствуют в нативной конформации. Хотя мы и не ставили перед собой задачу проанализировать все найденные Крайтоном метастабильные состояния молекулы, тем не менее в свете полученных нами результатов становится очевидно, что их образование на начальных этапах самоорганизации вполне вероятно. Так, расчет показал, что, например, структуры метастабильных продуктов с «неправильными» дисульфидными связями Cys⁵—Cys¹⁴ и Cys⁵—Cys³⁸ включают в основном те же микроглобулы, что и нативная конформация БПТИ, по иному образцу упаковки. В начале ренатурации они являются даже самыми низкоэнергетическими, однако быстро теряют свою предпочтительность по мере реализации дальних взаимодействий.

ЛИТЕРАТУРА

1. Kassel B., Laskowski M. Biochem. and Biophys. Res. Commun., 1965, v. 20, № 4, p. 463—468.
2. Huber R., Kukla D., Bode W., Schwager P., Bartels K., Deisenhofer J., Steigemann W. J. Mol. Biol., 1974, v. 89, № 1, p. 73—101.
3. Deisenhofer J., Steigemann W. Acta cryst., 1975, v. B31, № 2, p. 238—250.
4. Попов Е. М. Молекулярн. биология, 1975, т. 9, № 4, с. 578—593.
5. Попов Е. М. Int. J. Quant. Chem., 1979, v. 16, № 5, p. 707—737.
6. Попов Е. М. Физика молекул. Киев: Наукова думка, 1980, вып. 8, с. 69—113.
7. Pauling L. The Nature of the Chemical Bond., 3rd ed. N. Y.: Cornell Univ. Press, 1960.
8. Momany F. A., McGuire R. F., Burgess A. W., Scheraga H. A. J. Phys. Chem., 1975, v. 79, № 7, p. 2361—2384.
9. Ponnuswamy P. K., McGuire R. F., Scheraga H. A. Int. J. Pept. Prot. Res., 1973, v. 5, № 1, p. 73—84.
10. Scott R. A., Scheraga H. A. J. Chem. Phys., 1966, v. 45, № 9, p. 2091—2101.
11. Липкинд Г. М., Архипова С. Ф., Попов Е. М. Ж. структ. химии, 1970, т. 11, № 1, с. 121—126.
12. Попов Е. М., Дашевский В. Г., Липкинд Г. М., Архипова С. Ф. Молекулярн. биология, 1968, т. 2, № 4, с. 612—620.
13. Schellman I. A. Compr. rend. trav. lab. Carlsberg. ser. chim., 1955, v. 29, № 3, p. 223—228.
14. Klotz T. M., Fransen I. S. J. Amer. Chem. Soc., 1962, v. 84, № 12, p. 3461—3465.
15. Mizushima S., Shimanouchi T., Tsuboi M. Nature, 1950, v. 166, № 4, p. 406—409.
16. Попов Е. М., Швырков В. Н., Спасов В. З. Биоорган. химия, 1982, т. 8, № 1, с. 61—71.

17. Завильный А. А., Попов Е. М. Молекулярн. биология, 1982, т. 16, № 1, с. 129–141.
18. Попов Е. М., Ахмедов Н. А., Липкинд Г. М. Биоорган. химия, 1975, т. 1, № 9, с. 1268–1279.
19. Липкинд Г. М., Ахмедов Н. А., Попов Е. М. Биоорган. химия, 1975, т. 1, № 9, с. 1280–1290.
20. Мельников П. Н., Ахмедов Н. А., Липкинд Г. М., Попов Е. М. Биоорган. химия, 1976, т. 2, № 1, с. 28–42.
21. Ахмедов Н. А., Липкинд Г. М., Мельников П. Н., Попов Е. М. Биоорган. химия, 1976, т. 2, № 6, с. 746–761.
22. Архипова С. Ф., Севастьянова Н. Н., Липкинд Г. М., Попов Е. М. Биоорган. химия, 1977, т. 3, № 3, с. 335–347.
23. Севастьянова Н. Н., Попов Е. М. Биоорган. химия, 1977, т. 3, № 4, с. 473–484.
24. Севастьянова Н. Н., Попов Е. М. Биоорган. химия, 1978, т. 4, № 8, с. 997–1016.
25. Попов Е. М., Мельников П. Н. Биоорган. химия, 1979, т. 5, № 6, с. 828–847.
26. Попов Е. М., Мельников П. Н. Биоорган. химия, 1979 т. 5, № 10, с. 1471–1494.
27. Creighton T. E., J. Mol. Biol., 1975, v. 95, № 2, p. 167–199.
28. Creighton T. E., Kalej E., Arnon R. J. Mol. Biol., 1978, v. 123, № 2, p. 129–147.
29. Anfinsen C. B. Science, 1973, v. 181, № 2, p. 223–230.
30. Low J. P. Progr. Phys. Org. Chem., 1968, v. 6, № 1, p. 1–10.
31. Sutter D., Dreizler H., Rudolph H. Z. Naturforsch., 1965, v. A20, № 10, p. 1676–1684.
32. Jones D. D., Bernal I., Frey M. N., Koetzle T. F. Acta cryst. 1974, v. B30, № 9, p. 1220–1229.
33. Diamond R. Acta cryst., 1971, v. A27, № 3, p. 436–452.

Поступила в редакцию
6.1.1982

A PRIORI CALCULATION OF SPATIAL STRUCTURE OF BOVINE PANCREATIC TRYPSIN INHIBITOR

POPOV E. M., GODJAEV N. M., ISMAILOVA L. I., MUSAEV Sh. M.,
ALIEV R. E., AKHMEDOV N. A., MAKSUMOV I. S.

*M. M. Shemyakin Institute of Bioorganic Chemistry, Academy of Sciences
of the USSR, Moscow; Azerbaidjan State University, Baku*

A priori conformational analysis of the bovine pancreatic trypsin inhibitor (BPTI) has been carried out on the basis of the quantitative theory of protein structure. The conformation calculated starting exclusively from the BPTI amino acid sequence is compared to the available X-ray data. Least square deviations of the calculated values from those of the X-ray structure for φ and ψ angles amounted to 13° , and for ω and χ to 9° and 18° , respectively; the difference estimated for distances between all pairs of the backbone C^α atoms was 1.4 \AA , and for hydrogen bond lengths it amounted to 0.5 \AA . Conformational aspects of disulphide bond formation and the mechanism of spontaneous folding are discussed. The role of short, medium and long-range interactions in the spatial organization of a protein globule is analyzed.