



УДК 547.962:541.63

## КОНФОРМАЦИОННЫЕ ВОЗМОЖНОСТИ ПЕПТИДНЫХ ФРАГМЕНТОВ, СОДЕРЖАЩИХ ДИСУЛЬФИДНЫЕ МОСТИКИ

*Ширцхалава М. К.*

*Институт физиологии Академии наук ГССР, Тбилиси*

*Поройков В. В.*

*НИИ по биологическим испытаниям химических соединений,  
Купавна Московской области*

*Туманян В. Г.*

*Институт молекулярной биологии Академии наук СССР, Москва*

Рассмотрены конформационные аспекты образования дисульфидной связи в пептидных фрагментах. Показано, что для пептида  $-Cys-X-Cys-$  замыкание дисульфидного мостика невозможно, если пептидные группы имеют плоскую *транс*-конформацию. В случае тетрапептидов  $-Cys-X-Y-Cys-$  образование дисульфидного мостика резко ограничивает число возможных структур: из структур, возможных априори для циклических форм, могут реализоваться лишь 5. Предсказаны наиболее вероятные конформа-

ции для фрагментов  $-Cys-Glu-Ser-Cys-$ ,  $-Cys-Gln-Asn-Cys-$ ,  $-Cys-Gly-Pro-Cys-$  первичной структуры ДНКазы, пептида задней доли гипофиза и тиоредоксина *E. coli* соответственно.

Согласно общепринятой гипотезе, вся информация о пространственной структуре пептида или белка содержится в его аминокислотной последовательности. Линейный полимер принимает компактную нативную структуру, отвечающую минимуму свободной энергии и стабилизированную вторичными взаимодействиями. В случае восстановления цистеинов до цистинонов эти взаимодействия имеют ковалентный характер.

Естественно пытаться трактовать биологическую роль дисульфидных мостиков в конформационном аспекте. Наличие дополнительных жестких связей ограничивает конформационную подвижность, выделяет или, напротив, запрещает некоторые конформации, наконец, может менять термостабильность молекулы. Как следствие можно ожидать, например, изменения резистентности к протеиназам, усиления (ослабления) ферментативной активности, стабилизации (дестабилизации) четвертичной структуры. Такого рода факторы могут появляться и исчезать в процессе эволюции.

Рассмотрим доступный в настоящее время материал о встречаемости дисульфидных связей в белках с расшифрованной первичной структурой [1]. На рис. 1 представлен график распределения числа дисульфидных связей в зависимости от количества аминокислот между остатками цисти-

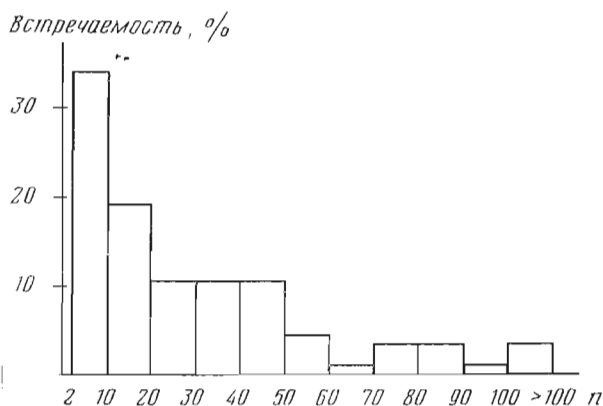


Рис. 1. Гистограмма встречаемости S-S-мостиков между остатками цистеина в зависимости от числа аминокислотных остатков между ними (данные по 45 негомологичным белкам)

на. Выборка включала в себя 45 негомологичных белков, содержащих в сумме 131 дисульфидную связь. Из приведенных данных видно, что более 50% фрагментов с дисульфидными мостиками имеют длину менее 20 ами-

нокислотных остатков, а фрагменты  $\begin{array}{c} \text{—S—S—} \\ | \quad | \end{array}$ ,  $\begin{array}{c} \text{—S—S—} \\ | \quad | \quad | \end{array}$  не встречаются вовсе.

Конформационные возможности фрагмента  $\begin{array}{c} \text{—S—S—} \\ | \quad | \end{array}$  изучали Чандрасекхаран и Баласубрамания [2]. Было показано, что такая дисульфидная связь может образоваться лишь в том случае, если амидная группа между цистеинами находится в *цис*-конформации, к тому же неплохой.

Настоящая работа посвящена конформационному анализу циклических фрагментов  $\begin{array}{c} \text{—S—S—} \\ | \quad | \end{array}$  и  $\begin{array}{c} \text{—S—S—} \\ | \quad | \quad | \end{array}$  (X и Y — остатки любых аминокислот). Фрагмент второго типа встречается в белках относительно часто, тогда как первый до настоящего времени не обнаружен.

Согласно современным представлениям, образование дисульфидного мостика происходит в уже сформировавшейся глобуле, соответствующей ненапряженной, энергетически выгодной структуре [3] (исключение, вероятно, составляют молекулы, имеющие предшественника, например инсулин). Учитывая это обстоятельство, исследование перечисленных фрагментов можно ограничить поиском энергетически выгодных конформаций, для которых замыкание дисульфидной связи происходит без существенного напряжения структуры.

Конформационные расчеты проводили в рамках попарно-аддитивного приближения [4]. Минимизация энергии выполнялась по методу сопряженных градиентов, все углы  $\phi$  и  $\psi$  рассматривались как независимые переменные. Учитывались невалентная, электростатическая и торсионная составляющие. Потенциал цевалентных взаимодействий включал в себя энергию водородной связи. Более подробно методика расчета и использованные параметры описаны в работах [5, 6].

Была введена мера возможности замыкания дисульфидного мостика для некоторой конформации как функция величины относительной конформационной энергии, ширины потенциальной ямы и геометрических условий замыкания цикла. Учет этих трех факторов приводит к следующему

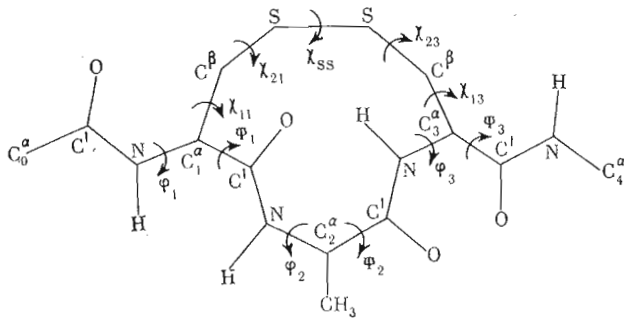


Рис. 2. Структура фрагмента  $\begin{array}{c} | \text{---S---S---} | \\ \text{-Cys-Ala-Cys-} \end{array}$

выражению для вероятности образования циклической структуры:

$$p_j = \frac{\sum_{i=1}^k A_i e^{-E_{ij}/RT}}{\sum_{i=1}^k e^{-E_{ij}/RT}},$$

где  $A_i$  — коэффициент, равный 1, если цикл замыкается, и нулю в противном случае;  $E_{ij}$  — конформационная энергия;  $R$  — универсальная газовая постоянная;  $T$  — абсолютная температура (310 К);  $j$  — номер конформации;  $i$  — индекс, характеризующий вероятность данного конформационного состояния. Для контроля замыкания дисульфидной связи мы использовали аналитический метод Гоу и Шераги [7]. Применение этого метода позволяет сократить число независимых переменных и исследовать геометрические возможности замыкания цикла независимо от расчета энергии

молекулы, а в случае фрагмента  $\begin{array}{c} | \text{---S---S---} | \\ \text{-Cys-Ala-Cys-} \end{array}$  удалось избежать такого расчета.

Фрагмент  $\begin{array}{c} | \text{---S---S---} | \\ \text{-Cys-Ala-Cys-} \end{array}$ . Возможность замыкания S—S-связей в линейном пептиде определяется величиной девяти углов:  $\psi_1$ ,  $\psi_2$ ,  $\psi_3$ ,  $\chi_{11}$ ,  $\chi_{21}$ ,  $\chi_{SS}$ ,  $\chi_{23}$  и  $\chi_{13}$  (рис. 2); после замыкания цикла число независимых переменных уменьшается до 3 [7]. В нашем случае в качестве независимых переменных были выбраны углы  $\psi_2$ ,  $\psi_3$  и  $\psi_1$ . Каждый из этих углов варьировал в пределах от  $-180^\circ$  до  $180^\circ$  с шагом  $10^\circ$ ; значения зависимых торсионных углов рассчитывали по формулам, приведенным в работе [7], при этом проверяли выполнение условия замыкания цикла. Точки, для которых замыкание дисульфидной связи возможно, были нанесены на  $\psi_2$ - $\psi_3$ -карту (рис. 3). На эту же карту нанесены области допустимых конформаций дипептидной единицы глицина, рассчитанные с использованием потенциала жестких сфер [8] и потенциала 6—12 [9]. Из представленных на рис. 3 данных видно, что стерически допустимая область для дипептидной единицы глицина находится вне области, в которой замыкание цикла геометрически возможно. Поскольку дипептидная единица глицина обладает наибольшей конформационной свободой по сравнению с остальными аминокислотными остатками, можно предположить, что замыкание дисульфидной связи во фрагменте -Cys-X-Cys- приведет к крайне напряженной конформации и, следовательно, циклические структуры типа

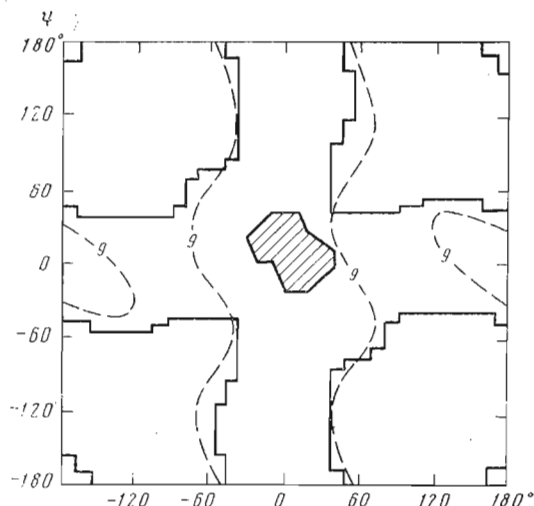


Рис. 3. Конформационная карта стерически разрешенных областей для дипептидной единицы глицина. Сплошная линия — потенциалы жестких сфер, штриховая — потенциалы б-12. Заштрихована область, в которой замыкание дисульфидной связи для фрагмента  $\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---Ala---Cys---} \end{array}$  геометрически возможно.

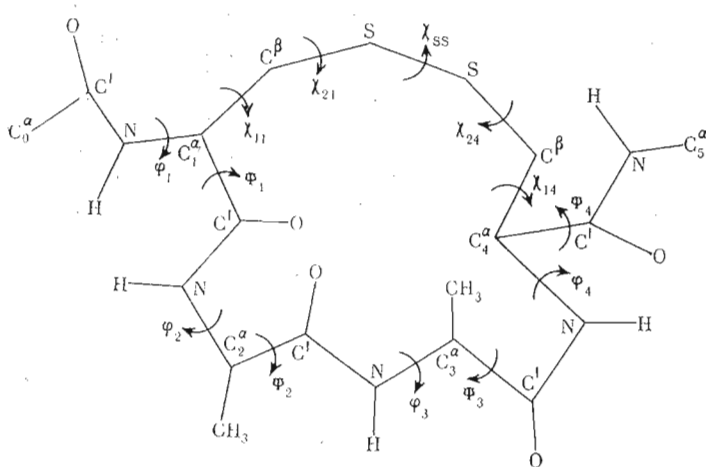


Рис. 4. Структура фрагмента  $\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---Ala---Ala---Cys---} \end{array}$

$\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---X---Cys---} \end{array}$  маловероятны. Это заключение находится в полном согласии с имеющимся экспериментальным материалом: среди расшифрованных к настоящему времени первичных структур ни одна не содержит фрагмента  $\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---X---Cys---} \end{array}$  (см. рис. 1).

Фрагменты  $\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---Ala---Ala---Cys---} \end{array}$ ,  $\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---Ala---Thr---Cys---} \end{array}$ . Для фрагмента  $\begin{array}{c} | \text{---S---S---} | \\ \text{---Cys---Ala---Ala---Cys---} \end{array}$  число переменных, независимо влияющих на

Стартовые конформации аминокислотных остатков, использованные в расчетах \*

Конформационная область	Аланин		Цистеин			Треонин		
	$\varphi$	$\psi$	$\varphi$	$\chi_1$	$\psi$	$\varphi$	$\chi_1$	$\psi$
<i>B</i>	-134	137	-140	60	145	-130	55	135
			-132	-175	140	-145	-175	145
			-130	-50	132	-130	-50	132
<i>R</i>	-115	-45	-108	62	-32	-107	60	-31
			-112	180	-45	-112	-57	-44
			-122	-57	-44			
<i>L</i>	45	42	47	-174	94	30	50	60
			45	-60	40	45	-60	42
<i>H</i>	70	-60						

\* Угол  $\chi_2$  имел значения 180 и 135° для цистеина и треонина соответственно.

замыкание дисульфидной связи, равно 5 и перебор по сетке становится практически невозможным (рис. 4). Поэтому для анализа этого фрагмента мы использовали несколько иной подход. На первой стадии рассматривали фрагмент -Ala-Ala-Ala-Ala-, для которого нашли все локальные минимумы. Точки, из которых проводилась минимизация энергии, получали, комбинируя минимумы *B*, *R*, *L*, *H* дипептидной единицы аланина (табл. 1). Из локальных минимумов тетрааланина выбирали те, для которых расстояние между  $C_1^\alpha$  и  $C_4^\alpha$  не превышало 8,5 Å. Необходимым геометрическим условием замыкания цикла через S-S-мостик при длине связи S-S 2,04 Å. и валентных углах  $\angle C_1^\alpha SS = \angle SSC_4^\alpha = 104^\circ$  является требование, чтобы расстояние между  $C^\alpha$ -атомами остатков цистина не превышало 7 Å. Мы использовали большее значение (8,5 Å) для исследования окрестности каждого локального минимума и вычисления вероятности замыкания цикла. При вычислении величин  $p_j$  варьировались лишь независимые (с точки зрения замыкания цикла) углы  $\varphi_2, \varphi_3, \varphi_4, \psi_2, \psi_3, \psi_4$ . К каждому из углов  $\varphi_2, \varphi_3, \varphi_4, \psi_2, \psi_3, \psi_4$  прибавляли и вычитали 20°, комбинируя получаемые значения между собой. Угол  $\varphi_4$  перебирали по сетке с шагом 20° в интервалах  $-150^\circ \leq \varphi_4 \leq -90^\circ$  (*B*- и *R*-минимумы) и  $30^\circ \leq \varphi_4 \leq 90^\circ$  (*L*- и *H*-минимумы). Углы  $\varphi_1$  и  $\psi_4$  не влияли на замыкание цикла и принимали значения  $\varphi_1 -130^\circ, \varphi_1 140^\circ$  при  $\varphi_4 < 20^\circ$  или  $\psi_4 50^\circ$  при  $\varphi_4 \geq 20^\circ$ . Угол  $\psi_1$  принимал значение 140° (вытянутая конформация), если цикл не замыкался, или значение, которое он получал как зависимый угол при замыкании цикла. Таким образом, в окрестности каждого локального минимума тетрааланина было рассмотрено около 320 точек.

В табл. 2 приведены локальные минимумы, для которых вероятность замыкания цикла превосходит 0,01. Торсионные углы для первого остатка не указаны, так как их значения не влияют независимым образом на образование цикла. Из торсионных углов четвертого остатка лишь угол  $\varphi_4$  влияет на замыкание цикла, что дает возможность объединить для него области *B* и *R*, а также *L* и *H*. Из данных табл. 2 видно, что замыкание дисульфидной связи наиболее вероятно для конформации  $B_2L_3$  (*B*+*R*)<sub>4</sub>. Для иных конформаций образование дисульфидной связи, хотя и возможно, представляется маловероятным, поскольку для структур (2) и (3) необходимы формы *H*, которые обычно значительно проигрывают по энергии и могут реализоваться лишь при наличии глицина в соответствующих положениях, а вероятность образования цикла для структур (4) и (5) пренебрежимо мала. Образование дисульфидного мостика для конформаций  $R_2R_3$  (*B*+*R*)<sub>4</sub> и  $R_2R_3$  (*L*+*H*)<sub>4</sub> гораздо менее вероятно (значения  $p_j$  равны

Вероятность образования циклической структуры ( $P_i$ ) для различных конформаций -Ala-Ala-Ala-Ala-

$P_j$	Номер конформации	Конформационная область		
		Ala <sup>2</sup>	Ala <sup>3</sup>	Ala <sup>4</sup>
0,86	1	<i>B</i>	<i>L</i>	<i>B+R</i>
0,12	2	<i>B</i>	<i>H</i>	<i>L+H</i>
0,09	3	<i>H</i>	<i>R</i>	<i>L+H</i>
0,04	4	<i>R</i>	<i>B</i>	<i>R+B</i>
0,03	5	<i>L</i>	<i>H</i>	<i>L+H</i>

$2,5 \cdot 10^{-3}$  и  $1,2 \cdot 10^{-4}$  соответственно), что позволяет исключить соответствующие конформации из дальнейшего рассмотрения.

Результаты, полученные для тетрааланина, были использованы как исходные данные для конформационного анализа фрагмента -Cys-Ala-Ala-Cys-. Конформации остова, для которых возможно замыкание цикла, могут быть непосредственно перенесены на фрагмент, содержащий цистеины, а конформации бокового радикала цистеина были взяты из таблицы локальных минимумов (табл. 1). Затем с помощью метода Гоу и Шераги для каждого набора углов проверялась возможность замыкания цикла. Входными параметрами служили углы  $\varphi_2, \varphi_3, \varphi_4$ . Как видно из табл. 3, замыкание цикла наиболее вероятно для конформации (1) (см. табл. 2). Значения независимых углов в этом случае остаются в пределах локальных минимумов, значение зависимого параметра  $\varphi_1$  ( $-80^\circ < \varphi_1 < 40^\circ$ ) также попадает в разрешенную область *B*-минимума.

Для конформаций (2) и (3) замыкание цикла представляется вероятным лишь при наличии глицина в цепи.

На последнем этапе работы был выполнен конформационный анализ

фрагмента  $\begin{array}{c} \text{—S—S—} \\ | \quad | \\ \text{-Cys-Ala-Thr-Cys-} \end{array}$  первичной структуры тиоредоксин-редуктазы [1]. Как и выше, стартовые точки брали из табл. 1, 2. В табл. 3 приведены предпочтительные конформации этого фрагмента. Они сравнительно мало проигрывают по энергии, отвечающей глобальному минимуму вытянутой конформации (все остатки в *B*-области). Следует отметить,

что результаты, полученные для фрагмента  $\begin{array}{c} \text{—S—S—} \\ | \quad | \\ \text{-Cys-Ala-Ala-Cys-} \end{array}$  практически совпали с данными, полученными для фрагмента

$\begin{array}{c} \text{—S—S—} \\ | \quad | \\ \text{-Cys-Ala-Thr-Cys-} \end{array}$ . Это не удивительно, поскольку для фрагментов типа -Cys-X-Y-Cys- уже при наличии боковых радикалов аланина, содержащих лишь метильную группу, вероятность замыкания дисульфидного мостика резко различается для различных конформаций. Из 64 ( $2 \times 4 \times 4 \times 2$ ) теоретически возможных структур лишь для пяти вероятность замыкания цикла превосходит 0,01 (табл. 2). При этом три из них реализуются лишь при наличии глицина, а для структуры (4) вероятность замыкания цикла составляет всего 0,04. Следует ожидать, что при увеличении размеров боковых радикалов наблюдаемая резкая дифференциация структурных форм (с точки зрения возможности замыкания цикла) может лишь усиливаться. Поэтому полученные нами результаты справедливы для любых боковых радикалов X и Y.

Полученные в настоящей работе результаты дают возможность предсказать наиболее вероятные конформации входящих в белки фрагментов

Таблица 3

$\overline{\text{rS-S}}$   
 Предпочтительные конформации фрагментов -Cys-Ala-Ala-Cys- и -Cys-Ala-Thr-Cys-

Фрагмент	Конформационная область	Торсионные углы, град *											E <sub>отн</sub> , ккал/моль	
		φ <sub>1</sub>	χ <sub>11</sub>	ψ <sub>1</sub>	φ <sub>2</sub>	ψ <sub>2</sub>	φ <sub>3</sub>	χ <sub>13</sub>	χ <sub>23</sub>	ψ <sub>3</sub>	φ <sub>4</sub>	χ <sub>14</sub>		ψ <sub>4</sub>
-Cys-Ala-Ala-Cys-	1 (лок. мин.)	-129	71	135	-127	128	48	-	-	43	-110	-71	133	3
$\overline{\text{rS-S}}$ -Cys-Ala-Ala-Cys-	1 (замкн. конф.)	-129	-137	163	-127	128	48	-	-	43	-105	-87	133	6
-Cys-Ala-Ala-Cys-	2 (лок. мин.)	-128	-71	132	-127	127	49	-	-	39	-103	-71	-39	6
$\overline{\text{rS-S}}$ -Cys-Ala-Ala-Cys-	2 (замкн. конф.)	-128	-119	153	-127	127	49	-	-	39	-93	-120	-39	6
-Cys-Ala-Thr-Cys-	1 (лок. мин.)	-133	-168	148	-108	156	41	53	151	46	-107	60	-38	6
$\overline{\text{rS-S}}$ -Cys-Ala-Thr-Cys-	1 (замкн. конф.)	-133	-169	150	-108	156	41	53	151	46	-107	-20	-38	7
-Cys-Ala-Thr-Cys-	2 (лок. мин.)	-136	-178	106	-141	122	35	-33	154	65	-146	58	149	7
$\overline{\text{rS-S}}$ -Cys-Ala-Thr-Cys-	2 (замкн. конф.)	-136	-150	178	-141	122	35	-33	154	65	-146	14	149	7

\* Значение угла χ<sub>3</sub> листеново было ~180°.

$\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-X-Y-Cys-} \end{array}$  . Так, для фрагментов  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Gly-Ser-Cys-} \end{array}$  из ДНКазы и  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Gln-Asn-Cys-} \end{array}$  из пептида задней доли гипофиза [1] наиболее вероятно структура (1). Наличие в цепи остатка пролина вносит дополнительные конформационные ограничения. Для пептидов типа  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Pro-X-Cys-} \end{array}$  и  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-X-Pro-Cys-} \end{array}$  (последовательность такого типа  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Gly-Pro-Cys-} \end{array}$  имеется в тиоредоксине *E. coli*) возможны структуры (1), (2), (4) и (4), (3) соответственно. К сожалению, сопоставить эти предсказания с экспериментом в настоящее время не представляется возможным из-за отсутствия соответствующих рентгенографических данных.

Таким образом, проведенные расчеты показали, что для фрагмента  $-\text{Cys-X-Cys-}$  замыкание дисульфидной связи невозможно при плоской *транс*-пептидной группе. Для фрагмента  $-\text{Cys-X-Y-Cys-}$  образование дисульфидной связи накладывает жесткие ограничения на число возможных структур. Если X или/и Y — глицин, могут реализоваться 5 из 64 возможных конформаций, если другие аминокислоты — лишь 2.

Авторы выражают искреннюю признательность С. А. Шерману за предоставление комплекса программ, с использованием которых была выполнена часть расчетов.

#### ЛИТЕРАТУРА

- Dayhoff M. O. (1972) Atlas of Protein Sequence and Structure, vol. 5; (1973) suppl. 1; (1976) suppl. 2; National Biomed. Res. Found., Georgetown University Med. Centre, Washington.
- Chandrasekharan R., Balasubramanian B. (1969) Biochim. et biophys. acta. 188, 1-9.
- Geisow M. J. (1977) Nature, 268, 9-10.
- Дашевский В. Г. (1974) Конформации органических молекул, «Химия», М.
- Галактионов С. Г., Никифорович Г. В., Перельман Т. Л. (1974) Диффузия в сложных молекулярных структурах, «Наука и техника», Минск.
- Шерман С. А. (1977) Канд. дис. «Расчет пространственных структур олигопептидных фрагментов белков и биологических пептидов», Ин-т биоорганич. химии АН БССР, Минск.
- Go N., Scheraga H. A. (1970) Macromolecules, 3, 178-187.
- Scheraga H. A. (1968) Adv. Phys. Org. Chem., 6, 103-184.
- Zimmerman S. S., Pottle M. S., Nemethy G., Scheraga H. A. (1977) Macromolecules, 10, 1-9.

Поступила в редакцию  
28.IX.1979

После доработки  
28.XI.1979

### CONFORMATIONAL POSSIBILITIES OF PEPTIDE FRAGMENTS CONTAINING DISULFIDE BRIDGES

PIRTSKHALAVA M. K., POROIKOV V. V., TUMANYAN V. G.  
*Institute of Physiology, Academy of Sciences of the Georgian SSR, Tbilisi;*  
*Institut for Biological Testing of Chemical Compounds, Kupavna;*  
*Institut of Molecular Biology, Academy of Sciences*  
*of the USSR, Moscow*

Conformational aspects of disulfide bond formation in peptide fragments were analyzed. In peptides Cys-X-Cys the closure of disulfide bridge is impossible if the peptide group have a planar *транс*-configuration. For Cys-Y-Cys tetrapeptides (where X and Y stand for any amino acid residues), the disulfide bond formation considerably limits the number of possible conformations. Only five conformational states may be realized for cyclic peptides out of 64 (2×4×4×2) a priori possible. If X and/or Y are proline residues, additional conformational constraints appear. The most probable con-

$\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Glu-Ser-Cys-} \end{array}$ ,  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Gln-Asn-Cys-} \end{array}$ ,  $\begin{array}{c} \text{S} - \text{S} \\ | \quad | \\ -\text{Cys-Gly-Pro-Cys-} \end{array}$   
 formations were predicted for the fragments of deoxyribonuclease, posterior pituitary peptide and *E. coli* thioredoxin, respectively.