



УДК 547.962:541.63

ПОДХОД К РАСЧЕТУ СТАБИЛЬНЫХ КОНФОРМАЦИЙ  
ОЛИГОПЕПТИДОВ

## I. ДИ- И ТРИПЕПТИДЫ С ОСТАТКАМИ АЛАНИНА И ФЕНИЛАЛАНИНА

*Попов Е. М., Ахмедов Н. А., Липкинд Г. М.**Институт биоорганической химии им. М. М. Шемлякина  
Академии наук СССР, Москва*

Исследованы конформационные возможности фрагментов  $C^{\alpha}$ -CO-*L*-Ala-*L*-Ala-*L*-Ala-NH- $C^{\alpha}$ ,  $C^{\alpha}$ -CO-*L*-Phe-*L*-Ala-NH- $C^{\alpha}$ ,  $C^{\alpha}$ -CO-*L*-Ala-*L*-Phe-NH- $C^{\alpha}$ . Показано, что взаимодействия между соседними аминокислотными остатками приводят к заметной энергетической дифференциации пространственных форм. Для молекулы, состоящей из остатков аланина предпочтительны свернутые формы основной цепи *R-R-R* и *R-R-B*. При наличии остатков фенилаланина, благодаря значительным дисперсионным взаимодействиям между ароматическим боковым радикалом и пептидными единицами, в растянутых формах основной цепи предпочтительными становятся конформации *B-B* и *B-R*. Проведено сравнение конформационных состояний изученных фрагментов и соответствующих ди- и трипептидных фрагментов белков.

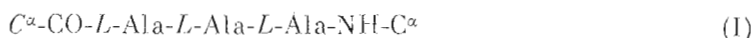
Известна серия работ [1—8], посвященных теоретическому анализу пространственного строения метиламидов *N*-ацетил- $\alpha$ -аминокислот. Сопоставление результатов расчета свободных монопептидов с пространственным строением остатков в более сложных соединениях выявило чрезвычайно важную роль в формировании структуры пептидной цепи взаимодействий атомов в пределах одного остатка (назовем их ближними взаимодействиями). В белках аминокислотные остатки имеют конформации, которые соответствуют предпочтительным формам метиламидов *N*-ацетил- $\alpha$ -аминокислот. Влияние взаимодействий между соседними в последовательности остатками (средние взаимодействия) и между остатками, удаленными по цепи, но пространственно сближенными в глобуле (дальние взаимодействия), проявляется в выборе конформационного состояния остатка из набора оптимальных форм монопептида [9, 10].

Многие природные остатки обладают большим числом близких по энергии стабильных конформаций. (Для оценки общей ситуации можно принять, что в среднем один остаток имеет 10 низкоэнергетических оптимальных форм.) Это обстоятельство исключает решение структурных задач сложных пептидов на основе прямого использования данных по монопептидам. В настоящем и следующем сообщениях мы рассматриваем вопросы, касающиеся конформационного анализа олигопептидных молекул, требующего учета ближних и средних взаимодействий.

Результаты экспериментальных [11—12] и теоретических [13—14] работ показали, что данный остаток в пептидной цепи ( $n$ ) может испытывать заметное влияние 3—4 предшествующих ( $n-1$ ) — ( $n-4$ ) и последующих ( $n+1$ ) — ( $n+4$ ) остатков, причем взаимодействия с ближайшими

соседями, как правило, являются самыми сильными. Ввиду существенных расчетных трудностей в качестве первого шага представляет интерес исследовать характер взаимодействий между соседними остатками и тенденцию к уменьшению числа конформационных состояний с удлинением цепи. Этот вопрос прежде всего целесообразно рассмотреть на примере ди- и трипептидов [15]. В том случае, если взаимодействия между двумя остатками приводят к заметной энергетической дифференциации простейших форм дипептидов [16] и если взаимодействия между остатками  $n - 1$  и  $n$ ,  $n$  и  $n + 1$  оказывают на конформационное состояние трипептида более существенное влияние, чем взаимодействия между его крайними звеньями ( $n - 1$ ) и ( $n + 1$ ), то предварительный расчет ди- и трипептидных фрагментов может оказаться плодотворным в решении конформационной проблемы сложных олигопептидных молекул.

Нами выполнен конформационный анализ следующих ди- и трипептидных фрагментов:



Расчет фрагментов (II) и (III) дает возможность выяснить по отдельности влияние предшествующего и последующего остатков фенилаланина (или аланина) на конформационное состояние аланина (или фенилаланина). В трипептидных фрагментах воздействие на состояние центрального остатка осуществляется одновременно с обеих сторон. Анализ всех фрагментов выполнен на основе полученных ранее данных по монопептидам [6, 8]. Включение остатка фенилаланина обусловлено наличием у него объемной и вместе с тем лабильной боковой цепи, склонной вступать во взаимодействия с соседними остатками. Выбор аланина связан со значительной конформационной свободой его основной цепи, что позволяет осуществлять эффективные взаимодействия между концевыми звеньями трипептида.

Предварительно для выяснения характера взаимодействий атомов основной цепи рассмотрен трипептидный аланиновый фрагмент (I). Анализ аланинового дипептида был сделан ранее [18].

Кроме того, в настоящем сообщении изложены и обсуждены результаты расчета двух дипептидов с остатками фенилаланина и аланина.

*Параметры и потенциальные функции.* Длины связей в основной цепи приняты равными параметрам Полинга-Кори [19]; валентные углы пептидной группы соответствуют значениям, предложенным в [5]. Длины связей C—N и C $\equiv$ C бензольного кольца равны соответственно 1,0 и 1,4 Å; угол C $\alpha$ C $\beta$ C $\gamma$  = 113° [2].

Расчет выполнен при учете невалентных и электростатических взаимодействий, торсионного вклада. Невалентные взаимодействия оценивались по потенциалу Букингема (6-*exp*) с параметрами Дашевского [20]. Электростатическая энергия рассчитывалась по закону Кулона с использованием распределения зарядов на атомах, предложенного Поландом и Шерагой [21]. Конформационные состояния молекул исследованы применительно к полярной среде. В связи с этим величина диэлектрической проницаемости принята равной 10 [6]. По этой же причине не учитывалась возможность образования сильно ослабленных в этих условиях внутримолекулярных водородных связей. В работе рассмотрены только развернутые конформации молекул. Торсионные потенциалы и величины барьер-

\* См. сообщение II в этом же номере журнала, стр. 1280—1290.

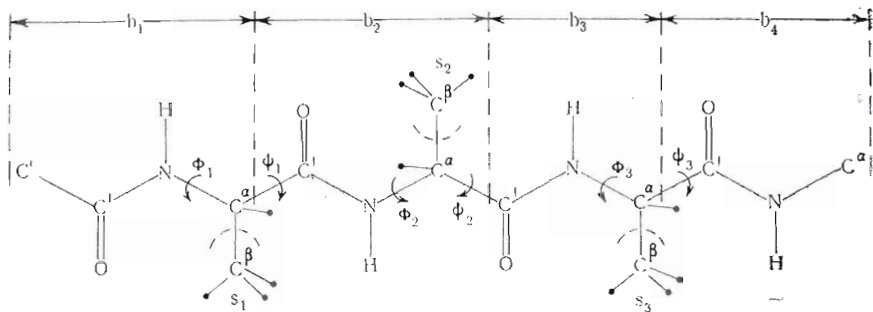


Рис. 1. Модель фрагмента  $C^\alpha\text{-CO-L-Ala-L-Ala-L-Ala-NH-C}^\alpha$

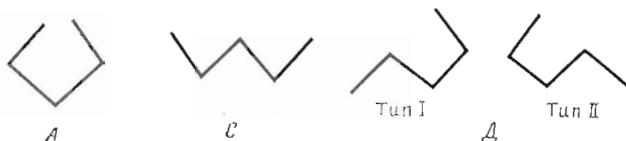


Рис. 2. Геометрия основной цепи в формах *A*, *C*, *D*

ров взяты такими же, как в работе [22]. Наличие небольшого барьера вращения вокруг связи  $C^\beta-C^\gamma$  ( $\sim 0,2$  ккал/моль) не принималось во внимание. Это позволило исключить из рассмотрения большое число локальных конформеров по  $\chi^2$ \*, не представляющих интереса.

Для боковой цепи остатков фенилаланина в качестве нулевых приближений при минимизации потенциальной энергии использованы параметры

Таблица 1

Геометрические и энергетические параметры оптимальных конформаций молекулы *Ac-L-Ala-NHMe*

Конформация	Углы вращения, град		Энергия, ккал/моль
	$\varphi$	$\psi$	
<i>B</i>	-115	140	0
<i>R</i>	-90	-54	0,3
<i>L</i>	52	60	1,4

предпочтительных конформеров молекулы *Ac-L-Phe-NHMe*, в которых угол  $\chi^1$  имеет значения  $\sim 60, 180$  и  $-60^\circ$ , а  $\chi^2$  — около  $90^\circ$  [8]. Для основной цепи исходными параметрами служили результаты расчета *Ac-L-Ala-NHMe* [6] и *Ac-L-Phe-NHMe* [8]. Для основной цепи каждого остатка учитывались три возможных состояния, находящиеся на конформационной карте  $\varphi-\psi$  в низкоэнергетических областях *R* ( $\varphi, \psi -180, -0^\circ$ ), *B* ( $\varphi -180, -0^\circ, \psi, 0, -180^\circ$ ) и *L* ( $\varphi, \psi 0, -180^\circ$ ).

Поиск минимума осуществлялся по вторым производным [24]. Наряду с общей энергией рассчитывались отдельные вклады от скелетных взаимодействий ( $E_{b_i-b_j}$ ), от взаимодействий в пределах одного остатка ( $E_{(b_i-s_i)+(b_{i+1}-s_i)}$ ), от взаимодействий между боковыми цепями ( $E_{s_i-s_j}$ ) и между основными и боковыми цепями разных остатков ( $E_{b_i-s_j(j \neq i, i+1)}$ ).

Расчет конформационных состояний фрагмента  $C^\alpha\text{-CO-L-Ala-L-Ala-L-Ala-NH-C}^\alpha$  (*I*). Модель фрагмента при значениях углов  $\varphi$  и  $\psi$ , равных  $180^\circ$ , дана на рис. 1;  $b_1 - b_4$  — звенья пептидного скелета,  $s_1 - s_3$  — боковые цепи. В исходных для минимизации приближениях геометрия каждого остатка соответствовала одной из оптимальных форм *B*, *R* или *L* молекулы *Ac-L-Ala-NHMe* (табл. 1). В табл. 2 представлены результаты минимизации — геометрические параметры всех 27 возможных оптимальных форм, величины общей энергии ( $E_{\text{общ}}$ ) и ее составляющие. Все конформации аланинового трипептида в зависимости от формы основной цепи разделены

\* Отсчет углов  $\varphi, \psi$  и  $\chi$  соответствует номенклатуре [23].

Геометрические и энергетические параметры оптимальных конформаций фрагмента  $C^{\alpha}-CO-L-Ala-L-Ala-L-Ala-NH-C^{\alpha}$

Конформация		Углы вращения, град					Энергии взаимодействия, ккал/моль			$E^0$ ккал/моль	Формы сменных дипетидов (табл. 3)	Конформации трипептид- ных фрагмен- тов в белках	
обо- значе- ние	тип	$\varphi_1$	$\psi_1$	$\varphi_2$	$\psi_2$	$\varphi_3$	$\psi_3$	$E_{b_i-b_j}$	$E_{b_i-s_j} (i \neq j, i+1)$				$E_{s_i-s_j}$
Свернутые формы основной цепи													
$A_1^0$	$R-R-R$	-68	-44	-60	-37	-65	-44	11,8	21,4	-1,3	0	$a_1^0-a_1^0$	308
$A_2^0$	$R-R-B$	-55	-51	-78	-42	-103	72	10,9	22,1	-1,1	0	$a_1^0-a_2^0$	47
$A_3^0$	$R-B-L$	-78	-51	-82	93	52	58	11,5	22,6	-0,6	1,5	$a_2^0-a_3^0$	4
$A_4^0$	$B-L-L$	-64	139	51	45	52	54	11,7	22,6	-0,9	1,7	$a_3^0-a_4^0$	2
$A_5^0$	$L-L-L$	53	46	49	51	52	53	12,3	23,1	-0,1	2,3	$a_4^0-a_4^0$	—
Полусвернутые формы основной цепи													
$D_1^0$	$R-B-R$	-57	-51	-100	51	-135	-59	12,2	21,8	-1,5	0,4	$a_1^0-c_3^0$	34
$D_2^0$	$B-R-R$	-108	133	-55	-44	-80	-44	13,4	21,0	-1,7	0,9	$c_3^0-a_1^0$	49
$D_3^0$	$R-B-B$	-56	-50	-102	66	-111	148	12,7	21,8	-1,7	0,9	$a_2^0-c_1^0$	41
$D_4^0$	$B-R-B$	-108	148	-58	-46	-112	52	12,8	21,5	-1,5	0,9	$c_2^0-a_2^0$	42
$D_5^0$	$B-L-R$	-66	119	51	42	-120	-60	12,9	21,9	-1,8	1,0	$a_3^0-c_5^0$	1
$D_6^0$	$B-L-B$	-65	124	51	44	-112	146	13,4	21,9	-2,0	1,3	$a_3^0-c_3^0$	8
$D_7^0$	$L-L-R$	53	46	48	51	-147	-61	13,6	22,2	-1,5	2,1	$a_4^0-c_3^0$	—
$D_8^0$	$B-B-L$	-87	101	-66	103	53	45	12,3	22,9	-1,2	2,3	$c_1^0-a_3^0$	6
$D_9^0$	$R-R-L$	-56	-48	-65	-42	52	48	14,2	21,6	-1,8	2,4	$a_1^0-c_4^0$	4
$D_{10}^0$	$L-L-B$	53	45	60	54	-114	147	13,9	22,1	-1,5	2,5	$a_4^0-c_3^0$	3
$D_{11}^0$	$L-R-R$	52	69	-52	-43	-81	-42	14,3	21,9	-1,4	2,7	$c_5^0-a_1^0$	—
$D_{12}^0$	$R-R-B$	51	69	-54	-48	-107	57	13,4	22,2	-1,1	2,8	$c_3^0-a_2^0$	2
$D_{13}^0$	$R-L-L$	-95	-50	50	55	54	49	14,4	22,4	-1,4	3,2	$c_4^0-a_4^0$	1
$D_{14}^0$	$L-B-L$	51	61	-65	106	53	48	13,1	23,1	-0,1	3,6	$c_3^0-a_3^0$	2
Развернутые формы основной цепи													
$C_1^0$	$B-B-B$	-88	93	-88	87	-111	148	12,6	22,2	-0,8	2,1	$c_1^0-c_1^0$	164
$C_2^0$	$B-B-R$	-88	93	-102	122	-102	-54	13,4	21,5	-0,6	2,4	$c_1^0-c_2^0$	52
$C_3^0$	$R-L-B$	-96	-51	51	41	-119	147	15,2	21,2	-1,5	2,8	$c_4^0-c_3^0$	2
$C_4^0$	$R-L-L$	-95	-50	51	37	-160	-59	15,8	21,3	-2,2	2,9	$c_4^0-c_5^0$	—
$C_5^0$	$B-R-L$	-88	90	-97	-52	52	49	14,0	22,4	-1,0	3,3	$c_2^0-c_4^0$	—
$C_6^0$	$L-B-B$	50	48	-115	153	-100	114	14,6	21,5	-0,9	3,4	$c_3^0-c_1^0$	9
$C_7^0$	$L-B-R$	50	49	-106	115	-101	-54	14,7	21,7	-0,1	3,5	$c_3^0-c_2^0$	1
$C_8^0$	$L-R-L$	50	48	-110	-56	52	49	15,7	22,1	-1,2	4,7	$c_5^0-c_4^0$	—

Таблица 3

## Геометрические и энергетические параметры оптимальных конформаций молекулы Ac-L-Ala-L-Ala-NHMe

Конформация		Углы вращения, град				$E_{\text{общ}}$ ккал/моль
обозначение	тип	$\varphi_1$	$\psi_1$	$\varphi_2$	$\psi_2$	
Свернутые формы						
$a_1^0$	R-R	-54	-47	-78	-44	0
$a_2^0$	R-B	-87	-54	-110	50	0,2
$a_3^0$	B-L	-73	125	50	40	0,8
$a_4^0$	L-L	51	56	52	44	1,3
Развернутые формы						
$c_1^0$	B-B	-92	106	-94	82	0,3
$c_2^0$	B-R	-98	76	-90	-56	0,6
$c_3^0$	L-B	50	44	-114	158	1,3
$c_4^0$	R-L	-103	-47	50	44	1,6
$c_5^0$	L-R	56	44	-125	-53	1,7

Таблица 4

## Геометрические и энергетические параметры оптимальных конформаций молекулы Ac-L-Phe-NHMe

Конформация	Углы вращения, град			$E$ , ккал/моль
	$\varphi$	$\psi$	$\chi^1$	
B	-115	164	60	0
	-110	152	180	0
	-124	150	-60	0
R	-99	-33	60	0,4
	-113	-60	180	0,3
	-137	-60	-60	0,1
L	40	44	60	2,6
	54	76	180	2,7
	53	60	-60	1,3

на три группы. Группу  $A^0$  составляют конформации, в которых звено  $b_1$  сближено с  $b_3$  и  $b_4$ , а  $b_2$  — с  $b_4$ ; они являются комбинациями свернутых форм смежных дипептидов R-R, R-B, B-L и L-L (табл. 3,  $a^0$ ). Конформации в группе  $C^0$  образованы из развернутых дипептидных форм B-B, B-R, L-B, R-L и L-R (табл. 3,  $c^0$ ). В группу  $D^0$  входят конформации, которые являются комбинациями свернутых и развернутых дипептидных форм ( $a^0$  и  $c^0$ ). Целесообразность такого разделения обусловлена наличием у конформаций каждой группы специфических средних взаимодействий. На рис. 2 представлено схематическое изображение геометрии основной цепи в указанных формах.

Самыми выгодными конформациями алаиногого трипептида как по величине  $E_{\text{общ}}$ , так и по  $E_{b_i-b_j}$  являются R-R-R и R-R-B ( $A_1^0$ ,  $A_2^0$ ). Сближенность звеньев основной цепи приводит к возникновению дисперсионных взаимодействий между  $b_1$  и  $b_3$ ,  $b_2$  и  $b_4$ ,  $b_1$  и  $b_4$ , энергия которых составляет около  $-4$  ккал/моль. Дополнительная стабилизация на трипептидном фрагменте за счет взаимодействия между первым и четвертым пептидными звеньями  $E_{b_1-b_4}$  равна  $-1,5$  ккал/моль. Такие же дисперсионные взаимодействия характерны и для других конформаций этой группы — R-B-L, B-L-L и L-L-L. Меньшая предпочтительность последних связана с невыгодными в форме L близкими взаимодействиями  $E_{(b_i-s_i)+(b_{i+1}-s_i)}$  (табл. 2).

Дисперсионные взаимодействия в основной цепи имеют место и в конформациях группы  $D^0$  (2—3 ккал/моль). Однако в отличие от  $A^0$  они осуществляются в пределах одного из дипептидных фрагментов. Низкоэнергетические конформации в  $A^0$  и  $D^0$ , помимо дисперсионных взаимодействий в основной цепи, стабилизированы также за счет взаимодействий  $b_1 - s_3$  и  $b_4 - s_1$  ( $\sim -1$  ккал/моль).

Полностью вытянутые конформации аланинового трипептида ( $C^0$ ) имеют максимальные значения  $E_{\text{общ}}$ . Так,  $B-B-B$  и  $B-B-R$  проигрывают  $R-R-R$  и  $R-R-B$  приблизительно 2 ккал/моль. Это происходит из-за отсутствия дополнительных стабилизирующих взаимодействий между звеньями основной цепи. Вытянутая форма скелета ослабляет в аланиновом трипептиде взаимодействия типа  $b_i - s_j$  ( $j \neq i, i + 1$ ). Таким образом, в случае  $C^0$  конформационная энергия трипептида складывается аддитивно из энергий ближних взаимодействий.

Наименьшая потенциальная энергия  $R-R-R$  и  $R-R-B$  не означает, что вес этих конформаций должен быть подавляющим. Во-первых, энергия некоторых других форм лишь незначительно выше  $R-R-R$  и  $R-R-B$ . Во-вторых, инициация витка  $\alpha$ -спирали связана с большим проигрышем энтропии. Только начиная с декапептида [11, 12] энталпийный выигрыш становится достаточным для ее преимущественной стабилизации. Поэтому представляется неправомерным непосредственное сравнение результатов энталпийного расчета аланинового тетрапептида с экспериментом, как это описано в работах [25, 26]. Величина энергии  $E_{\text{общ}}$  является наиболее объективным критерием при оценке относительной стабильности конформеров в пределах однотипных геометрических форм (свернутых, полусвернутых и развернутых, табл. 2).

Сравним конформационные состояния остатков аланина в дипептидных и трипептидных фрагментах (табл. 2 и 3). Глобальные конформации аланинового трипептида  $R-R-R$  и  $R-R-B$  представляют собой комбинации наиболее выгодных форм дипептида —  $R-R$  ( $a_1^0$ ) и  $R-B$  ( $a_2^0$ ), которые дополнительно стабилизируются взаимодействиями между  $b_1$  и  $b_4$ . Свернутые конформации трипептида с большей энергией являются сочетаниями менее выгодных свернутых дипептидных форм. В табл. 2 приведены формы дипептидов, образующих конформацию трипептида. Во всех случаях ситуация, отмеченная для конформаций группы  $A^0$ , характерна также для группы  $C^0$  и  $D^0$ .

Сопоставим конформации рассчитанного свободного трипептидного фрагмента с геометрией основных цепей белковых фрагментов, включающих три остатка аланинового типа (все остатки за исключением глицина и пролина). В табл. 2 приведены также данные о частоте встреч данной конформации основной цепи у трипептидных фрагментов, X-Y-Z — в белках \*. Прежде всего отметим, что все наблюдаемые в белках формы основной цепи X-Y-Z соответствуют конформации аланинового трипептида. Большинство фрагментов (98%) попадает в интервал 0—2,5 ккал/моль. Белковые фрагменты достаточно полно представлены во всех группах (361 —  $A^0$ , 228 —  $C^0$ , 193 —  $D^0$ ). Наиболее часто встречаются формы  $R-R-R$  и  $B-B-B$ . Поскольку сопоставляемые участки белков и аланиновый трипептид имеют лишь сходные основные цепи (включая атомы  $C^\beta$ ), а в остальном существенно различаются, то сравнительно малый энергетический интервал свидетельствует о том, что реализующиеся в белках конформационные состояния фрагментов -X-Y-Z- являются близкими к оптимальным как в отношении скелетных взаимодействий, так и в отношении взаимодействий между основными и боковыми цепями соседних и удаленных остатков.

\* Использованы значения углов вращения  $\varphi$  ( $C^\alpha - N$ ) и  $\psi$  ( $C^\alpha - C'$ ) в миоглобине [27], лизоциме [28], инсулине [29],  $\alpha$ -хамотрипсине [30], цитохроме *c* [31] и карбоксипептидазе [32].

Геометрические и энергетические параметры оптимальных конформаций фрагмента  $C^{\alpha}-CO-L-Rhe-L-Ala-NH-C^{\alpha}$ 

Конформация	Углы вращения, град						Энергия взаимодействия, ккал/моль				$E_{общ}$ , ккал/моль	Конформация дисперсных фрагментов в белках	
	тип	$\varphi_1$	$\psi_1$	$\chi_1^1$	$\chi_1^2$	$\varphi_2$	$\psi_2$	$E_{b_i-b_j}$	$E(b_i-s_i) + (b_{i+1}-s_i)$	$E_{b_i-s_j} (i \neq i, i+1)$			$E_{s_i-s_j}$
Развернутые формы основной цепи													
$c_1'$		-109	149	179	87	-94	88	9,0	13,7	-2,8	-0,3	0	7
$c_2'$	$B_2-B$	-109	147	176	88	-119	-59	10,1	12,8	-2,7	-0,2	0,3	3
$c_3'$	$B_1-B$	-129	178	69	94	-85	92	9,4	13,5	-2,6	-0,2	0,6	3
$c_4'$	$B_1-R$	-130	173	67	93	-98	-57	10,3	12,6	-2,1	-0,2	1,0	3
$c_5'$	$B_3-B$	-130	162	-60	90	-111	149	10,2	12,5	-0,3	-0,1	2,6	16
$c_6'$	$B_3-R$	-130	159	-60	90	-95	-51	10,2	12,7	-0,5	-0,1	2,7	4
$c_7'$	$R_2-L$	-100	-49	176	64	51	58	10,5	14,1	-2,2	-0,1	3,3	1
$c_8'$	$R_1-L$	-109	-20	56	90	53	77	11,8	14,0	-2,4	-0,1	3,6	—
$c_9'$	$L_3-B$	53	44	-56	90	-115	457	10,3	13,9	-0,6	-0,1	4,0	—
$c_{10}'$	$R_3-L$	-105	-53	-60	90	51	45	10,6	14,2	-0,8	-0,1	4,2	—
$c_{11}'$	$L_3-R$	52	44	-57	91	-124	-58	10,6	14,0	-0,6	-0,1	4,3	—
$c_{12}'$	$L_2-B$	52	47	-173	67	-115	56	9,7	14,9	-0,6	-0,1	5,4	—
$c_{13}'$	$L_1-B$	38	40	51	93	-114	157	12,1	14,0	-0,9	-0,2	5,6	—
$c_{14}'$	$L_2-R$	52	45	-172	66	-123	-58	10,7	15,2	-0,5	-0,1	5,8	—
$c_{15}'$	$L_1-R$	37	43	49	93	-125	-56	12,3	14,2	-0,9	-0,1	5,8	—
Свернутые формы основной цепи													
$a_1'$	$R_1-R$	-55	-37	60	89	-89	-44	9,4	13,2	-0,5	-1,2	1,4	7
$a_2'$	$R_1-B$	-111	-39	63	92	-122	151	10,4	12,7	-0,5	-1,4	1,5	3
$a_3'$	$R_2-R$	-56	-52	176	89	-86	-42	9,3	13,4	-0,4	-0,6	2,0	15
$a_4'$	$R_2-B$	-111	-61	173	81	-122	154	10,0	12,8	-0,3	-1,0	2,0	2
$a_5'$	$B_2-L$	-71	130	177	90	51	43	9,0	14,4	-0,5	-1,3	2,2	1
$a_6'$	$R_3-B$	-138	-59	-60	91	-106	159	10,1	12,5	-0,3	-0,2	2,5	3
$a_7'$	$R_3-R$	-138	-59	-60	91	-100	-51	10,4	12,6	-0,4	-0,2	2,6	11
$a_8'$	$B_1-L$	-74	157	70	86	52	43	9,8	14,4	-0,4	-1,2	3,1	—
$a_9'$	$B_3-L$	-76	123	-63	93	52	43	9,0	14,0	-0,3	-0,1	4,0	—
$a_{10}'$	$L_3-L$	51	57	-59	90	52	48	9,4	15,0	-0,4	-0,1	4,3	—
$a_{11}'$	$L_2-J$	52	61	-174	68	52	47	9,4	16,1	-0,5	-0,2	5,4	—
$a_{12}'$	$L_1-L$	38	54	47	94	53	62	10,8	15,7	-0,9	-0,1	5,9	—

Геометрические и энергетические параметры оптимальных конформаций фрагмента  $C^{\alpha}CO-L-Ala-L-Phe-NH-C^{\beta}$ 

Конформация		Углы вращения, град					Энергия взаимодействия, ккал/моль					Конформации диглицидных фрагментов в белках
обозначение	тип	$\varphi_1$	$\varphi_2$	$\psi_2$	$\chi_2^1$	$\chi_2^2$	$E_{b_i-b_j}$	$E(b_i-s_i) + (b_{i+1}-s_i)$	$E_{b_i-s_j} (i \neq j, i+1)$	$E_{s_i-s_j}$	$L_{общ}$ ккал/моль	
Развернутые формы основной цепи												
$c_1''$	$B-B_3$	-101	143	-145	133	-58	89	40,4	44,8	-0,2	0	15
$c_2''$	$B-R_3$	-94	127	-140	-59	-57	92	10,0	12,1	-0,1	0,3	2
$c_3''$	$B-B_1$	-92	84	-143	153	54	89	9,6	12,3	-0,1	0,5	3
$c_4''$	$B-R_1$	-92	90	-161	-45	57	92	10,4	12,5	-0,1	1,3	1
$c_5''$	$R-L_3$	-100	-48	55	-44	-55	92	10,5	13,5	-0,2	1,9	—
$c_6''$	$L-B_3$	56	49	-107	157	-66	100	10,0	13,5	-0,1	2,0	—
$c_7''$	$L-R_3$	53	54	-112	-56	64	103	10,3	13,4	-0,1	2,2	—
$c_8''$	$B-B_2$	-82	93	-142	145	180	88	8,9	12,8	-0,1	2,3	7
$c_9''$	$B-R_2$	-87	90	-108	-56	179	87	8,9	13,0	-0,1	2,5	4
$c_{10}''$	$L-B_1$	49	54	-117	161	64	87	10,4	13,0	-0,1	2,8	—
$c_{11}''$	$L-B_2$	50	47	-116	146	180	89	10,2	12,5	-0,1	3,3	—
$c_{12}''$	$L-R_1$	51	50	-110	-37	62	90	11,0	12,9	-0,1	3,4	—
$c_{13}''$	$R-L_1$	-105	-52	38	44	47	92	12,1	13,8	-0,2	3,5	—
$c_{14}''$	$L-R_2$	51	45	-124	-59	179	87	10,7	12,9	-0,1	3,8	—
$c_{15}''$	$R-L_2$	-101	-52	54	47	-172	66	10,6	14,4	-0,1	5,4	—
Свернутые формы основной цепи												
$a_1''$	$R-B_3$	-98	51	-131	150	-57	89	10,0	41,9	-1,4	0,9	5
$a_2''$	$R-B_1$	-104	-56	-136	164	58	90	10,2	11,6	-1,0	4,2	5
$a_3''$	$R-R_1$	-65	-40	-54	-38	64	89	9,6	12,8	-0,2	4,8	8
$a_4''$	$R-R_3$	-58	-45	-120	-45	-58	87	9,6	12,8	-1,0	4,9	14
$a_5''$	$R-B_2$	-99	-54	-141	146	180	88	9,6	12,0	-0,2	2,2	3
$a_6''$	$B-L_3$	-71	124	53	42	-55	93	8,9	13,9	-0,9	2,3	—
$a_7''$	$R-R_2$	-66	-47	-95	-53	179	77	9,5	12,7	-0,1	2,8	13
$a_8''$	$L-L_3$	50	57	53	47	-56	90	9,4	14,6	-0,2	3,6	—
$a_9''$	$B-L_1$	-72	125	39	42	50	91	10,4	14,5	-1,3	3,9	—
$a_{10}''$	$B-L_2$	-69	122	52	45	179	68	9,8	14,5	-0,4	4,8	—
$a_{11}''$	$L-L_1$	53	61	37	46	47	94	11,0	14,8	-0,7	5,0	—
$a_{12}''$	$L-L_2$	49	57	53	48	-172	64	9,6	14,9	-0,1	5,6	—



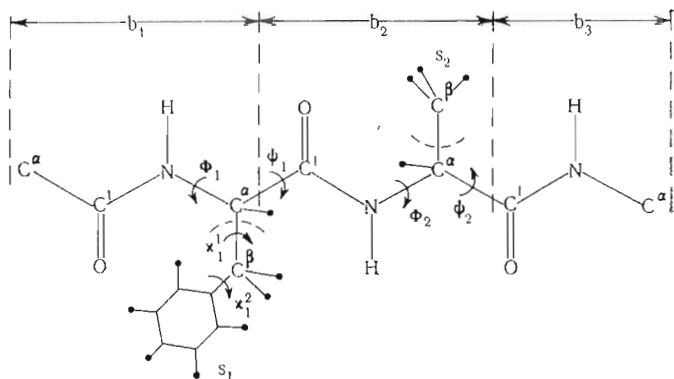


Рис. 3. Модель фрагмента  $C^\alpha\text{-CO-L-Phe-L-Ala-NH-C}^\alpha$

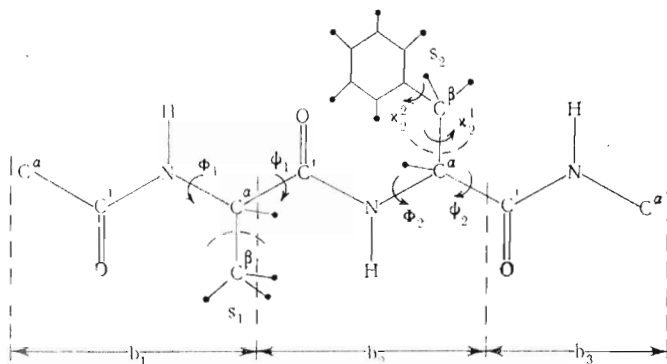


Рис. 4. Модель фрагмента  $C^\alpha\text{-CO-L-Ala-L-Phe-NH-C}^\alpha$

Расчет конформационных состояний фрагментов  $C^\alpha\text{-CO-L-Phe-L-Ala-NH-C}^\alpha$  и  $C^\alpha\text{-CO-L-Ala-L-Phe-NH-C}^\alpha$ , (II) и (III). Исходные конформации дипептидов формировались исходя из девяти ротамеров фенилаланина (табл. 4) и трех ротамеров аланина (табл. 1). Найденные после минимизации геометрические параметры оптимальных форм, величины потенциальной энергии и ее составляющие представлены в табл. 5, 6. Каждая конформация обозначена двумя буквами, которые отвечают состояниям основных цепей двух остатков; индекс у одной из букв отражает положение боковой цепи фенилаланина (1 —  $\chi^1 \sim 60^\circ$ , 2 —  $\chi^1 \sim 180^\circ$ , 3 —  $\chi^1 \sim -60^\circ$ ). Исходное значение угла  $\chi^2$  во всех случаях принималось равным  $90^\circ$ . Конформации в табл. 5, 6 разбиты на группы свернутых (a) и развернутых (c) форм основной цепи. Модели дипептидных фрагментов даны на рис. 3, 4.

Наиболее низкоэнергетическими формами основной цепи у обоих фрагментов являются *B-B*, *B-R*, *R-R* и *R-B*. Энергия этих форм у аланинового дипептида находится в интервале 0—0,6 ккал/моль (табл. 3). Близки между собой, независимо от положения боковой цепи, величины энергии *R* и *B* у *Ac-L-Phe-NHMe* (табл. 4). Поэтому при аддитивном формировании структуры рассматриваемых дипептидов различие в конформационной энергии не должно было бы превышать 1 ккал/моль. Однако на самом деле разброс составляет  $\sim 3$  ккал/моль. Таким образом, уже на уровне дипептида только за счет дополнительных взаимодействий, возникающих вследствие увеличения боковой цепи у одного остатка, наблюдается значительное детерминирование конформационных состояний.

В отличие от аланинового дипептида глобальными оказываются конформации не со свернутой, а с развернутой основной цепью *B-B*. По-преж-

нему здесь очень малы ( $\sim 0,2$  ккал/моль) взаимодействия между  $b_1$  и  $b_3$  (рис. 3, 4), однако имеются благоприятные условия для стабилизирующих взаимодействий между ароматическим кольцом и удаленной от фенилаланина пептидной группой. У фрагмента (II) они максимальны ( $-2,5$  ккал/моль) при  $\chi^1 \sim 60^\circ$  и  $180^\circ$ ; у фрагмента (III) параллельные расположения плоскостей  $s_2$  и  $b_1$  достигаются при  $\chi^1 \sim 60$  и  $-60^\circ$ . Энергетическая разность среди конформеров  $B-B$  и  $B-R$  обусловлена в основном вкладом  $E_{b_1-s_2}$  или  $E_{b_1-s_1}$ . У дипептида с N-концевым остатком фенилаланина все формы с  $s$   $B$ - и  $R$ -состояниями предпочтительнее форм, включающих  $L$ . У дипептида с C-концевым остатком фенилаланина формы  $B-B_2$  и  $B-R_2$  ( $c_8''$  и  $c_9''$ ) имеют несколько большую энергию, чем  $R-L_3$ ,  $L-B_3$  и  $L-R_3$  ( $c_5''-c_7''$ ). Это объясняется тем, что в  $c_5''-c_7''$  при  $\chi^1 \sim -60^\circ$  осуществляются дисперсионные взаимодействия между  $s_2$  и  $b_1$ , которые перекрывают проигрыш в энергии ближних взаимодействий. Формы  $c_8''$  и  $c_9''$  являются самыми выгодными по энергии взаимодействия между звеньями основной цепи ( $E_{b_i-b_j}$ ), но уступают большинству форм по величинам  $E_{(b_i-s_i)+(b_{i+1}-s_i)}$  и  $E_{b_i-s_j(j \neq i, i+1)}$ . Преимущество развернутых дипептидных конформаций перед свернутыми определяется главным образом энергетической составляющей  $E_{b_i-s_j(j \neq i, i+1)}$ . Например, в  $B_2-B$  величина  $E_{b_2-s_1} = -2,7$  ккал/моль, а в лучшей свернутой форме  $R_1-R$  она равна  $-0,3$  ккал/моль.

Если у метиламида N-ацетил-L-аланил-L-аланина (см. табл. 3) энергия более компактных по геометрии основной цепи форм  $R-R$  и  $R-B$  совпадает с энергией форм  $B-B$  и  $B-R$ , то в случае фрагментов (II) и (III) они менее выгодны на  $1,5-3,0$  ккал/моль. Энергетический диапазон среди конформеров  $R-R$  и  $B-B$  невелик. Свернутые конформации  $a$  характеризуются дисперсионными взаимодействиями между пептидными звеньями  $b_1$  и  $b_3$  ( $\sim -1,5$  ккал/моль). Однако выигрыш энергии компенсируется более высокой, чем у форм с энергией отталкивания между  $b_1$  и  $b_2$ ,  $b_2$  и  $b_3$ . Поэтому соответствующие конформации обеих групп по своим скелетным взаимодействиям практически изоэнергетичны. У наиболее выгодных конформаций  $a$  боковые цепи  $s_1$  и  $s_2$  взаимодействуют между собой. Однако из-за небольших размеров боковой цепи в аланине эффект таких контактов не так велик, как в случае взаимодействия бензольного кольца с пептидным звеном в группе  $c$ . При замене остатка аланина на другой остаток дисперсионные взаимодействия между  $s_1$  и  $s_2$  в свернутых конформациях могут оказаться более значительными [16].

Из сравнения энергий дипептидных фрагментов (табл. 5 и 6) можно заключить, что энергии однотипных форм у фрагмента (III) при  $\chi^1 -60^\circ$  и у фрагмента (II) при  $\chi^1 180^\circ$ , а также в обоих случаях при  $\chi^1 60^\circ$  примерно равны. Отсюда следует, что если у рассматриваемых фрагментов ротамеры  $\chi^1 180^\circ$  и  $\chi^1 -60^\circ$  существенно различны по стабильности, то на трипептидном фрагменте с центральным ароматическим остатком они будут близкими по энергии.

Рассмотрим вопрос о вероятности встретить рассчитанные конформации у соответствующих свободных дипептидов и у дипептидных фрагментов, входящих как в олигопептиды, так и в белки. В отношении дипептидов положение достаточно очевидно — критерием предпочтительности форм может служить в данном случае величина  $E_{\text{общ}}$ . Согласно полученным данным, молекулы Ac-L-Phe-L-Ala-NHMe и Ac-L-Ala-L-Phe-NHMe в полярной среде будут находиться главным образом в развернутых конформациях  $c_1-c_4$ . Это тем более вероятно, что эти конформации предпочтительнее свернутых форм  $a$  не только по энтальпии, но и по энтропии. Что же касается оценки конформационных состояний фрагментов -L-Phe-L-Ala- и -L-Ala-L-Phe- в более сложных олигопептидных молекулах, то величина  $E_{\text{общ}}$  имеет здесь относительную ценность. Малые величины энергии  $c_1-c_4$ , как уже отмечалось, обусловлены в основном

большей, чем у  $a_1 - a_4$ , стабилизацией за счет  $E_{b_i-s_j (j \neq i, i+1)}$ . В то же время конформация  $a$  предпочтительнее  $c$  по взаимодействиям между боковыми цепями ( $E_{s_1-s_2}$ ). В отношении суммарной энергии скелетных взаимодействий ( $E_{b_i-b_j}$ ) свернутые и развернутые конформации приблизительно равноценны. Поэтому ни одна из низкоэнергетических конформаций как в группе  $c$ , так и в  $a$  не обладает такой предпочтительностью, которая осталась бы доминирующей при изменении внешних условий и увеличении длины цепи. Величина  $E_{\text{общ}}$  также не может быть определяющей при выборе набора вероятных конформаций встроеного в олигопептид дипептидного фрагмента. Например, конформация  $B_2 - B$  предпочтительнее  $B_3 - B$  на 2,6 ккал/моль (табл. 5). Различие в  $E_{\text{общ}}$  обусловлено лишь отсутствием у  $B_3 - B$  дисперсионного взаимодействия между  $b_3$  и  $s_1$ , которое может быть скомпенсировано средними и дальними взаимодействиями. К конформациям, уступающим не только по  $E_{\text{общ}}$ , но и по целому ряду видов взаимодействий, у фрагмента (II) можно отнести  $c_7' - c_{15}'$  и  $a_8' - a_{12}'$  (табл. 5), а у фрагмента (III)  $c_{10}'' - c_{15}''$  и  $a_8'' - a_{12}''$  (табл. 6).

Сопоставим теперь результаты расчета свободных дипептидов с конформационными состояниями соответствующих дипептидных фрагментов в белках \*. Для получения более общих заключений сравним теоретические данные не со структурами белковых фрагментов  $-L\text{-Phe-}L\text{-Ala-}$  и  $-L\text{-Ala-}L\text{-Phe}$ , которых очень мало, а со структурами фрагментов вида  $-R\text{-X-}$  и  $-X\text{-R-}$ , где  $R$  — не только фенилаланин, но и все его конформационные аналоги (тирозин, триптофан, гистидин);  $X$  — остаток аланинового типа (все, кроме глицина и пролина). (Количество фрагментов  $-R\text{-X-}$  составило 79, а  $-X\text{-R-}$  равнялось 80.) Распределение по конформационным состояниям дипептидных белковых фрагментов дано в табл. 5 и 6.

Конформации  $-R\text{-X-}$  и  $-X\text{-R-}$ , как и можно было ожидать, приблизительно равномерно распределены по группам  $a$  и  $c$ . В каждой группе они соответствуют низкоэнергетическим формам свободных дипептидов. Встречающиеся в белках структуры представляют собой все комбинации конформационных состояний остатков  $R$  и  $X$  с формами основной цепи  $R$  и  $B$ . Из 159 случаев лишь в двух реализуются конформации с формой  $L$ , это  $R_2-L$  и  $B_2-L$  (табл. 5), имеющих самую низкую энергию среди  $L$ -конформаций дипептидов. Таким образом, все дополнительные взаимодействия, которым подвержены фрагменты в белковой глобуле, реализуются при низкоэнергетических конформациях  $a$  и  $c$ . Вместе с тем распределение белковых фрагментов по конформациям в каждой группе не соответствует порядку расположения по величинам  $E_{\text{общ}}$  конформаций свободных дипептидов. Причины несоответствия достаточно очевидны из проведенного выше анализа (см. также [18]).

#### ЛИТЕРАТУРА

1. Ramachandran G. N., Sasisekharan V. (1968) *Advances Protein Chem.*, **23**, 283—457.
2. Scheraga H. A. (1968) *Advances Phys. Org. Chem.*, **6**, 103—184.
3. Lewis P. N., Momany F. A., Scheraga H. A. (1973) *Isr. J. Chem.*, **11**, 121—152.
4. Попов Е. М., Дашевский В. Г., Липкинд Г. М., Архипова С. Ф. (1968) *Молекулярн. биология*, **2**, 612—621.
5. Попов Е. М., Липкинд Г. М., Архипова С. Ф., Дашевский В. Г. (1968) *Молекулярн. биология*, **2**, 622—630.
6. Липкинд Г. М., Архипова С. Ф., Попов Е. М. (1970) *Ж. структурн. химии*, **11**, 121—126.
7. Lipkind G. M., Arkhipova S. F., Popov E. M., (1973) *Int. J. Peptide Protein Res.*, **5**, 381—397.
8. Липкинд Г. М., Архипова С. Ф., Попов Е. М. (1970) *Изв. АН СССР. Сер. хим.*, № 2, 315—322.
9. Попов Е. М., Липкинд Г. М. (1971) *Молекулярн. биология*, **5**, 624—636.
10. Липкинд Г. М., Попов Е. М. (1971) *Молекулярн. биология*, **5**, 667—679.

\* Использованы рентгеноструктурные данные по глобулину [27],  $\alpha$ -химотрипсину [30] и карбоксипептидазе [32].

11. Goodman M., Naider F., Toniolo C. (1971) *Biopolymers*, **10**, 1719—1730.
12. Yaron A., Katchalski E., Berger A., Fasman G. D., Sober H. A. (1971) *Biopolymers*, **10**, 1107—1120.
13. Ponnuswamy P. K., Warne P. K., Scheraga H. A. (1973) *Proc. Nat. Acad. Sci. U. S. A.*, **70**, 830—833.
14. Lipkind G. M., Popov E. M. (1973) *Int. J. Peptide Protein Res.*, **5**, 371—379.
15. Попов Е. М. (1975) *Молекулярн. биология*, **9**, 578—593.
16. Kreissler M. A., Akhmedov N. A., Arkhipova S. F., Lipkind G. M., Popov E. M. (1974) *J. Chimie Phys.*, 913—919.
17. Липкинд Г. М., Ахмедов Н. А., Попов Е. М. (1975) *Биоорган. химия*, **1**, 1280—1290.
18. Липкинд Г. М., Архипова С. Ф., Попов Е. М. (1970) *Молекулярн. биология*, **4**, 331—338.
19. Pauling L. (1960) *The Nature of the Chemical Bond*. 3rd. ed., Cornell Univ. Press., N. Y.
20. Дашевский В. Г. (1970) *Высокомолекулярные соединения*, сер. «Итоги науки», 208—212.
21. Poland D., Scheraga H. A. (1967) *Biochemistry*, **6**, 3791—3800.
22. Scott R. A., Scheraga H. A. (1966) *J. Chem. Phys.*, **45**, 2091—2101.
23. IUPAC — IUB Commission on Biochemical Nomenclature. Recommendations 1970 (1971) *Biochim. et biophys. acta*, **229**, 1—17.
24. Дашевский В. Г. (1974) Конформации органических молекул, «Химия», М.
25. Галактионов С. Г. (1973). Исследование пространственной молекулярной структуры пептидов и белков, докт. дис., Минск.
26. De Coen J. L. (1970) *J. Mol. Biol.*, **49**, 405—414.
27. Watson H. C. (1969) *Progr. Stereochemistry*, **4**, 229—280.
28. Blake C. C. F., Mair G. A., North A. C. T., Phillips D. C., Sarma V. R. (1967) *Proc. Roy. Soc. London*, **B167**, 365—391.
29. Blundell T. L., Dodson G. G., Hodgkin D. C., Mercola D. A. (1972) *Advances Protein Chem.*, **26**, 279—402.
30. Birktoft J. J., Blow D. M. (1972) *J. Mol. Biol.*, **68**, 187—240.
31. Wu T. T., Kabat E. A. (1973) *J. Mol. Biol.*, **75**, 13—31.
32. Quiocho F. A., Lipscomb W. N. (1971) *Advances Protein Chem.*, **25**, 1—78.

Поступила в редакцию  
21.I.1975

## APPROACH TO CALCULATIONS OF STABLE CONFORMATIONS OF OLIGOPEPTIDES. I. DI- AND TRIPEPTIDES; CONTAINING Ala- AND Phe-RESIDUES

POPOV E. M., AKHMEDOV N. A., LIPKIND G. M.

*M. M. Shemyakin Institute of Bioorganic Chemistry,  
Academy of Sciences of the USSR, Moscow*

Conformational properties of the following fragments  $C^\alpha$ -CO-*L*-Ala-*L*-Ala-*L*-Ala-NH- $C^\alpha$ ,  $C^\alpha$ -CO-*L*-Phe-*L*-Ala-NH- $C^\alpha$ ,  $C^\alpha$ -CO-*L*-Ala-*L*-Phe-NH- $C^\alpha$  are analyzed. The interactions between neighboring amino acid residues result in a considerable energy differentiation of the space structures. Folded *R-R-R* and *R-R-B* backbone forms are shown to be preferred for a molecule composed of Ala-residues. When Phe-residues are present, *B-B* and *B-R* conformations are most favorable because of considerable dispersion interactions between an aromatic radical and peptide units in stretched forms of the backbone. Comparisons are made with conformational states of corresponding di- and tripeptide fragments in proteins.