

УДК 576.8:577.21

© 1992 г. Г. И. Чинелс

## СВОРАЧИВАНИЕ ПЕПТИДНЫХ ЦЕПЕЙ ПРИ ОБРАЗОВАНИИ ПРОСТРАНСТВЕННЫХ СТРУКТУР МОЛЕКУЛ БЕЛКОВ ОПРЕДЕЛЯЕТСЯ КОДОМ КОРНЕЙ КОДОНОВ АМИНОКИСЛОТ

*Институт органического синтеза Латвийской академии наук, Рига*

На основе анализа обширной базы данных первичных структур белков показано, что структура ближайших по последовательности пар аминокислот, частоты встречаемости которых имеют максимальные положительные отклонения от среднестатистических, в большинстве случаев соответствует коду корней кодонов аминокислот. Наиболее ярко это проявляется для аминокислотных пар в положениях  $n$  и  $n+5$  полипептидных цепей. Основной вклад при сворачивании пептидных цепей вносят аминокислоты семейства A/U.

Специфическую биологическую активность молекул белков однозначно определяет их уникальная пространственная структура. В молекулах ДНК закодирована лишь линейная аминокислотная последовательность полипептидных цепей. Составлением полного словаря кодонов и соответствующих им аминокислот была решена первая часть проблемы генетического кода, но нерешенной осталась вторая — каким путем первичная структура полипептидной цепи определяет образование трехмерной структуры белка [1]. В настоящее время наука еще не в состоянии дать исчерпывающий ответ на этот вопрос (см. обзор [2] и цитированную в нем литературу). Принципиально новый подход в решении второй части проблемы генетического кода был дан в работах [3, 4]. Предложенный подход предусматривает существование специфического кода взаимодействия аминокислот и «антиаминокислот» в соответствии с алгоритмом а-п-п̄-а̄ (аминокислота — кодон — антикодон — антиаминокислота) [5].

Исследование скрытой симметрии генетического кода [5] и эволюция структур семейств аминокислот [6, 7] привело к выявлению другого, более вырожденного универсального кода, определяющего первичное структурообразование пептидных цепей в водном окружении [8]. Этот код был назван кодом корней (вторых оснований) кодонов аминокислот (далее код к-к). От линейного кода аминокислота — антиаминокислота (код а-а̄), определяющего взаимодействие sense- и антисенс-пептидов, код к-к отличается значительно более выраженной вырожденностью (код а-а̄ охватывает лишь около 30% взаимодействующих пар аминокислот, предусмотренных кодом к-к) и применим для исследования взаимодействия пептидных цепей, закодированных некомплемментарными цепями ДНК. Согласно коду к-к, взаимодействуют аминокислоты, имеющие комплементарные корни (вторые основания) кодонов A/U и G/C [8]. В настоящей работе с использованием данных анализа первичных структур белков [9, 10] показано, что специфическое взаимодействие аминокислотных остатков при сворачивании полипептидной цепи в процессе самосборки молекул белков определяется кодом к-к.

В некотором приближении процесс сворачивания полипептидной цепи при образовании белковой структуры можно сопоставить с непрерывной

цепью реакций лиганд-рецепторного взаимодействия. N-Концевой фрагмент растущей на рибосоме пептидной цепи (длиной 5–7 аминокислотных остатков) можно рассматривать как лиганд, который, изгибаясь назад, ищет в пептидной цепи себе комплементарный рецепторный участок и ассоциирует с ним. Образованную структуру в свою очередь можно рассматривать как новый лиганд, который взаимодействует со следующим комплементарным участком на растущей пептидной цепи и т. д. Предполагается, что при таком процессе самосборки белка происходит непрерывное образование и перегруппировка промежуточных «лиганд-рецепторных» комплексов [4].

Не любая, например, случайным образом подобранная аминокислотная последовательность приводит к образованию компактной пространственной структуры. Первичные структуры природных белков являются продуктами длительного процесса эволюции. Результаты исследований прежних лет свидетельствуют, что частоты появления разных парных комбинаций аминокислот — ближайших соседей в пептидных цепях — не случайны и что первичные структуры белков в процессе эволюции проходят обучение [9, 10].

В работе [9] приведены результаты анализа первичных структур 4612 разных белков, содержащих 1062 149 аминокислотных остатков. Для 20×20 пар природных аминокислот были рассчитаны частоты встречаемости для положений 1–1 (соседние аминокислоты) 1–2, 1–3... 1–30. Анализ полученных результатов показал [9, 10], что уникальность первичных структур белков определяет специфичность близких парных нековалентных взаимодействий аминокислот. Эти взаимодействия становятся случайными (равновероятными) лишь начиная с пары 1–10, т. е. с 10-го соседнего положения. Таким образом, в белках по всей длине полипептидных цепей существует определенная селекция аминокислотной последовательности нонпептидных фрагментов. Это хорошо согласуется с моделью самосборки молекул белков типа лиганд-рецептор. Средняя длина природных пептидных лигандов (семейств брадикинина, тахикининов, нейромедининов, нейротензинов, нейрогипофизарных гормонов и т. п.), а также фрагментов, формирующих активные центры более длинных пептидных объектов, близка к этой величине. По-видимому, 9 аминокислот — это оптимальная длина фрагмента цепи, необходимая для обеспечения достаточно надежной (селективной) передачи информации при осуществлении различных биологических функций, в том числе и самосборки молекул белков. Сходство аминокислот, выраженное в предпочтении к определенным парным сочетаниям с ближайшими по цепи соседями в первичных структурах белков, в работе [10] определяется термином «neighbourhood selectivity» (далее сокращенно NS). Поскольку величина NS отражает нековалентные взаимодействия аминокислот [9, 10], в этом феномене должно проявляться действие кода корней кодонов аминокислот [8].

Согласно коду к-к [8], парные контакты образуют боковые радикалы аминокислот, корни кодонов которых комплементарны согласно структуре двучепочечной ДНК или ДНК-РНК-взаимодействий, т. е. A/T(U) и G/C. Исходя из существующих типов корней кодонов, природные аминокислоты можно разделить на 4 группы: G, A, C и U [5]. Аминокислоты одной группы, характеризующиеся одинаковыми корнями, потенциально эквивалентны и могут заменять друг друга в реакциях первичного структурообразования белков. Таким образом, если код к-к определяет образование специфических пространственных структур белков, то аминокислоты внутри групп G, A, C и U должны иметь одинаковую или близкую величину NS. Для проверки этого положения удобно использовать представления об изоморфных заменах аминокислот [10] и оценки

Корреляционные коэффициенты NS (численные значения указаны после символа аминокислоты), отражающие сходство природных аминокислот\* (стандартное отклонение 4) [10]

1 D	E/34	N/32	T/11	H/10	K/10	Q/8	G/8	S/7	A/3	R/2	P/1	M/-16	C/-20	W/-20	Y/-22	F/-31	V/-32	I/-32	L/-40
2 E	Q/37	D/34	K/24	N/13	A/12	R/11	G/-5	P/-7	H/-8	S/-10	W/-10	T/-12	M/-17	C/-18	V/-20	Y/-28	I/-36	F/-40	L/-40
3 K	Q/26	E/24	R/24	N/17	D/10	A/2	P/-2	H/-4	T/-8	S/-10	W/-10	G/-11	V/-14	C/-14	I/-17	M/-19	Y/-21	F/-26	L/-33
4 N	D/32	K/17	E/13	Q/12	T/12	R/11	G/10	H/8	P/8	S/1	W-6	C/-7	A/-12	M/-12	Y/-13	F/-23	V/-30	I/-32	L/-35
5 Q	E/37	K/26	R/26	N/12	D/8	A/0	P/0	W/-3	H/-4	S/-7	C/-8	G/-10	T/-10	M/-12	Y/-12	I/-20	F/-22	Y/-23	L/-26
6 H	Y/12	D/10	N/8	C/3	I/3	R/0	P/-2	S/-3	K/-4	Q/-4	W/-6	T/-7	E/-8	M/-8	G/-8	L/-9	A/-10	V/-10	F/-10
7 Y	I/29	F/24	V/17	L/16	H/12	M/8	W/6	C/5	R/0	T/-2	P/-9	N/-13	S/-16	A/-18	G/-18	K/-21	D/-22	Q/-23	E/-28
8 F	L/36	I/32	Y/24	M/20	V/16	C/11	W/8	T/4	S/0	P/-2	A/-8	H/-10	R/-11	D/-11	G/-12	Q/-22	N/-23	K/-26	E/-40
9 L	I/46	F/36	V/26	M/24	Y/16	C/15	W/8	A/-6	H/-9	S/-11	T/-11	G/-16	P/-16	R/-24	Q/-26	K/-33	N/-35	D/-40	E/-40
10 I	V/46	L/46	F/32	Y/29	M/18	W/13	C/12	H/3	T/-5	A/-17	S/-17	K/-17	P/-18	Q/-20	G/-21	R/-25	N/-32	E/-36	D/-39
11 M	L/24	F/20	I/18	W/12	V/10	Y/8	C/2	R/0	A/-2	T/-8	H/-8	S/-10	P/-10	G/-10	Q/-12	N/-12	D/-16	E/-17	K/-19
12 V	I/46	L/26	Y/17	F/16	W/14	M/10	C/8	T/-4	H/-10	A/-12	Q/-12	G/-14	K/-14	S/-16	P/-18	R/-18	E/-20	N/-30	D/-32
13 T	S/24	N/12	D/11	F/4	P/2	Y/-2	A/-3	C/-4	V/-4	I/-5	R/-5	H/-7	M/-8	W/-8	K/-8	Q/-10	L/-11	E/-12	G/-13
14 A	E/12	S/10	C/4	D/3	K/2	Q/0	P/-2	R/-2	M/-2	T/-3	G/-3	L/-6	F/-8	N/-8	H/-10	V/-12	W/-12	I/-17	Y/-18
15 P	S/15	N/8	G/4	T/2	D/1	Q/0	A/-2	F/-2	H/-2	K/-2	C/-3	Q/-4	E/-7	V/-9	W/-9	M/-10	L/-16	P/-18	I/-18
16 S	T/24	P/15	A/10	G/9	D/7	N/1	F/-0	H/-3	C/-6	Q/-7	M/-10	E/-10	K/-10	R/10	L/-11	V/-16	Y/-16	I/-17	W/-18
17 R	Q/26	K/24	E/11	N/11	D/2	H/0	Y/0	M/0	A/-2	P/-4	T/-5	G/-6	W/-7	S/-10	F/-11	C/-12	V/-18	L/-24	I/-25
18 G	N/10	S/9	D/8	P/4	C/-1	A/-3	E/-5	R/-6	H/-8	M/-10	Q/-10	K/-10	F/-12	W/-13	T/-13	V/-14	L/-16	Y/-18	I/-21
19 W	V/14	I/12	M/12	C/8	L/8	F/8	Y/6	Q/-3	N/-6	H/-6	R/-7	T/-8	P/-9	K/-10	A/-12	G/-13	S/-18	D/-20	D/-20
20 C	L/15	I/12	F/11	V/8	W/8	Y/5	A/4	H/3	M/2	G/-1	P/-3	T/-4	S/-6	N/-7	Q/-8	R/-12	K/-14	E/-18	D/-20

\* Аминокислоты, имеющие одинаковые корни с аминокислотой сопоставления, выделены жирным шрифтом.

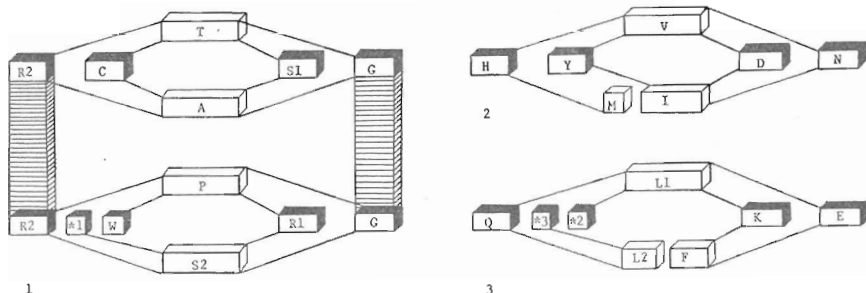


Рис. 1. Структуры семейств природных аминокислот. Полярные аминокислоты отмечены черным цветом. 1 – семейство аминокислот G/C, 2 и 3 – полусемейства A/U-1 и A/U-2. Для получения максимального сходства с семействами A/U-1 и A/U-2 семейство G/C условно разделено по аминокислотам R2 и G на две части. Размеры «кирпичиков» пропорциональны числу кодонов, определяющих данную аминокислоту. Связь, соединяющая две аминокислоты, соответствует паре «кодон-антикодон», согласно алгоритму генетического кода а-п-а-а [5]

их сходства, выраженные посредством корреляционных коэффициентов NS.

В табл. 1 на основе рассчитанных корреляционных коэффициентов NS [10] аминокислоты расположены в ряду в порядке понижения их сходства с объектом сопоставления – аминокислотой, расположенной в крайней левой колонке. Аминокислоты сопоставления сгруппированы согласно структурам корней кодонов: 1–7 – группа аденина или А, 8–12 – группа U, 13–16 – группа С и 16–20 – группа G; серин, имея полярно-аполярную природу, входит в состав двух групп – С и G [5].

Табл. 1, полученная на основе данных по более чем 1 миллиону пар аминокислот [9, 10], оказалась достаточно информативной. Как и ожидалось, в феномене NS действительно проявляется действие кода к-к (и более общего кода сигнатур) [11–13]. В семействе аминокислот аденина (ряды 1–7, табл. 1), судя по величинам NS, структурное сходство и близкую реакционную способность по выбору ближайших соседей имеют D, E, K, N, Q, в меньшей степени H. Исключение составляет Y. Большинство аминокислот этой группы имеют максимальные численные значения коэффициентов и расположены в левой части табл. 1.

Отклонение поведения тирозина от законов кодового взаимодействия аминокислот было предметом обсуждения в работе [13]. Данные, полученные на основе коэффициентов NS, подтверждают сделанные выводы. Тирозин в большинстве случаев ведет себя как неполярная аминокислота группы U. Из аминокислот группы А наиболее близкие с тирозином свойства проявляет лишь гистидин (см. табл. 1). В ряду тирозина большинство аминокислот группы А расположено в правой части таблицы. Зато в группе U (ряды 8–12, табл. 1) тирозин расположен среди ее членов в левой части таблицы. Для других представителей группы А необходимо отметить сходство лизина, глутаминовой кислоты и аргинина. E и R имеют одинаковые значения коэффициентов NS по отношению к лизину. Сходство K и E определяется кодом к-к, несмотря на противоположные заряды, а сходство K и R – общими положительными зарядами (как носителями сигнатур), несмотря на разные корни кодонов.

Выраженное стремление выбрать одинаковых ближайших по цепи партнеров проявляют также представители группы U (ряды 8–12, табл. 1). В этом плане это самая монолитная, судя по значениям корреляционных коэффициентов NS, группа аминокислот, все члены которой плотно расположены в левом фланге таблицы. Аминокислоты группы А и U, согласно структуре генетического кода, образуют два полусемейства:

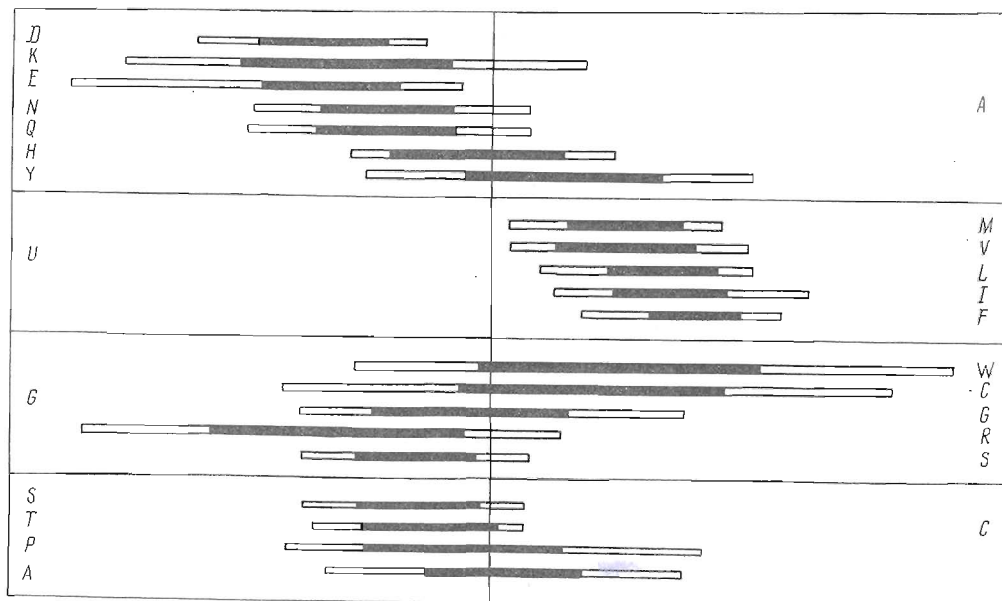


Рис. 2. Характеристики гидрофильных и гидрофобных свойств аминокислот группы А, U, G и С. Для сопоставления данных нескольких десятков известных шкал индексы гидропатичности аминокислот (ИГ) нормированы (среднее значение для каждой шкалы — 0, дисперсия  $\pm 1$ ), черными полосами отмечены районы, в которых предполагаются 75% значений ИГ, определенных разными методами [14]

A/U-1 и A/U-2 (рис. 1) [5–7], члены которых характеризуются выраженными гидрофильными и гидрофобными свойствами [14].

Совершенно другие свойства проявляют аминокислоты семейства G/C. Аминокислоты группы С (ряды 13–16, табл. 1), которые по данным анализа точечных мутаций в структурах гомологических белков имеют высокие частоты взаимного обмена (например, A/T 590, A/P 345, S/T 696, S/A 772, S/P 269 и T/P 345, см. рис. 80 в работе [15]), согласно данным по корреляционным коэффициентам NS (за исключением 16-го ряда серина), не проявляют ярко выраженного сходства в отношении предпочтения к каким-либо определенным типам аминокислот-соседей. То же самое можно сказать и о группе аминокислот G. Аминокислоты этой группы в ходе эволюции редко заменяют друг друга в структурах гомологичных белков. Причиной этого является яркая индивидуальность их поведения [5]. Аминокислоты семейства G/C в отличие от A/U обладают слабо выраженными гидропатичными свойствами (рис. 2). Все это позволяет предположить, что в процессе самосборки белков решающее значение в формировании структуры имеют взаимодействия между аминокислотами групп А и U.

Сворачивание полипептидной цепи в белковую структуру происходит уже в процессе ее синтеза на рибосоме, т. е. контрансляционно. N-Концевая часть растущего пептида после синтеза 30–40 остатков оказывается свешенной с рибосомы в окружающую водную среду [16]. При этом для изучения процесса сворачивания пептидной цепи может быть применена модель «ледокола» [17]. Согласно этой модели, в основе действия кода к-к при взаимодействиях пептидных участков в водной среде лежит образование связей между противоположными по полярности боковыми радикалами аминокислот. Как известно, свободная энергия системы «аполярная молекула — кластер молекул воды» ниже свободной энергии системы «аполярная молекула — несвязанная вода» [17]. Поэтому вокруг аполярных углеводородных частей молекул происходит образование льдоподоб-

Частота встречаемости (в "единицах отклонения" [9]) лизина по отношению к ближайшим по положению аминокислотам\* в пептидных цепях белков

Поло- жение**	"Единицы отклонения" аминокислот																			
	V/3,2	Y/2,5	H/1,9	N/1,1	R/1,0	A/0,8	K/0,7	E/0,6	T/0,5	I/0,3	G/0,0	D/-0,2	L/-0,5	C/-0,5	P/-1,2	Q/-1,7	F/-2,2	W/-2,5	M/-2,5	S/-3,4
1	V/3,2	Y/2,5	H/1,9	N/1,1	R/1,0	A/0,8	K/0,7	E/0,6	T/0,5	I/0,3	G/0,0	D/-0,2	L/-0,5	C/-0,5	P/-1,2	Q/-1,7	F/-2,2	W/-2,5	M/-2,5	S/-3,4
2	I/3,4	H/2,6	Y/2,2	L/2,1	V/1,9	W/1,9	C/0,8	K/0,5	G/0,3	F/-0,1	T/-0,1	M/-0,9	R/-1,0	N/-1,5	Q/-1,7	S/-1,9	R/-2,1	P/-2,4	E/-2,5	D/-3,6
3	I/2,4	E/2,4	V/2,3	Q/1,9	G/1,1	P/0,7	C/0,4	D/0,3	W/0,2	K/-0,5	F/-0,6	S/-0,6	H/-0,7	A/-1,1	L/-1,1	N/-1,1	T/-1,1	R/-1,9	M/-3,3	Y/-4,4
4	D/4,8	E/4,4	Q/1,1	W/1,1	P/1,0	T/0,9	H/0,8	N/0,4	K/0,0	S/-0,6	Y/-0,6	I/-0,7	G/-0,8	C/-0,9	V/-1,1	A/-1,3	L/-1,4	F/-1,5	R/-2,7	M/-4,8
5	I/1,4	K/1,1	P/0,9	F/0,8	H/0,8	C/0,7	V/0,7	L/0,2	W/0,2	G/0,0	Q/0,0	E/-0,1	T/-0,3	F/-0,5	S/-0,6	A/-0,8	M/-1,4	Y/-1,9	N/-2,2	D/-2,3
6	Y/3,9	C/2,1	W/2,1	L/1,3	I/0,6	F/0,5	K/0,5	M/0,3	Y/0,3	E/-0,2	P/-0,3	H/-0,3	Q/-0,7	T/-0,8	D/-0,8	R/-1,4	A/-1,8	G/-1,9	N/-2,2	S/-2,3
7	G/3,2	W/2,9	D/2,3	E/2,0	Y/0,5	K/0,4	N/0,1	V/-0,1	H/-0,3	L/-0,3	Q/-0,3	C/-0,8	I/-0,8	S/-1,0	M/-1,1	A/-1,2	P/-1,4	R/-1,4	T/-1,5	F/-1,9
8	V/1,9	E/1,5	K/1,3	S/0,8	T/0,7	Q/0,6	D/0,5	M/0,3	C/0,1	L/0,0	F/-0,3	H/-0,5	G/-0,6	I/-0,7	A/-0,7	N/-0,1	R/-0,9	W/-1,2	Y/-2,1	P/-3,0
9	V/3,4	I/1,7	P/1,4	K/1,2	L/0,2	C/0,1	M/-0,1	F/-0,2	H/-0,2	R/-0,3	N/-0,4	W/-0,4	T/-0,5	Q/-0,6	C/-0,8	Y/-0,8	S/-0,9	D/-1,3	G/-1,4	A/-1,7

\* Комплементарные лизину аминокислоты группы U выделены жирным шрифтом.

\*\* Положения аминокислот - соседней соответствующей нумерации миниматрицы в табл. 3.

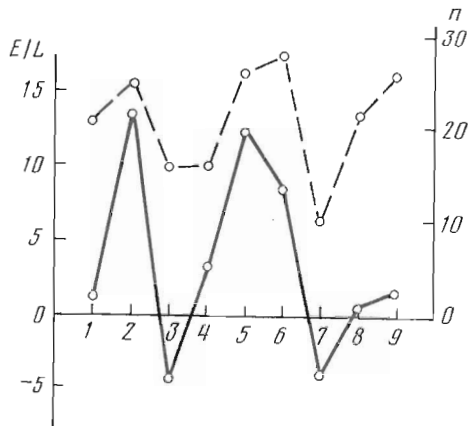


Рис. 3. Зависимости величины «единиц отклонения» для пары аминокислот  $E/L$  (сплошная линия) и числа положительных значений «единиц отклонения» ( $n$ ) для аминокислот семейства  $A/U$  ( $n$ ) (штриховая линия) от взаимного расположения аминокислот в пептидной цепи. Порядковые номера на оси  $x$  соответствуют номерам миниматриц в табл. 3

ных структур воды. Вода вокруг неполярных боковых цепей аминокислот как будто замерзает, образуя «айсберги». В случае гидрофобных взаимодействий организованные структуры воды должны быть удалены из пространства между радикалами аминокислот. При реализации этого процесса взаимодействие противоположных по полярности радикалов имеет определенные преимущества по сравнению с взаимодействием одноименных радикалов, например аполярных, так как полярные (ионы или диполи) группы разрушают организованную структуру воды, а аполярные, наоборот, усиливают образование структурированной воды. При этом предполагается, что при сближении двух аполярных радикалов на расстоянии, достаточные для перекрывания их упорядоченных гидратных оболочек, образуется единая кооперативная система связанной воды. В результате наблюдается своеобразный эффект усиления, который приводит к возрастанию стабильности упорядоченных структур воды между радикалами, препятствуя их сближению [17].

Дополнительно представляло интерес выяснить, на каком расстоянии в пептидных цепях наиболее часто расположены взаимно комплементарные (согласно коду к-к) аминокислоты. В качестве тестовой модели была выбрана аминокислота лизин (группа А) и рассмотрены частоты ее встречаемости по отношению к ближайшим соседям в положениях цепи от 1 до 9 (табл. 2). Для количественной оценки были использованы «единицы отклонения» [9] от средних значений частот при случайном распределении аминокислот по цепи. Табл. 2 разделена ломаной линией на две части, включающие в себя соответственно положительные и отрицательные значения величин «единиц отклонения». В положительной части таблицы все пять аминокислот группы  $U$  (F, I, L, M, V) встречаются лишь в шестом ряду (относительное расположение аминокислот 1—6 или  $n$  и  $n+5$ ), т. е. при изгибе пептидной цепи такое взаимное расположение аминокислот создает оптимальные стерические условия для взаимодействия комплементарных боковых радикалов. Характерно, что в случае расположения 1 и 7 (ряд 7, табл. 2) все без исключения численные значения «единиц отклонения» для аминокислот группы  $U$  отрицательные.

Для дополнительной проверки отмеченного факта были составлены соответствующие миниматрицы для всех аминокислот семейства  $A/U$



Миниматрицы\* "единиц отклонения" [9], характеризующие встречаемость аминокислот членов семейства A/U в положениях от 1—2 до 1—10

1. Матрица AA

A	F	I	L	M	V
D	0,6	5,7	3,2	-0,5	0,6
E	-1,3	1,5	1,2	2,6	-1,0
H	1,6	-2,1	-0,1	-1,6	-0,3
K	-2,2	0,3	-0,5	-2,8	3,2
N	0,8	3,1	1,9	2,0	0,6
Q	-1,0	-0,2	0,2	1,1	0,7
Y	2,8	-0,4	0,3	0,1	-1,3

2. Матрица A x A

A	F	I	L	M	V
D	1,0	5,2	2,2	1,2	5,4
E	2,7	4,5	13,0	5,8	2,5
H	-0,1	1,8	-0,6	-0,9	0,8
K	-0,1	3,4	2,1	-0,9	1,9
N	-1,0	1,9	1,2	2,4	0,1
Q	1,5	3,2	3,3	3,3	2,4
Y	-3,5	-1,5	-1,1	1,5	-2,6

3. Матрица A 2x A

A	F	I	L	M	V
D	0,0	2,0	-1,1	-0,1	0,0
E	0,6	-0,3	-4,3	-3,3	-3,4
H	-0,7	1,8	-1,0	-0,7	-1,0
K	-0,6	2,4	-1,1	-3,3	2,3
N	0,8	3,2	3,6	-2,7	2,3
Q	0,8	-1,6	2,3	-1,7	-1,1
Y	0,8	-3,3	0,3	1,7	-3,2

4. Матрица A 3x A

A	F	I	L	M	V
D	1,9	1,1	-0,8	-2,3	1,2
E	-1,3	-3,2	3,6	-1,7	-0,7
H	-0,6	1,0	0,6	0,6	3,8
K	-1,5	-0,7	-1,4	-4,8	-1,1
N	0,0	2,8	2,4	-1,5	-0,6
Q	-0,8	-0,4	0,1	-3,0	-0,4
Y	0,2	-0,4	0,0	2,2	0,7

5. Матрица A 4x A

A	F	I	L	M	V
D	0,1	3,4	3,0	-1,2	1,6
E	0,7	1,8	11,6	4,6	4,0
H	-1,5	1,2	-0,8	-0,9	0,0
K	0,8	1,4	0,2	-1,4	0,7
N	2,5	4,1	0,6	-0,4	2,8
Q	-0,2	0,9	1,5	0,7	0,7
Y	1,1	-2,3	-1,5	-1,2	-1,7

6. Матрица A 5x A

A	F	I	L	M	V
D	0,7	0,4	2,5	1,8	3,0
E	2,0	0,8	8,0	-1,0	1,9
H	0,0	0,2	-1,1	-0,7	0,7
K	0,5	0,6	1,3	0,3	3,9
N	0,0	3,1	0,8	-0,5	0,2
Q	0,5	0,7	3,6	0,2	0,2
Y	-2,0	-1,5	-0,3	-0,8	-0,2

7. Матрица A 6x A

A	F	I	L	M	V
D	0,8	0,4	1,0	-1,1	-1,1
E	-1,7	1,2	-3,8	-4,2	-1,7
H	0,1	-1,4	-0,9	0,4	-0,4
K	-1,9	-0,5	-0,3	-1,1	-0,1
N	-2,6	0,0	-1,2	0,0	-0,7
Q	0,0	-2,8	-1,5	-1,6	-0,3
Y	0,9	-0,1	-0,6	-0,7	-0,4

8. Матрица A 7x A

A	F	I	L	M	V
D	0,4	0,5	1,1	2,0	2,7
E	-0,1	1,1	0,4	-2,0	2,4
H	-0,0	-1,2	0,3	0,5	0,3
K	-0,3	-0,7	0,0	0,3	1,9
N	-1,2	1,7	4,0	1,4	-0,5
Q	-0,6	-2,1	0,9	-0,5	1,0
Y	-0,9	0,5	-0,7	1,7	-2,1

9. Матрица A 8x A

A	F	I	L	M	V
D	-0,2	-0,9	2,3	1,9	-0,3
E	0,8	-1,5	1,3	2,5	1,6
H	0,5	-0,9	0,6	0,2	1,2
K	-0,2	1,7	0,2	-0,1	3,4
N	0,4	1,1	2,9	1,2	1,7
Q	0,2	-0,4	2,9	-2,0	1,2
Y	1,9	0,7	-1,5	0,0	1,2

\* Символами "x" обозначены аминокислоты, разделяющие члены семейства A/U в пептидных цепях белков.

(табл. 3). Кроме основного показателя — чисел «единиц отклонения» с положительным знаком, подсчитанных для всего семейства по каждой матрице (рис. 3), дополнительно, в качестве специфического показателя, были выбраны величины «единиц отклонения» для аминокислотной пары E/L. В сополимерах эти аминокислоты участвуют в специфических взаимодействиях [18]; о их взаимном сродстве свидетельствуют высокие значения соответствующих коэффициентов (матрицы № 2 и 5, табл. 3), т. е. можно полагать, что при благоприятных условиях остатки E и L всегда будут стремиться к взаимодействию с образованием парных контактов. Как видно из графика (рис. 3), оба выбранных показателя имеют общие максимумы при расстояниях 5—6 аминокислотных остатков. Вторым максимум имеет место для пары E/L, разделенной одним аминокислотным



остатком (матрица № 2, табл. 3). При этом, по-видимому, обеспечиваются условия для образования характерного комплекса [19].

Таким образом, данные, полученные при анализе первичных структур большого числа белков, убедительно свидетельствуют, что важным фактором, определяющим специфичность сворачивания пептидных цепей при образовании белковых глобул, является код к-к. Основной вклад в этом процессе вносят аминокислоты семейства А/У, аминокислоты же семейства G/C имеют менее специфические взаимодействия.

#### СПИСОК ЛИТЕРАТУРЫ

1. Kolata G. // Science. 1986. V. 233. № 4748. P. 1037-1039.
2. Creighton T. E. // Biochem. J. 1990. V. 270. № 1. P. 1-16.
3. Меклер Л. Б. // Биофизика. 1969. Т. 14. Вып. 4. С. 581-584.
4. Меклер Л. Б., Идлис Р. Г. // Биофизика. 1981. Т. 26. Вып. 3. С. 574. и депонированный в ВИНИТИ за № 1476-81 от 3 апреля 1981 г. полный текст.
5. Чипенс Г. И. // Биоорганич. химия. 1991. Т. 17. № 9. С. 1284-1288.
6. Чипенс Г. И. // Журн. эволюц. биохим. и физиол. 1991. Т. 27. № 4. С. 513-521.
7. Чипенс Г. И. // Журн. эволюц. биохим. и физиол. 1991. Т. 27. № 4. С. 522-529.
8. Чипенс Г. И., Иевиня Н. Г., Рудзис Р. В. // Биоорганич. химия. 1991. Т. 17. № 11. С. 1582-1584.
9. Cserzo M., Simon I. // Int. J. Peptide and Protein Res. 1989. V. 34. № 3. P. 184-195.
10. Tüdös E., Cserzo M., Simon I. // Int. J. Peptide and Protein Res. 1990. V. 36. № 3. P. 236-239.
11. Чипенс Г. И. // Структурные основы действия пептидных и белковых иммунорегуляторов / Ред. Чипенс Г. И. Рига: Зинатне, 1990. С. 10-13.
12. Chipens G. I. // Adv. Drug Deliv. Rev. 1988. V. 2. Issue 2. P. 167-206.
13. Чипенс Г. И. // Биоорганич. химия. 1991. Т. 17. № 10. С. 1335-1346.
14. Чипенс Г. И., Балодис Ю. Ю., Гилломедова Л. Е. // Укр. биохим. журн. 1991. Т. 63. № 4. С. 20-29.
15. Dayhoff M. O., Schwartz R. M., Orcutt B. C. // Atlas of Protein Structure. 1978. V. 5. Suppl. 3. P. 345-352. Nat. Biomed. Res. Found Georgetown University, Washington.
16. Спириг А. С. Структура рибосомы и биосинтез белка. М.: Высшая школа, 1986. С. 272-274.
17. Чипенс Г. И., Рудзис Р. В. // Биоорганич. химия. 1991. Т. 17. № 11. С. 1445-1448.
18. Poland D., Scheraga H. A. // Poly  $\alpha$ -Amino Acids / Ed. Fasman G. N. Y.: Marcel Dekker, 1967. P. 481-483.
19. Root-Bernskin R. S. // J. Theoret. Biol. 1982. V. 94. № 24. P. 885-994.

Поступила в редакцию  
30.VI.1991

G. CHIPENS

#### FOLDING OF PEPTIDE CHAINS IS STIPULATED BY THE CODE OF AMINO ACID CODON ROOTS DURING THE FORMATION OF SPATIAL STRUCTURES OF PROTEIN MOLECULES

*Institute of Organic Synthesis, Latvian Academy of Sciences, Riga*

Basing on the analysis of a large number of protein sequences (Cserzo M., Simon I., 1989), the structure of the amino acid nearest neighbour pair whose occurrence has a maximal positive deviation from the mean statistical value, is shown to correspond in most cases to the code of the amino acid codon roots. It reveals particularly amino acid pairs in  $n$  and  $n+5$  positions of polypeptide chains. Amino acids belonging to A/U family contribute mostly to the folding of peptide chains.