



УДК 547.96+548.737

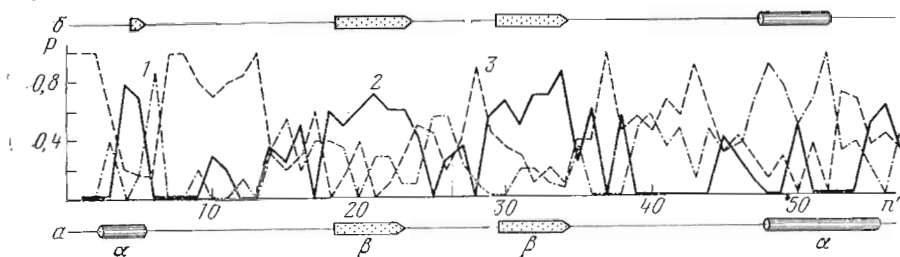
ПРЕДСКАЗАНИЕ СТРУКТУРИРОВАННЫХ УЧАСТКОВ
ГЛОБУЛЯРНЫХ БЕЛКОВ ПО ИХ ПЕРВИЧНОЙ СТРУКТУРЕ*Барковский Е. В., Бандарин В. А.**Минский государственный медицинский институт*

Предложен метод предсказания структурированных участков в глобулярных белках по их первичной структуре. Метод правильно предсказывает спиральное или не-спиральное состояние 87% и β -структурное или не- β -структурное состояние 94% всех входящих в белок аминокислотных остатков. Предсказана локализация элементов вторичной структуры ряда белков с неизвестной пространственной структурой.

В настоящее время широкое распространение и признание получили теории формирования вторичной структуры в развернутых белковых цепях и самоорганизации третичной структуры глобулярных белков [1—5]. Согласно представлениям, развиваемым в этих теориях, уже на первых стадиях самоорганизации белковой цепи образуются флуктуирующие α -спирали и β -структуры, определенные участки которых на последующих стадиях самоорганизации закрепляют свою вторичную структуру дальними взаимодействиями, а вторичная структура остальных участков при этом разрушается.

Вместе с тем использование данных о нативной вторичной структуре белка при моделировании укладки его полипептидной цепи также может оказаться эффективным [6]. Это оправдывает попытки теоретического предсказания α -спиральных и β -структурных участков белков, входящих в уже сформировавшуюся глобулу.

В большинстве из ныне существующих методов предсказания вторичной структуры белков учитываются лишь ближние взаимодействия. Эти методы чаще всего используют параметры, полученные на основании статистического анализа особенностей аминокислотного состава участков вторичной структуры белков [7—14], хотя в некоторых случаях для предсказания используются термодинамические параметры, полученные в опытах на синтетических полипептидах [15]. Важнейшим выводом этих работ является утверждение, что конформационное состояние данного аминокислотного остатка в белке в значительной степени определяется ближними взаимодействиями. Однако использование предсказательных методов, основанных на учете лишь ближних взаимодействий, часто не дает корректной картины [16—17]. Важную роль в выборе конформационного состояния данного аминокислотного остатка из числа оптимальных форм, дозволенных по ближним взаимодействиям, играют средние и дальние взаимодействия [18]. Анализ средних взаимодействий показывает, что наиболее сильное влияние на конформационное состояние остатка (n) оказывают его ближайшие соседи — $(n - 1)$ и $(n + 1)$ [18].



Вероятность образования (P) α -спиральной (1), β -структурной (2) и нерегулярной (3) конформации в панкреатическом ингибиторе трипсина (n' — номер дипептида). a — локализация α -спиральных и β -структурных участков в нативном белке [28], b — предсказанная вторичная структура

На основании этого можно предположить, что разработка метода, учитывающего парные корреляции между соседними аминокислотными остатками, позволит улучшить точность предсказания элементов вторичной структуры глобулярных белков по сравнению с методами, основанными на статистическом анализе отдельных аминокислотных остатков.

Распределение дипептидных фрагментов в элементах вторичной структуры глобулярных белков. Вероятности встречаемости всевозможных дипептидов в α -спиральной, β -структурной и неупорядоченной конформациях (табл. 1), необходимые для предсказания вторичной структуры, получены путем статистического анализа данных по пространственной структуре 20 белков: α -цепи гемоглобина [19], термолизина [20], папаина [21], цитохрома b_5 [22], кальцийсвязывающего белка [23], рубредоксина [24], цитохрома c [25], химотрипсина [26], трипсиногена [27], ингибитора трипсина [28], миоглобина [29], эрабутоксина b [30], нуклеазы [31], рибонуклеазы S [32], эластазы [33], субтилизина [34], А-цепи инсулина, В-цепи инсулина [35], лизоцима [36], глобина миноги [37].

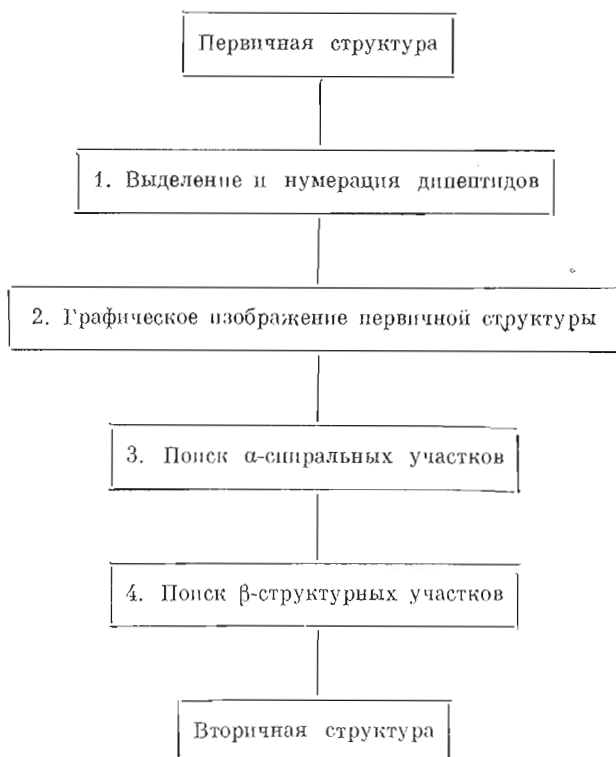
Всего проанализировано 2550 дипептидов, из них: 825 в α -спиральных, 360 в β -структурных и 1365 в неупорядоченных участках. Из 400 теоретически возможных дипептидов в элементах вторичной структуры расщатриваемых белков не встретилось 25 дипептидов, что, очевидно, является следствием недостаточной выборки. Следует отметить, что в состав этих дипептидов в основном входят аминокислотные остатки метионина, триптофана, гистидина, частота встречаемости которых в белках невелика сама по себе [38]. Клетки, соответствующие этим дипептидам, не заполнены.

Алгоритм предсказания вторичной структуры глобулярных белков. Сопоставление графического изображения первичной структуры белков, полученного на основании вероятностей встречаемости дипептидов в α -спиральной, β -структурной и неупорядоченной конформациях, и элементов их нативной вторичной структуры позволило разработать алгоритм поиска вторичной структуры глобулярного белка (см. схему). Рассмотрим более подробно схему алгоритма по стадиям.

Стадия 1. Первичная структура белка условно разбивается на ряд перекрывающихся дипептидов, которые последовательно нумеруются от N-конца цепи. Например, в трипептиде Val-Leu-Lys следует различать два дипептида: Val-Leu и Leu-Lys.

Стадия 2. Строится график зависимости между первичной структурой белка и вероятностью нахождения составляющих ее перекрывающихся дипептидов в α -спиральной, β -структурной и неупорядоченной конформациях. Для дипептидов, не встречающихся в выборке, вероятность пребывания в любой конформации условно принята равной 0,33. На рисунке в качестве примера представлено графическое изображение такой зависимости для панкреатического ингибитора трипсина.

Блок-схема алгоритма предсказания вторичной структуры глобулярного белка



Стадия 3. α -Спиральным считается участок полипептидной цепи, на котором суммарная вероятность встречаемости дипептидов в α -спиральной конформации больше таковой в других конформациях. α -Спираль прерывается дипептидами, вероятность пребывания которых в α -спиральной конформации равна нулю, а также двумя и более перекрывающимися дипептидами, вероятности пребывания которых в нерегулярной конформации больше, чем в α -спиральной конформации. Один дипептид не прерывает α -спираль, как бы ни была велика его вероятность встречаемости в нерегулярной конформации.

Стадия 4. Локализация β -структурных участков проводится только на тех отрезках полипептидной цепи, на которых не были предсказаны α -спирали. β -Структурным считается участок полипептидной цепи, на котором суммарная вероятность встречаемости дипептидов в β -структурной конформации превышает более чем в 2 раза суммарную вероятность встречаемости в нерегулярной конформации. β -Структура прерывается дипептидами, вероятность пребывания которых в β -структурной конформации равна нулю, а также любым другим дипептидом, вероятность встречаемости которого в нерегулярной конформации больше, чем в β -структурной.

Предсказание вторичной структуры глобулярных белков. Для получения количественных параметров, характеризующих точность предсказания вторичной структуры при помощи вышеописанного алгоритма, нами были локализованы α -спиральные и β -структурные участки в 10 глобулярных белках (все белки взяты из множества, используемого для получения алгоритма). Полученные результаты представлены в табл. 2.

Как видно из представленных данных, предлагаемый метод правильно предсказывает спиральное или неспиральное состояние 87% всех аминокислотных остатков, входящих в рассмотренные белки, и β -структурное

Таблица 1

Вероятности встречаемости дипептидов в α -спиральной (левый верхний угол), β -структурной (правый верхний угол) и неупорядоченной (нижняя строка) конформациях

	Ala	Phe	Leu	Pro	Met	Val	Ile	His	Lys	Arg
Ala	0,7 0,2	0,6 0,1	0,7 0,2 0,1	0,38 0,33	0,5 0	0,5 0,2 0,3	0,7 0,2 0,1	0,6 0,40	0,55 0,45	0 0,2 0,5
Phe	0,4 0,6	0,5 0,4 0,1	1,0 0	0,45 0,55	0,1 0,2	0,7 0,5	0,25 0,25	0 1,0	0,3 0,3	0,4 0,1 0,8
Leu	0,7 0,3	0,85 0,15	0,1 0,1	0,55 0,45	0,35 0	0,4 0,5 0,1	0,25 0,65 0,1	0,5 0,2	0,4 0,45 0,15	0,3 0,6 0,1
Pro	0,7 0,3	0	0,15 0,35	0 1,0	0 0	0,4 0,6	0,25 0,15	0,6 0,4	0,4 0,6	0 0 1,0
Met	0 1,0	0,35 0	0 1,0	0 0	1,0 0	0,2 0,8	0,25 0,55	0 0	1,0 0	0 1,0 0
Val	0,45 0,15	0 0,1	0,5 0,45	0 1,0	0 0,75 0,25	0,05 0,9 0,05	0,3 0,6 0,1	0,15 0,75	0,4 0,35	0,35 0,20
Ile	0,75 0,05	0,45 0,20	0,5 0,1	0 1,0	0 0	0,1 0,8 0,1	0,4 0,6 0	0,65 0,35	0,3 0,3	0,4 0,35 0,5
His	0,9 0,1	0 1,0	0 1,0	0,25 0,75	1,0 0	1,0 0	0 0,2	0 1,0	0,2 0,4	0,4 0,4
Lys	0,55 0,20	0,9 0,1	0,15 0,15	0,25 0,75	0,6 0,4	0,4 0,2	0,35 0,45 0,20	0,35 0,3	0,6 0,25	0,15 0,65
Arg	0,45 0,55	0 1,0	0,1 0,2	0 1,0	1,0 0	0 0,65 0,35	0,4 0,6	1,0 0	0,6 0,4	0 1,0

(продолжение таблицы)

	Tyr	Trp	Ser	Thr	Asp	Asn	Glu	Gln	Cys	Gly
Tyr	0,25 0,2	0 1,0	0,35 0,35	0 0,7	0,15 0,20	0,6 0,3	0 0,2	0,8 0,3	0 1,0	0,35 0,4
Trp	0 1,0	0 1,0	0,25 0,15	0 1,0		0 1,0	0 1,0		0 1,0	0 1,0
Ser	0 1,0	0,25 0,75	0,25 0,15	0,2 0,35	0,25 0,30	0,60 0,40	0,45 0,55	0,55 0,45	0,2 0,8	0,2 0,15
Thr	0,3 0,1	0 0,75	0,25 0,45	0,05 0,25	0,35 0,4	0,45 0,20	0,65 0,35	0,45 0,2	0,35 0,65	0 0,8
Asp	0,5 0	0 1,0	0 1,0	0,15 0,42	0,25 0,75	0 1,0	0,75 0,25	0 0,5	0,5 0	0,2 0,8
Asn	0,35 0,65	0,5 0,5	0,1 0,9	0,15 0,45	0,3 0,7	0,1 0,9	0,8 0,2	0,2 0,55	0,45 0,1	0,1 0,9
Glu	0,85 0,15	0 1,0	0,6 0,4	0,6 0,1	0,75 0,25	0,45 0,55	0,75 0,25	0,9 0,1	0 1,0	0,3 0,7
Gln	0,35 0,1	0,35 0,65	0,35 0,25	0,2 0,3	0,2 0,8	0,55 0,45	0,75 0,25	0,15 0,85	0,8 0,2	0,45 0,55
Cys	0 0,2	0,8	0,55 0,45	0 1,0	0 1,0	0,25 0,75	0,65 0,35	0 0,35	0 1,0	0,15 0,35
Gly	0 0,75	0 0,1	0,05 0,6	0,2 0,45	0,25 0,75	0,1 0,9	0,4 0,35	0,15 0,85	0 1,0	0 0,4

(продолжение таблицы)

	Tyr	Trp	Ser	Thr	Asp	Asn	Glu	Gln	Cys	Gly
Ala	0,3 0,2 0,5	0,6 0,1 0,3	0,45 0,3 0,25	0,1 0,1 0,8	0,6 0,25 0,15	0,4 0,3 0,3	0,9 0,1 0	0,7 0,3 0	0,3 0,1 0,6	0,25 0,35 0,4
Phe	0,3 0,6 0,1	0 1,0 0	0,6 0,4 0	0,4 0,5 0,1	0,25 0,75 0	0 0,6 0,4	0,5 0,1 0,4	0 1,0 0	0 0,2 0,8	0,4 0,6 0
Leu	0,5 0,35 0,15		0,6 0,15 0,25	0,5 0,5 0	0,65 0,35 0	0 0,35 0,65	0,85 0,15 0	0,2 0,2 0,6	0 0,45 0,55	0,45 0,35 0
Pro	0,2 0,8 0	0 1,0 0	0 1,0 0	0,2 0,3 0,5	0 1,0 0	0 1,0 0	0,65 0,35 0	0,6 0,4 0	0 1,0 0	0 1,0 0
Met	0 1,0 0		0,25 0,55 0,2	0 1,0 0		0,6 0,4 0	1,0 0 0	0 1,0 0	0 1,0 0	0 1,0 0
Val	0,07 0,85 0,08	1,0 0 0	0,1 0,6 0,3	0,35 0,4 0,25	0,45 0,55 0	0,7 0,3 0,3	0,5 0,25 0,25	0,65 0,35 0	0,2 0,25 0,55	0,1 0,3 0,6
Ile	0 0,9 0,1	0 1,0 0	0,3 0,55 0,15	0,1 0,2 0,7	0,3 0,6 0,1	0,4 0,35 0,25	0,15 0,75 0,1	0,4 0,6 0	0,4 0,2 0,4	0 0,5 0,5
His	0,4 0,6 0		0,45 0,55 0	0 1,0 0		0 1,0 0	0,6 0,4 0		0,2 0,35 0,45	0,75 0,25 0
Lys	0,35 0,65 0		0,35 0,25 0,4	0,45 0,2 0,35	0,55 0,45 0	0,20 0,4 0,4	0,45 0,55 0	0,45 0,55 0	0 1,0 0	0 1,0 0
Arg	0 0,4 0,6	0,65 0,35 0	0,4 0,25 0,35	0,3 0,7 0	0 0,2 0,8	0,45 0,55 0	0,45 0,35 0,20	0,55 0,45 0	0,8 0,2 0	0 0,6 0,4

(окончание таблицы)

	Ala	Phe	Leu	Pro	Met	Val	Ile	His	Lys	Arg
Tyr	0,15 0,05	0,3 0,7 0	0,6 0,1 0,3	0 1,0 0	0	0 0,6 0,4	0,1 0,1 0,8	0 0,25 0,75	0,25 0,15 0,6	0,2 0,8 0
Trp	0,8 0,2	0	0 1,0 0	0 1,0 0	0	0,2 0,65 0,15	1,0 0 0	1,0 0 0	0 1,0 0	0,65 0,35 0
Ser	0,65 0,25	0,1 1,0	0,15 0,7 0,15	0,6 0,4 0	0,25 0,2 0,55	0,25 0,4 0,35	0,15 0,55 0,30	0,6 0,4 0	0,45 0,55 0	0,35 0,65 0
Thr	0,6 0,4	0,2 0,1 0,7	0,3 0,2 0,5	0,25 0,75 0	0	0,3 0,45 0,25	0,4 0,2 0,4	0,4 0,1 0,5	0,45 0,45 0,1	0 0,5 0,5
Asp	0,6 0,25	0,15 0,5	0,45 0,35 0,2	0,35 0,65 0	0	0,5 0,3 0	1,0 0 0	0 1,0 0	0,55 0,45 0	0,2 0,8 0
Asn	0,1 0,5	0,45 0,55	0,4 0,6	0 1,0	0 0,85 0,15	0,5 0,5	0 0,6 0,4	0 1,0 0	0,35 0,35 0,3	0,25 0,15 0,6
Glu	0,85 0,15	0,9 0,1	0,85 0,15	0 1,0	0,35 0,65 0	0,15 0,75 0,1	0 1,0 0	0 1,0 0	0,6 0,4 0	0,5 0,4 0,1
Gln	0,25 0,15	0,6 0,2	0,65 0,35	0 0,5	1,0 0	0,35 0,5 0,15	0 0,65 0,35	0,25 0,6 0,15	0,1 0,6 0,3	0 1,0 0
Cys	0,3 0,1	0,6	0,15 0,7 0,15	0 1,0	0,65 0,35 0	0 1,0 0	1,0 0 0	1,0 0 0	0,35 0,35 0,3	0 0,55 0,45
Gly	0,4 0,3	0,3 0,4	0,1 0,9	0,15 0,85	0 1,0	0,15 0,45 0,4	0,2 0,4 0,4	0,4 0,6	0,4 0,6	0 0,5

Сравнение экспериментальных данных с результатами расчета вторичной структуры белков *

Белок	α-Спиральные участки **		Участки с β-структурой **		Количество аминокислотных остатков в белке	Точность расчета, %		
	Эксперимент	Расчет	Эксперимент	Расчет		α-структура (Qα)	β-структура (Qβ)	(α+β)-структура (Q)
Миоглобин [29]	3-18, 20-35, 36-42, 51-57, 58-77, 86-94, 100-118, 124-148 Нет	5-22, 27-36, 37-43, 49-79, 82-87, 89-95, 100-119, 124-149 27-32	Нет	Нет	153	87	100	87
Рубредоксин [24]	Нет	27-32	4-6, 11-13, 17-19, 42-44, 49-52	4-6, 11-14, 17-19, 36-38, 48-52	53	89	85	74
Рибонуклеаза S [32]	3-13, 28-34, 50-57	2-13, 28-32, 49-57	44-48, 71-75, 79-87, 96-100	45-48, 69-75, 79-87, 94-104, 106-111	124	97	94	91
Нуклеаза [31]	54-67, 99-106, 122-134	56-67, 97-106, 108-113, 120-136	12-19, 21-27, 30-36, 71-75, 91-95	12-19, 22-28, 31-39, 71-75, 90-94, 114-116	149	91	92	83
Папаин [21]	26-41, 50-56, 69-78, 116-126	26-35, 50-56, 69-78, 119-127, 159-164	163-167, 170-173	130-133, 134-136, 149-151, 170-174, 186-188	212	92	87	83
Глобин многи [37]	12-29, 31-44, 45-52, 67-88, 91-105, 112-127, 132-148 2-6, 47-56	11-24, 36-43, 46-52, 63-88, 91-98, 112-128, 133-148 47-53	Нет	Нет	148	83	100	83
Ингибитор трипси-на [28]	3-18, 20-35, 36-42, 52-71, 80-88, 94-112, 118-138	2-17, 20-34, 40-45, 47-64, 78-89, 96-112, 119-130	18-24, 29-35	4-6, 18-24, 29-35	58	86	95	86
Гемоглобин, α-цепь [19]	5-10, 14-20, 64-73, 103-117, 132-145, 223-238, 242-252, 269-275	12-19, 63-69, 71-75, 88-95, 103-109, 112-117, 132-145, 174-180, 194-198, 223-236, 242-248, 267-275	Нет	Нет	141	74	100	74
Сублилизин [34]	8-18, 26-33, 40-51, 60-70, 79-89, 99-107	8-22, 27-33, 39-51, 59-69, 76-88, 100-108	Нет	56-58	275	84	89	76
Миоген [23]	8-18, 26-33, 40-51, 60-70, 79-89, 99-107	8-22, 27-33, 39-51, 59-69, 76-88, 100-108	Нет	56-58	108	87	97	84
Всего по 10 белкам					1421	87	94	82

* Приведены литературные экспериментальные данные о вторичной структуре белков
 ** Приведены порядковые номера аминокислотных остатков в полипептидной цепи.

Таблица 3

Теоретическая локализация α -спиральных и β -структурных участков в белках с неизвестной пространственной структурой

Белок *	α -Спиральные участки	Участки с β -структурой	Количество аминокислотных остатков в белке	Спиральность, %	β -Структурированность, %
LiV-связывающий белок [40]	2-6, 25-36, 56-67, 84-88, 100-109, 113-117, 135-139, 141-147, 176-187, 209-213, 236-242, 256-270, 274-278, 300-307, 339-344	7-12, 43-47, 89-91, 95-99, 132-134, 148-150, 169-171, 188-190, 196-199, 231-233, 244-246, 279-282, 284-289	344	35	14
Нейротоксин II [41]	6-10, 19-22	12-14, 23-25, 34-38	61	15	18
α -Токсин [42]	6-10, 19-22, 24-29	12-14, 34-38, 55-57	61	24	18
Адренодоксин [43]	38-41, 49-53, 57-61, 63-75, 78-82, 108-117	6-10, 11-15, 28-30, 32-35, 92-96, 102-107	117	36	24
Фрагмент I протромбина [44]	3-10, 13-16, 18-23, 25-33, 45-50, 55-62, 64-67, 106-111	36-38, 39-41, 78-80, 84-87, 102-104, 147-152	155	33	14
Фрагмент II протромбина [45]	12-15, 27-33, 37-48, 58-61, 93-101, 105-108, 110-115	21-24	116	39	3
Гистон H2A [46]	12-16, 38-42, 50-61, 62-66, 84-88, 90-93, 99-104, 118-127	23-25, 29-31, 70-72, 95-97, 115-117	129	40	12
Гистон H2B [47]	15-25, 68-78, 103-111, 116-120	39-44, 45-49, 88-90, 99-102	125	29	14
Гистон H3 [48]	17-29, 45-50, 58-65, 74-83, 91-101, 103-108, 111-115, 122-127	69-71, 118-120	135	48	4
Гистон H4 [49]	18-24, 62-68, 76-86	50-53, 59-61, 87-92, 97-99	102	25	16

* Приведены литературные источники данных о первичной структуре белка.

или не- β -структурное состояние 94% всех остатков. При одновременном определении всех элементов вторичной структуры (α -спираль, β -структура и нерегулярные участки) метод правильно предсказывает структурное состояние 82% всех аминокислотных остатков, входящих в рассмотренные белки.

Для α -спиралей метод правильно предсказывает структурное состояние 85% $\{[(605 - 93)/605] \cdot 100\}$ всех аминокислотных остатков, входящих в α -спиральные участки. Для исследованных белков, содержащих в спиральном состоянии 605 аминокислотных остатков, метод исключает из α -спиралей 93 «спиральных» аминокислотных остатка и включает в спиральные участки 93 «лишних» аминокислотных остатка. Таким образом, при определении спиральных участков погрешность метода, равная 13% (100-87), на 6,5% обусловлена предсказанием «лишних спиральных» аминокислотных остатков и на 6,5% — недопредсказанием «спиральных» остатков.

В случае расчета β -структурных участков метод правильно предсказывает структурное состояние 82% $\{[(132 - 24)/132] \cdot 100\}$ всех аминокислотных остатков, входящих в β -структурные участки. Метод неправильно определяет структурное состояние 24 аминокислотных остат-

ков из 132, входящих в β -структурные участки, и включает в эти участки 64 «лишних» остатка. Таким образом, погрешность при определении β -структурных участков, равная 6% (100—94), на 1,7% обусловлена исключением « β -структурных» аминокислотных остатков и на 4,3% включением в β -структурные участки «лишних» аминокислотных остатков.

Некоторое расхождение между рассчитанными и экспериментальными литературными данными, по-видимому, обусловлено влиянием дальних взаимодействий на формирование элементов вторичной структуры белков, а у глобинов, кроме того, также влиянием гема.

Сравнение количественных параметров точности предсказания элементов вторичной структуры глобулярных белков, полученных нашим методом и наиболее признанным методом, основанным на общих стереохимических соображениях о возможности встраивания участков вторичной структуры в гидрофобное ядро [39], дает основание полагать, что предлагаемый метод также может быть эффективно применен для предсказания структурированных участков белков с неизвестной пространственной структурой. В табл. 3 приведены предсказанные предлагаемым методом структурированные участки некоторых белков с неизвестной пространственной структурой.

ЛИТЕРАТУРА

1. Птицын О. Б., Финкельштейн А. В. [(1978) *Биоорганич. химия*, 4, 349—353.
2. Финкельштейн А. В. (1978) *Биоорганич. химия*, 4, 345—348.
3. Лим В. И., Мазанов А. Л., Ефимов А. В. (1978) *Молекулярн. биология*, 12, 206—213.
4. Лим В. И. (1978) *Молекулярн. биология*, 12, 214—217.
5. Мазанов А. Л., Лим В. И. (1978) *Молекулярн. биология*, 12, 219—232.
6. Tanaka S., Scheraga H. A. (1977) *Proc. Nat. Acad. Sci. USA*, 74, 1320—1323.
7. Guzzo A. V. (1965) *Biophys. J.*, 5, 809—822.
8. Prothero J. W. (1966) *Biophys. J.*, 6, 367—370.
9. Ptitsyn O. B. (1969) *J. Mol. Biol.*, 42, 501—510.
10. Птицын О. Б., Финкельштейн А. В. (1970) *Докл. АН СССР*, 195, 221—224.
11. Птицын О. Б., Финкельштейн А. В. (1970) *Биофизика*, 15, 757—768.
12. Chou P. Y., Fasman G. D. (1974) *Biochemistry*, 13, 211—222.
13. Maxfield F. R., Scheraga H. A. (1975) *Macromolecules*, 8, 491—493.
14. Fasman G. D., Chou P. Y., Adler A. J. (1976) *Biophys. J.*, 16, 1201—1238.
15. Денесюк А. И., Птицын О. Б., Финкельштейн А. В. (1974) *Биофизика*, 19, 549—561.
16. Green N. M., Flanagan M. T. (1975) *Biochem. J.*, 153, 729—732.
17. Lenstra J. A., Hofsteenge J., Beintema I. I. (1977) *J. Mol. Biol.*, 109, 185—191.
18. Есипова Н. Г., Туманян В. Г. (1972) *Молекулярн. биология*, 6, 840—851.
19. Peutz M. F. (1969) *Encyclor. Polymer. Sci. and Technol.*, 11, 646.
20. Matthews B. W., Weaver L. H., Kester W. R. (1974) *J. Biol. Chem.*, 249, 8030—8044.
21. Dreuth J., Jansonius J. N., Koekoek R., Swen H. M. (1968) *Nature*, 218, 929—932.
22. Mathews F. S., Levine M., Argos P. (1971) *Nature New Biol.*, 233, 15—16.
23. Kretsinger R. H., Nockolds C. E. (1973) *J. Biol. Chem.*, 248, 3313—3326.
24. Herriott J. R., Sicker L. C., Jensen L. H., Lovenberg W. (1970) *J. Mol. Biol.*, 50, 391—406.
25. Dickerson R. E., Takano T., Eisenberg D., Kallai O. B., Samson L., Cooper A., Margoliash E. (1971) *J. Biol. Chem.*, 246, 1511—1535.
26. Blow D. M. (1969) *Biochem. J.*, 112, 261—268.
27. Stroud R. M., Kay L. M., Dickerson R. E. (1972) *Cold Spring Harbor Sympos. Quant. Biol.*, 36, 125.
28. Huber R., Kukla D., Ruhlmann A., Steigeman W. (1971) *Proc. Int. Res. Conf. Protease Inhibitors. München, 1970* (Fritz H., Tschesche H., eds.), pp. 56—65, Walter de Gruyter, Berlin.
29. Watson H. C. (1969) *Progr. Stereochem.*, 4, 299—333.
30. Low B. W., Preston H. S., Sato A., Rosen L., Scarl J. E., Rudko A., Richardson J. S. (1976) *Proc. Nat. Acad. Sci. USA*, 73, 2991—2994.
31. Arnone A., Bier C. J., Cotton F. A., Day V. W., Hasen E. E., Richardson J. S., Yonath A. (1971) *J. Biol. Chem.*, 246, 2302—2316.
32. Wyckoff H. W., Tsernoglou D., Hanson A. W., Knox I. R., Lee B., Richards F. M. (1970) *J. Biol. Chem.*, 245, 305—328.
33. Shotton D. M., Watson H. C. (1970) *Nature*, 225, 811—816.
34. Wright C. S., Alden R. A., Kraut J. (1969) *Nature*, 221, 235—242.

35. Blundell T. L., Cutfield J. F., Cutfield S. M., Dodson E. J., Dodson G. G., Hodgkin D. C., Mercola D. A., Vijayan M. (1971) *Nature*, **231**, 506—511.
36. Braut D. A., Schimmel P. R. (1967) *Proc. Nat. Acad. Sci. USA*, **58**, 428—435.
37. Heudrickson W. A., Love W. E. (1971) *Nature New Biol.*, **232**, 197—203.
38. Rajan S. S., Srinivasan R. (1976) *Curr. Sci.*, **45**, 859—860.
39. Лим В. И. (1974) *Биофизика*, **19**, 562—575.
40. Ovchinnikov Yu. A., Aldanova N. A., Grinkevich Y. A., Arzamazova N. M., Moroz I. N. (1977) *FEBS Lett.*, **78**, 313—316.
41. Grishin E. V., Sukhikh A. P., Lukuanchuk N. N., Lipkin V. M., Ovchinnikov Yu. A., Sorokin V. M. (1973) *FEBS Lett.*, **36**, 77—78.
42. Eaker D. L., Porath I. (1967) *Jap. J. Microbiol.*, **11**, 353—355.
43. Ахрем А. А., Лапко А. Г., Лапко В. Н., Морозова Л. А., Репин В. А., Тищенко И. В., Чащин В. Л. (1978) *Биоорг. химия*, **4**, 462—475.
44. Walz D. A., Hewett-Emmett D., Seegers W. H. (1977) *Proc. Nat. Acad. Sci. USA*, **74**, 1969—1972.
45. Butkowski R. J., Elion J., Downing M. R., Mann K. G. (1977) *J. Biol. Chem.*, **252**, 4942—4957.
46. Yeoman L. C., Olson M. O. J., Sugano N., Jordon J. J., Taylor C. W., Starbuck W. C., Busch H. (1972) *J. Biol. Chem.*, **247**, 6018—6023.
47. Iwai K., Ischikawa K., Hayashi H. (1970) *Nature*, **226**, 1056—1058.
48. De Lange R. J., Hooper J. A., Smith E. L. (1972) *Proc. Nat. Acad. Sci. USA*, **69**, 882—884.
49. De Lange R. J., Fambrough D. M., Smith E. L., Bonner J. (1969) *J. Biol. Chem.*, **244**, 319—334.

Поступила в редакцию
28.VI.1978

SECONDARY STRUCTURE PREDICTION FOR GLOBULAR PROTEINS FROM THEIR AMINO ACID SEQUENCE

BARKOVSKY E. V., BANDARIN V. A.

Minsk State Medical Institute, Minsk

A method for localization of secondary structural regions (α -helices and β -structures) by the amino acid sequence is proposed for globular proteins. The method predicts the helical or non-helical, and β -structural or non- β -structural state for all the residues with 87 and 94% accuracy, respectively. α -Helical and β -structural regions are predicted for a number of proteins with unknown three-dimensional structure.
