



УДК 577.152.31\*273:577.322.4

Me+

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ АМИНОКИСЛОТНЫХ  
ПОСЛЕДОВАТЕЛЬНОСТЕЙ И РАСЧЕТ ВТОРИЧНЫХ СТРУКТУР  
ГУАНИЛСПЕЦИФИЧНЫХ РИБОНУКЛЕАЗ *PENICILLIUM  
BREVICOMPACTUM*, *ASPERGILLUS ORYZAE* (T<sub>1</sub>)  
И *ASPERGILLUS CLAVATUS* (C<sub>2</sub>)**

**Шляпников С. В., Буликов В. А., Яковлев Г. И.**

*Институт молекулярной биологии Академии наук СССР, Москва*

По модифицированному методу Чоу и Фасмана проведен расчет вторичных структур внеклеточных гуанилспецифичных РНКаз аскомицетов грибов *Penicillium brevicomactum*, *Aspergillus oryzae* (T<sub>1</sub>) и *Aspergillus clavatus* (C<sub>2</sub>). Сравнительный анализ первичных и вторичных структур РНКаз свидетельствует о высокой степени гомологии ферментов (64–68%), повышенной инвариантности аминокислотных остатков гидрофобного характера и о подобии в распределении участков регулярных конформаций ( $\alpha$ -спирали,  $\beta$ -структуры,  $\beta$ -повороты) по ходу полипептидной цепи белков. Выявлены консервативные и переменные области структур РНКаз. Проведено сравнение теоретически предсказываемых вторичных структур РНКаз с данными рентгеноструктурных исследований.

В настоящее время принято считать, что третичная структура белков, а следовательно, и их функция однозначно определены аминокислотной последовательностью. Вместе с тем по-прежнему актуальной остается проблема анализа информации, заложенной в первичной структуре, и способов ее реализации при пространственной упаковке белков.

Согласно современной физической концепции самоорганизации белков [1–4], процесс компактизации белковых глобул, хотя и несколько условно, может быть разделен кинетически на два этапа: формирование локально упорядоченных участков вторичной структуры и последующее взаимодействие их между собой с образованием уникальных функционально активных конформаций. Несмотря на то что структура каждого белка уникальна, описание ее в первом приближении сводится к анализу элементов вторичных структур, присутствующих в молекуле, и определению их взаимного расположения. При этом отмечены новые уровни структурной организации белковых глобул и найдены некоторые общие эмпирические закономерности их топологии [2, 5–8].

Ответ на вопросы, каким образом различные аминокислотные последовательности могут образовывать подобные трехмерные структуры и как белки адаптируются к мутациям в ходе природной селекции, сохраняя при этом «эволюционную память», в той или иной степени дает сравнительное изучение семейств белков с известной структурой и функцией, в частности сериновых протеиназ [9], гемоглобинов [10, 11], цитохромов с [12], иммуноглобулинов [13] и NAD-зависимых дегидрогеназ [14]. В какой-то мере к уровню исследования этой группы белков приближается изучение обособленного семейства внеклеточных рибонуклеаз (РНКаз) микроорганизмов, объединяющих две родственные подгруппы ферментов эукариотического и прокариотического происхождения [15]. Различаясь некоторыми свойствами и специфичностью, РНКазы микроорганизмов подобно панкреатическим РНКазам осуществляют тождественную функцию — реакцию расщепления полинуклеотидной цепи РНК путем трансфосфорилирования с образованием олиго- и мононуклеозид-2', 3'-циклофосфатов и дальнейший гидролиз последних до соответствующих производных 3'-фосфорной кислоты [16]. Для некоторых представителей этого класса РНКаз определена аминокислотная последовательность и ус-

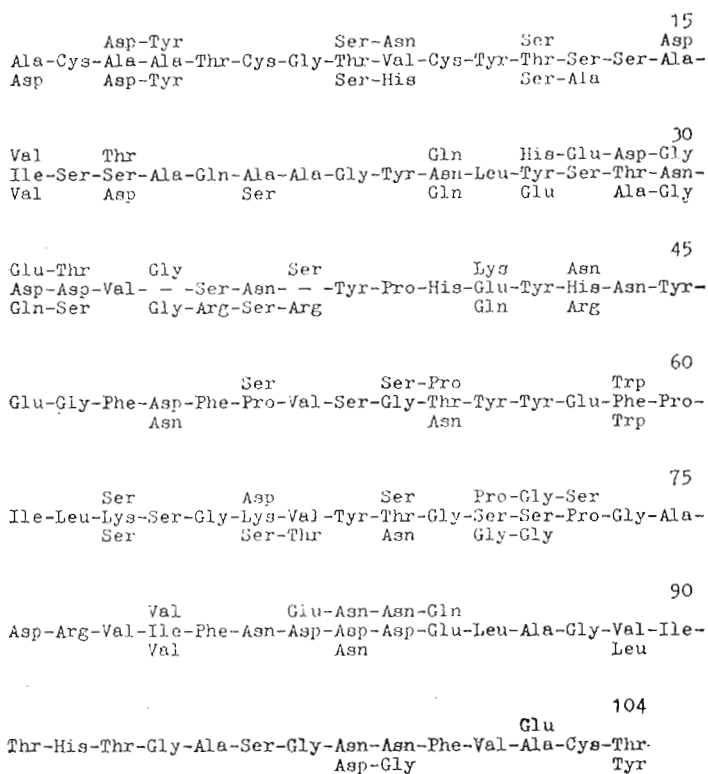


Рис. 1. Первичные структуры гуанилспецифичных РНКаз грибов. Представлена аминокислотная последовательность *P. brevicom pactum* [23] и аминокислотные остатки РНКаз  $T_1$  и  $C_2$  вверху и внизу соответственно, отличные от остатков РНКазы *P. brevicom pactum*

таповлена трехмерная структура с атомным разрешением [15, 17—22].

Недавно нами определена аминокислотная последовательность еще одной внеклеточной РНКазы — гуанилспецифичной РНКазы из гриба *Penicillium brevicom pactum* [23]. В настоящей работе проводится анализ первичной и вторичной структур РНКазы *P. brevicom pactum*, при этом соответствующие структуры анализируются сравнительно со структурами РНКаз  $T_1$  и  $C_2$  (*Aspergillus oryzae* и *A. clavatus*) [18, 22]. Существенная гомология первичных структур гуанилспецифичных РНКаз значительно упрощает проведение их сравнения. Кроме того, наличие данных рентгеноструктурного изучения РНКаз  $T_1$  и  $C_2$  позволяет проанализировать результаты расчета вторичных структур молекул на основании их аминокислотных последовательностей. В этой связи следует отметить, что расчетный метод — единственный, позволяющий предсказать картину распределения элементов вторичной структуры исследуемого белка по ходу полипептидной цепи. Сопоставление экспериментальных и расчетных данных может, таким образом, способствовать решению одной из задач, встающих при изучении трехмерной структуры белка, — сопоставления наблюдаемых при этом фрагментов белковой молекулы с конкретными участками полипептидной цепи.

Гомология аминокислотных последовательностей РНКазы *P. brevicom pactum* и РНКазы  $T_1$  (с учетом двух делеций в положениях 34 и 37 молекулы РНКазы *P. brevicom pactum*) составляет 66%, а с РНКазой  $C_2$  — 64% (рис. 1). Степень гомологии РНКаз  $T_1$  и  $C_2$  равна 68%\*.

Анализ единичных замещений аминокислотных остатков в первичных структурах РНКаз  $T_1$  и  $C_2$  сравнительно с последовательностью РНКазы

\* Здесь и далее приведена нумерация остатков по аминокислотной последовательности РНКаз  $T_1$  и  $C_2$ ; две делеции в положениях 34 и 37 введены в первичную структуру РНКазы *P. brevicom pactum*.

*P. brevicompactum* (рис. 1) показывает, что из 35 замен, наблюдаемых для РНКазы  $T_1$ , и 37 — для РНКазы  $C_2$ , только в 10 и 15 случаях соответственно необходимо изменение двух нуклеотидов в триплетных кодонах, кодирующих данные остатки. Остальные варианты аминокислотных замен удовлетворяют требованию одноточечных нуклеотидных мутаций. Идентичные относительно РНКазы *P. brevicompactum* замещения в РНК-азах  $T_1$  и  $C_2$ , связанные с заменой одного или двух нуклеотидов в кодонах, локализованы в полипептидных цепях в положениях 3, 8, 12, 16, 34, 72, 79, 83 и 4, 25, 30, 59, 63 соответственно.

Максимальная вариабельность в структуре всех трех гуанилспецифичных РНКаз наблюдается главным образом в двух участках полипептидной цепи: 27—37 и 66—71; кроме того, имеются различия также в позициях 9, 18, 41, 43 и 55. Инвариантные остатки РНКаз в свою очередь сгруппированы в четырех сегментах полипептидной цепи ферментов. Эти участки структуры представляют собой последовательности, состоящие из 5 (44—48), 7 (56—62), 8 (74—81) и 12 (86—97) аминокислотных остатков, причем в последних трех пептидных сегментах наблюдается замена остатков с ярко выраженными гидрофобными свойствами: остатки  $\text{Phe}^{59}$  и  $\text{Ile}^{79}$  в РНКазе *P. brevicompactum* заменены в гомологичных белках соответственно остатками  $\text{Trp}^{59}$  и  $\text{Val}^{79}$ , тогда как место остатка  $\text{Val}^{89}$  в РНКазе *P. brevicompactum* в РНКазе  $C_2$  занимает остаток  $\text{Leu}^{89}$ .

Существенным в этой связи является тот факт, что перечисленные консервативные сегменты структуры РНКаз включают в себя остатки активного центра ферментов —  $\text{Glu}^{38}$  и  $\text{His}^{92}$ , которые непосредственно участвуют в процессе ферментативного катализа, а также остатки  $\text{Arg}^{77}$ ,  $\text{Asn}^{44}$  и  $\text{Tyr}^{45}$ , играющие важную роль в специфическом узнавании субстрата, его связывании и стабилизации промежуточных соединений [15, 20].

Анализ профилей гидрофобности ( $H$ ) и доступности аминокислотных остатков РНКазы *P. brevicompactum* растворителю ( $B$ ) при условии отнесения остатков к гидрофобным значениям  $H \geq 12$  и  $B \geq 0,3$  [24] (рис. 2а) дает в целом схожую картину. Локальные замещения в структурах гомологичных ферментов не сказываются существенным образом на общем характере амфифильности в полипептидных цепях сравниваемых молекул. РНКазы *P. brevicompactum* содержит в сумме 56 остатков гидрофобного и 46 остатков гидрофильного характера. Соответствующие значения для РНКазы  $T_1$  и РНКазы  $C_2$  — 53/51 и 58/46. Вместе с тем при рассмотрении амфифильных свойств инвариантных в структуре трех ферментов аминокислотных остатков обнаруживается значительный консерватизм остатков гидрофобного характера по сравнению с гидрофильными остатками. Так, из 58 аминокислотных остатков, идентичных для всех трех РНКаз, 41 представляют собой остатки гидрофобного и 17 — гидрофильного характера, что составляет в среднем 74 и 36% соответственно от общего содержания их в молекулах.

Анализ амфифильных свойств переменных аминокислотных остатков РНКаз также свидетельствует о высоком консерватизме остатков гидрофобного характера. При сравнении первичных структур трех РНКаз только в пяти случаях обнаруживаются замены гидрофобных остатков на гидрофильные, тогда как для гидрофильных аминокислотных остатков число соответствующих по характеру аминокислотных замен равно 22.

Инвариантные по амфифильным свойствам аминокислотные остатки в первичных структурах сравниваемых РНКаз расположены в основном в сегментах полипептидной цепи 38—53, 55—66 и 73—98. Эти участки включают в себя ранее выделенные консервативные области аминокислотных последовательностей РНКаз, частично расширяют их, а в последнем случае объединяют две близко расположенные консервативные области в единое целое.

Консерватизм в распределении гидрофобных остатков в аминокислотных последовательностях гомологичных белков — важная черта первичных структур, предполагающая существенную роль аминокислотных остатков гидрофобного характера в формировании пространственной архитектуры белковых молекул. Согласно современным представлениям,

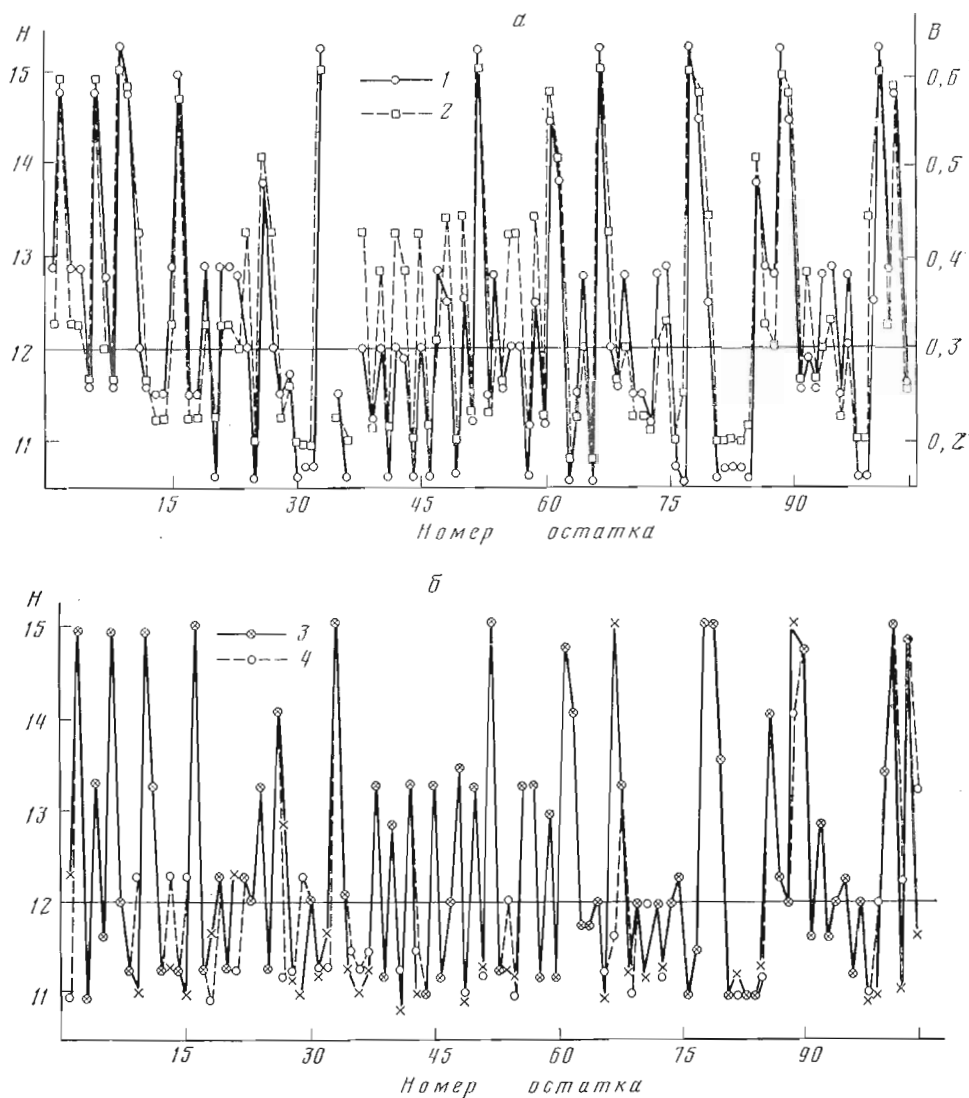
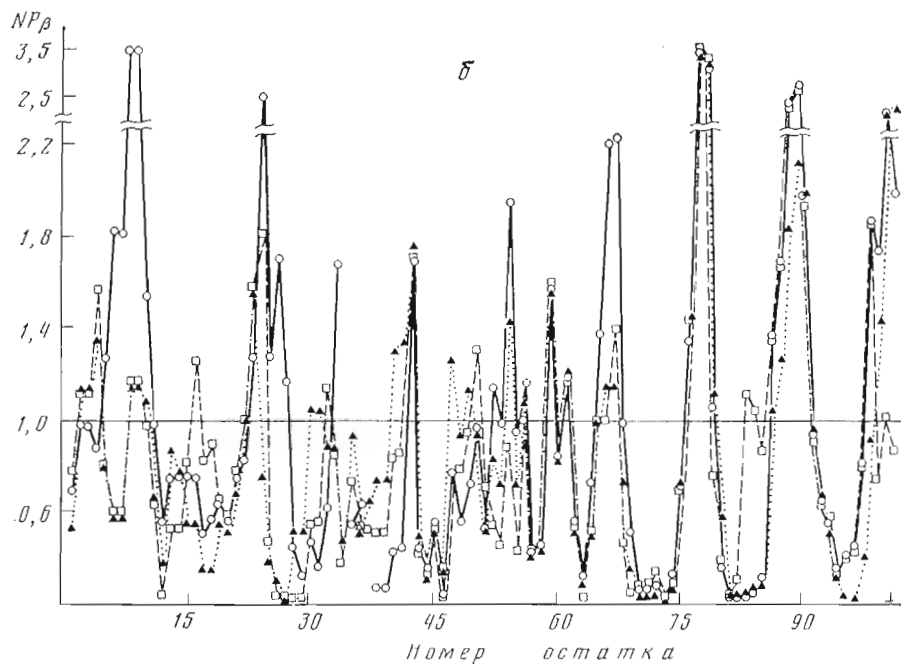
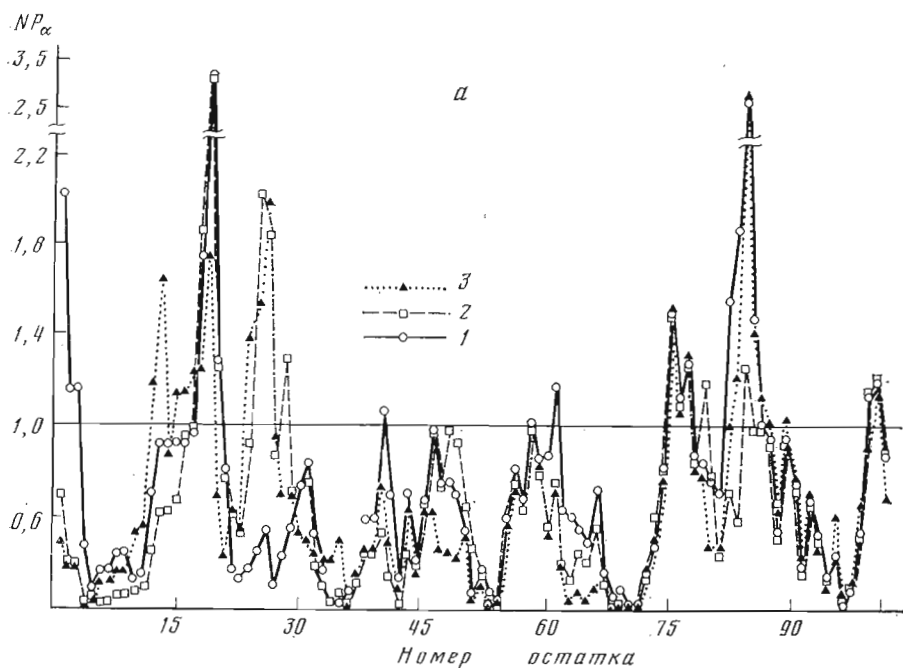


Рис. 2. Профили гидрофобности ( $H$ ) и доступности аминокислотных остатков РНКаз-растворителю ( $B$ ). а - РНКазы *P. brevicompactum* (1 -  $H$ , 2 -  $B$ ), б - РНКазы  $T_1$  (3) и  $C_2$  (4)

укладка полипептидной цепи белка определяется сравнительно небольшим количеством ( $\sim 30-40\%$ ) структурно эквивалентных аминокислотных остатков, которые участвуют во внутримолекулярных контактах, находятся внутри белковой глобулы и обеспечивают связывание субстрата или простетических групп ферментов. Большая часть структурно эквивалентных остатков включена в образование блоков вторичной структуры белков, тогда как переменные аминокислотные остатки, как правило доступные растворителю, находятся в областях неупорядоченной структуры, полипептидных петлях и изгибах цепи. В свою очередь для остатков, недоступных растворителю, существенно сохранение в первом приближении лишь их общей гидрофобности, при этом размеры боковых радикалов остатков могут значительно меняться [25-27].

Кристаллографическое изучение РНКаз  $T_1$  и  $C_2$  свидетельствует о подобии пространственной организации этих ферментов. Согласно данным рентгеноструктурного исследования, РНКазы  $T_1$  и  $C_2$  являются, по классификации Левитта, ( $\alpha$ - $\beta$ )-белками [5]. Топология их упаковки описывается расположенным в центральной части молекул, скрученным из че-



четырёх антипараллельных  $\beta$ -цепей  $\beta$ -слоем (остатки 40–42, 57–61, 76–81 и 84–90/92), который примыкает к единственной в белках  $\alpha$ -спирали, состоящей из 4, 5 витков (12/13–29) [17, 20].

На рис. 3 (а–в) для перекрывающихся тетрапептидных сегментов гуанилспецифичных РНКаз представлено распределение конформационных потенциалов двух типов вторичной структуры ( $\alpha$ -спиралей,  $\beta$ -структур), а также вероятностей  $\beta$ -поворотов, полученное в соответствии с модифицированным [28] методом расчета вторичных структур белков Чоу и Фасмана [29], исходя из аминокислотных последовательностей (см. раздел «Расчетная часть»).

Детальная схема локализации предсказываемых структур по ходу полипептидных цепей РНКаз показана на рис. 4а. Дополнительно к рассмотренным значениям конформационных потенциалов здесь проведено

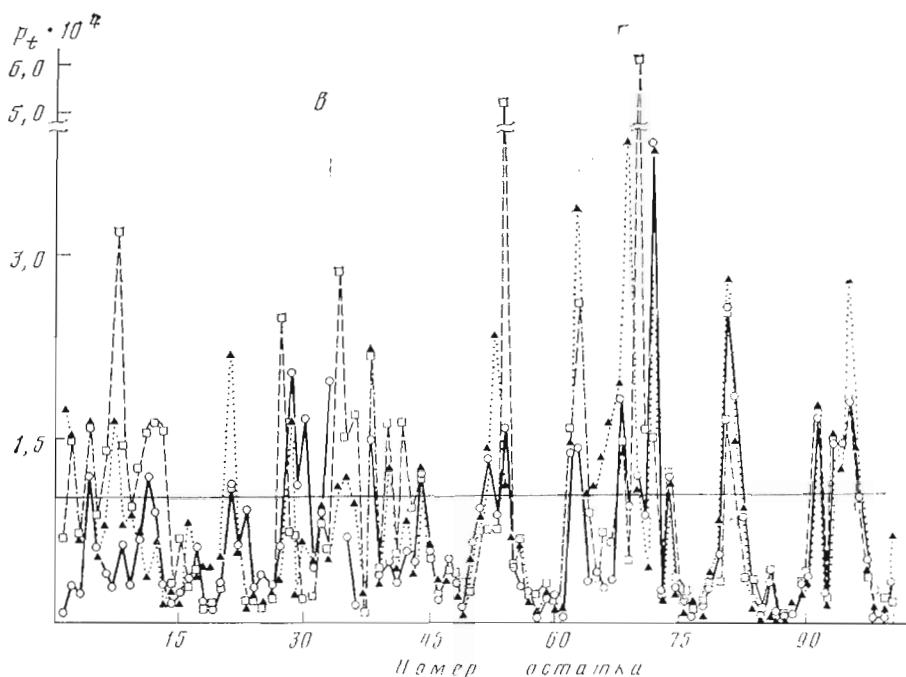


Рис. 3. Произведения конформационных потенциалов ( $NP_\alpha$  и  $NP_\beta$ ) аминокислотных остатков ( $a$  —  $\alpha$ -спираль,  $b$  —  $\beta$ -структура) и вероятность  $\beta$ -поворотов ( $\sigma - \pi$ ) в тетрапептидах РНКаз *P. brevicompactum* (1),  $T_1$  (2) и  $C_2$  (3). Горизонтальные линии — граничные значения предсказываемых конформаций

определение границ сегментов вторичных структур. При этом, однако, не учтено частичное или полное их совмещение и конкурентное перекрытие. Анализ последних в значительной степени основан на сопоставлении относительных величин конформационных потенциалов того или иного элемента, вторичной структуры, в частности  $\beta$ -поворотов (см. таблицу), и наряду с некоторыми другими ограничениями предсказаний [29, 30] сводится к следующему.

1. Во всех трех белках N-концевая область РНКаз (1/2—10/12) не может целиком представлять вытянутую  $\beta$ -конформацию, что объясняется наличием внутримолекулярной S—S-связи остатков Cys<sup>2</sup> и Cys<sup>10</sup> в нативных ферментах [22, 23]. Вместе с тем нельзя исключить полностью короткого  $\beta$ -структурного сегмента, например 8—11 для РНКазы *P. brevicompactum* или 4—7 для РНКаз  $T_1$  и  $C_2$ , стабилизируемого, в частности, за счет взаимодействия с предсказываемым для всех анализируемых белков элементом  $\beta$ -структуры в С-концевой части их молекул (96/97—101). Профиль гидрофобности в полипептидных цепях РНКаз (рис. 2) свидетельствует о том, что С-концевой участок  $\beta$ -структуры молекул частично экранирован от растворителя (см. [24]). На этом основании можно предполагать вовлечение его в формирование  $\beta$ -складчатой структуры. Потенциальными партнерами С-концевого участка  $\beta$ -структуры в  $\beta$ -слое могли бы оказаться в таком случае ближайшие две также экранированные от растворителя (рис. 2)  $\beta$ -цепи 74—81 и 83/85—92/93 (рис. 4) и/или некий сегмент  $\beta$ -структуры в N-концевой части молекул РНКаз. Топологическая сближенность N- и С-концевых областей РНКаз обеспечивается и дисульфидным мостиком, связывающим остатки Cys<sup>9</sup> и Cys<sup>103</sup>. Хотя сам факт наличия  $\beta$ -структуры в С-концевой части молекул не следует однозначно из опубликованных результатов рентгеноструктурного исследования гуанилспецифичных РНКаз грибов, тем не менее данная область прослеживается в ряде гомологичных бактериальных РНКаз [15, 19, 31, 32].

2. Исключаются из предсказания расчетные  $\beta$ -конформации в участках 15—17 РНКазы  $T_1$ , 21—22 РНКазы *P. brevicompactum* и 21—24/26

Значения средних потенциалов в конформациях  $\alpha$ -спирали ( $P_\alpha$ ),  $\beta$ -структуры ( $P_\beta$ ) и  $\beta$ -поворота ( $P_t$ ) для тетрапептидов РНКаз

Номер остатка	РНКазы T <sub>1</sub>			РНКазы C <sub>2</sub>			РНКазы <i>P. brevicompactum</i>		
	$P_\alpha$	$P_\beta$	$P_t$	$P_\alpha$	$P_\beta$	$P_t$	$P_\alpha$	$P_\beta$	$P_t$
1	—	—	—	0,85	0,93	1,31	—	—	—
2	0,81	1,10	1,19	0,81	1,10	1,19	—	—	—
4	0,69	1,15	1,21	0,69	1,15	1,21	0,88	0,99	1,09
6	0,68	0,89	1,43	—	—	—	—	—	—
7	0,68	0,89	1,43	0,76	0,89	1,28	—	—	—
8	0,71	1,07	1,33	—	—	—	—	—	—
9	0,71	1,07	1,33	—	—	—	—	—	—
10	0,73	1,04	1,30	—	—	—	—	—	—
11	0,75	0,93	1,36	—	—	—	0,76	1,04	1,24
12	0,83	0,70	1,44	—	—	—	—	—	—
13	0,90	0,93	1,20	—	—	—	—	—	—
21	1,02	0,97	1,00	0,86	0,95	1,20	1,02	0,97	1,00
27	1,02	0,63	1,18	—	—	—	—	—	—
28	—	—	—	0,97	0,86	1,16	0,82	0,84	1,35
29	—	—	—	—	—	—	0,88	0,79	1,36
30	—	—	—	—	—	—	0,94	0,92	1,24
33	—	—	—	—	—	—	0,80	1,20	1,16
34	0,69	0,78	1,49	0,82	0,84	1,22	—	—	—
35	0,72	0,96	1,39	0,85	1,02	1,12	—	—	—
36	0,67	0,91	1,41	—	—	—	—	—	—
38	0,85	0,91	1,15	0,84	0,99	1,15	0,94	0,81	1,09
40	0,88	0,99	1,16	0,94	1,09	1,01	—	—	—
42	0,68	1,18	1,35	—	—	—	—	—	—
44	0,86	0,87	1,25	0,86	0,87	1,25	0,86	0,87	1,25
52	—	—	—	0,77	1,02	1,26	0,81	1,10	1,10
53	—	—	—	0,67	0,96	1,42	—	—	—
54	0,68	1,06	1,31	0,65	1,14	1,35	0,69	1,22	1,20
62	0,83	0,89	1,25	0,83	0,89	1,25	0,93	0,88	1,15
63	0,78	0,69	1,47	0,72	0,75	1,46	0,91	0,74	1,25
64	—	—	—	0,73	0,86	1,34	—	—	—
65	—	—	—	0,71	1,04	1,27	—	—	—
66	—	—	—	1,34	1,07	1,27	—	—	—
67	—	—	—	0,69	1,07	1,30	—	—	—
68	0,83	0,77	1,30	0,62	0,96	1,45	0,71	1,04	1,27
69	—	—	—	0,59	0,78	1,56	—	—	—
70	0,62	0,76	1,16	0,57	0,70	1,55	0,67	0,70	1,48
71	0,60	0,70	1,52	—	—	—	—	—	—
72	0,83	0,77	1,30	0,78	0,72	1,32	0,83	0,72	1,29
74	0,99	0,76	1,16	0,99	0,76	1,16	0,99	0,76	1,16
81	0,88	0,76	1,35	0,84	0,71	1,51	0,92	0,63	1,48
82	—	—	—	1,05	0,58	1,30	1,13	0,50	1,28
92	0,95	0,91	1,03	0,95	0,91	1,03	0,95	0,91	1,03
94	0,83	0,77	1,30	0,83	0,77	1,30	0,83	0,77	1,30
95	0,86	0,81	1,30	0,94	0,72	1,28	0,86	0,81	1,30
96	0,67	0,82	1,53	0,73	0,70	1,50	0,67	0,82	1,53
97	0,76	0,98	1,32	0,82	0,85	1,29	0,76	0,98	1,32

РНКазы T<sub>1</sub> и C<sub>2</sub>, которые перекрываются соответствующими областями протяженных  $\alpha$ -спиралей с большими  $P_\alpha$ -конформационными потенциалами.

3. Не учитываются в качестве  $\alpha$ -спиралей сегменты полипептидной цепи, общие для С-концевой части молекул РНКаз. Для участка 73/74—79/80 РНКаз T<sub>1</sub>, C<sub>2</sub> и *P. brevicompactum* и участка 98—101 РНКазы T<sub>1</sub>, значения  $\alpha$ -конформационных потенциалов меньше соответствующих  $\beta$ -потенциалов (см. рис. 3). То же справедливо для фрагмента  $\alpha$ -спиральной структуры в участке полипептидной цепи РНКазы T<sub>1</sub> 83—87.  $\alpha$ -Спирали в РНК-азах C<sub>2</sub> и *P. brevicompactum* 80/81—87/88 исключены из-за наличия в этом сегменте молекулы дестабилизирующего  $\alpha$ -спираль кластера трех и четырех остатков дикарбоновых аминокислот. По этой же причине  $\alpha$ -спираль в РНКазе T<sub>1</sub> 15—29 ограничена в С-концевой части остатком His<sup>27</sup>. Один виток  $\alpha$ -спирали 1—4 в РНКазе *P. brevicompactum* не прогнозируется ввиду большей предпочтительности в этом районе молекулы конформации  $\beta$ -структуры.

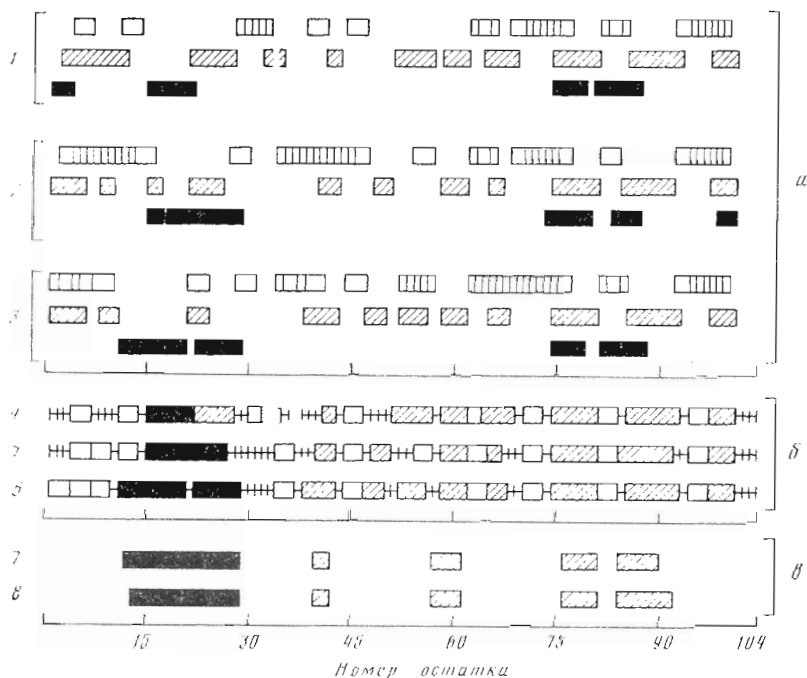


Рис. 4. Вторичные структуры гуанилспецифических РНКаз грибов. а — локализация всех выделенных расчетным методом элементов вторичных структур РНКаз *P. brevicomactum* (1),  $T_1$  (2) и  $C_2$  (3); б — конечный вариант предсказания вторичных структур РНКаз *P. brevicomactum* (4),  $T_1$  (5) и  $C_2$  (6); в — данные кристаллографических исследований РНКаз  $T_1$  (7) и  $C_2$  (8). Светлые прямоугольники — участки  $\beta$ -поворотов, заштрихованные —  $\beta$ -структуры, зачерненные —  $\alpha$ -спирали

4. Не предсказываются общие для анализируемых РНКаз конформации  $\beta$ -поворотов в позициях полипептидных цепей 27/28, 38, 63, 72, 74, 95, 96 и 97, а также в позициях РНКазы  $T_1$  — 13, 40, 42, РНКазы  $C_2$  — 21, 52, 53, 54, 64, 65, 66, 67 и 92 и в позиции 92 РНКазы *P. brevicomactum* — соответствующие аминокислотные остатки молекул белков являются компонентами других типов вторичных структур. Вместе с тем для всех гуанилспецифических РНКаз прогнозируются общие  $\beta$ -повороты полипептидных цепей в позициях 4, 44, 62, 70, 81 и 94, а также в позиции 7 — для РНКаз  $T_1$  и  $C_2$ , 11 — для РНКаз  $T_1$  и *P. brevicomactum*, 30 — для РНКаз *P. brevicomactum* и 34 — для РНКаз  $T_1$  и  $C_2$ . В этой связи исключаются дублирующие их  $\beta$ -повороты в соответствующих местах индивидуальных молекул: 2 — для РНКаз  $T_1$  и  $C_2$ , 6, 8—10 и 12 — для РНКазы  $T_1$ , 29 — для РНКазы *P. brevicomactum*, 35 — для РНКаз  $T_1$  и  $C_2$  и 36 — для РНКазы  $T_1$ , 68 — для всех трех белков, 69 — для РНКазы  $C_2$ , 71 — для РНКазы  $T_1$ , 82 — для РНКаз  $C_2$  и *P. brevicomactum* и, наконец, 92 — для РНКазы  $T_1$  (см. таблицу, рис. 4).

Таким образом, суммируя приведенные выше расчеты и соображения, окончательный вариант предсказания вторичных структур гуанилспецифических РНКаз грибов на основании анализа их аминокислотных последовательностей может быть проиллюстрирован схемами, представленными на рис. 4б.

В нижней части рис. 4 показаны области полипептидных цепей, образующие  $\alpha$ -спирали и  $\beta$ -структуры в кристаллах РНКазы  $C_2$  и комплекса РНКазы  $T_1$  с гуанозин-2'-фосфатом [17, 20].

Отмечая общее качественное подобие в распределении и локализации элементов вторичных структур гуанилспецифических РНКаз, полученных расчетным методом, следует тем не менее выделить некоторое несовпадение в структурах, предсказываемых для N-концевой половины их молекул. В отличие от расчетных конформаций C-концевой половины молекул РНКаз (58—104), которые практически идентичны для трех сравниваемых



мых полипептидных цепей, за исключением размера небольшого сегмента — структуры 66—69, включающего частично одну из двух ранее отмеченных областей максимальной варибельности ферментов 66—71, совпадающими структурами в N-концевой половине молекулы РНКаз (1—57) являются лишь два  $\beta$ -поворота в позициях 4—7 и 44—47, участок  $\alpha$ -спирали 15—22 и  $\beta$ -структуры в районе 41—43. Вторая область максимальной варибельности РНКаз, 27—37, как видно из рис. 4, представляет собой основную область нерегулярной структуры ферментов. В N-концевой части РНКаз локализованы и все другие, общие для трех молекул замещения в аминокислотных остатках; существенным при этом является сохранение конформации полипептидной цепи в сегментах, включающих в себя остатки Tyr<sup>42</sup> и Tyr<sup>43</sup>, Asn<sup>44</sup>, возможно, His<sup>40</sup>, а также при замещении в РНКазх остатков в положении 43 [20].

Сравнение элементов вторичных структур РНКаз, полученных расчетным методом, с данными рентгеноструктурного анализа РНКазы С<sub>2</sub> и комплекса РНКазы Т<sub>1</sub> с гуанозин-2'-фосфатом (см. рис. 4б, в) показывает, что все четыре области вытянутой  $\beta$ -конформации, в том числе короткий трехчленный сегмент полипептидной цепи в районе 40—42, выделенные в кристаллографических исследованиях, прогнозируются теоретически с высокой степенью точности. Вместе с тем для трех сравниваемых белков имеются различия в предсказании границ единственного  $\alpha$ -спирального участка.

В РНКазе *P. brevicompactum* короткий участок  $\alpha$ -спирали 15—22 граничит с участком  $\beta$ -структуры 22—28, тогда как в других гомологичных РНКазх более вероятной (см. выше и рис. 3) в этой области представляется  $\alpha$ -спиральная конформация.

В РНКазе *P. brevicompactum* на основании расчетных данных предполагается также наличие  $\beta$ -структур в районе 51—57, 64—69 и 98—101. Аналогичные и частично совпадающие элементы структур прогнозируются и в соответствующих сегментах полипептидных цепей РНКаз Т<sub>1</sub> и С<sub>2</sub> (рис. 4), хотя, согласно результатам рентгеноструктурного анализа, конформация белков в этих участках нерегулярна.

Выявленные несоответствия экспериментальных и расчетных данных могут быть объяснены несколькими причинами. Одна из них, по-видимому наиболее существенная, заложена, очевидно, в несовершенстве самой методологии расчета, учитывающей в основном локальные взаимодействия остатков аминокислот, расположенных по длине аминокислотной цепи в непосредственной близости друг от друга, и не принимающей во внимание важную роль дальних взаимодействий в формировании глобулярной структуры белков [28—30]. Кроме того, метод основан на статистически усредненных параметрах соответствующих конформаций, найденных экспериментально для ограниченного количества белков, и не учитывает возможные отклонения геометрических параметров структур от стандартных значений.

В этой связи представляется интересным, что, согласно расчетным данным,  $\alpha$ -спираль в РНКазе С<sub>2</sub>, предсказываемая сравнительно с рентгеноструктурными данными с достаточной степенью точности, претерпевает некоторое искажение в средней своей части (рис. 3). Это искажение характерно не только для данного белка, но может быть прослежено и в случае РНКазы Т<sub>1</sub> (см. рис. 3) и особенно для РНКазы *P. brevicompactum*, в молекуле которой полипептидная цепь после излома  $\alpha$ -спирали принимает конформацию вытянутой  $\beta$ -структуры.

Другая категория возможных причин различий экспериментально и расчетно выявленных вторичных структур РНКаз может быть обусловлена также различиями между структурами белка в комплексе РНКазы Т<sub>1</sub> с ингибитором и без него. Возможны также различия, связанные с неадекватностью отнесения авторами рентгеноструктурных исследований к  $\beta$ -структурным областям только фрагментов полипептидных цепей с двугранными углами, обеспечивающими образование внутримолекулярных водородных связей, и определением границ соответствующих структур.

Очевидно, что внеклеточные гуанилспецифичные РНКазы аскомицетов

являются близкородственными ферментами, существенно подобными как в функциональном, так и в структурном отношении. Вместе с тем прослеживается структурная гомология с другими РНКазами грибов с той же или отличающейся специфичностью, а также внеклеточными бактериальными РНКазами [16, 33], хотя в последнем случае филогенетическая дистанция между двумя семействами РНКаз проявляется в большей вариабельности их структур.

Гомология первичных и вторичных структур РНКаз микроорганизмов (и в том числе наблюдаемое подобие их пространственной упаковки [15]) — веский довод в пользу наиболее популярной среди исследователей версии их дивергентного происхождения [15, 16, 33]. Тем не менее не может быть полностью исключена возможность конвергентной эволюции РНКаз, в ходе которой происходит отбор наиболее стабильных и термодинамически выгодных функциональных состояний белка. Не является исчерпанным также вопрос о структурной и функциональной взаимосвязи группы РНКаз микроорганизмов с семейством панкреатических РНКаз [16, 34].

В любом случае представляется очевидным, что, несмотря на «вырожденность» трехмерной структуры РНКаз, тождественность функционирования ферментов обеспечивается совершенно определенной конформацией полипептидной цепи и точной пространственной локализацией аминокислотных остатков, существенных для каталитического процесса.

### Расчетная часть

Предсказание вторичных структур РНКаз по данным их аминокислотной последовательности проводили в соответствии с модифицированным Дурфтоном и Хайдером [28] методом расчета Чоу и Фасмана [29, 30]. Основные изменения в методологии предсказания:

1) замена тенденциозно-произвольной и неколичественной системы поиска центров формирования элементов упорядоченной структуры и их расширения по Чоу и Фасману числовым анализом конформационных потенциалов аминокислотных остатков;

2) вместо средних арифметических значений конформационных потенциалов, определяющих вероятность какой-либо конформации, но сглаживающих предпочтительность пептидных сегментов соответствующему типу структуры при низких величинах конформационных потенциалов ( $\sim 1,03-1,05$ ), использован принцип перемножения потенциалов, расширяющий степень градации и соотношения аминокислот по принципу их индифферентности или соответствия двум основным категориям вторичных структур белков —  $\alpha$ -спиралям и  $\beta$ -структурам;

3) расчет элементов вторичных структур базировался на анализе конформационных потенциалов последовательно перекрывающихся друг с другом по ходу полипептидной цепи тетрапептидных сегментов как минимальных звеньев цепи, способных адекватно отражать искомую конфигурацию.

Алгоритм предсказания вторичной структуры РНКаз включал в себя расчет конформационных потенциалов  $\alpha$ -спиралей и  $\beta$ -структур тетрапептидов по данным соответствующих значений для аминокислот [28]

$$NP_{\alpha} = P_{\alpha, i} \times P_{\alpha, i+1} \times P_{\alpha, i+2} \times P_{\alpha, i+3},$$

$$NP_{\beta} = P_{\beta, i} \times P_{\beta, i+1} \times P_{\beta, i+2} \times P_{\beta, i+3},$$

а также вероятности  $\beta$ -поворота в  $i$ -м аминокислотном остатке:

$$p_i = f_i \times f_{i+1} \times f_{i+2} \times f_{i+3},$$

где  $f_i-f_{i+3}$  — частота нахождения в белках соответствующего аминокислотного остатка в 1, 2, 3 и 4-м положениях  $\beta$ -поворота [28]. Любой кластер тетрапептидов с  $NP_{\alpha} > 1$  и  $NP_{\beta} > 1$  рассматривался в качестве кандидата соответствующей конформации.

Следующая стадия расчета сводилась к определению протяженности и границ  $\alpha$ -спиралей и  $\beta$ -структур на основании анализа величин  $NP_{\alpha}$  и

$NP_{\beta}$  перекрывающихся тетрапептидов. Наличие средних значений  $NP_{\alpha}(NP_{\beta}) > 1$  для перекрывающихся сегментов свидетельствовало о включении соответствующих аминокислотных остатков в конкретный тип конформации. Элементы вторичной структуры ограничивались аминокислотными остатками, для которых средняя величина  $NP_{\alpha}(NP_{\beta})$  перекрывающихся тетрапептидов была меньше 1. При этом минимальное число аминокислотных остатков, необходимое для образования  $\beta$ -структур и  $\alpha$ -спиралей, составляло соответственно 3 и 4.

В качестве предельной нижней величины вероятности  $\beta$ -поворота в РНКазах, достаточной для его предсказания, нами выбрано значение  $p_i = 1,0 \cdot 10^{-4}$ , приблизительно вдвое превышающее среднюю вероятность любого тетрапептида в белках находиться в этой конформации —  $0,55 \cdot 10^{-4}$  [28]. Необходимым условием предсказания  $\beta$ -поворотов в тетрапептидах было также достаточное значение среднего конформационного потенциала ( $P_i > 1,0$ ), причем  $\beta$ -поворот прогнозировался только при соблюдении соотношения  $P_{\alpha} < P_i > P_{\beta}$ . Значения  $p_i$  для аминокислотных остатков во всех четырех позициях  $\beta$ -поворота взяты из работы [28]. В ряде случаев, несмотря на соблюдение необходимых и достаточных условий предсказания  $\beta$ -поворотов, последние не прогнозировались, если выделенные фрагменты полипептидных цепей оказывались занятыми элементами других структур или более предпочтительными считали соответствующие конформации в соседних, частично перекрывающихся сегментах.

Решение вопроса о предпочтительности той или иной конформации при перекрывании  $\alpha$ -спиралей и  $\beta$ -структур основывалось на анализе относительных величин их рассчитанных конформационных потенциалов с учетом определенных запретов на локализацию некоторых аминокислотных остатков в данном элементе вторичной структуры [28, 30].

#### ЛИТЕРАТУРА

1. Levitt M. J. Mol. Biol., 1976, v. 104, № 1, p. 59–107.
2. Ptitsyn O. B. FEBS Lett., 1981, v. 131, № 2, p. 197–202.
3. Лим В. И., Мазанов А. Л., Ефимов А. В. Молекулярн. биология, 1978, т. 12, № 1, с. 201–213.
4. Ptitsyn O. B., Finkelstein A. V. In: Protein folding/Ed. Jaenicke R. Amsterdam — New York: Elsevier/North-Holland, 1980, p. 101–115.
5. Levitt M., Chothia C. Nature, 1976, v. 261, № 5561, p. 552–558.
6. Wetlaufer D. B. Proc. Nat. Acad. Sci. USA, 1973, v. 70, № 3, p. 697–701.
7. Kirscher K., Bisswanger H. Annu. Rev. Biochem., 1976, v. 45, p. 143–166.
8. Schulz G. E., Schirmer R. H. Nature, 1974, v. 250, № 5462, p. 142–144.
9. Hartley B. S. Symp. Soc. Gen. Microbiol., 1974, v. 24, p. 151–182.
10. Love W. E., Klock P. A., Lattman E. E., Padlan E. A., Ward K. B., Jr. Cold Spring Harbor Symp. Quant. Biol., 1974, v. 36, p. 349–356.
11. Perutz M. F., Kendrew J. C., Watson H. C. J. Mol. Biol., 1965, v. 13, № 3, p. 669–678.
12. Almasy R. J., Dickerson R. E. Proc. Nat. Acad. Sci. USA, 1978, v. 75, № 6, p. 2674–2678.
13. Shinoda T., Takahashi N., Takayasu T., Okuyama T., Shimizu A. Proc. Nat. Acad. Sci. USA, 1981, v. 78, № 2, p. 785–789.
14. Rossmann M. G., Liljas A., Branden G.-I., Banaszak L. J. In: The Enzyme/Ed. Boyer P. D. N. Y.: Acad. Press, 1975, v. 11, p. 61–102.
15. Hill G., Dodson G., Hienemann U., Saenger W., Mitsuo Y., Nakamura K., Borisov V., Tischenko G., Polyakov K., Pavlovsky A. Trends in Biochem. Sci., 1983, v. 8, № 10, p. 364–369.
16. Безбородова С. И., Безбородов А. М. В кн.: Биосинтез микроорганизмами нуклеаз и протеаз. М.: Наука, 1979, с. 92–145.
17. Поляков К. М., Вагин А. А., Тищенко Г. Н., Безбородова С. И., Вайнштейн Б. К. Докл. АН СССР, 1983, т. 273, № 6, с. 1383–1386.
18. Безбородова С. И., Ходова О. М., Степанов В. М. Биоорган. химия, 1983, т. 9, № 8, с. 1136–1138.
19. Manguen Y., Hartley R. W., Dodson E. J., Dodson G. G., Bricogne G., Chothia C., Jack A. Nature, 1982, v. 297, № 5862, p. 162–164.
20. Hienemann U., Saenger W. Nature, 1982, v. 299, № 5878, p. 27–31.
21. Hartley R. W., Barker E. A. Nature New Biol., 1972, v. 235, № 53, p. 15–16.
22. Takahashi K. J. Biol. Chem., 1965, v. 240, № 10, p. 4117–4119.
23. Куликов В. А., Шляпников С. В., Яковлев Г. И. Биоорган. химия, 1985, т. 11, № 3, с. 334–346.
24. Cid H., Bunster M., Arriagada E., Campos M. FEBS Lett., 1982, v. 150, № 1, p. 247–254.
25. Lesk A. M., Chothia C. J. Mol. Biol., 1980, v. 136, № 3, p. 225–270.
26. Лим В. И., Птицын О. В. Молекулярн. биология, 1970, т. 4, № 3, с. 372–383.
27. Lenstra J. A., Hofsteenge J., Veintema J. J. J. Mol. Biol., 1977, v. 109, № 2, p. 185–193.

28. Dujton H. J., Hider R. C. J. *Mol. Biol.*, 1977, v. 115, № 2, p. 177-193.
29. Chou P. Y., Fasman G. D. *Adv. in Enzymol.*, 1978, v. 47, p. 45-148.
30. Chou P. Y., Fasman G. D. *Biochemistry*, 1974, v. 13, № 2, p. 222-245.
31. Pavlovsky A. G., Vagin A. A., Vainstein B. K., Cherpurnova N. K., Karpeisky M. Ya. *FEBS Lett.*, 1983, v. 162, № 1, p. 167-170.
32. Nakamura K., Iwahashi K., Yamamoto Y., Itaka Y., Yoshida N., Mitsui Y. *Nature*, 1982, v. 299, № 5883, p. 564-566.
33. Hartley R. W. J. *Mol. Evol.*, 1980, v. 15, № 4, p. 355-358.
34. Ycás M. J. *Mol. Evol.*, 1976, v. 7, № 3, p. 215-244.

Поступила в редакцию  
4.IX.1984

COMPARATIVE ANALYSIS OF THE AMINO ACID SEQUENCE AND  
CALCULATION OF THE SECONDARY STRUCTURES FOR GUANYL-SPECIFIC  
RIBONUCLEASES FROM *PENICILLIUM BREVICOMPACTUM*,  
*ASPERGILLUS ORYZAE* (T<sub>1</sub>) AND *ASPERGILLUS CLAVATUS* (C<sub>2</sub>)

SHLYAPNIKOV S. V., KULIKOV V. A., YAKOVLEV G. I.

*Institute of Molecular Biology, Academy of Sciences  
of the USSR, Moscow*

The secondary structures of extracellular guanyl-specific RNases from fungi *Penicillium brevicompactum*, *Aspergillus oryzae* (T<sub>1</sub>) and *Aspergillus clavatus* (C<sub>2</sub>) have been calculated using a modified Chou - Fasman procedure. A comparison of the primary and secondary structures of the RNases revealed a high degree of homology (64-68%), and a similar distribution of predicted structural regions ( $\alpha$ -helices,  $\beta$ -structure and  $\beta$ -turns). Conspicuous conservatism was observed in the distribution of hydrophobic amino acid residues. The predicted secondary structure patterns were discussed in light of the RNase X-ray structures.