

ОСНОВЫ СПЕЦИФИЧНОСТИ ВЗАИМОДЕЙСТВИЯ БЕЛКОВ И НУКЛЕИНОВЫХ КИСЛОТ

Д.В. Штадлер, Н.Л. Миронова, М.А. Зенкова, В.В. Власов

Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск, Россия,
e-mail: marzen@niboch.nsc.ru

В данном обзоре рассмотрены типы контактов РНК с аминокислотами и белками разной функциональной направленности, обеспечивающие специфическое связывание с отдельными гетероциклическими основаниями и последовательностями нуклеотидов. Рассмотрено применение «кода узнавания» для создания искусственных рибонуклеаз со специфичностью РНКазы Т1 на основе олигонуклеотид-пептидных конъюгатов.

Введение

Термин «искусственные», или «химические», рибонуклеазы был введен более 10 лет назад для описания низкомолекулярных органических соединений, способных в физиологических (или близких к ним) условиях катализировать расщепление фосфодиэфирных связей в РНК специфично к последовательности или случайным образом (Morrow, 1994). Первые химические рибонуклеазы обладали низкой активностью, и заметное расщепление РНК наблюдалось лишь при длительной инкубации РНК-субстрата в присутствии искусственного фермента.

Позднее был сформулирован общий принцип строения искусственных рибонуклеаз: объединение в одной молекуле структур, обладающих сродством к определенным нуклеотидным последовательностям или элементам структуры РНК, с молекулами, способными катализировать расщепление фосфодиэфирных связей. В качестве РНК-связывающих фрагментов (или доменов, по аналогии с природными катализаторами-ферментами) были использованы поликатионы различного строения, интеркаляторы, пептиды и антисмысловые олигонуклеотиды. При конструировании каталитических доменов известных на настоящий момент искусственных рибонуклеаз использовались или металлокомп-

лексы, или органические структуры, содержащие функциональные группы – имидазольные (Зенкова и др., 2000, Zenkova *et al.*, 2001), карбоксильные (Endo *et al.*, 1996), гуанидиниевые и аминокислотные группы в разных комбинациях (Oost *et al.*, 1997), характерные для каталитических центров природных рибонуклеаз – РНКазы А и РНКазы Т1.

Основная проблема, с которой сталкиваются при конструировании химических рибонуклеаз, состоит в создании конструкций, проявляющих различную специфичность расщепления РНК, подобно природным рибонуклеазам. В природных РНКазах проблема специфичности решается организацией субстрат-связывающего центра, в котором специфическое взаимодействие аминокислотных остатков с РНК обеспечивает связывание гетероциклического основания и оптимальную конформацию фосфодиэфирной связи, которая подвергается расщеплению. Большинство исследованных искусственных рибонуклеаз расщепляют РНК преимущественно по связям в мотивах Руг-А, тогда как другая специфичность расщепления РНК химическими рибонуклеазами была показана лишь в единичных работах (Endo *et al.*, 1997).

При конструировании искусственных рибонуклеаз учитывают особенности строения каталитического центра природных ферментов

и вводят в структуру соединения функциональные группы, например, такие, как имидазол, аминогруппы, которые в природных рибонуклеазах участвуют в реакции трансэтерификации. Моделирование специфического связывания, происходящего в природном ферменте, является чрезвычайно непростой задачей в связи со сложной пространственной организацией субстрат-связывающего центра и многоочечным контактом фермент-субстрат. Тем не менее неоднократно предпринимались попытки такого моделирования: в структуру искусственных рибонуклеаз вводили ароматические аминокислотные остатки, например, фенилаланин (Tung *et al.*, 1992), способствующие стабилизации гетероциклического основания в нужной конформации путем стэкинга. Однако это не влияло на специфичность расщепления РНК: расщеплению подвергались связи в Руг-А-мотивах.

Теоретически узнавание нуклеотидов данного типа А, С, G или U в последовательности РНК может обеспечиваться высокоселективными взаимодействиями между определенными функциональными группами, аминокислотами или пептидными мотивами и определенным видом гетероциклических оснований РНК, а также с сахарофосфатным остовом. Для создания искусственных рибонуклеаз с разной специфичностью необходимо понимание правил универсального кода узнавания – некой формулы соответствия определенного набора функциональных групп и их пространственного расположения в искусственной рибонуклеазе от последовательности оснований в РНК. Такие формулы соответствия были найдены природой и реализованы в ряде белков и ферментов.

В данном обзоре будут рассмотрены типы контактов РНК с аминокислотами и белками разной функциональной направленности, обеспечивающих специфическое связывание с отдельными гетероциклическими основаниями и последовательностями нуклеотидов. Будет рассмотрено применение «кода узнавания» для создания искусственных рибонуклеаз со специфичностью РНКазы T1 на основе олигонуклеотид-пептидных конъюгатов.

1. Правила узнавания: «последовательность РНК– последовательность белка». **Специфическое взаимодействие аминокислот с гетероциклическими основаниями в РНК**

Вторичная структура РНК представляет собой комбинацию коротких двуспиральных участков, соединенных петлями различного размера. Двойная спираль (А-форма) РНК имеет глубокую и узкую большую бороздку, которая обычно не доступна для взаимодействия с белками (Drapez, 1999). Однако вторичная структура РНК в основном состоит из коротких двуцепочечных участков и поворотов между ними, что делает большую бороздку доступной для взаимодействия с концев спирали. Кроме того, структура РНК содержит значительное число мисматчей, боковых и внутренних петель различной длины, которые дестабилизируют двойную спираль, и способствует взаимодействию белков с РНК со стороны большой бороздки. Напротив, широкая и плоская поверхность малой бороздки РНК доступна для взаимодействия с белками (Hermann, Patel, 1999).

Ввиду больших различий между вторичными структурами РНК и ДНК многие правила узнавания белками ДНК не могут быть применены к РНК (Pabo, Nekludova, 2000; Luscombe *et al.*, 2001). Однако один из важных определяющих элементов специфичности, как в ДНК, так и в РНК-белковых комплексах – это водородные связи, образованные между полярными группами в белках и основаниями в нуклеиновых кислотах.

РНК-белковые взаимодействия находятся на ранней стадии изучения. Самое очевидное отличие между РНК и ДНК комплексами – это участие 2' ОН группы рибозы в образовании более чем четверти всех водородных связей, обнаруженных в РНК-белковых комплексах (Treger, Westhof, 2001). В комплексах РНК-аминокислота водородные связи играют менее важную роль, чем в комплексах ДНК-аминокислота (Allers, Shamo, 2001; Jones *et al.*, 2001), так как в РНК основания часто находятся вне спирали, выведены из стэкинг-взаимодействий и участвуют в Ван-дер-ваальсовых взаимодействиях с белком (Cheng *et al.*, 2003). Тем не менее для

специфического узнавания РНК водородные связи так же важны, как и для узнавания ДНК. В работе (Cheng *et al.*, 2003) были проанализированы возможные водородные связи между аминокислотами и основаниями и парами оснований в РНК. Были рассмотрены взаимодействия с участием неспаренных оснований в боковых и апикальных петлях шпилек; с неканоническими парами оснований, которые способны образовывать множественные водородные связи и, таким образом, могут являться специфическими элементами узнавания, и проведены расчеты возможных взаимодействий, между аминокислотами и неспаренными основаниями (А, С, G, U, A⁺ и C⁺), или парами оснований в РНК в планарных конформациях.

1.1. Водородные связи между аминокислотами и основаниями в РНК

Из 32 возможных контактов, в которых участвуют 2 или более водородных связей, между неспаренными основаниями в РНК и аминокислотами – в 12 участвуют остатки Asn(Gln) и в 8 – Ser (Thr/Тур). Боковые радикалы этих аминокислот способны образовывать связи со всеми основаниями, кроме протонированного аденина А⁺. Боковые радикалы Asn(Gln), содержащие две акцепторные группы, способны взаимодействовать с основаниями А, G, А⁺ и C⁺, но не с уридином U. Для аргинина, содержащего пять донорных групп, было обнаружено формирование 4 контактов с С и G.

В табл. 1 представлены данные по специфическому взаимодействию оснований в РНК и ДНК с аминокислотами, проанализированные на основе данных о таких взаимодействиях в ДНК- и РНК-белковых комплексах. Из 32 типов контактов, предсказанных теоретически, в составе комплексов было найдено 17 типов контактов. 12 типов контактов в ДНК-белковых комплексах, включающих 5 типов контактов со стороны большой бороздки, и два типа контактов, протекающих со стороны малой бороздки: Arg•G и Asn(Gln) А. Было обнаружено 6 типов контактов ДНК с боковыми радикалами аминокислот, в которых наблюдалось образование двух водородных связей с «наружной» стороны

Уотсон-Криковской пары. Например, образование водородной связи между Ser и U наблюдали в составе комплекса транскрипции (Najmudin *et al.*, 2000), Lys-C – в комплексах одноцепочечной ДНК в составе нуклеокапсида (Morellet *et al.*, 1998); Arg-C – в комплексе эксцизионной репарации (Norman *et al.*, 2001); Asn(Gln)-А – в комплексах РНКазы В/ДНК и ДНК/метилтрансфераза (Ko *et al.*, 1996; Goedecke *et al.*, 2001); Arg-G – в двух белках, участвующих в связывании теломерных участков хромосомы (Horvath *et al.*, 1998; Ding *et al.*, 1999) и Asp (или Glu) - C⁺ – в комплексах ДНК/метиلاза HhaI или HaeIII (Klimasauskas *et al.*, 1994; Reinisch *et al.*, 1995) (табл. 1).

Данные по РНК-белковым комплексам относительно малочисленны, однако удалось провести сравнение взаимодействий между РНК и боковыми радикалами аминокислот в белках. Было найдено 16 типов контактов между аминокислотами и основаниями в РНК, в 8 из которых аминокислоты образуют по две водородные связи с краевыми атомами оснований (табл. 1). Некоторые из контактов, такие, как Glu-U(T), Ser-C и Ser-C⁺, были найдены только в РНК-белковых комплексах, тогда как взаимодействия Ser-U(T), Lys-C, Arg-C, Asn(Gln)-А и Asp-G были найдены как в РНК, так и в ДНК-белковых комплексах. При узнавании основания РНК, находящегося в составе Уотсон-Криковской пары, наиболее часто встречались взаимодействия Arg-G, а взаимодействия Asn(Gln)-А встречались реже всего.

1.2. Взаимодействие аминокислот с обоими основаниями в РНК в составе комплементарной или неканонической пары

Наибольшая специфичность узнавания уникальных участков РНК достигается в случае, если аминокислота образует водородные связи с обоими основаниями неканонической пары (Walberer *et al.*, 2003). Специфическое опознавание неправильной пары G•А было обнаружено в комплексе Rev пептид/RRE РНК (Tan *et al.*, 1993). Теоретически в РНК может быть 53 возможных пары оснований (Cheng *et al.*, 2003). Так же, как и в случае не спаренных оснований, наиболее частым типом взаимодействий

Таблица 1

Водородные связи между аминокислотой и основанием,
обнаруженные в ДНК- и РНК-белковых комплексах

Взаимодействие		Тип взаимодействия ¹	Найдено		Комплекс ²
Аминокислота	Основание		ДНК	РНК	
Ser Thr Tyr	T U	W-C	1	1	ДНК: обратная транскриптаза (1d0e) Ser67; РНК: Sxl (1b7f) Tyr164
Asn Gln	U	W-C	0	6	РНК: AspRS (1asy, 1asz) Gln138, (1c0a) Gln46, (1efw) Gln47, (1il2) Gln1046; GlnRS (1euq) Gln517
Lys	C	W-C	1	1	ДНК: нуклеокапсид (1bj6) Lys34; РНК: нуклеолин (1fje) Lys94
Arg	C	W-C	1	4	ДНК: оксоG гликозилаза (1fn7) Arg204; РНК: S15/S16/S18 (1ekc, 1g59) Arg74; AspRS (1il2) Arg225; GluRs (1g59) Arg358
Ser Thr Lys	C	W-C	0	1	РНК: AspRS (1asz) Ser329
Asn Gln	A	M-G	83	1	РНК: 50S рибосомы (1jj2) Asn44
Ser Thr Tyr	A	M-G	11	7	РНК: U1A (1aud) Tyr12, (1dz5) Ser45, Thr88, Tyr12; MS2 coat (5msf, 6msf, 7msf) Thr45
Asn Gln	A	W-C	3	2	ДНК: РНКазы В (1rbj) Gln69, Asn71; метилтрансфераза (1g38) Asn105 РНК: U4 (1e7k) Ser96; рибозим (1hp6) Ser91
Ser Thr Tyr	A	M-G	0	1	РНК: U2B''/A' (1a9n) Ser91
Asn Gln	G	m-G	7	3	ДНК: теломера BP (1otc) Gln135
Lys	G	MG	20	6	РНК: L30 (1ck8, 1cn9) Lys28; 50 S (1jj2) Lys35; PpoRS (1h4g, 1h4s) Lys369
Arg	G	M-G	164	19	ДНК: теломера BP (1otc) Arg274; РНК: AspRS (1c0a, 1il2) Arg222; Rev (1ull) Arg6, (484d) Arg41; нуклеолин (1fje) Arg49
Asp Glu	G	W-C	10	15	ДНК: теломера BP (1otc) Asp225, Glu45; (1jb7) Asp223, Asp225, Glu45, (1k8g, 1kix) Asp223, Asp25; UP1 (2up1) Asp42; РНК: TRAP (1c9s, 1gtf, 1gtn) Asp39, Glu36; AspRS (1il2) Glu93; ThrRS (1gf6, 1kog) Glu600; PpoRS (1h4g, 1h4s) Asp354, (1h4s) Glu340; 50s (1jj2) Asp105, Glu71
Asp Glu	C ⁺	W-C	3	0	ДНК: HaeIII (1dct) Glu109; HhaI (1mht, 4mht) Glu119
Ser Thr Tyr	C ⁺	W-C	0	2	РНК: U1A (1aud, 1dz5) Tyr12

¹ сокращение W-C соответствует паре Уотсона-Крика; M-G – большой бороздке; m-G – малой бороздке;

² в скобках приведен номер в соответствии с PDB (Protein Data Bank) аналогично ДНК, встречали контакты Arg-G, а взаимодействие Asn(Gln)-A встречали реже всего.

было взаимодействие оснований с Asn(Gln) (77 случаев). При образовании контактов с обоими основаниями в паре наиболее распространенными были контакты с Arg (64 случая), который практически не участвовал во взаимодействиях с неспаренными основаниями. Было обнаружено три типа возможных контактов Asn(Gln) с парами пурин-пиримидин: два с G•C и одно с C•A⁺; одно с парой C•C⁺ и ни одного с парами пурин-пурин. Из 186 возможных контактов 9 происходят с образованием трех водородных связей между основанием и боковыми радикалами Asn(Gln) или Arg. 4 взаимодействия вовлекали контакты между Arg и воббл-парой G•U или обратными воббл-парами; 5 взаимодействий включали контакты между Asn(Gln) и четырьмя неканоническими парами оснований – двумя G•G и двумя G•C⁺. Эти пары оснований наиболее часто участвуют во взаимодей-

ствиях с аминокислотами в РНК. Кроме того, эти пары могут участвовать во взаимодействии с еще одним основанием, причем третье основание выполняет специфическое узнавание пары, так же, как и аминокислота. Одновременное связывание аминокислотой обоих оснований в паре часто встречается как элемент специфического распознавания в РНК-белковых комплексах. Так, например, в комплексе белка сплайсомеры U2B''-U2A' с U2 шпилькой snРНК Lys20 специфически узнает U•U пару, находящуюся в петле (рис. 1а). При образовании комплекса Rev белок (HIV) связывается с RRE шпилькой РНК посредством образования водородных связей между Asn40 и обоими основаниями пары G•A (рис. 1б). Положение Asn40 в двух комплексах Rev пептид-РНК было определено с помощью ЯМР, однако эти комплексы отличались по ориентации водородных связей (рис. 1в). Пока

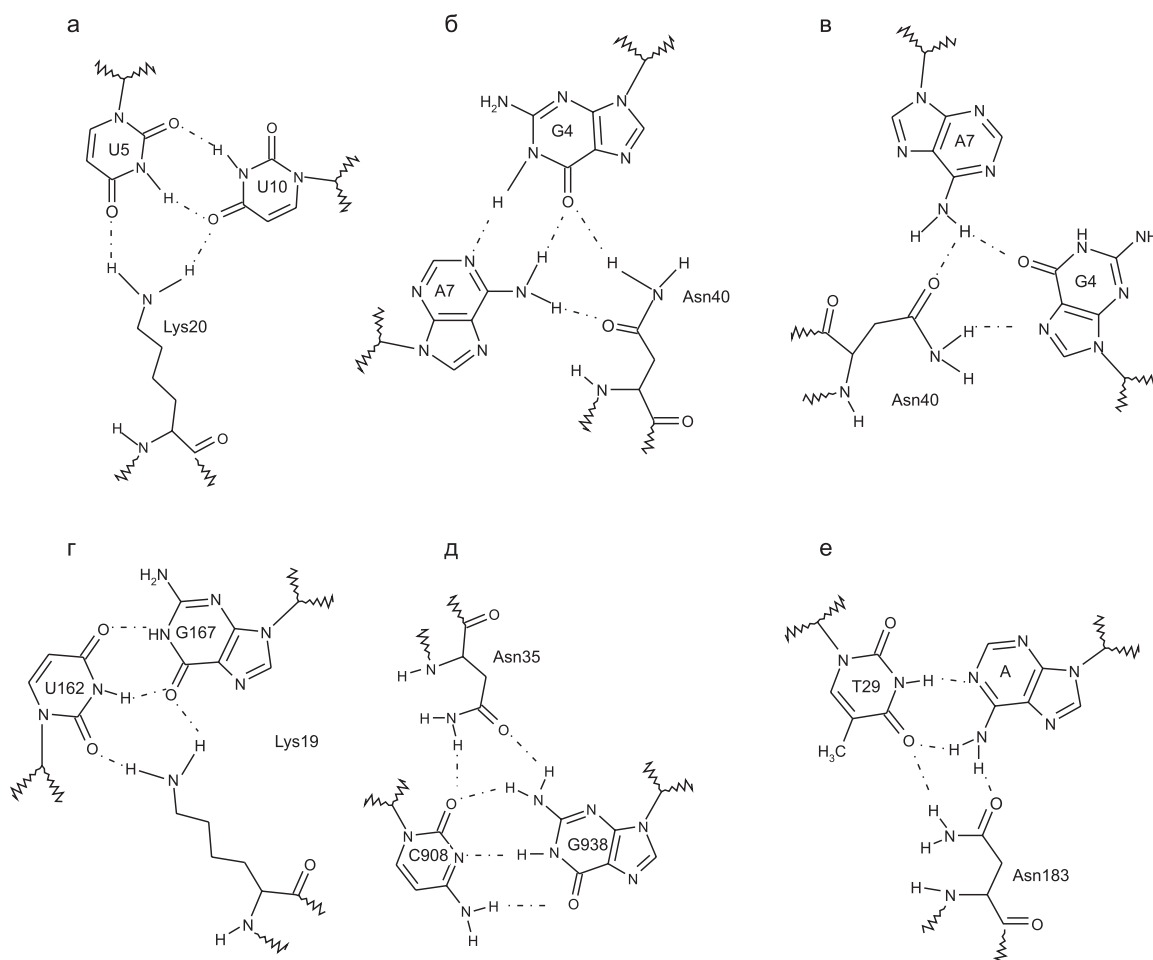


Рис. 1. Опознавание аминокислот пар оснований в РНК- и ДНК- белковых комплексах. а – Lys20/РНК, б – Asn40/РНК, в – Asn40/РНК, г – Lys19/РНК, д – Asn35/ДНК, е – Asn183/ДНК.

не ясно, связаны ли эти различия с разными контекстами РНК, в которых располагаются G•A пары. RRE-связывающий пептид из комбинаторной библиотеки, обладающий наибольшей аффинностью к субстрату, связывал G•A пару путем образования водородной связи боковым радикалом Glu вместо Asn, расположенным в окружении остатков аргинина (Ye *et al.*, 1996).

В составе РНК-белковых комплексов обнаружено четыре возможных типа контактов аминокислот с классическими Уотсон-Криковскими парами оснований. Так, боковой радикал Asn (Gln) взаимодействует как со стороны большой, так и со стороны малой бороздки с G•C парой, и со стороны большой бороздки с A•U парой. Два из этих взаимодействий наблюдались в составе комплексов Asn-A•T взаимодействие через большую бороздку в комплексе с-Myb-ДНК (Ogata *et al.*, 1994); и Asn-G•C, связывающийся через малую бороздку в случае комплексов Endo IV – ДНК и Gln-РНК синтетаза – тРНК (Ogata *et al.*, 1994, Arnez, Steitz, 1996) (рис. 1д). В комплексе с-Myb-ДНК Asn183 образует водородные связи с обоими основаниями A•T пары (рис. 1е). В кристаллической структуре комплекса Endo IV – ДНК Asn183 взаимодействует с G•C парой посредством образования прямых водородных связей между боковым радикалом и основаниями. Однако Endo IV является эндонуклеазой эксцизионной репарации и опознает не содержащие оснований нуклеотиды, поэтому роль специфического взаимодействия с основаниями не совсем ясна.

Для воббл-пары G•U наиболее часто встречается взаимодействие с аминокислотами Arg, Lys, Asn(Gln) и Ser(Thr/Tyr). Arg и Lys взаимодействуют с G•U парой только со стороны большой бороздки, тогда как Ser (Thr/Tyr) – со стороны малой бороздки, Asn(Gln) взаимодействует с G•U со стороны обеих бороздок. Для 7 из 11 Arg-G•U водородных связей необходима непланарная ориентация гуанидиниевой группы относительно пары оснований. Взаимодействие между Lys28 и G•U парой наблюдали в его собственной тРНК (Mao *et al.*, 1999). Другим примером такого взаимодействия является взаимодействие между Lys19 и G•U парой в комплексе SRP 19-7SL/РНК (рис. 1г) (Hainzl *et al.*, 2002).

2. Правила узнавания «последовательность РНК – структура белка»: основы специфичности белков с разной функциональной направленностью

2.1. Основы специфичности узнавания природных рибонуклеаз

Рибонуклеазы являются природными катализаторами расщепления фосфодиэфирных связей в РНК. Каталитические функции РНКаз, важнейших ферментов метаболизма РНК, заключаются в расщеплении мРНК, превращении предшественников РНК в зрелые формы, продукции малых регуляторных РНК, деградации определенных типов РНК.

В этой части обзора будут рассмотрены три группы РНКаз, отличающихся по молекулярному весу, аминокислотной последовательности и специфичности действия: семейство РНКазы А (Bientema *et al.*, 1997) (специфичность расщепления 5'-Pyr↓A-3'), семейство РНКазы T1 (Pace *et al.*, 1991) (специфичность расщепления 5'-Gp↓X-3' и 5'-ApX-3') и семейство РНКазы T2 (Irie, Ohga, 2001) (рибонуклеазы с низкой специфичностью). Структура всех природных РНКаз включает два типа активных центров: центр связывания субстрата и каталитический центр, осуществляющий связывание фосфатной группы и гидролиз фосфодиэфирной связи. В природных РНКазах проблема специфичности решается путем организации субстрат-связывающего центра, в котором взаимодействие аминокислотных остатков с РНК обеспечивает связывание гетероциклического основания и оптимальную конформацию фосфодиэфирной связи, по которой протекает гидролиз.

Субстрат-связывающий центр рибонуклеаз семейства РНКазы А (панкреатическая рибонуклеаза А, RC-РНКазы (Liao, 1992), онконаза (Ardelt *et al.*, 1991) образован боковыми радикалами треонина, фенилаланина и серина. Серин и фенилаланин находятся в частичном стэкинге с пиримидиновым кольцом субстрата, а гидроксильная группа и амидная группа треонина участвуют в связывании пиримидинового гетероцикла.

Специфическое связывание гуанина в случае рибонуклеаз семейства РНКазы T1 (РНКазы

T1 (Pace *et al.*, 1991), РНКазы F1 (Vassilyev *et al.*, 1993) осуществляется путем закрепления этого гетероциклического основания между двумя ароматическими боковыми радикалами тирозина (Tyr42 и Tyr45), а остатки аспарагиновой кислоты (Asp43 и Asp44) вместе с Tyr45 образуют сеть водородных связей с O6 и N7 атомами гуанина. Роль аргинина в субстрат-связывающем центре РНКазы T1 до сих пор до конца не установлена, но для белков, мутантных по Arg77, была показана полная потеря функциональной активности (Backmann *et al.*, 1994). Предполагают, что Arg77 участвует в стабилизации переходного пентакоординированного состояния фосфора. Связывание аденина (РНКазы U2) (Uchida *et al.*, 1970) осуществляется боковыми радикалами тирозина, глутаминовой и аспарагиновой кислот: Tyr44 и Glu69 образуют водородные связи с N7 и N1 атомами аденина соответственно.

Третье семейство, семейство РНКазы T2, включает рибонуклеазы с низкой специфичностью, такие, как РНКазы T2 (Irie, Ohga, 2001) и РНКазы Rh (Hogiuchi *et al.*, 1998) преимущественно расщепляют связи после остатков аденина, РНКазы LE (Jost *et al.*, 1991) (связи после остатков гуанина), РНКазы CL1 и РНКазы Mc1 (связи после остатков уридина). Для РНКаз этого семейства в субстрат-связывающем центре характерно присутствие кластера гидрофобных аминокислот (триптофана и тирозина), закрепляющих гетероциклическое основание между двумя ароматическими боковыми радикалами.

Одним из интересных примеров рибонуклеаз является рибонуклеаза Р (РНКазы Р), катализирующая созревание пула тРНК (пре-тРНК) во всех организмах от археобактерий до эукариот путем расщепления фосфодиэфирной связи между гуанозином с 5' конца тРНК и остатком уридина на 3'-конце пре-тРНК. В результате образуется молекула зрелой тРНК, содержащая на 5'-конце фосфатную группу, и выщепляется лидерная последовательность, содержащая на 3'-конце гидроксильную группу. Таким образом, рибонуклеаза Р расщепляет пре-тРНК с образованием 5'-фосфатной и 3'-гидроксильной групп, чем отличается от многих РНКаз и рибозимов, которые катализируют расщепление с образо-

ванием 2',3'-циклофосфата и 5'-гидроксильной группы (Hseih *et al.*, 2004). Почти все РНКазы Р из трех главных царств (археобактерии, бактерии и эукариоты) содержат в своем составе РНК и белковые субъединицы. Исключение составляют РНКазы Р из митохондрий, хлоропластов и Aquifex, которые не содержат РНК компонент (Brown, 1999; Doudna, Cech, 2002; Xiao *et al.*, 2002).

РНК в составе РНКазы Р из различных организмов значительно отличаются по длине и последовательности. РНК компонент РНКазы Р является главным каталитическим доменом этого фермента, тогда как белковый компонент РНКазы Р отвечает за сродство к пре-тРНК и увеличивает локальную концентрацию ионов магния, требующуюся для оптимальной активности фермента. Изучение структуры РНКазы Р из *Bacillus subtilis* показало, что белок представляет собой глобулу, часть которой гомологична двум РНК связывающим белкам: С-концевому домену рибосомального белка 5S и домену VI фактора элонгации G (Hansen *et al.*, 2001). Белок РНКазы Р из *Bacillus subtilis* содержит три возможных РНК-связывающих участка: (1) центральная впадина, формируемая одной α -спиралью и четырьмя β -складками, (2) отрицательно заряженная металлсвязывающая петля и (3) богатый основными аминокислотами РНК-связывающий мотив (RNR-мотив), консервативный у всего семейства РНКаз Р. Центральная впадина белка прямо взаимодействует с одноцепочечной 5'-лидерной последовательностью пре-тРНК на расстоянии 4–8 нуклеотидов от расщепляемой фосфодиэфирной связи, и, таким образом, играет роль в субстратном узнавании. Таким образом, главная функциональная роль белкового компонента в бактериальной РНКазе Р – это усиление сродства РНК-домена к субстрату и к ионам металла, а не связывание с РНК-субстратом.

2.2. Специфическое опознавание определенных нуклеотидных последовательностей в РНК аргинин-богатыми белками

В этой части обзора будет рассмотрено опознавание определенных нуклеотидных

последовательностей РНК белками на примере трех комплексов: РНК-Rev пептид; РНК-N белок фага γ и TAR РНК – Tat белок BIV (bovine immunodeficiency virus). Эти белки, связывающиеся с большой бороздкой РНК, относятся к классу аргинин-богатых белков (Drapez, 1999).

Структура комплекса Rev-RRE была изучена методом ЯМР (Battiste *et al.*, 1996) и приведена на рис. 2а. Было показано, что два мисматча пурин-пурин и основание внутренней петли (U72) создают необычный S-образный остов, в котором G71 «выпрыгивает» и располагается параллельно с G48, с которым он образует неканоническую пару. «Эффект сетки» приводит к значительному расширению большой бороздки и облегчает взаимодействие РНК с α -спиралью белка. Rev пептид принимает альфа-спиральную конформацию при связывании с РНК (Tan *et al.*, 1993) (рис. 2а).

N белок фага λ предотвращает терминацию транскрипции некоторых собственных оперонов фага посредством связывания с транскрибируемыми РНК. N-концевая часть пептида не упорядочена, но упаковывается в α -спираль при связывании с боксом В-шпильки РНК (Tan, Frankel, 1995; Van Glist *et al.*, 1997). При изучении комплекса методом ЯМР было показано, что α -спираль пептида связывается с GNRA-подобной четырехзвенной петлей, и делает большую бороздку прилегающей спирали доступной для взаимодействия (Legault *et al.*, 1998). α -спираль белка простирается на 4 пары оснований в шпильке РНК и «экспонирует» четырехзвенную петлю посредством стэкинга остатка триптофана с основанием петли (рис. 2б).

Белок Tat BIV содержит аргинин-богатый участок, который связывает шпильку TAR в вирусной мРНК. TAR район содержит два выпетленных основания, одно из которых, U12, не участвует в стэкинг взаимодействии, а второе основание, U10, располагается в большой бороздке, образуя Хугстинговские связи с парой A13•G11 (рис. 2в). Такой тип взаимодействия приводит к «застегиванию» спирали, при этом большая бороздка открывается достаточно широко для взаимодействия с пептидом, который связывается как нерегулярная β -шпилька. Обе β -шпильки Tat белка контактируют со спиралью

РНК. Специфические контакты с основаниями формируются посредством образования водородных связей с тремя остатками Arg и двумя положениями в белковом остове. Изолейцин располагается напротив Хугстиновской тройки A13•U10•G11 и играет важную роль в формировании комплекса. α -карбонильные атомы Gly71 и Gly74 располагаются близко к рибозе A21 и основанию U21, соответственно, и играют важную роль в осуществлении специфических взаимодействий; Gly76 важен для поворота шпильки пептида.

3. Применение «правил узнавания» для конструирования искусственных рибонуклеаз, расщепляющих РНК по связям после остатков гуанина

В недавних работах нами было показано, что конъюгаты на основе олигонуклеотидов, не комплементарных РНК-субстрату, и пептида (LR)₄G-NH₂, расщепляют РНК с высокой эффективностью по фосфодиэфирным связям в последовательностях двух типов: Pур-А и G-Х (Миронова, 2003; Mironova *et al.*, 2004а, 2006).

Пептидный остаток отвечает за каталитическую активность и сродство к РНК, а олигонуклеотид влияет на структуру пептида и модулирует специфичность конъюгата в структуре внутримолекулярного комплекса (Mironova *et al.*, 2006). Рибонуклеазная активность таких конъюгатов и специфичность расщепления РНК зависела от длины и последовательности олигонуклеотида (Mironova *et al.*, 2004b, 2006). В результате скрининга длины и последовательности олигонуклеотида, пептида и структуры линкерных групп был идентифицирован конъюгат девятизвенного олигонуклеотида и пептида (LR)₄G-NH₂, соединенных линкером из трех остатков дезоксирибозы, расщепляющий РНК по связям в G-Х последовательностях (per-9) (Неопубл. данные).

Исходя из литературных данных по сродству аминокислот к определенным основаниям в РНК было сделано предположение, что именно остатки аргинина в пептиде отвечают за сродство конъюгата к остаткам гуанина. Для под-

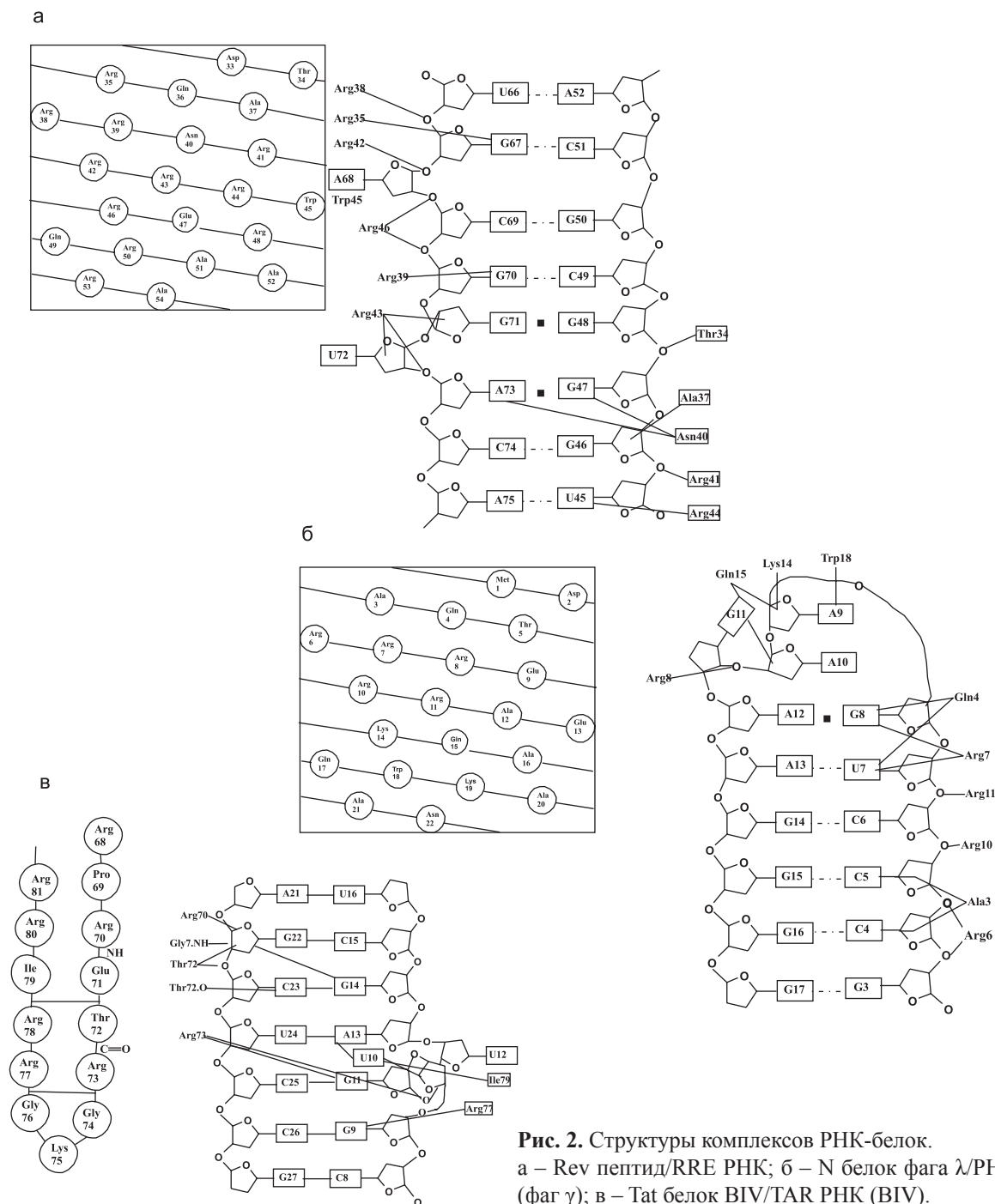


Рис. 2. Структуры комплексов РНК-белок. а – Rev пептид/RRE РНК; б – N белок фага λ/РНК (фаг γ); в – Tat белок BIV/TAR РНК (BIV).

тверждения этой гипотезы использовали метод химической модификации РНК с последующим расщеплением модифицированного субстрата конъюгатом пер-9. В качестве модифицирующих реагентов использовали диметилсульфат (ДМС), метилирующий остатки гуанина по N7 атому, и глиоксаль, модификация которым приводит к циклизации N1 атома гуанина и экзоциклической аминогруппы второго атома, в результате чего

эти два положения становятся закрытыми от взаимодействия (рис. 3).

Модификацию проводили согласно методикам, описанным в работах (Brunel, Romby, 2000; Zhang *et al.*, 2001). В экспериментах по модификации ДМС и глиоксалем использовали 21-звенный олигонуклеотид ON21, содержащий шесть остатков гуанина (рис. 3а). На рис. 3в представлен анализ расщепления ON21 и

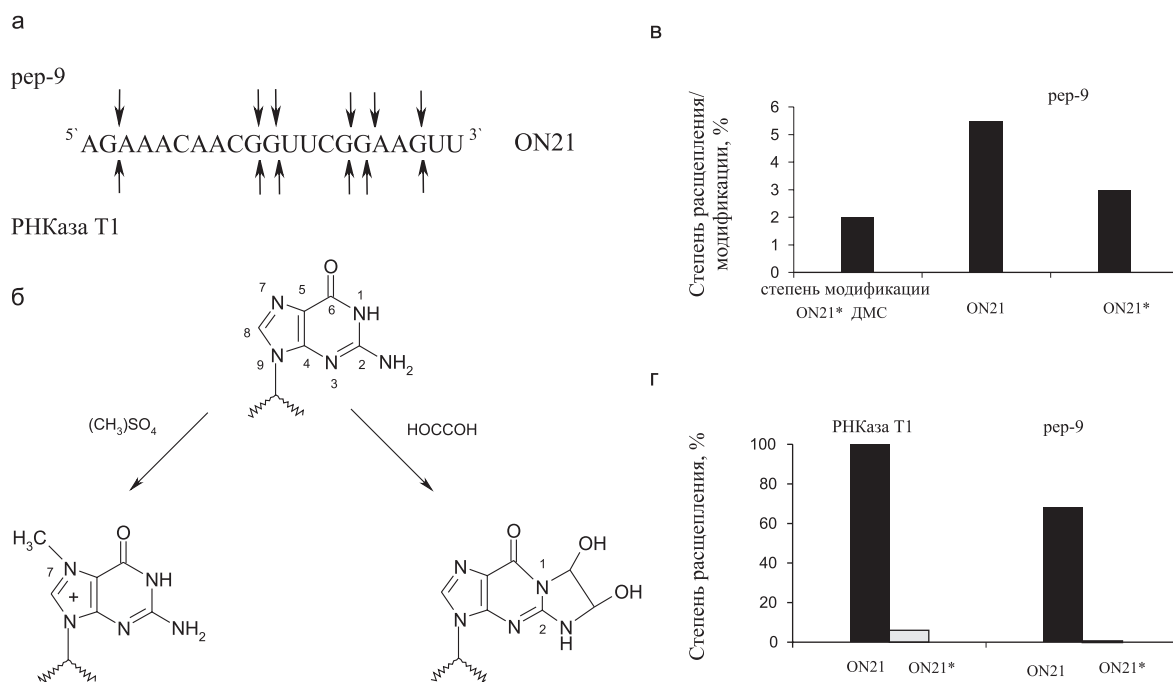


Рис. 3. Остатки гуанина в олигодезоксирибонуклеотиде ON21, расщепляемые РНКазой Т1 и конъюгатом пер-9 (а). Функциональные группы, блокируемые ДМС и глиоксалем в гуанине (б). Сравнение эффективности расщепления ON21 и ON21, модифицированного диметилсульфатом, конъюгатом пер-9 (в). Сравнение эффективности расщепления ON21 и ON21, модифицированного глиоксалем, конъюгатом пер-9 и РНКазой Т1 (г).

ON21, модифицированного диметилсульфатом, конъюгатом пер-9. Для оценки степени модификации ON21 диметилсульфатом использовали реакцию восстановления боргидридом натрия с последующим расщеплением анилином. Степень модификации составила 2 %. Конъюгат пер-9 расщеплял модифицированный ON21 по связям в G-X-последовательностях с эффективностью в 2 раза меньшей, чем немодифицированный ON21, и снижение эффективности расщепления было прямо пропорционально степени модификации.

На рис. 3д представлен анализ расщепления рибоолигонуклеотида ON21 и ON21, модифицированного глиоксалем, конъюгатом пер-9 и РНКазой Т1.

Известно, что РНКазы Т1 не расщепляют остатки гуанина, содержащие модификации в 1-м и во 2-м положениях (Brunel, Romby, 2000), и, таким образом, снижение степени расщепления модифицированного ON21 этой рибонуклеазой может служить контролем степени модифика-

ции. Как видно из представленных данных, РНКазы Т1 в концентрации 10 ед. акт. расщепляла немодифицированный ON21 на 100 %, тогда как в случае ON21, модифицированного глиоксалем, степень расщепления составляла 68 %. Таким образом, степень модификации ON21 глиоксалем составила 32 %. Конъюгат пер-9 расщеплял немодифицированный ON21 с эффективностью 6 % за 8 ч. Модификация ON21 глиоксалем приводила к полному исчезновению расщепления конъюгатом пер-9 по связям в G-X-последовательностях.

Полученные данные продемонстрировали, что N1 атом гуанина и экзоциклическая аминогруппа при C2 атоме гуанина являются функционально важными и, возможно, участвуют в образовании комплекса с олигонуклеотид-пептидным конъюгатом. Модификация по N7 положению гуанина может оказывать влияние на расщепление конъюгатом по двум причинам: 1) N7 положение участвует в образовании водородных связей с конъюгатом и в этом случае

его метилирование препятствует связыванию с конъюгатом и, следовательно, расщеплению; 2) появление положительного заряда у остатка гуанина, модифицированного ДМС, приводит к электростатическому отталкиванию положительно заряженных остатков аргинина. Таким образом, специфическое опознавание гуанина в РНК осуществляется остатками аргинина в пептиде, взаимодействующем с N7 и N1 положениями основания.

Заключение

Таким образом, несмотря на то, что не существует единой формулы опознавания аминокислотами нуклеиновых кислот (НК), можно говорить о некоторых правилах связывания, определяющих специфичность взаимодействия определенных оснований или пар оснований с аминокислотами. При опознавании единичного основания в НК множественные водородные связи являются ключевыми и определяют универсальную специфичность аргинина и лизина к гуанину. Глутаминовая и аспарагиновая кислоты также специфично связывают остатки гуанина в РНК, а для аспарагина и глутаминина показано преимущественное связывание с уридином. Специфичное связывание с короткими нуклеотидными последовательностями в РНК достигается за счет взаимодействия аминокислоты более чем с двумя основаниями, последовательно расположенными в цепи. Наибольшая специфичность при узнавании уникальных участков в РНК достигается при взаимодействии аминокислоты с обоими основаниями неканонической пары (например, Dгарег, 1999).

Специфическое взаимодействие рибонуклеаз с основаниями РНК достигается более сложными способами, чем с помощью опознавания основания одной аминокислотой. Важной характерной особенностью субстрат-связывающих центров сиквенс-специфических рибонуклеаз (семейства РНКазы А и РНКазы Т1) является связывание гетероциклического основания путем его стэкинга с боковыми радикалами ароматических аминокислотных остатков и за счет формирования уникальной сети водородных связей между гетероциклическим

основанием в составе РНК и аминокислотами (аспарагиновой кислотой, аргинином, лизином и т. д.) в составе фермента. Отличительной чертой неспецифических рибонуклеаз является отсутствие аминокислотного «кода узнавания» для основания, а закрепление основания в активном центре достигается с помощью кластера гидрофобных аминокислот.

На искусственной конструкции применение кода узнавания определенных нуклеотидов определенными аминокислотами удалось продемонстрировать впервые: аргинин-богатый пептид в структуре конъюгата пер-9 способен опознавать остатки гуанина в РНК и катализировать расщепление фосфодиэфирной связи, расположенной рядом.

Работа поддержана грантами программ фундаментальных исследований РАН «Молекулярная и клеточная биология» и «Фундаментальные науки – медицине», РФФИ 05-04-49109, грантом CRDF Y1-B-08-08 и грантом Министерства образования Российской Федерации.

Литература

- Зенкова М.А., Власов А.В., Коневец Д.А. и др. Химические рибонуклеазы. Синтез и РНК-гидролизующая активность химических рибонуклеаз – конъюгатов 1,4-диазабцикло[2.2.2]-октана, несущих различный суммарный положительный заряд // Биорг. химия. 2000. Т. 26. С. 610–616.
- Миронова Н.Л. Катализаторы расщепления РНК – пептидолигонуклеотиды и органические соединения, включающие основные аминокислоты и катионные структуры: Дис. ... канд. биол. наук. Новосибирск: Новосибирский ин-т биоорганической химии СО РАН, 2003. 151 с.
- Allers J., Shamo Y. Structure-based analysis of protein-RNA interactions using the program ENTANGLE // J. Mol. Biol. 2001. V. 311. P. 75–86.
- Ardelt W., Mikulski S.M., Shogen K. Amino acid sequence of an anti-tumor protein from *Rana pipiens* oocytes and early embryos // J. Biol. Chem. 1991. V. 266. P. 245–251.
- Arnez J.G., Steitz T.A. Crystal structures of three miscylating mutants of *Escherichia coli* glutamyl-tRNA synthetase complexed with tRNA(Gln) and ATP // Biochemistry. 1996. V. 35. P. 14725–14733.
- Backmann J., Doray C.C., Grunert H.P. *et al.* Extended kinetic analysis of ribonuclease T1 variants leads to an improved scheme for the reaction mechanism //

- Biochem. Biophys. Res. Commun. 1994. V. 199. P. 213–219.
- Battiste J.L., Mao H., Rao N.S. *et al.* α Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex // *Science*. 1996. V. 273. P. 1547–1551.
- Beintema J.J., Breukelman H.J., Carsana A., Furia A. // *Ribonucleases, Structures and Function* / Ed. G. D'Alessio, J.F. Riordan. N.Y.: Academic, 1997. P. 265.
- Brown J.W. The Ribonuclease P Database // *Nucl. Acids Res.* 1999. V. 27. № 1. P. 314.
- Brunel C., Romby P. Probing RNA structure and RNA-ligand complexes with chemical probes // *Methods Enzymol.* 2000. V. 318. P. 3–2.
- Cheng A.C., Chen W.W., Fuhramann C.N., Frankel A.D. Recognition of nucleic acid bases and base-pairs by hydrogen bonding to amino acid side-chains // *J. Mol. Biol.* 2003. V. 327. P. 781–796.
- Ding J., Hayashi M.K., Zhang Y. *et al.* Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA // *Genes Dev.* 1999. V. 13. P. 1102–1115.
- Doudna J.A., Cech T.R. The chemical repertoire of natural ribozymes // *Nature*. 2002. V. 418 (6894). P. 222–228.
- Draper D.E. Themes in RNA-protein recognition // *J. Mol. Biol.* 1999. V. 293. P. 255–270.
- Endo M., Azuma Y., Saga Y. *et al.* Molecular design for a RNA scission. interposition of oligoamines between two DNA oligomers // *J. Org. Chem.* 1997. V. 62. P. 846–848.
- Endo M., Hirata K., Ihahra T. *et al.* RNA hydrolysis by the cooperation of carboxylate ion and ammonium // *J. Am. Chem. Soc.* 1996. V. 118. P. 5478–5479.
- Goedecke K., Pignot M., Goody R.S. *et al.* Structure of the N6-adenine DNA methyltransferase M. TaqI in complex with DNA and a cofactor analog // *Nature Struct. Biol.* 2001. V. 8. P. 121–125.
- Hainzl T., Huang S., Sauer-Eriksson A.E. Structure of the SRP19 RNA complex and implications for signal recognition particle assembly // *Nature*. 2002. V. 417. P. 767–771.
- Hansen A., Pfeiffer T., Zuleeg T. *et al.* Exploring the minimal substrate requirements for trans-cleavage by RNase P holoenzymes from *Escherichia coli* and *Bacillus subtilis* // *Mol. Microbiol.* 2001. V. 41. № 1. P. 131–143.
- Hermann T., Patel D.J. Stitching together RNA tertiary architectures // *J. Mol. Biol.* 1999. V. 294. P. 829–849.
- Horiuchi H., Yanai K., Takagi M. *et al.* Primary structure of a base nonspecific ribonuclease from *Rhizopus niveus* // *J. Biochem.* 1988. V. 103. P. 408–418.
- Horvath M.P., Schweiker V.L., Bevilacqua J.M. *et al.* Crystal structure of the *Oxytricha nova* telomere and binding protein complexed with single strand DNA // *Cell*. 1998. V. 95. P. 963–974.
- Hsieh J., Andrews A.J., Fierke C.A. Roles of protein subunits in RNA-protein complexes: lessons from ribonuclease P // *Biopolymers*. 2004. V. 73. P. 79–89.
- Irie M., Ohga K. Primary structure of a base nonspecific ribonuclease from *Rhizopus niveus* // *Methods Enzymol.* 2001. V. 341. P. 42–55.
- Jones S., Daley D.T.A., Luscombe N.M. *et al.* Protein–RNA interactions: a structural analysis // *Nucl. Acids Res.* 2001. V. 29. P. 943–954.
- Jost W., Bak H., Glund K. *et al.* Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esulentum*) cells // *Eur. J. Biochem.* 1991. V. 198. P. 1–6.
- Klimasauskas S., Kumar S., Roberts R.J., Cheng X. HhaI methyltransferase flips its target base out of the DNA helix // *Cell*. 1994. V. 76. P. 357–369.
- Ko T.P., Williams R., McPherson A. Structure of a ribonuclease B/pd(pA)₄ complex // *Acta Crystallog. sect. D.* 1996. V. 52. P. 160–164.
- Legault P., Li J., Mogridge J., Kay L.E., Greenblatt J. NMR structure of the bacteriophage 268 RNA-Protein Recognition lambda N peptide/box B RNA complex: recognition of a GNRA fold by an arginine-rich motif // 1998. V. 93. № 2. P. 289–299.
- Liao Y.D. A pyrimidine-guanine sequence-specific ribonucleases from *Rana catesbeiana* (bullfrog) oocytes // *Nucl. Acids Res.* 1992. V. 20. P. 1371–1377.
- Luscombe N.M., Laskowski R.A., Thornton J.M. Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level // *Nucl. Acids Res.* 2001. V. 29. P. 2860–2874.
- Mao H., White S.A., Williamson J.R. A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex // *Nature Struct. Biol.* 1999. V. 6. P. 1139–1147.
- Mironova N.L., Pyshnyi D.V., Ivanova E.M. *et al.* Covalently attached oligodeoxyribonucleotides induce RNase activity of a short peptide and modulate its base-specificity // *Nucl. Acids Res.* 2004a. V. 32. P. 1928–1936.
- Mironova N.L., Pyshnyi D.V., Ivanova E.M. RNA-cleaving oligonucleotide-peptide conjugates // *Artificial Ribonuclease* / Ed. M.A. Zenkova. *Nucleic Acids and Molecular Biology.* 2004b. V. 13. P. 151–172.
- Mironova N.L., Pyshnyi D.V., Stadler D.V. *et al.* G-specific RNA-cleaving conjugates of short peptides and oligodeoxyribonucleotides // *J. Biomol. Struct. Dyn.* 2006. V. 23. № 6. P. 591–602.
- Morellet N., Demene H., Teilleux V. *et al.* Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide

- d(ACGCC) // *J. Mol. Biol.* 1998. V. 283. P. 419–434.
- Morrow J.R. Artificial ribonucleases // *Adv. Inorg. Biochem.* 1994. V. 9. P. 41–74.
- Najmudin S., Cote M.L., Sun D. *et al.* Crystal structures of an N-terminal fragment from *Moloney murine leukemia virus reverse transcriptase* complexed with nucleic acid: functional implications for template-primer binding to the fingers domain // *J. Mol. Biol.* 2000. V. 296. P. 613–632.
- Norman D.P., Bruner S.D., Verdine G.L. Coupling of substrate recognition and catalysis by a human base-excision repair protein // *J. Am. Chem. Soc.* 2001. V. 123. P. 359–360.
- Ogata K., Morikawa S., Nakamura H. *et al.* Solution structure of a specific DNA complex of the Myb DNA binding domain with cooperative recognition // *Cell.* 1994. V. 79. № 4. P. 639–648.
- Oost T., Kalesse M. Synthesis of RNase active site model systems using a steroid template // *Tetrahedron.* 1997. V. 53. P. 8421–8438.
- Pabo C.O., Necludova L. Geometric analysis and comparison of protein–DNA interfaces: why is there no simple code for recognition? // *J. Mol. Biol.* 2000. V. 301. P. 597–624.
- Pace C.N., Heinemann U., Hahn U., Saenger W. Ribonuclease T1: structure, function and stability // *Angew. Chem.* 1991. V. 66. P. 403–421.
- Reinisch K.M., Chen L., Verdine G.L., Lipscomb W.N. The crystal structure of HaeIII methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing // *Cell.* 1995. V. 82. P. 143–153.
- Tan R., Chen L., Buettner J.A. *et al.* RNA recognition by an isolated α helix // *Cell.* 1993. V. 73. P. 1031–1040.
- Tan R., Frankel A.D. Structural variety of arginine-rich RNA-binding peptides // *Proc. Natl Acad. Sci. USA.* 1995. V. 92. P. 5282–5286.
- Treger M., Westhof E. Statistical analysis of atomic contacts at RNA–protein interfaces // *J. Mol. Recog.* 2001. V. 14. P. 199–214.
- Tung C.H., Wei Z., Leibowitz M.J., Stein S. Design of peptide-acridine mimics of ribonuclease activity // *Proc. Natl Acad. Sci. USA.* 1992. V. 89. P. 7114–7118.
- Uchida T., Arima T., Egami F. Specificity of RNase U2 // *J. Biochem (Tokio).* 1970. V. 67. P. 91–102.
- Van Gilst M.R., Rees W.A., Das A., von Hippel P.H. Complexes of N antitermination protein of phage lambda with specific and nonspecific RNA target sites on the nascent transcript // *Biochemistry.* 1997. V. 36. P. 1514–1524.
- Vassilyev D.G., Katayanagi K., Ishikawa K. *et al.* Crystal structure of ribonuclease F1 of *Fusarium moniforme* in its free form and in complex with 2'GMP // *J. Mol. Biol.* 1993. V. 230. P. 979–996.
- Walberer B.J., Cheng A.C., Frankel A.D. Structural diversity and isomorphism of hydrogen bonded base interactions in nucleic acids // *J. Mol. Biol.* 2003. V. 327. № 4. P. 767–780.
- Xiao S., Scott F., Fierke C.A., Engelke D.R. Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes // *Annu. Rev. Biochem.* 2002. V. 71. P. 165–189.
- Ye X., Gorin A., Ellington A.D., Patel D.J. Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex // *Nature Struct. Biol.* 1996. V. 3. P. 1026–1033.
- Zenkova M., Beloglazova N., Sil'nikov V. *et al.* RNA cleavage by 1,4-diazabicyclo[2.2.2]octan-imidazole conjugates // *Methods Enzymol.* 2001. V. 341. P. 468–490.
- Zhang C., Trottier M., Chen C., Guo P. Chemical modification patterns of active and inactive as well as procapsid-bound and unbound DNA-packaging RNA of bacterial virus Phi29 // *Virology.* 2001. V. 281. P. 281–293.

BASIS OF SPECIFICITY IN INTERACTIONS BETWEEN NUCLEIC ACIDS AND PROTEINS

D.V. Stadler, N.L. Mironova, M.A. Zenkova, V.V. Vlassov

Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia,
e-mail: marzen@niboch.nsc.ru

Summary

In this review we described contacts between RNA and amino acids and proteins with various functions providing specific recognition with individual heterocyclic bases and nucleotide sequences. It was described application of «recognition code» for development of artificial ribonucleases on the base of oligonucleotide-peptide conjugates displaying the specificity of RNase T1.