

МНОЖЕСТВЕННЫЕ НЕКАНОНИЧЕСКИЕ СТЭКИНГ-ВЗАИМОДЕЙСТВИЯ ГЕТЕРОЦИКЛИЧЕСКИХ ОСНОВАНИЙ КАК ОДИН ИЗ ГЛАВНЫХ ФАКТОРОВ ОРГАНИЗАЦИИ ТРЕТИЧНОЙ СТРУКТУРЫ РНК

© 2023 В.Г. Метелев¹, Е.Ф. Баулин², А.А. Богданов^{1,3,4*}

¹ Московский государственный университет имени М.В. Ломоносова,
химический факультет, 119991 Москва, Россия

² Институт математических проблем биологии РАН – филиал ИПМ имени М.В. Келдыша РАН,
142290 Пущино, Россия

³ НИИ физико-химической биологии имени А.Н. Белозерского,
Московский государственный университет имени М.В. Ломоносова,
119992 Москва, Россия; электронная почта: bogdanov@belozersky.msu.ru

⁴ Государственный научный центр Федеральное государственное бюджетное учреждение науки
Институт биоорганической химии имени академиков М.М. Шемякина и Ю.А. Овчинникова РАН,
117997 Москва, Россия

Поступила в редакцию 21.03.2023

После доработки 18.04.2023

Принята к публикации 20.04.2023

Межплоскостные (стэкинг) взаимодействия гетероциклических оснований нуклеотидных остатков (н.о.) РНК являются одним из важнейших факторов организации её вторичной и третичной структуры. Большая часть таких (канонических) взаимодействий осуществляется между непосредственными соседями в полинуклеотидных цепях РНК. Однако по мере накопления данных об атомных третичных структурах самых разнообразных РНК и их комплексов с белками стало ясно, что с помощью стэкинга оснований (неканонического) могут взаимодействовать н.о. РНК, не являющиеся соседями в их полинуклеотидных цепях и иногда разделённые в первичной структуре РНК десятками или сотнями н.о. В настоящей работе представлена исчерпывающая база данных о таких элементах и их окружении в макромолекулах природных и синтетических РНК. Они названы NA-BSE (nonadjacent base-stacking elements). Анализ этих данных показал, что формирующие NA-BSE н.о. составляют в среднем около четверти всех н.о. той или иной РНК и должны рассматриваться как полноценные мотивы в их третичной структуре. В работе проведена классификация NA-BSE по типам локализации в макромолекулах РНК. Показано, что структурообразующая роль NA-BSE состоит в компактном сворачивании однотяжевых петель РНК, в превращении двутяжевых и более сложноорганизованных петель в несовершенные спирали, а также в связывании удалённых в первичной и вторичной структуре районов РНК.

КЛЮЧЕВЫЕ СЛОВА: стэкинг-взаимодействия, третичная структура РНК, рибосомные РНК, рибонуклеаза Р, РНК-мотивы.

DOI: 10.31857/S0320972523060076, EDN: EFLFMJ

ВВЕДЕНИЕ

Известно, что макромолекулярная структура РНК определяет её функциональные свойства. И хотя общие принципы организации вторичной и третичной структуры одноцепочечных РНК были сформулированы Fresco et. al. [1] и Спириным [2] более 60 лет назад, сегодня мы

всё ещё не можем, располагая только нуклеотидной последовательностью некой достаточно протяжённой РНК, достоверно предсказать её пространственную структуру.

За последние 20–25 лет в изучении макромолекулярной структуры РНК произошла настоящая революция: в руках исследователей оказались тысячи третичных структур разно-

Принятые сокращения: н.о. – нуклеотидные остатки; РНКазы Р – рибонуклеаза Р; BIE – base-intercalated element; BWE – base-waged element; DC – double-crossing; NA-BSE – nonadjacent base-stacking element.

* Адресат для корреспонденции.

образных природных РНК и их синтетических аналогов, полученных с атомным разрешением с помощью рентгеноструктурного анализа, криоэлектронной микроскопии или ЯМР-спектроскопии. Следует подчеркнуть, что многие из этих структур описывают детали пространственной организации макромолекул РНК в комплексе со специфическими белками, то есть в состоянии, близком к тому, в котором они находятся в клетке. Это особенно важно, поскольку пространственная структура свободной РНК, за редкими исключениями, существенно отличается от таковой в природных РНК-белковых комплексах. В настоящее время общедоступные банки данных (PDB, NDB) содержат более 5000 таких структур. Анализ этих данных постоянно приносит информацию об общих структурных элементах, т.н. мотивах, которые повторяются с разной степенью регулярности в третичной структуре широкого спектра РНК [3]. Так, например, уже при анализе первых атомных структур в рибосомах (рибосомных РНК, рРНК) были идентифицированы такие элементы, как А-мины [4] и К-повороты [5], и продемонстрирована их общность для РНК других классов. Следует подчеркнуть, что открытие новых мотивов постоянно продолжается. Например, относительно недавно был обнаружен и охарактеризован широко распространённый мотив, в котором свободная пара электронов О4'-атома остатка рибозы одного нуклеотида взаимодействует с плоскостью гетероциклического основания другого нуклеотида [6]. Мы, в свою очередь, описали мотивы, которые мы назвали BIE и BWE (base-intercalated element и base-waged element соответственно) [7]. В этих структурных элементах гетероциклическое основание удалённого нуклеотидного остатка (н.о.) либо интеркалирует между двумя соседними основаниями определённого динуклеотидного сегмента РНК (BIE), либо вытесняет среднее основание из тринуклеотидного сегмента, занимая его место (BWE). BIE и BWE участвуют не только в сворачивании полинуклеотидных цепей РНК в уникальные пространственные структуры, но также важны для организации функциональных центров РНК и РНК-белковых комплексов. В ходе анализа BIE и BWE мы заметили, что эти структурные элементы являются представителями гораздо более обширного семейства мотивов. Общим для этого семейства является стэкинг-взаимодействие между основаниями н.о., которые расположены в первичной структуре РНК на разных расстояниях друг от друга, но не являются непосредственными соседями. Мы объединили

их под сокращённым обозначением NA-BSE (nonadjacent base-stacking elements). Соответственно, межплоскостные взаимодействия нуклеиновых оснований в NA-BSE мы называем здесь неканоническими, чтобы отличить их от канонических стэкинг-взаимодействий соседних н.о. в РНК и полинуклеотидах.

В этой статье мы представляем базу данных, включающую все разновидности мотивов этого семейства, и рассматриваем их структурную роль. Они обнаруживаются во всех природных и многих синтетических РНК, для которых на сегодняшний день установлена третичная структура с высоким разрешением. Следует отметить, что существование NA-BSE некоторых типов ранее эпизодически фиксировалось в третичных структурах ряда РНК и рибонуклеопротеинов (РНП), см., например, работы Dallas et al., Cate et al. и Teplova et al. [8–10]. Однако их никогда не рассматривали как повторяющиеся элементы (мотивы) в пространственной структуре РНК, их не классифицировали по типам локализации в макромолекулах РНК, а также не анализировали их структурообразующие и функциональные свойства. Такой систематический анализ проводится впервые в данной работе. Он показывает, что NA-BSE не только широко распространены во всех классах РНК, но и играют значительную роль в организации их третичной структуры.

МАТЕРИАЛЫ И МЕТОДЫ

Мы определили NA-BSE как максимальный набор из k нуклеотидов $N_1, \dots, N_k$, такой, что для каждого $i < k$ N_i и N_{i+1} образуют неканоническое стэкинг-взаимодействие оснований, т.е. они не являются соседями по последовательности, и их основания уложены в стопку.

Для построения набора данных NA-BSE мы использовали неизбыточный набор РНК-содержащих структур из PDB с ограничением среднего разрешения 3,0 Å (версия 3.166 [11]). Для аннотирования NA-BSE использовалась программа DSSR (версия 2.0 [12]). Результирующий набор данных включал только NA-BSE из цепей РНК длиной ≥ 25 нуклеотидов, при этом анализировались стэкинг-взаимодействия оснований, для которых угол основание-основание был $< 30^\circ$, и расстояние между ближайшими атомами двух оснований составляло $< 4,0$ Å. Другой набор данных был создан с использованием неизбыточного набора структур с ограничением разрешения 3,5 Å.

Столбцы подготовленных таблиц были такими же, как в работе Baulin et al. [7], и включали параметры стэкинга, параметры спариваний, в которых участвуют основания, и другие (см. вкладку readme в табл. S1 Приложения). Мы также аннотировали третичные мотивы РНК, имеющие общие нуклеотиды с NA-BSE, включая коаксиальный стэкинг, А-миноры, К-повороты, рибозные молнии, BIE, BWE, динуклеотидные платформы, внутренние петли с тандемными GA-спариваниями и внутренние петли типа UAA/GAN (см. табл. S2 Приложения). Внутренние петли были аннотированы с использованием пакета URSLIB [13], а все остальные мотивы были аннотированы с использованием DSSR.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Базы данных и статистика. В качестве основной в этой работе использовалась база данных, разработанная на основании анализа третичных структур РНК, полученных с разрешением не ниже 3,0 Å (база данных «3,0»). Она включает в себя 11 733 NA-BSE различных размеров, образованных за счёт стэкинг-взаимодействий между гетероциклическими основаниями 14 862 н.о., не являющихся непосредственными соседями в полинуклеотидных цепях РНК (табл. S1 Приложения). При этом 9525 NA-BSE состоят из двух н.о., 1500 мотивов — из трёх, 512 мотивов — из четырёх, 179 мотивов — из пяти и 17 мотивов — из шести н.о. В основной базе данных NA-BSE, образованные н.о., принадлежащими разным молекулам РНК, составляют всего 3,3%.

Из 11 733 NA-BSE 60,5% представлены только пуриновыми н.о., 32,4% — пуриновыми и пиримидиновыми и 5,9% — только пиримидиновыми н.о.; 1,2% NA-BSE содержат модифицированные основания.

Мотивы, описываемые в этой работе, обнаруживаются во всех элементах вторичной структуры РНК. Это хорошо иллюстрирует рис. 1, на котором в классической модели вторичной структуры сегмента (49-360) рРНК малой субъединицы рибосом *E. coli* отмечены н.о., образующие NA-BSE в этом сегменте рРНК в составе малой субъединицы рибосомы. Они составляют примерно четверть (25,6%) от общего числа н.о. в этом сегменте РНК. Похожая картина наблюдается и для NA-BSE 23S рРНК этих рибосом, доля которых равна примерно 26% от их общего числа н.о. в её макромолекуле.

Более 98% NA-BSE, входящих в основную базу данных, дополнительно стабилизированы

взаимодействием основания с фосфатными группами, 22,7% мотивов стабилизированы взаимодействием основания одного н.о. с O4'-атомом рибозы другого н.о. В большинстве NA-BSE (56,1%) их основания находятся в стэкинге с основаниями фланкирующих н.о., расположенных с обеих сторон мотива; у 35,1% мотивов взаимодействующий с ними фланкирующий нуклеотид расположен только с одной стороны, а 8,8% мотивов вообще не имеют фланкирующих н.о. Кроме того, гетероциклические основания в большинстве NA-BSE образуют с основаниями соседних районов РНК как уотсон-криковские, так и (гораздо чаще) неутсон-криковские пары (см. табл. S1 Приложения и рассматриваемые ниже примеры).

6903 NA-BSE (58,8%) не участвуют в формировании каких-либо других известных мотивов, обнаруженных в третичных структурах РНК; в то же время 4830 NA-BSE (41,2%), входящих в базу данных «3,0», имели общие нуклеотиды по крайней мере с одним из таких мотивов.

Типы NA-BSE. Структурные мотивы, описываемые в этой работе, можно подразделить на три типа.

NA-BSE первого типа (NA-BSE I) образованы н.о. одного и того же одноцепочечного участка или (крайне редко) остатками одной из цепей двухцепочечного участка РНК. В первичной структуре РНК они разделены всего несколькими (часто одним) н.о., а образование мотива сопровождается выходом из петли или из двойной спирали одного или нескольких остатков (рис. 2, а). Эти остатки часто играют важную роль в узнавании РНК белками (см., например, работы Yogesh et al. и Laura et al. [14, 15]).

Ко второму типу (NA-BSE II) мотивов этого семейства относятся структурные элементы, в которых стэкинг-взаимодействия осуществляются между гетероциклическими основаниями н.о., находящихся в разных цепях одной двойной спирали (как, например, на рис. 2, б) или более сложноорганизованной петли, а также в участках противоположных полинуклеотидных цепей, связанных друг с другом, в частности, благодаря образованию неканонических (неутсон-криковских) пар оснований. (Следует заметить, что в традиционных схемах вторичных структурах РНК они часто представлены в виде петель.)

К третьему типу NA-BSE (NA-BSE LR, long-range) относятся мотивы, образованные за счёт стэкинг-взаимодействий нуклеотидов, расположенных в разных петлях РНК, в разных цепях одной и той же РНК (при этом расстояние

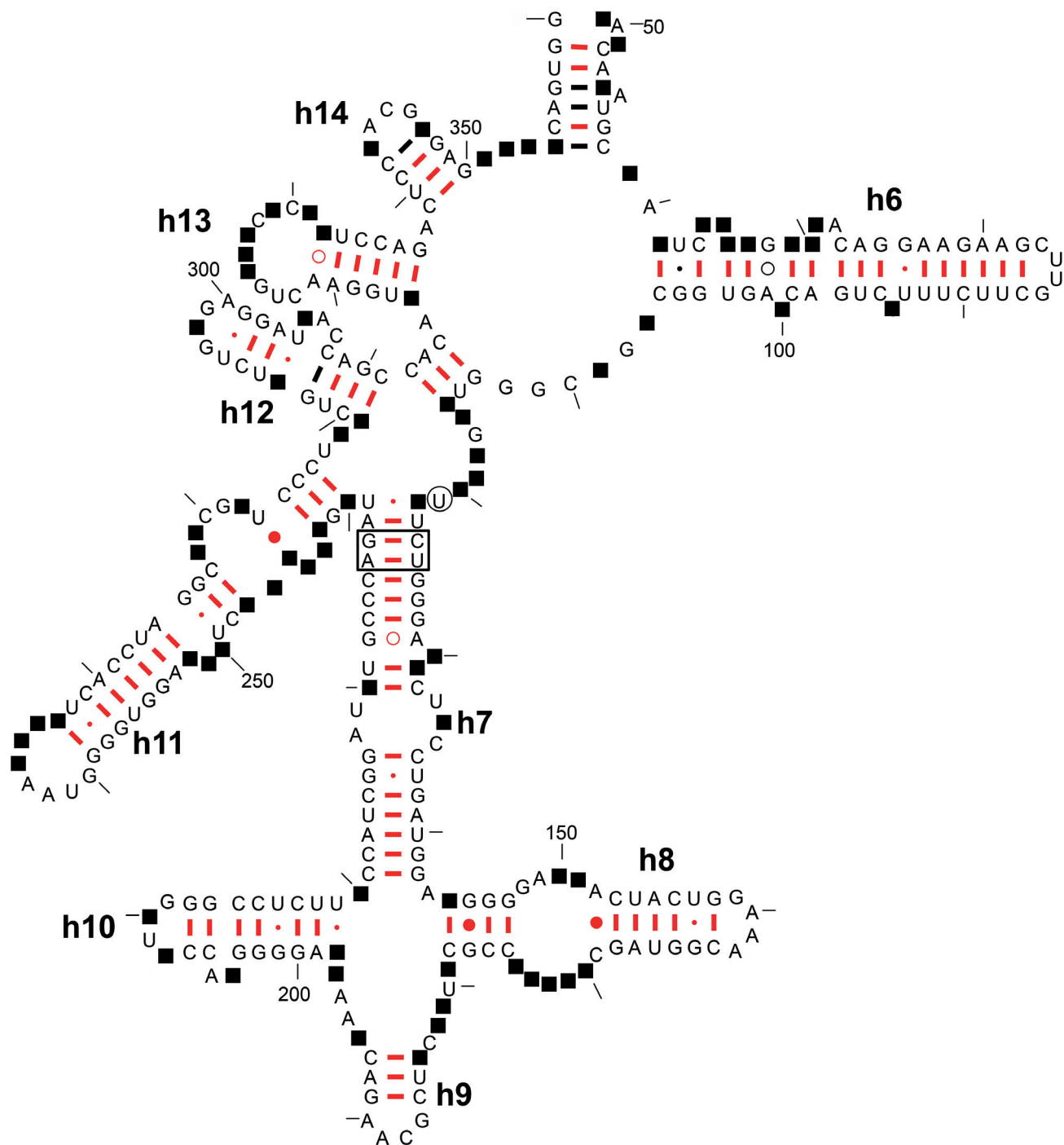


Рис. 1. Схема канонической вторичной структуры сегмента (49-360) 16S рРНК *E. coli* (http://rna.bgsu.edu/rna3dhub/motifs/2ds/escherichia_coli_16s). Чёрными квадратами отмечены н.о., образующие NA-BSE в этом сегменте рРНК в составе малой субъединицы рибосомы

между взаимодействующими нуклеотидами в первичной структуре РНК может достигать нескольких сотен н.о. (рис. 3)) или в разных молекулах РНК. Так, в частности, в связывании 5,8S рРНК с остальной частью рРНК большой субъединицы рибосом эукариот участвуют семь NA-BSE третьего типа (см., например, файл 6ZMI в табл. S1 Приложения).

Следует подчеркнуть, что BIE и BWE, подробно описанные ранее [7], могут относиться к любому из трёх типов NA-BSE. В основной базе данных, представленной в настоящей работе, они, как правило, входят в состав NA-BSE, состоящих из четырёх н.о.

Наиболее распространённым типом NA-BSE оказался тип II, на долю которого приходится

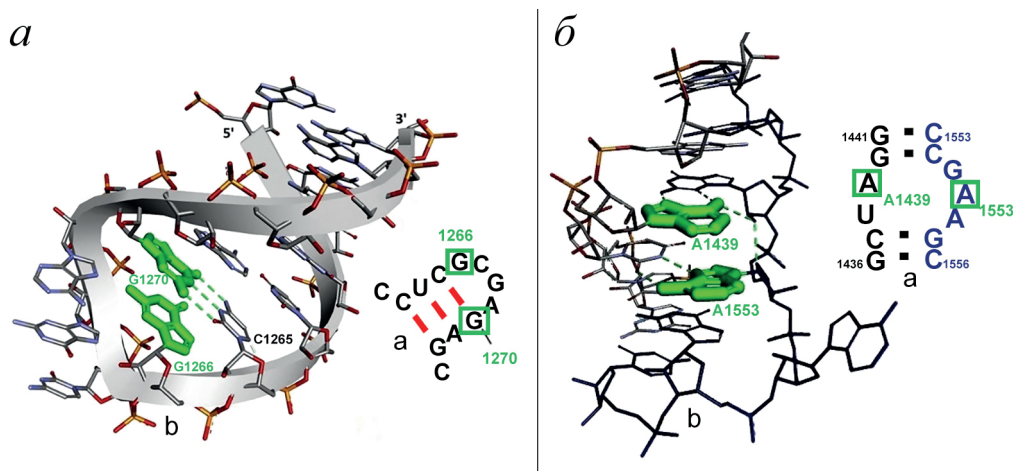


Рис. 2. Локализация н.о., образующих NA-BSE первого и второго типа. *а* – Локализация н.о., образующих NA-BSE первого типа, во фрагменте 16S рРНК в составе рибосом *E. coli*. *б* – Локализация н.о., образующих NA-BSE второго типа, во фрагменте 23S рРНК в составе рибосом *Thermus thermophilus*. Панели (*а*) – локализация н.о. во вторичной структуре, панели (*б*) – пространственная организация этих мотивов в третичной структуре рРНК (PDB ID 7K00 и 4Y40 соответственно)

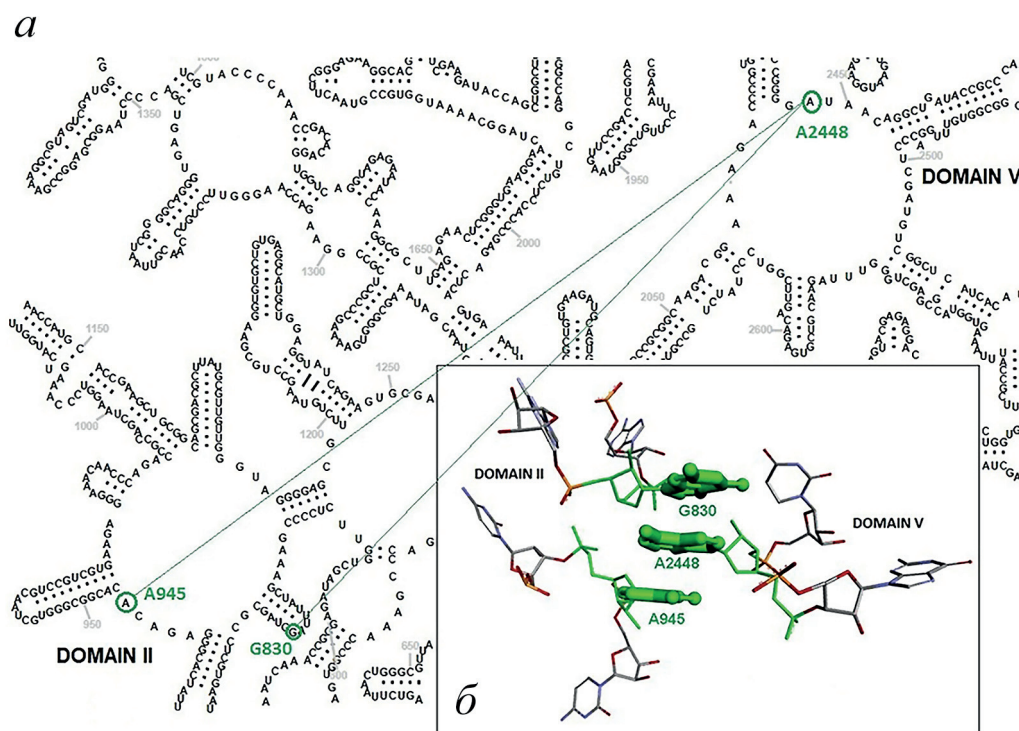


Рис. 3. Локализация н.о., образующих NA-BSE третьего типа, в структуре 23S рРНК рибосом *E. coli*. *а* – Локализация н.о. во вторичной структуре. *б* – Пространственная организация этого мотива во внутририбосомной третичной структуре рРНК (PDB ID 7K00)

45,5% мотивов, при этом доли мотивов типа I и типа LR в основной базе данных составили 31,8% и 22,7% соответственно.

Структурообразующие функции NA-BSE. Разнообразие структурообразующих функций NA-BSE мы продемонстрируем здесь на примере рибонуклеазы Р (РНКазы Р). Структура рибозимов этого класса из различных источников, а также их комплексов с тРНК, процессинг

5'-концевого района которых они осуществляют, была детально изучена (для обзора см. работу Mondragón et al. [16]). Макромолекула РНКазы Р состоит из двух доменов: S-домена, специфически узнающего её субстрат, предшественник тРНК, и С-домена, обладающего каталитической активностью. В настоящей работе была проанализирована атомная структура S-домена РНКазы Р из *T. thermophilus*, кото-

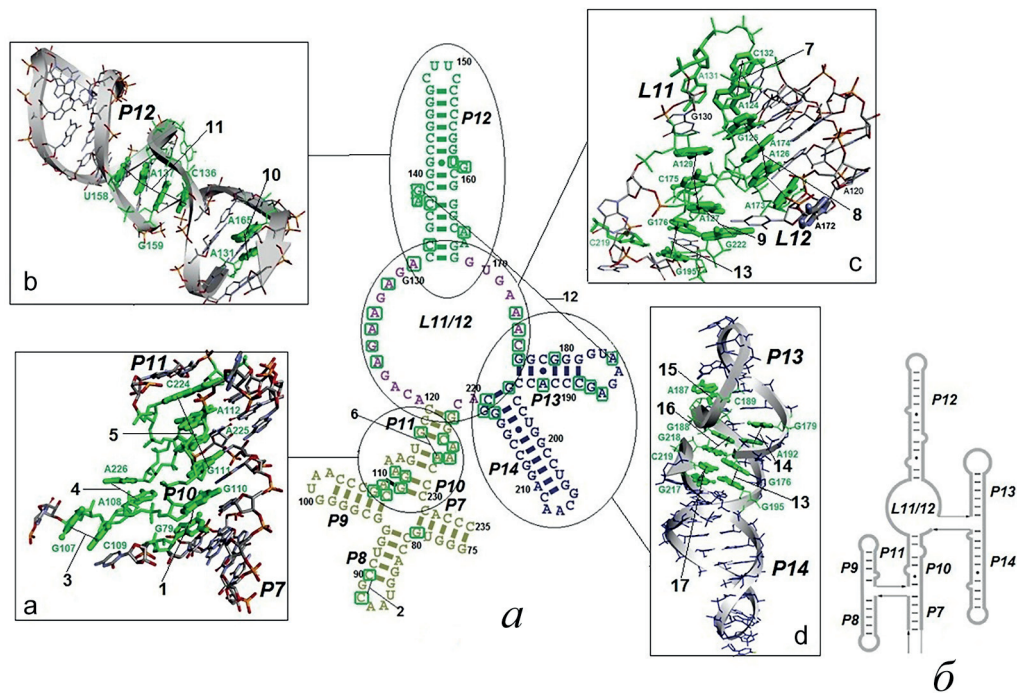


Рис. 4. Участие NA-BSE в организации третичной структуры РНКазы Р *T. thermophilus*. *a* – В центре – схема канонической вторичной структуры S-домена рибозима [17]; нуклеотиды, образующие NA-BSE, помещены в квадраты; во вставках справа и слева (*a*, *b*, *c*, *d*) – третичная структура фрагментов рибозима. *б* – Схема взаимного расположения спиральных участков РНК S-домена РНКазы Р в его третичной структуре [19]

NA-BSE, участвующие в формировании третичной структуры S-домена РНКазы Р *T. thermophilus* [17, 18]

№	NA-BSE	Тип	Роль в формировании третичной структуры РНК
1	G79.G110	II	коаксиальный стэкинг спиралей P7 и P10
2	C88.C90	I	компактная укладка петли с экспонированием A89
3	A107.C109	I	внутриспиральный стэкинг с экспонированием A108
4	A108.A226	LR	межспиральный контакт спиралей P9 и P10
5	G111.A225.A112.C224	LR-LR-LR	формирование межпетлевого NA-BSE-ZIP мотива
6	G117.A227	II	коаксиальный стэкинг спиралей P10 и P11
7	A124.C132	I	компактная укладка петли L11
8	G125.A174.A126.A.173	II-II-II	формирование внутриветлевого (L11/L12) NA-BSE-ZIP мотива
9	A129.C175.A127.G.222	II-II-II	формирование внутриветлевого (L11/L12) NA-BSE-ZIP мотива
10	A131.A165	LR	формирование компактной структуры основания спирали P12
11	C136.G159.A137.U158	LR-LR-LR	формирование NA-BSE-ZIP мотива в составе спирали P12
12	G138.A184	LR	формирование контакта между спиральями P12 и P13
13	G176.G195	II	коаксиальный стэкинг спиралей P13 и P14
14	G179.A192	II	сшивание двух цепей в «дефектном» участке спирали P13
15	A187.C189	I	компактная укладка петли в шпильке с экспонированием G188
16	G188.G218	LR	формирование межпетлевого контакта
17	G217.C219	I	внутриспиральный стэкинг с экспонированием G218

рую Krasilnikov et al. [17, 18] удалось получить в изолированном виде с максимально высоким (2,9 Å) разрешением для РНК этого класса (PDB ID 1U9S). Кроме того, S-домен в составе полноценной РНКазы Р, в отличие от С-домена, не связан ни с какими белками, модулирующими структуру этого рибозима, и можно считать, что его третичная структура в изолированном состоянии максимально приближена к внутриклеточной.

На рис. 4, а приведена вторичная структура S-домена РНКазы Р в канонической форме. Этот домен состоит из 161 н.о., ровно 100 из которых находятся в спиральных участках, образуя уотсон-криковские пары. Из оставшихся 61 н.о. 42 (т.е. примерно четверть от общего числа н.о. в этой РНК) входят в состав NA-BSE. Их нуклеотидный состав, типы структуры и возможная роль в формировании третичной структуры РНК приведены в таблице. Рассмотрим эти данные подробнее.

Как видно из рис. 4, б, восемь двуспиральных элементов канонической вторичной структуры S-домена РНКазы Р в её третичной структуре в действительности представлены четырьмя достаточно протяжёнными спиральями. При этом в трёх случаях (спирали P8-P9, P7-P10 и P13-P14) это осуществляется с помощью коаксиального стэкинга, одного из самых детально изученных способов организации макромолекулярной структуры РНК (для обзора см. работу Butcher et al. [3]). В этом процессе принимают участие три NA-BSE (№№ 1, 6 и 13 в таблице). Следует заметить, что участие этих мотивов в коаксиальном стэкинге спиралей вполне ожидаемо, т.к. он, как правило, осуществляется с помощью уотсон-криковских пар оснований, находящихся на концах спиральных элементов (см., например, стэкинг-контакт нуклеотидов G176 и G195, входящих в состав концевых уотсон-криковских пар в спиральях P13 и P14 соответственно; рис. 4, а, вставка «d»).

Далее, роль мотива № 14 (G179.A192 в таблице) в стабилизации спирали P13 также вполне очевидна: известно, что стэкинг пуриновых оснований н.о., расположенных в разных цепях несовершенного участка спирали, повышает устойчивость неуотсон-криковских пар оснований [20], в данном случае, пары A192-C178. Его следует назвать межцепочечным поперечным стэкингом (cross-strand stacking). Образующийся здесь NA-BSE и подобные ему мотивы в подавляющем большинстве случаев входят в состав E-мотивов, первый из которых был обнаружен в т.н. E-«петле» 5S рРНК [8].

Другой интересный пример участия NA-BSE в организации протяжённого спирального участка РНК — это образование блока из двух пиримидиновых и двух пуриновых нуклеотидов в спирали P12 (рис. 4, а, вставка «b», элемент 11; № 11 в таблице). Центральная ось этого блока фактически параллельна основной оси спирали P12. Важно, что в центре блока находятся «петлевые» (в терминах вторичной структуры) остатки A137 и G159, а на его концах — остатки C136 и U158, дополнительно образующие в спирали P12 уотсон-криковскую CG- и неканоническую UG-пары соответственно. При этом «петлевые» пуриновые остатки занимают место в центре блока с максимальным перекрытием плоскостей всех четырёх оснований. Расположенный по соседству G138 в формировании блока не участвует, но, образуя NA-BSE с A184 (№ 12 в таблице), стабилизирует контакт между спиральями P12 и P13. Сходным образом организованы блок, формирующий протяжённую структуру спиралей P10-P11 (№ 5 в таблице; рис. 4, а, вставка «a»), а также два блока, ответственных за достаточно компактную конформацию крупного петлевого участка S-домена L11/L12 (№№ 8 и 9 в таблице; рис. 4, а, вставка «c»). В целом, структура каждого такого блока напоминает застёжку «молния», и эти мотивы мы в дальнейшем обозначаем здесь как NA-BSE-ZIP (от англ. zipper).

Контакт между спиральями P9 и P10 стабилизируется также NA-BSE A108.A126 (№ 4 в таблице). При этом образование данного мотива становится возможным благодаря внутриспиральному стэкингу оснований н.о. A107 и C109, который происходит с экспонированием A108 (№ 3 в таблице; рис. 4, а, вставка «a»). Как уже отмечалось выше, такие стэкинг-взаимодействия важны не только для образования третичной структуры РНК (как в данном случае), но и их контактов с белками.

Спираль P12 не образует коаксиального стэкинг-контакта ни с одной из спиралей S-домена. Однако её сочленение с петлёй L11 организовано при помощи NA-BSE A131.A165, который расположен в основании этой спирали (№ 10 в таблице; рис. 4, а, вставка «b»). При этом он оказывается сближенным (но прямо не контактирует) с NA-BSE A124.C132, который делает более компактным один из участков большой «петли» L11 (№ 7 в таблице; рис. 4, а, вставка «c»).

NA-BSE (№№ 15–17 в таблице) образуют целый ансамбль, выполняющий разнообразные функции в организации третичной структуры S-домена РНКазы Р. Действительно, об-

разование NA-BSE первого типа (A187.C189, № 15 в таблице) в петле классической шпильки P13 сопровождается экспонированием G188, который вступает в стэкинг-взаимодействие с G218 в основании шпильки P14 и создаёт контакт между этими двумя спиральными участками, отчётливо видный на рис. 4, а (вставка «d»). В то же время у G218 появилась возможность образовать этот контакт благодаря стэкингу оснований в н.о. G217 и C219 (NA-BSE первого типа, № 17 в таблице).

Таким образом, анализируя третичную структуру сравнительно небольшой РНК, можно проследить разнообразие участия всех типов NA-BSE в формировании её макромолекулы. Они способствуют компактному сворачиванию однопетлевых петель РНК (а иногда и определяют его), участвуют в превращении двупетлевых петель в несовершенные спирали, играют важную роль в связывании удалённых в первичной и вторичной структуре районов РНК.

Мотивы типа NA-BSE-ZIP, столь широко представленные в S-доме РНКазы Р, встречаются также в тРНК, в рРНК в составе рибосом и в некоторых других РНК (см. табл. S1 Приложения). Более того, при анализе структур этих РНК видно, что они входят в довольно обширное подсемейство NA-BSE с характерным внедрением соседних в первичной структуре РНК н.о. в противоположные по направлению фосфодиэфирных связи участки её полинуклеотидной цепи. При этом основания каждого из этих остатков находятся в стэкинге с н.о., которые также являются соседями в первичной структуре РНК. То есть здесь мы

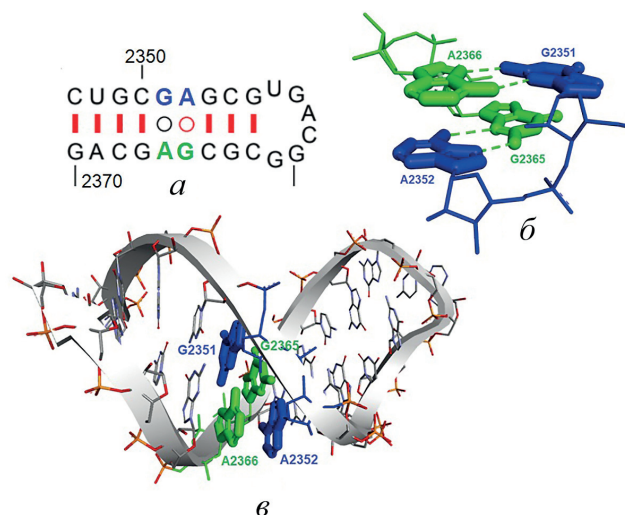


Рис. 5. Пример наиболее распространённого варианта NA-BSE-DC. а – Мотив-образующие GA-последовательности во вторичной структуре 23S рРНК *E. coli* (отмечены кружками). б – Пространственная структура мотива [PDB ID 7K00]. в – Положение мотива в третичной структуре фрагмента 23S рРНК в составе рибосомы

имеем дело с двойным поперечным межцепочечным стэкингом оснований. Мы обозначаем далее эти мотивы как NA-BSE-DC (double-crossing). Впервые NA-BSE-DC был замечен в кристаллической структуре другого рибозима, а именно в Р4-Р6 домене интрона группы I [9].

На рис. 5 приведён типичный пример NA-BSE-DC. В спирали H86 23S рРНК большой субъединицы рибосом *E. coli* он образован двумя динуклеотидными GA-сегментами этой рРНК, расположенными друг над другом в её канонической вторичной структуре (рис. 5, а).

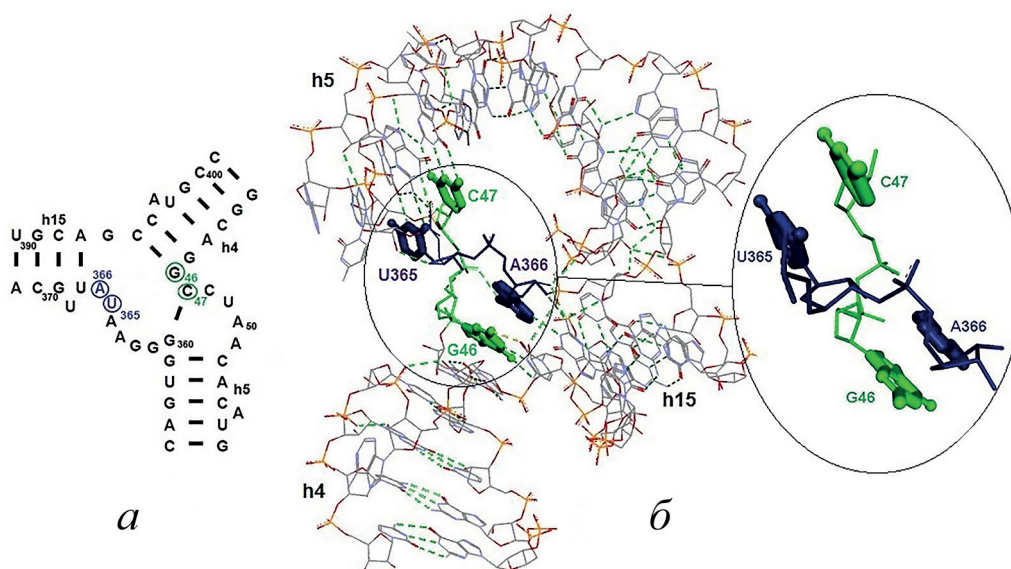


Рис. 6. Пример NA-BSE-DC с максимальным расстоянием между соседними основаниями в первичной структуре РНК. Показано расположение мотив-образующих н.о. во вторичной структуре (а) и в третичной структуре (б) 16S рРНК в составе рибосомы *E. coli* (PDB ID 7K00)

Поперечный стэкинг одноимённых пуриновых оснований позволяет им образовывать полноценные хугстиновские GA-пары (trans-Hoogsteen/SugarEdge пары оснований по номенклатуре Леонтиса–Вестхофа [21]). При этом общая форма спирали не искажается, и она лишь несколько изгибается в районе расположения NA-BSE-DC (рис. 5, б и в).

Другой пример NA-BSE-DC, иллюстрирующий участие мотивов этого подсемейства в формировании третичной структуры РНК (рис. 6), заметно отличается от предыдущего. Здесь непосредственные соседи в первичной структуре РНК располагаются в достаточно удалённых друг от друга спиральных участках, и их фосфодиэфирные связи служат для необычного соединения этих спиралей. При этом основания н.о., образующих NA-BSE-DC, спарены не друг с другом (как в предыдущем случае), а с функциональными группами н.о. спиралей, в которые они включены.

ЗАКЛЮЧЕНИЕ

Итак, NA-BSE способствуют компактному сворачиванию однотяжевых петель РНК (а иногда и определяют его), участвуют в превращении двутяжевых петель в несовершенные спирали, играют важную роль в связывании удалённых в первичной или вторичной структуре районов РНК (например, отдельных доменов рРНК или её сегментов, если рРНК транскрибируется в виде фрагментов).

В заключение отметим ещё одну потенциально важную черту рассмотренных здесь РНК-мотивов. Недавно Noller et al. [22], проанализировав 2000 нуклеотидных последовательностей рРНК малой субъединицы рибосом прокариот и цитоплазматических рибосом эукариот, установили, что примерно десятая часть н.о. этих РНК (140 н.о. в 16S рРНК *E. coli*) являются абсолютно консервативными. Анализируя все известные на сегодня экспериментальные данные, авторы смогли приписать с некоторой долей вероятности прямую или косвенную роль в функционировании рибосомы 82 консервативным н.о. 16S рРНК. Почему мутации гетероциклических оснований в

оставшихся 58 нуклеотидах этой РНК неблагоприятны, остаётся не ясным.

Сопоставляя результаты работы Noller et al. с данными, полученными в настоящей работе, можно высказать следующие соображения. Из 140 абсолютно консервативных остатков, выявленных этими авторами, 41 участвует в образовании NA-BSE. При этом 23 таким н.о. можно приписать определённую функциональную роль. Поиск ответа на вопрос, почему основания в оставшихся 18 н.о. не могут быть заменены на другие, с нашей точки зрения, принципиально важен. Предварительный анализ показывает, что NA-BSE, построенные из этих н.о., находятся внутри «головки» и «тела» малой субъединицы рибосомы, т.е. располагаются вне её функциональных центров. С другой стороны, в литературе накоплено большое число фактов, которые говорят о том, что и в рибосомах, и в других РНК-содержащих макромолекулярных комплексах происходит передача функциональных (конформационных) сигналов на большие расстояния [23–25]. Вполне логично предположить, что NA-BSE в рРНК, а также во многих других РНК участвуют в построении системы передачи таких сигналов.

Вклад авторов. Е.Б. выполнил вычислительную работу и конформационный анализ. В.М. проанализировал данные и подготовил все рисунки. А.Б. разработал проект. Все авторы внесли значительный вклад в написание рукописи.

Финансирование. В.М. и А.Б. благодарят за поддержку Московский государственный университет имени М.В. Ломоносова и Минобрнауки (Соглашение 1075-15-2021-1949 от 28.09.21).

Благодарности. Авторы благодарят рецензентов за внимательное прочтение рукописи и конструктивные предложения.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

Дополнительные материалы. Приложение к статье опубликовано на сайте журнала «Биохимия» (<https://biochemistrymoscow.com>).

СПИСОК ЛИТЕРАТУРЫ

1. Fresco, J. R., Alberts, B. M., and Doty, P. (1960) Some molecular details of the secondary structure of ribonucleic acids, *Nature*, **188**, 98–101, doi: 10.1038/188098a0.
2. Spirin, A. S. (1960) On macromolecular structure of native high-polymer ribonucleic acid in solution, *J. Mol. Biol.*, **2**, 436–446, doi: 10.1016/S0022-2836(60)80054-X.

3. Butcher, S. E., and Pyle, A. M. (2011) The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks, *Acc. Chem. Res.*, **44**, 1302-1311, doi: 10.1021/ar200098t.
4. Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B., and Steitz, T. A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif, *Proc. Natl. Acad. Sci. USA*, **98**, 4899-4903, doi: 10.1073/pnas.081082398.
5. Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2001) The kink-turn, *EMBO J.*, **20**, 4214-4221, doi: 10.1093/emboj/20.15.4214.
6. Chawla, M., Chermak, E., Zhang, O., Bujnicki, J. M., Oliva, R., and Cavallo, L. (2017) Occurrence and stability of lone pair–stacking interactions between ribose and nucleobases in functional RNAs, *Nucleic Acids Res.*, **45**, 11019-11032, doi: 10.1093/nar/gkx757.
7. Baulin, E., Metelev, V., and Bogdanov, A. (2020) Base-intercalated and base-wedged stacking elements in 3D-structure of RNA and RNA–protein complexes, *Nucleic Acids Res.*, **48**, 8675-8685, doi: 10.1093/nar/gkaa610.
8. Dallas, A., and Moore, P. B. (1997) The loop E–loop D region of *Escherichia coli* 5S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins, *Structure*, **5**, 1639-1653, doi: 10.1016/s0969-2126(97)00311-0.
9. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing, *Science*, **273**, 1678-1685, doi: 10.1126/science.273.5282.1678.
10. Teplova, M., Malinina, L., Darnell, J. C., Song, J., Lu, M., Abagyan, R., Musunuru, K., Teplov, A., Burley, S. K., Darnell, R. B., and Patel, D. J. (2011) Protein–RNA and protein–protein recognition by DualKH1/2 domains of the neuronal splicing factor Nova-1, *Structure*, **19**, 930-944, doi: 10.1016/j.str.2011.05.002.
11. Leontis, N. B., and Zirbel, C. L. (2012) Nonredundant 3D structure datasets for RNA knowledge extraction and benchmarking, *RNA 3D Structure Analysis and Prediction* (Leontis, N., and Westhof, E., eds), *Nucleic Acids and Molecular Biology*, **27**, 282-298, Springer-Verlag Berlin Heidelberg, doi: 10.1007/978-3-642-25740-7_13.
12. Lu, X.-J., Bussemaker, H. J., and Olson, W. K. (2015) DSSR: an integrated software tool for dissecting the spatial structure of RNA, *Nucleic Acids Res.*, **43**, e142, doi: 10.1093/nar/gkv716.
13. Shalybkova, A. A., Mikhailova, D. S., Kulakovskiy, I. V., Fakhranurova, L. I., and Baulin, E. F. (2021) Annotation of the local context of RNA secondary structure improves the classification and prediction of A-minors, *RNA*, **27**, 907-919, doi: 10.1261/rna.078535.120.
14. Yogesh, K., Gupta, Y. K., Nair, D. T., Wharton, R. P., Aggarwal, A. K. (2008) Structures of human Pumilio with noncognate RNAs reveal molecular mechanisms for binding promiscuity, *Structure*, **16**, 549-557, doi: 10.1016/j.str.2008.01.006.
15. Guogas, L. M., Filman, D. J., Hogle, J. M., and Gehrke, L. (2004) Cofolding organizes alfalfa mosaic virus RNA and coat protein for replication, *Science*, **306**, 2108-2111, doi: 10.1126/science.1103399.
16. Mondragón, A. (2013) Structural studies of RNase P, *Annu. Rev. Biophys.*, **42**, 537-557, doi: 10.1146/annurev-biophys-083012-130406.
17. Krasilnikov, A. S., Xiao, Y., Pan, T., and Mondragón, A. (2004) Basis for structural diversity in homologous RNAs, *Science*, **306**, 104-107, doi: 10.1126/science.1101489.
18. Krasilnikov, A. S., Yang, X., Pan, T., and Mondragón, A. (2003) Crystal structure of the specificity domain of ribonuclease P, *Nature*, **421**, 760-764, doi: 10.1038/nature01386.
19. Reiter, N. J., Osterman, A., Torres-Larios, A., Swinger, K. K., Pan, T., and Mondragón, A. (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA, *Nature*, **468**, 784-789, doi: 10.1038/nature09516.
20. Mignon, P., Loverix, S., Steyaert, J., and Geerlings, P. (2005) Influence of the π – π interaction on the hydrogen bonding capacity of stacked DNA/RNA bases, *Nucleic Acids Res.*, **33**, 1779-1789, doi: 10.1093/nar/gki317.
21. Leontis, N. B., and Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs, *RNA*, **7**, 499-512, doi: 10.1017/s1355838201002515.
22. Noller, H. F., Donohue, J. P., and Gutell, R. R. (2022) The universally conserved nucleotides of the small subunit ribosomal RNAs, *RNA*, **28**, 623-644, doi: 10.1261/rna.079019.121.
23. Sergiev, P. V., Kiparisov, S. V., Burakovsky, D. E., Lesnyak, D. V., Leonov, A. A., Bogdanov, A. A., and Dontsova, O. A. (2005) The conserved A-site finger of the 23S rRNA: just one of the intersubunit bridges or a part of the allosteric communication pathway? *J. Mol. Biol.*, **353**, 116-123, doi: 10.1016/j.jmb.2005.08.006.
24. Walkera, A. S., Russ, W. P., Ranganathanc, R., and Schepartz, A. (2020) RNA sectors and allosteric function within the ribosome, *Proc. Natl. Acad. Sci. USA*, **117**, 19879-19887, doi: 10.1073/pnas.1909634117.
25. Peselis, A., and Serganov, A. (2021) Cooperativity and allostery in RNA systems, *Methods Mol. Biol.*, **2253**, 255-271, doi: 10.1007/978-1-0716-1154-8_15.

MULTIPLE NON-CANONICAL BASE-STACKING INTERACTIONS AS ONE OF THE MAJOR DETERMINANTS OF RNA TERTIARY STRUCTURE ORGANIZATION

V. G. Metelev¹, E. F. Baulin², and A. A. Bogdanov^{1,3,4*}

¹ Faculty of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia

² Institute of Mathematical Problems of Biology of the Russian Academy of Sciences –
the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences,
142290 Pushchino, Moscow Region, Russia

³ A. N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University,
119992 Moscow, Russia; e-mail: bogdanov@belozersky.msu.ru

⁴ Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, 117997 Moscow, Russia

Interplane (stacking) interactions of heterocyclic bases of nucleotide residues (n.t.) of RNA are one of the most important factors in the organization of its secondary and tertiary structure. Most of these (canonical) interactions are carried out between neighbors in the polynucleotide chains of RNA. However, with the accumulation of data on the atomic tertiary structures of a wide variety of RNAs and their complexes with proteins, it became clear that RNA nucleotide residues that are not neighbors in their polynucleotide chains and are sometimes separated in the RNA primary structure by tens or hundreds of n.t. can interact with the help of base stacking (non-canonical). This paper presents an exhaustive database of such elements and their environment in the macromolecules of natural and synthetic RNAs. They were called nonadjacent base-stacking elements (NA-BSE). The analysis of these data showed that the NA-BSE forming nucleotides, on average, account for about a quarter of all nucleotides of a particular RNA, therefore, they should be considered as real motifs in their tertiary structure. The classification of NA-BSE by types of localization in RNA macromolecules is carried out. It is shown that the structure-forming role of NA-BSE consists in the compact folding of single-stranded RNA loops, in the transformation of double-stranded bulges into imperfect helices, as well as in the binding of RNA regions removed in their primary and secondary structure.

Keywords: stacking interactions, tertiary structure of RNA, ribosomal RNA, ribonuclease P, RNA motifs