

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 548.73

СТРУКТУРА БЕЛКА Hfq ИЗ *Chromobacterium haemolyticum* ВЫЯВИЛА НОВЫЙ ВАРИАНТ РЕГУЛЯЦИИ СВЯЗЫВАНИЯ РНК С БЕЛКОМ

© 2023 г. Н. В. Леконцева¹, А. Д. Никулин^{1,*}

¹Институт белка РАН, Пущино, Россия

*E-mail: nikulin@vega.protres.ru

Поступила в редакцию 27.07.2023 г.

После доработки 27.07.2023 г.

Принята к публикации 28.08.2023 г.

Определена структура белка Hfq из бактерии *Chromobacterium haemolyticum*, формирующая кристаллы в двух различных пространственных группах. В обоих случаях белок имеет характерную четвертичную структуру кольцевого гексамера. Полученная структура показала ранее не описанный вариант взаимодействия С-концевой неструктурированной части Hfq с аминокислотными остатками проксимального РНК-связывающего участка белка. Такой контакт может вносить вклад в регуляцию связывания молекул РНК с белком Hfq.

DOI: 10.31857/S0023476123700339, EDN: FVWDIA

ВВЕДЕНИЕ

Многие бактерии используют малые регуляторные РНК (мРНК) для эффективного отклика на экстремальные внешние воздействия и стрессы [1–3]. Они влияют на синтез многих факторов транскрипции и вирулентности, регулируя уровень трансляции и время жизни мРНК, что приводит к быстрому изменению метаболических процессов в клетках [4–7]. Существует особый класс белков, взаимодействующих с мРНК, расплетающих их вторичную структуру и способствующих взаимодействию мРНК с комплементарными участками мРНК [8–11]. Такие белки имеются как в грамотрицательных бактериях, так и в грамположительных, однако отсутствуют в термофильных и экстремально термофильных бактериях. Благодаря свойству этих белков разрушать вторичную структуру молекул РНК и способствовать приобретению ими необходимых для функционирования конформаций их часто называют РНК-шаперонами.

Наиболее характерным представителем РНК-шаперонов является белок Hfq [12]. Это высококонсервативный белок длиной от 80 до 100 аминокислотных остатков (а.о.), формирующий характерную четвертичную структуру в виде тороидальных гексамеров. Каждый мономер белка состоит из пятипятевого β -бочонка с предшествующей N-концевой α -спиралью. Эта основная часть белка включает в себя ~65 а.о. и составляет консервативный структурный мотив, относящийся к семейству LSM доменов (InterPro IPR005001). Последующие а.о. составляют струк-

турно неупорядоченную неконсервативную С-концевую часть переменной длины (С-концевой домен или СТД). В полученных ранее структурах белков Hfq из различных организмов эта область либо полностью не описана, либо имеет описание только небольшой части. α -fold также характеризует С-концевую часть белка как неструктурированную.

Hfq имеет три области связывания РНК: проксимальную, дистальную и латеральную [12–15]. Проксимальная область белка расположена в центральной поре гексамера со стороны α -спиралей и специфично связывает уридинсодержащие участки мРНК. Дистальная область Hfq расположена с противоположной стороны гексамера и специфически узнает полиадениновые 3'-участки мРНК или AAN-триплеты мРНК. Положительно заряженные латеральные участки связывания РНК на белке Hfq расположены на боковой поверхности каждого мономера в районе петли между N-концевой α -спиралью и первым β -тяжем. Они имеют повышенное сродство к аденин/уридинсодержащим участкам РНК и считаются местом встречи мРНК с комплементарными участками мРНК. В зависимости от пути взаимодействия мРНК с белком Hfq было предложено относить их к двум различным классам [16]. К первому относят мРНК, которые связываются с проксимальным и латеральными участками Hfq уридинсодержащими участками и взаимодействуют с мРНК, связанной с дистальной областью белка. Ко второму классу относят мРНК, для которых необходимо связывание с

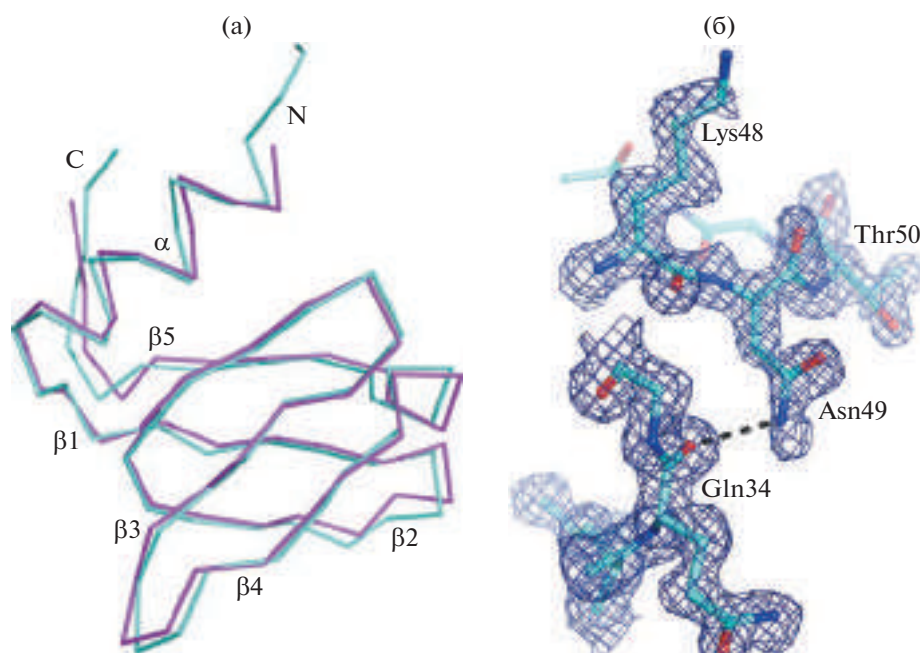


Рис. 1. Ленточное представление структуры мономера белка Hfq из *C. haemolyticum* и наложение на нее структуры мономера белка Hfq из *E. coli* (PDB 6BDG). Обозначены N- и C-концы белков, элементы вторичной структуры (а). Фрагмент участка полипептидной цепи ChaHfq в области аспарагина Asn49 и соответствующая карта электронной плотности $2mF_o - DF_c$ со срежкой 1σ . Водородная связь между атомом азота боковой цепи Asn49 и атомом кислорода главной цепи Gln34 показана пунктирной линией (б).

Pseudomonas aeruginosa (PDB 1U1S) в качестве исходной модели. Асимметричные части ячеек содержали мономер белка (пр. гр. $P6$) или гексамер белка (пр. гр. $P2_12_12_1$). Структуры уточнены с использованием программного комплекса PHENIX. Для проверки возможных погрешностей на структуру определяемого белка за счет влияния расчетных фаз с ориентированной исходной модели белка Hfq использовали карты электронной плотности *Composite Omit Map* с моделируемым отжигом структуры. Ручная правка и модификация модели осуществлялись с помощью программного комплекса COOT [29]. После завершения уточнения структур белка ChaHfq координаты структур и дифракционные данные депонировали в банк данных белковых структур PDB. Статистика сбора данных и кристаллографического уточнения представлена в табл. 1.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Структура белка ChaHfq в пр. гр. $P6$. Благодаря внутренней симметрии шестого порядка гексамер Hfq может кристаллизоваться в пр. гр. $P6$ с одним мономером в асимметричной ячейке. В банке данных PDB имеются 11 депонированных структур белка Hfq из разных организмов (PDB 5SZE, 2YLC, 5I2I, 4MML, 7OG8, 4PNO, 7OGW, 4RCB, 4RCC, 6BDG, 2Y90), имеющих пр. гр. $P6$. В полной ячейке кристалла содер-

жится гексамер, являющийся активной формой белка Hfq.

Мономер белка ChaHfq повторяет структуры гомологичных белков (рис. 1а). В электронную плотность были вписаны а.о. с 5 по 71. Первые четыре N-концевых и четырнадцать C-концевых остатков белка в модели отсутствуют, поскольку имеют высокую подвижность, а соответствующие участки карты электронной плотности не позволяют провести полипептидную цепь. Такое поведение концевых участков Hfq наблюдается во всех известных структурах этого белка, полученных с помощью рентгеноструктурного анализа. Наиболее интересным в структуре ChaHfq является хорошее качество электронной плотности для аспарагина Asn49 (рис. 1б), находящегося в петле между β -тяжами $\beta 3$ и $\beta 4$. Углы ϕ и ψ этого высококонсервативного аспарагина лежат в запрещенной области карты Рамачандрана, причем это характерно для всех известных структур белков Hfq. Исходя из полученной структуры можно предположить, что положение боковой цепи аспарагина фиксируется водородной связью с кислородом карбоксильной группой остатка глутамина Gln34. Других причин, стерических или функциональных, для сохранения Asn49 своей конформации не наблюдается.

Структура белка ChaHfq в пр. гр. $P2_12_12_1$. Асимметричная часть ячейки кристалла белка ChaHfq в этой пространственной группе содержит пол-

Таблица 1. Статистика обработки дифракционных данных и уточнения структур белка Hfq из *Chromobacterium haemolyticum*

	Гексамер в асимметричной части ячейки	Мономер в асимметричной части ячейки
Код PDB	8P9M	8P9I
Длина волны, Å	1.5418	1.5418
Диапазон разрешения, Å	22.2–2.0 (2.1–2.0)	21.1–1.4 (1.45–1.4)
Пр. гр.	$P2_12_12_1$	$P6$
Параметры ячейки, Å; град	$a = 58.5, b = 71.6, c = 96.6;$ $\alpha = \beta = \gamma = 90$	$a = b = 65.1, c = 27.7;$ $\alpha = \beta = 90, \gamma = 120$
Общее число рефлексов	53 153 (5300)	68 876 (4985)
Уникальных рефлексов	27 718 (2743)	13 424 (1343)
Избыточность	1.9 (1.9)	5.1 (3.7)
Полнота, %	98.4 (99.7)	99.8 (99.6)
Среднее отношение $I/\sigma(I)$	8.2 (0.9)	14.3 (1.6)
В-фактор Вильсона	27.5	7.3
R_{merge}	0.068 (0.76)	0.095 (0.65)
R_{meas}	0.096 (1.07)	0.105 (0.76)
R_{pim}	0.068 (0.76)	0.045 (0.39)
CC1/2	1.0 (0.58)	1.0 (0.48)
CC*	1.0 (0.86)	1.0 (0.80)
Число рефлексов, использованное для расчета R_{free}	1123 (119)	682 (64)
R_{work}	0.22 (0.33)	0.15 (0.22)
R_{free}	0.28 (0.36)	0.19 (0.30)
Число неводородных атомов	3535	661
белка	3281	545
лигандов	6	1
воды	248	115
Среднеквадратичные отклонения		
длин связей, Å	0.008	0.006
валентных углов, град	0.93	0.91
Положение остатков на карте Рамачандрана		
предпочтительные области, %	95.0	97.0
допустимые области, %	3.8	3.0
запрещенные области, %	1.2	0
Средний температурный В-фактор	32.75	9.16
для белка	32.35	7.21
для молекул воды	41.54	18.27

Данные в скобках соответствуют слою наиболее высокого разрешения.

ный гексамер. Структура основной части гексамера аналогична структурам гомологичных белков Hfq (рис. 2). Полученная электронная плотность позволила для большинства мономеров белка описать положение остатков с 5 по 72, а для двух мономеров удалось идентифицировать электронную плотность С-концевой части до 76 (цепь F)

и до 79 остатков (цепь В) из 85. Это так называемые области “С-концевого домена” или STD белка Hfq. Аминокислотные остатки 70–72 цепи F контактируют с остатками 49–50 (область петли $\beta 3$ – $\beta 4$) симметричной молекулы, что фиксирует их положение в ячейке кристалла и позволяет получить интерпретируемую электронную плотность.

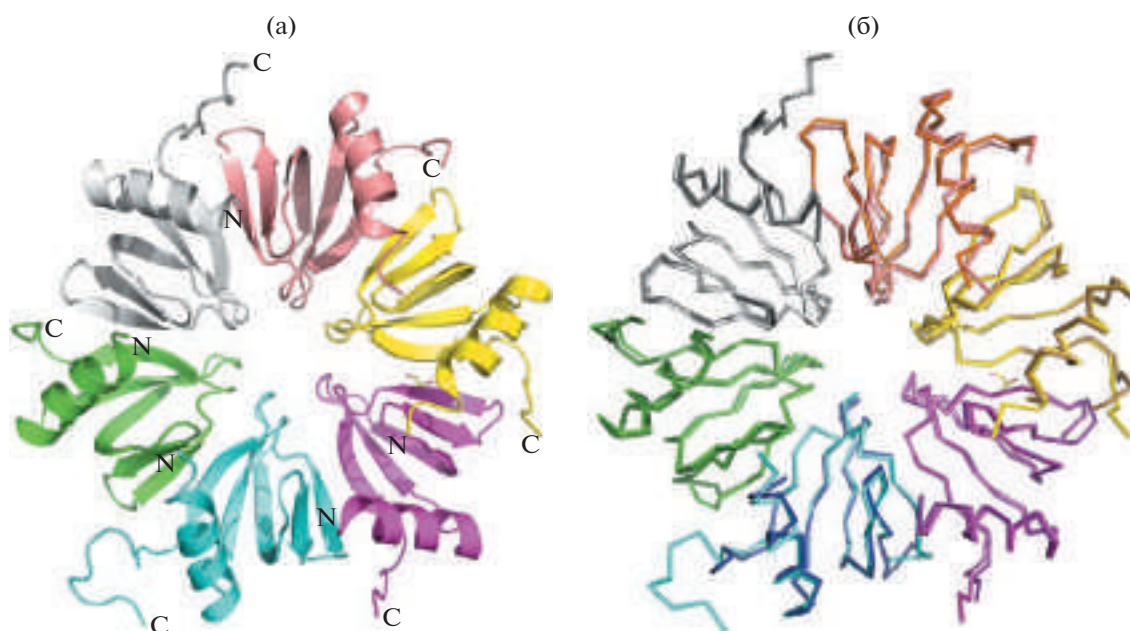


Рис. 2. Схематическое представление гексамера белка Hfq из *C. haemolyticum*. Вид со стороны α -спиралей белка (проксимальной стороны гексамера). Указаны N- и C-концы (а). Наложение главных цепей гексамера белка ChaHfq и гексамера белка Hfq из *Pseudomonas aeruginosa* (PDB 1UIS) (б).

Аминокислотные остатки С-концевой части цепи В контактируют с проксимальной областью симметричной молекулы белка, что позволило уверенно вписать остатки с 73 по 79 в экспериментальную электронную плотность (рис. 3а). Сравнение полученной структуры со структурами белка Hfq в комплексах с РНК показало, что остатки 73–79 занимают положение олиго(У) РНК в этих комплексах (рис. 3б, 3в). В комплексе Hfq-AU₆A РНК важную роль играет позиционирование оснований нуклеотидов между консервативными остатками тирозина/фенилаланина (Phe42 в Hfq из *E. coli*, Tyr43 в ChaHfq) с образованием их стэкинг-взаимодействия и формирования гидрофобных контактов [30]. В полученной структуре белка ChaHfq аминокислоты Pro72, Pro76 и His74 располагаются между боковыми цепями тирозинов и участвуют в гидрофобных контактах. Вместо стэкинга оснований нуклеотидов с ароматическими боковыми цепями имеют место σ – π -взаимодействия, такие как контакт атома кислорода боковой цепи Gln78 с ароматическим кольцом Tyr43 (рис. 3б). Для комплексов Hfq с РНК характерна развитая сеть водородных связей между боковыми цепями а.о. белка с атомами оснований и сахарофосфатного остова. В случае белка ChaHfq таких контактов всего четыре: между атомами кислорода главной цепи пролина Pro72 и гистидина His75 с боковой цепью глутамина Gln42, между боковой цепью глутамина Gln78 с карбоксильной группой глутамина Gln42 (рис. 3б), а также N ξ атомом Lys79 с кислородом

главной цепи остатка His58. В результате наблюдаемые контакты участка 73–79 ChaHfq не позволяют говорить о его значительном сродстве к проксимальной области белка Hfq, однако положение STD в проксимальном участке белка может составить некоторую конкуренцию связыванию РНК в этой области.

Текущая модель функционирования белка Hfq как регулятора трансляции посредством мРНК основана на одновременном (или последовательном) связывании с белком двух разных РНК: мРНК и мРНК [31]. При этом белок Hfq служит не только местом встречи двух молекул РНК с комплементарными участками, но и способствует разрушению (плавлению) их вторичной структуры, ускоряя формирование нового дуплекса мРНК–мРНК. Для его образования мРНК и мРНК первоначально связываются в разных областях белка — проксимальном и дистальном сайтах Hfq. Известны данные об участии STD белка Hfq из *E. coli* в регуляции связывания мРНК и их взаимодействия с мРНК [18, 19, 32–34]. Так, было показано, что С-концевой домен белка играет роль в специфическом узнавании мРНК белком Hfq и высвобождении сформированного дуплекса мРНК с мРНК, причем как *in vitro*, так и *in vivo* [18, 19]. Предполагается, что STD участвует в дискриминации РНК по специфичности к белку Hfq, контактируя с латеральным участком связывания основной части белка и не позволяя неспецифичным РНК связываться с белком [17]. В такой модели а.о. С-концевой части белка по-

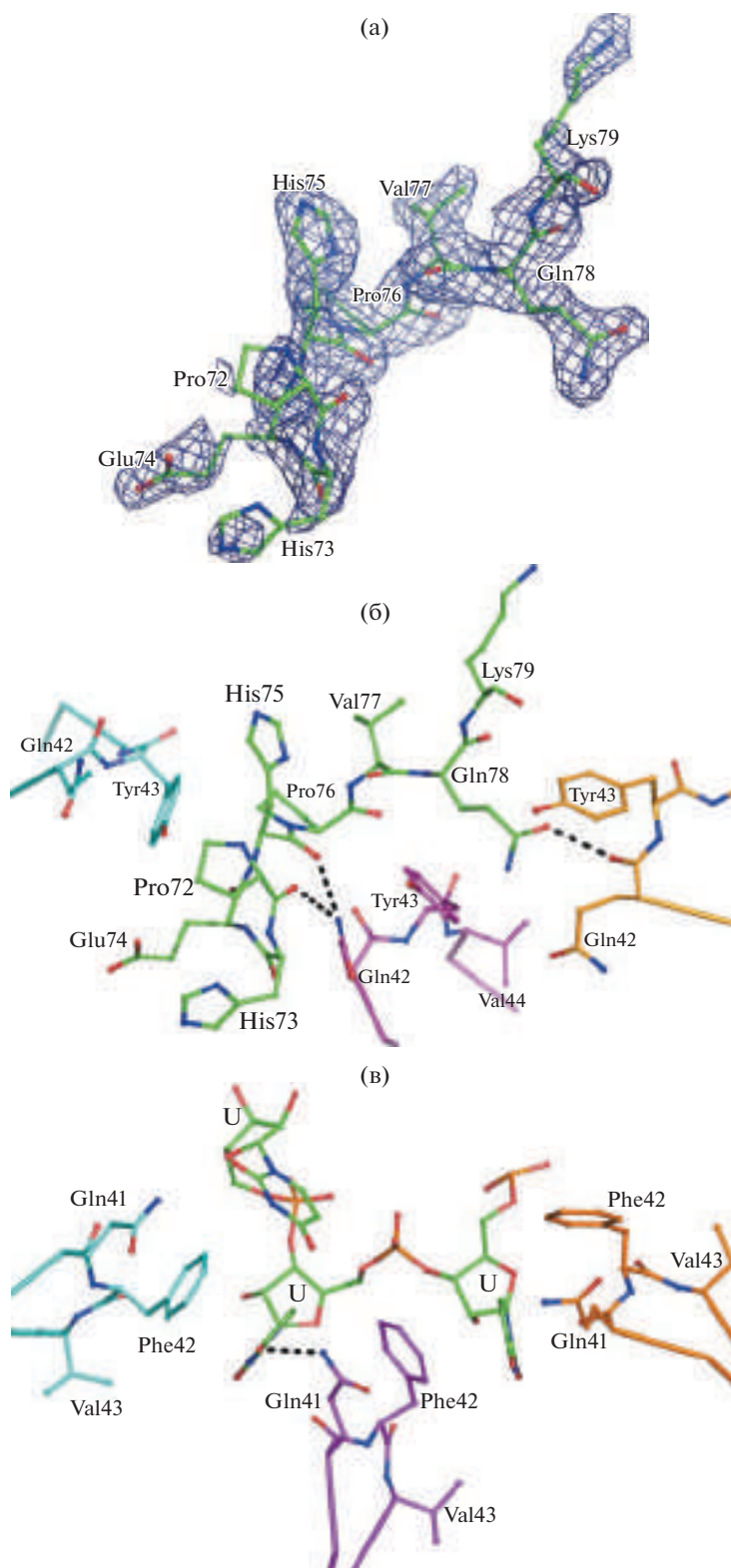


Рис. 3. Участок цепи В белка Hfq из *C. haemolyticum* с 73 по 79 аминокислотный остаток с соответствующим фрагментом карты электронной плотности $2mF_o - DF_c$ со срезкой 1σ (а). Контакты участка 73–79 с аминокислотными остатками симметрично связанной молекулы белка ChaHfq. Водородные связи показаны пунктирными линиями (б). Фрагмент структуры белка Hfq из *E. coli* в комплексе с AU₆A РНК (PDB 3RER). Шаростержневой моделью представлены три уридина из связанной молекулы РНК. Обозначены консервативные а.о. белка глутамин Gln41 (соответствует Gln42 белка ChaHfq), фенилаланина Phe42 (соответствует Tyr43) и валина Val43 (соответствует Val44) белка из *E. coli* (в) (ориентация участков белка Hfq из *E. coli* соответствует структуре ChaHfq панели б).

сле 72 взаимодействуют с остатками латеральной области гексамера и могут закрывать ее от связывания РНК. Структура белка Hfq из *C. haemolyticum* демонстрирует еще один вариант участия STD в регуляции связывания мРНК с белком. Гидрофобные области STD связываются с проксимальным участком белка, имеющим выраженную специфичность к олиго(У)-последовательностям РНК, ингибируя связывание Hfq с неспецифичными РНК. Однако данное предположение строится на полученной нами структуре белка Hfq из *C. haemolyticum* в комбинации с опубликованными ранее данными и требует экспериментальной проверки.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (госзадание АААА-А18-118120690135-9).

СПИСОК ЛИТЕРАТУРЫ

1. Jørgensen M.G., Pettersen J.S., Kallipolitis B.H. // Biochim. Biophys. Acta — Gene Regul. Mech. 2020. V. 1863. P. 194504. <https://doi.org/10.1016/j.bbagr.2020.194504>
2. Holmqvist E., Wagner G.H. // Biochem. Soc. Trans. 2017. V. 45. P. 1203. <https://doi.org/10.1042/BST20160363>
3. Dutta T., Srivastava S. // Gene. 2018. V. 656. P. 60. <https://doi.org/10.1016/j.gene.2018.02.068>
4. Wagner E.G.H., Romby P. // Adv. Genet. 2015. V. 90. P. 133. <https://doi.org/10.1016/bs.adgen.2015.05.001>
5. Pecoraro V., Rosina A., Polacek N. // Non-Coding RNA. 2022. V. 8. P. 22. <https://doi.org/10.3390/ncrna8020022>
6. Miyakoshi M. et al. // Mol. Microbiol. 2022. V. 117. P. 160. <https://doi.org/10.1111/mmi.14814>
7. Antoine L. et al. // Genes (Basel). 2021. V. 12. P. 1125. <https://doi.org/10.3390/genes12081125>
8. dos Santos R.F., Arraiano C.M., Andrade J.M. // Curr. Genet. 2019. V. 65. P. 1313. <https://doi.org/10.1007/s00294-019-00990-y>
9. Stenum T.S., Holmqvist E. // Mol. Microbiol. 2022. V. 117. P. 4. <https://doi.org/10.1111/mmi.14785>
10. Katsuya-Gaviria K. et al. // RNA Biol. 2022. V. 19. P. 419. <https://doi.org/10.1080/15476286.2022.2048565>
11. Woodson S.A., Panja S., Santiago-Frangos A. // Microbiol. Spectr. / Ed. Storz G., Papenfort K. 2018. V. 6. <https://doi.org/10.1128/microbiolspec.RWR-0026-2018>
12. Updegrove T.B., Zhang A., Storz G. // Curr. Opin. Microbiol. 2016. V. 30. P. 133. <https://doi.org/10.1016/j.mib.2016.02.003>
13. Murina V., Lekontseva N., Nikulin A. // Acta Cryst. D. 2013. V. 69. P. 1504. <https://doi.org/10.1107/S090744491301010X>
14. Park S. et al. // Elife. 2021. V. 10. P. 1. <https://doi.org/10.7554/eLife.64207>
15. Kavita K., de Mets F., Gottesman S. // Curr. Opin. Microbiol. 2018. V. 42. P. 53. <https://doi.org/10.1016/j.mib.2017.10.014>
16. Schu D.J. et al. // EMBO J. 2015. V. 34. P. 2557. <https://doi.org/10.15252/embj.201591569>
17. Santiago-Frangos A. et al. // Proc. Natl. Acad. Sci. U. S. A. 2019. V. 166. P. 10978. <https://doi.org/10.1073/pnas.1814428116>
18. Kavita K. et al. // Nucl. Acids Res. 2022. V. 50. P. 1718. <https://doi.org/10.1093/nar/gkac017>
19. Santiago-Frangos A. et al. // Proc. Natl. Acad. Sci. U. S. A. 2016. V. 113. P. E6089. <https://doi.org/10.1073/pnas.1613053113>
20. Han X.Y., Han F.S., Segal J. // Int. J. Syst. Evol. Microbiol. 2008. V. 58. P. 1398. <https://doi.org/10.1099/ijls.0.64681-0>
21. Lima-Bittencourt C.I. et al. // BMC Microbiol. 2007. V. 7. P. 58. <https://doi.org/10.1186/1471-2180-7-58>
22. Takenaka R. et al. // Jpn. J. Infect. Dis. 2015. V. 68. P. 526. <https://doi.org/10.7883/yoken.JJID.2014.285>
23. Okada M. et al. // BMC Infect. Dis. 2013. V. 13. P. 406. <https://doi.org/10.1186/1471-2334-13-406>
24. Tanpowpong P., Charoenmuang R., Apiwatanakul N. // Pediatr. Int. 2014. V. 56. P. 615. <https://doi.org/10.1111/ped.12301>
25. Teixeira P. et al. // Mol. Genet. Genomics. 2020. V. 295. P. 1001. <https://doi.org/10.1007/S00438-020-01676-8>
26. Winn M.D. et al. // Acta Cryst. D. 2011. V. 67. P. 235. <https://doi.org/10.1107/S0907444910045749>
27. McCoy A.J. et al. // J. Appl. Cryst. 2007. V. 40. P. 658. <https://doi.org/10.1107/S0021889807021206>
28. Afonine P. V et al. // Acta Cryst. D. 2012. V. 68. P. 352. <https://doi.org/10.1107/S0907444912001308>
29. Emsley P. et al. // Acta Cryst. D. 2010. V. 66. P. 486. <https://doi.org/10.1107/S0907444910007493>
30. Wang W. et al. // Genes Dev. 2011. V. 25. P. 2106. <https://doi.org/10.1101/gad.16746011.2004>
31. Santiago-Frangos A., Woodson S.A. // Wiley Interdiscip. Rev. RNA. 2018. V. 9. P. e1475. <https://doi.org/10.1002/wrna.1475>
32. Sonnleitner E. et al. // Biochem. Biophys. Res. Commun. 2004. V. 323. P. 1017. <https://doi.org/10.1016/j.bbrc.2004.08.190>
33. Vecerek B. et al. // Nucl. Acids Res. 2008. V. 36. P. 133. <https://doi.org/10.1093/nar/gkm985>
34. Panja S. et al. // J. Mol. Biol. 2015. V. 427. P. 3491. <https://doi.org/10.1016/j.jmb.2015.07.010>