

УДК 577.21

ПРОСТРАНСТВЕННАЯ СТРУКТУРА И РНК-СВЯЗЫВАЮЩИЕ СВОЙСТВА Lsm-БЕЛКА ИЗ *Halobacterium salinarum*

© 2021 М.С. Фандо, А.О. Михайлина, Н.В. Леконцева,
С.В. Тищенко, А.Д. Никулин*

Институт белка Российской академии наук, 142290 Пущино,
Московская обл., Россия; электронная почта: nikulin@vega.protres.ru

Поступила в редакцию 09.03.2021

После доработки 13.04.2021

Принята к публикации 13.04.2021

В настоящей работе определена структура и исследованы РНК-связывающие свойства Lsm-белка археи *Halobacterium salinarum*, отличительной особенностью которого является короткая петля L4, соединяющая тяжи $\beta 3$ и $\beta 4$. Поскольку бактериальные белки семейства Lsm, называемые Hfq, имеют короткую петлю L4 и образуют гексамеры, тогда как Lsm-белки архей, SmAP, имеют длинную петлю L4 и образуют гептамеры, было высказано предположение, что длина петли L4 может влиять на четвертичную структуру белков. Более того, петля L4 прикрывает область SmAP, соответствующую одному из сайтов связывания РНК в белках Hfq и, таким образом, может влиять на РНК-связывающие свойства белка. Наши результаты показывают, что SmAP из *H. salinarum* формирует гептамеры и обладает теми же РНК-связывающими свойствами, что и гомологичные белки с длинной петлей L4. Следовательно, длина L4 не влияет на количество мономеров в Lsm-белках и на их РНК-связывающие свойства.

КЛЮЧЕВЫЕ СЛОВА: Lsm-белки, SmAP, четвертичная структура белков, РНК-связывающие свойства белков, малые регуляторные РНК, мРНК, *Halobacterium salinarum*.

DOI: 10.31857/S0320972521070034

ВВЕДЕНИЕ

Белки архей SmAP (Sm Archaeal Protein) относятся к семейству Lsm (Sm-like) белков, представители которого имеют характерную пространственную структуру пятитяжевого β -листа с N-концевой α -спиралью [1–3]. Тяжи $\beta 1$, $\beta 2$ и $\beta 3$ составляют структурный мотив Sm1, а тяжи $\beta 4$ и $\beta 5$ – мотив Sm2. Эти участки белка соединены петлей L4, длина которой варьирует в Lsm-белках из разных организмов [4]. Lsm-белки формируют четвертичную структуру в виде тороидальных колец, как правило, из 6 или 7 мономеров. Бактериальные Lsm-белки, называемые Hfq, образуют стабильные в растворе гомогексамеры [5]. Эукариотические Lsm-белки характеризуются наличием ряда парало- гов в различных клетках организма; они формируют гетерологичные гептамеры (Lsm1–7 и Lsm2–8) [6–8]. В то же время Sm-белки эукари- от организмов формируют гетерогептамеры только в

присутствии специфически связываемых малых ядерных РНК и выполняют роль каркасных белков малых ядерных рибонуклеопротеинов (мяРНП) [2, 9, 10]. Архейные белки семейства Lsm имеют от одного до трех парало- гов (SmAP1, SmAP2, SmAP3) в зависимости от организма и существуют в виде гомологичных гептамеров, независимо от присутствия РНК [2]. Тем не менее существуют примеры архейных белков, которые имеют другую степень олигомеризации, например, SmAP-белок из *Methanococcus (Methanocaldococcus) jannaschii* формирует гомогексамер и описывается как Hfq-подобный белок [11, 12], а SmAP2 из *Archaeoglobus fulgidus* (AF-Sm2) может формировать как гексамеры, так и гептамеры в зависимости от присутствия РНК и pH среды [13]. Полагают, что зависимость структуры от кислотности среды обусловлена изменением зарядов контактирующих в межмономерном пространстве аминокислотных остатков глутаминовой кислоты (Glu) [14]. Несмотря на значительный объем накопленных экспериментальных данных о природе Lsm-белков, механизм формирования четвертичной структуры с различным числом мономеров до сих пор не установлен. На основании сравнения первичных и четвертичных структур Lsm-бел-

Принятые сокращения: ДСН – додецилсульфат натрия; ДТТ – дитиотреитол; ПААГ – полиакриламидный гель; Mant-AMP – 2'/3'-(N-метил-антраилоил)-аденозин-5'-монофосфат; Mant-GMP – 2'/3'-(N-метил-антраилоил)-гуанозин-5'-монофосфат.

* Адресат для корреспонденции.

ков из бактерий и архей/эукариот была выдвинута гипотеза о взаимосвязи длины петли L4 и степени олигомеризации Lsm-белков [4], однако ограниченность структурных данных касательно белков с различной длиной петли не позволяла подтвердить или опровергнуть её. Установление гексамерной структуры Hfq-подобного SmAP-белка из археи *M. jannaschii* с короткой петлей L4 вновь заставило исследователей задуматься о влиянии этой петли на олигомеризацию белка. Следует, однако, учитывать, что аминокислотная последовательность этого белка ближе к гексамерным белкам Hfq, поэтому указанный пример не может служить основанием для оценки выдвинутой гипотезы. Для прояснения этого вопроса в представленной работе нами определена и проанализирована структура SmAP-белка из галофильной археи *Halobacterium salinarum*, в котором петля L4 является короткой и составляет всего 3 аминокислотных остатка (а.о.).

Недавно нами была продемонстрирована высокая специфичность ряда SmAP-белков к уридинмонофосфату (UMP) и олиго(U)-РНК при низком сродстве этих белков к другим рибонуклеотидам [15]. Сравнение структур белков SmAP и их бактериальных гомологов Hfq в комбинации с биохимическими данными позволило предположить, что белки SmAP имеют только один, универсально консервативный уридин-специфический сайт, характерный для всего семейства Lsm. Это значительно отличает белки SmAP от бактериальных белков Hfq, у которых существенную роль в функционировании играет второй, аденин-специфический РНК-связывающий сайт. В SmAP-белках этот регион прикрыт от внешних контактов протяжённой петлей L4, к тому же, структура соответствующего участка в архейных белках значительно отличается от бактериальных. В белке SmAP из *H. salinarum* петля L4 имеет минимальную длину и не мешает доступу к участку, соответствующему аденин-связывающему сайту белка Hfq. Измерение сродства такого SmAP-белка к олиго(A)-РНК позволит оценить влияние длины петли L4 на функциональность этого участка белка. Таким образом, определение структуры и измерение РНК-связывающих свойств белка SmAP из *H. salinarum* позволит оценить влияние петли L4 на структуру и функцию архейных белков семейства Lsm.

МАТЕРИАЛЫ И МЕТОДЫ

Клонирование генов белка SmAP из *H. salinarum*. В базе данных NCBI имеется информа-

ция об аминокислотной последовательности белка SmAP из *H. salinarum* длиной 60 а.о. (<https://www.ncbi.nlm.nih.gov/protein/167727293>), а в БД UniProt – протяженностью 69 а.о. (<https://www.uniprot.org/uniprot/Q9HPS2>). В работе использовались оба варианта белка SmAP. Далее вариант длиной 60 а.о. будет обозначаться как HsaSmAP60, а длиной 69 а.о. – HsaSmAP69. Гены обоих вариантов нарабатывали с помощью метода полимеразной цепной реакции (ПЦР) с использованием хромосомной ДНК археи *H. salinarum* R1 в качестве матрицы. ПЦР проводили с использованием следующих олигонуклеотидных праймеров:

HsaFor 5'-CATCACCATGGGCGGCCGACC-3' и HsaRev 5'-GGGATAAGCTTTCATGGTTTGATG-GTGACG-3' для HsaSmAP60;

HsaFor 5'-CCATGGATGCCACCACCGG-3' и HsaRev 5'-GGGATAAGCTTTCATGGTTTGATG-GTGACG-3' для HsaSmAP69.

Полученные ПЦР-фрагменты клонировали в экспрессионный вектор pProExHtb («Invitrogen», США) с использованием эндонуклеаз рестрикции Bsp19I и HindIII («СибЭнзим», Россия). Полученные конструкции проверяли секвенированием («Синтол», Россия) и трансформировали ими клетки *E. coli* BL21(DE3), дополнительно содержащими плазмиду pLacIRARE («Novagen», Индия), кодирующую пять тРНК, узнающих редкие для *E. coli* кодоны соответствующих аминокислот.

Выделение и очистка белков. Клетки культивировали в колбах емкостью 2 л при температуре 37 °С и интенсивном перемешивании (170 об./мин) на среде LB (Lysogeny broth), содержащей 100 мкг/мл ампициллина и 10 мкг/мл хлорамфеникола до достижения значения оптической плотности $OP_{600} \approx 0,7$ о.е. Для активации T7 РНК-полимеразы в среду добавляли изопропил-β-D-1-тиогалактопиранозид (ИПТГ) до конечной концентрации 0,5 мМ. После добавления индуктора клетки инкубировали при 20 °С в течение 12 ч для предотвращения агрегации нарабатываемых белков. Клетки собирали центрифугированием при 8 000 g в течение 20 мин при 4 °С, затем клеточную массу ресуспендировали в 30 мл лизирующего буфера (20 мМ Tris-HCl, pH 8,0, 0,5 М NaCl, 1 мМ фенолметилсульфонил фторид (PMSF), 1 мМ дитиотреитол (ДТТ), 0,1% Triton X-100) и разрушали на проточном прессе EmulsiFlex-C3 («Avestin», Канада). Клеточный дебрис осаждали центрифугированием в течение 30 мин при 14 000 g. Супернатант прогревали при 70 °С в течение

10 мин. Денатурировавшие термолабильные белки *E. coli* осаждали низкоскоростным центрифугированием при 14 000 g и температуре +4 °C в течение 30 мин. Супернатант наносили на колонку с агарозой, модифицированной Ni-NTA («Qiagen», Германия), уравновешенную буфером, содержащим 20 mM Tris-HCl, pH 8,0, 0,5 M NaCl, и промывали тем же буфером с добавлением 1 M LiCl. Для элюции белка использовали линейный градиент имидазола от 0 до 300 mM. Удаление олигогистидиновой последовательности осуществляли инкубацией препарата белка с TEV-протеазой в соотношении 10 : 1 соответственно, переводя диализом в буфер, содержащий 0,3 M NaCl, 20 mM Tris-HCl, pH 8,0, 1 mM ДТТ, 0,5 mM ЭДТА. После протеолиза раствор прогревали в течение 10 мин при 70 °C, денатурировавшую протеазу осаждали центрифугированием при 14 000 g в течение 30 мин при температуре +4 °C. Разделение белков по наличию олигогистидиновой концевой последовательности осуществляли путём нанесения смеси белков на колонку с агарозой, модифицированной Ni-NTA («Invitrogen»). На финальной стадии целевые белки очищали гель-фильтрацией на колонке Superdex 75 («Cytiva», США), уравновешенной буфером, содержащим 20 mM Tris-HCl, pH_{25 °C} 8,0, 300 mM NaCl. Полученные препараты белков хранили при -20 °C.

Выделение антител, специфичных к белку SmAP из *H. salinarum*. Сыворотка крови кролика, содержащая антитела, специфичные к белку SmAP из *H. salinarum* (AT_{SmAP}), была получена в Институте биофизики клетки РАН (Пушино, Россия). Белки сыворотки кролика осаждали 2 M сульфатом аммония и отделяли центрифугированием при 20 000 g и 4 °C в течение 15 мин. Далее сыворотку очищали с помощью двух последовательных хроматографий. На первом этапе использовали аффинную хроматографию на Protein-A Sepharose («Pharmacia LKB», США). Сыворотку наносили на смолу в буфере PBS (137 mM NaCl, 2,7 mM KCl, 10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, pH 7,4) и промывали тем же буфером. Элюцию суммарных антител проводили глициновым буфером с концентрацией 250 mM, pH 3,0. Затем антитела, специфичные к белку SmAP из *H. salinarum*, очищали на смоле из агарозы, модифицированной Ni-NTA, с иммобилизованным на ней рекомбинантным белком SmAP, содержащим 6 остатков гистидина на N-конце. Суммарные очищенные антитела наносили на Ni-NTA-SmAP в буфере PBS, промывали буфером PBS с 2 M LiCl. Элюцию AT_{SmAP} проводили градиентом буфера PBS с содержанием MgCl₂ от 1 до 4 M. Фракции, содержащие чистые AT_{SmAP}, не загрязненные белком SmAP,

хранили в буфере PBS с 50% глицерина при температуре -20 °C.

Анализ вестерн-блот. Фракции белков разделяли при помощи электрофореза с 14%-ным ПААГ-ДСН, после чего переносили на PVDF-мембрану («Bio-Rad», США) с использованием Trans-Blot Turbo Blotting System того же производителя (1 A, 25 B, 30 мин). Блокировали мембрану в буфере TBS (50 mM Tris-HCl, pH 7,5, 150 mM NaCl). Далее обрабатывали мембрану буфером TBSTM (TBS pH 7,5, 0,05% Tween 20, 5%-ное обезжиренное сухое молоко) в течение 1 ч при комнатной температуре. Затем мембрану инкубировали с антисывороткой кролика, разведенной буфером TBST (TBS pH 7,5, 0,05% Tween 20) (1 : 2000) в течение 1 ч при комнатной температуре. После трехкратного промывания (по 10 мин каждое) в TBST мембрану инкубировали в течение 1 ч при комнатной температуре с козьими антителами к IgG кролика, конъюгированными с пероксидазой хрена (Goat Anti-Rabbit IgG H&L (HRP) (ab205718); «Abcam», Великобритания) (1 : 20 000). Промывали мембрану в буфере TBST 2 раза по 5 мин, затем 1 раз в буфере TBS в течение 2 мин. Инкубировали мембрану с хемилюминесцентным субстратом (Clarity™ Western ECL Substrate, «Bio-Rad») в темноте в течение 5 мин. Анализ проводили с использованием системы детекции ChemiDoc MP Imaging System («Bio-Rad»).

Анализ сродства белков к РНК по изменению подвижности в ПААГ при электрофорезе в неденатурирующих условиях. РНК-олигонуклеотиды синтезируются химически и предоставлены фирмой «Синтол» (Россия). В работе использовали РНК-олигонуклеотиды с 18 уридинами или аденинами на 3'-конце. На 5'-конце находится 16 нуклеотидов, не имеющих специфического сродства к белкам SmAP/Hfq [15]; они служат спейсерами при иммобилизации олигонуклеотидов на чипы в экспериментах по анализу сродства исследуемых белков к РНК методом поверхностного плазмонного резонанса (ППР, см. ниже):

U18-РНК: 5'-GUGGUCAGUCGAGUGG-(U)₁₈-3'

A18-РНК: 5'-GUGGUCAGUCGAGUGG-(A)₁₈-3'.

Для оценки сродства белков к РНК по изменению подвижности олигонуклеотидов в ПААГ в неденатурирующих условиях образец РНК (3 мкг) прогревали в течение 10 мин при 60 °C, добавляли белок в различных молярных соотношениях, затем полученную смесь инкубировали при 37 °C в течение 10 мин. После этого проводили электрофоретический анализ в 10%-ном

ПААГ (акриламид : метилен-бис-акриламид = 19 : 1) в 90 мМ трис-ацетатном буфере (ТАЕ) с рН 7,8, содержащем 2 мМ $MgCl_2$. В образцы перед нанесением добавляли глицерин и бромфеноловый синий до конечных концентраций 5 и 0,1% соответственно. Электродный буфер имел следующий состав: 90 мМ ТАЕ (рН 7,8), 10 мМ $MgCl_2$. После окончания электрофореза РНК в геле окрашивали раствором, содержащим 0,25%-ный толуидиновый синий, 5%-ную уксусную кислоту, 10%-ный этанол.

Анализ сродства исследуемых белков к РНК методом поверхностного плазмонного резонанса.

Анализ осуществляли с использованием олигонуклеотидов А18-РНК и U18-РНК, биотинилированных по 5'-концу. Связывание белка с целевой РНК проводили при 25 °С с использованием системы ProteOn XPR36 («Bio-Rad»). Биотинилированные РНК прогревали 5 мин при 95 °С и наносили со скоростью потока 30 мкл/мин на сенсорный чип ProteOn NLC с иммобилизованным авидином до заполнения поверхности чипа, характеризующегося значением 500–700 единиц ответа (response unit, RU). Поверхность чипа дважды промывали восстанавливающим раствором, содержащим 0,05% w/v додецилсульфат натрия для удаления не связавшейся РНК.

Для каждого набора сенсограмм были приготовлены серии разведений белков в буфере, содержащем 50 мМ Tris-HCl (рН 7,5), 10 мМ $MgCl_2$, 350 мМ NaCl, 2 мМ ДТТ с 0,05% v/v Tween 20 буфером (TMN). Нанесение образца включает фазу ассоциации длиной 300–500 сек и фазу диссоциации 700–3000 сек. В конце каждого цикла поверхность чипа промывали 0,1–0,2%-ным ДСН, чтобы добиться полной диссоциации белка с поверхности РНК. Кинетический анализ проводили при использовании набора из 4–5 сенсограмм с применением модели для одностадийной реакции (по Ленгмюру) с коррекцией на масс-транспорт. Были определены константы скоростей ассоциации (k_a) и диссоциации (k_d) комплексов и рассчитана равновесная константа диссоциации ($K_D = k_d/k_a$).

Анализ сродства исследуемых белков к флуоресцентно-меченым рибонуклеотидам по анизотропии флуоресценции. Эксперименты проводили на флуоресцентном спектрометре Cary Eclipse («Agilent», США) при температуре 22 °С в кювете 0,3 × 0,3 см согласно описанной ранее методике [16, 17]. Для проведения измерений раствор Mant-AMP или Mant-GMP («Jena Bioscience», Германия) с концентрацией 1 мМ в буфере состава 0,2 М NaCl, 10 мМ $MgCl_2$, 50 мМ Tris-HCl (рН 8,0) титровали раствором белка с увеличением его концентрации от 10^{-6} до 10^{-4} М. Ани-

зотропию флуоресценции измеряли на длине волны 448 нм при возбуждении УФ на длине волны 335 нм. Для каждой точки проводили 5 независимых измерений величины интенсивности флуоресценции и затем автоматически рассчитывали среднюю величину и погрешность в программе QtiPlot («Iondev Srl», Румыния). Результат отображали графически с помощью программы QtiPlot.

Кристаллизация белка SmAP из *H. salinarum*.

Для экспериментов по кристаллизации использовали метод диффузии паров в висящей капле на 24-луночных плашках (VDX greased plates, «Hampton Research», США). Кристаллизацию проводили в каплях объемом 2–3 мкл, помещенных на силиконизированные стеклянные пластины («Hampton Research») при 24 °С. Объем противораствора составлял 400 мкл.

Кристаллы белка HsaSmAP69, пригодные для рентгеноструктурного анализа, получены не были.

Кристаллы белка HsaSmAP60 были получены в условиях D9 (15% w/v ПЭГ 8000; 100 мМ MES, рН 6,5; 200 мМ ацетат кальция) набора Nuc-Pro 2 («Jena Bioscience») при соотношении белок : противораствор = 2 : 1 и концентрации белка в образце 9,8 мг/мл. Кристаллы вырастали в течение недели до размера 100–150 мкм в длину и 20–30 мкм в ширину. Кристаллы комплекса HsaSmAP60 с UMP были получены инкубацией кристаллов белка в криопротекторном растворе (0,17 М ацетат натрия, 0,085 М какодилат натрия, рН 6,5, 25,5% ПЭГ 8000, 15% глицерин) с добавлением рибонуклеотида в концентрации 20 мМ.

Получение наборов дифракционных данных. С кристаллов белка HsaSmAP60 в комплексе с UMP были собраны дифракционные данные на экспериментальной станции ID23-1 источника синхротронного излучения ESRF (Гренобль, Франция) (таблица). Первичную обработку данных проводили с помощью программы XDS (Институт медицинских исследований им. Макса Планка, Гейдельберг, Германия) [18], затем набор структурных факторов интегрировали и приводили к единой шкале в программе Aimless комплекса CCP4 (Лаборатория молекулярной биологии MRC, Кэмбридж, Великобритания) [19].

Определение и уточнение структуры белка.

Структуру белка HsaSmAP60 в комплексе с UMP определяли методом молекулярного замещения с использованием программы MoRDa [20] комплекса CCP4 (<https://ccp4online.ccp4.ac.uk/ccp4online/>). Автоматический поиск наиболее подходящей модели для решения структуры показал наилучший результат для мономера белка AF-Sm2 из *Archaeoglobus fulgidus* (PDB

ID: 1jjo). Уточнение модели проводили в программе *phenix.refine* [21] комплекса Phenix (Лаборатория молекулярной биологии MRC) [22] с последовательным использованием протоколов *rigid body refinement* (уточнение положения каждого мономера белка как отдельного твердого тела) и *simulated annealing – Cartesian* (уточнение координат индивидуальных атомов и индивидуальных В-факторов с использованием симуляции молекулярной динамики). После первого цикла уточнения использовали возможность автоматического дополнения модели связанными молекулами воды с помощью программы *phenix.refine*. На этом шаге также дополняли незаполненные участки электронной плотности аминокислотными остатками согласно аминокислотной последовательности. В электронной плотности, не принадлежащей полипептидным цепям белка, было идентифицировано 14 молекул UMP, 1 молекула глицерина и 28 ионов кальция. Для них проведено уточнение индивидуальных координат, В-факторов, а также величин замещения (*occuранcy*). После каждого этапа уточнения соответствие структуры и расчетной карты электронной плотности проверяли в программе *Coot* (Лаборатория молекулярной биологии MRC) [23]. На финальных стадиях уточнения использовали протокол без применения симуляции отжига с проверкой соответствия параметров геометрии полипептидной цепи и лигандов идеальным значениям, что позволило получить структуру высокого качества с минимальными отклонениями по геометрии и низкими значениями R-фактора (таблица).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Идентификация варианта белка SmAP археи *H. salinarum*. В ходе поиска последовательности белка SmAP из *H. salinarum*, а также кодирующего его гена, в базах данных NCBI и UniProt были обнаружены две записи с отличающейся длиной N-концевой части белка на 9 аминокислотных остатков (рис. 1). Оба варианта гена были клонированы, затем белки наработаны в *E. coli* – HsaSmAP60 (60 а.о.) и HsaSmAP69 (69 а.о.). Отличие в первичной структуре белков не сказалось на их свойствах при выделении и очистке. С помощью анализа вестерн-блот нами было показано, что в клетках *H. salinarum* присутствует белок длиной 69, а не 60 а.о. (рис. 2).

РНК-связывающие свойства SmAP из *H. salinarum*. Сродство белков к РНК-олигонуклеотидам оценивали двумя методами: по изменению подвижности РНК-белкового комплекса при электрофорезе в ПААГ в неденатурирующих ус-

Статистика набора дифракционных данных и уточнения структуры белка SmAP из *H. salinarum* (HsaSmAP60) в комплексе с UMP

Характеристика	Значение
Код PDB	6tfl
Длина волны (Å)	0.9724
Область разрешения (Å)	50.0-2.4 (2.5-2.4)
Пространственная группа	P 41 21 2
Параметры ячейки (Å, °)	156.45; 156.45; 106.63; 90; 90; 90
Общее число рефлексов	1318697 (82803)
Число уникальных рефлексов	52360 (5118)
Избыточность	25.2 (16.0)
Полнота данных (%)	99.8 (98.4)
Среднее значение I/Σ(I)	7.3 (2.3)
В-фактор Вильсона	42.7
R-merge	0.348 (0.908)
R-meas	0.355 (0.937)
R-pim	0.069 (0.232)
CC1/2	0.99 (0.89)
CC	1.00 (0.97)
Число рефлексов, использованных для уточнения	52247 (5101)
Число рефлексов, использованных для расчета R-free	2700 (279)
R-work	0.178 (0.270)
R-free	0.230 (0.346)
CC(work)	0.94 (0.72)
CC(free)	0.91 (0.65)
Число атомов общее	7187
В молекулах белка и нуклеотидах	6841
В лигандах	34
В молекулах воды	312
СКО длин связей	0.007
СКО валентных углов	0.818
СКО двугранных углов	24.76
СКО хиральных углов	0.059
СКО планарных углов	0.004
Число остатков в областях карты Рамачандрана	
Предпочтительной (%)	98.57
Допустимой (%)	1.43
Запрещенной (%)	0.00
Число нежелательных ротамеров (%)	0.00
Оценка нежелательных контактов (clashscore)	5.61
Средний В-фактор атомов (общий)	47.50
Белка	47.40
Лигандов	73.52
Воды	49.21

Примечание. Данные в скобках соответствуют слою высокого разрешения 2,5–2,4 Å.

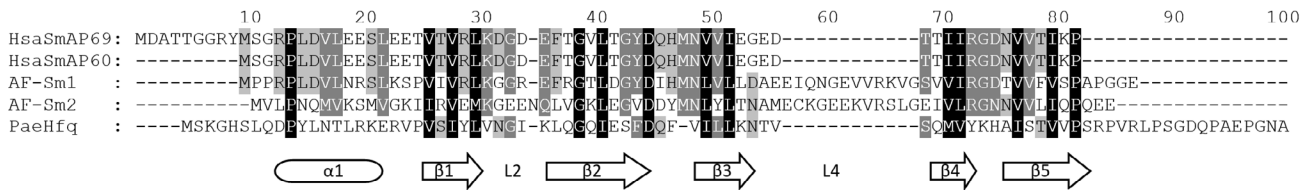


Рис. 1. Сравнение последовательностей мономеров белков HsaSmAP69 (SmAP из *H. salinarum* длиной 69 а.о. согласно записи UniProt Q9HPS2), HsaSmAP60 (SmAP из *H. salinarum* длиной 60 а.о., UniProt B0R5R2), AF-Sm1 (SmAP1 из *A. fulgidus*, UniProt O29386), AF-Sm2 (SmAP2 из *A. fulgidus*, UniProt O29885), PaeHfq (Hfq из *Pseudomonas aeruginosa*, UniProt Q9NUM0). Черным цветом выделены аминокислотные остатки со степенью гомологии 100%, темно-серым – от 80 до 100%, светло-серым – от 60 до 80%. Внизу приведена соответствующая вторичная структура белков. Выравнивание последовательностей проводили в программе ClustalX с учетом пространственной структуры белков. Анализ гомологии проводили в программе GeneDoc

ловиях («гель-шифт») и методом поверхностного плазмонного резонанса. На рис. 3 представлены результаты «гель-шифта»: показано, что оба варианта белка HsaSmAP не связываются с A18-РНК. Комплекс с U18-РНК формировался только с удлинённым вариантом белка, HsaSmAP69 (рис. 3).

Использование более чувствительного по сравнению с «гель-шифтом» метода ППР показало, что с U18-РНК способны взаимодействовать оба белка (рис. 4, а, б), причём константы диссоциации комплексов отличаются менее, чем на порядок: $K_D(\text{HsaSmAP69} - \text{U18-РНК}) = 0,12 \pm 0,05$ нМ, $K_D(\text{HsaSmAP60} - \text{U18-РНК}) = 0,93 \pm 0,11$ нМ. Сродство белков к U18-РНК оказалось близким к имеющимся в литературе

данным для комплексов гомологичных белков с такой же РНК [15]. Как указано выше, вариант HsaSmAP69, в отличие от HsaSmAP60, имеет дополнительную *N*-концевую последовательность, содержащую преимущественно положительно заряженные а.о. (рис. 1), которые, по всей видимости, способствуют связыванию белка с РНК. Аналогичный эффект *N*-концевой последовательности ранее наблюдался нами для Hfq-подобного SmAP-белка из археи *M. jannaschii*: при удалении этого участка K_D комплекса белка с олиго(U) также увеличивалась примерно на порядок [12].

Данные, полученные методом ППР, подтвердили отсутствие связывания белков с A18-РНК (рис. 4, в, г).

Проведенный ранее структурно-кинетический анализ РНК-белковых взаимодействий SmAP из других архей показал, что их сродство к олиго(A)-РНК является незначительным [15]; согласно нашим данным, для HsaSmAP связывание не детектируется даже при высоких концентрациях белков (рис. 4).

Нуклеотид-связывающие свойства SmAP из *H. salinarum*. Ранее мы успешно использовали анализ нуклеотид-связывающих свойств белков Hfq и SmAP для оценки их сродства к РНК по изменению флуоресценции [15–17]. Аналогичные эксперименты для HsaSmAP69 показали, что этот белок не связывает нуклеотиды AMP и GMP (рис. 5). Эти данные подтверждают, что белок HsaSmAP не имеет участков специфического связывания аденинов и гуанинов.

Пространственная структура белка SmAP из *H. salinarum*. Структура белка HsaSmAP определена методом молекулярного замещения с использованием сервиса MoRDa («BioMEX», Великобритания) [20]. Этот сервис использует набор инструкций для решения задачи молекулярного замещения с автоматическим перебором пространственных структур из базы данных с наивысшей степенью гомологии относительно

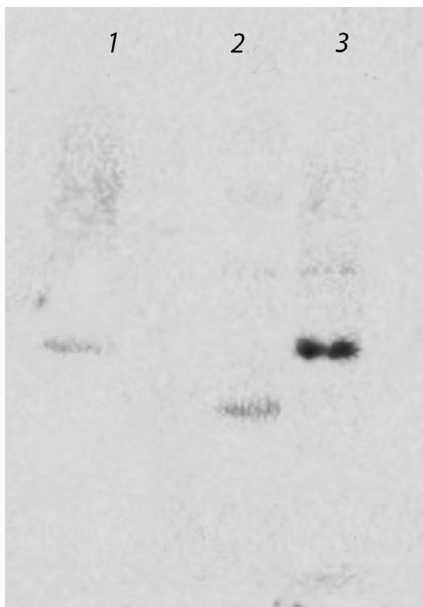


Рис. 2. Вестерн-блот: 1 – лизат клеток *H. salinarum*; 2 – рекомбинантный SmAP из *H. salinarum* длиной 60 а.о. (HsaSmAP60); 3 – рекомбинантный SmAP из *H. salinarum* длиной 69 а.о. (HsaSmAP69)

заданной последовательности белка. Каждая выбранная структура используется для независимого решения задачи молекулярного замещения, что позволяет оптимизировать поиск решения и улучшить его точность. На выходе выдается наилучшее решение из возможных для всего пула первоначальных моделей. Для нашего случая наилучшей моделью оказалась структура мономера SmAP-белка из *A. fulgidus* AF-Sm2 (PDB ID: 1ljo) [13, 24].

В асимметричной части ячейки кристалла найдено 14 мономеров, образующие два гепта-

мера белка HsaSmAP (рис. 6, а и б). Структура мономера полностью соответствует структурам других белков этого класса и содержит *N*-концевую α -спираль и β -бочонок из пяти тяжей, причём тяж $\beta 2$ имеет большую длину и излом посередине. Сравнение со структурами белков Hfq из *E. coli* и AF-Sm2 из *A. fulgidus* показало небольшое отклонение между координатами $S\alpha$ -цепи консервативной части белков – 1,05 Å и 0,84 Å соответственно (рис. 6, в и г). Интересно отметить, что третичная структура мономера HsaSmAP несколько ближе к структуре бактери-

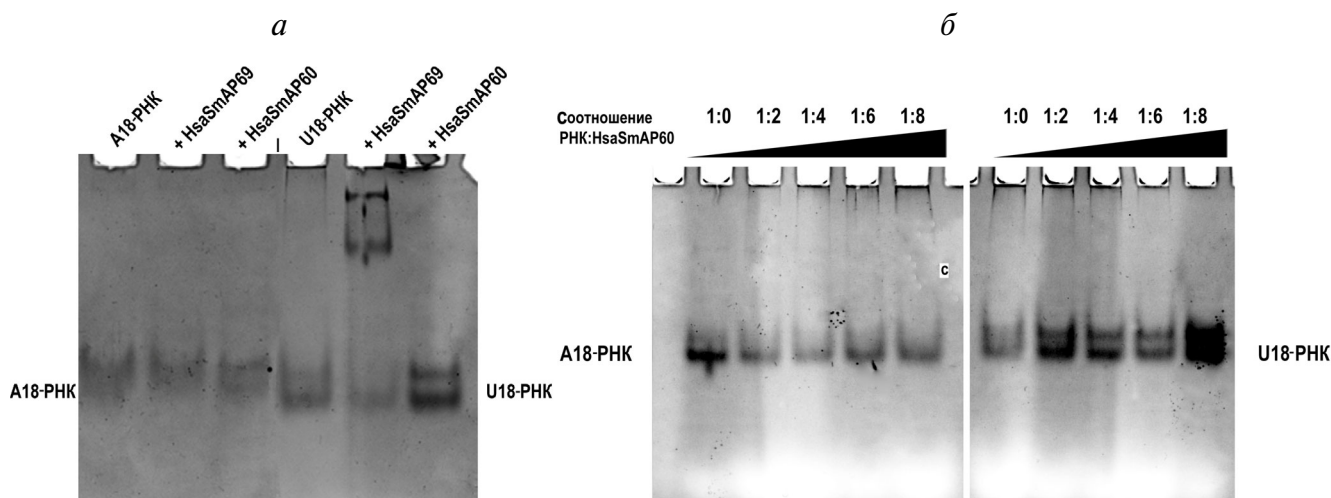


Рис. 3. Проверка способности SmAP из *H. salinarum* образовывать комплексы с A18-РНК и U18-РНК по изменению подвижности в ПААГ при электрофорезе в неденатурирующих условиях: а – при молярном соотношении РНК : белок = 1 : 2; б – для HsaSmAP60 при различных молярных соотношениях РНК : белок

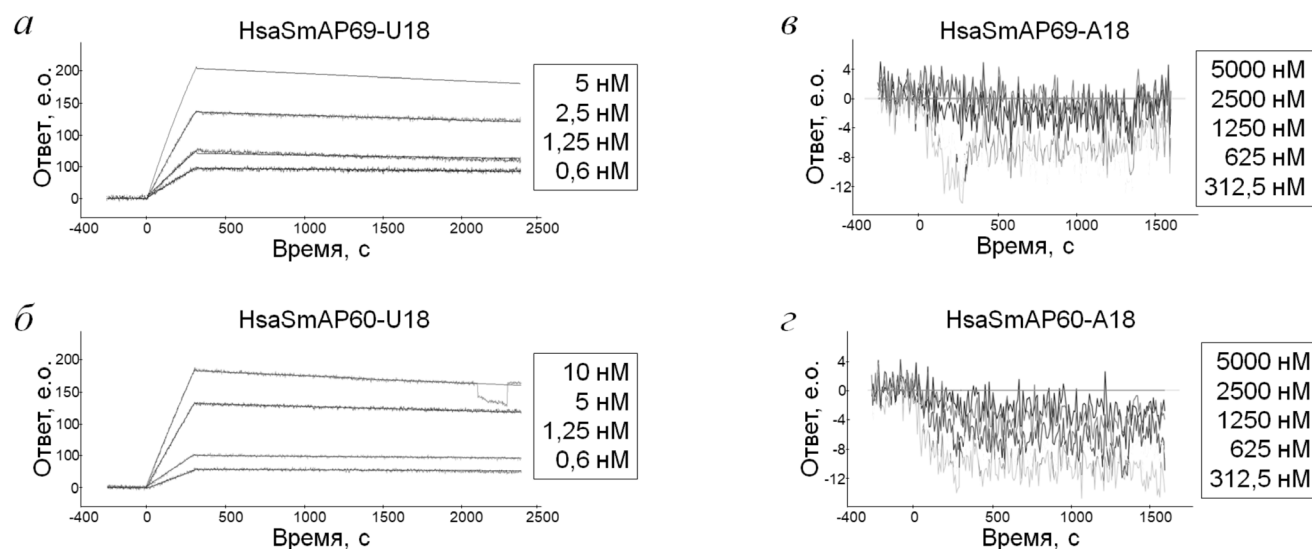


Рис. 4. Сенсограммы, отображающие кинетический анализ взаимодействия HsaSmAP69 (а, в) и HsaSmAP60 (б, г) с U18-РНК (а, б) и A18-РНК (в, г). Для каждого набора сенсограмм указаны используемые концентрации белков. Ровные линии соответствуют теоретической аппроксимации полученных экспериментальных данных

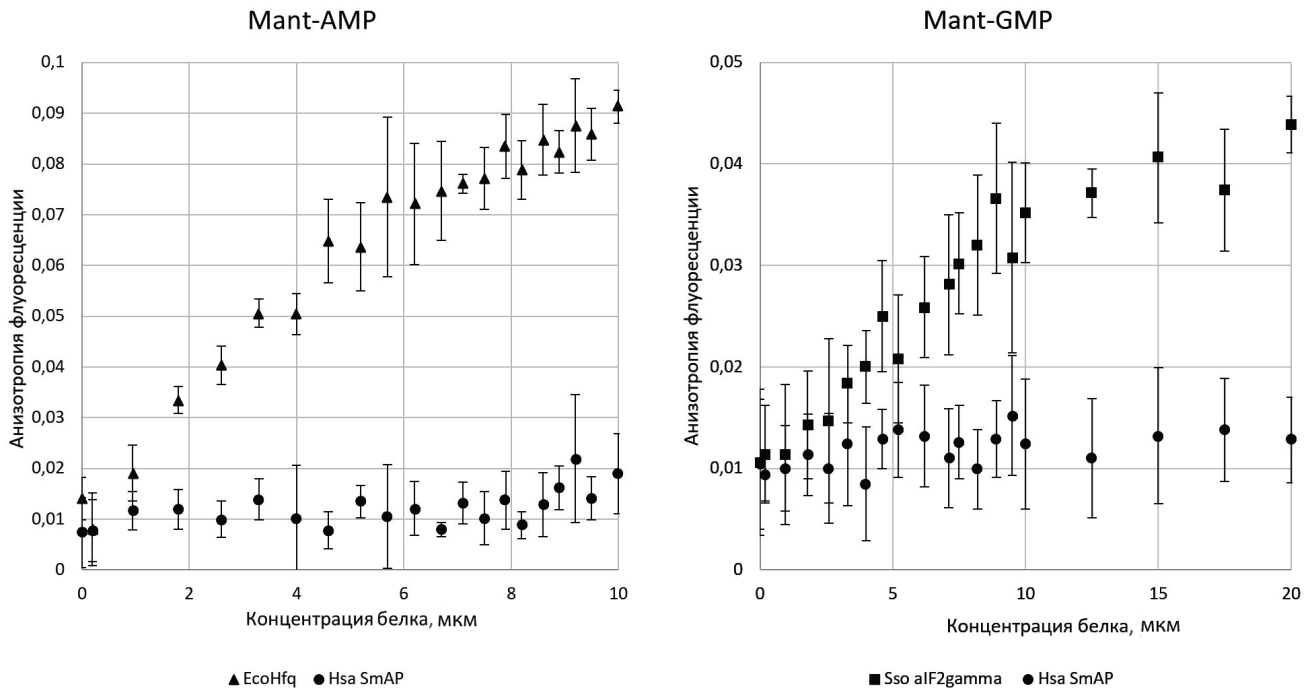


Рис. 5. Кривые титрования белка HsaSmAP69 флуоресцентно-мечеными нуклеотидами Mant-AMP и Mant-GMP. В качестве контроля (специфически связывающихся с нуклеотидами белков) использованы белок Hfq из *E. coli* (EcoHfq) для Mant-AMP и гамма-субъединица aIF2 из *Sulfolobus solfataricus* (Sso aIF2gamma) для Mant-GMP

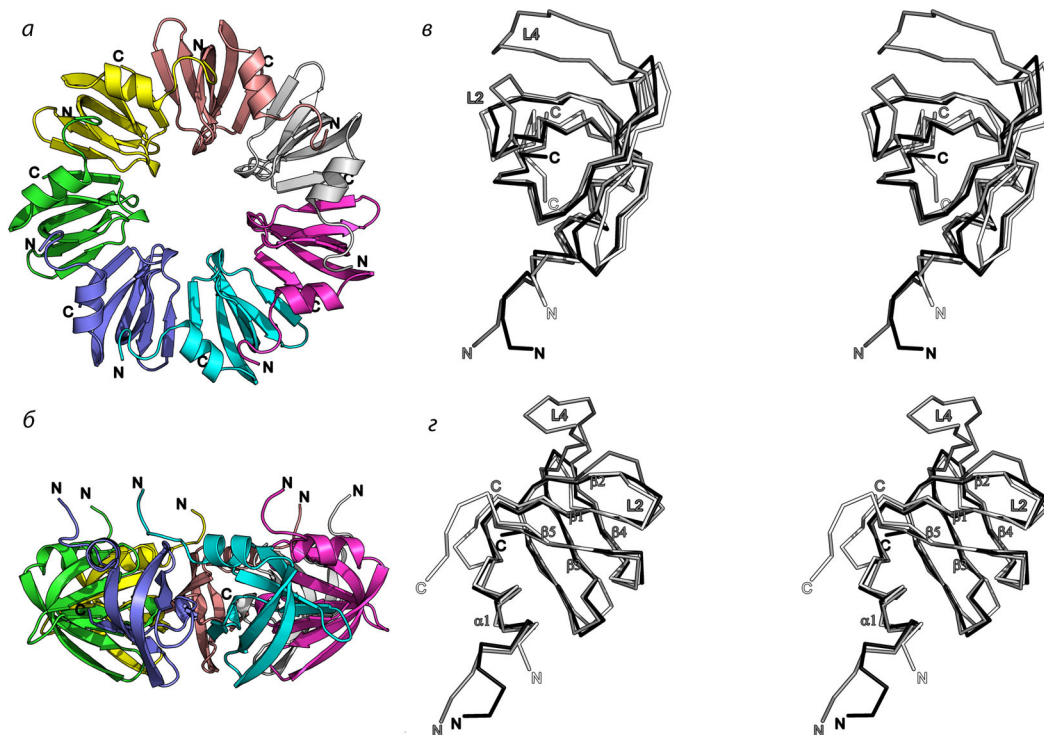


Рис. 6. Структура белка SmAP из *H. salinarum*: *a* – общий вид структуры гептамера со стороны α -спиралей мономеров белка (проксимальная сторона). Обозначены *N*- и *C*-концы мономеров; *b* – общий вид структуры с боковой стороны гептамера; *v* и *z* – наложение α -цепей мономеров SmAP из *H. salinarum* (PDB ID: 6tfl) (черная), AF-Sm2 из *A. fulgidus* (PDB ID: 1ljo) (серая) и Hfq из *E. coli* (PDB ID: 1hk9) (белая). Стереопары: мономер на панели *z* повернут относительно панели *v* на 90° по вертикальной оси. Обозначены *N*- и *C*-концы белков, элементы вторичной структуры, петли L2 и L4. (С цветными вариантами рисунков 6–8 можно ознакомиться в электронной версии статьи на сайте: <http://sciencejournals.ru/journal/biokhsm/>)

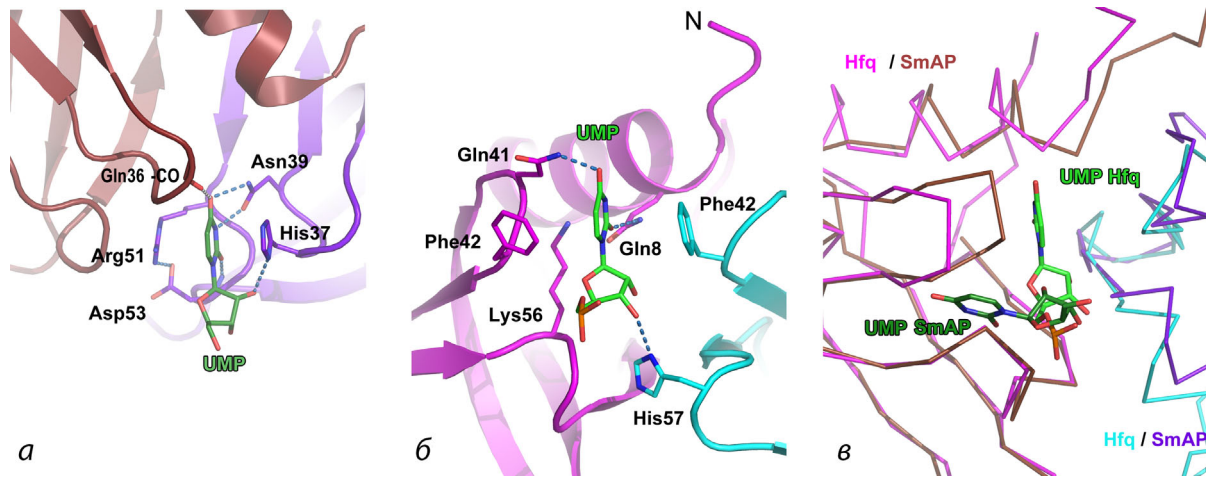


Рис. 7. Сравнение уридин-специфического РНК-связывающего участка белков: *a* – SmAP из *H. salinarum* (PDB ID: 6tfl); *б* – Hfq из *P. aeruginosa* (PDB ID: 4mml). Водородные связи показаны пунктирными линиями. *в* – Изменение ориентации связанных со SmAP и Hfq нуклеотидов относительно друг друга. Соседние мономеры белка Hfq показаны малиновым и голубым, SmAP – коричневым и фиолетовым. Наложение структур произведено в программе Coot

ального белка Hfq, чем к структуре AF-Sm2: оба белка имеют короткую последовательность петли L4, что сказывается на ориентации петли L2, соединяющей тяжи $\beta 1$ и $\beta 2$. Тем не менее белок Hfq имеет больше витков спирали $\alpha 1$ по сравнению с архейными белками, что влияет на конформацию петли L1, соединяющую спираль и тяж $\beta 1$ (рис. 6, *з*). Несмотря на короткую последовательность петли L4, белок HsaSmAP формирует гептамер, а не гексамер, что указывает на независимость четвертичной структуры Lsm-белков от длины петли L4.

В каждом гептамере белка HsaSmAP найдено 7 молекул UMP (рис. 7). Они связаны в стандартном для Lsm-белков уридин-специфическом РНК-связывающем участке белка на проксимальной стороне мультимера. Дополнительных мест связывания UMP, как и в случае белка Hfq из *P. aeruginosa* [25], зафиксировано не было. Основание нуклеотида расположено между боковыми цепями аминокислотных остатков Arg51 и His37 одного мономера белка, причём, как и в аналогичных комплексах белков SmAP с уридинами, боковая цепь аргинина находится в стэкинге с основанием. Атомы Arg51 не формируют водородные связи с атомами основания — их образуют атомы боковой цепи Asn39 с N4 и O4 уридина, а также атомы кислорода главной цепи Asp53 с O2 и Gln36 соседнего мономера с O4 (рис. 7, *а*). Такая система водородных связей характерна для известных структур SmAP-белков в комплексах с олиго(U)-РНК и UMP [15, 24, 26–28]. Важную роль в дискриминации рибонуклеотидов от дезоксирибонуклеотидов играет контакт боковой цепи His37 с O2-атомом UMP, аналогичный контакту O2 с His57 в бакте-

риальных белках Hfq. Сравнение структуры комплексов белков Hfq и SmAP показывает, что модификация области контактов с рибонуклеотидом вызывает изменение положения основания относительно межмономерного интерфейса белков (рис. 7, *в*). UMP в SmAP-белках развернут примерно на 90° по сравнению с его ориентацией в комплексах с белками Hfq при минимальном изменении локальной структуры белков. Возможно, такое изменение положения уридина на поверхности SmAP-белков связано с изменением роли белка в клетках архей по сравнению с бактериальными гомологами.

Сравнение структуры области белка HsaSmAP, соответствующей аденин-специфическому сайту связывания белка Hfq, показало её полную непригодность для взаимодействия с рибонуклеотидами, несмотря на то что эта часть дистальной поверхности гептамера HsaSmAP не прикрыта петлей L4, как в большинстве SmAP-белков (рис. 8). В бактериальных белках Hfq аденин-связывающий сайт сформирован несколькими гидрофобными остатками, что создает благоприятные условия для фиксации основания аденина в правильном положении благодаря стэкингу с консервативным остатком Phe/Тур (в Hfq из *E. coli* – Тур25). При этом атомы боковых цепей консервативных Thr и Gln формируют водородные связи с атомами основания (рис. 8, *б*). В аналогичном участке белка HsaSmAP не наблюдается такого кармана из гидрофобных остатков, а доступ к нему закрыт Arg20, положение боковой цепи которого зафиксировано тремя водородными связями (рис. 8, *а*). Аналогичная картина наблюдается для архейного белка AF-Sm1 (рис. 8, *в*) и для

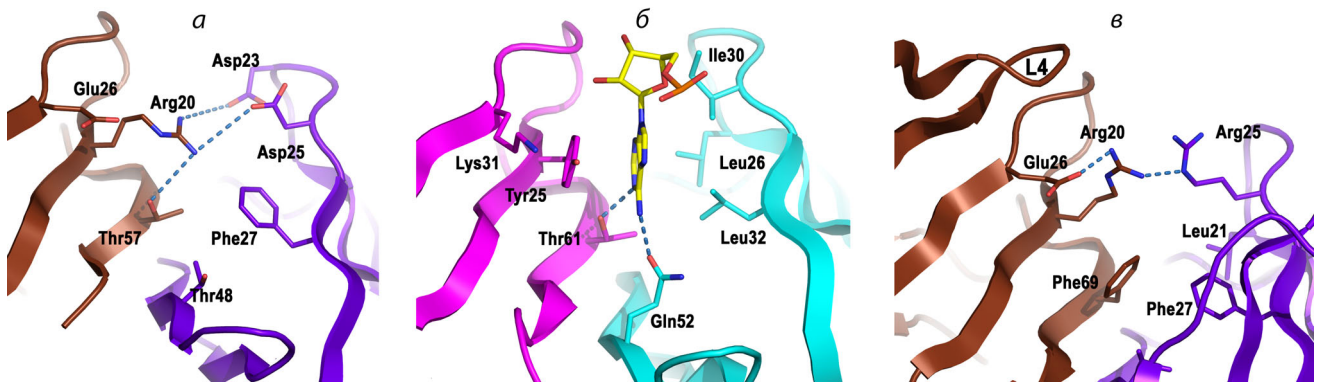


Рис. 8. Сравнение областей белков, соответствующих аденин-специфическому РНК-связывающему участку Hfq: *a* – SmAP из *H. salinarum* (PDB ID: 6tfl); *б* – Hfq из *E. coli* (PDB ID: 3gib); *в* – AF-Sm1 из *A. fulgidus* (PDB ID: 1i5l). Соседние мономеры белка Hfq показаны малиновым и голубым, SmAP-белков – коричневым и фиолетовым

других SmAP-белков [15]. Такое расположение аминокислотных остатков, по-видимому, не позволяет этому участку SmAP-белков выполнять роль РНК-связывающего сайта независимо от длины петли L4, что также подтверждается нашими данными по измерению сродства HsaSmAP к A18-РНК и Mant-AMP.

Суммируя полученные нами данные по структуре белка HsaSmAP и его сродству к олигоРНК и отдельным нуклеотидам, можно утверждать, что этот белок имеет только один, уридин-специфический, РНК-связывающий сайт. Несмотря на то что короткая петля L4 открывает доступ к дистальной поверхности гептамера, белок HsaSmAP не обладает сродством к олиго(A)-РНК, поскольку такой сайт в их структуре не формируется.

Финансирование. Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 18-04-00222).

Благодарности. Авторы выражают благодарность В. Балобанову (ИБ РАН) за помощь в измерениях, В.В. Врублевской (ИБК РАН ФИЦ ПНЦБИ РАН) за предоставление антител и сотрудникам ИБП РАН ФИЦ ПНЦБИ РАН за предоставление возможности проведения измерений методом поверхностного плазмонного резонанса.

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.

СПИСОК ЛИТЕРАТУРЫ

- Murina, V. N., and Nikulin, A. D. (2011) RNA-binding Sm-like proteins of bacteria and archaea. Similarity and difference in structure and function, *Biochemistry (Moscow)*, **76**, 1434-1449, doi: 10.1134/S0006297911130050.
- Mura, C., Randolph, P. S., Patterson, J., and Cozen, A. E. (2013) Archaeal and eukaryotic homologs of Hfq, *RNA Biol.*, **10**, 636-651, doi: 10.4161/rna.24538.
- Lekontseva, N. V., Stolboushkina, E. A., and Nikulin, A. D. (2021) Diversity of LSM Family proteins: similarities and differences, *Biochemistry (Moscow)*, **86**, S38-S49, doi: 10.1134/S0006297921140042.
- Sauter, C., Basquin, J., and Suck, D. (2003) Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*, *Nucleic Acids Res.*, **31**, 4091-4098, doi: 10.1093/nar/gkg480.
- Sauer, E. (2013) Structure and RNA-binding properties of the bacterial LSm protein Hfq, *RNA Biol.*, **10**, 610-618, doi: 10.4161/rna.24201.
- Sharif, H., and Conti, E. (2013) Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover, *Cell Rep.*, **5**, 283-291, doi: 10.1016/j.celrep.2013.10.004.
- Zhou, L., Hang, J., Zhou, Y., Wan, R., Lu, G., et al. (2013) Crystal structures of the Lsm complex bound to the 3'-end sequence of U6 small nuclear RNA, *Nature*, **506**, 116-120, doi: 10.1038/nature12803.
- Roth, A. J., Shuman, S., and Schwer, B. (2018) Defining essential elements and genetic interactions of the yeast Lsm2-8 ring and demonstration that essentiality of Lsm2-8 is bypassed via overexpression of U6 snRNA or the U6 snRNP subunit Prp24, *RNA*, **24**, 853-864, doi: 10.1261/rna.066175.118.
- Pomeranz Krummel, D. A., Oubridge, C., Leung, A. K. W., Li, J., and Nagai, K. (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution, *Nature*, **458**, 475-480, doi: 10.1038/nature07851.
- Wilusz, C. J., and Wilusz, J. (2005) Eukaryotic Lsm proteins: lessons from bacteria, *Nat. Struct. Mol. Biol.*, **12**, 1031-1036, doi: 10.1038/nsmb1037.
- Nielsen, J. S., Bøggild, A., Andersen, C. B. F. F., Nielsen, G., Boysen, A., et al. (2007) An Hfq-like protein in archaea: crystal structure and functional characterization of the Sm protein from *Methanococcus jannaschii*, *RNA*, **13**, 2213-2223, doi: 10.1261/rna.689007.

12. Nikulin, A., Mikhailina, A., Lekontseva, N., Balobanov, V., Nikonova, E., and Tishchenko, S. (2017) Characterization of RNA-binding properties of the archaeal Hfq-like protein from *Methanococcus jannaschii*, *J. Biomol. Struct. Dyn.*, **35**, 1615-1628, doi: 10.1080/07391102.2016.1189849.
13. Törö, I., Basquin, J., Teo-Dreher, H., and Suck, D. (2002) Archaeal Sm proteins form heptameric and hexameric complexes: crystal structures of the sm1 and sm2 proteins from the hyperthermophile *Archaeoglobus fulgidus*, *J. Mol. Biol.*, **320**, 129-142, doi: 10.1016/j.jbiochi.2020.05.001.
14. Kilic, T., Sanglier, S., Van Dorsselaer, A., and Suck, D. (2006) Oligomerization behavior of the archaeal Sm2-type protein from *Archaeoglobus fulgidus*, *Protein Sci.*, **15**, 2310-2317, doi: 10.1016/j.jbiochi.2020.05.001.
15. Lekontseva, N., Mikhailina, A., Fando, M., Kravchenko, O., Balobanov, V., et al. (2020) Crystal structures and RNA-binding properties of Lsm proteins from archaea *Sulfolobus acidocaldarius* and *Methanococcus vannielii*: similarity and difference of the U-binding mode, *Biochimie*, **175**, 1-12, doi: 10.1016/j.jbiochi.2020.05.001.
16. Balobanov, V., Lekontseva, N., Mikhaylina, A., and Nikulin, A. (2020) Use of fluorescent nucleotides to map RNA-binding sites on protein surface, *Methods Mol. Biol.*, **2113**, 251-262, doi: 10.1007/978-1-0716-0278-2_17.
17. Nemchinova, M., Balobanov, V., Nikonova, E., Lekontseva, N., Mikhaylina, A., et al. (2017) An experimental tool to estimate the probability of a nucleotide presence in the crystal structures of the nucleotide-protein complexes, *Protein J.*, **36**, 157-165, doi: 10.1016/j.jbiochi.2020.05.001.
18. Kabsch, W. (2010) XDS, *Acta Crystallogr.*, **D66**, 125-132, doi: 10.1107/S0907444909047337.
19. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., et al. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr.*, **D67**, 235-242, doi: 10.1107/S0907444910045749.
20. Vagin, A., and Lebedev, A. (2015) MoRDa, an automatic molecular replacement pipeline, *Acta Crystallogr.*, **A71**, s19-s19, doi: 10.1107/S2053273315099672.
21. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., et al. (2012) Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr.*, **D68**, 352-367, doi: 10.1107/S0907444912001308.
22. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr.*, **D66**, 213-221, doi: 10.1107/S0907444909052925.
23. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr.*, **D60**, 2126-2132, doi: 10.1107/S0907444904019158.
24. Törö, I., Thore, S., Mayer, C., Basquin, J., Séraphin, B., and Suck, D. (2001) RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex, *EMBO J.*, **20**, 2293-2303, doi: 10.1093/emboj/20.9.2293.
25. Murina, V., Lekontseva, N., and Nikulin, A. (2013) Hfq binds ribonucleotides in three different RNA-binding sites, *Acta Crystallogr.*, **D69**, 1504-1513, doi: 10.1107/S090744491301010X.
26. Thore, S., Mayer, C., Sauter, C., Weeks, S., and Suck, D. (2003) Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya, *J. Biol. Chem.*, **278**, 1239-1247, doi: 10.1074/jbc.M207685200.
27. Mura, C., Phillips, M., Kozhukhovskiy, A., and Eisenberg, D. (2003) Structure and assembly of an augmented Sm-like archaeal protein 14-mer, *Proc. Natl. Acad. Sci. USA*, **100**, 4539-4544, doi: 10.1073/pnas.0538042100.
28. Mura, C., Kozhukhovskiy, A., Gingery, M., and Phillips, M. (2003) The oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs), *Protein Sci.*, **12**, 832-847, doi: 10.1110/ps.0224703.ensure.

STRUCTURE AND RNA-BINDING PROPERTIES OF LSM PROTEIN FROM *Halobacterium salinarum*

M. S. Fando, A. O. Mikhaylina, N. V. Lekontseva, S. V. Tishchenko, and A. D. Nikulin*

*Institute of Protein Research of the Russian Academy of Sciences, 142290 Pushchino,
Moscow Region, Russia; e-mail: nikulin@vega.protres.ru*

The structure and the RNA-binding properties of the Lsm protein from *Halobacterium salinarum* have been determined. A distinctive feature of this protein is the presence of a short L4 loop connecting the $\beta 3$ and $\beta 4$ strands. Since bacterial Lsm proteins (also called Hfq proteins) have a short L4 loop and form hexamers, whereas archaeal Lsm proteins (SmAP) have a long L4 loop and form heptamers, it has been suggested that the length of the L4 loop may affect the quaternary structure of Lsm proteins. Moreover, the L4 loop covers the region of SmAP corresponding to one of the RNA-binding sites in Hfq, and thus can affect the RNA-binding properties of the protein. Our results show that the SmAP from *H. salinarum* forms heptamers and possesses the same RNA-binding properties as homologous proteins with the long L4 loop. Therefore, the length of the L4 does not govern the number of monomers in the protein particles and does not affect the RNA-binding properties of Lsm proteins.

Keywords: Lsm proteins, SmAP, protein quaternary structure, protein RNA-binding properties, small regulatory RNA, sRNA, *Halobacterium salinarum*