

УДК: 577.21

Белок бактериального нуклеоида Dps связывается со структурированными РНК

**Быков А.А.^{1,2}, Шавкунов К.С.^{1,3}, Панюков В.В.^{3,4},
Озолин О.Н.*^{1,3}**

¹*Институт биофизики клетки РАН, Пущино, Московская область, Россия*

²*Пушинский государственный естественно-научный институт, Пущино, Московская область, Россия*

³*Пушинский научный центр РАН, Пущино, Московская область, Россия*

⁴*Институт математических проблем биологии РАН — филиал Института прикладной математики им. М.В. Келдыша РАН, Пущино, Московская область, Россия*

Аннотация. Для взаимодействия с сахаро-фосфатным остовом ДНК архитектурный белок бактериального нуклеоида Dps использует боковые группы лизинов, расположенные на его N-концевых модулях. Электростатический характер взаимодействия предполагает способность Dps связываться с любой нуклеотидной последовательностью, в том числе с РНК. Имеющиеся данные указывают также на то, что Dps имеет повышенное сродство к разветвлённым структурам в ДНК. В РНК такие структуры формируются гораздо чаще, чем в ДНК. Поэтому в работе исследована способность очищенного и иммобилизованного на акрилатных сферах белка Dps связываться с изолированными из бактериальных клеток короткими РНК и установлено, что предпочтительными мишенями для взаимодействия являются транспортные и малые регуляторные РНК, формирующие стабильные вторичные структуры. Среди РНК, обнаруженных в комплексе с Dps, оказалось 8 транскриптов, соответствующих межгенным пространствам, что может свидетельствовать о наличии в них новых генов. Кроме этого, были зарегистрированы продукты длиной 9–13 нуклеотидов, принадлежащие малым нетранслируемым РНК SdsR и RyeA, транскрибируемым по обоим нитям из одного геномного локуса. Так как число более длинных транскриптов из этой области оказалось, по крайней мере, в пять раз меньшим, можно предположить, что два встречных продукта формируют частично комплементарный дуплекс, подвергающийся контролируемому процессингу. Избирательность Dps к этим молекулам, также как к другим структурированным РНК, указывает на возможность его участия не только в конденсации бактериального генома, но и в поддержании функционального состояния клеточного транскриптома.

Ключевые слова: *Dps, комплексы Dps-РНК, pull-down assay, RNA-seq.*

ВВЕДЕНИЕ

Структурное состояние бактериального генома контролируется специальными архитектурными белками, которые поддерживают его в рабочем состоянии во время быстрого роста и защищают от деструктивных повреждений в условиях различных стрессов [1–10]. На стационарной фазе роста основную защитную функцию у кишечной палочки выполняет белок Dps. Для взаимодействия с отрицательно-

* ozoline@rambler.ru

заряженной ДНК он использует положительно-заряженные N-концевые модули своих 12 субъединиц [11–13]. Несмотря на электростатическую природу связывания, эффективность взаимодействия с разными фрагментами ДНК может отличаться [14], что свидетельствует о некоторой селективности белка. Согласно данным, полученным методом атомно-силовой микроскопии, эта избирательность может быть обусловлена повышенным сродством Dps к разветвлённым структурам в ДНК [14]. Такие участки предоставляют возможность для образования множественных контактов с N-концевыми модулями белка, что неизбежно стабилизирует взаимодействие. В таком случае, структурированные молекулы РНК, имеющие разветвлённые вторичные структуры и предоставляющие равную с ДНК возможность для электростатического взаимодействия, должны эффективно связываться с Dps. В литературе имеются данные, свидетельствующих о способности Dps связываться с РНК, например, в работе [15] авторы обнаружили зависимость формы кристаллов, формируемых Dps с ДНК от присутствия РНК. Однако способ взаимодействия Dps и РНК пока совсем не исследован. Поэтому основными целями данной работы стали: оценка способности Dps взаимодействовать с РНК и идентификация молекул РНК, преимущественно с ним взаимодействующих. Для адекватной характеристики структурообразующей способности выявленных мишеней, в качестве потенциальных партнёров для связывания была использована фракция коротких РНК (длиной от 25/9 до 250 нуклеотидов). Этот диапазон включает практически все типы малых структурированных РНК, в том числе тРНК, sРНК, 5S-рРНК, а также небольшую фракцию совсем коротких РНК с ограниченной способностью к формированию вторичных структур.

МАТЕРИАЛЫ И МЕТОДЫ

Бактериальные штаммы, условия культивирования клеток и очистки белка Dps

Для выделения РНК был использован лабораторный штамм *Escherichia coli* K-12 MG1655 (U00096.3, *E. coli*). Культивирование проводили на жидкой среде Luria-Bertani (LB) [16] при 37 °С с перемешиванием. Для получения Dps использовали клетки *E. coli* BL21* (DE3), трансформированные плазмидой pGEM_dps, несущей нативный ген *dps*, и метод очистки, предложенный в работе [17].

Выделение микроРНК

Тотальную фракцию РНК получали из клеток *E. coli* K-12 MG1655, выращенных до $OD_{600} = 0.65$. Клетки собирали центрифугированием (12000 об./мин, 10 мин), после чего добавляли 1 мл реактива Trizol (Thermo Fisher Scientific, США) и помещали в жидкий азот на 15 сек. После полного замораживания, пробу переносили в термостат на 37 °С и инкубировали до полного размораживания биомассы. Процесс замораживания – размораживания повторяли дважды. После повторного размораживания добавляли 200 мкл хлороформа (Реахим, Россия), интенсивно перемешивали 15 сек и инкубировали 5 минут, периодически встряхивая пробирку. Водную фракцию отделяли центрифугированием при 12000 об./мин в течение 10 мин при +4 °С и повторно обрабатывали хлороформом. РНК из водной фазы осаждали 500 мкл изопропилового спирта, и инкубировали 10 мин при комнатной температуре. Осадок собирали центрифугированием и дважды промывали 70 % этиловым спиртом по 400 и 200 мкл, соответственно. После подсушивания при 37 °С (10 мин) осадок растворяли в 42.5 мкл воды, обработанной диэтилпирикарбонатом (PanReac AppliChem, США). Пробу обрабатывали ДНКазой I (New England Biolabs, Великобритания) согласно протоколу производителя. Фракцию коротких РНК выделяли из полученного препарата тотальной РНК с использованием набора mirVana™ miRNA Isolation Kit (Thermo Fisher Scientific, США) в полном соответствии с инструкциями производителя. Конечную

концентрацию РНК в образце определяли на спектрофотометре ND-1000 (NanoDrop Technologies Inc., США).

Иммобилизация Dps на акрилатных сферах и эксперимент «pull-down»

Анти-Dps антитела были получены и очищены, как описано в работе [18]. Их концентрация перед осаждением 60 % сульфатом аммония была 3 мг/мл. Суспензию осаждённых антител хранили при 4 °С. Для иммобилизации Dps из полученного препарата антител отбирали 100 мкл взвеси, и осадок собирали центрифугированием при 12000 об./мин (10 мин). Его растворяли в 50 мкл буфера IP, смешанного в соотношении 3:1 из буфера, содержащего 100 мМ NaCl, 50 мМ Трис-HCl, pH 8, 5мМ ЭДТА, 0.2 % NaN₃ и 0.5 % додецилсульфата натрия, а также буфера, содержащего 100 мМ Трис-HCl, pH 8.6, 100 мМ NaCl, 5 мМ ЭДТА, 0.2 % NaN₃ и 5 % Тритон-X 100. Все операции проводили при +4 °С. К 120 мкл препарата белка Dps с концентрацией 0.6 мг/мл добавляли 680 мкл буфера IP, перемешивали и добавляли 14 мкл антител, растворённых в буфере IP. В качестве контроля использовали 800 мкл буфера IP, к которому добавляли 14 мкл антител, но не добавляли Dps. Оба образца инкубировали 14 часов при постоянном перемешивании с использованием автоматического ротора (Biosan, Латвия). Далее к ним добавляли 60 мкл Protein A/G UltraLink® Resin Beads (Thermo Fisher Scientific, США) и инкубировали 2 часа для иммобилизации белка на акрилатных шариках. Шарик собирали центрифугированием (2500 об./мин, 2.5 мин), дважды промывали 200 мкл буфера ТВ (50 мМ Трис-HCl (pH 8.0), 10 мМ MgCl₂, 0.1 мМ ЭДТА, 50 мМ NaCl) и ресуспендировали в 50 мкл буфера ТВ с добавлением 20 мкл препарата коротких РНК. РНК-белковые комплексы формировали в течение 2-х часов при 37 °С. Затем акрилатные шарики осаждали и дважды промывали 200 мкл 50 мМ NaCl. Для экстракции связавшихся РНК использовали 50 мкл 200 мМ NaCl. Сразу после его добавления и перемешивания пробы наносили на микроколоники из набора для очистки продуктов полимеразной цепной реакции (Евроген, Россия). Препарат очищенных РНК собирали после центрифугирования при 13400 об./мин в течение 1 минуты.

К полученным образцам РНК добавляли 50 мкл изопропилового спирта и инкубировали 14 часов при -20 °С. Осадок собирали центрифугированием и два раза промывали 200 мкл 70 % этанола. После подсушивания его растворяли в 25 мкл безнуклеазной воды (Promega, США). Концентрацию полученных РНК определяли на спектрофотометре ND-1000 (NanoDrop Technologies Inc., США).

Приготовление библиотек и секвенирование

При приготовлении библиотек для секвенирования были использованы короткие РНК, полученные в двух независимых экспериментах «pull-down», как описано выше, но способ приготовления библиотек немного отличался. В первом случае, к полученным РНК с использованием набора реактивов Ion Total RNA-Seq Kit v2 (Thermo Fisher Scientific, США) были лигированы универсальные адаптеры и последующая подготовка образца производилась с использованием Ion OneTouch™ System (Thermo Fisher Scientific, США) и набора реактивов Ion PGM Hi-Q OT2 Kit (Thermo Fisher Scientific, США). Во втором случае, перед присоединением адаптеров Ion Torrent, была использована система лигирования Illumina (набор реактивов NEBNext Multiplex Small RNA Library Prep Set, New England Biolabs, Великобритания). Это позволило увеличить размер последовательностей, прочитываемых секвенатором Ion Torrent, т. е. идентифицировать в комплексах с Dps не только длинные, но и совсем короткие, РНК. Далее секвенирование производилось на платформе Ion PGM™ System (Thermo Fisher Scientific, США) с использованием чипов Ion 314™ Chip v2. Все операции были выполнены в строгом соответствии с протоколами производителя.

Обработка полученных данных

Предварительно, исходный BAM формат данных инвертировали в FASTQ с использованием пакета Samtools v 1.3.1 [19]. Для удаления дополнительных адаптеров во втором эксперименте был использован алгоритм Cutadapt [20]. Чтобы максимально снизить вклад ошибок секвенирования, исходные последовательности были отфильтрованы по качеству прочтения. Для этого был использован инструмент FASTQ Quality filter, входящий в программный пакет FastX-Toolkit v 0.0.14 [21], применяя при этом пороговый уровень качества $Q \geq 13$, обеспечивающий 95% достоверность для 90 % нуклеотидов в каждом прочтении.

Прошедшие фильтрацию прочтения длиной более 25 нуклеотидов (н.) были нарезаны на фрагменты стандартного размера в 25 н. и картированы на геном *E. coli* MG1655 (v.3) с использованием программы Matcher [22], которая находится в свободном доступе по адресу <http://www.mathcell.ru/DnaRnaTools/Matcher.zip>. Картирование на геном последовательностей длиной 9–14 н. осуществлялось независимо для каждой длины, а прочтения, попадающие в одну позицию, суммировались. При выравнивании анализируемых последовательностей на геном программой Matcher учитывалась возможность их соответствия нескольким участкам бактериальной хромосомы, и допускалось присутствие двух ошибок в 25 нуклеотидах последовательности. Профили, полученные для экспериментального и контрольного образцов, были нормированы с помощью метода, предложенного в работе [23].

РЕЗУЛЬТАТЫ

Иммобилизованный на акрилатных сферах Dps способствует сорбции структурированных РНК

Используя прямое лигирование коротких РНК, смытых с иммобилизованного Dps и с сорбента, содержащего только анти-Dps антитело, мы получили приблизительно одинаковое количество прочтений (93658 и 106753 соответственно). После фильтрации по качеству для анализа осталось 57929 последовательностей в экспериментальном и 63902 прочтения в контрольном образцах. Их длина в обоих случаях варьировала от 25 до 250 нуклеотидов, хотя средний размер отличался, и для контрольного образца составлял 60.1 н., а для экспериментального – 86.1 н. После нарезания на фрагменты длиной 25 нуклеотидов экспериментальный набор содержал 182085 фрагментов, а контрольный – 135765. Однако в контрольном образце оказалось много технических прочтений и посторонних последовательностей, не относящихся к бактериальному геному, на который картировалась только 14141 последовательность, а из 182095 образцов экспериментального набора в геноме было обнаружено 92322 последовательности.

В принципе, эта большая количественная разница свидетельствует о том, что иммобилизованный Dps эффективно взаимодействует с РНК, выделенными из бактериальных клеток. Однако распределение контрольных кДНК по геному оказалось не случайным (внутренняя окружность на рис. 1), более того, их профиль был очень похож на профиль распределения по геному для РНК экспериментального набора (вторая окружность от центра на рис. 1). Поэтому качественная нормировка оказалась совершенно необходимой для идентификации транскриптов с повышенным сродством к Dps. Было опробовано несколько методов, начиная с самого простого на число картированных прочтений (рис 2,А, $K = 6.52$), до более сложных способов, учитывающих профиль сигналов контрольного образца. При общей схожести, наиболее убедительные результаты были получены методом, изначально предложенным фирмой Affymetrix для анализа данных, полученных на микроматрицах [23], а затем реализованным в ряде других подходов [25].

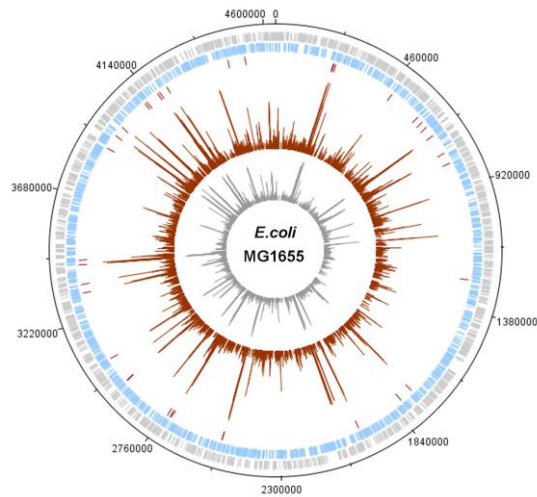


Рис. 1. Распределение в геноме мест синтеза коротких РНК, сорбирующихся на иммобилизованный белок *Dps* (вторая окружность от центра) и на акрилатные сферы, покрытые анти-*Dps* антителом (центральная окружность). Показаны значения $\log_{10}(N + 1)$, где N – сумма прочтений в бине длиной 50 н.п. На внешних двух окружностях схематически показано распределение генов по верхней и нижней нитям генома. Штрихами на третьей окружности показаны пики, амплитуда которых в опытном образце была, по крайней мере, в пять раз больше, чем в контрольном, а также пики, обнаруженные только в опытном образце. График построен с использованием программы DNAPlotter v. 1.3. [24].

Этот метод для расчёта нормировочного коэффициента K использует скорректированные средние значения, удаляя ~2 % сигналов с самой высокой и самой низкой интенсивностью из обоих сравниваемых наборов. На рисунке 2,Б показан результат использованной нормировки, которую получили при удалении 0.5 % экстремальных сигналов. График отражает наличие прочтений с одинаковой представленностью в обоих образцах (лежат на биссектрисе графика) и позволяет идентифицировать РНК, сорбция которых увеличивается в присутствии *Dps*. РНК, обнаруженные только в экспериментальном образце в количестве не меньшем, чем 10 молекул, также как РНК, число которых в этом образце было, по крайней мере, в пять раз больше, чем в контрольном, были отобраны как потенциальные РНК-партнёры *Dps* (таблица 1).

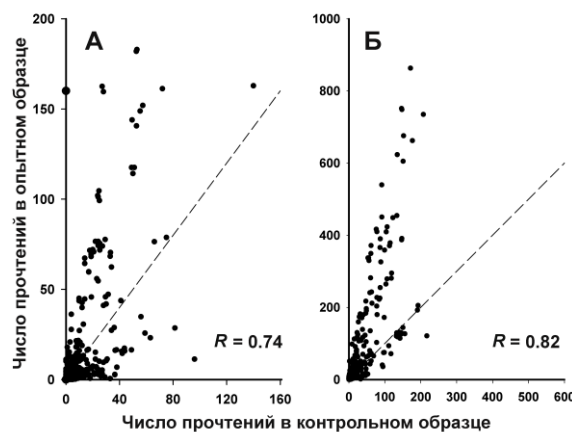


Рис. 2. Корреляция между числом картированных прочтений, зарегистрированных в экспериментальном и контрольном образцах в бинах длиной 50 н.п. **А.** Нормировочный коэффициент рассчитан по общему числу картированных на геном последовательностей ($K = 6.52$). **Б.** Нормировочный коэффициент ($K = 2.94$) рассчитан по скорректированным средним значениям числа картированных прочтений после удаления 0.5 % позиций с самым большим и самым маленьким их количеством и только по позициям, в которых число зарегистрированных последовательностей было больше 1. На панелях приведены значения коэффициента корреляции Пирсона R . Пунктирные линии показывают биссектрису координатного угла.

Таблица 1. РНК-продукты, взаимодействующие с *Dps in vitro*

Область синтеза обнаруженной РНК			Ассоциированные гены			нить	*****Тип РНК-продукта
*Левая граница	*Правая граница	**Длина (н.п.)	Ген	***Левая граница	****Правая граница		
225499	225582	84	<i>alaV</i>	225500	225575	+	Транспортная РНК
228921	229009	89	<i>aspU</i>	228928	229004	+	Транспортная РНК
236930	237014	85	<i>aspV</i>	236931	237007	+	Транспортная РНК
476447	476567	121	<i>ffs</i>	476448	476561	+	4.5S-RNA
607736	607793	58	<i>sokE</i>	607734	607792	-	Малая РНК
631993	632068	76	<i>cstA</i>	629894	631999	+	Потенциальный продукт из межгенного спейсера (+) РНК, имеющая промотор в <i>cspE</i> и терминатор в межгенном спейсере (+)
			<i>ybdD</i>	632182	632379	+	
			<i>cspE</i>	657292	657501	+	
657466	657541	76	<i>crcB</i>	657555	657938	-	
696430	696507	78	<i>glnX</i>	696430	696504	-	Транспортная РНК
774309	774363	55	<i>ybgE</i>	774309	774602	+	мРНК
816914	817023	110	<i>ybhK</i>	815739	816647	-	Потенциальный продукт из межгенного спейсера (+)
			<i>moaA</i>	817044	818033	+	
1270399	1270449	51	<i>rdlC</i>	1270393	1270460	+	Малая РНК
1746435	1746595	161	<i>valV</i>	1746435	1746511	+	Транспортная РНК
			<i>valW</i>	1746516	1746592	+	Транспортная РНК
			<i>rpmI</i>	1799802	1799999	-	РНК, имеющая промотор в <i>infC</i> и терминатор в межгенном спейсере (-)
1800017	1800156	140	<i>infC</i>	1800096	1800638	-	
1991887	1992049	163	<i>leuZ</i>	1991815	1991901	-	Транспортная РНК
			<i>cysT</i>	1991914	1991987	-	Транспортная РНК
			<i>glyW</i>	1992042	1992117	-	Транспортная РНК
2518015	2518117	103	<i>alaX</i>	2518041	2518116	-	Транспортная РНК
2518130	2518232	103	<i>alaW</i>	2518156	2518231	-	Транспортная РНК
2729340	2729445	106	<i>gltW</i>	2729369	2729444	-	Транспортная РНК
2734153	2734270	118	<i>ryfD</i>	2734153	2734295	-	Малая РНК
2746150	2746255	106	<i>ZipA5'</i>	2746183	2746432	-	Малая РНК
2947385	2947466	82	<i>metZ</i>	2947387	2947463	+	Транспортная РНК
2947494	2947580	87	<i>metW</i>	2947497	2947573	+	Транспортная РНК
2947604	2947686	83	<i>metV</i>	2947607	2947683	+	Транспортная РНК
3049082	3049157	76	<i>invR</i>	3046901	3046993	+	Потенциальный продукт из межгенного спейсера (-)
			<i>gcvH</i>	3049160	3049549	-	
3318190	3318290	101	<i>metY</i>	3318213	3318289	-	Транспортная РНК
3350588	3350699	112	<i>arcZ</i>	3350577	3350697	+	Малая РНК
3426928	3427164	237	<i>alaU</i>	3426958	3427033	-	Транспортная РНК
			<i>ileU</i>	3427076	3427152	-	Транспортная РНК
3443650	3443712	63	<i>secY</i>	3442766	3444097	-	мРНК
3888411	3888534	124	<i>tnaC</i>	3888435	3888509	+	Аттенуатор <i>tnaCAB</i>
3943383	3943516	134	<i>gltU</i>	3943435	3943510	+	Транспортная РНК
3982508	3982592	85	<i>hisR</i>	3982509	3982585	+	Транспортная РНК
4106329	4106390	62	<i>cpxP</i>	4105820	4106320	+	Потенциальный продукт из межгенного спейсера (+)
			<i>fieF</i>	4106469	4107371	+	
4115666	4115718	53	<i>glpX</i>	4114569	4115579	-	Потенциальный продукт из межгенного спейсера (-)
			<i>glpK</i>	4115714	4117222	-	
4168371	4168453	83	<i>gltT</i>	4168372	4168447	+	Транспортная РНК
4175663	4175836	174	<i>glyT</i>	4175673	4175747	+	Транспортная РНК
			<i>thrT</i>	4175754	4175829	+	Транспортная РНК
4209772	4209855	84	<i>gltV</i>	4209774	4209849	+	Транспортная РНК
4463026	4463054	29	<i>nrdD</i>	4460522	4462660	-	Потенциальный продукт из межгенного спейсера (-)
			<i>treC</i>	4463054	4464709	-	
4528057	4528119	63	<i>ryjB</i>	4527977	4528066	+	Потенциальный продукт из межгенного спейсера (+)
			<i>sgcQ</i>	4528111	4528917	-	

Красным цветом отмечены транскрипты, соответствующие межгенным областям; зелёным цветом отмечены области, не имеющие прочтений в контрольном образце.

Выделение жирным шрифтом:

* – есть промотор рядом с РНК;

** – потенциальная РНК способна формировать вторичные структуры;

*** – продукты длиной 25 н. зарегистрированы в двух экспериментах;

**** – во втором эксперименте зарегистрированы продукты длиной 9–14 н.;

***** – в скобках показано направление транскрипции для межгенных продуктов.

Стало очевидно, что эффективнее всего с Dps связываются транспортные и малые регуляторные РНК. Все такие молекулы формируют стабильные шпилечные структуры, что подтверждает предположение о повышенном сродстве Dps к разветвлённым структурам в нуклеиновых кислотах [14]. Кроме этого, в комплексе с Dps было обнаружено семь РНК-продуктов, синтезируемых из межгенных участков (отмечены красным в таблице 1). Это может свидетельствовать о наличии в этих участках новых генов для нетранслируемых РНК. В двух других случаях картированные олигонуклеотиды, кроме межгенного участка, захватывали конец соседнего гена. Такими продуктами, в частности являются РНК, синтезируемые между конвергентными генами *cspE* и *crcB* (рис. 3). Соответствующая РНК заканчивается, по-видимому, на терминаторе гена *cspE* (кодирует ДНК- и РНК-связывающий белок, функционирующий как антитерминатор на ρ -независимых терминаторах транскрипции [26]), но внутри этого гена есть два промотора для независимой инициации транскрипции. Не исключено, поэтому, что зарегистрированные РНК синтезируются независимо от основной мРНК. Подвергаясь специфическому процессингу, они могут функционировать в клетках как нетранслируемые РНК. Промоторы для независимой от основной мРНК транскрипции найдены ещё для трёх продуктов, простирающихся в межгенные области (отмечены жирным шрифтом в первом и втором столбцах таблицы 1), а шесть потенциально новых РНК-продуктов способны образовать вторичные структуры со свободной энергией фолдинга меньше -21.8 кКал/моль (отмечены жирным шрифтом в третьем столбце таблицы 1).

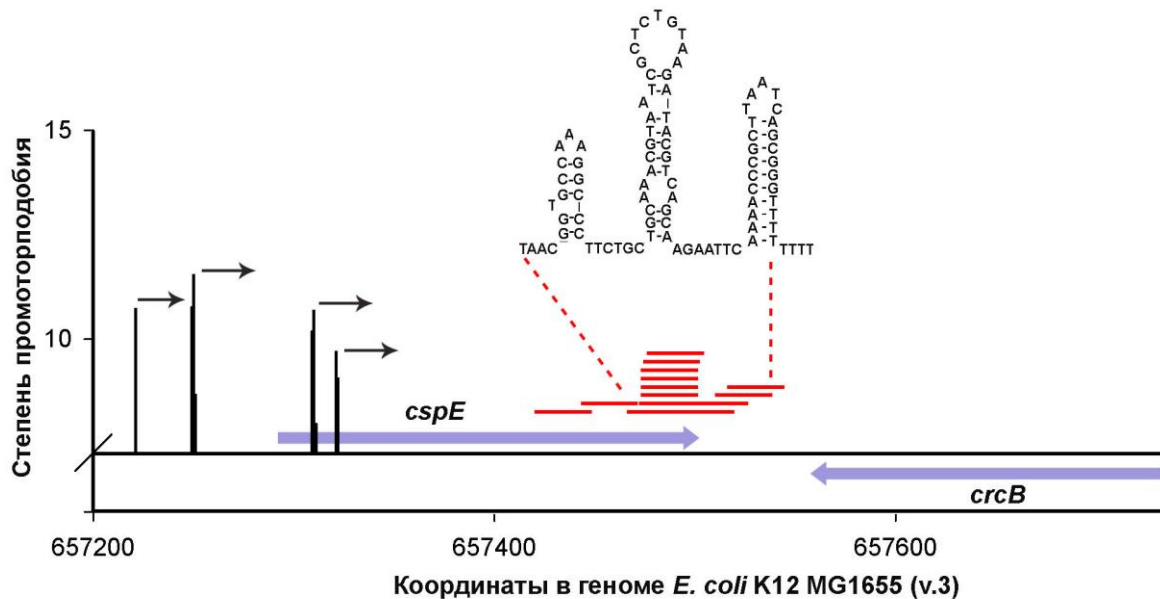


Рис. 3. Расположение зарегистрированных прочтений (красные линии) в геномной области *cspE/crcB*. Гены и направление их транскрипции показаны синими стрелками. Предсказанные PlatProm промоторы – чёрными столбиками со стрелками [27]. Сверху показана вторичная структура РНК, предсказанная программой RNA structure [28].

Таким образом, лучше всего с иммобилизованным Dps взаимодействуют тРНК и малые регуляторные РНК, но большинство связавшихся с Dps транскриптов, не принадлежащих к этим категориям, тоже способны формировать шпилечные структуры с высокой энергией фолдинга.

Иммобилизованный на акрилатных сферах Dps может сорбировать короткие РНК

Почти все РНК, отобранные по высокой степени представленности в экспериментальном наборе (таблица 1), были длинее 50 нуклеотидов (таблица 1). Прочтения короче 25 нуклеотидов вообще не были зарегистрированы в этом эксперименте. Это могло быть следствием того, что короткие РНК просто не

связываются с Dps и не сорбируются на акрилатные сферы, использованные для иммобилизации. Возможно такие олигонуклеотиды являются плохими субстратами для присоединения адаптеров ферментной системой Ion Torrent. Поэтому во втором эксперименте были использованы двойные адаптеры. Вначале к РНК, смытым с акрилатных сфер и с иммобилизованного Dps, были присоединены адаптеры, используемые для секвенирования на платформе Illumina (набор реактивов NEBNext Multiplex Small RNA Library Prep Set, New England Biolabs). После получения первой и второй копий ДНК к ним были присоединены адаптеры Ion Torrent и полученные фрагменты отсеквенированы. Такая стратегия позволяла обнаружить совсем короткие транскрипты, расположенные между внутренними адаптерами.

В этом эксперименте было получено всего 3694 прочтений с опытного образца, а в контрольном препарате количество РНК было недостаточным для последующего секвенирования. Длина последовательностей, снятых с иммобилизованного Dps варьировала от 2 до 231 нуклеотидов (средний размер 33.3 н.). После фильтрации по качеству прочтения и сортировки по размерам мы получили 2055 фрагментов длиной 25 нуклеотидов и 363 фрагмента длиной от 9 до 14 нуклеотидов. На второй от центра окружности рисунка 4 показано распределение в геноме длинных фрагментов. Видно, что большая часть пиков соответствует РНК, обнаруженным в комплексе с Dps в первом эксперименте (совпадающие области отмечены жирным шрифтом в пятой и шестой колонках таблицы 1). К ним, в частности, относится и РНК из геномной области *cspE/crcB*.

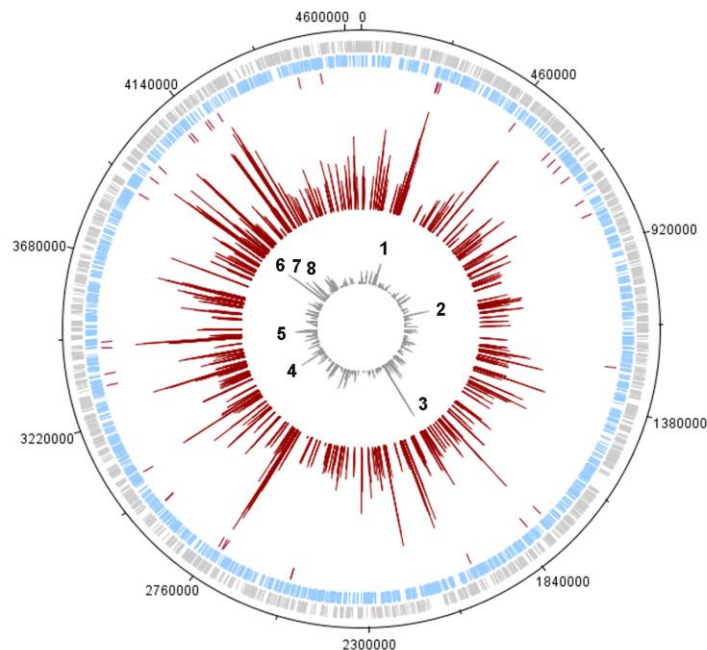


Рис. 4. Распределение в геноме мест синтеза РНК длиной 25 нуклеотидов и 9–14 нуклеотидов, сорбирующихся на иммобилизованный белок Dps (центральная и вторая окружность от центра, соответственно). Показаны значения $\log_{10}(N + 1)$, где N – сумма прочтений в бине длиной 50 н.п. На внешних двух окружностях схематически показано распределение генов по верхней и нижней нитям генома. Штрихами на третьей окружности показаны участки генома, отобранные по первому эксперименту (рис. 1 и таблица 1). График построен с использованием программы DNAPlotter v. 1.3. [24].

Распределение по геному прочтений длиной 9–14 нуклеотидов отличалось от более длинных РНК (центральная окружность на рис. 4), хотя некоторое соответствие и прослеживается. Доминирующими являются восемь пиков. Семь из них тоже соответствуют транспортным и регуляторным РНК. Исключением является пик 2, который расположен в середине гена *asnS*, кодирующего аспарагинил-тРНК синтетазу – фермент биосинтеза тРНК, но фрагмент с зарегистрированными прочтениями длиной

63 н. тоже формирует шпильчатые структуры со свободной энергией фолдинга -17.5 кКал/моль. Особое значение имеют прочтения, формирующие самый большой пик 3. Он соответствует двум малым нетранслируемым РНК RyeA и SdsR с плохо изученными свойствами, транскрибируемым по обоим нитям из одного геномного локуса (рис. 5).

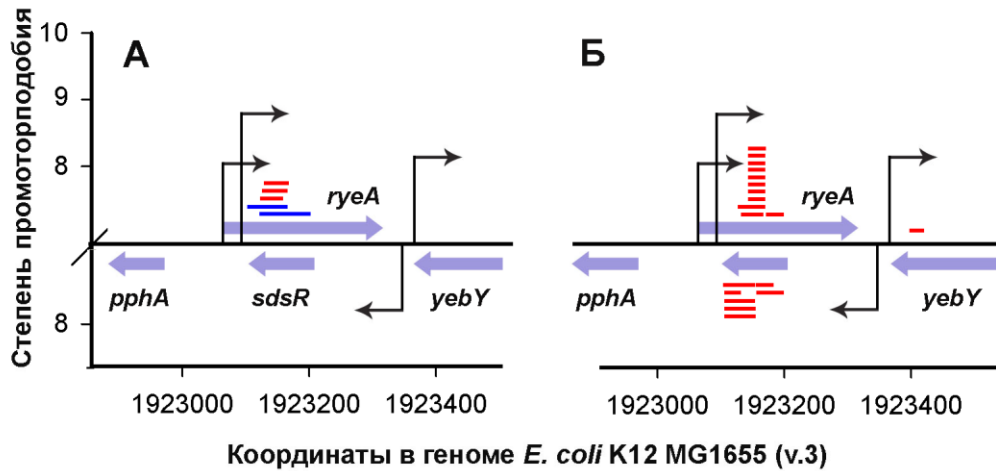


Рис. 5. Расположение прочтений длиной 25 н. (А) и 9–14 н. (Б) в геномной области *ryeA/sdsR*, зарегистрированных в первом и втором экспериментах (синие и красные линии, соответственно). Гены и направление их транскрипции показаны голубыми стрелками. Предсказанные PlatProm промоторы – чёрными столбиками со стрелками [27].

Взаимодействия с 25 н. РНК из этого локуса были зарегистрированы в обоих экспериментах (рис. 5,А), но только короткие прочтения второго эксперимента сформировали значимый пик в этой области (рис. 5,Б). Короткие прочтения были зарегистрированы по обоим нитям (рис. 5,Б), что предполагает их принадлежность разным транскриптам, хотя использование двойных адаптеров не позволяет это утверждать с полной уверенностью. Однако явное преобладание коротких продуктов свидетельствует о посттранскрипционном процессинге исходных РНК. Известно, что обе РНК имеют нестабильный размер: от 249 до 150 н. для RyeA и от 104 до 74 н. для SdsR [29, 30], что предполагает их посттранскрипционный процессинг, который может быть опосредован образованием частично комплементарного дуплекса не внутри одной РНК, а с участием двух независимо транскрибируемых продуктов [30, 31]. Не исключено, что конечной целью этого процессинга является образование нетипичных для бактерий совсем коротких РНК-дуплексов, напоминающих микроРНК высших организмов.

ОБСУЖДЕНИЕ

Основными целями данной работы были проверка способности Dps взаимодействовать с РНК и структурная характеристика транскриптов, связывающихся с иммобилизованным белком. Dps является крупным белком, состоящим из 12 одинаковых субъединиц с молекулярной массой 18.695 кДа. Для взаимодействия с нуклеиновыми кислотами он использует положительно заряженные боковые группы лизинов, находящиеся на экспонированных на поверхность шарообразной глобулы гибких N-концевых модулях. Вся остальная поверхность белковой глобулы заряжена отрицательно. Поэтому опасность неспецифической сорбции нуклеиновых кислот на поверхность белка казалась минимальной. Тем не менее, мы столкнулись с проблемой неспецифической сорбции РНК на акрилатную матрицу с присоединённым к ней анти-Dps антителом. Скорее всего, это обусловлено положительным зарядом на поверхности антитела, который вполне мог отобраться иммунной системой к отрицательно заряженному эпитопу.

Менее понятна схожесть профилей распределения по геному прочтений контрольного и опытного образцов (рис. 1), которая в значительной степени определяется наличием в обоих наборах транспортных РНК. Так как среди пиков, отобранных по отсутствию прочтений в контрольном наборе, нет ни одного, соответствующего тРНК (отмечены зелёным цветом в таблице 1), проще всего их доминирование объяснить высокой концентрацией этих молекул в клетках. Рибосомные белки и РНК неизбежно «засоряют» любые биологические препараты и для удаления рРНК и тРНК имеются специальные методы, которые нельзя было использовать в поисковой работе. Однако не очень понятно, почему среди отобранных партнёров совсем не оказалось продуктов деградации больших рибосомных РНК.

В первом эксперименте количество сорбированных молекул на контрольные сферы было намного меньше, чем на сферы, покрытые Dps, а во втором эксперименте фракцию неспецифически связавшихся РНК вообще не удалось получить. Поэтому мы считаем, что жёсткие критерии, использованные для идентификации РНК, способных связываться с Dps, позволяют сделать вывод о способности Dps связываться с РНК, по крайней мере, *in vitro*, причём основными мишенями для этого белка являются структурированные молекулы, формирующие стабильные шпилечные структуры. Если это так, то возникает необходимость в переосмыслении биологической роли Dps. Являясь одним из основных структурообразующих белков бактериального нуклеоида, Dps конденсирует геном при недостатке источников питания в кристалло-подобную структуру [32] и использует ферроксидазную активность для защиты ДНК от деструктивного воздействия активных форм кислорода [33]. Особое сродство Dps к структурированным РНК увеличивает функциональный потенциал этого белка, но биологическую целесообразность сочетания этих функций ещё только предстоит понять.

Работа выполнялась при финансовой поддержке Российского научного фонда (грант № 14-14-00985, ОНО, АСБ, КСШ). Авторы выражают благодарность В.В. Врублевской и О.С. Моренкову за получение анти-Dps антител.

СПИСОК ЛИТЕРАТУРЫ

1. Dorman C.J. Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. *J. Mol. Microbiol. Biotechnol.* 2014 V. 24. P. 316–331. doi: [10.1159/000368850](https://doi.org/10.1159/000368850).
2. Azam T.A., Ishihama A. Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* 1999. V. 274. P. 33105–33113. doi: [10.1074/jbc.274.46.33105](https://doi.org/10.1074/jbc.274.46.33105).
3. Azam T.A., Iwata A., Nishimura A., Ueda S., Ishihama A. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* 1999. V. 181. P. 6361–6370.
4. Azam T.A., Hiraga S., Ishihama A. Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes to Cells.* 2000. V. 5. P. 613–626. doi: [10.1046/j.1365-2443.2000.00350.x](https://doi.org/10.1046/j.1365-2443.2000.00350.x).
5. Grainger D.C., Hurd D., Goldberg M.D., Busby S.J.W. Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Research.* 2006. V. 34. P. 4642–4652. doi: [10.1093/nar/gkl542](https://doi.org/10.1093/nar/gkl542).
6. Kahramanoglou C., Seshasayee A.S.N., Prieto A.I., Ibberson D., Schmidt S., Zimmermann J., Benes V., Fraser G.M., Luscombe N.M. Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Research.* 2011. V. 39. P. 2073–2091. doi: [10.1093/nar/gkq934](https://doi.org/10.1093/nar/gkq934).
7. Vora T., Hottes A.K., Tavazoie S. Protein occupancy landscape of a bacterial genome. *Molecular Cell.* 2009. V. 35. P. 247–253. doi: [10.1016/j.molcel.2009.06.035](https://doi.org/10.1016/j.molcel.2009.06.035).

8. Prieto A.I., Kahramanoglou C., Ali R.M., Fraser G.M., Seshasayee A.S.N., Luscombe N.M. Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Research*. 2012. V. 40. P. 3524–3537. doi: [10.1093/nar/gkr1236](https://doi.org/10.1093/nar/gkr1236).
9. Dorman C.J. H-NS, the genome sentinel. *Nat. Rev. Microbiol.* 2007. V. 5. P. 157–161. doi: [10.1038/nrmicro1598](https://doi.org/10.1038/nrmicro1598).
10. Wang W., Li G-W., Chen C., Xie X.S., Zhuang X. Chromosome organization by a nucleoid-associated protein in live bacteria. *Science*. 2011. V. 333. P. 1445–1449. doi: [10.1126/science.1204697](https://doi.org/10.1126/science.1204697).
11. Almirón M., Link A.J., Furlong D., Kolter R. A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev.* 1992. V. 6. P. 2646–2654. doi: [10.1101/gad.6.12b.2646](https://doi.org/10.1101/gad.6.12b.2646).
12. Grant R.A., Filman D.J., Finkel S.E., Kolter R., Hogle J.M. The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat. Struct. Biol.* 1998. V. 5. P. 294–303. doi: [10.1038/nsb0498-294](https://doi.org/10.1038/nsb0498-294).
13. Ceci P., Cellai S., Falvo E., Rivetti C., Rossi G.L., Chiancone E. DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Research*. 2004. V. 32. P. 5935–5944. doi: [10.1093/nar/gkh915](https://doi.org/10.1093/nar/gkh915).
14. Melekhov V.V., Shvyreva U.S., Timchenko A.A., Tutukina M.N., Preobrazhenskaya E.V., Burkova D.V., Artiukhov V.G., Ozoline O.N., Antipov S.S. Modes of *Escherichia coli* Dps interaction with DNA as revealed by atomic force microscopy. *PLoS ONE*. 2015 V. 10. doi: [10.1371/journal.pone.0126504](https://doi.org/10.1371/journal.pone.0126504).
15. Ghatak P., K. Karmakar S., Kasetty D.C. Unveiling the role of Dps in the organization of mycobacterial nucleoid. *PLoS ONE*. 2011. V. 6. № 1. doi: [10.1371/journal.pone.0016019](https://doi.org/10.1371/journal.pone.0016019).
16. Sambrook J., Fritsch E.F., Maniatis. T. *Molecular cloning: a laboratory manual, 2nd ed.* New York: Cold Spring Harbor Laboratory Press, 1989.
17. Покусаева В.О., Антипов С.С., Швырева У.С., Тутукина М.Н., Озолин О.Н. Суперпродукция, выделение и очистка функционально активного бактериоферритина Dps *E. coli*. *Сорбционные и хроматографические процессы*. 2012. Т. 12. С. 1011–1017.
18. Oppermann M. Anion exchange chromatography for purification of monoclonal IgG antibodies. In: *Monoclonal antibodies*. Ed. Peters J.P., Baumgarten H. Heidelberg: Springer, 1992. P. 271–275.
19. Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. The sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009. V. 25. № 16. P. 2078–2079. doi: [10.1093/bioinformatics/btp352](https://doi.org/10.1093/bioinformatics/btp352).
20. Marcel M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*. 2011. V. 17. № 1. P. 10–12. doi: [10.14806/ej.17.1.200](https://doi.org/10.14806/ej.17.1.200).
21. FastX-Toolkit URL: http://www.hannolab.cshl.edu/fastx_toolkit (accessed 23.11.16).
22. Panyukov V.V., Kiselev S.S., Shavkunov K.S., Masulis I.S., Ozoline O.N. Mixed promoter islands as genomic regions with specific structural and functional properties. *Mathem. Biol. Bioinf.* 2013. V. 8. P. 432–448. doi: [10.17537/2013.8.432](https://doi.org/10.17537/2013.8.432).
23. *Statistical Algorithms Description Document*. URL: http://media.affymetrix.com/support/technical/whitepapers/sadd_whitepaper.pdf (accessed 23.11.16).
24. Carver T., Thomson N., Bleasby A., Berriman M., Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*. 2009. V. 25. № 1. P. 119–120. doi: [10.1093/bioinformatics/btn578](https://doi.org/10.1093/bioinformatics/btn578).
25. Aleksic J., Carl S., Fryel M. Beyond library size: a field guide to NGS normalization. *bioRxiv*. 2014. doi: <https://doi.org/10.1101/006403>

26. Bae W., Xia B., Inouye M., Severinov K. *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc. Natl. Acad. Sci.* 2000. V. 9714. P. 7784–7789. doi: [10.1073/pnas.97.14.7784](https://doi.org/10.1073/pnas.97.14.7784).
27. Shavkunov K.S., Masulis I.S., Tutukina M.N., Deev A.A., Ozoline O.N. Gains and unexpected lessons from genome-scale promoter mapping. *Nucleic Acids Res.* 2009. V. 37. № 15. P. 4919–4931. doi: [10.1093/nar/gkp490](https://doi.org/10.1093/nar/gkp490).
28. Mathews D.H. RNA secondary structure analysis using RNAstructure. *Current Protocols in Bioinformatics.* 2014. V. 46. doi: [10.1002/0471250953.bi1206s46](https://doi.org/10.1002/0471250953.bi1206s46).
29. Wassarman K.M., Repoila F., Rosenow C., Storz G., Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 2001. V. 15. № 13. P. 1637–1651. doi: [10.1101/gad.901001](https://doi.org/10.1101/gad.901001).
30. Vogel J., Bartels V., Tang T.H., Churakov G., Slagter-Jäger J.G., Hüttenhofer A., Wagner E.G. RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* 2003. V. 31. № 22. P. 6435–6443. doi: [10.1093/nar/gkg867](https://doi.org/10.1093/nar/gkg867).
31. Argaman L., Hershberg R., Vogel J., Bejerano G., Wagner E.G., Margalit H., Altuvia S. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* V. 11. № 12. P. 941–950. doi: [10.1016/S0960-9822\(01\)00270-6](https://doi.org/10.1016/S0960-9822(01)00270-6).
32. Frenkiel-Krispin D., Levin-Zaidman S., Shimoni E., Wolf S.G., Wachtel E.J., Arad T. Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection. *EMBO.* 2001. V. 20. P. 1184–1191. doi: [10.1093/emboj/20.5.1184](https://doi.org/10.1093/emboj/20.5.1184).
33. Zhao G., Ceci P., Ilari A., Giangiacomo L., Laue T., Chiancone E., Emilia C., Chasteen D.N. Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells. A ferritin-like DNA-binding protein of *Escherichia coli*. *J. Biol. Chem.* 2002. V. 277. P. 27689–27696. doi: [10.1074/jbc.M202094200](https://doi.org/10.1074/jbc.M202094200).

Рукопись поступила в редакцию 24.11.2016.

Дата опубликования 30.11.2016.