================= **TRANSLATIONS OF PUBLISHED ARTICLES** =================
UDC: 577.21

Bacterial nucleoid protein Dps binds structured RNA molecules

Bykov A.A.[1,2], Shavkunov K.S.[1,3], Panyukov V.V.[3,4], Ozoline O.N.*[1,3]

[1]*142290, Institute of Cell Biophysics of Russian Academy of Sciences, Pushchino, Moscow region, Russia*
[2]*142290, Pushchino State Institute of Natural Sciences, Pushchino, Moscow region, Russia*
[3]*142290, Pushchino Scientific Center of Russian Academy of Sciences, Pushchino, Moscow region, Russia*
[4]*142290, Institute of Mathematical Problems of Biology - the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino, Moscow region, Russia*

Abstract. Architectural protein Dps of bacterial nucleoid employs side lysine groups at its N-terminal modules for interacting with the sugar-phosphate backbone of the DNA. Electrostatic nature of interaction assumes the potential ability of Dps to bind with any nucleotide sequence, including RNA. The available data also indicate that Dps exhibits enhanced affinity to branched DNA structures. In RNA molecules such structures are formed more frequently than in DNA. Hence, the aim of this investigation was to study the ability of purified Dps immobilized on acrylate spheres to bind short RNAs isolated from bacterial cells. It appeared that transport and small regulatory RNAs forming stable secondary structures are the preferred targets for such interaction. Among RNAs identified in complexes with Dps 8 transcripts corresponded to intergenic spaces, which might indicate the presence of novel genes. Moreover, 9–13 nucleotide long products, belonging to small untranslated RNAs SdsR and RyeA and transcribed from both strands of the same locus, were registered. Since the amount of longer transcripts from this region was at least five-fold lower, it can be presumed that pairs of counter-synthesized products form partly complementary duplexes subjected to the controlled processing. The selectivity of Dps to these molecules, as well as to other structured RNAs, indicates a possibility of its involvement not only in bacterial genome condensation, but also in maintaining the functional state of the transcriptome.

Key words: Dps, Dps-RNA complexes, pull-down assay, RNA-seq.

INTRODUCTION

The structural state of the bacterial genome is controlled by special architectural proteins, maintaining its functional state during quick growth and protecting the genome from the destructive damage under various stress conditions [1–10]. The major protection of the genome in *E. coli* at stationary growth phase is mainly realized by Dps protein. Its molecules utilize positively charged N-terminal modules of the 12 subunits for interaction with negatively charged DNA [11–13]. Despite the electrostatic nature of binding, the efficiency of interaction with different DNA fragments may vary [14], indicating certain sequence or

---

*ozoline@rambler.ru

structure selectivity of the protein. According to the data obtained by atomic force microscopy, this selectivity can be stipulated by an increased affinity of Dps to branched DNA structures [14]. Such regions can potentially promote stabilizing contacts with additional N-terminal modules of the protein. In such a case, structured RNA molecules, which possess branched elements of the secondary structure and are as efficient for electrostatic interaction as the DNA, are expected to be perfect targets for Dps. The ability of Dps to interact with RNAs is supported by the data [15], which demonstrated a dependency of the form of crystals composed of Dps and the DNA from the presence of RNA. However, the mode of interaction between Dps and RNA has not been studied so far. Thus, the main objectives of the present work were assessment of the ability of Dps to interact with RNA and identification of predominantly involved RNA molecules. A fraction of short RNAs (from 25/9 to 250 nucleotides long) was used as potential binding partners for adequate characterization of the structural capability of the identified targets. The used length range assumes the presence of nearly all types of small structured RNAs, including tRNAs, sRNAs, 5S-rRNAs, as well as a small fraction of extremely short RNAs with a limited potentiality for forming secondary structures.

MATERIALS AND METHODS

Bacterial strains, culturing conditions and Dps purification

Cells of a laboratory strain *Escherichia coli* K-12 MG1655 (U00096.3, *E. coli*) were used for RNA isolation. Bacteria were cultured in liquid Luria-Bertani (LB) medium [16] at 37 °C with shaking. Dps was isolated and purified from *E. coli* BL21* (DE3) cells transformed with pGEM_dps plasmid, which carries the native gene *dps*, applying the purification method proposed in [17].

Isolation of microRNAs

The fraction of total RNA was isolated from *E. coli* K-12 MG1655 cells grown to $OD_{600} = 0.65$. The cells were harvested by centrifugation (12000 rpm, 10 min), the precipitated pellet was dissolved in 1 ml Trizol reagent (Thermo Fisher Scientific, USA) and placed in liquid nitrogen for 15 sec. Following complete freezing, the sample was placed at 37°C and incubated until fully thawed. The freezing-thawing was repeated twice. After the second thawing, 200 mkl of chloroform (Reakhim, Russia) was added, followed by intense shaking for 15 sec and 5-minute incubation at room temperature with periodic shaking. The aqueous phase was separated by centrifugation at 12000 rpm for 10 min at 4 °C and again treated with chloroform. RNA from the aqueous phase was precipitated by the addition of 500 mkl of isopropanol and incubated for 10 min at room temperature. The sample was collected by centrifugation and washed twice with 400 and 200 mkl of 70 % ethanol, respectively. The pellet was dried at 37 °C (10 min) and dissolved in 42.5 mkl of deionized water treated with diethyl pyrocarbonate (PanReac AppliChem, USA). The sample was then incubated with DNAse I (New England Biolabs, UK) according to the manufacturer's protocol. The fraction of short RNAs was obtained from the total RNA sample using mirVana™ miRNA Isolation Kit (Thermo Fisher Scientific, USA) in full compliance with the instructions of the manufacturer. The final concentration of RNA was measured using ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA).

Immobilization of Dps on acrylate spheres and pull-down assay

Anti-Dps antibodies were obtained and purified as described in [18]. Their concentration prior to precipitation with 60 % ammonium sulfate was 3 mg/ml. The suspension of precipitated antibodies was stored at 4 °C. For Dps immobilization, 100 mkl of antibody suspension was precipitated by centrifugation at 12000 rpm (10 min). It was subsequently dissolved in 50 mkl IP buffer composed of a buffer containing 100 mM NaCl, 50 mM Tris-

HCl, pH 8, 5 mM EDTA, 0.2 % NaN$_3$ and 0.5 % sodium dodecyl sulfate, and a buffer containing 100 mM Tris-HCl, pH 8.6, 100 mM NaCl, 5 mM EDTA, 0.2 % NaN$_3$ and 5 % Tritox-X 100, at a 3:1 ratio. All operations were conducted at 4 °C. A total of 680 mkl IP buffer was added to the Dps sample (120 mkl, 0.6 mg/ml), mixed and supplemented with 14 mkl antibodies dissolved in IP buffer. IP buffer (800 mkl) supplemented with 14 mkl antibodies, without the addition of Dps, was used as a control. Both samples were incubated for 14 hours with constant agitation using an automated rotator (Biosan, Latvia), followed by the addition of 60 mkl Protein A/G UltraLink® Resin Beads (Thermo Fisher Scientific, USA) and 2-hour incubation for protein immobilization on acrylate spheres. Spheres were collected by centrifugation (2500 rpm, 2.5 min), twice washed with 200 mkl of TB buffer (50 mM Tris-HCl (pH 8.0), 10 mM MgCl$_2$, 0.1 mM EDTA, 50 mM NaCl) and resuspended in 50 mkl TB buffer with subsequent addition of 20 mkl of the short RNA sample. RNA-protein complexes were allowed to form for 2 hours at 37 °C. After that acrylate spheres were precipitated and washed twice with 200 mkl 50 mM NaCl. The bound RNA was extracted with 50 mkl 200 mM NaCl. Immediately after its addition and mixing the samples were purified using microcolumns from the kit for purification of PCR products (Evrogen, Russia). Purified RNAs were precipitated by addition of 50 mkl isopropanol after 14-hour incubation at –20 °C. The pellet was washed twice with 200 mkl 70 % ethanol, dried and dissolved in nuclease-free water (Promega, USA). The concentration of the resulting RNA samples was measured using ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA).

Library preparation and sequencing

Short RNAs obtained in two independent pull-down assay experiments as described above were used for preparation of sequencing libraries, with two different techniques of library preparation applied. In the first case, universal adapters were ligated to the isolated RNAs using Ion Total RNA-Seq Kit v2 (Thermo Fisher Scientific, USA) and the subsequent sample processing was done using Ion OneTouch™ System (Thermo Fisher Scientific, USA) and Ion PGM Hi-Q OT2 Kit (Thermo Fisher Scientific, USA). In the second case, Illumina ligation system (NEBNext Multiplex Small RNA Library Prep Set, New England Biolabs, UK) was used prior to the attachment of Ion Torrent adapters. This additional step allowed increasing the length of sequences read by Ion Torrent, i.e. providing identification of not only long, but also very short RNAs in complexes with Dps. Subsequently, sequencing was done using an Ion PGM™ System (Thermo Fisher Scientific, USA) with Ion 314™ Chip v2. All operations were carried out in full accordance with the manufacturer's protocols.

Processing of the obtained data

The initial data in BAM format were converted into FASTQ using Samtools v 1.3.1 [19]. Removal of additional adapters in the second experiment was done using Cutadapt algorithm [20]. To maximally reduce the impact of sequencing errors the initial sequences were filtered by reading quality. This was implemented with FASTQ Quality filter, a component of FastX-Toolkit v0.0.14 [21], using a threshold quality level of $Q \geq 13$ which ensures 95 % reliability for 90 % nucleotides in each read.

Reads longer than 25 nucleotides (nt) that passed quality filtering were sliced by fragments with a standard length of 25 nt and mapped on the genome of *E. coli* MG1655 (v.3) using software [22] freely available at http://www.mathcell.ru/DnaRnaTools/Matcher.zip. Mapping 9–14 nt sequences on the genome was done independently for each length, and reads falling into the same position were summarized. Alignment of analyzed sequences on the genome with Matcher accounted their correspondence to several regions of the bacterial chromosome permitting two mismatches over a 25 nt sequence. The profiles obtained for the experimental and control samples were normalized using the approach proposed in [23].

RESULTS

t3

Dps immobilized on acrylate spheres promotes sorption of structured RNAs

Direct ligation of short RNAs, washed off from immobilized Dps and from the adsorbent with only anti-Dps antibody, yielded approximately equal amount of reads (93658 and 106753, respectively). Following quality filtering, the experimental and control sets included 57929 and 63902 reads, respectively. In both cases their lengths varied from 25 to ~ 250 nucleotides, though the mean sizes differed, being 86.1 and 60.1 nt for the experimental and control sets, respectively. After slicing by 25 nt fragments, the experimental set included 182085 fragments, while the control set was composed of 135765 fragments. However, the control set contained many technical reads and alien sequences not related to the bacterial genome, with only 14141 sequences mapped, while 92322 out of 182095 nucleotide sequences from the experimental set were found in the genome.

In theory, this large quantitative difference evidences that immobilized Dps efficiently interacts with RNAs isolated from the bacterial cells. At the same time, the genome-wide distribution of cDNAs in the control sample was not random (internal circle in Fig. 1). Moreover, their distribution profile was highly similar to that of RNAs in the experimental set (second outermost circle in Fig. 1). Thus, the qualitative normalization was absolutely necessary for identification of transcripts with high affinity to Dps.
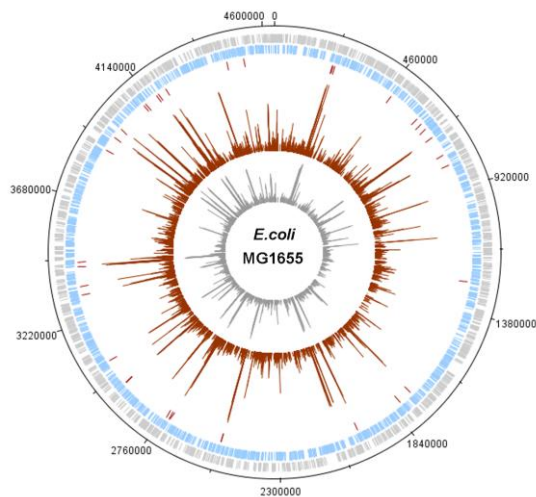

Fig. 1. Genome-wide distribution of genes for short RNAs adsorbed on immobilized Dps (second outermost circle) and acrylate spheres coated with anti-Dps antibody (central circle). $\mathrm{Log}_{10}(N + 1)$ values are shown, where N is the sum of reads in a 50 bp long bin. The two outer circles schematically demonstrate the distribution of genes along the upper and lower genomic strands. Dashes in the third circle denote peaks with at least five times higher amplitude in the experimental sample, as compared to the control, as well as peaks detected only in the experimental set. The scheme was constructed using DNAPlotter v. 1.3. [24].

Several approaches were tested, from the simplest normalization on the number of mapped reads (Fig. 2,A, $K = 6.52$), to more complicated methods, which take into account the profile of the control signals. Despite their overall similarity, the most convincing results were derived according to the procedure originally proposed by Affymetrix for analysis of microarray data [23], and later realized in a number of other methods [25].

In this method the corrected mean values for the control and experimental data sets were calculated by removing ~ 2 % of signals with the highest and lowest intensity. Fig. 2,B shows the results of the optimized normalization after removal of 0.5 % extreme signals. The graph reflects the presence of reads with similar representation in both samples (laying on the bisector line of the graph) and allows for identification of RNAs, the sorption of which is increased in the presence of Dps. RNAs detected only in the experimental sample in the amount of no less than 10 molecules, as well as RNAs present in this sample at least at a five-

fold higher level compared to the control, were selected as potential RNA partners of Dps (Table 1).
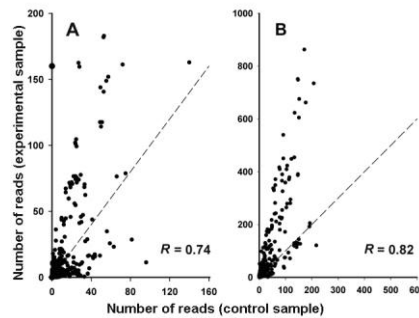
Fig. 2. Correlation between the numbers of mapped reads registered in the experimental and control samples in 50 bp long bins. **A.** The normalization coefficient ($K = 6.52$) was calculated using the total number of sequences mapped on the genome. **B.** The normalization coefficient ($K = 2.94$) was calculated using corrected mean values of the numbers of mapped reads after removal of 0.5 % positions with the highest and lowest amount and only for positions with more than 1 registered sequences. Panels display the values of Pearson correlation coefficient R. Dashed lines show the bisector line of the coordinate angle.

It became clear, that Dps exhibits the highest affinity to transport and small regulatory RNAs. All molecules of these types form stable looped structures, thus confirming the presumed increased affinity of Dps to branched structures in nucleic acids [14]. Moreover, seven RNA products originated from intergenic sites (marked by red in Table 1) were identified in complexes with Dps. This may indicate the presence of novel genes encoding untranslated RNAs. In two other cases the mapped oligonucleotides covered not only an intergenic region, but also the end of a neighboring gene. Such products, in particular, include RNAs synthesized from between convergent genes *cspE* and *crcB* (Fig. 3). The corresponding RNA apparently ends at the terminator of *cspE* gene (encodes a DNA- and RNA-binding protein functioning as an anti-terminator on ρ-independent transcription terminators [26]), but inside its coding sequence this gene carries two promoters for independent initiation of transcription.
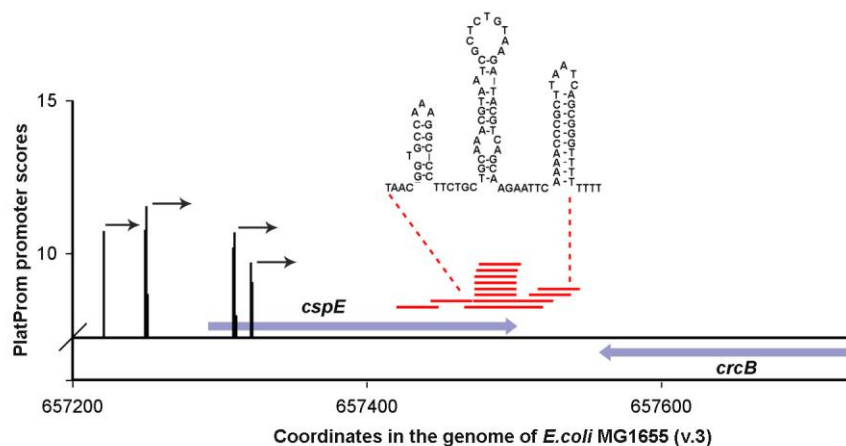
Fig. 3. Disposition of registered reads (red lines) in the genomic locus *cspE/crcB*. Genes and the direction of their transcription are marked by blue arrows. Promoters predicted by PlatProm are denoted by black bars and arrows [27]. The secondary structure of the RNA product shown in the upper part was predicted using RNA Structure [28].

Table 1. RNA products interacting with Dps *in vitro*

t5

Genomic region corresponding to the selected RNAs			Associated genes			Strand	*****Type of RNA product
*Left border	*Right border	**Length (bp)	Gene	***Left border	****Right border		
225499	225582	84	*alaV*	**225500**	**225575**	+	Transport RNA
228921	229009	89	*aspU*	**228928**	229004	+	Transport RNA
236930	237014	85	*aspV*	**236931**	237007	+	Transport RNA
476447	476567	121	*ffs*	**476448**	476561	+	4.5S-RNA
607736	**607793**	58	*sokE*	607734	607792	–	Small RNA
631993	632068	**76**	*cstA*	629894	631999	+	Potential product from the intergenic spacer (+)
			ybdD	632182	632379	+	
657466	657541	**76**	*cspE*	**657292**	657501	+	RNA possessing a promoter in *cspE* and a terminator in the intergenic spacer (+)
			crcB	**657555**	657938	–	
696430	**696507**	78	*glnX*	**696430**	696504	–	Transport RNA
774309	774363	55	*ybgE*	774309	774602	+	mRNA
816914	817023	**110**	*ybhK*	815739	816647	–	Potential product from the intergenic spacer (+)
			moaA	817044	818033	+	
1270399	1270449	51	*rdlC*	1270393	1270460	+	Small RNA
1746435	1746595	161	*valV*	1746435	1746511	+	Transport RNA
			valW	1746516	1746592	+	Transport RNA
1800017	**1800156**	**140**	*rpmI*	**1799802**	**1799999**	–	RNA possessing a promoter in *infC* and a terminator in the intergenic spacer (–)
			infC	**1800096**	1800638	–	
1991887	**1992049**	163	*leuZ*	**1991815**	1991901	–	Transport RNA
			cysT	**1991914**	1991987	–	Transport RNA
			glyW	**1992042**	1992117	–	Transport RNA
2518015	**2518117**	103	*alaX*	**2518041**	**2518116**	–	Transport RNA
2518130	**2518232**	103	*alaW*	**2518156**	**2518231**	–	Transport RNA
2729340	**2729445**	106	*gltW*	2729369	**2729444**	–	Transport RNA
2734153	**2734270**	118	*ryfD*	**2734153**	2734295	–	Small RNA
2746150	**2746255**	106	*ZipA5'*	2746183	2746432	–	Small RNA
2947385	2947466	82	*metZ*	**2947387**	2947463	+	Transport RNA
2947494	2947580	87	*metW*	**2947497**	2947573	+	Transport RNA
2947604	2947686	83	*metV*	**2947607**	2947683	+	Transport RNA
3049082	3049157	**76**	*invR*	**3046901**	3046993	+	Potential product from the intergenic spacer (–)
			gcvH	**3049160**	3049549	–	
3318190	**3318290**	101	*metY*	**3318213**	3318289	–	Transport RNA
3350588	3350699	112	*arcZ*	**3350577**	**3350697**	+	Small RNA
3426928	**3427164**	237	*alaU*	**3426958**	**3427033**	–	Transport RNA
			ileU	**3427076**	**3427152**	–	Transport RNA
3443650	3443712	63	*secY*	3442766	3444097	–	mRNA
3888411	3888534	124	*tnaC*	3888435	3888509	+	*tnaCAB* attenuator
3943383	3943516	134	*gltU*	**3943435**	**3943510**	+	Transport RNA
3982508	3982592	85	*hisR*	**3982509**	**3982585**	+	Transport RNA
4106329	4106390	62	*cpxP*	4105820	4106320	+	Potential product from the intergenic spacer (+)
			fieF	4106469	4107371	+	
4115666	4115718	53	*glpX*	4114569	4115579	–	Potential product from the intergenic spacer (–)
			glpK	4115714	4117222	–	
4168371	4168453	83	*gltT*	4168372	**4168447**	+	Transport RNA
4175663	4175836	174	*glyT*	**4175673**	4175747	+	Transport RNA
			thrT	**4175754**	4175829	+	Transport RNA
4209772	4209855	84	*gltV*	4209774	**4209849**	+	Transport RNA
4463026	**4463054**	29	*nrdD*	4460522	4462660	–	Potential product from the intergenic spacer (–)
			treC	4463054	4464709	–	
4528057	4528119	**63**	*ryjB*	4527977	4528066	+	Potential product from the intergenic spacer (+)
			sgcQ	4528111	4528917	–	

Transcripts corresponding to intergenic regions are marked by red; green denotes the regions not represented among the reads of the control sample.

Bolding:
*– presence of a promoter close to the encoded RNA;
**– the potential RNA is capable of forming secondary structures;
***– 25 nt long products were registered in two experiments;
****– 9–14 nt long products were registered in the second experiment;
***** – the direction of transcription for intergenic products is shown in parentheses.

Therefore, it cannot be excluded that the registered RNAs are synthesized independently from the main mRNA. Being subjected to specific processing, they can function as untranslated RNAs in the cells. Potential promoters for transcription independent from the main mRNAs were found for three other products spanning into intergenic regions (marked by bold font in the first and second columns in Table 1), and six potential novel RNA products can form secondary structures with free energy of folding < –21.8 kKal/mol (marked by bold font in the third column in Table 1).

Thus, tRNAs and small regulatory RNAs have the highest efficiency of binding with immobilized Dps, but most transcripts of other types bound with Dps are also capable of forming looped structures with high free energy of folding.

Dps immobilized on acryate spheres can adsorb short RNAs

Almost all RNAs selected by their high abundance in the experimental set (Table 1) were longer than 50 nucleotides (Table 1). No reads shorter than 25 nt were registered in this experiment. This could have been due either to the fact that short RNAs simply do not bind to Dps and are not adsorbed on acrylate spheres used for immobilization or because such oligonucleotides are poor substrates for the attachment of adapters using Ion Torrent enzyme system. Hence, in the second experiments we used double adapters. At the first step, RNAs washed off from acrylate spheres and immobilized Dps were ligated with adapters used for sequencing on the Illumina platform (reagent kit NEBNext Multiplex Small RNA Library Prep Set, New England Biolabs). Following the obtainment of the first and second DNA copies, Ion Torrent adapters were attached, and the resulting fragments were sequenced. This strategy allowed identifying extremely short transcripts between the internal adapters.
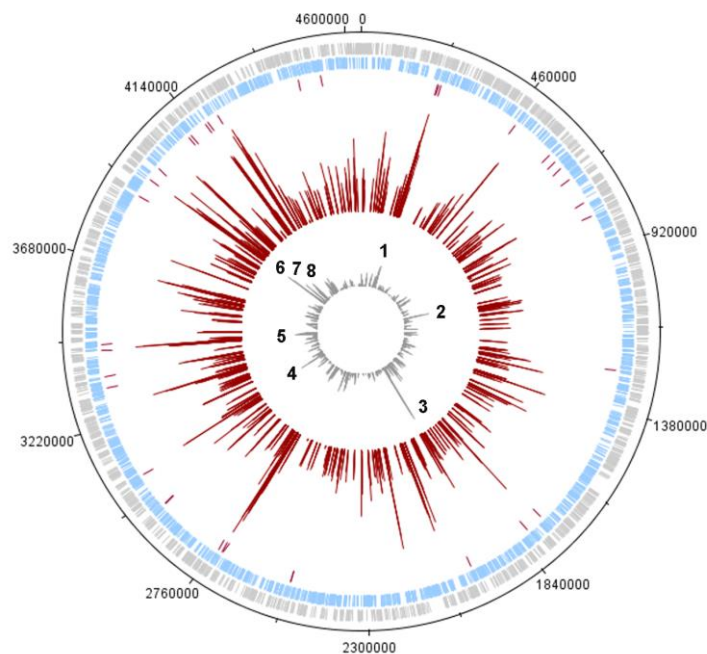

Fig. 4. Genome-wide distribution of long (25 nt) and short (9–14 nt) RNAs adsorbed on immobilized Dps (the central and second outermost circles, respectively). $Log_{10}(N + 1)$ values are shown, where N is the sum of reads in a 50 bp long bin. The two outer circles schematically demonstrate the distribution of genes along the upper and lower genomic strands. Dashes in the third circle denote genomic areas selected based on the results of the first experiment (Fig. 1 and Table 1). The scheme was constructed using DNAPlotter v. 1.3. [24].

This experiment provided 3694 sequence reads in the immunoprecipitated sample, whereas the amount of RNAs in the control specimen was insufficient for subsequent sequencing. The length of sequences eluted from immobilized Dps varied from 2 to 231

nucleotides (mean length 33.3 nt). After filtering of reads by quality and sorting by size we obtained 2055 fragments with a length of 25 nt and 363 sequences ranging in size from 9 to 14 nucleotides. The second outermost circle in figure 4 demonstrates the distribution of long fragments over the genome. It is obvious that most peaks correspond to RNAs detected in complex with Dps in the first experiment (corresponding regions are marked by bold font in the fifth and sixth columns in Table 1). They include, in particular, the RNAs from the genomic locus *cspE*/*crcB*.

The distribution of 9–14 nt long reads over the genome differed from that observed for longer RNAs (the central circle in Fig. 4), though they are congruent to each other to a certain extent. Eight dominant peaks can be observed. Seven of them also correspond to transport and regulatory RNAs. Peak 2 makes an exception, falling into the middle part of *asnS* gene which encodes asparaginyl-tRNA synthase, an enzyme involved in tRNA biosynthesis. However, the registered reads corresponded to the 63 nt long sequence also capable to form looped structures with free energy of folding –17.5 kKal/mol. Reads constituting the largest peak 3 are of special interest. They correspond to two small untranslated RNAs RyeA and SdsR with poorly studied properties, which are transcribed from both strands of the same genomic locus (Fig. 5).
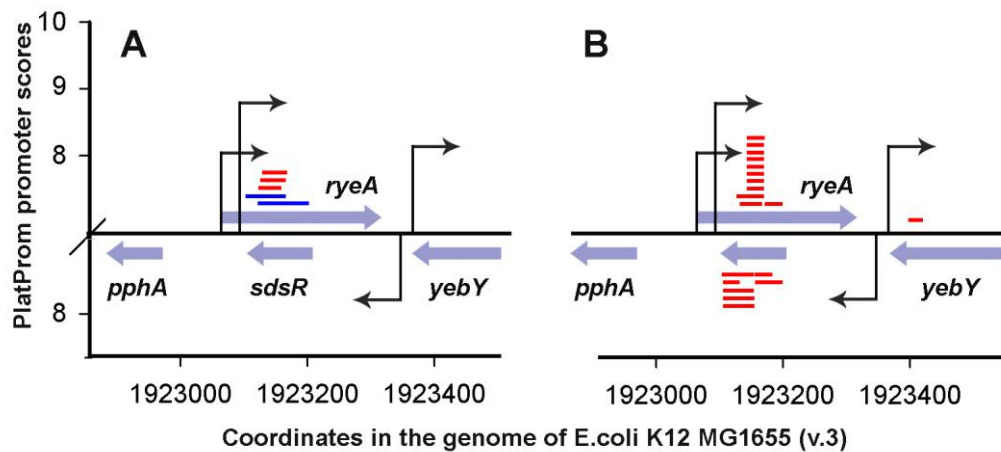
Fig. 5. Distribution of 25 nt long (**A**) and 9–14 nt long (**B**) reads in the genomic region *ryeA*/*sdsR* registered in the first and second experiments (blue and red lines, respectively). Genes and the direction of their transcription are marked by blue arrows. Promoters predicted by PlatProm are indicated by black bars with arrows [27].

Interaction with 25 nt long RNAs originating from this locus was registered in both experiments (Fig. 5,A), though only short reads in the second experiment formed a significant peak in this region (Fig. 5,B). Short reads were registered at both strands (Fig. 5,B), assuming their belonging to different transcripts, though the use of double adapters does not allow for absolute certainty. However, the obvious predominance of short products testifies posttranscriptional processing of the initial RNAs. Both RNAs are known to be of unstable length (249 and 150 nt long isomers were previously registered for RyeA, and two transcripts with 104 and 74 nucleotides were detected for SdsR [29, 30]), suggesting their posttranscriptional processing. This processing may be mediated by the formation of a partly complementary duplex not inside a single RNA, but rather with the involvement of two independently transcribed products [30, 31]. It cannot be excluded that the ultimate objective of this processing is the formation of extremely short RNA duplexes resembling microRNAs of higher organisms, which are not typical for bacteria.

DISCUSSION

The main objectives of this study consisted in testifying the ability of Dps to interact with RNAs and characterizing the transcripts binding to the immobilized protein. Dps is a large

t8

protein composed of 12 identical subunits each weighing 18.695 kDa. Its interaction with nucleic acids is mediated by positively charged side groups of lysine residues located at flexible N-terminal modules exposed on the negatively charged surface of the spherical protein globule. Thus, the risk of unspecific sorption of nucleic acids on the protein surface appeared to be minimal. Nevertheless, we faced the problem of unspecific RNA sorption on the acrylate matrix with the bound anti-Dps antibody molecules. Probably this is due to the positive charge on the antibody surface, which could have possibly been selected to a negatively charged epitope by the immune system.

The similarity between the genome-wide distribution profiles for the reads of the control and experimental samples (Fig. 1), which is largely stipulated by the presence of transport RNAs in both sets, is less clear. Since not a single peak standing for a tRNA (marked by green in Table 1) was selected for the absence of reads in the control set, their dominance can be most easily explained by a high concentration of tRNA molecules in the cells. Ribosomal proteins and RNAs inevitably contaminate all biological specimens, and rRNA and tRNA removal may be done using special techniques, which however, were not applicable in this exploratory study. In view of this, the reasons for the complete absence of degradation products of large ribosomal RNAs among the selected partners are not quite clear.

The amount of molecules adsorbed on the control spheres in the first experiment was much lower, than in the case of Dps-coated spheres, and in the second experiment we failed to obtain the fraction of unspecifically bound RNAs at all. Thus, we presume that strict criteria, applied for identification of RNAs capable to bind Dps, allow for concluding that Dps can interact with RNAs, at least *in vitro*, and the main targets of this protein are structured molecules that form stable looped structures. If this is the case, the biological role of Dps needs to be reconsidered. Being one of the main structural proteins of the bacterial nucleoid, Dps packs the genome to a crystal-like structure under nutrient-deficient conditions [32] and utilizes ferroxidase activity to defend the DNA from the destructive impact of reactive oxygen species [33]. Specific affinity of Dps to structured RNAs increases its functional potential, though the biological rationale of the combination of these functions is yet to be comprehended.

REFERENCES

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.

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.

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.

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.

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.

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.

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.

9. Dorman C.J. H-NS, the genome sentinel. *Nat. Rev. Microbiol.* 2007. V. 5. P. 157–161. doi: 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.

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.

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.

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.

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.

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.

16. Sambrook J., Fritsch E.F., Maniatis. T. *Molecular cloning: a laboratory manual, 2nd ed*. New York: Cold Spring Harbor Laboratory Press, 1989.

17. Pokusaeva V.O., Antipov S.S., Shvyreva U.S., Tutukina M.N., Ozoline O.N. Overexpression, isolation and purification of functionally active *E. coli* bacterioferritin Dps. *Sorption and Chromatographic Processes* (in Russian). 2012. V. 12. P. 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.

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.

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.

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.

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.

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.

28. Mathews D.H. RNA secondary structure analysis using RNAstructure. *Current Protocols in Bioinformatics*. 2014. V. 46. doi: 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.

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.

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.

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.

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.