



Excessive Promoters as Silencers of Genes Horizontally Acquired by *Escherichia coli*

Aleksandr Bykov^{1,2}, Olga Glazunova¹, Olga Alikina¹, Natalia Sukharicheva¹, Irina Masulis¹, Konstantin Shavkunov¹ and Olga Ozoline^{1*}

¹ Laboratory of Functional Genomics and Cellular Stress, Institute of Cell Biophysics of the Russian Academy of Sciences, Pushchino Scientific Center for Biological Research, Moscow, Russia, ² Research Institute of Molecular Pathology, Vienna Biocenter Campus, Vienna, Austria

OPEN ACCESS

Edited by:

Tatiana Venkova,
Fox Chase Cancer Center,
United States

Reviewed by:

Dale Eugene Alexis Lewis,
National Institutes of Health (NIH),
United States
Akira Ishihama,
Hosei University, Japan
Gloria Del Solar,
Spanish National Research Council,
Spain

*Correspondence:

Olga Ozoline
ozoline@rambler.ru

Specialty section:

This article was submitted to
Molecular Recognition,
a section of the journal
Frontiers in Molecular Biosciences

Received: 18 October 2019

Accepted: 10 February 2020

Published: 26 February 2020

Citation:

Bykov A, Glazunova O, Alikina O, Sukharicheva N, Masulis I, Shavkunov K and Ozoline O (2020) Excessive Promoters as Silencers of Genes Horizontally Acquired by *Escherichia coli*. *Front. Mol. Biosci.* 7:28. doi: 10.3389/fmolb.2020.00028

Horizontally acquired genes are usually transcriptionally inactive, although most of them are associated with genomic loci enriched with promoter-like sequences forming “promoter islands.” We hypothesized that lateral DNA transfer induces local mutagenesis, accumulating AT base pairs and creating promoter-like sequences, whose occupancy with RNA polymerase and a specific silencer H-NS suppresses the transcription of foreign genes. Error-prone mutagenesis was implemented for the “promoter island” of a foreign gene *appY* and the promoter region of an inherent gene *dps*. Derivatives with changed transcriptional activity were selected using a reporter plasmid pET28_eGFP. Only one cycle of mutagenesis with negative selection suppressed the activity of the main *dps* promoter to the background level due to a single substitution in its -10 element, while positive selection gave a sequence with improved -35 element, thus testifying feasibility of the approach. The same suppression for *appY* was achieved by three cycles, while eightfold transcription activation required nine iterations of mutagenesis. In both cases, the number of potential start points decreased resulting in an ordinary regulatory region with only one dominant promoter in the case of positive selection. Efficiency of H-NS binding remained virtually unchanged in all mutant constructs. Based on these findings we conclude that excessive promoters can adversely affect transcription by providing a platform for interference between several RNA polymerase molecules, which can act as a silencer at promoter-dense regions.

Keywords: horizontal gene transfer, gene silencing, promoter islands, transcription, H-NS, error-prone mutagenesis, *gfp* reporter assay

INTRODUCTION

Horizontal gene transfer plays a pivotal role in bacterial evolution assisting in the adaptation of microbes to the environment and increasing the diversity of their populations (Gogarten et al., 2002; Wiedenbeck and Cohan, 2011; Cordero and Polz, 2014). At least five mechanisms allow bacteria to capture alien genetic material, including conjugation (Bañuelos-Vazquez et al., 2017; Delavat et al., 2017), transduction (Keen et al., 2017), transformation (Overballe-Petersen et al., 2013) and transport within either outer membrane vesicles (Tran and Boedicker, 2017) or phage-like particles (Lang et al., 2012; Grüll et al., 2018). Escaping bacterial defense systems, fragments

of alien DNA with a certain probability can incorporate into the genome of a new host, where they can be identified based on the contextual difference from the rest of the nucleotide sequence (Lawrence and Ochman, 1998; Nakamura et al., 2004; Price et al., 2008; Langille and Brinkman, 2009; Huang et al., 2012). Although it is still unknown how bacteria integrate foreign genes into their regulatory networks, the recombinant areas turned out to be enriched with AT base pairs (Daubin and Ochman, 2004), and increased frequency of promoter-like sequences has been already regarded as a signature of foreign genes (Huang et al., 2012).

A typical regulatory region of bacterial genes contains one or several overlapping promoters with one or several transcription start points (TSPs) in each (Gama-Castro et al., 2016). However, there is a tendency to initiate transcription from a single site, which is sometimes regulated by superimposed promoters recognized by different σ -factors (Panyukov and Ozoline, 2013). Hence, it was surprising to find 78 “promoter islands” with extremely high density of potential TSPs, using a promoter finder PlatProm (Shavkunov et al., 2009). All these “islands” formed complexes with RNA polymerase and initiated synthesis of short oligonucleotides, whereas full-length transcription was barely detected (Panyukov and Ozoline, 2013; Panyukov et al., 2013). The biological expediency of such suppression became clear after it turned out that 75 out of 78 “islands” were associated with genes horizontally acquired by *Escherichia coli* (*E. coli*).

Most transferred genes are useless for the cell, and bacteria have elaborated mechanisms for their silencing. In *E. coli* this function is performed by a specific sentinel, a histone-like protein H-NS (Lucchini et al., 2006; Dorman, 2007), which inhibited transcription from all of the tested “promoter islands” (Purtov et al., 2014). Given that association with RNA polymerase is a general mode of transcription repression by H-NS (Oshima et al., 2006), excessive promoters may emerge near alien genes due to spontaneous mutagenesis evolutionarily aimed to create a platform for this combined binding. In this case, an increase in the transcriptional activity of the “islands” should be accompanied by a decrease in the number of potential binding sites for RNA polymerase and/or H-NS. Here we confirmed this hypothesis for RNA polymerase using error-prone PCR and a reporter plasmid for the selection of mutated genomic regions with increased and decreased promoter activity.

MATERIALS AND METHODS

Bacterial Strains

Transcriptional activity of promoter mutants was estimated in *E. coli* Top 10 cells. Model DNA fragments were amplified from the genome of *E. coli* K12 MG1655 (GenBank NC_000913.3). Cells of *E. coli* BL21(DE3) or *E. coli* BL21(DE3) Δ *hns* transformed with pGEM_H-NS-His expression vector (Tutukina et al., 2015) were used to purify recombinant H-NS or to obtain cellular lysates enriched with H-NS-His.

Plasmids and DNA Fragments

Plasmid pET28b-EGFP with the *gfp* gene encoding green fluorescent protein (Masulis et al., 2015) was used as

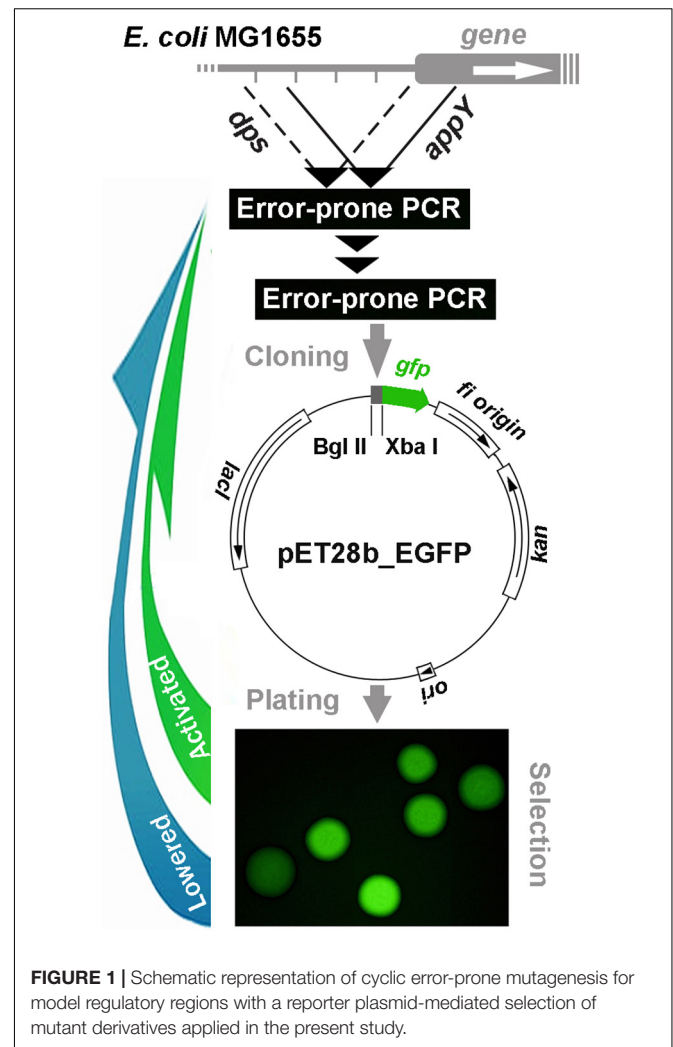


FIGURE 1 | Schematic representation of cyclic error-prone mutagenesis for model regulatory regions with a reporter plasmid-mediated selection of mutant derivatives applied in the present study.

a reporter vector (Figure 1). At the preliminary step, the fragments subjected to mutagenesis were amplified with Taq DNA polymerase (Evrogen, Russia) from the purified genomic DNA of *E. coli* MG1655 using Biometra T1 thermocycler (Germany). A fragment containing the regulatory region of the *dps* gene (402 bp) was obtained with primers: 5'-TCCTCTAGATGTTATGTCCCAGT-3' and 5'-GGAAGATCTTCCTCGGAGAAACACT-3' (underlined are restriction sites for *Xba*I and *Bgl*II, respectively). A fragment with the *appY*-associated “promoter island” (423 bp) was amplified with primers: 5'-GATAAGATCTGCAAGTAAAAATGATACTC-3' and 5'-CCCTCTAGATTTGTGCGCTTACAATAAA-3'. All PCR reactions were carried out using a standard protocol: 2 min melting at 95°C followed by 35 cycles: 95°C, 30 s; 55°C, 30 s; 72°C, 1 min.

Error-Prone PCR Mutagenesis

After purification of DNA samples with QIAquick PCR Purification Kit (Qiagen, Germany), error-prone PCR was carried out with GeneMorph II Random Mutagenesis Kit (Agilent, United States). Without purification, 0.1 ng of amplicon

samples were used for a second stage of PCR. Thereafter, amplicons were purified with QIAquick PCR Purification Kit and prepared for cloning into pET28_EGFP vector using FastDigest *Xba*I and FastDigest *Bgl*II nucleases (Thermo Fisher Scientific, United States). Restriction was carried out in universal FastDigest buffer for 1 h at 37°C. Restriction fragments were purified with QIAquick PCR Purification Kit, while the pET28_EGFP vector processed in parallel was fractionated by gel electrophoresis in 1% agarose and purified with MinElute Gel Extraction Kit (Qiagen, Germany). Restriction fragments were ligated with T4 Ligase (Thermo Fisher Scientific, United States) immediately upstream of the *gfp* gene (12 h at 4°C) and the library of recombinant pET28b_eGFP (**Figure 1**) was used for transformation. Transformed cells were plated on 1.5% LB agar and incubated at 37°C overnight in the presence of kanamycin (80 µg/ml). GFP fluorescence was measured for individual colonies (exemplified in **Figure 1**) using Leica DM6000B fluorescent microscope (Germany) with excitation/emission at 480/510 nm and ImageJ software¹. Several colonies with the highest and lowest level of fluorescence were selected and cultured for 12 h in LB broth, followed by isolation of plasmid with Plasmid Miniprep kit (Evrogen, Russia). Amplicons of target sequences were obtained with Taq polymerase (Evrogen, Russia) and sequenced. The mutated inserts with the largest number of new spontaneous mutations were used for a subsequent cycle of random mutagenesis, and the whole procedure was repeated until the GFP fluorescence was not distinguishable from the previous round.

Fluorescence Measurements

The expression level of *gfp* in up-regulated constructs was compared to the *E. coli* cells transformed with corresponding pET28b-eGFP plasmids. For visualization, cells in LB medium (1 µl) freshly transformed with each of the constructed vectors were plated on one Petri dish and grown overnight before imaging. For quantitative measurements, cells were cultured for 12 h in 4 ml of LB medium, sedimented, suspended in 1 ml of TE buffer (pH 7.5) and sonicated (Misonix, United States). Cellular debris was sedimented and fluorescence intensity of GFP was measured in supernatant using a fluorescent spectrophotometer Cary Eclipse (Varian, Australia). The total protein concentration of the supernatant measured with a NanoDrop ND-1000 spectrophotometer (United States) was used for normalization.

In silico Analysis

The distribution of potential TSPs was analyzed by a promoter finder PlatProm² (Shavkunov et al., 2009). The binding sites for 109 transcription factors functioning in *E. coli* (listed in **Supplementary Image 1**) were searched by using RSAT full options matrix-scan software with *E. coli* K12 residue probabilities³ (Turatsinze et al., 2008). The position weight matrices for motifs recognized by 82 transcription factors were built using aligned sequences of their binding sites

collected in RegPrecise⁴ (Novichkov et al., 2013). Matrices for 26 transcription factors, including H-NS, were taken from the Virtual Footprint collection⁵ (Münch et al., 2005), while the matrix for the nucleoid protein Dps was obtained using original ChIP-seq data (Antipov et al., 2017). Sequence Logos for **Supplementary Image 1** were generated by WebLogo 3⁶ (Crooks et al., 2004).

Primer Extension

The total RNA was isolated from bacterial cells transformed with pET28b-EGFP containing different inserts as described (Masulis et al., 2015). Ten micrograms of RNA and 2 pmol of a ³²P labeled *gfp*-specific primer 5'-CTCTGGTCAGGCAGATACTCTGGTCAG-3' were used for the synthesis of cDNA by RevertAid Premium reverse transcriptase (Thermo Fisher Scientific, United States). Samples were treated with RNase A (Thermo Fisher Scientific; 10 U, 37°C, 30 min), precipitated with threefold volume of 96% ethanol and 0.3M sodium acetate and washed with 70% ethanol. The precipitate was dissolved in 5 µl of 98% formamide with 8 mM NaOH and 4 mM EDTA, fractionated in 8% polyacrylamide gel (PAAG) with 8M urea and radioautographed.

SDS PAGE Electrophoresis

SDS electrophoresis was used to assess the purity of H-NS. Stacking gel contained 4% acrylamide, 0.125M Tris-HCl (pH 6.8), 0.1% SDS. Separating gel contained 12.5% acrylamide, 0.375 M Tris-HCl (pH 8.8), 0.1% SDS. Electrophoresis buffer included 25 mM Tris-HCl, 250 mM glycine and 0.1% SDS (pH 8.3). Ammonium persulfate and TEMED were added to a final concentration of 0.004 and 0.003%, respectively. Sample loading buffer contained 50 mM Tris-HCl (pH 6.8), 100 mM β-mercaptoethanol, 1% SDS, 0.01% bromophenol blue, and 10% glycerol. After sample loading the gels were run at constant current of 25 mA in stacking gel and 45 mA in separating gel. Molecular weight markers used were PageRuler Unstained Low Range Protein Ladder (Thermo Scientific, Lithuania) and Blue Prestained Protein Standard, Broad Range (NEB, United States). Following electrophoresis, the gels were stained in 0.25% Coomassie Brilliant Blue R-250 (Sigma, United States) with 45% ethanol and 10% acetic acid (Khimmed, Russia) and washed in solution containing 40% ethanol and 7% acetic acid.

Electrophoretic Mobility Shift Assays (EMSA)

The original and mutagenized DNA fragments were amplified from isolated plasmids, purified, and used (0.5 pmol per reaction) to form complexes with H-NS under two different experimental settings. Complexes with purified H-NS were formed in 20 µl of buffer containing 5 mM Tris-HCl (pH 8.0), 1 mM MgCl₂, 50 mM NaCl, 0.01 mM EDTA, 0.01 mM DTT and 3.3 mM imidazole (present in storage buffer of H-NS to prevent oligomerization)

¹<https://imagej.net/Welcome>

²<http://www.mathcell.ru/model6.php?l=en>

³http://embnet.ccg.unam.mx/rsat/matrix-scan-quick_form.cgi

⁴<http://regprecise.lbl.gov/RegPrecise>

⁵http://www.prodoric.de/vfp/vfp_promoter.php

⁶<http://weblogo.threeplusone.com/create.cgi>

for 40 min at 30°C. Following the addition of glycerol to the final concentration of 10%, the samples were loaded on pre-warmed 5% PAAG. The gels were run at constant voltage of 280 V. Alternatively, H-NS protein was obtained from IPTG-induced (100 μ M) *E. coli* BL21(DE3) Δ *hns* cells transformed with the pGEM_H-NS-His expression vector. Cells were harvested 4 h after the induction, washed, sonicated, and cellular debris was sedimented, while the supernatant containing 0.36–0.4 μ g/ μ l of total protein was used for complex formation. Bands were stained with AgNO₃ (pure protein) according to the protocol (Merril et al., 1979) or analyzed by Western blotting (cell lysates) using anti-His-tag antibodies (Cell Signaling Technology, United States). Cross-reactivity test with anti-His antibodies using a plasmid-less lysate of cells with deleted *hns* gene has been shown to be negative in a previous publication (Tutukina et al., 2015). The intensity of complexes formed by mutagenized derivatives was compared to that of native fragments using ImageJ. SigmaStat package of SigmaPlot was implemented for statistical analysis⁷.

RESULTS

Selected Genomic Regions

Two H-NS-dependent promoter regions were selected for error-prone mutagenesis. One of them belongs to the inherent *E. coli* gene encoding the nucleoid protein Dps (Figures 2A–C). It was selected to assess the capacity of random mutagenesis to affect the activity of normal promoters. Having no “promoter island,” this regulatory region contains the main promoter P_{dps} (Altuvia et al., 1994), three weak promoters P_1 , $P_{1'}$, P_2 and a distal promoter P_3 , which affects the strength of P_{dps} (Shvyreva et al., 2011). Nucleoid proteins Fis and H-NS, as well as local repressor MntR are inhibitors of P_{dps} , whereas OxyR and IHF activate this promoter upon oxidative stress and/or transition to stationary growth (Gama-Castro et al., 2016).

The main model sample (Figures 3A–C) was taken from the “promoter island” associated with the *appY* gene, whose lateral transfer was justified by four different *in silico* approaches (Lawrence and Ochman, 1998; Nakamura et al., 2004; Langille and Brinkman, 2009; Huang et al., 2012). The gene *appY* encodes a transcription activator of at least two horizontally transferred operons – *appCBXA* and *hyaABCDEF* (Gama-Castro et al., 2016). The fragment contains 171 TSPs (Figure 3B) predicted on both strands by the promoter finder PlatProm (Shavkunov et al., 2009), but only those with TSPs located 114 bp (*appYp*) and 25 bp ($P_{\sigma 38}$) upstream from the ATG codon were previously identified as functional by high-throughput techniques (Huerta and Collado-Vides, 2003; Maciag et al., 2011, respectively). The expression of the *appY* gene is inhibited by H-NS (Atlung et al., 1996; Purtov et al., 2014) and DpiA (CitB) (Ingmer et al., 1998) and can be activated by ArcA (Lynch and Lin, 1996). However, the binding site was experimentally identified only for ArcA (Lynch and Lin, 1996). Since PlatProm predicts TSPs, by taking into account the presence of promoter-specific

motifs in a wide flanking area (from –250 to +150 bp), the transfer of model regulatory regions into a plasmid slightly changed the profiles of potential TSPs at the borders of integration, which is shown in the main panels and inserts of Figures 2B, 3B. Due to the multiplicity of promoters in both model sequences, all positional coordinates are further referred in accordance with the ATG codons of the *dps* and *appY* genes. The sequences of both genomic regions are given in Supplementary Data Sheets 1, 2.

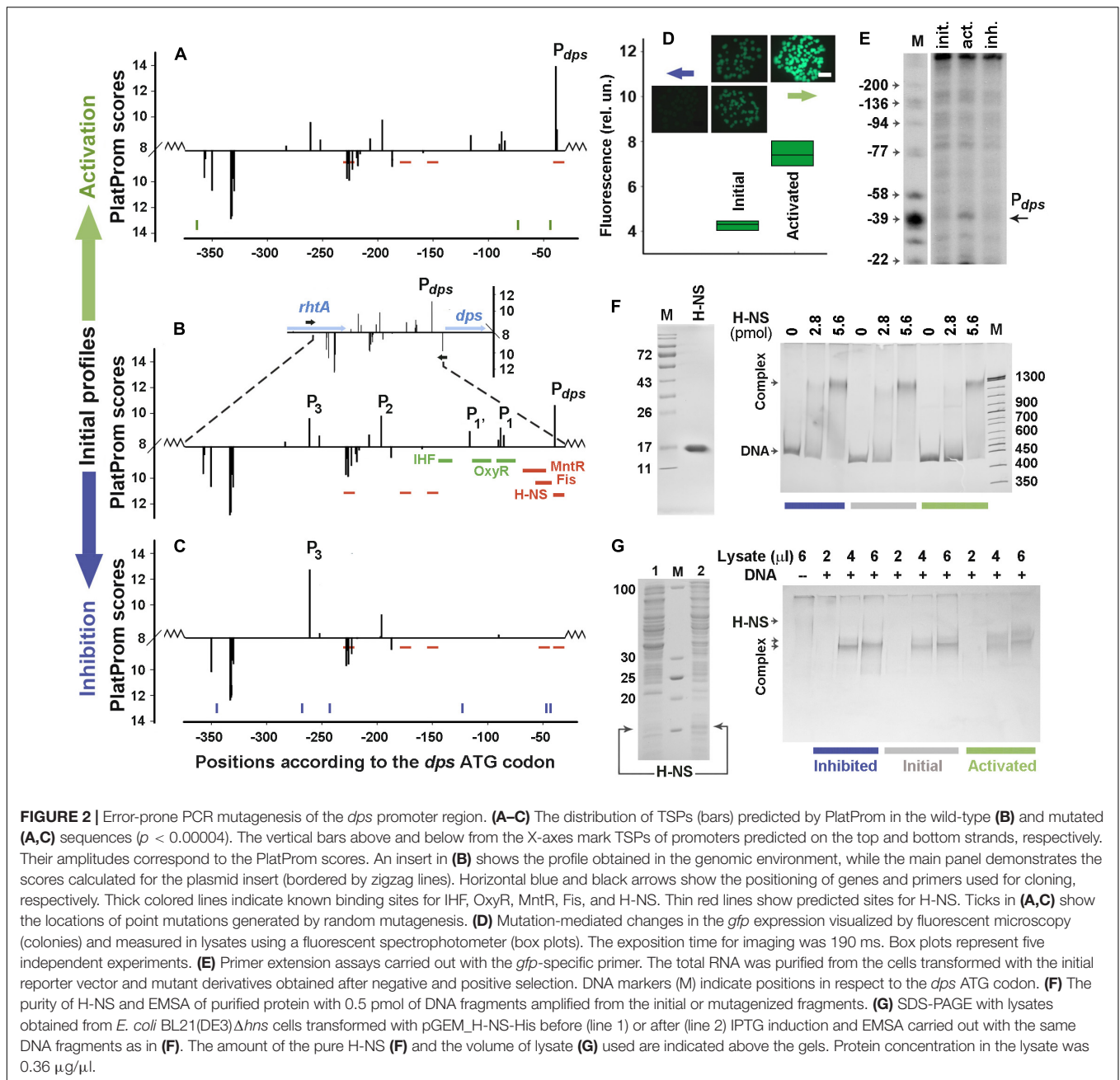
Error-Prone Mutagenesis Affected the Strength of P_{dps}

Only one round of negative or positive selection was sufficient to lower the transcriptional activity of the *dps* regulatory region to the background level or to activate the expression of the reporter gene (Figures 2A–D). Six mutations were obtained in the down-regulated derivative (blue ticks at the bottom of Figure 2C and sequences in Supplementary Data Sheet 1). One of them (position –47), which is 8 bp upstream of the P_{dps} TSP, lowered its score from 10.6 to 6.6 by replacing a conservative T in the –10 element of this promoter with A (TATACT \rightarrow TATACA, consensus: TATAAT). This also created a weak ($p = 9.0 \times 10^{-4}$) H-NS binding site in position –54 (Figure 2C). Another mutation (123 bp upstream of ATG) simultaneously disturbed the up-element of P_1 and the –10 element of $P_{1'}$ (TTTAGT TTTT \rightarrow TTTAGTTGT and TAGTTT \rightarrow TAGTTG, respectively). Substitution A \rightarrow T in position –268, on the contrary, improved the –10 element of P_3 (TAACCA \rightarrow TAACCT), but without a compensatory effect on the transcriptional activity reduced by the two aforementioned mutations. This independence of *gfp* expression from P_3 in the plasmid pET28b_EGFP corresponds to our previous data, indicating that P_3 and P_2 are required for maximum transcription of *gfp*, but cannot ensure its expression in the absence of $P_{1'}$, P_1 , and P_{dps} (Shvyreva et al., 2011).

Positive selection was attained by only three substitutions (green ticks in Figure 2A and sequences in Supplementary Data Sheet 1). One of them fell into the –35 element of P_{dps} (position –73) and shifted its context closer to the consensus sequence (TAGCGG \rightarrow TTGCGG, consensus: TTGACA). Primer extension assays confirmed the activation of P_{dps} (Figure 2E). The other two mutations (at positions –44 and –364) did not affect specific modules in any promoter. Moreover, the substitution at –44 turned out to be the same (A \rightarrow G) as in the suppressed construct (Figures 2A,C and Supplementary Data Sheet 1). Thus, the substitution in the –35 element of P_{dps} is most likely responsible for a twofold increase in the *gfp* fluorescence (Figure 2D).

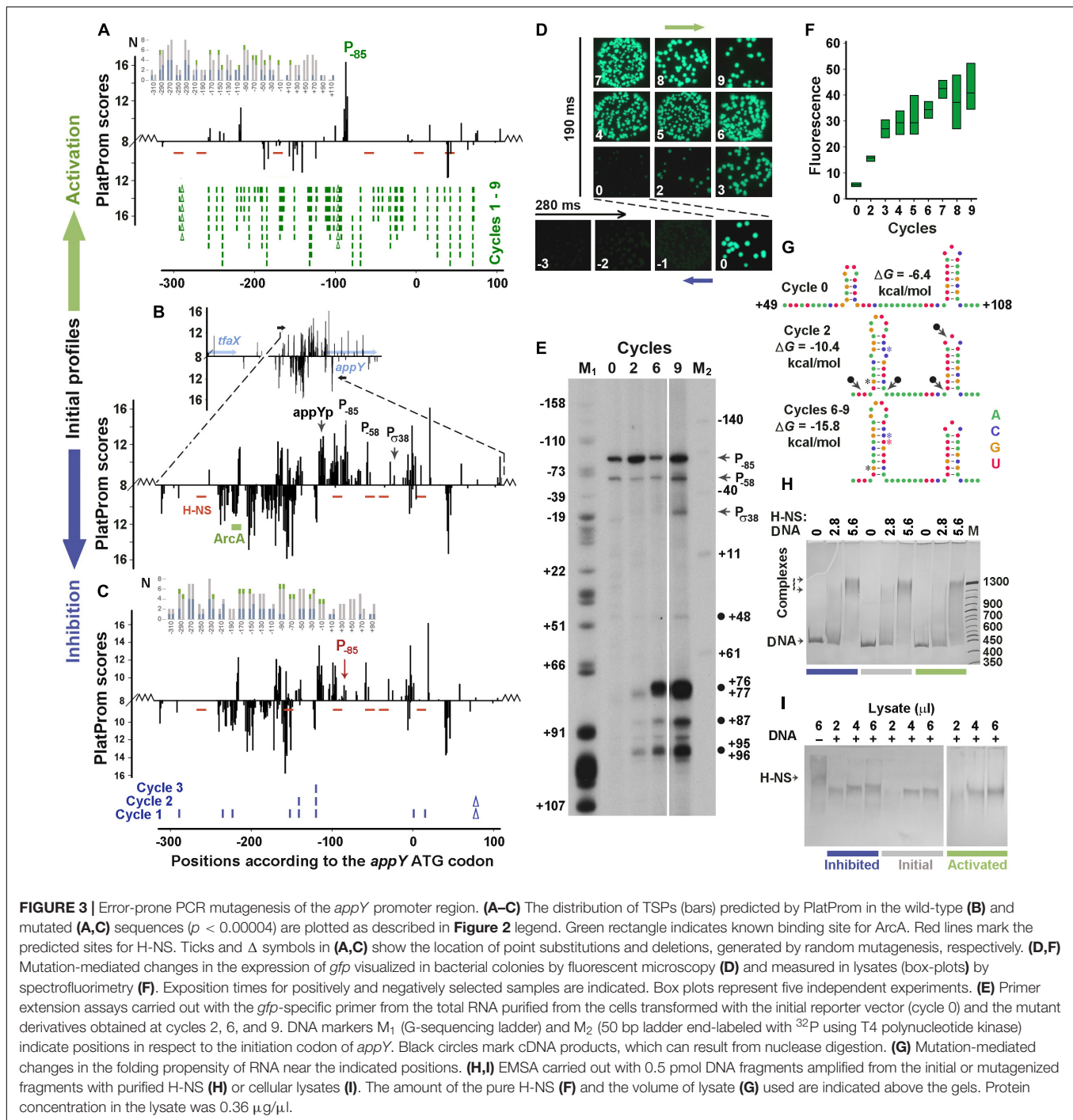
Since in the down-regulated derivative mutation at position –47 increased the conformity to the context of H-NS binding sites, we compared the H-NS affinity to both mutated constructs with that of the initial fragment by EMSA (Figures 2F,G). When pure H-NS protein (left panel in Figure 2F) was used, all fragments were retained in smeared complexes with electrophoretic mobility being highly dependent on the concentration of the protein. Densitometry analysis indicated

⁷<http://www.sigmaplot.co.uk/products/sigmaplot/statistics.php>



that the binding efficiency of the activated mutant remained the same as that of the initial fragment ($100.2 \pm 4.1\%$, $n = 6$), while interaction with the down-regulated construct showed only statistically insignificant tendency for increase ($117 \pm 13.7\%$, $n = 6$). As promoter mutants were selected *in vivo* (Figure 1), certain contribution to the expression of *gfp* can be provided by some cellular agents absent *in vitro*. We therefore performed parallel EMSA experiments using the same DNA samples with lysates obtained from bacterial cells overproducing H-NS (left panel in Figure 2G). In this case, all fragments tend to retain the protein in two complexes (Figure 2G) indicating that cellular components indeed affect the H-NS binding to the

model fragments, but the total amount of the protein bound to the transcriptionally inactive construct was again approximately the same as for the native fragment ($112.8 \pm 4.7\%$, $n = 6$). Surprisingly, the up-regulated derivative, whose substitutions did not change any of H-NS binding sites, reproducibly formed the largest complex with a higher efficiency than the two other fragments (Figure 2G) and demonstrated a slightly higher affinity to the protein ($128.9 \pm 9.1\%$, $p = 0.537$, $n = 6$). Assuming the slight difference between the smaller and larger complexes cannot be mediated by binding of a large protein, like RNA polymerase, we suggested that the mutation-mediated increase in binding may be due to some other transcription



factors. Thus, we searched for potential binding sites of 109 regulatory proteins with RSAT matrix-scan software and found them for Dps at position $-370/-363$ ($p = 4.7\text{e-}4$) and AraC at position $-94/-57$ ($p = 8.1\text{e-}4$). Of these binding sites, the one with increased affinity for the AraC dimer (monomer MW – 33.4 kDa) may explain the observed change in the mobility of the complexes. Thus, it became clear that the functionality of the ordinary promoter region is highly sensitive to random mutagenesis and quite expectedly accumulated changes in the

context of key promoter elements in response to positive or negative selection.

Mutation-Mediated Activation of the *appY* Regulatory Region Eliminated Excessive Promoters

Both negative and positive selection reduced the number of potential promoters in the *appY*-associated “promoter island”

(Figures 3A–C, sequences in **Supplementary Data Sheet 2** and PowerPoint presentation in **Supplementary Presentation 1**). The expression of *gfp* was suppressed to the background level by one deletion and eight point substitutions obtained in three cycles of negative selection (Figures 3A,C,D). Many potential promoters stayed unaffected, but the score of the promoter with a maximum yield (primer extension assay in Figure 3E, cycle 0) decreased due to the point substitution in its –35 element TTGCAA → CTGCAA (TSP at position –85). Thus, just as for the *dps* regulatory region (Figure 2), a mutation affecting the most active promoter was obtained in the first cycle of mutagenesis, indicating the highest dependence of transcriptional activity on the context of the dominant promoter, even in a promoter-dense region.

Nine rounds of mutagenesis were required to activate this area to the maximal level (Figures 3D,F). This generated 49 substitutions with two deletions and turned the “promoter island” into an ordinary regulatory region possessing only one dominant promoter cluster (Figure 3A and PowerPoint Presentation 1 in **Supplementary Materials**, step by step demonstrating all successive changes). Predominantly selected among many other promoter-like sequences in the native construction (Figure 3E, cycle 0), this cluster remains active in all derivatives. At the end of the experiment, its score increased from 14.76 to 16.8 due to two point mutations and one deletion that changed the –10 element to a nearly consensus sequence (TAAAAAT → TATACT), while the scores of many other promoter-like sequences decreased. At the same time, two mutations, obtained in the first ($T_{+70} \rightarrow C$) and the second ($A_{+54} \rightarrow G$) cycles, extended the hairpin originally formed in the region of +60/ +71 to +53/ +76, and the substitution $C_{+71} \rightarrow T$ ensured the perfect complementarity in this hairpin (Figure 3G and **Supplementary Data Sheet 2**). Being subjected to cleavage by cellular stem-loop – specific endonucleases (for instance RNase E or RNase III), such structures may give products marked by black circles in Figure 3E (near positions +48, +87 and +95/96). In addition, early termination of reverse transcription at the 3'-ends of the large hairpin can explain the appearance of the main product initiated from the plasmid-specific primer and terminated at positions +76/ +77. Since there were no substitutions in the first two cycles that simultaneously increased the scores of the three promoters (TSPs at positions +76/ +77, +87, and +95/96), it seems unlikely that these bands correspond to new transcription units. Therefore, the intensity of these bands may reflect an increase in the transcriptional output of the entire recombinant region, which roughly corresponds to the eightfold enhanced yield of GFP accumulated in the cells (Figures 3D,F).

As a foreign gene, *appY* is strongly inhibited by H-NS (Atlung et al., 1996; Purtov et al., 2014). Since the binding sites for this protein have not been identified experimentally, we searched them *in silico* and revealed five potential contact regions (red dashes in Figure 3A), although with a rather low reliability ($1.1e-3 \leq p \leq 5.0e-3$). Error-prone mutagenesis created one new site at position –160 in the down-regulated mutant (Figure 3C, $p = 3.8e-3$) that can contribute to the slightly increased affinity of H-NS to this construct: 110.1 ± 5.7 and $117.2 \pm 0.8\%$ in the

EMSA experiments with pure protein and lysates (Figures 3H,I, respectively) these differences are not statistically significant ($p = 0.78$ and 0.15 , respectively, $n = 6$). Multiple mutations in the activated construct created three additional sites (positions –296, –173 and +41, $2.1e-3 \leq p \leq 3.8e-3$) eliminating two binding modules at positions –102 and –45 (Figure 3A). Although their number turned out to be higher than in the control sample, the affinity of the mutant derivative stayed unchanged: 102.4 ± 6.1 and $97.2 \pm 5.3\%$ ($n = 6$) in assays with pure protein and lysates (Figures 3H,I, respectively). Thus, it became clear that the selection of derivatives with increased activity does not lead to the selection of mutants with a reduced affinity for H-NS.

To test the possibility that blindly selected mutations in the up-regulated derivative predominantly accumulated new motifs for activators or eliminated binding sites for inhibitors and gave the opposite effect for an inhibited construct, we grouped 109 transcription factors with known binding motifs into four functional categories (**Supplementary Image 1**). These include regulators known to act only as inhibitors (“inh”), only as activators (“act”), functionally dependent from promoter or ligand (“d”) and regulatory proteins with an uncharacterized mode of action (“?”). Using RSAT with $p \leq 1e-3$ as a threshold, we found 95 binding sites for 50 transcription factors in the initial sequence. Of these, 53 were targets for double regulators, 29 for known inhibitors, and only 7 sites for activators. Nine mutations of the down-regulated derivative left this proportion (54, 27, and 7, respectively) and the total number of transcription factor binding motives (96) almost the same as shown in Figure 3C. Multiple mutations of up-regulated construct reduced its AT-content from 74.9 to 69.1%. As a consequence, the total number of predicted binding sites decreased to 81, but the proportion of sites for regulators with different effects (45, 24, and 7, respectively), as well as their distribution along the recombinant area (insert in Figure 3A) again did not show any specific bias. Thus, we conclude, that the selection of mutants with altered expression of the reporter gene is primarily mediated by the selection of the derivative with a suppressed main promoter in the down-regulated mutant and the derivative without competing promoters in the activated mutant.

DISCUSSION

Exploiting foreign genes for rapid environmental adaptation, bacteria also have to evolve the mechanisms for their acquisition or suppression. We hypothesized that recombination increases local mutagenesis with predominant accumulation of AT base pairs. This can be achieved by deamination of cytidines or oxidation of guanines (Panyukov et al., 2013; Kiselev et al., 2017), if nucleotide sequences in the recombinant area are more sensitive for such modifications. As a result, AT-rich H-NS binding sites capable of suppressing transcription and promoter-like regions suitable for controlling expression of beneficial genes can be evolved. Enhanced mutagenesis in the recombinant area and an assumption that excessive RNA polymerase binding sites can negatively affect transcription, are the two key points in our hypothesis. To test the first assumption, we transferred two

foreign genes into the *E. coli* MG1655 genome and started a long-term evolutionary experiment that already confirmed an increased frequency of spontaneous GC to AT substitutions in the recombinant region (Glazunova et al., 2016). Here, we confirmed the second assumption by obtaining an up-regulated *appY* mutant with dramatically decreased number of potential binding sites for RNA polymerase.

The idea of the bacterial RNA polymerase to be a repressor of transcription was implicitly or directly announced before. Though transcription inhibition, as opposed to the main synthetic role, is not regarded as its general function, there are a number of studies indicating that the enzyme complex formation with DNA can suppress transcription from neighboring promoters (Nieuwkoop and Bender, 1988; Masulis et al., 2015). For instance, in the promoter region of the hutUH operon investigated in the genome of *Klebsiella aerogenes* (Nieuwkoop and Bender, 1988), the binding of RNA polymerase to a mutated P_{UH} promoter with a strengthened -10 region was shown to block the transcription from the overlapping weaker P_C site. Another way of transcription blockade is related to the inactivation of a downstream promoter under active functioning of an upstream promoter. This phenomenon can be exemplified by intragenic promoters required for the synthesis of an RNA product from within the gene of 16S rRNA and activated in the stationary phase upon repression of transcription of the major promoters of the *rrn* operon (Takada et al., 2016). A total of 2,701 sites for binding of RpoD holoenzyme were identified (Shimada et al., 2014) by Genomic SELEX screening. Of these, 1626 promoters turned out to be located inside 777 open reading frames, thus implying a potential functional “conflict” between the promoters located outside of coding regions and adjacent to them intragenic sites of interaction, and there are numerous examples of RNA polymerase complex formation with intragenic promoters (Tutukina et al., 2007; Masulis et al., 2019).

Starting these experiments with promoter regions regulated by H-NS, we expected to observe changes in the number of its potential binding sites, types of complexes formed or efficiency of interaction. However, only the activated construct with the *dps* regulatory region demonstrated some mutation-driven alterations in complex formation (Figure 2G). Thus, we

conclude that the decrease in the number of promoter-like sites is the main driving force for the enhanced transcriptional activity in the *appY* construct. If the population control over horizontally transferred genes primarily operates at the level of randomly created promoters, then a new function of RNA polymerase as a sentinel of foreign genes should be considered.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

AB and OG: isolation of plasmids for sequencing. AB: PCR mutagenesis and GFP fluorescence measurements for cell extracts. OA and KS: primer extension analysis. OA: electrophoretic mobility shift assay. NS and H-NS: isolation/purification. IM: *gfp*-encoding reporter plasmid. KS: initial plasmid construct. OO: bioinformatic analysis and manuscript preparation for publishing with a contribution of all other authors.

FUNDING

Error-prone mutagenesis in the study was supported by the Russian Foundation for Basic Research (Grant 16-04-01570), further experimental and bioinformatics analysis was done under the support of the Russian Science Foundation (Grant 18-14-00348).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2020.00028/full#supplementary-material>

REFERENCES

- Altuvia, S., Almirón, M., Huisman, G., Kolter, R., and Storz, G. (1994). The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol. Microbiol.* 13:265. doi: 10.1111/j.1365-2958.1994.tb00421.x
- Antipov, S. S., Tutukina, M. N., Preobrazhenskaya, E. V., Kondrashov, F. A., Patrushev, M. V., Toshchakov, S. V., et al. (2017). The nucleoid protein Dps binds genomic DNA of *Escherichia coli* in a non-random manner. *PLoS One* 12:e0182800. doi: 10.1371/journal.pone.0182800
- Atlung, T., Sund, S., Olesen, K., and Brøndsted, L. (1996). The histone-like protein H-NS acts as a transcriptional repressor for expression of the anaerobic and growth phase activator AppY of *Escherichia coli*. *J. Bacteriol.* 178:3418.
- Bañuelos-Vazquez, L. A., Torres Tejerizo, G., and Brom, S. (2017). Regulation of conjugative transfer of plasmids and integrative conjugative elements. *Plasmid* 91, 82–89. doi: 10.1016/j.plasmid.2017.04.002
- Cordero, O. X., and Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat. Rev. Microbiol.* 12, 263–272. doi: 10.1038/nrmicro3218
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Daubin, V., and Ochman, H. (2004). Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. *Genome Res.* 14, 1036–1042. doi: 10.1101/gr.2231904
- Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N., and van der Meer, J. R. (2017). The hidden life of integrative and conjugative elements. *FEMS Microbiol. Rev.* 41, 512–537. doi: 10.1093/femsre/fux008
- Dorman, C. J. (2007). H-NS, the genome sentinel. *Nat. Rev. Microbiol.* 5, 157–561. doi: 10.1038/nrmicro1598
- Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñoz-Rascado, L., García-Sotelo, J. S., et al. (2016). RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res.* 44, D133–D143. doi: 10.1093/nar/gkv1156

- Glazunova, O. A., Shavkunov, K. S., Tutukina, M. N., Panyukov, V. V., and Ozoline, O. N. (2016). Integration of foreign genetic material provokes local mutagenesis in the recipient genome. *Mat. Biolog. Bioinform.* 12, t12–t22. doi: 10.17537/2016.11.394
- Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G. (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* 19, 2226–2238. doi: 10.1093/oxfordjournals.molbev.a004046/
- Grüll, M. P., Mulligan, M. E., and Lang, A. S. (2018). Small extracellular particles with big potential for horizontal gene transfer: membrane vesicles and gene transfer agents. *FEMS Microbiol. Lett.* 365:fny192. doi: 10.1093/femsle/fny192
- Huang, Q., Cheng, X., Cheung, M. K., Kiselev, S. S., Ozoline, O. N., and Kwan, H. S. (2012). High density transcriptional initiation signals underline genomic islands in bacteria. *PLoS One* 7:e33759. doi: 10.1371/journal.pone.0033759
- Huerta, A. M., and Collado-Vides, J. (2003). Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. *J. Mol. Biol.* 333, 261–278. doi: 10.1016/j.jmb.2003.07.017
- Ingmer, H., Miller, C. A., and Cohen, S. N. (1998). Destabilized inheritance of pSC101 and other *Escherichia coli* plasmids by DpiA, a novel two-component system regulator. *Mol. Microbiol.* 29, 49–59. doi: 10.1046/j.1365-2958.1998.00895.x
- Keen, E. C., Bliskovsky, V. V., Malagon, F., Baker, J. D., Prince, J. S., Klaus, J. S., et al. (2017). Novel “superspreader” bacteriophages promote horizontal gene transfer by transformation. *mBio* 8:e02115. doi: 10.1128/mBio.02115-16
- Kiselev, S. S., Ozoline, O. N., Glazunova, O. A., Panyukov, V. V., and Shavkunov, K. S. (2017). Novel elements of bacterial genomes – promoter islands: intraspecies polymorphism and sequence stability. *Curr. Bioinform.* 12, 194–201. doi: 10.2174/1574893612666170321144558
- Lang, A. S., Zhaxybayeva, O., and Beatty, J. T. (2012). Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* 10, 472–482. doi: 10.1038/nrmicro2802
- Langille, M. G., and Brinkman, F. S. (2009). IslandViewer: an integrated interface for computational identification and visualization of genomic islands. *Bioinformatics* 25, 664–665. doi: 10.1093/bioinformatics/btp030
- Lawrence, J. G., and Ochman, H. (1998). Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. U.S.A.* 95, 9413–9417.
- Lucchini, S., Rowley, G., Goldberg, M. D., Hurd, D., Harrison, M., and Hinton, J. C. D. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* 3:e81. doi: 10.1371/journal.ppat.0020081
- Lynch, A. S., and Lin, E. C. (1996). Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli*: characterization of DNA binding at target promoters. *J. Bacteriol.* 178, 6238–6249.
- Maciag, A., Peano, C., Pietrelli, A., Egli, T., De Bellis, G., and Landini, P. (2011). In vitro transcription profiling of the σ S subunit of bacterial RNA polymerase: re-definition of the σ S regulon and identification of σ S-specific promoter sequence elements. *Nucleic Acids Res.* 39, 5338–5355. doi: 10.1093/nar/gkr129
- Masulis, I. S., Babaeva, Z. S., Chernyshov, S. V., and Ozoline, O. N. (2015). Visualizing the activity of *Escherichia coli* divergent promoters and probing their dependence on superhelical density using dual-colour fluorescent reporter vector. *Sci. Rep.* 5:11449. doi: 10.1038/srep11449
- Masulis, I. S., Sukharycheva, N. A., Kiselev, S. S., Andreeva, Z. S., and Ozoline, O. N. (2019). Between computational predictions and high-throughput transcriptional profiling: in depth expression analysis of the OppB transmembrane subunit of *E. coli* OppABCDF oligopeptide transporter. *Res. Microbiol.* S0923-2508, 30123–30128. doi: 10.1016/j.resmic.2019.10.005
- Merril, C. R., Switzer, R. C., and Van Keuren, M. L. (1979). Trace polypeptides in cellular extracts and human body fluids detected by two-dimensional electrophoresis and a highly sensitive silver stain. *Proc. Natl. Acad. Sci. U.S.A.* 76, 4335–4339.
- Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., et al. (2005). Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* 21, 4187–4189. doi: 10.1093/bioinformatics/bti635
- Nakamura, Y., Itoh, T., Matsuda, H., and Gojobori, T. (2004). Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.* 36, 760–766. doi: 10.1038/ng1381
- Nieuwkoop, A. J., and Bender, R. A. (1988). RNA polymerase as a repressor of transcription in the *hut(P)* region of mutant *Klebsiella aerogenes* histidine utilization operons. *J. Bacteriol.* 170, 4986–4990.
- Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., et al. (2013). RegPrecise 3.0 - A resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genom.* 14:745. doi: 10.1186/1471-2164-14-745
- Oshima, T., Ishikawa, S., Kurokawa, K., Aiba, H., and Ogasawara, N. (2006). *Escherichia coli* histone like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res.* 13, 141–153. doi: 10.1093/dnares/dsl009
- Overballe-Petersen, S., Harms, K., Orlando, L. A. A., Mayar, J. V. M., Rasmussen, S., Dahl, T. W., et al. (2013). Bacterial natural transformation by highly fragmented and damaged DNA. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19860–19865. doi: 10.1073/pnas.1315278110
- Panyukov, V. V., Kiselev, S. S., Shavkunov, K. S., Masulis, I. S., and Ozoline, O. N. (2013). Mixed promoter islands as genomic regions with specific structural and functional properties. *Math. Biol. Bioinform.* 8:432. doi: 10.17537/2013.8.432
- Panyukov, V. V., and Ozoline, O. N. (2013). Promoters of *Escherichia coli* versus promoter islands: function and structure comparison. *PLoS One* 8:e62601. doi: 10.1371/journal.pone.0062601
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2008). Horizontal gene transfer and the evolution of transcriptional regulation in *Escherichia coli*. *Genome Biol.* 9:R4. doi: 10.1186/gb-2008-9-1-r4
- Purto, Y. A., Glazunova, O. A., Antipov, S. S., Pokusaeva, V. O., Fesenko, E. E., Preobrazhenskaya, E. V., et al. (2014). Promoter islands as a platform for interaction with nucleoid proteins and transcription factors. *J. Bioinform. Comput. Biol.* 12:1441006. doi: 10.1142/S0219720014410066
- Shavkunov, K. S., Masulis, I. S., Tutukina, M. N., Deev, A. A., and Ozoline, O. N. (2009). Gains and unexpected lessons from genome-scale promoter mapping. *Nucleic Acids Res.* 37, 4919–4931. doi: 10.1093/nar/gkp490
- Shimada, T., Yamazaki, Y., Tanaka, K., and Ishihama, A. (2014). The whole set of constitutive promoters recognized by RNA polymerase RpoD holoenzyme of *Escherichia coli*. *PLoS One* 9:e90447. doi: 10.1371/journal.pone.0090447
- Shvyreva, U. S., Tutukina, M. N., and Ozoline, O. N. (2011). Bacterioferritin: properties, structural and functional organization of the *dps* gene regulatory region. *Biophysics* 56, 795–802. doi: 10.1134/S0006350911050204
- Takada, H., Shimada, T., Dey, D., Quyyum, M. Z., Nakano, M., Ishiguro, A., et al. (2016). Differential regulation of rRNA and tRNA transcription from the rRNA-tRNA composite operon in *Escherichia coli*. *PLoS One* 11:e0163057. doi: 10.1371/journal.pone.0163057
- Tran, F., and Boedicker, J. Q. (2017). Genetic cargo and bacterial species set the rate of vesicle-mediated horizontal gene transfer. *Sci. Rep.* 7:8813. doi: 10.1038/s41598-017-07447-7
- Turatsinze, J. V., Thomas-Chollier, M., Defrance, M., and van Helden, J. (2008). Using RSAT to scan genome sequences for transcription factor binding sites and cis-regulatory modules. *Nat. Protoc.* 3, 1578–1588. doi: 10.1038/nprot.2008.97
- Tutukina, M. N., Shavkunov, K. S., Masulis, I. S., and Ozoline, O. N. (2007). Intragenic promoter-like sites in the genome of *Escherichia coli* discovery and functional implication. *J. Bioinform. Comput. Biol.* 5, 549–560.
- Tutukina, M. N., Shvyreva, U. S., and Ozoline, O. N. (2015). Architectural factor of bacterial nucleoid, H-NS: overproduction and verification of its functional activity. *Sorpt. Chromatogr. Process.* 15:435.
- Wiedenbeck, J., and Cohan, F. M. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol. Rev.* 35, 957–976. doi: 10.1111/j.1574-6976.2011.00292.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Bykov, Glazunova, Alikina, Sukharicheva, Masulis, Shavkunov and Ozoline. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.