6456–6464 *Nucleic Acids Research, 2011, Vol. 39, No. 15* *Published online 13 May 2011*
doi:10.1093/nar/gkr307

The PurR regulon in *Escherichia coli* K-12 MG1655

Byung-Kwan Cho, Stephen A. Federowicz, Mallory Embree, Young-Seoub Park, Donghyuk Kim and Bernhard Ø. Palsson*

Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA

Received February 18, 2011; Revised and Accepted April 17, 2011

ABSTRACT

The PurR transcription factor plays a critical role in transcriptional regulation of purine metabolism in enterobacteria. Here, we elucidate the role of PurR under exogenous adenine stimulation at the genome-scale using high-resolution chromatin immunoprecipitation (ChIP)–chip and gene expression data obtained under *in vivo* conditions. Analysis of microarray data revealed that adenine stimulation led to changes in transcript level of about 10% of *Escherichia coli* genes, including the purine biosynthesis pathway. The *E. coli* strain lacking the *purR* gene showed that a total of 56 genes are affected by the deletion. From the ChIP–chip analysis, we determined that over 73% of genes directly regulated by PurR were enriched in the biosynthesis, utilization and transport of purine and pyrimidine nucleotides, and 20% of them were functionally unknown. Compared to the functional diversity of the regulon of the other general transcription factors in *E. coli*, the functions and size of the PurR regulon are limited.

INTRODUCTION

Metabolism enables a cell to assimilate exogenous nutrients both for energy generation and for macromolecular synthesis. A set of metabolic pathways directs the biosynthesis and utilization of nucleotides that is critical for virtually every aspect of cellular life. Purine and pyrimidine nucleotides constitute a part of the nucleic acids, cofactors in enzymatic reactions, intracellular and extracellular signals, phosphate donors and the major carriers of cellular energy (1,2). Since the biosynthesis and utilization of the nucleotides is demanding of cellular resources and plays a broad role in cellular processes, imbalances between the different nucleotide pools significantly perturb the normal cellular functions (2,3).

In general, metabolism is tightly controlled by feedback inhibition of enzyme activity by metabolites and by transcriptional regulation by DNA-binding proteins. The regulatory action of DNA-binding proteins is also modulated by the exogenous nutrients as stimuli. Therefore, there is great interest in not only elucidating the set of genes under regulation by the same stimuli (defined as a stimulon), but also identifying the collection of genes under regulation by the same regulatory protein (defined as a regulon). In the case of purine nucleotide metabolism in *Escherichia coli*, purine repressor (PurR) tightly regulates transcription of the enzymes involved in inosine 5′-monophosphate (IMP) biosynthesis and the conversion of IMP to adenosine monophosphate (AMP) and guanosine monophosphate (GMP)(2). The regulatory action of PurR on target genes is modulated by the binding of the small effector molecules [hypoxanthine (Hx) or guanine] and in effect endows PurR with the ability to affect transcriptional regulation (4). In other words, upon availability of purine nucleotides from the environment, the activity of PurR can be enhanced to repress the expression of target genes. However, little is known about *in vivo* PurR-binding events and their causal relationships with gene expression at the genome scale in the presence or absence of purine nucleotides. Such information is needed to reconstruct the PurR regulon and to understand purine metabolism.

The *E. coli* transcriptional regulatory network is believed to have a hierarchical topology with several global transcription factors (TFs) at the top-level (5–7). The global TFs were specified by the multiple functional categories of the genes regulated. By contrast, specific TFs restrict their target genes to the same metabolic pathways or the same functional categories (6). Previously, PurR was classified into the group of general TFs; however, due to the lack of information on target genes in its regulon, the understanding of the role of PurR is limited. In particular, it is unclear whether its effects on *E. coli* metabolism are direct or indirect. If the effects are indirect, it is also unclear whether the indirect effects are made through other TFs or other metabolites.

*To whom correspondence should be addressed. Tel: +1 858 534 5668; Fax: +1 858 822 3120; Email: palsson@ucsd.edu
Present addresses:
Byung-Kwan Cho, Department of Biological Sciences, KAIST, Daejoen 305-701, Korea.
Young-Seoub Park, Mogam Biotechnology Research Institute, Yongin 449-799, Korea.

In this study, we obtain and integrate genome-scale data from chromatin immunoprecipitation (ChIP)–chip and gene expression profiling to elucidate the regulatory role of PurR at a genome scale. First, using changes in transcript levels on a genome scale, we defined the adenine stimulon from comprehensively established sets of genes differentially expressed in response to exogenous adenine. Second, we used the *purR* deletion mutant to determine the PurR-dependent genes affected by the deletion mutant. Third, we set out to comprehensively establish the PurR-binding regions on the *E. coli* genome experimentally to further elucidate any DNA sequence motif correlated with the PurR regulatory action. Fourth, we determined the regulatory action of PurR based on the causal relationships between the association of PurR and changes in transcript levels. In the end, the reconstruction of the regulatory network of PurR allows us to understand the role of the PurR regulon as a part of the broader adenine stimulon. The results show that the role of PurR regulon is locally acting but its effect on the entire metabolism is critical in response to the exogenous purine stimulation.

MATERIALS AND METHODS

Bacterial strains and growth conditions

All strains used are *E. coli* K-12 MG1655 and its derivatives. The *E. coli* strain harboring PurR-8myc was generated as described previously (8). Deletion mutant (Δ*purR*) was constructed by a λ Red and FLP-mediated site-specific recombination system (9). Glycerol stocks of *E. coli* strains were inoculated into M9 minimal medium supplemented with 2 g/l glucose and cultured overnight at 37°C with constant agitation. The cultures were inoculated into 100 ml of the fresh M9 minimal medium in either the presence or absence of 100 μg/ml adenine and continued to culture at 37°C with constant agitation to mid-log phase.

Transcriptome analysis

Samples for transcriptome analyses were taken from exponentially growing cells. From the cells treated by 2 vol of RNAprotect Bacteria Reagent (Qiagen), total RNA was isolated using RNeasy kit (Qiagen) with DNaseI treatment in accordance with manufacturer's instruction. AffymetrixGeneChip*E. coli* Genome 2.0 arrays were used for genome-scale transcriptional analyses. Complementary DNA (cDNA) synthesis, fragmentation, end-terminus biotin labeling and array hybridization were performed as recommended by Affymetrix standard protocol. Raw CEL files were analyzed using robust multi-array average for normalization and calculation of probe intensities.

ChIP and microarray analysis

To identify PurR-binding regions *in vivo*, we isolated the DNA bound to PurR protein by ChIP. Cultures at mid-log phase were cross-linked by 1% formaldehyde at room temperature for 25 min. After cell lysis and sonication, the cross-linked DNA-PurR complex was immunoprecipitated by using the specific antibody against myc-tag (9E10, Santa Cruz Biotech) and Dynabeads Pan Mouse IgG magnetic beads (Invitrogen) followed by stringent washings as described previously (10). After reversal of the cross-links by incubation at 65°C overnight, the samples were treated by RNaseA (Qiagen) and proteaseK (Invitrogen) and then purified with a PCR purification kit (Qiagen). Then, the amplified ChIP DNA samples were labeled and hybridized onto whole-genome tiled microarrays (Roche-NimbleGen).

Data analysis

To identify PurR-binding regions, we used the peak finding algorithm built into the NimbleScan^{TM} software. Processing of ChIP–chip data was performed in three steps: normalization, IP/mock-IP ratio computation (log base 2), and enriched region identification. The $\log_2$ ratios of each spot in the microarray were calculated from the raw signals obtained from both Cy5 and Cy3 channels, and then the values were scaled by Tukey bi-weight mean. The $\log_2$ ratio of Cy5 (IP DNA) to Cy3 (mock-IP DNA) for each point was calculated from the scanned signals. Then, the bi-weight mean of this $\log_2$ ratio was subtracted from each point. Each log ratio dataset from triplicate samples was used to identify PurR-binding region using the software (width of sliding window = 300 bp). Our approach to identify the PurR-binding regions was to first determine binding locations from each data set and then combine the binding locations from at least five of the six data sets to define a binding region (11).

Motif searching

The PurR-binding motif analysis was completed using the MEME and FIMO tools from the MEME software suite (12). We first determined the proper binding motif and then scanned the full genome for its presence. The elicitation of the motif was done using the MEME program on the set of sequences defined by the PurR-binding regions. Using default settings, the previously determined PurR motif was recovered and then tailored to the correct size by setting the width parameter to 16 bp. We then used these motifs and the PSPM (position specific probability matrix) generated by MEME to rescan the entire genome with the FIMO program.

RESULTS

Determination of gene expression changes to the exogenous adenine

To characterize the changes in gene expression with exposure to exogenous adenine, global transcriptome analyses were performed using DNA microarrays. *Escherichia coli* strain K-12 MG1655 (wild type) was grown in M9 medium supplemented by 100 μg/ml adenine. Samples were removed from the culture with and without adenine stimulation and used for the extraction of total RNA. Analysis of the DNA microarray data

demonstrated that exogenous adenine stimulation led to changes in transcript level of about 10% of *E. coli* genes (Supplementary Table S1). Those include the upregulation of 144 genes and downregulation of 255 genes with more than 2-fold expression change and *P*-value <0.05. As previously well described (2), all of the genes associated with the purine biosynthesis pathway were repressed by the addition of adenine. Among them, *purD* encoding phosphoribosylamine–glycine ligase showed the highest repression factor (43.89-fold) (Supplementary Table S2). Interestingly, adenine addition led to the high downregulation of several transporter genes, including *codB*, *xanP*, *yeeF* and *uraA* encoding cytosine transporter (34.52-fold), xanthine NCS2 transporter (20.12-fold), amino acid APC transporter (18.61-fold) and uracil NCS2 transporter (17.64-fold).

A significant portion of the genes downregulated by adenine addition is associated with the pyrimidine and amino acid biosynthetic pathways. In particular, the downregulation of genes comprising the arginine biosynthetic pathway clearly demonstrates that the purine metabolism links to the *in vivo* level of arginine in order to regulate the biosynthesis of deoxyribonucleic and ribonucleic acid (13). On the other hand, the highly induced genes by adenine stimulation were involved in other various cellular processes (Supplementary Table S2). Among them, *ydhC* encoding drug MFS transporter had the highest activation factor (193.55-fold). Of interest, two non-coding RNAs, *gcvB* and *rybB,* were induced by a factor of 34.37 and 5.92, respectively. Previous studies showed that GcvB enhances the ability of *E. coli* to survive low pH by upregulating the levels of alternate sigma factor, RpoS (14). Another alternate sigma factor, RpoE-dependent RybB, regulates the synthesis of major porins in *E. coli* (15). Among genes in purine salvage pathways, *add* encoding adenosine deaminase was induced by adenine as well (17.61-fold) (2). The genes in thiamine and biotin biosynthesis pathways were also induced by the exogenous adenine. These observations demonstrate that a large number of the downregulated genes related with purine and pyrimidine biosynthesis and transport, arginine biosynthesis and ATP synthesis coupled with proton transport form a major portion (~64%) of the adenine stimulon.

PurR-dependent transcriptome response to the exogenous adenine

Next, we studied the global response of a *purR* deletion mutant to adenine stimulation to identify PurR-dependent genes (Supplementary Table S1). To address this issue, we isolated total RNA from the isogenic *purR* deletion mutant during exponential growth phase and hybridized the cDNA obtained from the total RNA onto Affymetrix microarrays. A comparison of the gene expression levels between cells grown in the presence and absence of the PurR protein in response to the exogenous adenine revealed that a total of 56 genes exhibit differential expression with more than 2-fold change and a false discovery rate (FDR) value <0.05 (*P*-value = 0.0056) from analysis of variance (ANOVA) analysis (Supplementary Table S3).

Nineteen genes (~34%) showed increased transcript levels in response to the exogenous adenine due to regulation by PurR (Supplementary Table S3). None of these 19 genes has been previously reported to be directly regulated by PurR. On the other hand, transcription of the 37 genes (~66%) was repressed by PurR. Eleven genes of the IMP (inosine 5′-monophosphate) biosynthetic pathway from PRPP (5-phosphoribosyl-1-pyrophosphate) clustered into this group. It has been previously determined that 10 of them were directly repressed by the PurR protein (16–24). Eight genes in pyrimidine biosynthesis and transport pathways were directly or indirectly regulated by the PurR protein, that include *carA*, *carB*, *pyrB*, *pyrI*, *pyrC*, *pyrD*, *codA* and *codB*. It has been experimentally determined that four of them (*carA*, *carB*, *pyrD* and *pyrC*) were directly repressed by the PurR (25–29). Transcription of *yieG* (–2.91-fold), *xanP* (–30.82-fold), and *uraA* (–2.04-fold) encoding adenine, xanthine and uracil transporter, respectively, were also affected by the *purR* deletion, indicating direct or indirect regulatory effect of the PurR protein on the adenine and uracil transport systems (30). Transcriptional repression of genes in arginine biosynthesis pathway (*argA*, *argB* and *argC*) is potentially mediated by the PurR protein. Interestingly, acid stress response genes (*hdeB*, *hdeA* and *hdeD*) decreased expression in a *purR* deletion mutant. Consistent with this observation, the level of messenger RNA (mRNA) transcript of *gadY,* a regulatory small RNA that is highly upregulated by low pH (31), was affected by the *purR* deletion. A hallmark of the *E. coli* response to the exogenous adenine is the rapid and strong repression of a set of genes in purine biosynthesis pathway. The observed repression of all of these genes in the wild-type strain, but not the *purR* deletion mutant, provided an internal validation of the microarray experiment.

Genome-wide identification of PurR regulon

PurR-binding regions have been characterized by *in vitro* DNA-binding experiments and mutational analysis; however, direct analysis of *in vivo* PurR binding is not available. We thus employed the ChIP coupled with microarrays (ChIP–chip) approach to determine the *in vivo* PurR-binding regions in *E. coli* cells under either the presence or the absence of exogenous adenine (Figure 1).

We performed a hybridization of the immunoprecipitated DNA (Cy5 channel) and mock immunoprecipitated DNA (Cy3 channel) onto the high-resolution whole-genome tiling microarrays, which contained a total of 371 034 oligonucleotides with 50-bp tiles overlapping every 25 bp on both forward and reverse strands (11,32). The normalized $\log_2$ ratios obtained from the hybridization identify the genomic regions enriched in the IP-DNA sample compared with the mock IP-DNA sample and thereby represent a genome-wide map of *in vivo* interactions between PurR protein and *E. coli* genome (Figure 1A). Using a peak finding algorithm, 35 and 13 unique and reproducible PurR-binding regions were identified from the

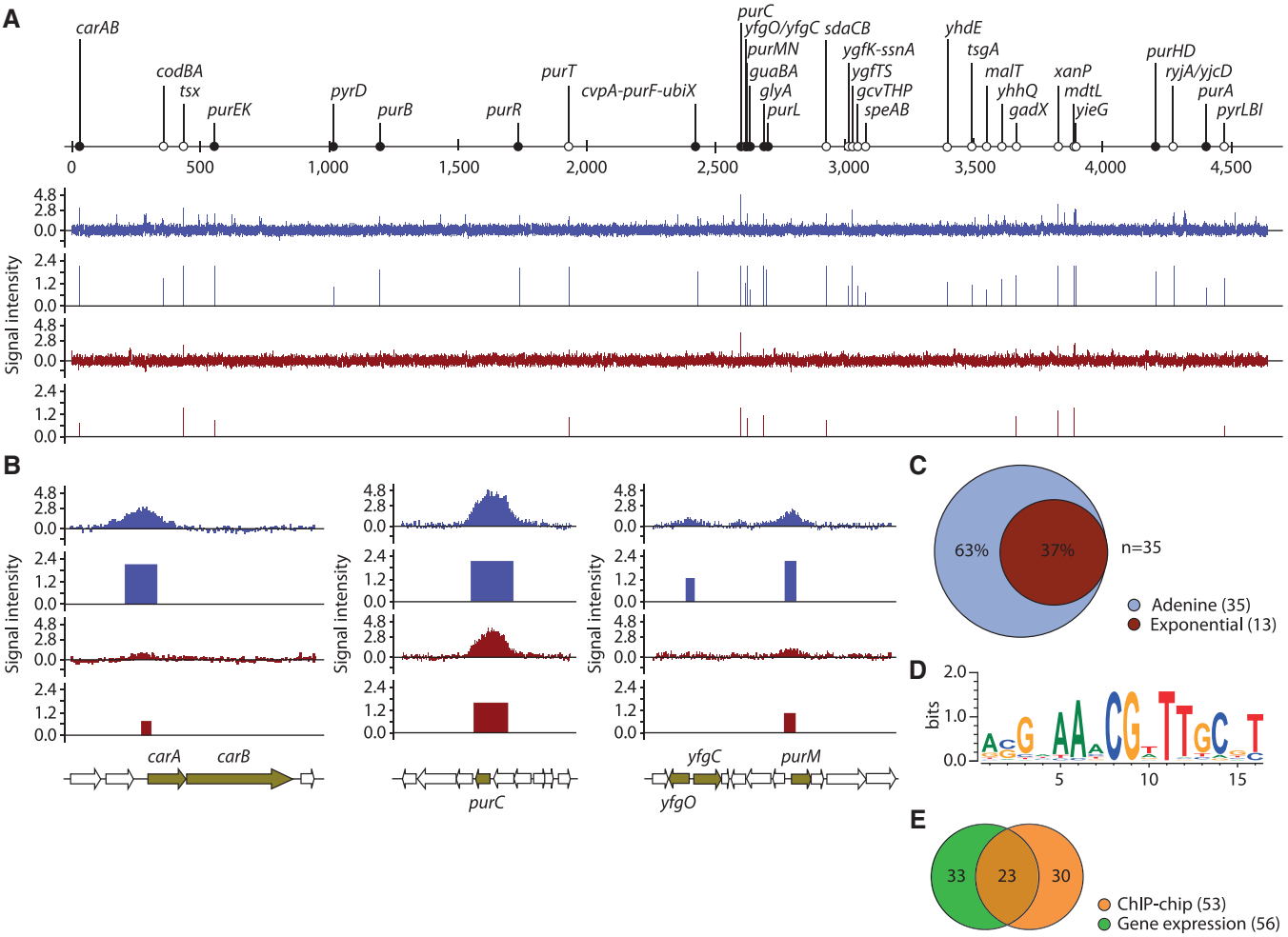

Figure 1. Genome-wide distribution of PurR-binding regions. (**A**) An overview of PurR-binding profiles across the *E. coli* genome at exponential growth phase in the presence (blue) or absence (red) of exogenous adenine. Black and white dots indicate previously known and newly found PurR-binding regions, respectively. (**B**) Determination of genuine PurR-binding regions on the selected regions. Promoter regions of *carAB*, *purC*, *yfgO*/*yfgC* and *purM* are occupied by PurR at exponential state. The peak height of the identified PurR-binding region is the log 2 enrichment ratio calculated from Cy5 (IP DNA) and Cy3 (mock IP DNA) signal intensity of the probe corresponding to the identified region. (**C**) Overlaps between PurR-binding regions of exponential phase in the presence (blue) and absence (red) of adenine. (**D**) Sequence logo representation of the PurR–DNA binding profile. (**E**) Comparison of ChIP–chip results and gene expression profiles.

hybridizations in exponential phase in the presence and absence of adenine, respectively (Table 1).

The genome-wide PurR-binding maps obtained from two different conditions, i.e. exponential growth phase in the presence and the absence of exogenous adenine, indicated that the PurR association on the *E. coli* genome is dramatically sensitive to the addition of adenine. For instance, PurR occupancy for the promoter regions of *carAB*, *purC* and *purMN* transcription units showed a great differential ratio between those two conditions (Figure 1B, Table 1). At the previously characterized PurR-binding promoter regions of *pyrD*, *purB*, *purR*, *cvpA-purF-ubiX*, *guaBA*, *purL* and *purHD*, we only observed the PurR-binding in the presence of exogenous adenine. Only 37% of binding sites (13 of 35) overlapped under the conditions in the absence and presence of exogenous adenine, and 62% of binding sites (22 of 35) were found in the presence of exogenous adenine (Figure 1C). Adenine can be converted to IMP through the intermediate formation of adenosine, inosine, and Hx catalyzed by purine nucleoside phosphorylase (*deoD*) and adenosine deaminase (*add*)(2). Thus, this observation indicates that the addition of exogenous adenine increased in the intracellular level of Hx, which led to the formation of the PurR–Hx complex that functions in transcriptional regulation (4). A total of 22 new PurR-binding regions were identified in this study, whose roles were involved in various cellular processes (Table 1). Prior to this study, 15 PurR-binding regions had been characterized by DNA-binding experiments *in vitro* and mutational analysis *in vivo*, 87% (13 of 15) of which were identified in this study (transcription units in bold characters in Table 1). The exceptions were *pyrC* and *glnB* promoters, whose cellular functions are related to the pyrimidine biosynthesis and nitrogen metabolism, respectively. It is unclear why those PurR-binding regions were missed from our analysis.

Table 1. Genome-scale identification of PurR-binding regions

Start	End	WA	WOA	Transcription unit	Motif position	Consensus sequence
28 951	29 929	2.08	0.70	***carA–carB***	29 415–29 430	TGGAAAACGCTTGCGC
169 607	170 394	1.00	0.73	*fhuC-fhuD-fhuB*	170 015–170 030	AGCAAAGCGTTTGCCC
354 056	354 323	1.41		*codB-codA*	354 064–354 079	ACGAAAACGATTGCTT
430 925	431 911	2.08	1.48	*Tsx*	431 376–431 391	ACGTAATCGATTGCGT
552 127	552 626	2.08	0.86	***purE–purK***	552 395–552 410	AGGAAAACGGTTGCGT
1 003 618	1 004 017	1.01		***pyrD***	1 003 889–1 003 904	CGGAAAACGTTTGCGT
1 190 891	1 191 196	1.90		***purB***	1 191 010–1 191 025	AGGTAACCGATTGCGT
1 735 528	1 736 095	1.99		***purR***	1 735 808–1 735 823	AGGTAAACGTTTGCCT
1 766 545	1 767 225	1.08		*ydiJ/ydiK*		
1 928 502	1 929 233	2.00	1.02	*purT*	1 928 852–1 928 867	ACGAAAACGTTTGCGT
2 226 706	2 227 311	1.13		*yohP*		
2 428 603	2 428 965	1.77		***cvpA–purF–ubiX***	2 428 842–2 428 857	AAGAAAACGTTTGCGT
2 439 185	2 439 884	0.97		*fabB-mnmC*	2 439 210–2 439 225	ACCAAAGCGATGGCAT
2 594 676	2 596 862	2.08	1.52	***purC***	2 595 794–2 595 809	ACGCACACGTTTGCGT
2 613 728	2 614 177	1.18		*yfgO/yfgC*	2 613 843–2 613 858	CCGGAAGCGATTGCCT
2 618 847	2 619 484	2.08	0.98	***purM–purN***	2 619 140–2 619 155	TCGCAAACGTTTGCTT
2 632 091	2 632 533	0.84		***guaB–guaA***	2 632 146–2 632 161	GCGTAACCGATTGCAT
2 683 277	2 683 995	2.08	1.13	***glyA***	2 683 648–2 683 663	AGGTAATCGTTTGCGT
2 693 490	2 693 714	1.88		***purL***	2 693 642–2 693 657	ACGAAACCGTTTGCGT
2 925 514	2 926 070	2.08	0.83	*sdaC-sdaB*	2 925 844–2 925 859	AGGGAAACGTTTGCGT
3 013 592	3 014 047	1.00		*ygfK-ssnA*	3 013 768–3 013 783	AGTTAAACGATTGCGT
3 028 215	3 029 251	2.08		*ygfT-ygfS*	3 028 661–3 028 676	GGGCAAGCGATTGCGC
3 048 883	3 049 539	1.06		*gcvT-gcvH-gcvP*	3 048 572–3 049 003	AAGAGAACGATTGCGT
3 396 153	3 396 727	1.26		*yhdE*		
3 490 230	3 490 492	1.08		*ppiA/tsgA*		
3 550 769	3 551 100	0.86		*malT*	3 551 001–3 551 016	ACGTCATCGCTTGCAT
3 607 049	3 607 298	1.39		*yhhQ*	3 607 081–3 607 096	AGGTTAACGATTGCGT
3 663 018	3 663 884	1.56	1.06	*gadX*	3 663 653–3 663 668	AGGGACACGCTTTCAT
3 826 187	3 827 555	2.08	1.36	*xanP*	3 826 900–3 826 915	TGGCAAACGTTTGCTT
3 889 003	3 889 940	2.08	1.52	*mdtL*		
3 894 397	3 895 190	2.08		*yieG*	3 894 732–3 894 747	ACGGCAACGATTGCGT
4 205 612	4 470 784	1.75		***purH–purD***	4 205 663–4 205 678	ACGAAAACGTTTGCGC
4 275 855	4 276 485	2.08		*ryjA/yjcD*	4 276 230–4 276 245	ACGATAACGTTTGCGC
4 402 483	4 402 483	1.19		***purA***		
4 470 473	4 470 609	1.43	0.58	*pyrL-pyrB-pyrI*		

We next assessed the locations of the PurR-binding regions against the current annotated genome information (11). The PurR-binding regions were observed only within intergenic (i.e., promoter and promoter-like) regions. Therefore, there exists a strong preference for the PurR-binding target to be located within the noncoding intergenic regions, similar to that observed for Lrp-binding sites (32). To identify common DNA sequence motifs of the PurR-binding regions, we used the MEME suite tool (12). The sequences of PurR-binding regions were used to generate the position specific probability matrix and to rescan the entire genome with the FIMO program. We then analyzed only those sites which were located in the PurR-binding regions and fell below a stringent cut-off (P-value <0.0001). This revealed a total of 28 conserved sequences spread across 35 binding regions (Table 1). The identified sequence motif (ACGNAAACGTTTGCNT) was consistent with the previously characterized 16-bp palindromic binding site of the PurR (Figure 1D) (2). Based on the fact that the increase in the intracellular adenine levels enhances PurR binding to its DNA targets and the coverage of the known binding regions in our data, we concluded that PurR-binding regions identified here are *bona fide* binding sites.

Genome-scale determination of causal relationship

Currently, a total of 22 genes have been characterized as members of PurR regulon to be directly repressed by PurR (33). From our ChIP–chip analyses, we significantly expanded the size of the PurR regulon to comprise 53 target genes (Table 1). To determine the causal relationships between the binding of PurR and the changes in RNA transcript levels of genes in the PurR regulon, we integrated the information on the binding regions of PurR with transcriptomic analysis. Among 53 target genes in PurR regulon determined by ChIP–chip analyses, we determined 23 genes (43%) differentially expressed in response to the *purR* deletion and the addition of exogenous adenine with more than a 2-fold change and an FDR value <0.05 (P-value = 0.0056) from ANOVA analysis (Figure 1E and Supplementary Table S4). The genes directly repressed by PurR in response to the exogenous adenine (23 genes) include *codB*, *codA*, *purT*, *xanP*, *yieG*,*pyrL*, *pyrB* and *pyrI* encoding cytosine NCS1 transporter, cytosine deaminase, phosphoribosylglycinamide-formyltransferase, xanthine NCS2 transporter, adenine transporter, PyrL leader peptide, aspartate carbamoyltransferase and aspartate carbamoyltransferase regulatory subunit, respectively, as newly found members.

The remaining 57% of the genes had a direct association with PurR, lacking significant changes in RNA transcript levels. The cellular functions of most of the remaining genes were not clustered in purine and pyrimidine metabolic pathways, indicating that the changes in their transcript levels require additional regulatory signals such as transcription factors. Surprisingly, none of the genes were directly activated by PurR. On the contrary, PurR completely represses target genes involved in the IMP biosynthetic pathway. This suggests that most of the repression is a direct interaction, but the transcriptional activation is indirect (Supplementary Tables S1, S3 and S4).

In general, the PurR-binding sites are located in the promoter region between position −35 and −10 promoter elements, indicating that the binding of PurR regulates transcription initiation (2). In the case of *purB* and *purR*, PurR binds to the open reading frame so that it blocks transcription elongation (34). Therefore, the binding position of PurR is of great interest in order to understand its regulatory mechanism. We calculated the distance between PurR-binding motifs and the transcription start sites (TSSs) based upon the TSSs recently published (11). Of 32 promoter regions directly regulated by PurR, 14 regions (44%) include the PurR-binding motif between −10 and −35 promoter elements.

Metabolic pathways directly regulated by PurR–Hx complex

Purine nucleotide metabolism plays a critical role in various cellular activities. To identify the metabolic pathways regulated by PurR–Hx complex, the members of PurR regulon were functionally classified and further mapped to the *E. coli* metabolic pathways (Figure 2). The genes with direct PurR association lacking changes in transcript levels were classified into other cellular functions. However, the genes directly repressed by PurR in response to the exogenous adenine clustered mainly into purine and pyrimidine metabolic pathways. First, PurR directly autoregulates itself and generates PurR–Hx complex in the presence of Hx or adenine (Figure 2A). Second, PurR completely regulates purine transport, biosynthesis, salvage and interconversion pathways (Figure 2B). Interestingly, PurR directly regulates serine transport and metabolic pathways to produce N^{10}-formyltetrahydrofolate (N^{10}-FTHF), which is an intermediate for the IMP biosynthetic pathway. In addition, we found that PurR directly represses xanthine (*xanP*), purine nucleoside (*tsx*) and adenine (*yieG*) transporters. Although the transporter for Hx is currently unknown, xanthine transporter is unable to transport Hx (35). Lastly, PurR downregulates the genes in pyrimidine biosynthetic and transport pathways (Figure 2C). Most of the genes having direct PurR association with differential gene expression were enriched into purine and pyrimidine metabolic pathways. Interestingly, none of genes involved with purine utilization, such as *apt*, *deoD* and *add*, are directly regulated by PurR.

DISCUSSION

We determined the PurR regulon in *E. coli* in response to the exogenous adenine stimuli by integrating genome-scale location analysis and gene expression profiles. The genome-wide map of PurR-binding sites presented here not only confirms previously characterized binding sites (15 regions) but also expands the number of known binding sites (35 regions) to a genome-wide assessment; similar to what we previously reported for Lrp- and Fis-binding sites (10,32). From the genome-wide

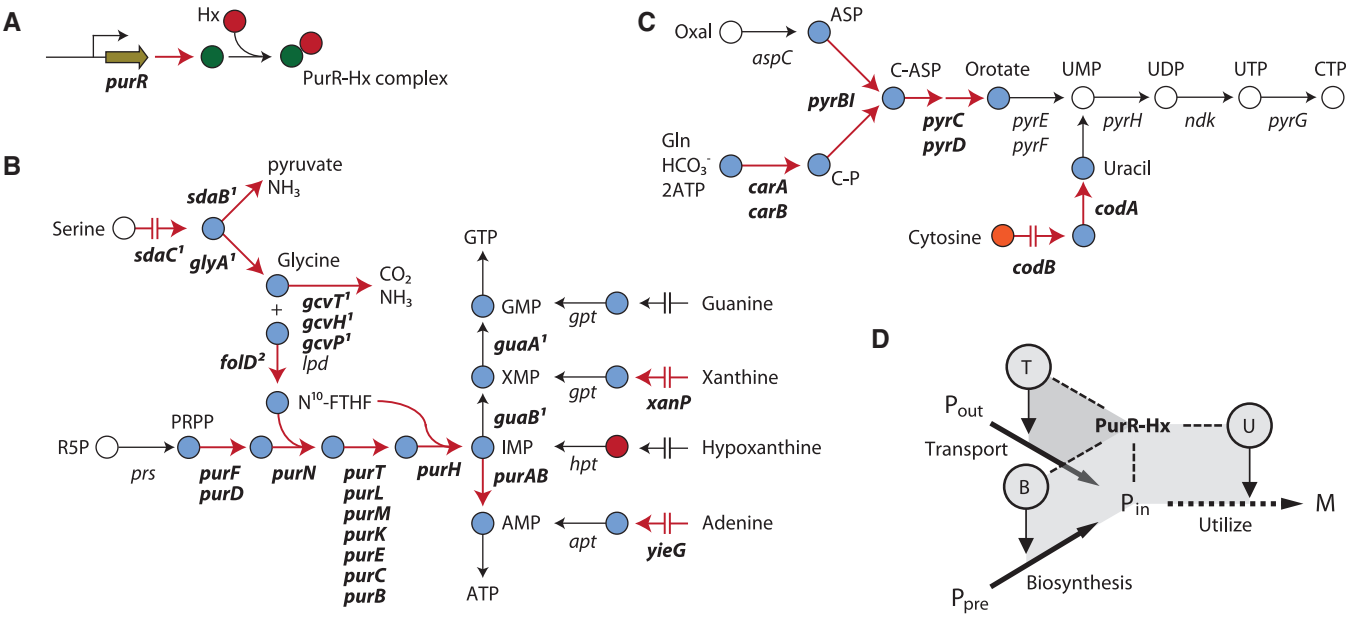

Figure 2. Metabolic pathways directly regulated by PurR and regulatory motif. (**A**) Formation of PurR–hypoxanthine complex. (**B, C**) The purine and pyrimidine biosynthesis pathways. The genes directly regulated by PurR are depicted by bold characters. Red arrows show the PurR-mediated repression. Broken arrows indicate the transporters. (**D**) Schematic diagram for the regulatory motif reconstruction in feedback loop.

mapping results, we were also able to show that: (i) a total of 35 PurR-binding regions were identified, all of which were located within noncoding regions, showing the strong binding preference of PurR-binding to the promoter and promoter-like regions; (ii) only 37% of binding sites (13 of 35) overlapped under the conditions in the absence and presence of exogenous adenine, indicating that PurR bindings to the *E. coli* genome are dramatically sensitive to the addition of exogenous adenine (or hypoxanthine); (iii) the integration of these results with mRNA transcript level information indicates that the functional assignment of the regulated genes is strongly enriched in the purine and pyrimidine metabolism-related functions. In addition, most of the genes were thoroughly repressed by PurR. Interestingly, the other genes directly bound by PurR lacking differential expression in response to the *purR* deletion or the exogenous adenine were functionally diverse; and (iv) the PurR-binding motifs were observed at the regions of −10 and −35 promoter elements, indicating PurR regulates transcription initiation.

We discovered PurR-binding regions from the promoter regions of *codBA*, *purT*, *xanP*, *yieG* and *pyrLBI* with the differential gene expression. First, PurR directly regulates the *de novo* biosynthesis of pyrimidine. Among the genes in the biosynthetic pathway, *codB* and *codA* encode a cytosine transporter belonging to the NCS1 family of purine and pyrimidine transporters and a cytosine deaminase metabolizing cytosine to uracil and ammonia, respectively. In addition, PurR directly regulates *pyrB* and *pyrI*, encoding catalytic and regulatory subunits of aspartate transcarbamylase (ATCase), respectively, catalyzing the first reaction of the *de novo* biosynthesis of pyrimidine nucleotides. Considering that PurR represses *carA*, *carB*, *pyrC* and *pyrD*, the very early steps of *de novo* pyrimidine biosynthesis and transport are tightly regulated by PurR in response to exogenous adenine. Interestingly, RutR, the uracil responsive transcription factor, binds to the promoter region of *carAB* (36). Although Shimada and co-workers demonstrated that the RutR binding site plays little or no role in the regulation of transcription, it may have an additional regulatory role along with other proteins. In the *carAB* promoter, at least five regulatory proteins (IHF, PepA, PurR, RutR and ArgR) are involved in the purine, pyrimidine and arginine-specific control of the promoter activity (28). The complexity of the multicomponent regulatory mechanisms modulating *carAB* transcription shows the need for a cellular balance between the synthesis of pyrimidine and purine residues. Second, PurR directly regulates the transport of purine nucleotides. We observed PurR-binding peaks at the upstream regions of *xanP*, *mdtL* and *yieG* with differential gene expression in response to the exogenous adenine. Interestingly, PurR-binding peaks were observed for *xanP* and *mdtL* in the absence of exogenous adenine but not for *yieG*, suggesting that the *yieG* encodes a high-affinity transport system for adenine, which is dispensable in the presence of excess substrate. It has been suggested that another adenine transport system close to the genomic position of *yieG* operates at low affinity and is not energy dependent (30). However, it has not been discovered which gene has the adenine transport function with low affinity. Here, we found that *mdtL* encoding drug MFS transporter is directly repressed by PurR and located close to *yieG* (~4.5 kb), indicating that *mdtL* might be responsible for the low-affinity adenine transport. Thus, the purine transport system of the PurR regulon can be composed of two high-affinity transporters (*xanP* and *yieG*) and one low-affinity transporter (*mdtL*).

Transcriptional regulatory systems often regulate the formation rates and the concentration of small molecules by feedback loops that regulate the transport, biosynthesis and metabolic enzymes (11,37). Since adenine can be utilized by *apt*, encoding adenine phosphoribosyltransferase, we were able to connect transport, biosynthesis and metabolic feedback loop pairs (Figure 2D). In the left loop, PurR–Hx complex represses the transcription of the transport proteins (T) for purine (*xanP* and *yieG*) and pyrimidine (*codB*), and biosynthetic proteins (B) for IMP (*purEK*, *purB*, *purT*, *purF*, *purC*, *purMN*, *purI* and *purHD*) and UMP (*carAB*, *codA*, *pyrBI* and *pyrCD*), reducing the influx of the purine or pyrimidine molecules (P_{in}) from the media (P_{out}) and precursors (P_{pre}). In the right loop, metabolic enzyme (U) responsible for converting P_{in} into metabolites (M) is not directly regulated by PurR–Hx complex; however, its transcript level is reduced by the exogenous adenine. Thus, the logical structure of the connected feedback loop (CFL) motif described by a notation that uses three signs indicating repression (R) or activation (A) for each of T, B, and U can be **R-R-R**. In the previous studies (32,37), the B component (i.e. biosynthesis) was not included in the logical structures. The **R-R-R** motif demonstrates that the influx and efflux are repressed for flow homeostasis(37), which means that the exogenous adenine cannot be utilized as nutrient molecules. In the case of nutrient molecules and homeostasis, the logical structures of CFL would have been **A-A/R-A** and **R-A/R-A**, respectively (37). Since the **R-R-R** motif is uncommon for the regulation of small molecules in living cells, transcriptional regulatory networks for maintaining the levels of the purine and pyrimidine molecules may be more complex than previously thought.

Previously, the PurR was classified into the group of general TFs based on the functional diversity of the genes in its regulon (6). Compared to the other general TFs in *E. coli* such as Fnr (38), Crp (39) and Lrp (32), the functions and size of the PurR regulon are limited. However, cellular functions of the genes are highly enriched into the purine and pyrimidine transport and biosynthesis, indicating that the direct effect of PurR on the *E. coli* metabolism is local, but via the balance of cellular purine content, it plays a critical role in metabolism. Now we may need to select a new list of global transcription factors in *E. coli*.

ACCESSION NUMBER

All raw data files have been deposited to Gene Expression Omnibus through accession numbers GSE26588 and GSE26589.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Marc Abrams for editing the manuscript.

FUNDING

National Institutes of Health Grant GM062791; The Office of Science-Biological and Environmental Research, U.S. Department of Energy, DE-FOA-0000143. Funding for open access charge: National Institutes of Health Grant GM062791.

Conflict of interest statement. None declared.

REFERENCES

1. Rolfes,R.J. (2006) Regulation of purine nucleotide biosynthesis: in yeast and beyond. *Biochem. Soc Trans*, **34**, 786–790.
2. Zalkin,H. and Nygaard,P. (1996) In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. American Society for Microbiology Press, Washington, DC, pp. 561–579.
3. Neuhard,J. and Kelin,R.A. (1996) In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. American Society for Microbiology Press, Washington, DC, pp. 580–599.
4. Houlberg,U. and Jensen,K.F. (1983) Role of hypoxanthine and guanine in regulation of *Salmonella typhimurium* pur gene expression. *J. Bacteriol.*, **153**, 837–845.
5. Ma,H.W., Kumar,B., Ditges,U., Gunzer,F., Buer,J. and Zeng,A.P. (2004) An extended transcriptional regulatory network of *Escherichia coli* and analysis of its hierarchical structure and network motifs. *Nucleic Acids Res.*, **32**, 6643–6649.
6. Seshasayee,A.S., Fraser,G.M., Babu,M.M. and Luscombe,N.M. (2009) Principles of transcriptional regulation and evolution of the metabolic system in *E. coli*. *Genome Res.*, **19**, 79–91.
7. Martinez-Antonio,A. and Collado-Vides,J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol.*, **6**, 482–489.
8. Cho,B.K., Knight,E.M. and Palsson,B.O. (2006) PCR-based tandem epitope tagging system for *Escherichia coli* genome engineering. *BioTechniques*, **40**, 67–72.
9. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA*, **97**, 6640–6645.
10. Cho,B.K., Knight,E.M., Barrett,C.L. and Palsson,B.O. (2008) Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Res.*, **18**, 900–910.
11. Cho,B.K., Zengler,K., Qiu,Y., Park,Y.S., Knight,E.M., Barrett,C.L., Gao,Y. and Palsson,B.O. (2009) The transcription unit architecture of the *Escherichia coli* genome. *Nat. Biotechnol.*, **27**, 1043–1049.
12. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.
13. Tabor,C.W. and Tabor,H. (1985) Polyamines in microorganisms. *Microbiol. Rev.*, **49**, 81–99.
14. Jin,Y., Watt,R.M., Danchin,A. and Huang,J.D. (2009) Small noncoding RNA GcvB is a novel regulator of acid resistance in *Escherichia coli*. *BMC Genomics*, **10**, 165.
15. Balbontin,R., Fiorini,F., Figueroa-Bossi,N., Casadesus,J. and Bossi,L. Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Mol. Microbiol.*, **78**, 380–394.
16. Makaroff,C.A. and Zalkin,H. (1985) Regulation of *Escherichia coli purF*. Analysis of the control region of a pur regulon gene. *J. Biol. Chem.*, **260**, 10378–10387.
17. Rolfes,R.J. and Zalkin,H. (1988) *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. Cloning, nucleotide sequence, and interaction with the *purF* operator. *J. Biol. Chem.*, **263**, 19653–19661.
18. Aiba,A. and Mizobuchi,K. (1989) Nucleotide sequence analysis of genes *purH* and *purD* involved in the *de novo* purine nucleotide biosynthesis of *Escherichia coli*. *J. Biol. Chem.*, **264**, 21239–21246.
19. Flannigan,K.A., Hennigan,S.H., Vogelbacker,H.H., Gots,J.S. and Smith,J.M. (1990) Purine biosynthesis in *Escherichia coli* K12: structure and DNA sequence studies of the *purHD* locus. *Mol. Microbiol.*, **4**, 381–392.
20. Meng,L.M., Kilstrup,M. and Nygaard,P. (1990) Autoregulation of PurR repressor synthesis and involvement of *purR* in the regulation of *purB*, *purC*, *purL*, *purMN* and *guaBA* expression in *Escherichia coli*. *Eur. J. Biochem.*, **187**, 373–379.
21. Tiedeman,A.A., DeMarini,D.J., Parker,J. and Smith,J.M. (1990) DNA sequence of the *purC* gene encoding 5′-phosphoribosyl-5-aminoimidazole-4-N-succinocarboxamide synthetase and organization of the *dapA-purC* region of *Escherichia coli* K-12. *J. Bacteriol.*, **172**, 6035–6041.
22. He,B., Smith,J.M. and Zalkin,H. (1992) *Escherichia coli purB* gene: cloning, nucleotide sequence, and regulation by *purR*. *J. Bacteriol.*, **174**, 130–136.
23. Watanabe,W., Sampei,G., Aiba,A. and Mizobuchi,K. (1989) Identification and sequence analysis of *Escherichia coli purE* and *purK* genes encoding 5′-phosphoribosyl-5-amino-4-imidazole carboxylase for *de novo* purine biosynthesis. *J. Bacteriol.*, **171**, 198–204.
24. Smith,J.M. and Daum,H.A. III (1986) Nucleotide sequence of the *purM* gene encoding 5′-phosphoribosyl-5-aminoimidazole synthetase of *Escherichia coli* K12. *J. Biol. Chem.*, **261**, 10632–10636.
25. Choi,K.Y. and Zalkin,H. (1990) Regulation of *Escherichia coli pyrC* by the purine regulon repressor protein. *J. Bacteriol.*, **172**, 3201–3207.
26. Wilson,H.R. and Turnbough,C.L. Jr (1990) Role of the purine repressor in the regulation of pyrimidine gene expression in *Escherichia coli* K-12. *J. Bacteriol.*, **172**, 3208–3213.
27. Larsen,J.N. and Jensen,K.F. (1985) Nucleotide sequence of the *pyrD* gene of *Escherichia coli* and characterization of the flavoprotein dihydroorotate dehydrogenase. *Eur. J. Biochem.*, **151**, 59–65.
28. Devroede,N., Thia-Toong,T.L., Gigot,D., Maes,D. and Charlier,D. (2004) Purine and pyrimidine-specific repression of the *Escherichia coli carAB* operon are functionally and structurally coupled. *J. Mol. Biol.*, **336**, 25–42.
29. Piette,J., Nyunoya,H., Lusty,C.J., Cunin,R., Weyens,G., Crabeel,M., Charlier,D., Glansdorff,N. and Pierard,A. (1984) DNA sequence of the *carA* gene and the control region of *carAB*: tandem promoters, respectively controlled by arginine and the pyrimidines, regulate the synthesis of carbamoyl-phosphate synthetase in *Escherichia coli* K-12. *Proc. Natl Acad. Sci. USA*, **81**, 4134–4138.
30. Burton,K. (1994) Adenine transport in *Escherichia coli*. *Proc. Biol. Sci.*, **255**, 153–157.
31. Opdyke,J.A., Kang,J.G. and Storz,G. (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.*, **186**, 6698–6705.
32. Cho,B.K., Barrett,C.L., Knight,E.M., Park,Y.S. and Palsson,B.O. (2008) Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **105**, 19462–19467.
33. Keseler,I.M., Bonavides-Martinez,C., Collado-Vides,J., Gama-Castro,S., Gunsalus,R.P., Johnson,D.A., Krummenacker,M., Nolan,L.M., Paley,S., Paulsen,I.T. *et al.* (2009) EcoCyc: a comprehensive view of *Escherichia coli* biology. *Nucleic Acids Res.*, **37**, D464–D470.
34. He,B. and Zalkin,H. (1992) Repression of *Escherichia coli purB* is by a transcriptional roadblock mechanism. *J. Bacteriol.*, **174**, 7121–7127.

35. Karatza,P. and Frillingos,S. (2005) Cloning and functional characterization of two bacterial members of the NAT/NCS2 family in *Escherichia coli*. *Mol. Membrane Biol.*, **22**, 251–261.
36. Shimada,T., Ishihama,A., Busby,S.J. and Grainger,D.C. (2008) The *Escherichia coli*RutR transcription factor binds at targets within genes as well as intergenic regions. *Nucleic Acids Res.*, **36**, 3950–3955.
37. Krishna,S., Semsey,S. and Sneppen,K. (2007) Combinatorics of feedback in cellular uptake and metabolism of small molecules. *Proc. Natl Acad. Sci. USA*, **104**, 20815–20819.
38. Grainger,D.C., Aiba,H., Hurd,D., Browning,D.F. and Busby,S.J. (2007) Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res.*, **35**, 269–278.
39. Grainger,D.C., Hurd,D., Harrison,M., Holdstock,J. and Busby,S.J. (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl Acad. Sci. USA*, **102**, 17693–17698.