

Novel Members of the Cra Regulon Involved in Carbon Metabolism in *Escherichia coli*^{∇†}

Tomohiro Shimada, Kaneyoshi Yamamoto, and Akira Ishihama*

Department of Frontier Bioscience and Micro-Nano Technology Research Center,
Hosei University, Koganei, Tokyo 184-8584, Japan

Received 10 October 2010/Accepted 12 November 2010

Cra (catabolite repressor activator) is a global regulator of the genes for carbon metabolism in *Escherichia coli*. To gain insights into the regulatory roles of Cra, attempts were made to identify the whole set of regulation targets using an improved genomic SELEX (systematic evolution of ligands by exponential enrichment) system. Surprisingly, a total of 164 binding sites were identified for Cra, 144 (88%) of which were newly identified. The majority of known targets were included in the SELEX chip pattern. The promoters examined by the *lacZ* reporter assay *in vivo* were all regulated by Cra. These two lines of evidence indicate that a total of as many as 178 promoters are under the control of Cra. The majority of Cra targets are the genes coding for the enzymes involved in central carbon metabolism, covering all the genes for the enzymes involved in glycolysis and metabolism downstream of glycolysis, including the tricarboxylic acid (TCA) cycle and aerobic respiration. Taken together, we propose that Cra plays a key role in balancing the levels of the enzymes for carbon metabolism.

Central carbon metabolism consists of three major pathways, the glycolysis (Embden-Meyerhof-Parnas [EMP]) pathway, the pentose-phosphate (PP) pathway, and the tricarboxylic acid (TCA) cycle, altogether leading to the production of metabolic energy. These three pathways also provide a number of building blocks that are utilized for the construction of cell architecture (for reviews, see references 26, 43, and 50). The catabolic functions and regulations of the individual enzymes involved in these processes have been characterized in detail for both prokaryotes and eukaryotes. In agreement with the essential functions for energy production, the genes coding for these enzymes are conserved among a wide variety of organisms. Carbon availability in the environment influences the expression pattern of these genes in various ways. Expression regulation has been studied in detail for a small number of individual genes, mainly using the model prokaryote *Escherichia coli* (for examples, see references 25, 30, and 35). However, the mechanisms underlying the balancing of the expression level between a number of genes for the utilization of various carbon sources remain uncharacterized.

One well-characterized regulator of the genes for carbon source utilization is CRP (cyclic AMP [cAMP] receptor protein) (sometimes called catabolite activator protein [CAP]), which is converted into the functional form after binding cAMP that is synthesized in the absence of glucose (10, 22). Besides CRP, another transcription factor, Cra (catabolite repressor activator), has been indicated to play a role in the global regulation of the genes for carbon metabolism (16, 46). Cra, a member of the GalR-LacI family, was originally iden-

tified as FruR, for repression of the genes for fructose catabolism (13). The induction of the fructose operon takes place when the repressor Cra is inactivated after interactions with inducers such as D-fructose-1-phosphate and D-fructose-1,6-bisphosphate (31). Cra consists of two functional domains, an N-terminal DNA-binding domain with an H-T-H motif and a C-terminal inducer-binding and subunit-subunit contact domain. In the presence of glucose, the intracellular concentrations of the inducers increase, which interact with Cra to prevent its binding to the target promoters.

Later, FruR was indicated to be involved in the regulation of other genes encoding enzymes involved in the catabolic pathways of both glycolysis and gluconeogenesis and was later renamed Cra (38–41, 45). Up to the present time, however, the whole set of regulation targets of Cra remains unidentified. Moreover, the differences in the regulation targets and regulatory roles of two global regulators, CRP and Cra, in the control of a number of genes for carbon source utilization are totally left undetermined. For the identification of as-yet-unidentified genes under the control of Cra, we developed the genomic SELEX (systematic evolution of ligands by exponential enrichment) system and isolated a number of Cra-binding sequences from mixtures of *E. coli* genome fragments (46). After the cloning and sequencing of SELEX fragments (the SELEX-clos method), we identified 10 different sites of Cra binding on the *E. coli* genome. Among the 10 regulation targets of Cra, including *gapA* (glyceraldehyde-3-phosphate dehydrogenase A), *eno* (enolase), *hydN* (formate dehydrogenase H)-*hypF* (carbamoyl phosphate phosphatase), and *yahA* (unidentified transcription factor), which were predicted from the Cra-binding sequences, 4 were found to be newly identified targets, as detected by mRNA analyses of both the wild type and the *cra* deletion mutant (note that these 4 Cra targets have already been included in RegulonDB). The improved genomic SELEX screening system has been employed for the identifi-

* Corresponding author. Mailing address: Hosei University, Department of Frontier Bioscience, Koganei, Tokyo 184-8584, Japan. Phone and fax: 81-42-387-6231. E-mail: aishiham@hosei.ac.jp.

† Supplemental material for this article may be found at <http://jbb.asm.org/>.

[∇] Published ahead of print on 29 November 2010.

cation of regulation targets of a number of transcription factors (for instance, see references 16, 49, and 50).

The SELEX-clos analysis allowed the identification of DNA sequences with high affinity to Cra that were enriched, but those with weak affinity were lost after repeated SELEX cycling. For instance, a total of 43 clones within 100 independently isolated clones included the same promoter region of the *fruB* gene (the IIA component of the fructose phosphotransferase system [PTS]) with a high affinity for Cra, which had previously been identified to be under the control of Cra (41). As a shortcut approach to identifying the whole set of targets under the control of Cra, we have subjected a mixture of genomic SELEX fragments to chip analysis using an *E. coli* tiling array and identified a total of 164 target candidates under the control of Cra, including 20 known targets. The regulation *in vivo* of some of the newly identified targets has been analyzed and indeed has been found to be under the control of Cra. Taken together with the results described herein, we propose a revised model of the regulatory roles of Cra in carbon metabolism.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* BW25113 (W3110 *lacI^q mmBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78*) and the isogenic *cra* disruptant JW0078, a product of the Keio collection (1), were obtained from the *E. coli* Stock Center (National Institute of Genetics, Mishima, Japan). Cells were grown in LB medium at 37°C under aeration with constant shaking at 140 rpm. Cell growth was monitored by measuring the turbidity at 600 nm. Plasmid pRS551 was used for the LacZ reporter assay of Cra-depending promoters (see below). For the maintenance of pRS551, ampicillin and kanamycin were added, each at a final concentration of 50 μg/ml.

Purification of the Cra protein. An *E. coli* BL21(DE3) transformant with expression plasmid pCra of His-tagged Cra was grown in LB broth, and cells were harvested and subjected to Cra purification. Protein purification was carried out according to standard procedures used in our laboratory (47, 53). In brief, lysozyme-treated cells were sonicated in the presence of 100 mM phenylmethylsulfonyl fluoride. After centrifugation of the cell lysate (30 ml) at 15,000 rpm for 60 min at 4°C, the resulting supernatant was mixed with 2 ml of 50% Ni-nitrilotriacetic acid (Ni-NTA) agarose solution (Qiagen) and loaded onto a column. The column was washed first with 10 ml of lysis buffer and then with 10 ml of washing buffer (50 mM Tris-HCl [pH 8.0 at 4°C], 100 mM NaCl). Ni-NTA-bound proteins were then eluted with 2 ml of an elution buffer (50 mM Tris-HCl [pH 8.0 at 4°C], 100 mM NaCl, and 200 mM imidazole) and dialyzed against a storage buffer (50 mM Tris-HCl [pH 7.6], 200 mM KCl, 10 mM MgCl₂, 0.1 mM EDTA, 1 mM dithiothreitol, and 50% glycerol). The Cra protein used throughout this study was >95% pure, as analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The Cra protein without a His tag was purified as described previously (32).

Genomic SELEX search for Cra-binding sequences. The genomic SELEX method was carried out as previously described (46). A mixture of DNA fragments of the *E. coli* K-12 W3110 genome was prepared after the sonication of purified genome DNA and cloned into multicopy plasmid pBR322. For each SELEX screening, the DNA mixture was regenerated by PCR. For SELEX screening, 5 pmol of the mixture of DNA fragments and 10 pmol His-tagged Cra were mixed in a binding buffer (10 mM Tris-HCl [pH 7.8 at 4°C], 3 mM magnesium acetate, 150 mM NaCl, and 1.25 mg/ml bovine serum albumin) and incubated for 30 min at 37°C. For the effector-positive reaction of Cra, 3 mM D-fructose-1,6-bisphosphate (Sigma) was added throughout the experiment. The DNA-transcription factor mixture was applied onto a Ni-NTA column, and after washing out the unbound DNA with binding buffer containing 10 mM imidazole, DNA-protein complexes were eluted with an elution buffer containing 200 mM imidazole. DNA fragments recovered from the complexes were PCR amplified. For SELEX-clos, PCR products were cloned into the pT7 Blue-T vector (Novagen) and transformed into *E. coli* DH5α cells. The sequencing of each clone was carried out by using the T7 primer (5'-TAATACGACTCACTATAGGG-3'). For SELEX chip assays, PCR-amplified products of the isolated DNA-protein complexes obtained in the presence and absence of the effector were labeled separately with Cy5 and Cy3 and then combined. The fluorescently

labeled DNA mixtures were hybridized to a DNA microarray consisting of 43,450 species of a 60-bp-long DNA probe, which are designed to cover the entire *E. coli* genome at 105-bp intervals (Oxford Gene Technology, Oxford, United Kingdom). In parallel with the test samples, the PCR product of the original DNA library was also analyzed under the same conditions. The fluorescent intensity of the test sample with each probe was normalized with that of the corresponding peak of the original library. After the normalization of each pattern, the Cy5/Cy3 ratio was measured and plotted along the *E. coli* genome.

Reporter assay of the promoter activity and regulation. To measure the activity and regulation of promoters, a LacZ reporter assay was employed by using plasmid vector pRS551 (20). Promoter fragments between position -500 and the start codon were amplified by PCR using a pair of primers (F and R) (for sequences, see Table S1 in the supplemental material). The fragments were digested with BamHI and EcoRI and then ligated into pRS551. The construction of the plasmids was confirmed by DNA sequencing. Plasmids were transformed into test *E. coli* strains, wild-type BW25113, and *cra* mutant strain JW0078. Cultures grown overnight in LB medium were diluted 1:1,000 into fresh medium, and cells were grown for 4 h to an optical density at 600 nm (OD₆₀₀) of 0.4 to 0.5. The activity of β-galactosidase was measured according to the standard Miller method (29). The measurement was performed four times to obtain average values.

RESULTS

Search for the regulation targets of Cra by using genomic SELEX screening. In order to gain insight into genome regulation by Cra for the control of carbon metabolism in *E. coli*, we tried to identify the whole set of target promoters, genes, and operons under the control of Cra. For this purpose, we employed “genomic SELEX” screening, in which purified His-tagged Cra was mixed with a collection of *E. coli* genome fragments of 200 to 300 bp in length and Cra-bound DNA fragments were affinity purified for the identification of Cra recognition sequences (46). The original substrate mixture of genomic DNA fragments used for genomic SELEX screening formed smear bands on PAGE gels. After two cycles of genomic SELEX, DNA fragments with a high affinity for Cra were enriched, forming sharper bands on PAGE gels (note that the numbers of SELEX cycles needed to obtain sharp bands in SELEX screening differ between test transcription factors, reflecting the affinity for targets). In our previous study, we tried to identify the regulation targets of Cra after the cloning and sequencing of SELEX fragments (46), but using the SELEX-clos method, it is practically difficult to identify the whole set of regulation targets, including those with a low affinity for Cra.

For the identification of the whole set of regulation targets of Cra, in this study we employed the SELEX chip method, in which genomic SELEX fragments were labeled with Cy5 while the original DNA library was labeled with Cy3, and the mixture of two fluorescently labeled samples was hybridized to a DNA tiling microarray (Oxford Gene Technology, Oxford, United Kingdom) (48). For an elimination of the bias of library DNA, the ratio of fluorescence intensity bound to each probe for the test sample and for the original library DNA was measured and plotted against the corresponding position on the *E. coli* genome (Fig. 1). Since the 60-nucleotide-long probes are aligned along the *E. coli* genome at 105-bp intervals, approximately 300-bp-long SELEX fragments should bind to two or more consecutive probes; we employed this criterion for the identification of specific peaks. We then carried out a SELEX chip analysis in the presence and absence of effectors. In the presence of effectors, either D-fructose-1-phosphate or D-fructose-1,6-bisphosphate, SELEX bands remained smeared on PAGE

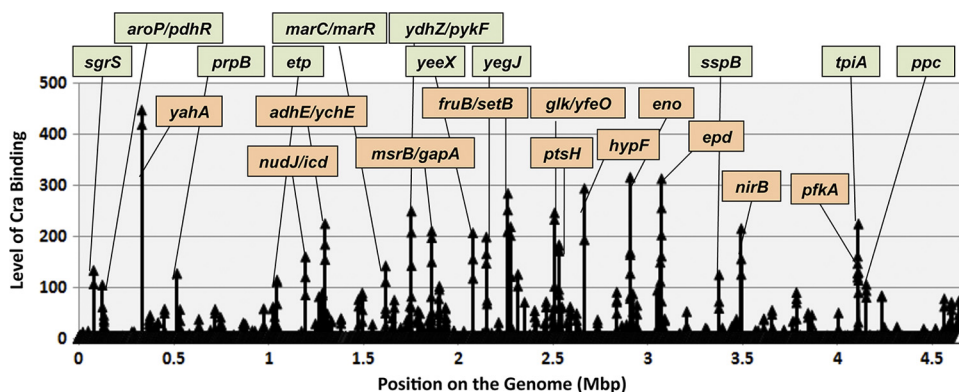


FIG. 1. Identification of Cra-binding sites on the *E. coli* genome. A genomic SELEX search of Cra-binding sequences was performed in the presence (open triangles) and absence (closed triangles) of D-fructose-1,6-bisphosphate using standard procedures (46). The resulting SELEX fragments were analyzed by using tiling DNA microarrays as described previously (48). The regulation target genes were predicted from the location of Cra-binding peaks. When the Cra-binding sequences were located within spacers, regulation targets were predicted based on the direction of transcription, either one direction (single gene) or both directions (two genes with a slash) (for details, see Table 2). The genes associated with some of the high-level peaks are indicated. Peaks shown in orange indicate the hitherto-identified targets, while peaks shown in green represent the newly identified targets.

gels even after 2 cycles of genomic SELEX, indicating that the sequence selectivity of Cra is higher without binding the effectors.

The SELEX chip analysis in the absence of an effector indicated a total of 194 peaks of Cra binding, but no clear peaks of Cra binding were identified in the presence of D-fructose-1,6-bisphosphate (Fig. 1). However, low but clear peaks were identified in the presence of effectors by expanding the y axis, indicating that the affinity for target sequences decreased for effector-bound Cra.

Of a total of 194 peaks, a group of 164 peaks is located upstream of open reading frames (ORFs), implying that these ORFs are the putative target genes or operons under the control of Cra. Some of those Cra-binding sites are located on ORFs of upstream genes. Among a total of 23 Cra targets listed in RegulonDB (12, 46), at least 20 were included in this group (Table 1; for details, see Table 2). Under the experimental conditions employed herein, we failed to identify SELEX peaks for only three known targets, *acnA* (aconitase), *cydA* (cytochrome *bd-I*), and *pck* (phosphoenolpyruvate [PEP] carboxykinase). Of these three targets, the involvement of Cra in expression regulation was experimentally identified only for the *pck* operon (6). Taken together, a total of 144 Cra-binding

sites were newly identified in this SELEX chip assay (Tables 1 to 3).

On the other hand, a group of 30 peaks (15.5%) was located downstream of ORFs, and thus, in this study, further examination was not performed for this group of Cra-binding sites. After genomic SELEX screening of a number of both hitherto-characterized and uncharacterized transcription factors from a total of about 300 species of *E. coli* transcription factors (16, 17), we realized that a specific set of transcription factors binds to the target sites located downstream of ORFs (T. Shimada et al., unpublished data), implying as-yet-unrecognized regulatory roles for such DNA signals.

Genomic SELEX screening was performed by using purified His-tagged Cra. The DNA recognition specificity of His-tagged Cra was considered to be the same with untagged Cra, as checked by gel shift experiments using some representative target promoters (data not shown).

Consensus recognition sequences of Cra. Cra forms two different conformations depending on the presence or absence of the effector D-fructose-1-phosphate or D-fructose-1,6-phosphate. Effector-free Cra has been indicated to bind the 14-bp-long Cra box with an imperfect palindromic sequence, GCTGAAACGTTTCA (16, 31, 46).

Since we obtained 6.3-fold-more targets than those hitherto identified (144 versus 23 targets), we reevaluated the consensus sequence of Cra binding using the whole set of predicted targets (167 targets), including 23 known targets. By using a collection of 500-bp sequences centered on each peak and BioProspector (<http://ai.stanford.edu/~xslu/Bio-Prospector/>), which was successfully employed previously for the identification of the RutR-box sequence (47), we identified a 14-bp Cra-box motif from all 167 Cra targets (Fig. 2). The Cra-box consensus sequence was identified to be 5'-GCTGAAAnCGnT TCA-3', in good agreement with previously reported sequences (16, 46). The G and C at positions 1 and 2 were indicated to be most important to exhibit high affinity for Cra (31, 39, 46; this study).

TABLE 1. Cra-binding sites on the *E. coli* genome^a

Target prediction	No. of binding sites	
	SELEX	Database
SELEX	144	
SELEX + RegulonDB	20	20
RegulonDB		3
Total	164	23

^a A collection of DNA fragments isolated by genomic SELEX was subjected to SELEX chip analysis. A total of 194 peaks were identified in the SELEX chip pattern, 164 of which were located within spacer regions. Among 23 known Cra targets, 20 were included in this library, and thus, a total of 144 Cra-binding sites were newly identified. There were 167 predicted targets.

TABLE 2. Cra-binding sites^a

Map position	Left gene	Cra site	Right gene	Cra-binding site determined by:	
				SELEX	RegulonDB
0.0	<i>yjiD</i> →		→ <i>thrLABC</i> ^c	S	
0.1	<i>thrC</i> →		→ <i>yaaX</i> ^c	S	
0.6	<i>fkpB</i> →	<i>ispH</i>	→ <i>rihC</i> ^c	S	
1.1	<i>kefC</i> →		→ <i>folA</i> ^c	S	
1.7	<i>tbpA-thiPQ</i> ← ^d	<i>sgrR</i>	→ <i>sgrST-setA</i> ^b	<i>sgrS</i>	
2.6	<i>aroP</i> ← ^b		→ <i>pdhR-aceEF-lpd</i> ^b	<i>pdhR</i>	<i>aroP</i>
2.7	<i>pdhR</i> →	<i>aceE</i>	→ <i>aceF-lpd</i> ^c	S	
2.8	<i>yacH</i> ← ^d		→ <i>acnB</i> ^b	S	<i>acnB</i>
2.9	<i>yacH</i> ← ^d	<i>acnB</i>	→ <i>yacL</i> ^d	S	
3.2	<i>panC</i> ← ^c	<i>panB</i>	← <i>yadC</i>	S	
5.3	<i>yafL</i> →		→ <i>yafM</i> ^c	S	
7.1	<i>betT</i> →		→ <i>yahA</i> ^b	S	<i>yahA</i>
7.4	<i>yahK</i> →	<i>yahL</i>	→ <i>yahM</i> ^c	S	
7.5	<i>prpR</i> ← ^d		→ <i>prpB</i> ^b	<i>prpB</i>	
7.9	<i>mhpR-lacI</i> ← ^d	<i>mhpA</i>	→ <i>mhpB</i> ^d	S	
8.1	<i>mhpE</i> →	<i>mhpT</i>	→ <i>yaiL</i> ^c	S	
8.7	<i>iraP</i> →		<i>phoA</i>	S	
9.0	<i>sbcD</i> ← ^d	<i>phoB</i>	→ <i>phoR</i> ^d	S	
9.7	<i>cyoABCDE</i> ← ^b		← <i>ampG</i>	<i>cyoA</i>	
11.5	<i>allR</i> →	<i>gcl</i>	→ <i>hyi-glxR-ybbW-allB-ybbY-glxK</i> ^c	S	
12.5	<i>nohB</i> →		→ <i>appY</i> ^c	S	
12.6	<i>envY-ompT</i> ← ^c	<i>ybcH</i>	← <i>nfrA</i>	S	
13.6	<i>ybdD</i> →	<i>ybdH</i>	→ <i>ybdL</i> ^c	S	
14.0	<i>citDEFXG</i> ← ^d	<i>citC</i>	→ <i>dpiBA</i> ^d	S	
15.0	<i>ybeZYX-lnt</i> ← ^d	<i>miaB</i>	→ <i>ubiF</i> ^d	S	
15.3	<i>fur</i> ← ^c		← <i>uof</i>	S	
15.5	<i>potE</i> ← ^d	<i>speF</i>	→ <i>ybjK</i> ^d	S	
15.8	<i>ybfB</i> →		→ <i>ybfC</i> ^c	S	
16.3	<i>glitA</i> ← ^d		→ <i>sdhCDAB-sucABCD</i> ^d	S	
16.6	<i>mngB</i> →		→ <i>cydAB</i> ^b		<i>cydA</i>
16.7	<i>tolR</i> →	<i>tolA</i>	→ <i>tolB-pal-ybgF</i> ^c	S	
17.9	<i>ybhF</i> ← ^c	<i>ybhG</i>	← <i>ybiH</i>	S	
18.7	<i>iaaA</i> →	<i>gsiA</i>	→ <i>gsiBCD</i> ^c	S	
18.9	<i>yliG</i> ← ^d	<i>bssR</i>	→ <i>yliI</i> ^d	S	
19.7	<i>hcp-hcr</i> ← ^c	<i>ybjE</i>	← <i>aqpZ</i>	S	
20.0	<i>aat</i> ← ^c	<i>cydC</i>	← <i>cydD</i>	S	
20.9	<i>msbA</i> →	<i>lpxK</i>	→ <i>ycaQ</i> ^c	S	
21.0	<i>smtA</i> →	<i>mukF</i>	→ <i>mukE</i> ^c	S	
22.1	<i>yccF</i> ← ^d		→ <i>helD</i> ^d	S	
22.2	<i>hspQ</i> ← ^d	<i>yccW</i>	→ <i>yccX</i> ^d	S	
22.2	<i>serT</i> ← ^d		→ <i>hyaABCDEFG</i> ^d	S	
22.5	<i>etp</i> ← ^c	<i>gfcE</i>	← <i>gfcD</i>	S	
22.5	<i>gfcE-etp-etk</i> ← ^c	<i>gfcD</i>	← <i>gfcC</i>	S	
24.9	<i>ycjH</i> →		→ <i>ptsG</i> ^c	S	
25.7	<i>nudJ</i> ← ^d	<i>rluE</i>	→ <i>icd</i> ^b	S	<i>icd</i>
26.9	<i>treA</i> ← ^c		← <i>dhaM</i>	S	
27.2	<i>yehM</i> ← ^c	<i>prs</i>	← <i>ispE</i>	S	
27.7	<i>narG</i> →	<i>narH</i>	→ <i>narII</i> ^c	S	
28.0	<i>adhE</i> ← ^b		→ <i>yehE</i> ^d	<i>adhE</i>	<i>adhE</i>
28.2	<i>kch</i> ← ^c		← <i>yciI</i>	S	
28.6	<i>yciN</i> ← ^d	<i>topA</i>	→ <i>cysB</i> ^d	S	
28.7	<i>yciX</i> →		→ <i>acnA</i> ^b		<i>acnA</i>
29.8	<i>ycjW</i> ← ^d		→ <i>ycjXF-tyrR</i> ^d	S	
31.8	<i>ydbC</i> →	<i>ydbD</i>	→ <i>ynbABCD</i> ^c	S	
32.2	<i>ydcI</i> ← ^d	<i>ydcJ</i>	→ <i>mdoD</i> ^d	S	
33.4	<i>yddG</i> ← ^d	<i>fdnG</i>	→ <i>fdnHI</i> ^d	S	
33.8	<i>gadC</i> ← ^c	<i>gadB</i>	← <i>pqqL</i>	S	
34.9	<i>marC</i> ← ^d		→ <i>marRAB</i> ^b	<i>marR</i>	
35.4	<i>ydfUT</i> ← ^c		← <i>rem</i>	S	
35.8	<i>ynfH</i> →	<i>dmsD</i>	→ <i>clbB</i> ^c	S	
35.9	<i>ynfK</i> ← ^c		← <i>dgsA</i>	S	
35.9	<i>dgsA</i> ← ^c		← <i>ynfL</i>	S	
36.0	<i>ynfL</i> ← ^d	<i>ynfM</i>	→ <i>asr</i> ^d	S	
36.1	<i>pntAB</i> ← ^d		→ <i>ydgH</i> ^d	S	
36.8	<i>rsxB</i> →	<i>rsxC</i>	→ <i>rsxDge-nth</i> ^c	S	

Continued on following page

TABLE 2—Continued

Map position	Left gene	Cra site	Right gene	Cra-binding site determined by:	
				SELEX	RegulonDB
37.0	<i>pdxY</i> ← ^c	<i>tyrS</i>	← <i>pdxH</i>	S	
37.6	<i>valW</i> →	<i>ydhR</i>	→ <i>ydhS</i> ^c	S	
37.8	<i>ydhZ</i> ← ^d		→ <i>pykF</i> ^b	<i>pykF</i>	<i>pykF</i>
38.5	<i>pps</i> ← ^b		→ <i>ydiA</i> ^d	<i>pps</i>	<i>pps</i>
39.1	<i>cedA</i> ← ^b	<i>katE</i>	← <i>chbG</i>	S	
39.2	<i>osmE</i> ← ^d	<i>nadE</i>	→ <i>cho</i> ^d	S	
39.3	<i>ves</i> ← ^c		← <i>spy</i>	S	
39.7	<i>ynjH</i> ← ^d		→ <i>gdhA</i> ^d	S	
39.9	<i>ydjE</i> ← ^c	<i>ydjF</i>	← <i>ydjG</i>	S	
40.1	<i>msrB</i> ← ^d		→ <i>gapA</i> ^b	S	<i>gapA</i>
40.4	<i>yeaO</i> →	<i>yoaF</i>	→ <i>yeaP</i> ^c	S	
40.9	<i>yoaE</i> ← ^d		→ <i>manXYZ</i> ^b	S	<i>manX</i>
41.1	<i>mgrB</i> ← ^c	<i>yobH</i>	← <i>kdgR</i>	S	
41.7	<i>edd-eda</i> ← ^b		← <i>zwf</i>	S	<i>edd</i>
41.7	<i>zwf</i> ← ^b	<i>yebK</i>	→ <i>pykA</i> ^d	<i>zwf</i>	
44.8	<i>yeex</i> ← ^c		← <i>yeaA</i>	S	
45.9	<i>wcaDEF-gmd-fcl-gmm-wcaI-cpsBG-wcaJ-wzxC</i> ← ^c	<i>wcaC</i>	← <i>wcaB</i>	S	
46.3	<i>yegD</i> →	<i>yegI</i>	→ <i>yegJ</i> ^c	S	
46.4	<i>ryeD</i> →		→ <i>mdtABCDbaeSR</i> ^c	S	
46.6	<i>ibsAB</i> ← ^d		→ <i>mdtABCDbaeSR</i> ^c	S	
46.9	<i>fbaB</i> ← ^b		→ <i>yegTUV</i> ^d	<i>fbaB</i>	
47.7	<i>yehUT</i> ← ^d		→ <i>mlrA</i> ^d	S	
48.7	<i>fruBKA</i> ← ^b		→ <i>setB</i> ^b	<i>fruB, setB</i>	<i>fruB</i>
51.3	<i>menDHBCE</i> ← ^c	<i>menF</i>	← <i>elaB</i>	S	
51.5	<i>nuoM</i> ← ^c	<i>nuoL</i>	← <i>nuoK</i>	S	
51.8	<i>nuoABCDEFGHIJKLMN</i> ← ^c		← <i>lrhA</i>	S	
52.0	<i>yfbUT</i> ← ^d	<i>yfbV</i>	→ <i>ackA</i> ^d	S	
53.1	<i>vacJ</i> ← ^d		→ <i>yfdC</i> ^d	S	
53.1	<i>vacJ</i> ← ^d	<i>yfdC</i>	→ <i>argW</i> ^d	S	
53.7	<i>yfdY</i> ← ^d		→ <i>lpxP</i> ^d	S	
53.8	<i>yfdY</i> ← ^c	<i>lpxP</i>	← <i>yfdZ</i>	S	
54.0	<i>glk</i> ← ^b		→ <i>yfeO</i> ^d	<i>glk</i>	<i>glk</i>
55.5	<i>maeB</i> ← ^b		→ <i>talA-tktB</i> ^d	<i>maeB</i>	
54.6	<i>cysK</i> →		→ <i>ptsHI-crr</i> ^b	<i>ptsH</i>	<i>ptsH</i>
55.8	<i>yffB</i> →	<i>dapE</i>	→ <i>ypfN</i> ^c	S	
56.8	<i>yfgJ</i> ← ^c	<i>der</i>	← <i>bamB</i>	S	
59.0	<i>bamD</i> ←		→ <i>raiA</i> ^c	S	
61.1	<i>hydN-hypF</i> ← ^b		← <i>ascG</i>	S	<i>hydN</i>
61.2	<i>ascG</i> ← ^d		→ <i>ascFB</i> ^d	S	
61.7	<i>ygbL</i> →	<i>ygbM</i>	→ <i>ygbN</i> ^c	S	
62.0	<i>ygbF</i> ← ^c	<i>ygbT</i>	← <i>ygcH</i>	S	
62.2	<i>cysH</i> ← ^c	<i>cysI</i>	← <i>cysJ</i>	S	
62.7	<i>eno</i> ← ^b	<i>pyrG</i>	← <i>mazG</i>	S	<i>eno</i>
62.9	<i>gudPXD</i> ← ^c		← <i>yqcA</i>	S	
63.1	<i>ygdH</i> →	<i>sdaC</i>	→ <i>sdaB</i> ^c	S	
63.2	<i>fucAO</i> ← ^d		→ <i>fucPIKUR</i> ^d	S	
63.4	<i>csdA</i> →		→ <i>csdE</i> ^c	S	
65.7	<i>ygfF</i> ← ^c	<i>gcvP</i>	← <i>gcvH</i>	S	
66.2	<i>epd-pgk-fbaA</i> ← ^b	<i>yggC</i>	← <i>yggD</i>	<i>epd</i>	<i>epd</i>
66.6	<i>yggI</i> →	<i>endA</i>	→ <i>rsmE-gshB</i> ^c	S	
68.3	<i>ygiV</i> ← ^d	<i>ygiW</i>	→ <i>qseBC</i> ^d	S	
69.1	<i>dan</i> ← ^d	<i>tttA</i>	→ <i>tttBT</i> ^d	S	
70.6	<i>garD</i> →	<i>sohA</i>	→ <i>yhaV</i> ^c	S	
71.3	<i>deaD</i> ← ^c		← <i>nlpI</i>	S	
71.4	<i>metY-rimP-nusA-intB-pnp</i> ← ^d		→ <i>argG</i> ^d	S	
72.7	<i>sspB</i> ← ^c	<i>sspA</i>	← <i>rpsI</i>	S	
73.1	<i>tldD</i> ← ^c	<i>yhdP</i>	← <i>rng</i>	S	
74.0	<i>fnt</i> ←	<i>rsmB</i>	→ <i>trkA</i> ^c	S	
74.2	<i>rplFRE-rpsE-rpmD-rplO-secY-rpmJ</i> ← ^c	<i>rpsH</i>	← <i>rpsN</i>	S	
74.4	<i>rplWB-rpsS-rplV-rpsC-rplP-rpmC-rpsQ</i> ← ^c	<i>rplD</i>	← <i>rplC</i>	S	
74.7	<i>bfd-bfr</i> ← ^c	<i>chiA</i>	← <i>tufA</i>	S	
74.9	<i>yheML</i> ← ^c	<i>yheN</i>	← <i>yheO</i>	S	
75.1	<i>yhfA</i> ← ^d		→ <i>crp</i> ^d	S	
75.3	<i>tsgA</i> →		→ <i>nirBDC-cysG</i> ^b	S	<i>nirB</i>
76.1	<i>yhgE</i> ← ^d		→ <i>pck</i> ^b		<i>pck</i>

Continued on following page

TABLE 2—Continued

Map position	Left gene	Cra site	Right gene	Cra-binding site determined by:	
				SELEX	RegulonDB
76.9	<i>glgCAP</i> ← ^b	<i>glgX</i>	→ <i>glgB</i> ^d	<i>glgC</i>	
77.8	<i>acpT</i> →		→ <i>nikABCDEF</i> ^c	S	
78.8	<i>hdeB-yhiD</i> ← ^d	<i>hdeA</i>	→ <i>hdeD</i> ^d	S	
78.8	<i>hdeD</i> →	<i>gadE</i>	→ <i>mdtE</i> ^c	S	
80.1	<i>yiaF</i> ← ^d		→ <i>yiaG</i> ^d	S	
81.1	<i>yibF</i> ← ^d	<i>rhsA</i>	→ <i>yibA</i> ^d	S	
81.3	<i>yibIH</i> ← ^d		→ <i>mtIADR</i> ^b	S	<i>mtIA</i>
81.4	<i>lldP</i> →	<i>lldR</i>	→ <i>lldD</i> ^c	S	
81.5	<i>yibN</i> ← ^d		→ <i>gpmM-envC-yibQ</i> ^b	<i>gpmM</i>	
81.9	<i>rfaJYZ-waaU</i> ← ^c	<i>rfaI</i>	← <i>rfaB</i>	S	
83.0	<i>uhpABC</i> ← ^c	<i>ilvN</i>	← <i>ilvB</i>	S	
83.1	<i>yidJ</i> ← ^d	<i>yidK</i>	→ <i>yidL</i> ^d	S	
83.3	<i>yidE</i> ← ^c		← <i>ibpB</i>	S	
83.7	<i>yidD</i> →	<i>yidC</i>	→ <i>mnmE</i> ^c	S	
85.0	<i>trpT</i> →	<i>hdfR</i>	→ <i>yifE</i> ^c	S	
86.3	<i>pldA</i> →	<i>recQ</i>	→ <i>rhtC</i> ^c	S	
86.5	<i>metR</i> ← ^c	<i>metE</i>	← <i>ysgA</i>	S	
88.5	<i>fieF</i> →		→ <i>pfkA</i> ^b	S	<i>pfkA</i>
88.5	<i>fieF</i> →	<i>pfkA</i>	→ <i>sbp</i> ^c	S	
88.6	<i>tpiA</i> ← ^b		← <i>yiiQ</i>	<i>tpiA</i>	
89.5	<i>ppc</i> ← ^b		← <i>argE</i>	<i>ppc</i>	
90.1	<i>rplL</i> →		→ <i>rpoBC</i> ^c	S	
90.8	<i>metA</i> →		→ <i>aceBAK</i> ^b	S	<i>aceB</i>
91.2	<i>lysC</i> ← ^d	<i>pgi</i>	→ <i>yjbEFGH</i> ^d	S	
91.5	<i>malEFG</i> ← ^d		→ <i>malK-lamB-malM</i> ^d	S	
92.0	<i>yjbR</i> →	<i>uvrA</i>	→ <i>ssb</i> ^c	S	
93.5	<i>adiA</i> ← ^d	<i>melR</i>	→ <i>melAB</i> ^d	S	
95.5	<i>ytfF</i> ← ^d	<i>ytfG</i>	→ <i>ytfH</i> ^d	S	
95.8	<i>ytfM</i> →	<i>ytfN</i>	→ <i>ytfP</i> ^c	S	
96.0	<i>fbp</i> ← ^d		→ <i>mpl</i> ^b	<i>mpl</i>	
96.4	<i>yjgH</i> ← ^d	<i>yjgI</i>	→ <i>yjgJ</i> ^d	S	
98.1	<i>uxuA</i> →	<i>uxuB</i>	→ <i>uxuR</i> ^c	S	
98.2	<i>yjiC</i> ← ^d		→ <i>yjiD</i> ^d	S	
98.3	<i>yjiK</i> ← ^c	<i>yjiL</i>	← <i>yjiM</i>	S	
98.9	<i>yjiY</i> ← ^d		→ <i>tsr</i> ^d	S	
99.2	<i>yjiB-dnaTC-yjiA</i> ← ^d	<i>yjiP</i>	→ <i>yjiQ-bglJ</i> ^d	S	
99.8	<i>nadR</i> →	<i>yjiK</i>	→ <i>slt</i> ^c	S	

^a A total of 164 Cra-binding sites are located either within spacers or on open reading frames, as shown in the “Cra site” column. The map positions of these sites on the *E. coli* K-12 genome are shown in the left-side “map position” column. Flanking genes are shown on both “left” and “right” sides of the Cra-binding sites. Arrows added in front of the gene designations show the direction of transcription. The Cra targets identified by SELEX chip are shown as S in the “SELEX” column. Regulation *in vivo* by Cra was confirmed for some of the predicted Cra targets, which are indicated by gene designations in the “SELEX” column. The Cra targets listed in RegulonDB are indicated in the “RegulonDB” column. Boldface indicates predicted Cra targets.

^b Cra targets experimentally identified in this study or listed in RegulonDB (group 1).

^c Single Cra targets located downstream of Cra-binding sites (group 2).

^d One of the divergently transcribed genes located downstream of Cra-binding sites (group 3) (note that the group 3 targets are counted as group 1/2 in Table 3).

Confirmation of regulation *in vivo* of the predicted target by Cra. Of the total of 23 known targets, we could identify clear SELEX chip peaks for at least 20 of them (shown in an orange background in Fig. 1). The high level of detection of known

targets itself indicates that the SELEX chip screening was successful for the detection of Cra targets and that the unidentified regulation targets of Cra must be included with a high probability in this genomic SELEX collection. To confirm that the newly identified target candidates are indeed under the

TABLE 3. Numbers of Cra-binding sites^a

Target	No. of Cra-binding sites determined by:	
	SELEX	RegulonDB
Total	164	23
Known targets of group 1	20	23
Newly identified targets of group 1	14	
Predicted targets of group 2	90	
Predicted targets of group 3	103	

^a The total number of binding sites predicted by SELEX and RegulonDB was 167. The total number of regulation target genes under the control of Cra was 178.5 (37 [group 1] + 90 [group 2] + 103 [group 3/2]).

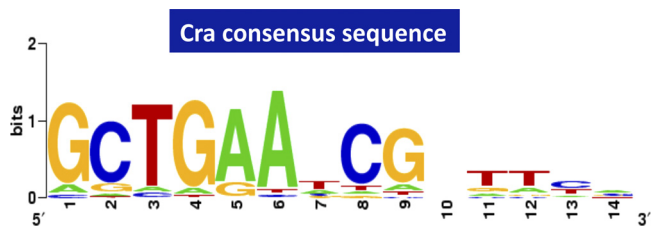


FIG. 2. Consensus sequences of Cra binding. Using the whole set of Cra-binding sequences identified by SELEX chip, the consensus sequences were reevaluated by using Logo analysis.

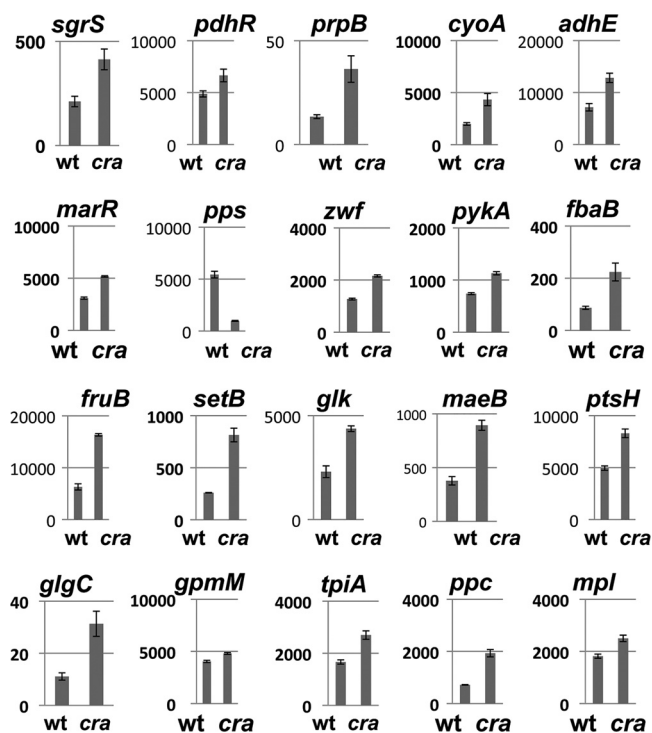


FIG. 3. Analysis of promoter regulation *in vivo* by Cra using LacZ fusions. The promoters carrying Cra-binding sites were fused to *lacZ*. The expression of LacZ was determined for wild-type (wt) and *cra*-defective strains (y axis, β -galactosidase activity in Miller units). The genes are aligned along the *E. coli* genome (from left to right in Fig. 1).

control of Cra, we constructed a set of promoter-*lacZ* fusions and performed the promoter assay in the presence and absence of Cra. When a single molecule of the transcription factor binds within the spacer between divergently transcribed genes in the *E. coli* genome, transcription toward one of the two opposite directions is regulated in most cases (16). The promoter assay also allowed the identification of the regulation target between the divergently transcribed genes and, moreover, to identify the mode of regulation *in vivo*, either activation or repression.

Here we focused on the set of genes that encode enzymes involved in central carbon metabolism. After the reporter assay of promoter-*lacZ* fusions in both wild-type *E. coli* and *cra* mutants was performed, all the genes tested were found to be under the control of Cra (Fig. 3; for the gene list, see Table 2), including the newly identified genes *sgrS* (*ptsG* stabilizer small RNA), *prpB* (2-methylisocitrate lyase), *cyoA* (cytochrome *o* ubiquinol oxidase), *marR* (multidrug resistance regulator), *zwf* (glucose-6-phosphate dehydrogenase), *fbaB* (fructose-bisphosphate aldolase), *setB* (lactose/glucose efflux system), *maeB* (NADP-linked malic enzyme), *glgC* (glucose-1-phosphate adenylyltransferase), *gpmM* (phosphoglyceromutase III), *tpiA* (triosephosphate isomerase), *ppc* (phosphoenolpyruvate carboxylase), and *mpl* (UDP-*N*-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase). The change in the promoter activity of the newly identified promoters in the *cra* mutant was as high as those of the known targets *pdhR* (*pdh* operon regulator), *adhE* (alcohol dehydrogenase), *pykA* (pyru-

vate kinase II), *pps* (phosphoenolpyruvate synthase), *fruB* (fructose-specific PTS IIA), *glk* (glucokinase), and *ptsH* (PTS system Hpr), which were analyzed as references (Fig. 3). The finding that all the target promoters examined are indeed under the control of Cra also supported the notion that genomic SELEX screening allowed the detection of specific Cra target promoters with a high level of accuracy.

The whole set of regulation targets of Cra. On the basis of two lines of evidence, i.e., the identification of about 90% of the known regulation targets by genomic SELEX screening and the confirmation of *in vivo* regulation by Cra for all the putative promoters tested, the promoters herein isolated by the SELEX chip screening could be targets under the control of Cra. A total of 164 Cra-binding sites within spacer regions can be classified into three groups: a total of 20 sites including the promoters for the known Cra-regulated genes (group 1) (Table 2), a total of 90 sites that are located upstream of specific promoters directing transcription toward one direction (group 2) (Table 2), and a total of 103 sites that are located within spacers between divergently transcribed genes (group 3) (Table 2). Since a single molecule of transcription factors associated between divergently transcribed genes in the *E. coli* genome generally regulates transcription toward one direction (16), we tentatively assumed that for the group 3 cases, one of the divergently organized genes is the regulation target of Cra (note that in Table 2, each of the divergently transcribed genes is counted as the target of 50% possibility). Based on these criteria, we propose that the total number of regulation target genes under the control of Cra is 178.5, as listed in Table 3. Thus, the total number of regulation targets of Cra increased 7.76-fold from 23 to 178.5.

In certain cases, such as those of *aroP*←Cra→*pdhR* and *fruB*←Cra→*setB* (Table 2), the spacer-bound Cra regulates transcription toward both directions. If such a gene organization exists in the uncharacterized group 3 targets, the number of regulation targets by Cra should be more than 178.5.

Classification of regulation modes. Among 20 species of the regulation targets shown in Fig. 3, the promoter activity of 19 species except for the *pps* promoter increased in the *cra* mutant, indicating that Cra is involved as a repressor for these operons. The promoter activity for the *pps* gene encoding phosphoenolpyruvate synthase, however, markedly decreased in the *cra* mutant (Fig. 3), indicating that Cra is involved as an activator for the *pps* operon. During growth on three-carbon substrates that require the gluconeogenesis pathway, the Pps enzyme provides PEP that is used for the synthesis of precursor metabolites for all cellular carbon compounds.

The binding site of Cra on some of the known targets listed in RegulonDB were experimentally identified by gel shift and DNase I footprinting assays, but in most cases, the Cra-binding sites were predicted on the basis of the location of the Cra consensus sequence. Since we obtained a refined consensus sequence of Cra binding (Fig. 2), the location of Cra-binding sites around the newly identified target promoters was also analyzed after an *in silico* search for the Cra-box sequence. For the promoters for which the activities were determined *in vivo* (Fig. 3), the locations of the Cra-binding site(s) relative to the transcription start site(s) are aligned along each target promoter (Fig. 4). The binding sites of Cra are widely distributed between positions -332.5 (*sgrS*) and $+7.5$ (*fruB*). Generally,

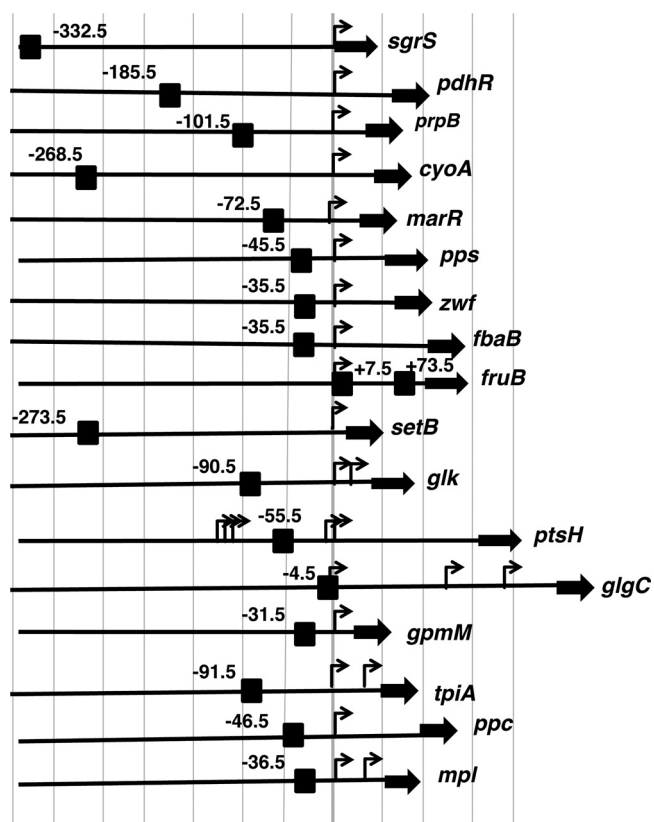


FIG. 4. Organization of the promoters under the control of Cra. The locations of Cra-binding sites (black box) between positions -350 and $+200$ relative to the transcription initiation site (arrow) are shown for the promoters, which were analyzed in Fig. 3. The number added at each Cra site represents the distance (in base pairs) from the respective transcription start site. Thick filled bars represent open reading frames of the target genes.

repressors bind within or downstream of target promoters (18, 19). For most of the targets identified herein, Cra represses transcription. In most cases, the binding sites of repressor Cra overlap with the target promoters (Fig. 4). Including the Cra targets analyzed herein (Fig. 3 and 4), the locations of the Cra box and promoter are aligned for all the known targets of Cra (Fig. 5). The mode of action of transcription activators has been established to be attributable to protein-protein interactions between upstream-bound transcription factors and RNA polymerase (18, 19). In good agreement with the proposed general rule, the binding sites of the repressor Cra are generally located near and downstream of the target promoters, while the activator Cra binds upstream of the target promoters, such as in the case of the transcription activation of the *pps* promoter by -45.5 -bound Cra (Fig. 3 and 4).

DISCUSSION

Regulation mode of Cra. Cra is a bifunctional transcription factor, acting as a repressor for most of the targets and as an activator for a small set of targets. In both modes, the level of binding of Cra to the target Cra box decreases in the presence of effectors. Thus, the regulation mode is determined by the location of Cra binding relative to the target promoters (Fig. 5). In several cases noted above, the repressor Cra was found to bind upstream from the target promoters, such as positions -332.5 of *sgrS*, -185.5 of *pdhR*, -101.5 of *prpB*, -268.5 of *cyoA*, and -273.5 of *setB* (Fig. 4). In certain special cases, transcription repression by upstream-bound transcription factors was suggested to be due to their tight binding to the α subunit of RNA polymerase, as detected for *aceBAK* repression by IclR (54). A similar mechanism may operate for the transcription repression of *glk* by -91.5 -bound Cra (28). The repression of the above-described four cases, however, takes place by Cra bound more than 100 bp upstream of the respec-

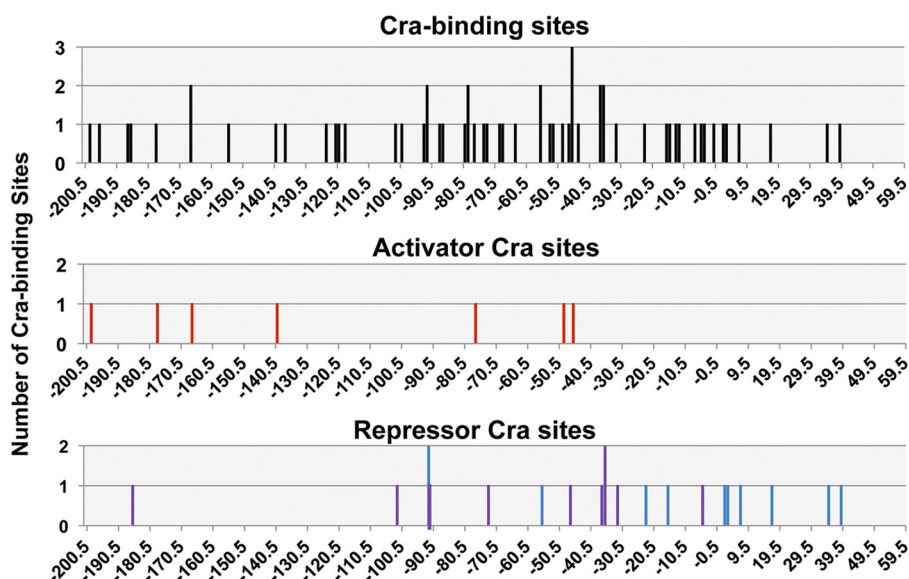


FIG. 5. Location of Cra-binding sites. The locations of Cra-binding sites (x axis) between positions -200 and $+60$ relative to transcription initiation sites for all the Cra targets so far analyzed are shown at the top. The activator Cra-binding sites are shown in the middle, while the repressor Cra-binding sites are shown at the bottom. The y axis represents the number of Cra-binding sites on the same site of the same promoter.

tive target promoters. The Cra-binding site (position -332.5) of *sgrS* is located on the structural gene of *sgrR*, the gene product of which regulates an overlapping *sgrST* operon encoding a small regulatory SgrS RNA and a small regulatory polypeptide, SgrT, with both being involved in the regulation of the *ptsG* operon (5, 36, 51, 52). Thus, Cra bound on *sgrS* may lead to the repression of *sgrR*, thereby resulting in a decrease in activator SgrR levels. The transcription of the *pdhR* gene is also repressed by upstream-bound Cra (position -185.5). On the *pdhR* promoter, multiple species of transcription factors associate, such as the activator CRP (positions -377.5, -110.5, -82.5, -70.5, -62.5, and -50.5), the activator Fnr (-50.5), and the repressor PdhR (-18) for the fine control of the key regulator of pyruvate dehydrogenase (PDH) complex formation (33, 37). Thus, the upstream-bound Cra may interact with one or more of these factors for effective repression. Likewise, the upstream-bound Cra on the *prpB* promoter may interact with the *prpB* promoter-interacting activators PrpR (positions -113.5 and -147.5) and CRP (position -106.5), leading to interference with activator functions (14). The Cra-binding site (position -273.5) of the *setB* gene corresponds to position -94.5 of the divergently transcribed *fruB* gene, on which two Cra sites have already been identified at positions +7.5 and +73.5 for effective repression (15). A possible interplay between Cra molecules on three Cra boxes within the 368-bp-long spacer may affect different modes of transcription regulation.

Novel regulatory roles of Cra. The most important regulatory role of Cra is believed to be the control of central carbon metabolism (10, 38). The newly identified targets were located within the metabolic map (Fig. 6). Cra-binding sites were found to be located on all the genes encoding the enzymes for central metabolism (Fig. 6). The promoter assay indicated that Cra represses the newly identified genes that are involved in glycolysis. In addition, Cra was found to repress the genes for the energy production pathway downstream of glycolysis. Both the *pykA* and *pykF* genes, encoding pyruvate kinase, are also repressed by Cra. Likewise, all three genes, *aceE* (PDH decarboxylase component E1), *lpd* (lipoamide dehydrogenase, PDH component E3), and *aceF* (PDH dihydrolipoyltransacetylase component E2), coding for the components of the PDH complex were all found to be repressed by Cra. The reaction of the PDH multienzyme complex encoded by *aceEF-lpd* is the gateway to the TCA cycle, producing acetyl coenzyme A (acetyl-CoA) (34). The promoter of the *pdhR-aceEF-lpd* operon is repressed by pyruvate-sensing PdhR (33) and Cra (this study) and is activated by CRP (55; T. Shimada et al., manuscript in preparation). All these three regulators also control the transcription of the key operon *cyoABCDE* for the terminal electron transport system downstream of the PDH cycle (32, 54; this study; Shimada et al., in preparation).

The pentose-phosphate (PP) pathway, also called the hexose monophosphate shunt, is one of three essential pathways, the EMP (Embden-Meyerhof-Parnas), PP, and TCA (tricarboxylic acid) pathways, of central metabolism as well as the pathway for providing building materials for nucleic acid synthesis. At the entry gate, D-glucose-6-phosphate is converted to D-glucono-lactone-6-phosphate by Zwf (44), which is metabolized to D-glyceraldehyde-3-phosphate by a set of enzymes, Pgl, Gnd, RpiAB, and TktAB, for reuse in central metabolism (15, 21).

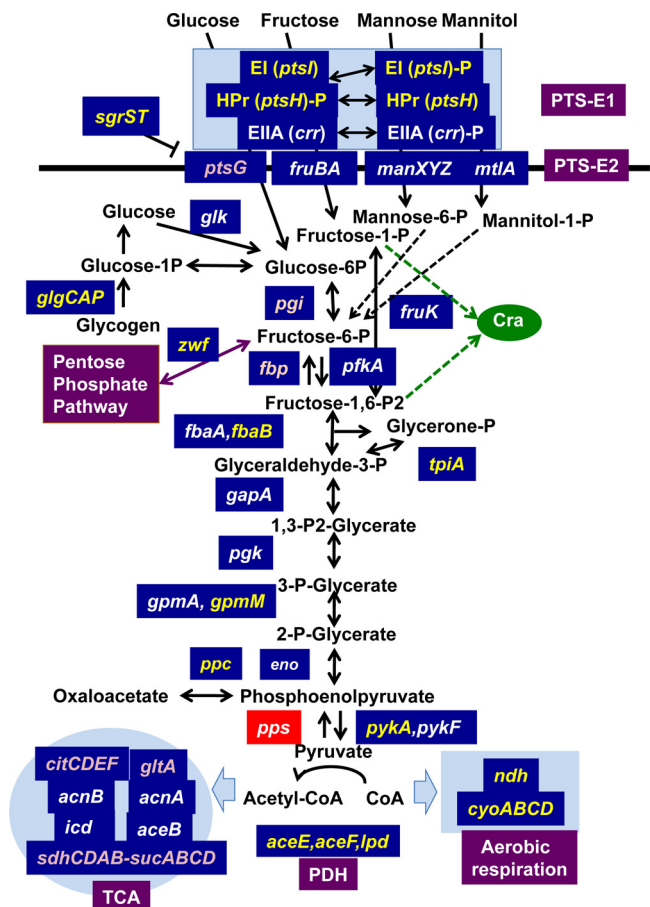


FIG. 6. Regulation targets of Cra. The genes shown within the metabolic map indicate the involvement of their gene products in the catalysis of the corresponding step reactions. The genes shown in blue boxes are under the control of Cra. The genes shown in yellow represent the Cra targets experimentally identified in this study, while the genes shown in purple represent Cra targets predicted solely on the basis of SELEX chip analyses.

After genomic SELEX, we found that Cra represses *zwf* involved in the initial step reaction.

Among a number of PTS transport systems for sugars (27), the *ptsI* gene, encoding PTS-E1, is under the control of Cra (Fig. 6). Among a total of 25 PTS-E2 systems for sugars, however, Cra is known to regulate only three systems, those for the transport of fructose, mannose, and mannitol. In addition, we identified PTS-E2 (PtsG) for glucose transport as an additional target of Cra. In contrast, most PTS-E2 systems are under the control of CRP (Shimada et al., in preparation).

Switching control between glycolysis and gluconeogenesis. Several step reactions of central carbon metabolism are apparently irreversible, with different enzymes being involved in the catalysis of forward and backward reactions. These irreversible reactions, sometimes called rate-controlling reactions, are considered to be involved in switching control between glycolysis and gluconeogenesis. There are 4 steps of the irreversible reactions in *E. coli* (Fig. 6): two at the entry gate of central metabolism for the supply of fructose-1,6-bisphosphate, i.e., the conversion of glucose to glucose-6-phosphate catalyzed by glucokinase (Glk) (28) and the synthesis of fructose-1,6-

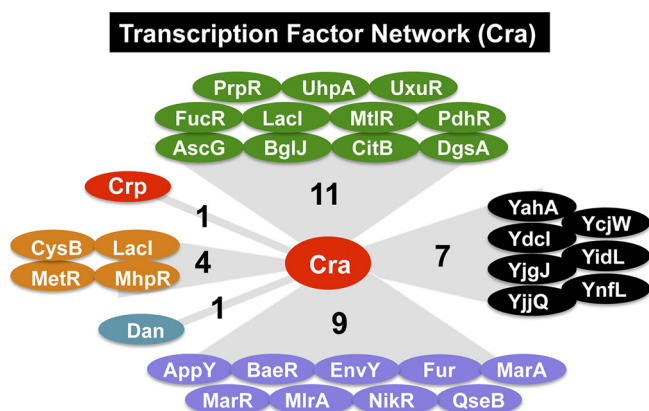


FIG. 7. Transcription factor network organized in the Cra regulon. The genes encoding transcription factors that are organized in the Cra regulon are shown. Transcription factors are classified based on their regulatory roles: green, transcription factors regulating the genes for carbon metabolism; red, global regulators of carbon metabolism; orange, transcription factors regulating the genes for nitrogen metabolism; blue, nucleoid protein with a regulatory role; purple, stress response regulators; black, uncharacterized transcription factors.

bisphosphate from fructose-6-phosphate catalyzed by four isozymes (Fbp, GlpX, YbhA, or YggF) (4, 9, 11, 24) (note that the reverse reaction is catalyzed by 6-phosphofruktokinase [PfkA] [8]), and two at the exit gate from the glycolysis pathway, i.e., the conversion of phosphoenolpyruvate to pyruvate by two pyruvate kinase isozymes (PykA and PykF) (3, 26) (the reverse reaction is catalyzed by phosphoenolpyruvate synthetase [PpsA] [7]) and the synthesis of phosphoenolpyruvate from oxaloacetate by phosphoenolpyruvate carboxykinase (Pck) (23). When the reactions leading to an accumulation of fructose-1,6-bisphosphate are activated, the overall reaction of carbon metabolism is directed toward gluconeogenesis.

Cra represses *pfkA*, *pykF*, and *glk*, thereby accelerating glycolysis (37). Cra also activates *pps* and *pck*, thus stimulating gluconeogenesis (38). Cra has been indicated to regulate all these rate-controlling genes and is thus considered to play a role in switching between glycolysis and gluconeogenesis. This regulatory function of Cra is controlled by the intracellular concentrations of key metabolites, including D-fructose-1-phosphate and D-fructose-1,6-bisphosphate. By SELEX chip analysis, a low-level peak of Cra binding was identified in the *pykA* promoter region. We then examined *pykA* promoter activity in the wild type and a *cra*-defective mutant and found that *pykA* is also repressed by Cra (Fig. 3). Glycogen, a glucose-containing polysaccharide, is synthesized for the storage of glucose. The *glgCAP* operon, encoding glycogen biosynthesis and degradation, is known to be regulated by CRP (2, 42). Here we also identified that the *glgC* promoter is repressed by Cra.

Transcription factor networks under the control of Cra. Promoters for the genes encoding the key metabolic enzymes and the essential cell architectures are under the control of multiple regulators, each monitoring different environmental conditions or metabolic states (16). After genomic SELEX screening, we discovered that a total of 33 genes encoding transcription factors are under the control of Cra, including 11 regulators of the genes for carbon metabolism, 4 regulators of

the genes for nitrogen metabolism, 9 regulators for stress response genes, and 7 putative regulators with unidentified regulatory roles (Fig. 7). In addition, Cra-binding sites of 4 regulators (MelR, HdfR, PhoB, and YdjF) were identified on the open reading frames but far from adjacent promoters. The integration of the genes encoding as many as 33 transcription factors into the Cra regulon indicates that a number of *E. coli* genes are under the indirect control of Cra. In particular, it is noteworthy that the *crp* gene is under the control of Cra because we identified more than 200 novel targets for cAMP-bound CRP (Shimada et al., in preparation).

The marked increase in the numbers of target genes under the control of Cra also indicates that the regulation targets for each of a total 300 species of transcription factors in *E. coli* (16, 17) must be more than those identified so far, and genomic SELEX screening should be a useful experimental system for searches of the whole set of regulation targets for each transcription factor.

ACKNOWLEDGMENTS

We thank Hiroshi Ogasawara and Jun Teramoto for discussion and Ayako Kori and Kayoko Yamada for preparation of proteins and technical support. *cra*-disrupted *E. coli* JW0078 was obtained from the *E. coli* Stock Center (National Institute of Genetics, Mishima, Japan).

This work was supported by grants-in-aid 21710198 to T.S. and 17076016, 8310133, and 21241047 to A.I. from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and by the Nano-Biology Project fund from the Micro-Nano Technology Research Center of Hosei University.

REFERENCES

- Baba, T., et al. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**:2006.0008. doi:10.1038/msb4100050.
- Ballicora, M. A., A. A. Iglesias, and J. Preiss. 2003. ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev.* **67**:213–225.
- Boiteux, A., M. Markus, T. Plesser, B. Hess, and M. Malcovati. 1983. Analysis of progress curves. Interaction of pyruvate kinase from *Escherichia coli* with fructose 1,6-bisphosphate and calcium ions. *Biochem. J.* **211**:631–640.
- Brown, G., et al. 2009. Structural and biochemical characterization of the type II fructose-1,6-bisphosphatase GlpX from *Escherichia coli*. *J. Biol. Chem.* **284**:3784–3792.
- Buhr, A., and B. Erni. 1993. Membrane topology of the glucose transporter of *Escherichia coli*. *J. Biol. Chem.* **268**:11599–11603.
- Chin, A. M., D. A. Feldheim, and M. H. Saier, Jr. 1989. Altered transcriptional patterns affecting several metabolic pathways in strains of *Salmonella typhimurium* which overexpress the fructose regulon. *J. Bacteriol.* **171**:2424–2434.
- Cooper, R. A., and H. L. Kornberg. 1967. The direct synthesis of phosphoenolpyruvate from pyruvate by *Escherichia coli*. *Proc. R. Soc. Lond. Biol. Sci.* **168**:263–280.
- Daldal, F. 1983. Molecular cloning of the gene for phosphofruktokinase-2 of *Escherichia coli* and the nature of a mutation, *pfkB1*, causing a high level of the enzyme. *J. Mol. Biol.* **168**:285–305.
- Donahue, J. L., J. L. Bownas, W. G. Niehaus, and T. J. Larson. 2000. Purification and characterization of *glpX*-encoded fructose 1,6-bisphosphatase, a new enzyme of the glycerol 3-phosphate regulon of *Escherichia coli*. *J. Bacteriol.* **182**:5624–5627.
- Fig, E., et al. 2009. cAMP receptor protein from *Escherichia coli* as a model of signal transduction in proteins—a review. *J. Mol. Microbiol. Biotechnol.* **17**:1–11.
- Fraenkel, D. G., and B. L. Horecker. 1965. Fructose-1,6-diphosphatase and acid hexose phosphatase of *Escherichia coli*. *J. Bacteriol.* **90**:837–842.
- Gama-Castro, S., et al. 2008. RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids. Res.* **36**:D120–D124.
- Geerse, R. H., J. van der Pluijm, and P. W. Postma. 1989. The repressor of the PEP-fructose phosphotransferase system is required for the transcription of the *pps* gene of *Escherichia coli*. *Mol. Gen. Genet.* **218**:348–352.
- Horswill, A. R., and J. C. Escalante-Semerena. 1997. Propionate catabolism

- in *Salmonella typhimurium* LT2: two divergently transcribed units comprise the *ppp* locus at 8.5 centisomes, *pppR* encodes a member of the sigma-54 family of activators, and the *pppBCDE* genes constitute an operon. *J. Bacteriol.* **179**:928–940.
15. Iida, A., S. Teshiba, and K. Mizobuchi. 1993. Identification and characterization of the *tktB* gene encoding a second transketolase in *Escherichia coli* K-12. *J. Bacteriol.* **175**:5375–5383.
 16. Ishihama, A. 2010. Prokaryotic genome regulation: multi-factor promoters, multi-target regulators and hierarchic networks. *FEMS Microbiol. Rev.* **34**: 628–645.
 17. Ishihama, A. 15 October 2009, posting date. Chapter 2.6, The nucleoid: an overview. In A. Böck et al. (ed.), *EcoSal—Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press, Washington, DC. <http://www.ecosal.org>.
 18. Ishihama, A. 1993. Protein-protein communication within the transcription apparatus. *J. Bacteriol.* **175**:2483–2489.
 19. Ishihama, A. 1992. Role of the RNA polymerase σ subunit in transcription activation. *Mol. Microbiol.* **6**:3283–3288.
 20. Ishii, D., A. Ishihama, and K. Yamamoto. 2009. Two modes of autoregulation of the *murR* repressor in *Escherichia coli*. *Biosci. Biotechnol. Biochem.* **73**:2528–2530.
 21. Katz, J., and R. Rognstad. 1967. The labeling of pentose phosphate from glucose-14C and estimation of the rates of transaldolase, transketolase, and the contribution of the pentose cycle, and ribose phosphate synthesis. *Biochemistry* **6**:2227–2247.
 22. Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya. 1993. Transcriptional regulation by cAMP receptor protein of *Escherichia coli*. *Annu. Rev. Biochem.* **62**:749–795.
 23. Krebs, A., and W. A. Bridger. 1980. The kinetic properties of phosphoenolpyruvate carboxylase of *Escherichia coli*. *Can. J. Biochem.* **58**:309–318.
 24. Kuznetsova, E., et al. 2006. Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**:36149–36161.
 25. Lemuth, K., et al. 2008. Global transcription and metabolic flux analysis of *Escherichia coli* in glucose-limited fed-batch cultivations. *Appl. Environ. Microbiol.* **74**:7002–7015.
 26. Malcovati, M., and G. Valentini. 1982. AMP- and fructose 1,6-bisphosphate-activated pyruvate kinases from *Escherichia coli*. *Methods Enzymol.* **90**(Pt. E):170–179.
 27. Mayer, C., and W. Boos. 29 March 2005, posting date. Chapter 3.4.1, Hexose/pentose and hexitol/pentitol metabolism. In A. Böck et al. (ed.), *EcoSal—Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press, Washington, DC. <http://www.ecosal.org>.
 28. Meyer, D., C. Schneider-Fresenius, R. Horlacher, R. Peist, and W. Boos. 1997. Molecular characterization of glucokinase from *Escherichia coli* K-12. *J. Bacteriol.* **179**:1298–1306.
 29. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
 30. Nanchen, A., A. Schicker, O. Revelles, and U. Sauer. 2008. Cyclic AMP-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*. *J. Bacteriol.* **190**: 2323–2330.
 31. Negre, D., et al. 1996. Definition of a consensus DNA-binding site for the *Escherichia coli* pleiotropic regulatory protein FruR. *Mol. Microbiol.* **21**:257–266.
 32. Negre, D., et al. 1998. FruR-mediated transcriptional activation at the *ppsA* promoter of *Escherichia coli*. *J. Mol. Biol.* **276**:355–365.
 33. Ogasawara, H., Y. Ishida, Y. Yamada, K. Yamamoto, and A. Ishihama. 2007. PdhR (pyruvate dehydrogenase complex regulator) controls the respiratory electron transport system in *Escherichia coli*. *J. Bacteriol.* **189**:5534–5541.
 34. Patel, M. S., and T. E. Roche. 1990. Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *FASEB J.* **4**:3224–3233.
 35. Perrenoud, A., and U. Sauer. 2005. Impact of global transcriptional regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on glucose catabolism in *Escherichia coli*. *J. Bacteriol.* **187**:3171–3179.
 36. Postma, P. W., J. W. Lengeler, and G. R. Jacobson. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**:543–594.
 37. Quail, M. A., and J. R. Guest. 1995. Purification, characterization and mode of action of the PdhR, the transcriptional repressor of the *pdhR*-*aceEF*-*lpd* operon of *Escherichia coli*. *Mol. Microbiol.* **15**:519–529.
 38. Ramseier, T. M. 1996. Cra and the control of carbon flux via metabolic pathways. *Res. Microbiol.* **147**:489–493.
 39. Ramseier, T. M., S. Bledig, V. Michotey, R. Feghali, and M. H. Saier, Jr. 1995. The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. *Microbiology* **16**:1157–1169.
 40. Ramseier, T. M., S. Y. Chien, and M. H. Saier, Jr. 1996. Cooperative interaction between Cra and Fnr in the regulation of the *cydAB* operon of *Escherichia coli*. *Curr. Microbiol.* **33**:270–274.
 41. Ramseier, T. M., et al. 1993. *In vitro* binding of the pleiotropic transcriptional regulatory protein, FruR, to the *fru*, *pps*, *ace*, *pts* and *icd* operons of *Escherichia coli* and *Salmonella typhimurium*. *J. Mol. Biol.* **234**:28–44.
 42. Romeo, T., and J. Preiss. 1989. Genetic regulation of glycogen biosynthesis in *Escherichia coli*: *in vitro* effects of cyclic AMP and guanosine 5'-diphosphate 3'-diphosphate and analysis of *in vivo* transcripts. *J. Bacteriol.* **171**: 2773–2782.
 43. Romeo, T., and J. L. Snoep. 7 October 2005, posting date. Chapter 3.5.1, Glycolysis and flux control. In A. Böck et al. (ed.), *EcoSal—Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press, Washington, DC. <http://www.ecosal.org>.
 44. Rowley, D. L., and R. E. Wolf. 1991. Molecular characterization of the *Escherichia coli* K-12 *zwf* gene encoding glucose 6-phosphate dehydrogenase. *J. Bacteriol.* **173**:968–977.
 45. Saier, M. H., Jr. 1998. Multiple mechanisms controlling carbon metabolism in bacteria. *Biotechnol. Bioeng.* **58**:170–174.
 46. Shimada, T., N. Fujita, M. Maeda, and A. Ishihama. 2005. Systematic search for the Cra-binding promoters using genomic SELEX system. *Genes Cells* **10**:907–918.
 47. Shimada, T., K. Hirao, A. Kori, K. Yamamoto, and A. Ishihama. 2007. RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines. *Mol. Microbiol.* **66**:744–757.
 48. Shimada, T., A. Ishihama, S. J. Busby, and D. C. Grainger. 2008. The *Escherichia coli* RutR transcription factor binds at targets within genes as well as intergenic regions. *Nucleic Acids Res.* **36**:3950–3955.
 49. Shimada, T., K. Yamamoto, and A. Ishihama. 2009. Involvement of the leucine response transcription factor LeuO in regulation of the genes for sulfa drug efflux. *J. Bacteriol.* **191**:4562–4571.
 50. Teramoto, J., S. H. Yoshimura, K. Takeyasu, and A. Ishihama. 2010. A novel nucleoid protein of *Escherichia coli* induced under anaerobic growth conditions. *Nucleic Acids Res.* **38**:3605–3618.
 51. Vanderpool, C. K. 2007. Physiological consequences of small RNA-mediated regulation of glucose-phosphate stress. *Curr. Opin. Microbiol.* **10**:146–151.
 52. Vanderpool, C. K., and S. Gottesman. 2004. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol. Microbiol.* **54**:1076–1089.
 53. Yamamoto, K., et al. 2005. Functional characterization *in vitro* of all two-component signal transduction systems from *Escherichia coli*. *J. Biol. Chem.* **280**:1448–1456.
 54. Yamamoto, K., and A. Ishihama. 2003. Two different modes of transcription repression of the *Escherichia coli* acetate operon by IclR. *Mol. Microbiol.* **47**:183–194.
 55. Zhang, Z., et al. 2005. Functional interactions between the carbon and iron utilization regulators, Crp and Fur, in *Escherichia coli*. *J. Bacteriol.* **187**:980–990.