Transcriptional regulation of a set of genes for synthesis and degradation of pyrimidines by RutR, the uracil/thymine sensing master regulator

SHIMADA, Tomohiro / 島田, 友裕

(発行年 / Year)
2008-03-24
(学位授与番号 / Degree Number)
32675甲第202号
(学位授与年月日 / Date of Granted)
2008-03-24
(学位名 / Degree Name)
博士(工学)
(学位授与機関 / Degree Grantor)
法政大学 (Hosei University)
(URL)
https://doi.org/10.15002/00005722

Doctoral Dissertation

2008, March

Transcriptional regulation of a set of genes for synthesis and degradation of pyrimidines by RutR, the uracil/thymine sensing master regulator

Tomohiro Shimada

Graduate School, Hosei University

Division of Material Chemistry

CONTENTS

SUMMARY

Using the genomic SELEX, a total of six *E. coli* DNA fragments have been identified, which formed complexes with transcription factor RutR, a putative regulator of the genes for pyrimidine degradation. The RutR regulon was found to include a large number of genes encoding components for not only degradation of pyrimidines but also transport of glutamate, synthesis of glutamine, conversion of glutamine to carbamoyl phosphate, synthesis of pyrimidine nucleotides and arginine from carbamoyl phsphate, and degradation of purines. DNase-I footprinting indicated that RutR recognizes a palindromic sequence of TTGACCAnnTGGTCAA. The RutR box in P1 promoter of *carAB* encoding carbamoyl phosphate synthetase, a key enzyme of pyrimidine synthesis, overlaps with the PepA (CarP) repressor-binding site, implying competition between RutR and PepA. RutR binding *in vitro* to the *carAB* P1 promoter was abolished by adding either uracil or thymine. Accordingly, in the *rutR* deletion mutant or in the presence of uracil, the activation *in vivo* of *carAB*

P1 promoter was markedly reduced. Northern blot analysis of the RutR target genes indicated that RutR represses the Gad system genes involved in glutamate-dependent acid resistance and allantoin degradation. We also analyzed RutR distribution on the *E. coli* genome in living cells by ChIP-chip methods. Cells were analyzed during growth phase and with or without its effecter, uracil. In addition of the RutR targets identified by genomic SELEX, a total of 19 new RutR targets have been identified. At least one of these putative RutR targets, *ves* (*ydjR*), was identified to be under the regulation of RutR and effector uracil. Altogether we propose that RutR is the pyrimidine sensor and the master regulator for a large set of the genes involved in the synthesis and degradation of pyrimidines.

INTRODUCTION

About 300 species of transcription factors exist in *Escherichia coli*, but the regulatory functions remain unidentified for almost one third of these factors (Perez-Rueda and Collado-Vides, 2000; Ishihama, 2000). The putative transcription factors with yet unidentified functions are considered to play regulatory roles in transcription of the genes needed for stress response in nature, which are not needed under laboratory culture conditions. Target genes under the control of each transcription factor could be predicted by Microarray analysis of mutants lacking a test transcription factor (Khodursky *et al.*, 2000; Wei *et al.*, 2001), but the discrimination between direct and indirect effects of the gene disruption is often difficult because *E. coli* transcription factors form the cascade of regulation network (for example see Ogasawara *et al.*, 2007a; 2007b). To overcome the difficulty arisen from the Microarray assay, we have initiated a systematic search for the recognition sequences by all the putative transcription factors using the newly developed genomic SELEX system

(Shimada *et al.*, 2005; Ogasawara *et al.*, 2007a; 2007b).

Along the systematic search for target genes and promoters under the control of the putative transcription factors from *E. coli*, we initiated the search for targets by a putative transcription factor YcdC, which consists of 212 amino acid residues and belongs to the TetR family. This group of DNA-binding transcription factors is composed of an N-terminal DNA-binding domain and a C-terminal ligand-binding domain (Ramos *et al.*, 2005). YcdC forms a homo-dimmer, and its crystal structure has been solved at the time when its regulatory function was not yet identified (see YcdC structure; Patskovdky *et al.*, 2006). In the middle of this research, Loh *et al.* (2006) reported that the undescribed genes, *ycdMLKJIHG*, are involved in the degradation of pyrimidines for reutilization as nitrogen sources [the operon was then renamed to *rutABCDEFG* (pyrimidine utilization)] and transcription of this *rut* operon could be repressed by YcdC [also renamed to RutR]. After the genomic SELEX search, we indeed identified a RutR-binding sequence in the intergenic region between *rutABCDEFG* and *rutR*. In addition, a number of novel targets have been identified in this research, including the

carAB operon encoding carbamoyl phosphate synthetase, which plays a key role in the *de novo* synthesis of pyrimidine nucleotides (Bouvier *et al.*, 1984), the *gadAX* operon encoding glutamate decarboxylase A and a global regulator GadX of the *gad* regulon (Mar *et al.*, 2002; 2003), the *gadBC* operon coding for glutamate decarboxylase B and glutamate-GABA (gamma-aminobutyrate) antiporter (Richard and Foster, 2003), the *ygiF-glnE* operon for glutamine synthesis from glutamate (van Heeswijk *et al.*, 1993), and the *gcl-hyi-glxR* operon for degradation of allantoin (Cusa *et al.*, 1999). After detailed analysis of transcription for each of the predicted RutR targets, we propose that RutR is not only a repressor for the genes for degradation of pyrimidines (Loh *et al.*, 2006) but an important master regulator for the supply of glutamate, the conversion of glutamate to glutamine, the conversion of glutamine to carbamoyl phosphate, the synthesis of pyrimidine nucleotides and arginine from carobamoyl phosphate, and the degradation of pyrimidines and purines for maintenance of metabolic balance between pyrimidines and purines. These genes are also involved in the maintenance of intracellular pH homeostasis under external acidic conditions.

To confirm the results of genomic SELEX analysis, we then performed ChIP-chip (chromatin immunoprecipitation in conjunction with high-density microarrays) experiments to study the distribution *in vivo* of RutR along the *E. coli* genome. ChIP-chip method on *E.coli* is well studied to measure protein-DNA interactions *in vivo* (Wade *et al.*, 2007). RutR is uracil/thymine sensing regulator, then Wild-type cells and *rutR* mutant cells were inoculated and compared with using ChIP-chip analysis under the three difference conditions; log phase, stationary phase and log phase with its effector uracil. This allowed us to identify 24 DNA targets for RutR in the log phase cells including 4 targets from genomic SELEX. In the stationary phase cells, 13 DNA targets were identified and all of them are overlapped with the log phase targets. Then we confirmed the binding activity of newly RutR targets with gel shift assay, both 4^{th} A and 13^{th} T of the RutR binding motif are important for high affinity to RutR *in vitro*.

RutR-binding activity of these newly identified targets was examined by gel shift assay and transcription profile in vivo of each of the target sequences were examined by Northern blot analyais. Physiological roles of RutR will be

discussed taking all the *in vitro* genomic SELEX and in vivo ChIP-chip results together.

MATERIALS AND METHODS

Bacterial strains and culture conditions

Escherichia coli BW25113 (W3110 *lacI*^{q} *rrnBT14* Δ *lacZWJ16* *hsdR514* Δ *araBADAH33,* Δ *rhaBADLD78)* was used as wild-type strain. Starting from BW25113 strain, a *rutR* disruptant strain JW0998 was constructed (Baba *et al.*, 2006). Cells were grown at 37^{o}C under aeration in M9 glucose (0.4%) medium. Cell growth was monitored by measuring the turbidity at 600 nm.

Purification of RutR protein

Plasmid pYcdC for expression of His-tagged RutR was constructed by inserting a PCR-amplified RutR-coding sequence DNA into pET21a(+) (Novagen) between *Nde*I and *Not*I sites. pYcdC was transformed into *E. coli* BL21 (DE3). Transformants were grown in LB broth and RutR expression was induced by adding 1 mM IPTG. After 3 hr induction, cells were harvested and

subjected to RutR purification. Protein purification was carried out according to the standard procedure in this laboratory. In brief, lysozyme-treated cells were sonicated in the presence of 100 mM PMSF. After centrifugation of cell lysate at 15,000 rpm for 20 min at 4°C, the resulting supernatant was mixed with 2 ml of 50% Ni-NTA agarose solution (Qiagen) and loaded onto a column. After washing with 10 ml of lysis buffer, the column was washed with 10 ml of washing buffer (50 mM Tris-HCl, pH 8.0 at 4°C, 100 mM NaCl). Proteins were then eluted with 2 ml of an elution buffer (200 mM imidazole, 50 mM Tris-HCl, pH 8.0 at 4°C, 100 mM NaCl), and dialyzed against a storage buffer (50 mM Tris-HCl, pH 7.6, 200 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, and 50% glycerol). RutR used throughout this study was more than 95% pure as analyzed by SDS-PAGE.

SELEX search for RutR-binding sequences

The substrate DNA for SELEX screening was prepared by PCR amplification of 200-300 bp-long *E. coli* DNA fragments using the plasmid library as template, each carrying an *E. coli* genome fragment of 200-300 bp in length

(Shimada *et al.*, 2004). For SELEX screening, 5 pmol of DNA fragments and 10 pmol His-tagged YcdC were mixed in a binding buffer (10 mM Tri-HCl, pH 7.8 at 4^{o}C, 3 mM Mg acetate, 150 mM NaCl, BSA 1.25 ug/ml) and incubated for 30 min at 37^{o}C. The mixture was applied onto Ni-NTA column and after washing unbound DNA with the binding buffer containing 10 mM imidazole, DNA-YcdC complexes were eluted with an elution buffer containing 100 mM imidazole. DNA fragments recovered from the complexes were ligated into pBR322 and PCR-amplified as above. If necessary, this SELEX cycle was repeated several times. For sequencing of RutR-bound DNA fragments, PCR products were cloned into pT7 Blue-T vector (Novagen) and transformed into *E. coli* DH5α. Sequencing was carried out using T7-primer (5'-TAATACGACTCACTATAGGG-3').

Gel mobility shift assay

All the DNA probes for the mobility shift assay were generated by PCR amplification of the promoter region of RutR target genes using a set of primers,

FITC-labelled T7-F primer (5'- TAATACGACTCACTATAGGG-3') and unlabelled T7-R primer (5'- GGTTTTCCCAGTCACACGACG-3'), cloned SELEX fragment (80 ng each) as template, and Ex-Taq DNA polymerase. After purification by PAGE, the FITC-labelled probes were incubated at 37 °C for 15 min with RutR in 0.01 ml of transcription buffer, which contained 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 3 mM Mg acetate, 0.1 mM EDTA, 0.1mM dithiothreitol, and 25 μg/ml of BSA. Reaction mixtures were directly subjected to 6% polyacrylamide gel electrophoresis. For search of the effector for RutR, the gel shift reaction was performed under the same conditions in the presence of various pyrimidine metabolites at the final concentration of 100 μM.

DNase-I footprinting

The probe was generated by PCR amplification of the RutR-binding sequence, isolated by genomic SELEX and identified by gel-shift assay, using ^{32}P-end-labeled primers, *E. coli* genome DNA (100 ng) as the template and Ex Taq DNA polymerase. PCR products were purified by PAGE. DNase-I

footprinting was performed essentially according to the standard procedure (Yamamoto and Ishihama, 2003). In brief, each labeled probe was incubated at 37°C for 15 min with RutR in 25 μl transcription buffer. After incubation for 15 min, DNA digestion was initiated by the addition of 5 ng of DNase I (Takara). After digestion for 45 s at 25°C, the reaction was terminated by the addition of 45 μl of DNase I stop solution (20 mM EDTA, 200 mM NaCl, 1% SDS, 250 μg of yeast tRNA per ml). Digested products were precipitated by adding ethanol, dissolved in formamide dye solution, and analyzed by electrophoresis on a 6% polyacrylamide gel containing 8 M urea.

Measurement of in vivo *promoter activity*

Promoter activity *in vivo* was measured using the newly developed DFP (double-fluorescent protein) vector (Makinoshima *et al.*, 2002; Shimada *et al.*, 2004). Briefly the test promoter of 300-500 bp starting from the initiation codon is inserted into TFP (two-fluorescent protein) vector so to adjust the initiation codon of GFP (green fluorescent protein) while RFP (red fluorescent protein) was

expressed by the reference promoter *lac*UV5. Fluorescent intensities of GFP and RFP were measured as described before (Makinoshima *et al.*, 2002; Shimada *et al.*, 2004).

S1 mapping of carA *mRNAs*

The ^{32}P-end-labeled *carA* promoter fragment was amplified by PCR using W3110 genomic DNA (100 ng) as template, ^{32}P-end-labeled *carA* R and unlabeled *carA* F primers (Table 1), and Ex Taq DNA polymerase (Takara). The ^{32}P-labeled fragment was purified by PAGE. Mixtures of the ^{32}P-end-labeled *carA* probe (10^4 cpm) and total RNA (50 μg) were incubated for 10 min at 75°C for denaturation, and then incubated at 37°C overnight for hybridization. After digestion with 50 units of S1 nuclease (Takara) at 37 °C for 10 min, undigested products were extracted with phenol, precipitated with ethanol, and analyzed by electrophoresis on polyacrylamide gels containing 8 M urea. The gel was dried and exposed on Image plate (Fuji). The scanning an image plate was carried out with BAS1000 (Fuji).

RNA preparation and Northern blot analysis

Total RNAs were extracted from exponentially growing *E. coli* cells (OD_{600}=0.3) or stationary-phase cells (OD_{600}=1.2) by the hot phenol method, precipitated with ethanol, and dissolved in RNase-free H_2O. After digestion with RNase-free DNase I (Takara), RNA was reextracted, precipitated with ethanol, dissolved in RNase-free water, and stored at −80°C until use. The purity of RNA was checked by electrophoresis on 2% agarose gel in the presence of formaldehyde followed by staining with EtBr. The DIG-labeled DNA fragments were amplified by PCR using W3110 genomic DNA (50 ng) as template, DIG-11-dUTP (Roche) and dNTP as substrates, gene-specific forward and reverse primers (Table 1), and Ex Taq DNA polymerase (Takara). Four μg of total RNAs were incubated in formaldehyde-MOPS (morpholinepropanesulfonic acid) gel-loading buffer for 10 min at 65 °C for denaturation, subjected to electrophoresis on formaldehyde-containing 2% agarose gel, and then transferred to nylon membrane (Roche). Hybridization was performed with DIG

easy Hyb system (Roche) at 50 ^{O}C over night with DIG labeled probe. For detection the DIG labeled probe, the membranes were treated with anti-DIG-AP Fab fragments and CDP-Star (Roche), and the images were scanned with LAS-1000 (Fuji Film). The products size on the membrane was estimated on the basis of migration of RNA markers (Toyobo).

ChIP-chip analysis

Bacterial cells were treated with formaldehyde, harvested, lysed and their nucleoprotein was extracted as described by Grainger *et al.* (2007). Immunoprecipitation was then performed using rabbit polyclonal antibodies against RutR. Immunoprecipitated DNA samples and total cell nucleoprotein samples were purified and labeled with Cy5 or Cy3 respectively, without amplification, as described by Grainger *et al.* (2007). Microarrays (Oxford Gene Technology) were designed and produced specifically to analyse DNA obtained from ChIP experiments with *E.coli* MG1655 and its derivatives (Grainger *et al.* 2005). Labelled DNA obtained from immunoprecipitations was

hybridized to the microarray as described previously (Grainger *et al.* 2005). Arrays were then scanned and probes with low Cy5 and Cy3 values and isolated probes with a high Cy5/Cy3 intensity ratio were removed.

ChIP-chip data analysis

The average Cy5/Cy3 intensity ratio calculated for each microarray spot was plotted against the corresponding position on the E.coli MG1655 chromosome, creating a profile of RutR binding (Fig. 10A). We then searched the profile for `peaks', formed by two or more consecutive probes, with a Cy5/Cy3 ratio clearly distinguishable from the background signal. A cut-off, both Cy5/Cy3 ratio greater than 2.5 in without uracil condition and (Cy5/Cy3 ratio of without uracil)/ (Cy5/Cy3 ratio of with uracil) ratio greater than 2-fold was set, and all probes that had an intensity ratio greater than this value were selected as RutR targets. When several adjacent probes (i.e. probes forming one peak) passed the cut-off, the target position was defined as the centre of the probe with the highest Cy5/Cy3 (listed in Table 3).

Table 1. Primers used for preparation of DNA probes for Gel-shift assay and Northern blot.

Primer name	Sequence 5' to 3'
1.Primers used in genomic SELEX study.	
carA-F	CCATCGTTCCGGCGCAAACTT
carA-R	TTACTTAGCGGTTTTACGGTACTG
rutR-F	TGGCGGGTGTTTCAAAAACCAATC
rutR-R	TTAACGTGGTCGAATCCCCTCAA
gadA-F	ATGGACCAGAAGCTGTTAACGGA
gadA-R	TGCCAGCAGATTTGTACCGGA
gadX-F	CAGAAATGCTACGTAAAAGAGCATTAAT
gadX-R	CTATAATCTTATTCCTTCCGCAGAAC
gadB-F	ATGGATAAGAAGCAAGTAACGGATTTAA
gadB-R	TGCCAGCAGATTTGTACCGGA
gadC-F	TGGTGATTACGTCTATCGCGTTG
gadC-R	TTAGTGTTTCTTGTCATTCATCACAATATAGT
gcl-F	CGCTAGGGGTTTGTGCCG
gcl-R	TTATTCATAGTGCATGAAGCAGGTTTC
glxR-F	GTGAAGGGACGTTGTCGATTATG
glxR-R	TCAGGCCAGTTTATGGTTAGCCA
ygiF-F	ATGGCTCAGGAAATCGAATTAAAGTTTATT
ygiF-R	TGTCGCCGCTAAGCAGTTCC
glnE-F	CGTGGGAACATCAGGCGCT
glnE-R	TCATTCTTCCACCAGCCACTTCT

2.Primers used in ChIP-chip study.

(A) Gel Shift Assay

carA-up	GGCTGCGAATTCCTTTTTGATATGCGAGATGTACTT
carA-down	CGCCCGAAGCTTCAAAACACCCTCCAGAGAATA
mraY-up	GGCTGCGAATTCGCAACGCTTGATGCTATCAAA
mraY-down	CGCCCGAAGCTTCATGTCCCATTCTCCTGTAAA
gcd-up	GGCTGCGAATTCCATCACCGTTAATGGTCAGAAA
gcd-down	CGCCCGAAGCTTTTGAGCGTGACACCATAC
yahA-up	GGCTGCGAATTCTGCTGAATGGATTCAGTCTTAAT
yahA-down	CGCCCGAAGCTTCATGAACACACCTTTATCTTTTATC
glxR-up	GGCTGCGAATTCGTATTGCCTGTATTCCTGG
glxR-down	CGCCCGAAGCTTCATAATTAACCTCTTTTAAATTTCGC
fepB-up	GGCTGCGAATTCGACGCTGGTGGAAACAATA
fepB-down	CGCCCGAAGCTTCATATCATCCTCCACAAAATGATA
fabA-up	GGCTGCGAATTCGTTTAAATGAGAAAATCGAAGGC
fabA-down	CGCCCGAAGCTTACCACCCGTTTCGGTCATTTT
rutA/rutR-up	GGCTGCGAATTCCGAAATCTTCACGAAACGCT
rutA/rutR-down	CGCCCGAAGCTTCATATCCTTTTCAGCCGC
ychN/ychO-up	GGCTGCGAATTCAGGACTCTACGTGCTGATTAA
ychN/ychO-down	CGCCCGAAGCTTCATTATTTACTCCTGTATTCAGG
feaB-up	GGCTGCGAATTCTCACATCCTCATGTTGCGAA
feaB-down	CGCCCGAAGCTTCATATATTCCGTGTCGTTTGCCA
ydhL/ydhM-up	GGCTGCGAATTCGGCGCAATAGTAATACGCT
ydhL/ydhM-down	CGCCCGAAGCTTCATCCCGGTGAATCCA

ves-up	GGCTGCGAATTCTGAATCTGTGGCGAAATGCT
ves-down	CGCCCGAAGCTTAGCGAGCCAGTTGATTTCA
yoaA/yoaB-up	GGCTGCGAATTCTGACTTATTTTGCTTCGATAAAGC
yoaA/yoaB-down	CGCCCGAAGCTTAAAGCCTGGTATCGCTTTC
pmrD-up	GGCTGCGAATTCGTTTCACTGGTTAATGACGAA
pmrD-down	CGCCCGAAGCTTAAGTACATGTCTGTTATCTTGTTT
mntH-up	GGCTGCGAATTCATAATCCTGGTCTATCAGAGA
mntH-down	CGCCCGAAGCTTCAGATTTTTACCGGTGGCAAT
yfiQ-up	GGCTGCGAATTCCTTTAACGGACCGGTACT
yfiQ-down	CGCCCGAAGCTTAGATAGTCAAGCAATTCATCAAC
glnE-up	GGCTGCGAATTCAGCACTCACTTACGTGATCT
glnE-down	CGCCCGAAGCTTCATAAGCGATTTTATCCTTGC
rpoD-up	GGCTGCGAATTCTTCTGAGCTTTCCCAGGAA
rpoD-down	CGCCCGAAGCTTTTTCTTCTTCAATCTGCTGCA
ebgC-up	GGCTGCGAATTCGTTTATCTGGTCGGAAAACAC
ebgC-down	CGCCCGAAGCTTAGTTGATATTGTGCTGCTTCAT
yhhX-up	GGCTGCGAATTCCTGCCACATATGATCGGAA
yhhX-down	CGCCCGAAGCTTAAGTACATACGGCAGATGG
yigB-up	GGCTGCGAATTCATTTACACACCGATGTAGAAC
yigB-down	CGCCCGAAGCTTACACTTCACCAGACTCCA
ulaC-up	GGCTGCGAATTCATCGAAATCTTTGGTTGCGT
ulaC-down	CGCCCGAAGCTTCATGGCGCGTCCTTACTT
yjiZ-up	GGCTGCGAATTCGCTTATCTCAATGCAGGTAG
yjiZ-down	CGCCCGAAGCTTCATATGAAGCTGGCAAAGTTCTG

fhuF-up	GGCTGCGAATTCCACATCTAATGCCTTTTCCT
fhuF-down	CGCCCGAAGCTTCATGATTTCATCTCTTTCATTGATAA

(B) Northern blot

carA-F	CCATCGTTCCGGCGCAAACTT
carA-R	TTACTTAGCGGTTTTACGGTACTG
*gcd*5-F	ATGGCAATTAACAATACAGGCTCG
*gcd*5-R	TCCTGATCGGCTACGGGG
*gcd*3-F	ATGCCAAAGGCACGGGTACG
*gcd*3-R	TTACTTCACATCATCCGGCAGC
*fabA*full-F	ATGGTAGATAAACGCGAATCCTATACA
*fabA*full-R	TCAGAAGGCAGACGTATCCTG
*ebgA*5-F	ATGAATCGCTGGGAAAACATTCAGC
*ebgA*5-R	TTAACCATCGCGCTGATGTCAAAC
*ebgA*3-F	AGGTGGCGTATTACGGTCGTG
*ebgA*3-R	TCATTGCTTATTCTCGCTGTGCAAATT
*ychO*3-F	AAAATAACCTCGGGCTGAATCTTAATTA
*ychO*3-R	AAACGACGAACTGCGCTGGG
*ychN*full-F	ATGCAAAAAATCGTGATCGTTGCCAAT
*ychN*full-R	TTAAAATGTGAGCACTTTATCGGCTGA
*ydhL*full-F	GTGGCGGAGCAATTAGAGTTCTT
*ydhL*full-R	TCAAAAGAGTGATGGTTGCTCCG
*ydhM*full-F	ATGGGGCTAAGCGAATTACTAAAAAC
*ydhM*full-R	CTAAACGGCAGGCGTCGCAr

*ves*full-F	ATGGAATACTTTGATATGCGTAAAATGTC
*ves*full-R	TTACTGAACTTGATCCGGCGAATG
*pmrD*full-F	ATGGAATGGCTGGTCAAAAAATCGTG
*pmrD*full-R	TTACTGAGTTTTCCCTGCCACTTTAC
*yfiP*3-F	CCAGTAATACCGGGCGTCTC
*yfiP*3-R	TTAAACGCTTTCTAACTGTTCTGCTGTr
*yfiQ*3-F	AAAACGGTGAACGCTGCTTGTTC
*yfiQ*3-R	TCATGATTCCTCGCGCTGGG
*yhhX*3-F	TGGACCAGATTATTTCTCTGTTCGG
*yhhX*3-R	TTACTTAGCGAGAGTTACTGTGGAG
*yhhY*full-F	ATGAGTGAGATAGTAATACGCCACG
*yhhY*full-R	TTACTTCACCCGCGCCATATAATATG
*xerC*3-F	TACGCGACCGTGCAATGCTG
*xerC*3-R	TTATTTCCCCCGTTTGGCGCG
*dapF*3-F	CCGATGATGATCTGGTCCGC
*dapF*3-R	TCATAGATGAATAAATCCGTCGTAGAC
*mraY*3-F	ATGGTCTCGACGGCCTGG
*mraY*3-R	TTAACGTACCTTCAGCGTTGCCA
*mraZ*full-F	ATGTTCCGGGGAGCAACGTTA
*mraZ*full-R	TTATAGAGACAAGTCCTGCAGTCG
*feaB*5-F	ATGACAGAGCCGCATGTAGCAG
*feaB*5-R	ACTACGCCAACCGGCTCTTTA
*feaB*3-F	AAGTGGGACCGGGGATGTC
*feaB*3-R	TTAATACCGTACACACACCGACTTAG

*feaR*3-F	TCTGCGCAGAAAGACTGGACG
*feaR*3-R	TTAGCGGAATTTACGTCGATACTCG
*yoaA*3-F	GTACCTCGCACGCCATGATG
*yoaA*3-R	TTACCTGGAGGATGGTATCGCAA
*yoaB*3-F	ATGACTATCGTTCGTATCGATGCTG
*yoaB*3-R	TTACACCGCAGCCACAATCTTAATTT
*yjjL*5-F	GTGGAAAAAGAAAATATCACCATCGATC
*yjjL*5-R	GGGCGTCCGCGCTCTTTG
*yjjL*3-F	ATAACCTGGATTTAAAAAGCACAGGGTT
*yjjL*3-R	TTAATCTTTACGTGGGTCGTTGATCG

RESULTS

Isolation of RutR-binding sequences by genomic SELEX

To get insights into the entire network of transcription regulation of the *E. coli* genome by transcription factor RutR, we employed the genomic SELEX system for isolation of the recognition DNA sequence(s) by purified RutR. In the original SELEX procedure, mixtures of synthetic oligonucleotides with all possible sequences were used as substrates (Ellington and Szostak, 1990; Tuerk and Gold, 1990), but in the improved SELEX method, a mixture of *E. coli* DNA fragments of 200-300 bp in length was used as substrates (Shimada *et al.*, 2005; Ogasawara *et al.*, 2007a; 2007b). For the screening of RutR-binding sequences, we repeated SELEX three cycles until several sharp bands were detected on PAGE (data not shown). After sequencing of SELEX fragments, a total of 68 independent sequences were identified, which could be located at 9 (or 10) regions of the *E. coli* genome (Table 2). The number of independent isolates was the highest (35 clones) for the sequence within the spacer region

upstream of the *carAB* operon. Since the binding site of a transcription factor is generally located upstream of the protein-coding sequence, we predicted that RutR is involved in regulation of the *carAB* operon encoding carbamylphosphate synthetase. Among 10 RutR-binding sequences herein isolated, we assumed that the affinity should be the highest for the *carAB* promoter because the number of independent clones was the largest.

A total of 9 independent clones was isolated, which all included a 241 bp-long sequence within the 3'-terminal 479 bp-long coding sequence of *gadA* or *gadB*, both encoding the glutamate decarboxylase (see Table 2) [Note that this 479-bp sequence is identical between *gadA* and *gadB*]. If the SELEX sequence is from the *gadA* locus, RutR might regulate the downstream *gadX* gene, which encodes the global regulator of at least 15 genes including *gadAXW* and *gadBC* involved in acid stress resistance. On the other hand, if the SELEX fragment is from the *gadB* locus, RutR should regulate the downstream *gadC* encoding glutamate-gamma aminobutyrate (GABA) antiporter. [This SELEX segment is hereafter described as *gadA/gadB*]. After DNase-I footprinting assay using

non-overlapping DNA probes, we found that the 3'-terminal proximal regions of both *gadA* and *gadB* carry the RutR-binding sequence (see below).

A total of 6 independent SELEX fragments were isolated, which are located on the 3'-terminal coding region of *ygiF* (predicted adenylate cyclase) and upstream of *glnE* encoding GlnA adenylyltransferase (see Table 2). Thus, RutR appears to regulate the *glnE* gene that is involved in the activity control of GlnE glutamine synthase through reversible adenylylation. A total of 6 independent clones were also isolated, which included the coding sequence of *hyi* (hydroxypyruvate isomerase) (see Table 2). This RutR-binding site is located upstream of *glxR* (tartronate semialdehyde reductase or 2-hydroxy-3-oxopropionate reductase) for degradation of purines, implying that *glxR* is under the control of RutR. Likewise, a total of 6 independent clones were isolated, which included the coding sequence of *yciR* that is located upstream of *deoL* (leader peptide) and *deoT* encoding a global regulator for the genes involved in a variety of metabolic pathways for maltose transport, fatty acid oxidation and peptide degradation (Elgrably-Weiss *et al.*, 2006). In good

agreement with the prediction by Loh *et al.* (2006), three SELEX fragments were isolated, which included the 3'-terminal proximal sequence of *rutA* that forms the *rutABCDEFG* operon for pyrimidine degradation (see Table 2). This RutR sequence may be involved in regulation of *rutABCDEFG* operon or *rutR* itself [this SELEX segment is tentatively described as *rutA*/*rutR*].

Table 2. RutR-binding DNA fragements isolated by genomic SELEX

No. clones	Left gene	SELEX fragment	Right gene	Size (bp)
35	*dapB* (-->)	S	(-->) ***carAB***	116
9	***gadWX*** (<--)	S (*gadA*)(<--)	(<--) *yhjA*	241
	gadC (<--)	S (*gadB*)(<--)	(<--) *pqqL*	241
6	***glnE*** (<--)	S (*ygiF*)(<--)	(-->) *ygiM*	233
6	*gcl* (-->)	S (*hyi*)(-->)	(-->) ***glxR***	216
6	***deoLT*** (<--)	S (*yciR*)(<--)	(<--) *rrb*	224
3	***rutA*** (<--)	S	(-->)***rutR***	168

A total of 68 DNA fragments have been isolated by the genomic SELEX screening. After sequencing, these fragments have been located at 9 (or 10) regions of on the *E. coli* genome. The list shows the SELEX fragments that were isolated more than 3 times independently. "S" indicates the location of each SELEX fragment while the direction of gene organization is shown by arrows. Bold indicates the putative regulation targets. Except for the SELEX fragment from the spacer region between *dapB* and *carAB*, all other SELEX fragments were located within the respective coding frames. The size indicates the length (bp) of DNA probes used in the gel shift assay (see Fig.1)

RutR binding in vitro *to SELEX DNA fragments*

To confirm RutR binding to the target sequences identified by genomic SELEX, we carried out DNA mobility shift assays. Six species of DNA probe of 116 to 241 bp in length were selected from the SELEX fragments (see Table 2), and PCR-amplified using FITC-labeled primers. Using the fluorescent-labeled probes, the gel shift assay was performed using increasing concentrations of the RutR protein. Results, summarized in Fig. 1A, indicated: (1) all six probes (*dapB-carA, rutA-rutR, ygiF, gadA/gadB, hyi* and *yciR*) formed RutR complexes; and (2) among these six probes, the RutR-binding affinity was higher for two intergenic spacer sequences, *dapB-carA* and *rutA-rutR*, than other four DNA probes within coding frames because the RutR complexes with the former probes were formed at lower protein concentrations. This finding indicates that the number of SELEX fragments isolated (see Table 2) correlates their RutR-binding affinity. Such a correlation has been observed in the genetic SELEX search with other transcription factors (Shimada *et al.*, 2005; Ogasawara *et al.*, 2007a; 2007b).

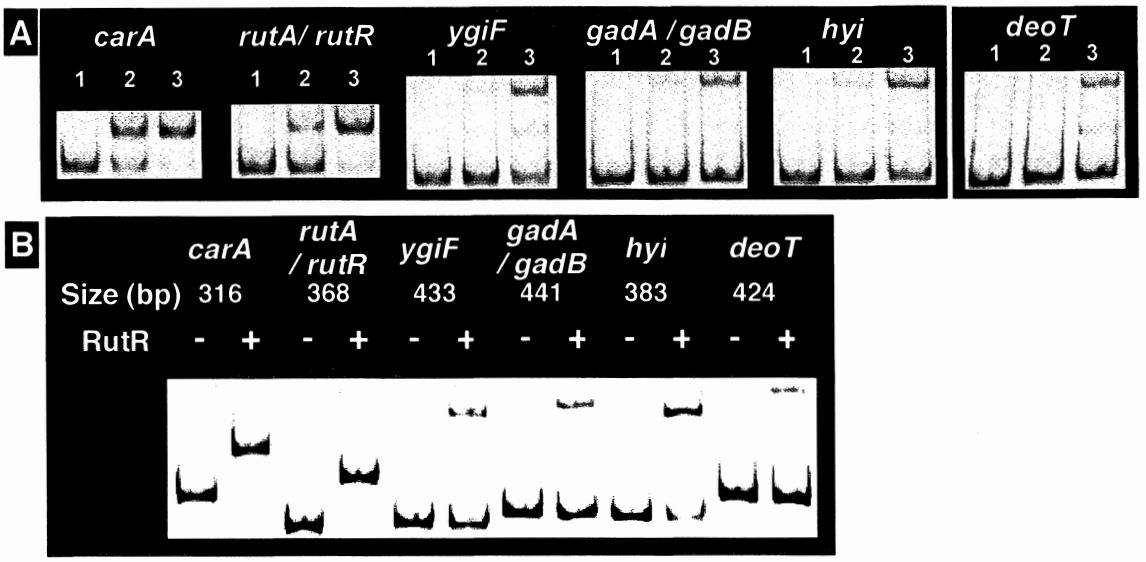

Fig. 1. Gel shift assay of RutR-SELEX fragment interaction. [A] FITC-labeled SELEX fragments were PCR-amplified as described in Experimental procedures. Mixture of 0.5 pM each of FITC-labeled DNA probes and three different concentrations of RutR (lane 1, 0; lane 2, 1 pM: lane 3, 2.5 pM) were incubated at 37℃ for 15 min and separated by gel electrophoresis on 5% polyacrylamide gel. [B] Six SELEX fragments were incubated with or without 2.5 pM RutR and separated on a 5% polyacrylamide gel. Size (bp) indicates the length of probes used for the gel shift assay. [Note that the probe sizes are longer than those of SELEX fragments, shown in Table 2, because of the attachment of PCR primer sequences that were used for probe amplification].

Fig.1B shows the migration pattern of RutR complexes for all six-test probes performed under the identical conditions. Compared with the probe size, the mobility of RutR complexes were faster for the *dapB-carA* and *rutA-rutR* spacer probes than other four coding frame probes. One possibility of this irregular mobility is that two molecules of RutR bind to these spacer probes, but this is unlikely because the DNase-I foot-printing assay indicated only a single protection site for these probes (see below). These two probes may have an intrinsic or RutR-induced curvature since the DNA curvature influences on its mobility on gel (Olivares-Zavaeleta *et al.*, 2006). The abnormal mobility might also be due to the difference in the position of RutR binding along the respective DNA probes. The results of gel shift assays altogether indicate that the RutR-binding sites can be classified into two groups: group-1 binding sites, located on the *dapB-carA* and *rutA-rutR* intergenic spacer sequences, have higher affinity to RutR and their complexes with RutR migrate faster on PAGE; group-2 binding sites, including *ygiF*, *gadA/gadB*, *hyi* and *yciR* coding sequences, have lower affinity to RutR, and their RutR complexes migrate

slower than group-1 complexes.

RutR box: a consensus palindromic sequence for RutR binding

Next we determined the RutR recognition sequences using DNase-I footprinting assay. Protection sites from DNase-I digestion were analyzed in the presence of RutR or RNA polymerase σ^{70} holoenzyme alone and in the simultaneous presence of both proteins. The sequence including in the *gadA/gadB* SELEX fragment was identical between *gadA* and *gadB*. For identification of the origin of this SELEX fragment, we performed the DNase-I protection assay separately for *gadA* and *gadB* using longer probes including the respective unique coding sequences. The DNase-I protection regions by RutR on the *dapB-carA, rutA-rutR, gadA, gadB, ygiF* and *hyi* ranged from 26 to 31 bp in length (Fig. 2, RutR lanes). With the *yciR* fragment, however, no clear protection site was detected (data not shown).

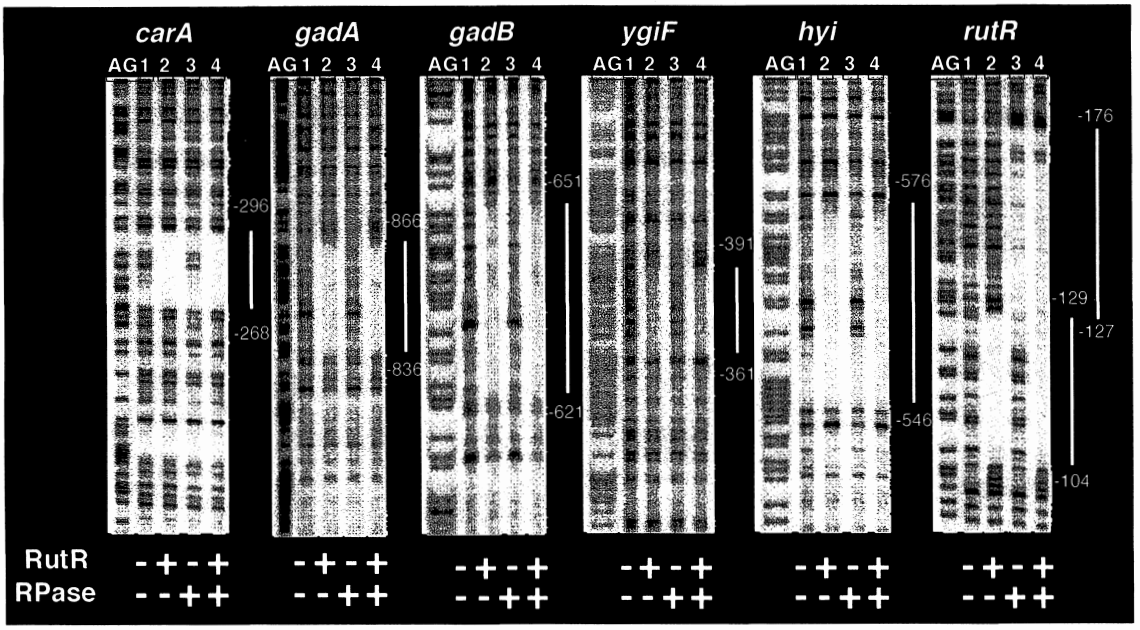

Fig. 2. DNase-I footprinting assay for determination of RutR-binding site. Each of 0.1 pM DNA ^{32}P–labeled probes was incubated in the presence or absence of RutR and/or RNA polymerase at 37°C for 15 min and then subjected to digestion with 5 ng DNase I. Mixtures were immediately subjected to separation by 6% urea polyacrylamide gel. Lane 1, no protein; lane 2, 5 pM RutR; lane 3, 5 pM RNA polymerase σ^{70} holoenzyme; lane 4, 5 pM RutR plus 5 pM RNA polymerase. AG shows AG ladder of probe DNA. Bars on right indicate the protected region from DNase-I digestion. Although the footprinting experiments were performed using the RutR-binding sequences, the numbers on right indicate the distance from the respective initiation codon of the predicted target genes (see Table 2). RNA polymerase-binding site, shown by dotted line, was detected only for *rutA/rutR* probe between -104 and -129 from the *rutR* initiation codon.

Among the six RutR-binding sequences, a palindromic consensus 7-2-7 sequence consisting of TTGACCAnnTGGTCAA was identified (Fig. 3A) [this RutR recognition sequence was hereafter designated as RutR box]. The 5'-proximal TTGACCC sequence is better conserved (7/7 for *carA*, *gadA* and *gadB*, 6/7 for *rutA* and *ygiF*, and 4/7 for *hyi*), suggesting that this is the core sequence for RutR recognition. The RutR box on the *gadA* and *gadB* sequences consists of 7-3-7 sequence with one extra base in the center. In concert with the failure of RutR complex formation by gel shift assay, the RutR box sequence could not be found in the *yciR* fragment.

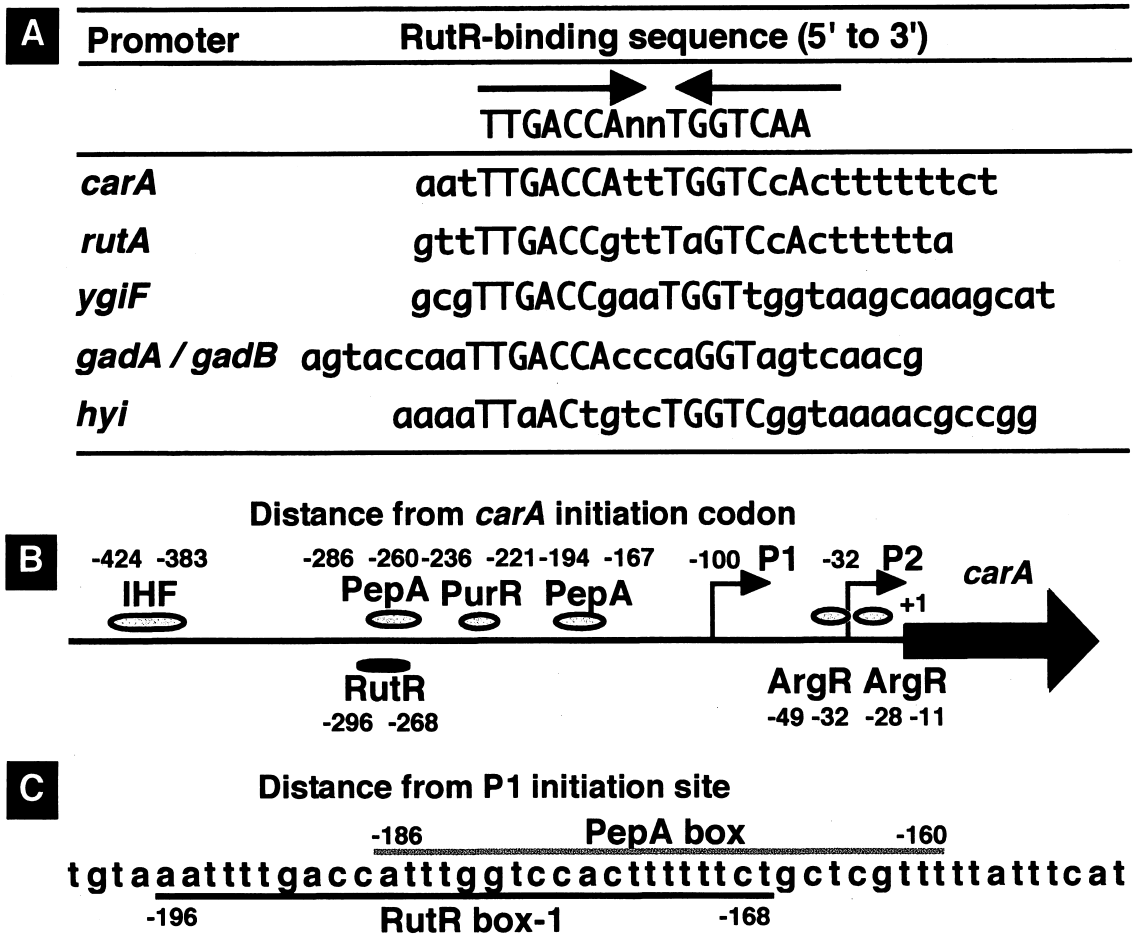

Fig. 3. RutR box sequences recognized *in vitro* by RutR protein. [A] RutR recognizes a palindromic sequence consisting of the consensus sequence TTGACCAnnTGGTCAA. [B] Location of the binding sites for RutR and other transcription factors on the *carA* promoter. Numbers indicate the distance from the *carA* initiation codon. [C] The sequence including the overlapping binding sites for RutR and PepA. The numbers represent the distance from transcription initiation site of *carA* P1.

In the presence of RNA polymerase σ^{70} holoenzyme alone, the promoter region protected by the RNA polymerase could be identified only for the *rutA-rutR* spacer probe, indicating that among the test probes, the RNA polymerase-binding sequence is included only for this probe (Fig. 2, RNA polymerase lanes). The binding sites of RutR on other probes are separated from the RNA polymerase-binding sites (or the target promoters). In the case of *rutA-rutR* probe, the RNA polymerase ($E\sigma^{70}$) bound to a site next to the RutR binding site, in agreement with the location of promoter sequence, -10 TAAAAT and -35 TTAATC, in this region. The *rutABCDEFG* (pyrimidine utilization) operon is transcribed by RNA polymerase containing RpoN sigma ($E\sigma^{54}$) and is not expressed under nitrogen-rich conditions (Loh *et al.,* 2006). Thus, the binding site of RNA polymerase RpoD holoenzyme ($E\sigma^{70}$) could be the promoter for the *rutR* gene, and thus the RutR binding to this site might represent the autogenous repression of the *rutR* gene. In the presence of both RutR and RNA polymerase, the RutR box and the *rutR* promoter were both protected from DNase-I digestion (Fig. 2, *rutR* panel), indicating that both proteins could bind side by side without

interference. In the autogenous repression of *rutR* transcription, RutR binds downstream of the RNA polymerase-binding site, implying that RutR plays as a road-blocker to inhibit the migration of RNA polymerase along the template.

Uracil- and thymine-induced dissociation of RutR from RutR boxes

In *E. coli*, carbamoyl phosphate is a common precursor of pyrimidines and arginine, and thus the expression of *carAB* might be under a complex and feedback regulation by various metabolites on the pathways for pyrimidine and arginine synthesis. Transcription of the *carAB* operon is under the control of two promoters (Bouvier *et al.*, 1984; Piette *et al.*, 1984): upstream P1 is controlled by pyrimidine while downstream P2 is controlled by purines and arginine. ArgR and PurR are responsible for the regulation by purines and arginine, but the sensor for pyrimidines and the corresponding regulator has not yet been identified. Here the RutR binding site was identified at -168 to -196 bp from P1 initiation site (see Fig. 1 and 2). The RutR-binding site overlaps with PepA (CarP)-binding site at -160 to -186 bp (Fig. 3B), raising a hypothesis that RutR is involved in *carA* P1

activation by competing with the binding of PepA (CarP) repressor.

If RutR is involved in the regulation of *carA* P1 promoter, it could be the as yet unidentified sensor for pyrimidines. To test this possibility, we performed the gel shift assay of RutR-*carA* P1 complex formation in the presence of pyrimidine bases, nucleosides or nucleotides. Fig. 4 shows the effect of 100 μM each of free pyrimidines and purines, their nucleosides and nucleotides on the RutR complex formation with three probes, *dapB-carA*, *rutA-rutR* and *ygiF*. The gel pattern clearly shows that both uracil and thymine abolished the formation of RutR box DNA-RutR protein complexes, but little inhibition was observed with cytosine, adenine, and guanine. The inhibitory activity of RutR binding to the respective target DNA is higher for uracil than thymine. The inhibitory activities of uracil nucleoside and nucleotides are lower than uracil base. Low-affinity probe *ygiF* is more sensitive to pyrimidines than high-affinity promoters, *dapB-carA* and *rutA-rutR*. Taken together we concluded that RutR is the sensor of pyrimidines, uracil and thymine.

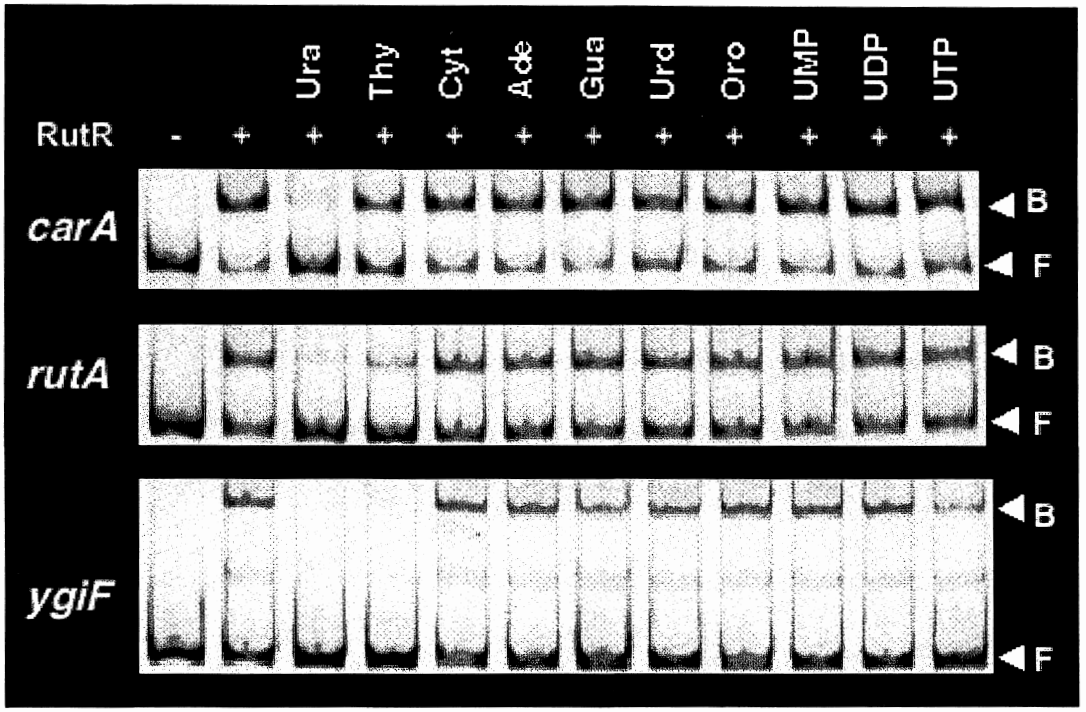

Fig. 4. Influence of bases, nucleosides and nucleotides on DNA-binding activity of RutR. The *carA*, *rutA* and *glnE* promoter fragments were mixed with RutR protein and incubated at 37℃ for 15 min in the presence of 100 μM each of the indicated base, nucleoside or nucleotide. After 15 min incubation, the reaction mixtures were subjected to 5% PAGE. Test components were: Ura, uracil; Thy, thymine; Cyt, cytosine; Ade, adenine; Gua, guanine; Urd, uridine; Oro, orotate; UMP. Uridine 5'-monophosphate; UDP, uridine 5'-diphosphate; UTP, uridine 5'-triphosphate. "B" indicates RutR-bound DNA while "F" indicates free DNA.

RutR and RNA polymerase are able to bind to the *rutR* promoter region simultaneously side by side as detected by DNase-I protection (see Fig. 2, *rutR* panel). Effect of uracil on the preformed RutR-RNA polymerase-DNA complex was then examined by gel shift assay. Under the conditions employed, a small amount of the ternary complex was identified besides the binary complexes, *i.e.*, RutR-DNA and RNA polymerase-DNA complexes (Fig. 5) [note that RNA polymerase forms a dimmer and oligomers, leading to form slowly migrating complexes]. By the addition of uracil, the level of both RutR-DNA and RutR-RNA polymerase-DNA complexes decreased (Fig. 5, lanes 3 and 7), indicating that RutR is preferentially dissociated from the ternary promoter complex. After dissociation of RutR from the ternary complex, the RNA polymerase remained associated with the promoter DNA. As a result, the level of RNA polymerase-DNA complexes containing more than two RNA polymerase molecules increased.

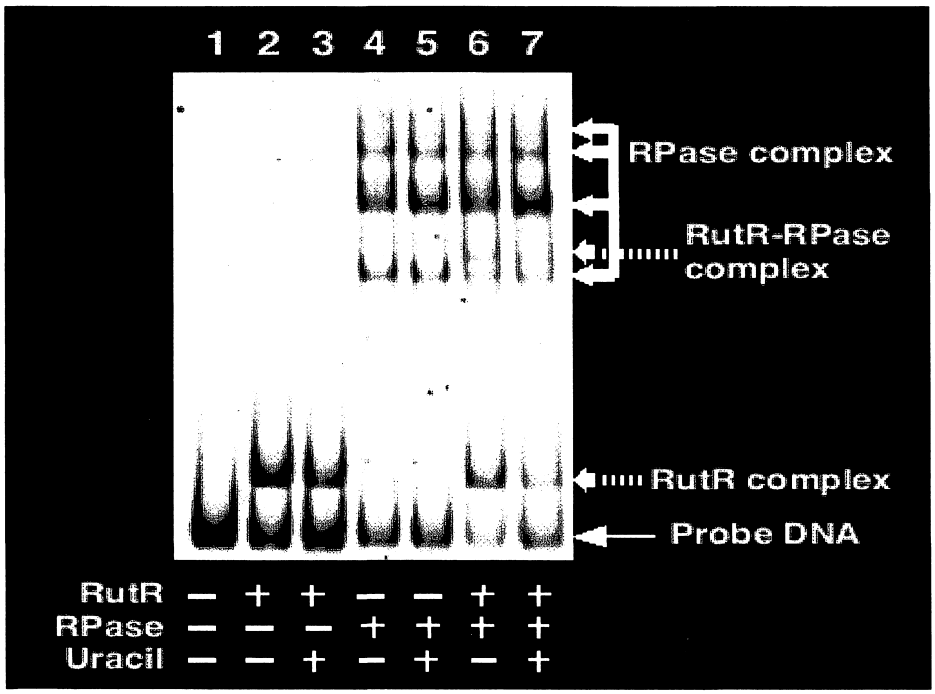

Fig. 5. Effect of uracil on RutR-RNA polymerase-DNA complex formation. The fluorescent-labeled *rutA* promoter fragment (0.5 pM) was mixed with RutR protein (1.25 pM), RNA polymerase (1.25 pM) or their combination. After incubation at 37℃ for 15 min, 100 μM uracil was added and the incubation was continued for additional 20 min. The reaction mixture was subjected to 3.5% PAGE.

Activation of carAB transcription in vivo *by RutR*

In order to get insight into the regulation *in vivo* of *carA* promoters, we first examined the *in vivo* activity of *carAB* promoter using the TFP (two fluorescent protein) promoter assay vector (Makinoshima *et al.*, 2002; Shimada *et al.*, 2004). The *carA* promoter fragment of 500 bp in length upstream from the *carA* initiation codon including the target sites for all the transcription factors, PepA, PurR, PepA, RutR and IHF (see Fig. 3B), was inserted into TFP vector so to adjust the initiation codon to that of GFP. In the wild-type *E. coli*, the GFP activity relative to RFP, which is under the control of internal reference promoter *lac*UV5, was strong at the exponential growth phase in M9-glucose medium (Fig. 6). In the mutant with *rutR* deletion, the promoter activity was markedly reduced as expected, but there was still a low level of GFP expression. The low level activity might be due to P1 activation by the decreased reiterative transcription of U clusters at P1 initiation site (Han and Turnbough, 1998) due to the decrease in UTP level or a change in the intracellular level of other four transcription factors, IHF, PepA (CarP), PurR and ArgR (see Fig. 3B). In the presence of

excess uracil, the promoter activity markedly decreased in agreement with the *in vitro* results that RutR binding to P1 is inhibited in the presence of uracil (see Fig. 4 and 5). The reduction of P1 activity might also be due to the increase in UTP-sensitive reiterative transcription (Han and Turnbrough, 1998). Taken both *in vitro* and *in vivo* observations together, we concluded that RutR activates transcription from *carA* P1 and pyrimidines are the effectors controlling RutR activity.

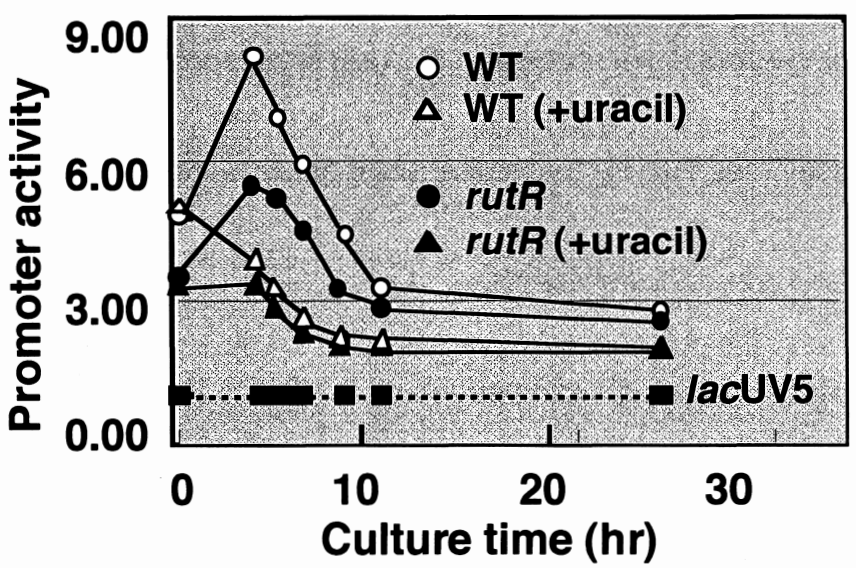

Fig. 6. *In vivo* assay of the *carA* promoter activity. Wild-type and *rutR*-lacking mutant *E. coli* were transformed with the *carA* promoter assay vector. Transformants were grown in M9-0.4% glucose medium and the promoter activity was determined by measuring the GFP/RFP ratio at 0, 4, 5, 6, 8, 10, 24 hr. For details see Experimental procedures.

In order to confirm the results of promoter assay, Northern blot analysis and S1 mapping were performed for direct measurement of RNA transcripts from the *carAB* operon. Both wild-type and the *rutR* mutant were grown in M9-glucose medium and RNA samples were prepared at exponential growth phase. By Northern blot analysis using a *carA* probe, one major transcript of about 4.0 kb in length was detected, which corresponds to the size of full-length *carAB* operon (Fig. 8). The level of *carAB* transcription in wild-type *E. coli* was abolished by the addition of uracil or in the *rutR* mutant, confirming the results of promoter assay (see Fig. 5). By Northern blot analysis, however, it was difficult for distinguish P1 and P2 transcripts.

Next S1 mapping was performed to detect P1 and P2 transcripts separately. In wild-type *E. coli*, the level of P1 activity was higher than P2 (Fig. 7, lanes 1 and 3). By the addition of uracil, the P1 activity was completely abolished (Fig. 7, lane 4). In the *rutR* mutant, the P1 activity significantly decreased, but the low level activity was retained, in agreement with the promoter assay (see Fig. 6).

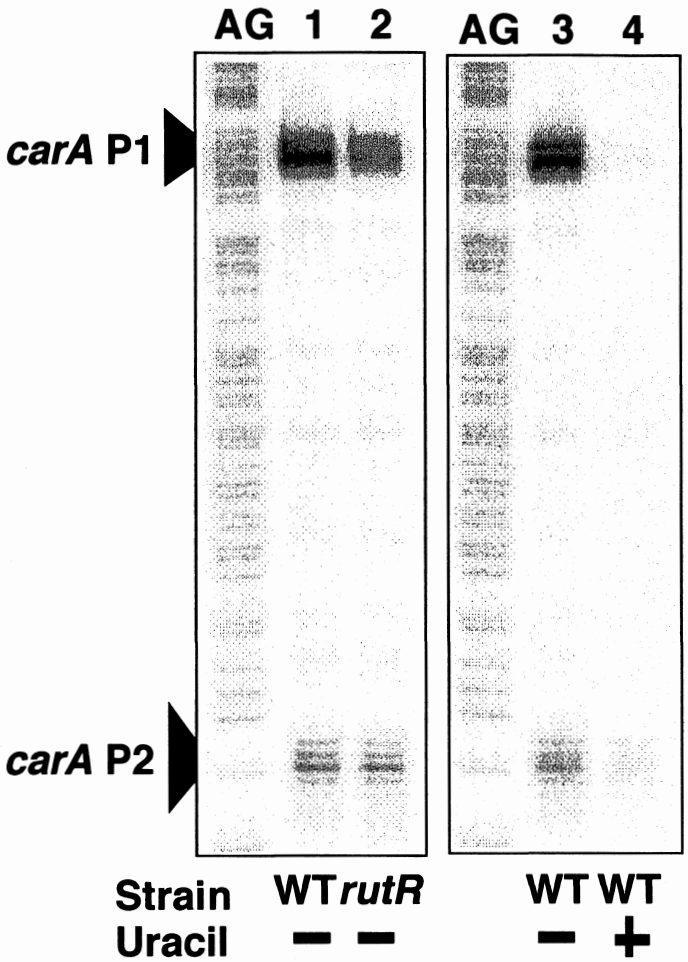

Fig. 7. S1 mapping of *carA* RNAs. *E. coli* wild-type BW25113 and its *rutR* deletion mutant JW0998 were grown in M9-0.4% glucose minimum medium in the absence (lane 1-3) or presence of uracil (lane 4). At the exponential growth phase, total RNA was extracted by the hot-phenol method. S1 nuclease protection assay was carried out for the *carA* promoter region. ^{32}P-labeled probes were prepared by PCR. Lanes AG indicate Maxam-Gilbert AG sequence ladders. Arrows on left indicate the P1 and P2 bands protected against S1 nuclease.

Regulation of glutamate-dependent acid resistance pathway by RutR

RutR-binding sites were identified within the coding frames of *gadA* and *gadB* (see Table 2 and Fig. 1), both encoding glutamate decarboxylase. The *gadA* gene forms an operon with *gadX* and *gadW*, while the *gadB* gene forms an operon with *gadC* encoding the glutamate/4-aminobutyrate (GABA) antiporter. GadX and GadW together participate the regulation of expression of the *gadAXW* and *gadBC* operons (Ma *et al.*, 2002; 2003). The Gad system is involved in glutamate transport and glutamate-dependent acid resistance of *E. coli* for survival under acidic conditions (Richard and Foster, 2003).

To identify the effect of RutR on these two *gad* operons, we performed Northern blot assay. In the stationary-phase wild-type *E. coli*, two bicistronic mRNAs, 3.0-kb *gadBC* RNA (Fig. 8C) and 2.7-kb *gadAX* RNA (Fig. 8D), were detected using a single and the same probe (*gadA/B* probe) with a sequence shared between *gadA* and *gadB*. The bicistronic nature of 3.0-kb *gadBC* RNA was proved because this band was also detected with use of a *gadC* probe (Fig. 8E). A smaller 1.4-kb RNA band was detected with the *gadA/B* probe (Fig. 8F)

but not with the *gadC/gadX* probes (data not shown), indicating that this RNA band contains monosictronic *gadA* and *gadB* RNAs. The sizes of these *gad* transcripts agreed well with the reported values (De Biase *et al.*, 1999; Ma *et al.*, 2002).

In the absence of RutR, the level of all these transcripts markedly increased, indicating that RutR bound on the coding frames of both *gadA* and *gadB* somehow represses transcription of the *gadAXW* and *gadBC* transcription. Accordingly transcription from these promoters was derepressed in the presence of a high concentration of uracil (Fig. 8C-8F). In exponential growth phase, however, none of these transcripts were detected (data not shown), in agreement with the finding that expression of the *gad* system depends on RpoS sigma (Castanie-Cornet *et al.*, 1999).

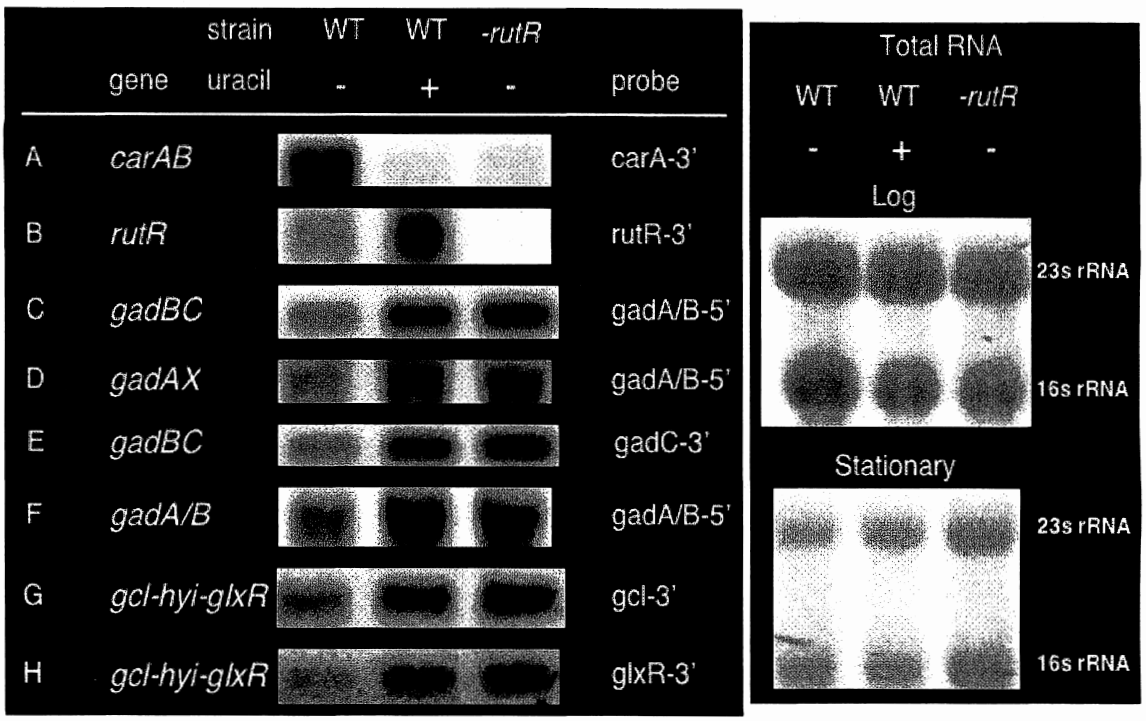

Fig. 8. Northern blotting of transcripts of the genes under the control of RutR. [Left side panel] Wild-type *E. coli* BW25113 was grown in M9-0.4% glucose minimum medium in the absence (column 1) or presence of uracil (column 2), and its *rutR* deletion mutant JW0998 was grown in M9-0.4% glucose minimum medium (column 3). Total RNA was extracted from both exponential-phase and stationary-phase cells and 4 μg each was subjected to Northern blot analysis using the indicated probes. Each probe contains 500-bp sequence starting from 5' end of the respective coding frame. Lanes A-B, *carAB* and *rutR* RNAs from exponential-phase cells; lanes C-H, *gadBC*, *gadAX* and *gcl-hyi-glxR* RNAs from stationary-phase cells. [Right panel] Total RNAs used for the Northern blot analysis.

Regulation of allantoin degradation pathway by RutR

RutR binding site was also identified within the coding sequence of *hyi* encoding hydroxypyruvate isomerase (see Table 2), which is organized in the allantoin utilization gene cluster consisting of 7 genes in the order of *gcl-hyi-glxR-ybbV-ybbW-allB-ybbY-glxK*. In *E. coli*, allantoin, a product of purine degradation, is converted into ureidoglycolate by AllB and AllC, which is then metabolized into two pathways, In the first pathway, ureidoglycolate is converted into 3-phosphoglycerate by AlllA, Gcl, GlxR and GlxK for its integration into the central energy metabolism. In the second pathway, ureidoglycolate is also catabolized into oxalureae by AllD, which is ultimately converted into oxamate and carbamoylphosphate (Cusa *et al.*, 1999). In order to get insight into the regulation of allantoin degradation gene cluster by RutR, we performed Northern blot analysis by using DIG-labeled *gcl* (the first gene in the *gcl* operon) and *glxR* (the third gene) probes. In the stationary phase, two probes detected the same size RNA product of about 3.4-kb (Fig. 8G and 8H), which was considered, based on the size, to correspond to *gcl-hyi-glxR* transcript. In the presence of a

high concentration of uracil or in the *rutR* mutant, the level of this RNA product significantly increased (Fig. 8G and 8H).

The *gcl-hyi-glxR* band was, however, not detected for RNA from exponential phase cells. The *gcl-hyi-glxR* gene cluster is involved in the first pathway of ureidoglycolate catabolism for its conversion into 3-phosphoglycerate. If RutR represses this glycerate pathway, ureidoglycolate may be catabolized through the second pathway leading to produce oxamate and carbamoylphosphate, the substrate of pyrimidine synthesis. We then concluded that RutR is also involved in the regulation of purine degradation downstream from allantoin for its reutilzation, through the supply of carbamoylphosphate, the substrate of pyrimidine synthesis. RutR was thus indicated to be the master regulator for the synthesis of glutamine via Gad system, the conversion of glutamine to carbamoyl phosphate, the synthesis of pyrimidines, the degradation of pyrimidines, and the downstream pathway of purine degradation from as well as the synthesis of arginine.

Finally we analyzed possible influence of RutR on transcription of

ygiF-glnE operon (Van Heeswijk *et al.*, 1993). GlnE encodes adenylating enzyme of glutamine synthetase, which supply the substrate for CarAB enzyme. We performed Northern blot analysis by using both *ygiF* (the first gene in the *ygiF-glnE* operon) and *glnE* (the second gene) probes, but no corresponding RNA product was detected (data not shown), implying that transcription level of the *ygiF-glnE* operon was low under the conditions employed.

Intracellular level of RutR protein

Transcription of the *rutR* gene was indicated to be under the autogenous control (see above). If this is the case, the intracellular concentration of RutR should stay constant under the steady-state of cell growth. We then examined the *rutR* RNA level by Northern blot analysis and the RutR protein level by the quantitative Western blot analysis.

The RutR level relative to that of RNA polymerase α subunit, which stays constant throughout the cell growth phase (Ishihama, 1990; Jishage *et al.*, 1996), stayed constant throughout the growth phase of cell cycle (Fig. 9, lanes 1-5), supporting the autogenous control model of RutR synthesis. By the addition of uracil, however, the synthesis of RutR is enhanced because of inactivation of RutR (Fig. 9, lanes 6-10). Assuming the number of RpoA as 5,000 molecules per genome equivalent of DNA (Ishihama, 1990; Jishage *et al.*, 1996), the number of RutR protein molecules was estimated to range approximately 150 molecules in the absence of uracil and 300 molecules in the presence of uracil. The Northern blotting confirmed that the *rutR* mRNA level increased in the presence of a high

concentration of uracil (Fig. 8B). The level of mRNA increase is apparently higher than that of RutR protein increase.

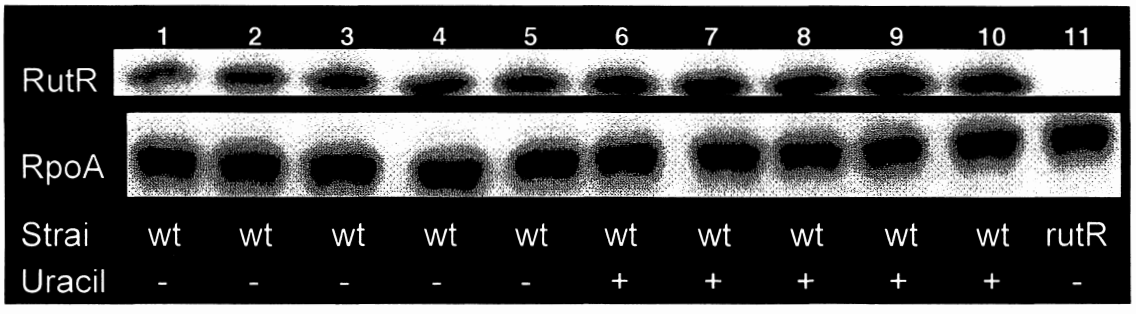

Fig. 9. Measurement of the intracellular concentration of RutR protein. Determination of the RutR level relative to RNA polymerase core subunit RpoA was carried out by quantitative immuno-blotting (Jishage *et al.*, 1996) using antibodies against purified RutR and RpoA. *E. coli* wild-type BW25113 (lane 1-10) and its *rutR* disruptant JW0998 (lane 11) were grown in M9-0.4% glucose medium in the absence (lane 1-5, 11) or presence of uracil (lane 6-10). Samples were taken at various growth phases (A600nm 0.3 for lanes 1, 6 and 11; 0.6 for lanes 2 and 7; 0.9 for lanes 3 and 8; 1.2 for lanes 4 and 9; 1.5 for lanes 5 and 10).

Isolation of genome DNA fragments associated with RutR in log-phase E. coli

To get insight into the RutR regulon *in vivo*, we next analyzed the distribution *in vivo* of RutR along the *E. coli* genome by using ChIP-chip system. Wild-type *E. coli* BW25113 and its *rutR* mutant JW0998 were grown in M9 glucose medium to OD_{600} of 0.3, and the cultures were treated with formaldehyde to make cross-linkage between genome DNA and DNA-bound proteins. The genome DNA was extracted and sonicated to yield DNA fragments of 500-1000 bp in length. Both free and DNA-crosslinked RutR were recovered after immuno-precipitation with RutR antibodies. DNA fragments recovered from the immuno-precipitates was labeled with Cy5 for wild-type DNA and Cy3 for mutant DNA, respectively, mixed and hybridized to an *E. coli* microarray. After washing and scanning of the fluorescent-labeled microarray, the ratio of Cy5/Cy3 signal intensity was calculated. This ChIP-chip experiment was repeated using cells grown in the presence of excess uracil. Fig. 10A shows a genome-wide profile of DNA-bound RutR. A number of peaks of DNA-bound RutR could be detected, which were significantly separated from the background signal.

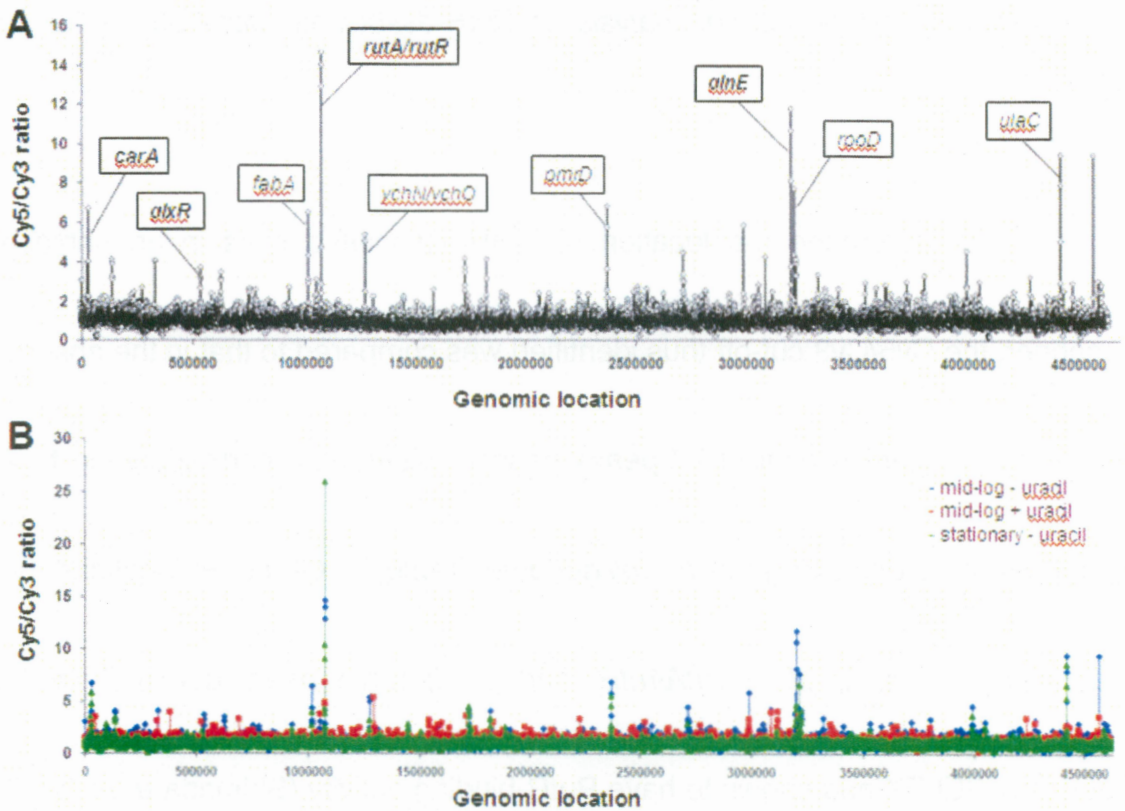

Fig. 10. Distribution of RutR binding across the *E.coli* chromosome. (A) The figure shows an overview of results from ChIP-chip three independent experiments that measure the profile of RutR binding across the *E.coli* chromosome during exponential growth in aerobic conditions. Binding signals (y-axis) are plotted against their location on the 4.64 Mb *E.coil* chromosome (x-axis). The locations of selected signals are labeled in small type (newly identified RutR targets) or in bold face (known RutR targets). A complete list of RutR targets identified is presented in Table 3. (B) The figure shows an overview of results from ChIP-chip under the difference conditions that measure the profile of RutR binding across the *E.coli*. Binding intensity is shown by Cy5/Cy3 signals in mid-log phase conditions (blue), stationary phase conditions (green) and mid-log phase with uracil conditions (red).

Identification and sequence analysis of RutR-associated sites on the E. coli *genome*

To determine the location of RutR-associated sites in an unbiased manner, the Cy5/Cy3 cut-off thus identified was compared to that in the absence of uracil addition. A total of 77 peaks passed this cut-off, corresponding to 24 separate locations along the *E. coli* genome (Table 3). Of the 24 locations thus identified, 4 (*carA*, *glxR*, *rutA/rutR*, and *glnE*) have been identified by the genomic SELEX and shown to have RutR-binding activity (Shimada *et al.* 2007). As an attempt to check the RutR binding activity for the rest of newly identified loci, we picked up a 500 bp-long DNA sequence from each of these peaks and searched for a RutR-box using Bioprospector program. As listed in Table 3, a common motif included in these sequences matches well with the known consensus motif, TTGACCAnnTGGTCAA, of RutR binding. This RutR-binding motif was identified in 19 out of the 20 putative RutR targets except for *yahA*. Among 19 targets, 4 (*fepB, ychN/ychO, dhL/ydhM*, and *yhhX*) are located in spacer regions. These 4 sites may be involved in regulation of the neighbouring

genes by RutR. On the other hand, other 15 targets are located within coding sequences. Some of these sites could be involved in regulation of the downstream genes as in the case of *glxR* and *glnE* (Shimada et al., 2007).

Possible influence of cell growth cycle on the RutR distribution within the genome was then analyzed. The ChIP-chip data for the stationary-phase cells are shown in Fig. 10B. Using the same cut-off level as employed for the analysis of log-phase data, a total of 48 RutR-binding sites was identified, which formed 13 separate peak locatioions. All of these 13 peaks in the stationary phase were included in 20 peaks identified for the log-phase cells. The results show that the profile of RutR association for a group of 13 RutR targets, including *carA, gcd, fabA, rutA/rutR, ydhL/ydhM, ves, pmrD, yfiQ, glnE, rpoD, ebgA, yigB, ulaC,* is independent of the growth phase, but that for another group of 7 targets associates RutR only in growing phase (Fig. 11A).

The binding activity of RutR to the target DNA is known to be controlled by uracil and thymine (Shimada *et al.* 2007). The ChIP-chip data, however, indicated that among 20 newly identified targets, the level of RutR association

deceased in the presence of excess uracil only for two targets, *fearR* and *ves* (renamed from *ydjR*) (Fig. 11A).

Table 3. RutR targets identified by ChIP-chip analysis

Peak centre	Gene	Function	Identified by SELEX?	Sequence motif identified by Bioprospector	Site centre	Distance from gene start codon	Gel shift
(1) Identified on non-coding region							
29360	*carA*	Carbamoyl phosphate synthetase	Yes	5'-TTGACCATTTGGTCCA-3'	29366.5	-284.5	Yes
624001	*fepB*	Ferric enterobactin ABC transporter	No	5'-ATGACAAATTCGACAA-3'	623778.5	-45.5	No
1073323	*rutA/rutR*	Pyrimidine utilization/ Transcription factor	Yes	5'-TGGACTAAACGGTCAA-3'	1073350.5	-116.5/-114.5	Yes
1272801	*ychN/ychO*	Unknown/ Unknown	No	5'-CTGACCAATCGGTCAC-3'	1272891.5	-256.5/-69.5	Yes
1723936	*ydhL/ydhM*	Unknown/ Unknown	No	5'-TAGACCGACTGGTCTA-3'	1724024.5	-80.5/-22.5	Yes
3578836	*yhhX*	Unknown	No	5'-ATGACCATGATTTCGT-3'	3578436.5	-0.5	Yes
(2) Identified on coding region							
96016	*mraY*	Peptidoglycan biosynthesis	No	5'-TTGACCGCGTGTTAAG-3'	95807.5	-194.5	Yes
139594	*gcd*	Glucose dehydrogenase	No	5'-CACACCAGTTGGTCAA-3'	139557.5	+1647.5	Yes
535353	*glxR*	Allantoin degradation	Yes	5'-TTAACTGTCTGGTCGG-3'	535245.5	-564.5	Yes
1015645	*fabA*	Fatty acid elongation	No	5'-TTGACCACACGGTCCA-3'	1015576.5	+117.5	Yes
1446529	*feaB*	Phenylthylamine degradation	No	5'-TTGAGCAGGCGGTAAA-3'	1446526.5	+993.5	Yes
1822567	*ves*	Unknown	No	5'-TTTACCACCTGGTCCG-3'	1822688.5	+336.5	Yes
1891016	*yoaA/yoaB*	Unknown/ Unknown	No	5'-CTGACCGTCTGGTGCA-3'	1891237.5	+22.5/-105.5	Yes
2371490	*pmrD*	Polymyxin B resistance	No	5'-TTGACCAGCCATTCCA-3'	2371549.5	+39.5	Yes
2510573	*mntH*	Manganese iron transporter	No	5'-ATCGCCATCAGGTTGG-3'	2510543.5	+182.5	No

2718213	*yfiQ*	Unknown	No	5'-TGGACCAAACAGTCTG-3'	2718345.5	+373.5	Yes
3197797	*glnE*	Glutamine synthetase adenylyltransferase	Yes	5'-CCAACCATTCGGTCAA-3'	3197662.5	-380.5	Yes
3211536	*rpoD*	RNA polymerase sigma70	No	5'-TTGACTACCTGGTCAA-3'	3211491.5	+802.5	Yes
3221325	*ebgA*	β-D-galactosidase	No	5'-TTTACCATCTGGTCAT-3'	3221158.5	+921.5	Yes
3994092	*yigB*	FMN phosphatase	No	5'-TTGACTGGCTGGTCAG-3'	3994157.5	-638.5	Yes
4418941	*ulaC*	L-ascorbate transporter	No	5'-TTGACCATACGGTAAA-3'	4419072.5	-213.5	Yes
4591395	*yjiZ*	L-galactonate MFS transporter	No	5'-TTGACCAGCCAGTCGG-3'	4591335.5	+957.5	Yes
4603338	*fhuF*	Ferrioxamine B iron reductase	No	5'-ATGACCATTGCGCCAG-3'	4603056.5	+175.5	No
(3) No binding motif							
331599	*yahA*	Phosphodiesterase, c-di-GMP-specific	No	-	-		No

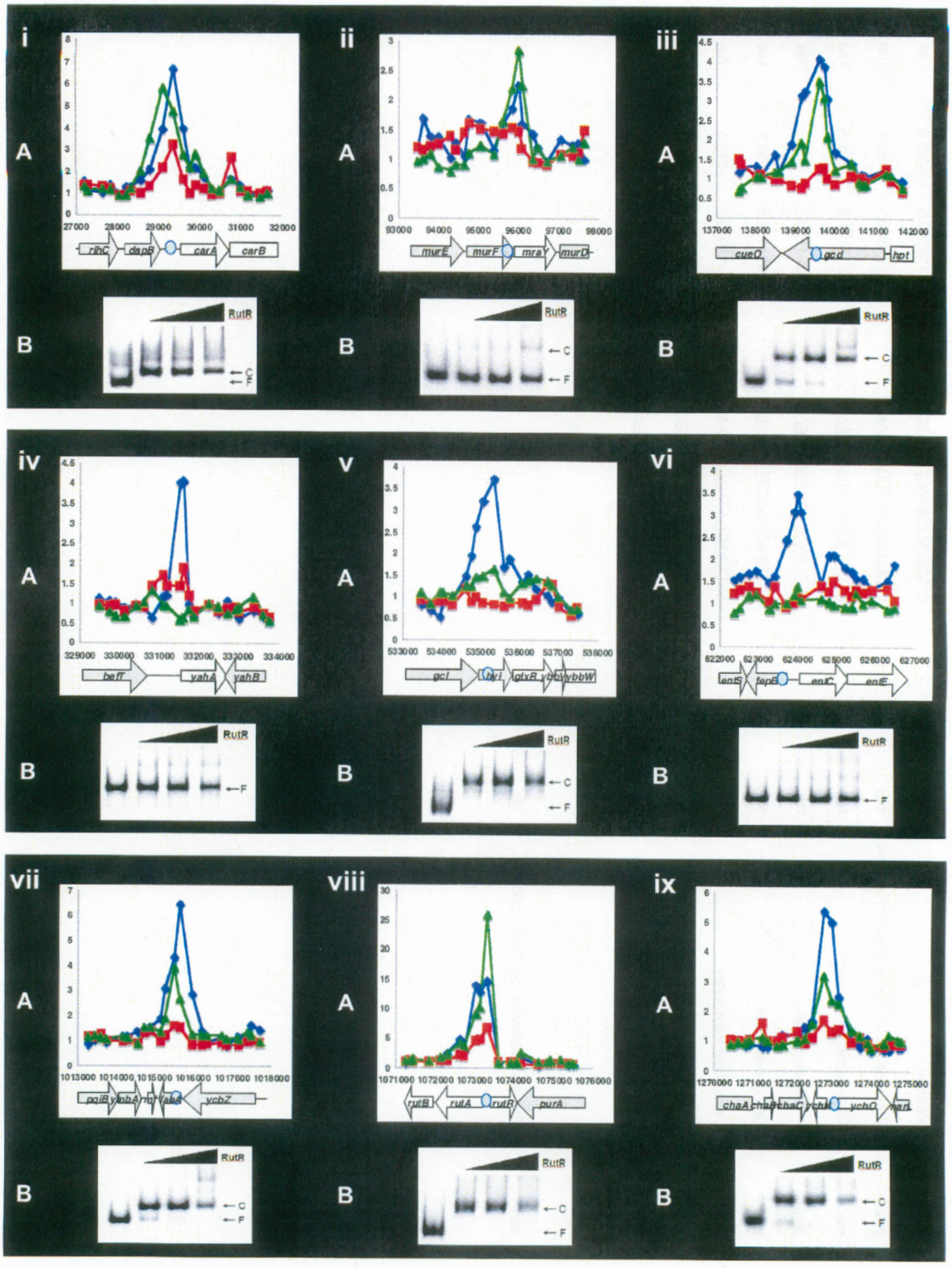
i
ii
iii
iv
v
vi
vii
viii
ix
A
B
RutR
C
F
ribC dapB carA carB
murE murF mraY murD
cueO gcd hpt
yahA yahB
gcl
entS fepB entC entE
ychZ
rutB rutA rutR purA
ychO

Fig. 11-1.

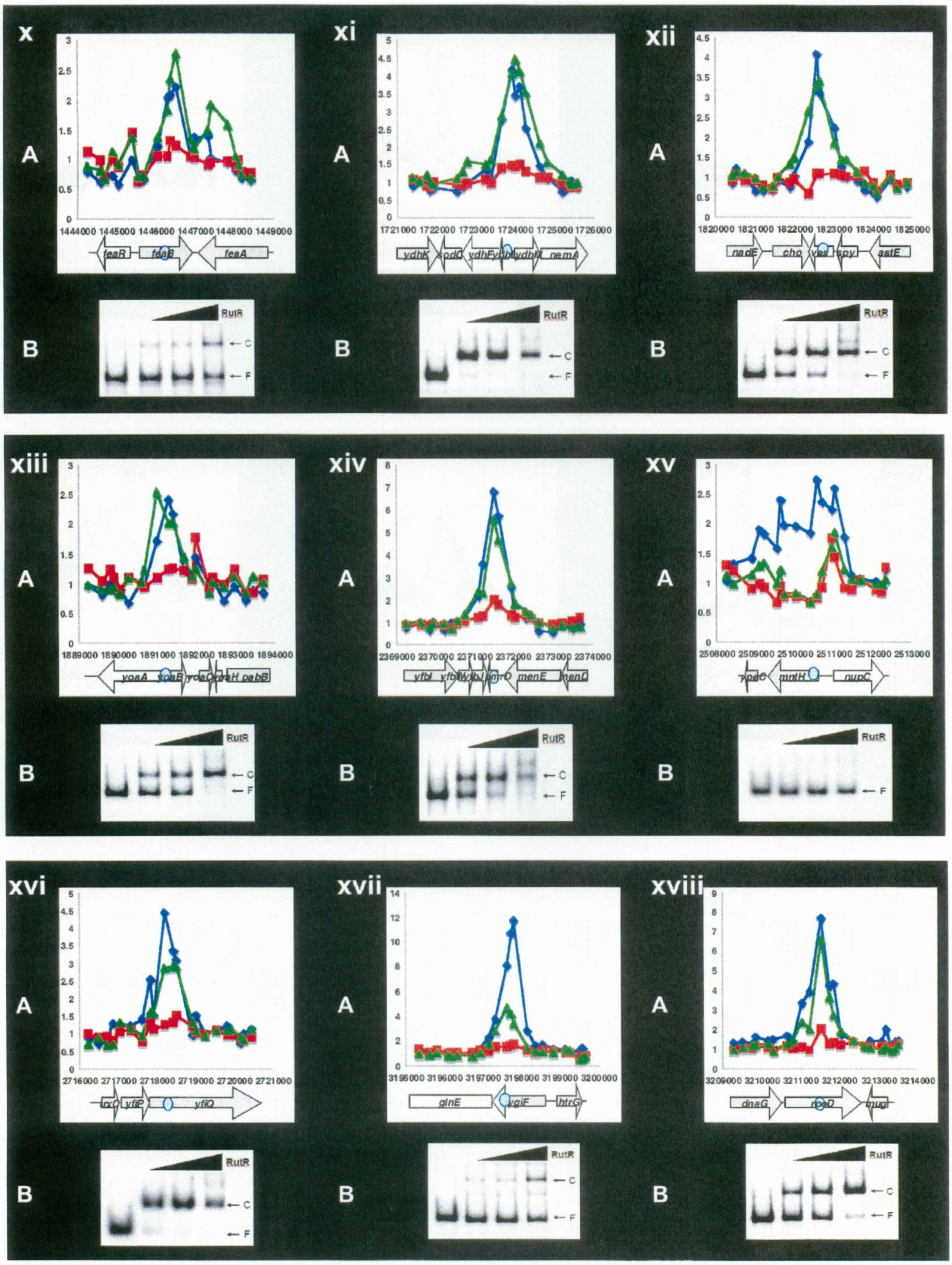
x
xi
xii
xiii
xiv
xv
xvi
xvii
xviii
A
B
RutR
C
F
feaR
feaB
feaA
ydhF
nemA
nadE
cho
astE
yoaA
menE
mntH
nupC
yfiQ
glnE
htrG
dnaG
rpoD

Fig. 11-2.

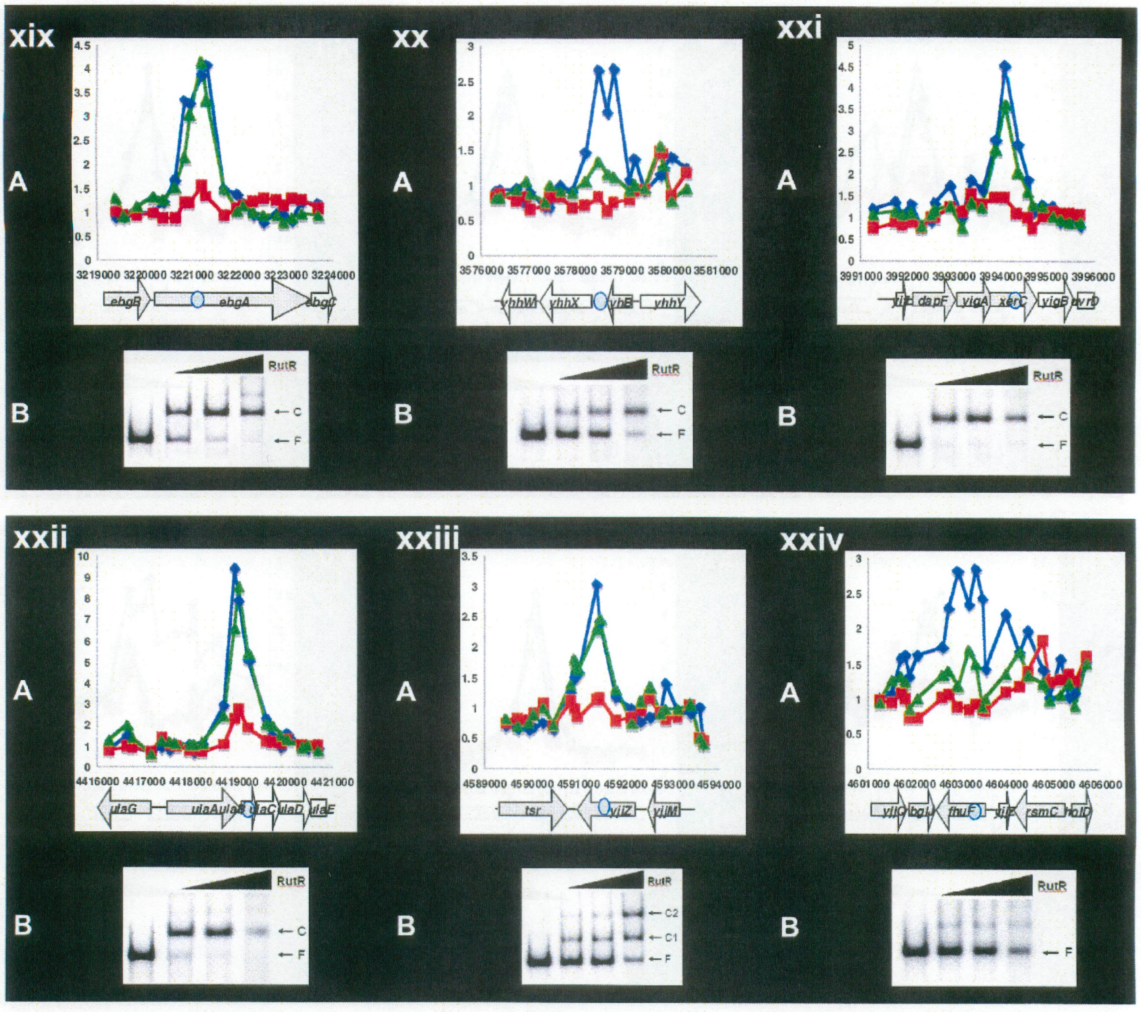
xix
A
B
RutR
C
F
xx
A
B
RutR
C
F
xxi
A
B
RutR
C
F
xxii
A
B
RutR
C
F
xxiii
A
B
RutR
C2
C1
F
xxiv
A
B
RutR
F

Fig. 11-3.

Fig. 11-1, 2, 3. Measurement of binding activity of all RutR targets identified by ChIP-chip assay. (A) The in vivo DNA binding profile of RutR obtained from ChIP-chip experiments and these signals are plotted against the corresponding features of the *E.coli* chromosome. Binding intensity is shown by Cy5/Cy3 signals in mid-log phase conditions (blue), stationary phase conditions (green) and mid-log phase with uracil conditions (red). (B) The in vitro DNA binding activity determine with purified RutR protein. The DNA fragments are amplified center of the each peaks by PCR and end-labelled. DNA fragments were incubated with 0, 10, 25, 50 nM RutR as indicated. RutR-DNA complex is shown as C, Free DNA is shown as F.

RutR-binding activity In vitro *to the RutR-associated genome sequences*

To confirm the RutR-binding activity for all of the RutR-associated sequences detected by ChIP-chip assay, we performed *in vitro* gel shift assays with purified RutR protein. DNA fragments including the RutR-associated sequences were PCR-amplified, radio-labeled at 5' termini, and were used for gel shift assay. A total of 16 out of 20 RutR targets *in vivo* showed RutR-binding activity *in vitro* (Fig. 11B), but 4 putative targets (*yahA, fepB, mntH,* and *fhuF*) did not show the RutR-binding activity under the conditions employed [note that the RutR-box sequence was not identified for *yahA*]. These results indicate that RutR binds these 4 putative targets indirectly, and RutR may form complex with other factors on these regions.

As an attempt to identify the essential bases within the RutR box sequence, a sequence logo representing the common motif was drawn using the set of RutR-associated sequences with the RutR-binding activity *in vitro* (Fig. 12). Both A at position 4 from 5' end and T a position 13 are completely conserved and ACCA between positions 4 and 7 is highly conserved for the

RutR box sequences with RutR-binding activity *in vitro*. On the other hand, the conserved A and T residues are not present in 3 sequences (*fepB, nupC,* and *fhuF*) with no RutR binding activity, and the *yahA* sequence with no RutR-binding activity *in vitro* lacks the RutR box-like sequence.

Fig. 12. RutR binding sequence motif present at newly identified RutR targets. The DNA sequences from each of the 19 targets detected RutR binding activity both in vivo and in vitro (see Table 3) were combined and analysed using Bioprospector (http://ai.stanford.edu/~xsliu/BioProspector/). The motifs identified were then aligned to create a sequence logo (http://weblogo.berkeley.edu/). Individual motifs are shown in Table 3.

Detection of RutR binding effect on neighbor genes transcripts

Both *in vitro* SELEX and *in vivo* ChIP-chip analyses indicated that RutR binds not only to promoter regions and also to sequences within coding regions. If the coding sequence-bound RutR plays a role in transcription regulation, the transcription level of those target sequences should increase or decrease in the absence of RutR binding after addition of effecter uracil. In some cases of transcription regulation in *E. coli*, transcription factors bind to coding sequences of upstream genes as in the case of regulation of *glxR* and *glnE* by RutR (Shimada et al., 2007).

Possible influence of RutR binding to the coding regions was then examined by Northern blot analysis of each target gene in the presence and absence of effecter uracil. Both BW25113 (wild-type) and JW0998 (*rutR* mutant) were grown in M9-glucose medium and RNA was extracted from both log- and stationary phases. For the wild-type culture, RNA was also extracted from the culture in the presence and absence of uracil. All these RNA samples were subjected to Northern analysis using DIG-labeled probes. Results are

summarized in Fig. 13. As described previously (Shimada et al., 2007), the level of *carAB* transcription decreases in the absence of activator RutR or by the addition of uracil to inactivate RutR (Fig. 13, lane carBA). Among the newly identified RutR targets, significant increase in RNA level in the absence of repressor RutR was observed only for *ves* (*ydjR*) (Fig. 13, lane ves). In agreement with this result, marked decrease in the *ves* RNA level was detected by the addition of uracil to inactivate RutR. In the presence of uracil addition, slight but significant decrease was observed for *gcd*, *fabA*, *ychN* and *feaR*, but the RNA levels for these targets were not affected in the *rutR* mutant. Taken all the ChIP-chip data togethere, we concluded that at least one gene, *ves* (*ydjR*), could be included in the RutR regulon.

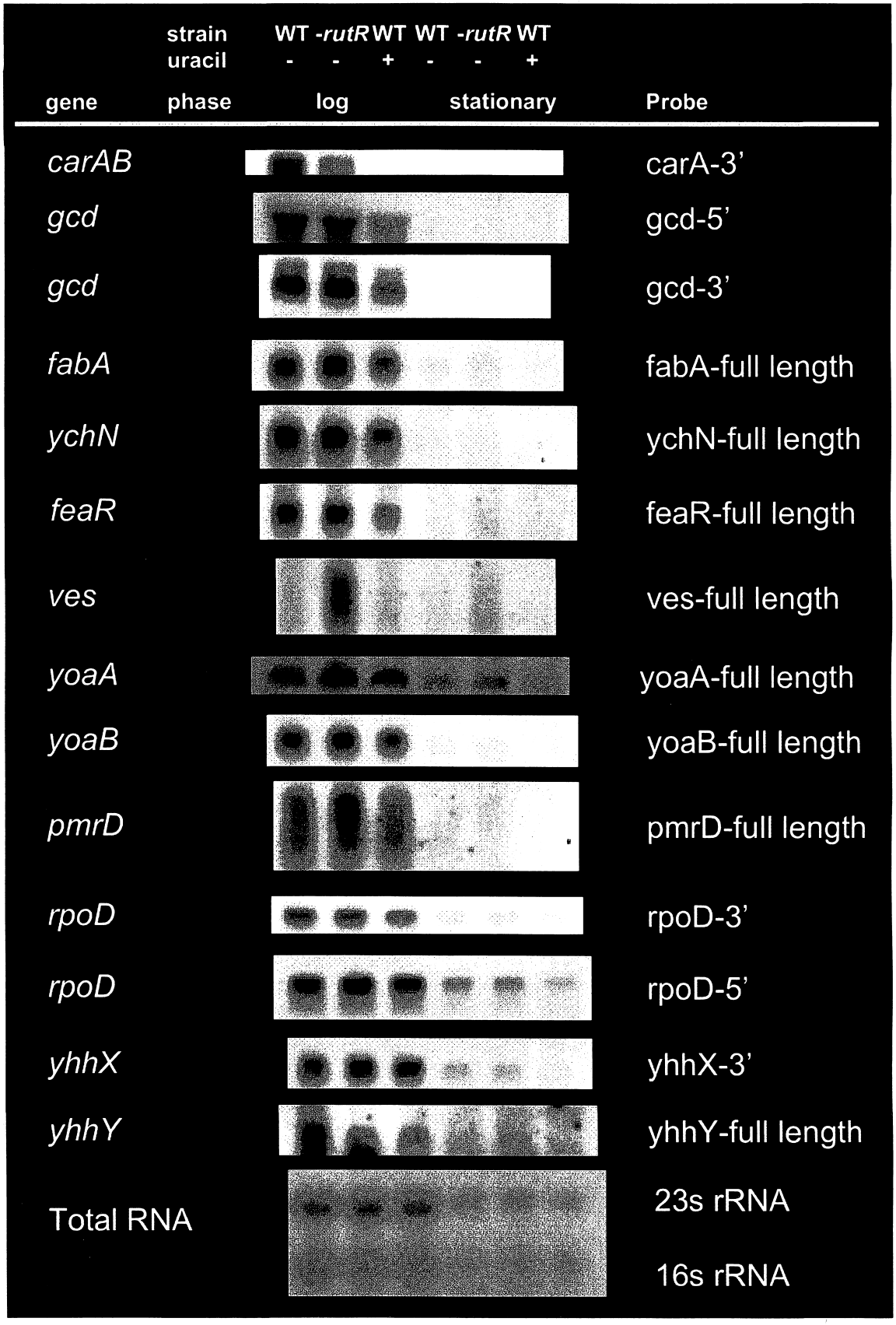
strain
uracil
WT -rutR WT WT -rutR WT
- - + - - +
gene
phase
log
stationary
Probe
carAB
carA-3'
gcd
gcd-5'
gcd
gcd-3'
fabA
fabA-full length
ychN
ychN-full length
feaR
feaR-full length
ves
ves-full length
yoaA
yoaA-full length
yoaB
yoaB-full length
pmrD
pmrD-full length
rpoD
rpoD-3'
rpoD
rpoD-5'
yhhX
yhhX-3'
yhhY
yhhY-full length
Total RNA
23s rRNA
16s rRNA

Fig. 13. Northern blotting of transcripts of the genes predicted RutR target. Wild-type (WT) *E.coli* BW25113 (column 1 and 4) and its *rutR*-deletion mutant (-*rutR*) JW0998 (column 3 and 6) were grown in M9-0.4% glucose minimum medium. Wild-type was also grown in presence of uracil (column 2 and 5). Total RNA was extracted from both exponential-phase (column 1-3) and stationary-phase (column 4-6) cells and 4 ug each was subjected to Northern blot analysis using the indicated probes.

DISCUSSION

Even for the well-characterized *E. coli*, the regulatory function remains unidentified for about one third of a total of 300 DNA-binding transcription factors. As an attempt for quick search of the target genes under the control of transcription factors with unidentified functions, we developed the genomic SELEX system (Shimada *et al.*, 2005; Ogasawara *et al.*, 2007a: 2007b). In this study, we employed this system for search of target DNA sequences recognized by RutR, and identified at least 7 target operons, *carAB, gadAX, gadBC, ygiF-glnE, gcl-hyi-glxR-ybbVW-allB-ybbY-glxK, yciR-deoLT, and rutR/rutABCDEFG* (see Table 2). By gel shift and DNase-I footprinting assays, RutR-binding sites were identified for six targets except for the *yciR-deoLT* operon. These six operons are directly or indirectly involved in the complex pathway of pyrimidine metabolism (Fig. 14).

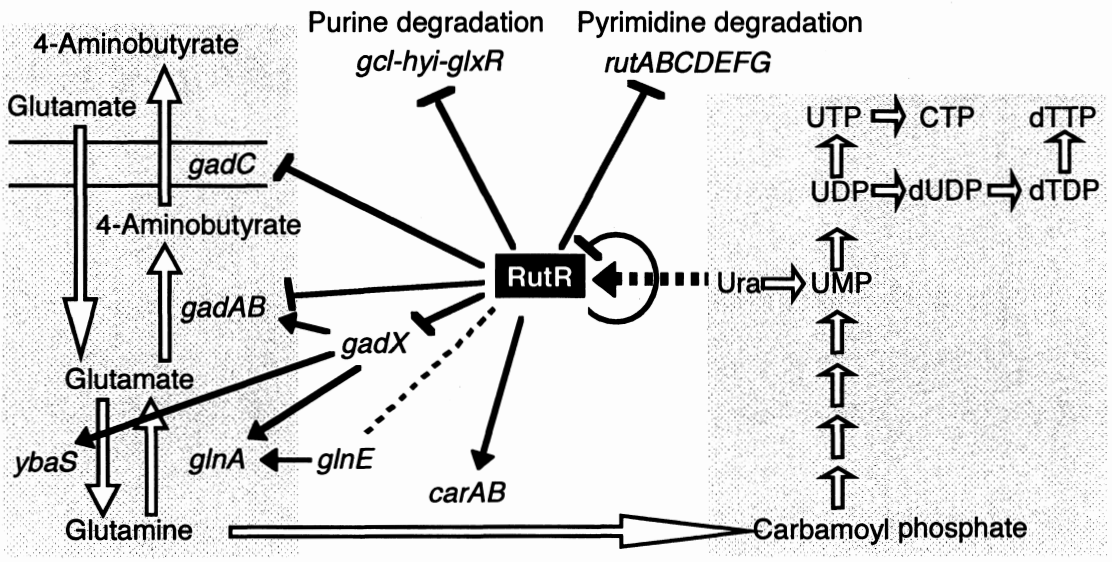

Fig. 10. Proposed model of the RutR regulon. On the bases of SELEX screening of RutR targets and *in vitro* and *in vivo* analyses of the target promoters, we propose that RutR is a global regulator controlling a set of genes (RutR regulon) involved in the synthesis and degradation of pyrimidines. The member genes are either activated or repressed by RutR transcription factor. Bold arrows indicate the activation targets of RutR while the lines with end mark indicate the repression targets. A thin dotted line indicates a predicted target of RutR. Uracil and thymine are the key effectors for controlling the RutR activity.

Regulation of pyrimidine degradation by RutR: During preparation of this report, Loh *et al.* (2006) reported that RutR is a repressor of the *rutABCDEFG* operon encoding a set of enzymes for pyrimidine degradation for its utilization (see Fig. 14). In this study, we indeed identified the RutR-binding site within the 190 bp-long spacer region betweem the *rutABCEFG* and *rutR* operons. Since transcription from the *rutABCEFG* operon is reported to be carried out by the RNA polymerase RpoN holoenzyme ($E\sigma^{54}$), the site (-120 to -174 from the *rutR* initiation codon) of RNA polymerase RpoD holoenzyme ($E\sigma^{70}$) herewith detected (see Fig. 2) represents the promoter for *rutR*. Along the *rutR* promoter, RutR binds to a site (-103 to -128 from the *rutR* initiation codon) downstream from and partially overlapping with the RNA polymerase-binding site, thereby repressing transcription of *rutR* in an autogenous manner. The transcription inhibition of the *rutABCEFG* operon might be attributable to the steric hindrance of RNA polymerase RpoN holoenzyme binding to its promoter by RutR (and RNA polymerase RpoD holoenzyme) on the *rutR* promoter.

Regulation of purine degradation by RutR: In addition to the control of

pyrimidine degradation, we found an RutR-binding site on the *hyi* gene within the *gcl-hyi-glxR-ybbVW-allB-ybbY-glxK* operon encoding the enzymes involved in the downstream pathway of purine degradation (see Table 2 and Fig. 1). In *E. coli*, allantoin, a product of purine degradation, is converted into ureidoglycolate by AllB and AllC, which is then metabolized into two pathways. In the first (glyoxylate) pathway, a set of enzymes, AllIA, Gcl, GlxR and GlxK, convert ureidoglycolate into 3-phosphoglycerate, which is then integrated into the central energy metabolism. In the second (oxaluric acid) pathway, ureidoglycolate is also catabolized into oxaluric acid by AllD, which is ultimately converted into oxamate and carbamoylphosphate (Cusa *et al.*, 1999). Northern blot analysis indicated that transcription of both upstream *gcl* and downstream *glxR* were indeed under the control of RutR (see Fig. 8), indicating that RutR blocks the first pathway of ureidoglycolate catabolism by repressing the *gcl* operon, leading to enhance the second pathway for production of carbamoylphosphate, which is reused for pyrimidine synthesis. The *gcl* operon is also regulated by the allantoin-response transcription factor AllR (Cusa *et al.*, 1999; Rintoul *et al*,

2003). Taken together we concluded that RutR controls not only pyrimidine degradation (Loh *et al.*, 2006) but also purine degradation.

Regulation of pyrimidine synthesis by RutR: Here we also found that RutR regulates the *carAB* genes encoding carbamoylphosphatase for synthesis of carbamoyl phosphate, the key substrate for the synthesis of pyrimidine (and arginine), from glutamine. In *E. coli*, carbamoyl phosphate is a common precursor of pyrimidines and arginine, and thus the expression of *carAB* is under a feedback regulation by various metabolites on the pathways for pyrimidine and arginine synthesis (see Fig. 3B). The *carAB* operon carries two promoters, P1 and P2 (Bouvier *et al.*, 1984; Piette *et al.*, 1984). The downstream promoter P2 responds to arginine and is regulated by the arginine sensor ArgR (Devroede *et al.*, 1998; Wang *et al.*, 1998). The upstream promoter P1 responds to pyrimidine and is under the control of at least four transcription factors, PepA (or CarP) (Charlier *et al.*, 1995), IHF (Charlier *et al.*, 1993), PyrH (Kholti *et al.*, 1998) and PurR (Devroede *et al.*, 2004). PyrH appears to modulate P1 activity indirectly by modifying the promoter sequence (Kholti *et al.*, 1998)], but the mechanism how

to respond to pyrimidine availability remains mystery. Here we identified that, in addition to these four factors, RutR is also involved in P1 regulation. Since the RutR-binding site overlaps with one of two PepA repressor-binding sites (see Fig. 3B) [bi-molecular interaction between two molecules of PepA might be needed for effective repression], transcription activation of *carAB* by RutR might be due to competitive inhibition of the repressor PepA-binding to the upstream operator of the P1 promoter (see Fig, 3C). RutR binding to *carAB* P1 is abolished in the presence of uracil or thymine (see Fig. 4). In the absence of RutR binding to P1 in the presence of excess uracil or thymine in wild-type *E. coli*, the PepA repressor might be able to bind to the target and inhibits *carAB* transcription. When the intracellular pyrimidine level decreases, the ligand-free RutR could be able to bind near the PepA operator site and interferes with the binding of PepA repressor. The effect of pyrimidine nucleosides and nucleotides is much weaker than pyrimidine bases, indicating that the effector molecules *in vivo* must be pyrimidine bases. This prediction was supported by *in vivo* analysis of the *carA* promoter (see Fig. 6) and *carA* RNA level in the presence and absence of uracil

(see Figs. 7 and 8). The P1 promoter activity might also be controlled in response to intracellular UTP level through the UTP-sensitive reiterative transcription of U residues near the transcription initiation site (Han and Turnbough, 1998).

Pyrimidine-sensing function of RutR: Here we identified that RutR is the sensor of uracil (and thymine). Two major metabolic routes exist for the synthesis of pyrimidine nucleotides, *i.e.*, the *de novo* synthesis of pyrimidines from carbamoylphosphate and the salvage pathway in the presence of free pyrimidines (see Fig. 14). Since the salvage of pyrimidines is far more energy efficient than *de novo* synthesis, it is therefore reasonable that uracil (and thymine) inactivates the RutR regulator for shut-off of the *de novo* synthesis pathway. In *E. coli*, the primary salvage pathway reuses free uracil to form UMP by the enzyme uracil phosphoribosyl transferase (Upp). Uracil can also be converted to UMP by the enzymes, uridine phosphorylase (Udp) and uridine kinase (Udk). Up to now, however, the primary salvage pathway of thymine has not been identified, but thymine is also converted to dTMP by the enzymes,

thymidine phosphorylase (DeoA) and thymidine kinase (Tdk). On the other hand, the effecter activity was not detected for cytosine, because cytosine nucleotides are always synthesized from uracil. Cytosine nucleotides are degraded to cytidine, which is then converted to uridine by cytidine aminohydrolase (cytidine deaminase; Cdd), ultimately leading to generate uracil. Thus we propose that the RutR regulator plays a key role in balancing the pyrimidine synthesis pathway between the *de novo* and the salvage pathways. In the presence of free uracil and thymine, the RutR regulator functions as a repressor for the *carAB* genes for inhibition of *de novo* synthesis of pyrimidines.

Regulation of glutamine supply by RutR: The supply of glutamine from externally added glutamate was also found to be indirectly regulated by RutR by controlling the synthesis of GadX and GadW, the regulators of the Gad system for glutamate transport. RutR-binding sites were identified within the coding frames of *gadA* and *gadB* (see Table 2 and Fig. 1), both encoding glutamate decarboxylase. GadA and GadB play major roles in supply of the substrate glutamine for *de novo* synthesis of pyrimidines. Glutamine synthetase, encoded

by *glnA,* provides the nitrogen source through assimilation of glutamate and ammonium into glutamine. GlnE is the enzyme that catalyzes adenylation of GlnA glutamine synthetase to control its enzyme activity (Rhee *et al.*, 1985). Taken together it turned out clear that RutR controls the supply of glutamine, the initial substrate for the *de novo* synthesis of pyrimidines, as well as the conversion of glutamine to carbamoylphosphate by CarAB carbamoylphosphate synthetase (Pierard and Wiame, 1964). The *ybaS* gene is now known to encode glutaminase, which also catalyzes the conversion of glutamate to glutamine. The expression of *ybaS* is also under the control of GadX transcription factor (Tucker *et al.*, 2003).

Taken altogether we propose that RutR is a master regulator of a large set of the genes for the synthesis and degradation of pyrimidines, including the genes for supply of glutamine, the initial substrate for pyrimidine synthesis. It is noteworthy that the master regulator, RutR, for pyrimidine metabolism also plays a role in regulation of the genes for degradation of purines. The interconnection between purine and pyrimidine metabolism may lead to establish the balance of

intracellular concentration between pyrimidines and purines.

Involvment of RutR for maintenance of pH homeostasis: RutR-binding sites were detected within the coding frames of the *ygiF, gadA, gadB* and *hyi* genes (see Table 2). A series of Northern blot analysis indicated that RutR is a repressor of the *gadBC* and *gadAX* operons (see Fig. 8). The *gadA* gene forms an operon with *gadX* and *gadW*, while the *gadB* gene forms an operon with *gadC* encoding the glutamate/4-aminobutyrate (GABA) antiporter. GadX and GadW together participate the regulation of expression of the *gadAXW* and *gadBC* operons (Ma *et al.*, 2002; 2003) (see Fig. 14). In *E. coli*, the physiologic anion is glutamate. The isoforms, GadA and GadB, of glutamate decarboxylase catalyze the conversion of glutamate and proton into gamma-aminobutyrate (GABA) and CO_2 (De Biase *et al.*, 1996; Tramonti *et al.*, 2002). Proton is exported through GadC transporter for maintain intracellular pH homeostasis, thereby leading to decrease environmental pH. On the other hand, GadX is an activator of the *gadAX* and *gadBC* operons. The Gad system is involved in glutamate-dependent acid resistance of *E. coli* for survival under acidic

conditions (Richard and Foster, 2003). Not only CarA and CarB but also GadX and GadC are required for acid resistance, for maintenance of the intracellular pH homeostasis under external acidic conditions. Glutamate is imported into *E. coli* cells through the GadC antiporter and then decarboxylated by GadA/GadB isoforms to GABA consuming a proton, thereby controlling the intracellular pH. GABA is exported through GadC while simultaneously importing the substrate glutamate. Within the *gadAXW* operon, two promoters have been identified, upstream of *gadA* and upstream of *gadX* (Ma *et al.*, 2002) while the *gadBC* operon carries two promoters, upstream of *gadB* and upstream of *gadC* (Biase *et al.*, 1999). The expression of the *gadAXW* and *gadBC* operons varies depending on the culture medium, growth phase and environmental pH. The balance of three transcription factors, GadE, GadX and GadW, is involved in this complex regulation.

Transcription regulation by a downstream-bound transcription factor. In the case of *gadAX*, *gadBC* and *ygiF-glnE* operons, RutR bound within the coding frames of the first genes regulate transcription of the entire operons (see Table 2). On

the other hand, in the case of *gcl* operon including a total of 7 genes for degradation of allantoin, the RutR-binding site was detected within the second *hyi* gene encoding hydroxypyruvate isomerase (see Table 2). The mechanism how the RutR repressor bound on the downstream genes is capable of regulating transcription of the entire operon awaits further studies.

RutR distribution on the genome in E. coli *cells.* For analysis of the RutR distribution along the genome of *E. coli* cells, we employed the ChIP-chip technology. The advantage of this approach to studying the RutR regulon is to identify RutR-associated sites *in vivo*. ChIP-chip technology will allow to detect the RutR targets of poorly transcribed genes. After ChIP-chip analysis of log-phase cells, we identified 24 RutR-associated loci, including 4 known RutR targets (*carA, rutA/rutR, glxR*, and *glnE*) previously identified by genomic SELEX. A total of 20 newly identified targets were tested for RutR-binding activity *in vitro* by gel shift assays, of which 19 DNA fragments showed high affinity binding to RutR and all these sequences contained RutR-box-like sequence with the TTGACCAnnTGGTCAA consensus. After comparison of these RutR binding

motifs with or without RutR binding activity *in vitro*, we identified that within the RutR-box sequence, 4th A and 13th T of binding motif is important for high affinity to RutR (see Fig. 12). Five targets with low affinity to RutR *in vitro* may bind to RutR cooperatively with another factor *in vivo*.

In the ChIP-chip analysis of log-phase cells, we failed to detect the association of RutR to the two known targets, *gadC* and *gadX*, that have been identified to be under the negative control of RutR by genomic SELEX analysis (Shimada et al., 2007). Under the culture conditions herein employed, the binding of RutR repressor to the *gadC* and *gadX* operons must be suppressed by yet unidentified factor(s). In stationary phase cells, 13 targets were detected and all of them are included in the targets in log-phase cells. This finding indicates that the distribution of RutR stays unaltered during the growth phase transition in good agreement with our finding that the intracellular concentration of RutR protein stays constant during growth phase transition of *E. coli* (Shimada *et al.* 2007). Intracellular level and distribution of transcription factors involved in regulation of genes for essential metabolism may be largely unaltered. Using

the ChIP-chip method, Grainger *et al.* (2005) reported that the distribution of FNR and IHF is unaltered during growth phase transition of *E. coli*.

ACKNOWLEDGMENTS

I wish to express my sincerest gratitude to Professor Akira Ishihama, Graduate School of Division of Material Chemistry, Hosei University, for kind guidance, valuable suggestions and discussions, and continuous encouragement throughout the course of this work and critical reading of the manuscript.

I am also deeply indebted to Dr. Kaneyoshi Yamamoto, Hosei University, and Dr. Michihisa Maeda, Meiji University for valuable suggestions and discussions throughout the work.

I wish to express my appresiation to Professor Stephen J. W. Busby and Dr. David C. Grainger, School of Biosciences, University of Birmingham in U.K, for valuable suggestions and discussions throughout the work of ChIP-chip assay in U.K. I would also thank to the members of Professor Busby's laboratory for kindness help. I would also like to thank to Hosei University to award a Scholarship for study ChIP-chip assay in U.K.

I also wish to appreciate for Dr. Dipankar Chatterji, Molecular Biophysics Unit,

Indian Institute of Science, and Dr. Ji Yang, Department of Microbiology and Immunology, University of Melbourne, for collaboration work.

I would also like to thank to Dr. Hirotada Mori for providing the *rutR* knock-out *E.coli* strain.

Many thanks are due to Miss Kiyo Hirao, Miss Ayako Kori, Mr. Yoshito Ogawa, Miss Kayoko Yamada, Dr. Hiroshi Ogasawara, Mr. Jun Teramoto, Mr. Koshiro Yabuki, Mr. Yoshimasa Umezawa and other members of Division of Material Chemistry, Hosei University.

Finally, very special thanks are forwarded to my parents for their continuous support and cooperation, without which this work was never possible.

REFERENCES

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M, Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knock-out mutants: Keio collection. *Mol Systems Biol* doi:10, 1038/msb4100050.

Bouvier, J., Patte, J.C., and Stragier, P. (1984) Multiple regulatory signals in the control region of the *Escherichia coli carAB* operon. *Proc Natl Acad Sci USA* **81**: 4139-4143.

Castanie-Cornet, M, P,. Penfound, T, A., Smith, D,. Elliott, J, F. and Foster, J,W. (1999) Control of acid resistance in *Escherichia coli. J Bacteriol* **181**: 3525-35.

Charlier, D., Hassanzadeh, G., Kholti, A., Gigot, D., Pierard, A., and Glansdorff, N. (1995) *carP*, involved in pyrimidine regulation of the *Escherichia coli* carbamoylphosphate synthetase operon, encodes a sequence-specific DNA-binding protein identical to XerB and PepA, also required for resolution of ColEI multimers. *J Mol Biol* **250**: 392-406.

Charlier, D., Roovers, M., Gigot, D., Huysveld, N., Pierard, A., and Glansdorff, N. (1993) Integration host factor (IHF) modulates the expression of the pyrimidine-specific promoter of the *carAB* operons of *Escherichia coli* K12 and *Salmonella typhimurium* LT2. *Mol Gen Genet* **237**: 273-286.

Cusa, E., Obradors, N., Baldoma, L., Badia, J., and Aguilar, J. (1999) Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli*. *J Bacteriol* **181**: 7479-7484.

De Biase, D., Tramonti, A., John, R.A., and Bossa, F. (1996) Isolation, overexpression, and biochemical characterization of the two isoforms of glutamic acid decarboxylase from *Escherichia coli*. *Protein Expr Purif* **8**: 430-438.

De Biase, D., Tramonti, A., Bossa, F., and Visca, P. (1999) The response to stationary-phase stress conditions in Escherichia coli: role and regulation of the glutamic acid decarboxylase system. *Mol Microbiol* **32**: 1198-1211.

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.

Elgrably-Weiss, M., Schlosser-Silverman, E., Rosenshine, I. and Altuvia, S. (2006) DeoR, a DeoR-type transcriptional regulator of multiple target genes. *FEMS Microbiol Lett* **254**:141-148.

Ellington, A.D., and Szostak, J.W. (1990) *In vitro* selection of DNA molecules that bind specific ligands. *Nature* **346**: 818-822.

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.

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.

Han, X. and Turnbough, Jr., C.L. (1998) Regulation of *carAB* expression in *Escherichia coli* occurs in part through UTP-sensitive reiterative transcription.

J. Bacteriol. **180**: 705-713.

Ishihama, A. (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol* **54**: 499-518.

Ishihama, A. (1990) Molecular assembly and functional modulation of *Escherichia coli* RNA polymerase. *Adv Biophys* **26**, 19-31.

Jishage, M., Iwata, A., Ueda, S. and Ishhama, A. (1996) Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: Intracellular levels of four species of sigma subunit under various growth conditions. *J. Bacteriol.* **178**: 2507-2513.

Khodursky, A.B., Peter, B.J., Cozzarelli, N.R., Botstein, D., Brown, P.O. and Yanofsky, C. (2000) DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Eschericha coli. Proc Natl Acad Sci USA* **97**: 122170-12175.

Kholti, A., Charlier, D., Gigot, D., Huysveld, N., Roovers, M., and Glansdorff, N. (1998) *pyrH*-encoded UMP-kinase directly participates in pyrimidine-specific modulation of promoter activity in *Escherichia coli. J Mol Biol* **280**: 5715-5782

Loh, K.D., Gyaneshwar, P., Papadimitriou, E.M, Fong, R., Kim, K.S., Parales, R., Zhou, Z., Inwood, W., and Kustu, S. (2006) A previously undescribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci USA* **103**: 5114-5119.

Ma, Z., Gong, S., Richard, H., Tucker, D.L., Conway, T., and Foster, J.W. (2002) Collaborative regulation of Escherichia coli glutamate-dependent acid resistance by two AraC-like regulators, GadX and GadW (YhiW). *J Bacteriol* **184**: 7001-12.

Ma, Z., Richard, H., and Foster, J. W. (2003) pH-dependent modulation of cyclic AMP levels and GadW-dependent repression of RpoS affect synthesis of the GadX regulator and *Escherichia coli* acid resistance. *J. Bacteriol* **185**:6852-6859.

Makinoshima, H., Nishimura, A., and Ishihama, A. (2002) Fractionation of *Escherichia coli* cell populations at different stages during growth transition to stationary phase. *Mol Microbiol* **43**: 269-279.

Ogasawara, H., Hasegawa, A., Kanda, E., Miki, T., Yamamoto, K. and Ishihama, A. (2007a) Genomic SELEX search for target genes under the

control of PhoQP-RstBA signal relay cascade. *J Bacteriol* **189**:4791-4899.

Ogasawara, H., Ishida, Y., Yamada, K., Yamamoto, K. and Ishihama, A. (2007b) PdhR (pyruvate dehydrogenase complex regulator) controls the respiratory electron transport system in *Escherichia coli. J Bacteriol* **189**:5534-5541.

Patskovsky, Y.V., Mennella, V., and Almo, S.C. (2006) Crystal structure of hypothetical transcriptional regulator *ycdC* from *Escherichia coli.* PDB data bank, 1pb6

Perez-Rueda, E., and Collado-Vides, J. (2000) The repertoire of DNA-binding transcriptional regulators in *Eschericha coli* K-12. *Nucleic Acids Res* **28**: 1838-1847.

Pierard, A., and Wiame, J.M. (1964) Regulation and mutation affecting a glutamine dependent formation of carbamyl phosphate in *Escherichia coli. Biochem Biophys Res Commun* **15**: 76-81.

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.

Ramos, J.L., Martinez-Bueno, M., Molina-Henares, A.J., Teran, W., Watanabe, K., Zhang, X., Gallegos, M.T., Brennan, R., and Tobes, R. (2005) The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* **69**: 326-356.

Rhee, S.G., Park, S.C., and Koo, J.H. (1985) The role of adenylyltransferase and uridylyltransferase in the regulation of glutamine synthetase in *Escherichia coli*. *Curr Top Cell Regul* **27**: 2212-2232.

Richard H.T., and Foster J.W. (2003) Acid resistance in *Escherichia coli*. *Adv Appl Microbiol* **52**: 167-186.

Rintoul, M.R., Cusa, E., Baldoma, L., Badia, J., Reitzer, L., and Aguilar, J. (2003) Regulation of the Escherichia coli allantoin regulon: coordinated function of the repressor AllR and the activator AllS. *J Mol Biol* **324**: 599-610.

Shimada, T., Makinoshima, H., Ogawa, Y., Miki, T., Maeda, M., and Ishihama, A.

(2004) Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector. *J Bacteriol.* **186**: 7112-7122.

Shimada, T., Fujita, N., Maeda, M., and Ishihama, A. (2005) Systematic search for the Cra-binding promoters using genomic SELEX system. *Genes Cells* **10**: 907-918.

Shimada, T., Hirao, K., Kori, A., Yamamoto, K. and Ishihama, A. (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.

Tramonti, A., Visca, P., De Cano, M., Falconi, M., and De Biase, D. (2002) Functional characterization and regulation of *gadX*, a gene encoding an AraC/XylS-like transcriptional activator of the *Escherichia coli* glutamic acid decarboxylase system. *J Bacteriol* **184**: 2603-2613.

Tucker, D.L., Tucker, N., Ma, Z., Foster, J.W., Miranda, R.L., Cohen, P.S., and Conway, T. (2003) Genes of the GadX-GadW regulon in *Escherichia coli*. *J Bacteriol* **185**: 3190-3201.

Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**: 505-510.

Van Heeswijk, W.C., Rabenberg, M., Westerhoff, H.V., and Kahn, D. (1993) The gees for the glutamine synthetase adenylation cascade are not regulated by nitrogen in *Escherichia coli. Mol Microbiol* **9**:443-457.

Wade, J.T., Struhl, K., Busby, S.J., and Grainger, D.C. (2007) Genomic analysis of protein-DNA interactions in bacteria: insights into transcription and chromosome organization. *Mol Microbiol* **65**: 21-26.

Wang, H., Glansdorff, N., and Charlier, D. (1998) The arginine repressor of *Escherichia coli* K-12 makes direct contacts to minor and major groove determinants of the operators. *J Mol Biol* **277**: 805-824.

Wei, Y., Lee, J.M., Riichmond, C., Blattner, F.R., Rafalski, J.A., and LaRossa, R.A. (2001) High-density microarray-mediated gene expression profiling of *Escherichia coli. J Bacteriol* **183**: 545-556.

Yamamoto, K., and Ishihama, A. (2003) Two different modes of transcription

repression of the *Escherichia coli* acetate operon by IclR. *Mol Microbiol* **47**: 183-194.

PUBLICATIONS

Academic papers

Shimada, T., Makinoshima, H., Ogawa, Y., Miki, T., Maeda, M., and Ishihama, A. "Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector." *J Bacteriol.* **186**: 7112-7122, 2004.

Shimada, T., Fujita, N., Maeda, M., and Ishihama, A. "Systematic search for the Cra-binding promoters using genomic SELEX system." *Genes Cells* **10**: 907-918, 2005.

Shimada, T., Hirao, K., Kori, A., Yamamoto, K. and Ishihama, A. "RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines." *Mol Microbiol* **66**: 744-757, 2007.

Chatterji, D., Ogawa, Y., Shimada, T. and Ishihama, A. "The role of the omega subunit of RNA polymerase in expression of the *relA* gene in *Escherichia coli*." *FEMS Microbiol Lett* **267**: 51-55, 2007.

Yang, J., Ogawa, Y., Camakaris, H., Shimada, T., Ishihama, A. and Pittard, A, J. "*folA*, a new member of the TyrR regulon in *Escherichia coli* K-12." *J Bacteriol* **189**: 6080-6084, 2007.

Ishihama, A., Shimada, T., Ogasawara, H., Teramoto, J. and Yamamoto, K. "Transcription factor-promoter interaction networks." In *Micro-Nano Mechatronics and Human Science* (Fukuda, T., ed.), pp. 165-169, 2005.

Ishihama, A., Ogasawara, H., Shimada, T., Teramoto, J., Hasegawa, A., Umezawa, Y., Yabuki, K., Ishida, Y., Inaba, T., Matsui, M., Kitai, Y., Kori, A., Yamada, K., Hirao, K. and Yamamoto K. "Multi-Scale Genetics towards

Understanding the Hierarchy of Transcription Factor Network in Genome Regulation." In *Micro-Nano Mechatronics and Human Science* (Fukuda, T., ed.), pp. 1-6, 2006.

Presentation at meetings

Shimada T, Hirao K, Fujita N, Yamamoto K, Maeda M, Ishihama A.

「Systematic SELEX search for transcription factor-binding sites on the E.coli Genome.」

Molecular Genetics of Bacteria & Phages.

Cold Spring Harbor, New York. August 2004.

Shimada T, Hirao K, Kori A, Yamamoto K, Grainger DC, Busby SJ and Ishihama A.

「RutR is the Uracil/Thymine Sensing Regulator of a Set of Genes for Synthesis and Degradation of Pyrimidines.」

RNA polymerase workshop.

Imperial College London, London. March 2007.

Shimada T, Hirao K, Kori A, Yamamoto K and Ishihama A.

「RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines.」

ACT-X, Thenth Asian Conference on Transcription

Indian Institute of Science, Bangalore. January 2008.

島田友裕, 牧野嶋秀樹, 山本兼由, 前田理久, 内海龍太郎, 石浜 明

「大腸菌全プロモーターの活性と特異性の網羅的解析」

大腸菌ゲノミクスワークショップー大腸菌のゲノミクスとリソースの開発

淡路島　2003 年 3 月

島田友裕, 牧野嶋秀樹,川島佑介, 山本兼由, 前田理久, 内海龍太郎, 石浜 明

「大腸菌全プロモーターの活性と特異性の網羅的解析」

第２６回日本分子生物学会年会　　神戸　2003年12月

島田友裕，藤田信之，小川宣仁，山本兼由，前田理久，内海龍太郎，石浜 明

「大腸菌転写因子の結合領域の同定と制御機構の解析」

日本農芸化学会　　広島　2004年3月

島田友裕，藤田信之，平尾貴世，山本兼由，前田理久，内海龍太郎、石浜 明

「大腸菌転写因子 Cra の結合領域の同定と制御機構の解析」

日本農芸化学会　　札幌　2005年3月

島田 友裕、平尾 貴世、郡 彩子、山本 兼由、石浜 明

「ウラシル/チミンを感知する転写因子 RutR によるピリミジン合成/分解経路に関与する遺伝子群の制御機構の解析」

第４回２１世紀大腸菌研究会　　静岡　2007年7月

島田 友裕、平尾 貴世、郡 彩子、山本 兼由、石浜 明

「ウラシル/チミンを感知する転写因子 RutR によるピリミジン合成/分解経路に関与する遺伝子群の制御機構の解析」

第 30 回日本分子生物学会年会・第 80 回日本生化学会

横浜　2007 年 7 月