The transcription regulator AllR senses both allantoin and glyoxylate and controls a set of genes for degradation and reutilization of purines

Akiko Hasegawa, Hiroshi Ogasawara, Ayako Kori, Jun Teramoto and Akira Ishihama

Department of Frontier Bioscience, Hosei University, Koganei, Tokyo 184-8584, Japan

Purines are degraded via uric acid to yield allantoin. Under anaerobic conditions, allantoin is further degraded via carbamoylphosphate to $\text{NH}_4^+$ to provide a nitrogen source and, under aerobic conditions, to 3-phosphoglycerate via glyoxylate for energy production. In this study, we found that a DNA-binding transcription factor AllR, together with AllS, plays a key role in switching control of two pathways, nitrogen assimilation and energy production. The repressor function of AllR is activated in the presence of allantoin, the common substrate for both pathways, leading to repression of the genes for energy production. On the other hand, when glyoxylate is accumulated, AllR is inactivated for derepression of the pathway for energy production. RutR, the master regulator for pyrimidines and arginine, is also involved in this pathway-switching control.

INTRODUCTION

In *Escherichia coli*, purines are degraded via uric acid to yield allantoin (Vogels & van der Drift, 1976). Under anaerobic growth conditions, allantoin is used as a nitrogen source. For nitrogen assimilation, allantoin is first converted to ureidoglycolate by AllB (allantoinase) and AllC (allantoate aminohydrolase) (Vogels & van der Drift, 1976; Chang *et al.*, 1993) (see Fig. 8a). Ureidoglycolate is then integrated into two metabolic pathways. In the first pathway, ureidoglycolate is converted into 3-phosphoglycerate by AllA (ureidoglycolate hydrolase), Gcl (glyoxylate carboligase), GlxR (tartronate semialdehyde reductase) and GlxK (glyceraldehyde kinase) for its integration into central energy metabolism. In the second pathway, ureidoglycolate is also converted into oxalurate by AllD (ureidoglycolate dehydrogenase), which is then converted into oxamate and carbamoylphosphate (Cusa *et al.*, 1999), and then processed to yield $3\text{NH}_4^+$ that is used for the synthesis of amino acids via glutamine (Vogels & van der Drift, 1976; Chang *et al.*, 1993). Carbamoylphosphate, an intermediate metabolite in this ureidoglycolate degradation pathway, is also used for the *de novo* synthesis of pyrimidines and arginine (Bouvier *et al.*, 1984; Shimada *et al.*, 2007). These genes are organized in four transcriptional units: allS, allA-allR, gcl-hyi-glxR-ybbV-allB-ybbY-glxK and allD-allG-ylbA.

Abbreviations: RFP, red fluorescent protein; SELEX, systematic evolution of ligands by exponential enrichment; TFP, two-fluorescent protein.

Two supplementary tables listing AllR-binding SELEX fragments, including those that were isolated only once, in the absence or presence of allantoin, are available with the online version of this paper.
carbamoyl phosphate, a common substrate of pyrimidine and arginine synthesis (Shimada et al., 2007). Together, these observations indicate the cross-talk between the catabolic pathways of pyrimidines and purines.

At present, however, the whole set of target genes and promoters for AllR and the effectors and regulatory modes of AllR action remain unsolved. In this study, we first searched for the target genes and promoters regulated by AllR by using the newly developed genomic systematic evolution of ligands by exponential enrichment (SELEX) system (Shimada et al., 2005). Using the target genes or promoters thus identified, we then analysed the possible influence of allantoin and glyoxylate on AllR function. In the case of IclR, the prototype of the AllR-group transcriptional regulators, the activity is controlled by two effectors, glyoxylate and pyruvate (Lorca et al., 2007). Based on the results described herein, we propose that in the case of AllR, allantoin is the co-repressor and glyoxylate is the inducer. Taking all the observations into account, we also propose a switching control model between the nitrogen-assimilation and energy-production pathways.

METHODS

Bacterial strains. E. coli KP7600 (W3110 lacQ lacZAM15 galK2 gal722) and JW0494 (allR deletion derivative of KP7600, a gift from T. Miki, Fukuoka Dental College) were used for analysis of the regulatory roles of AllR. For expression and purification of AllR, the expression plasmid pAllR was transformed into E. coli BL21(DE3) [F- ompT hsd (r5K1- m9A) dcm gal (DE3)]. For expression of AllR, transformants were grown in LB medium at 37 °C. For the promoter assay, a two-fluorescent protein (TFP) promoter assay vector was transformed into wild-type E. coli KP7600 or the allR-disrupted mutant JW0494 (KP7600 ΔallR). Cells were cultured in Luria–Bertani (LB) medium or M9–0.4% glucose medium at 37 °C. When necessary, 100 μg ampicillin ml⁻¹ was added to the medium.

Expression and purification of His-tagged AllR protein. For construction of plasmid pAllR for AllR expression, a DNA fragment corresponding to the AllR coding sequence was amplified by PCR using E. coli W3110 genomic DNA as a template and a pair of primers, which were designed to hybridize upstream or downstream of the AllR coding sequence. After digestion with NdeI and NotI, PCR products were cloned into pET21a(+) (Novagen) between the NdeI and NotI sites. The plasmid construct was confirmed by DNA sequencing. For protein expression, pAllR was transformed into E. coli BL21(DE3). Transformants were grown in 200 ml LB medium, and at OD₆₀₀ 0.6, IPTG was added to a final concentration of 1 mM. After 3 h incubation, cells were harvested by centrifugation, washed with lysis buffer (50 mM Tris-HCl, pH 8.0 at 4 °C, 100 mM NaCl), and then stored at −80 °C until use.

For protein purification, frozen cells were suspended in 3 ml lysis buffer containing 100 mM PMSF. Cells were treated with lysozyme and then subjected to sonication for cell disruption. After centrifugation at 18,000 g for 20 min at 4 °C, the resulting supernatant was mixed with 2 ml 50% Ni-nitritolriocetic acid (NTA) agarose solution (Qiagen) and loaded onto a column. After washing with 10 ml lysis buffer, the column was washed with 10 ml washing buffer (50 mM Tris-HCl, pH 8.0 at 4 °C, 100 mM NaCl), and then with 10 ml washing buffer containing 10 mM imidazole. Proteins were eluted with 2 ml elution buffer (lysis buffer plus 200 mM imidazole), and AllR peak fractions were pooled and dialysed against a storage buffer (50 mM Tris-HCl, pH 7.6 at 4 °C, 200 mM KCl, 10 mM MgCl₂, 0.1 mM EDTA, 1 mM DTT, 50 %, v/v, glycerol), and stored at −80 °C until use. Protein purity was checked by SDS-PAGE.

SELEX search for AllR-binding sequences. The genomic SELEX system was used as described previously (Shimada et al., 2005). Genomic DNA of E. coli W3110 was sonicated to generate fragments of 150–300 bp in length. The E. coli DNA library was constructed after cloning of these fragments into plasmid pBR322 at the EcoR site. A collection of 150–300 bp DNA fragments could be regenerated by PCR amplification using the E. coli DNA plasmid library as the template and a set of primers, EcoR-F (5′-CTGGTTATGCCCG-TACTGC-3′) and EcoR-R (5′-GCCATGCTGCGAATGGAC-3′), which hybridize to the pBR322 vector at EcoR junctions. PCR products thus generated were purified by 5 % PAGE. For the genomic SELEX screening of AllR-binding sequences, 5 pmol of DNA fragments and 20 pmol His-tagged AllR were mixed in a binding buffer (10 mM Tris-HCl, pH 7.8 at 4 °C, 3 mM magnesium acetate, 150 mM NaCl, 1.25 μg BSA ml⁻¹) and incubated for 30 min at 37 °C. The mixture was applied to a Ni-NTA column, and after washing unbound DNA with the binding buffer containing 10 mM imidazole, DNA–AllR complexes were eluted with an elution buffer containing 200 mM imidazole. When necessary, this SELEX cycle was repeated several times. For sequencing of AllR-bound DNA fragments, DNA fragments were dissociated and isolated from DNA–AllR complexes by PAGE and PCR-amplified. PCR products were cloned into the pT7 Blue-T vector (Novagen) using the blunt-end cloning kit (TakaRa) and transformed into E. coli DH5β. Sequencing analysis was carried out using the T7 primer (5′-TAATACGACTCACTATAGGG-3′) with an ABI DNA sequencer 3130x.

Gel mobility shift assay. Gel-shift assays were performed as described previously (Ogasawara et al., 2007a, b). In brief, probes were generated by PCR amplification of AllR-binding sequences from SELEX using a pair of primers, 5′-FITC-labelled T7-F primer (5′-TAAATACGACTCACTATAGGG-3′) and T7-R primer (5′-GGTTTTCCGCCATCAGCGAGG-3′), the genomic SELEX plasmids containing the respective AllR recognition sequences as templates, and Ex Taq DNA polymerase (TakaRa). PCR products with FITC at their termini were purified by PAGE. For gel-shift assays, 0.3 pmol each of the FITC-labelled probes was incubated at 37 °C for 30 min with various amounts of AllR in 12 μl gel shift buffer consisting of 10 mM Tris-HCl, pH 7.8 at 4 °C, 150 mM NaCl, 3 mM magnesium acetate. After addition of a DNA dye solution, the mixture was directly subjected to 6 % PAGE. Fluorescently labelled DNA in gels was detected using the Pharos FX plus system (Bio-Rad).

DNase I footprinting assay. The DNase I footprinting assay was carried out using FITC-labelled DNA fragments, as described previously (Ogasawara et al., 2007a, b). FITC-labelled probes (0.5 pmol each) was incubated at 37 °C for 30 min with various amounts of AllR in DNase I footprinting buffer consisting of 25 μl 10 mM Tris-HCl (pH 7.8), 150 mM NaCl, 3 mM magnesium acetate, 5 mM CaCl₂ and 25 μg BSA ml⁻¹. After incubation for 30 min, DNA digestion was initiated by the addition of 5 ng DNase I (TakaRa). After digestion for 30 s at 25 °C, the reaction was terminated by the addition of 25 μl phenol. Digested products were precipitated with ethanol, dissolved in formamide-dye solution, and analysed by electrophoresis on a 6 % polyacrylamide gel containing 8 M urea.

Measurement of promoter activity. Promoter strength was determined as described previously (Makinoshima et al., 2002; Shimada et al., 2004). In brief, GFP was expressed under the control of a test promoter while red fluorescent protein (RFP) was under the control of a reference promoter. The promoter assay vector pGFP

http://mic.sgmjournals.org
carries two types of fluorescent protein genes, one for RFP under the control of reference promoter lacUV5 and the other for GFP under the control of a test promoter (Makinoshima et al., 2002; Shimada et al., 2004). The gel primer sequence upstream from the initiation codon was amplified by PCR using the genomic DNA from E. coli KP7600 as a template and a pair of primers, H026T (5′-TGGGCAGAGATCGCATGCATGTCGTTGCGT-3′) and H026T (5′-TCATTITATGCATTATTATGTCCCTA-3′). These primers contain the recognition sequences for EcoT22I, BglII and BanHI, which can be used for promoter cloning. The PCR products were digested with the respective restriction enzymes and then ligated into pGRP at EcoT22I and BglII sites. The insertion in the promoter assay plasmid thus constructed was confirmed by sequencing, and the plasmid was named pGRPgel.

For the measurement of the fluorescence intensity of RFP or GFP expressed in E. coli, cells grown in LB medium or M9–0.4% glucose medium up to OD600 0.3–0.5 were harvested by centrifugation, and diluted with PBS (−) to obtain approximately the same cell density (OD600 0.6) for all samples. For the measurement of bulk fluorescence, aliquots of a 0.2 ml cell suspension were added to 0.4 ml flat-bottomed 96-well plates, and the fluorescence was measured with a Wallac 1420 ARVOx multilabel counter (Perkin-Elmer Life Sciences), in which GFP was measured using 485 nm excitation and 535 nm emission and RFP was measured using 544 nm excitation and 590 nm emission. The fluorescence intensity of GFP from the test promoter and RFP (lacUV5 promoter), respectively, while A and B represent the fluorescence intensities of GFP (lacUV5 promoter) and RFP (lacUV5 promoter), respectively.

**Primer extension analysis.** For preparation of total RNA for primer extension analysis, overnight cultures were diluted 100-fold in 100 ml LB medium and cells were grown to OD600 0.6–0.8 (exponential phase). RNA purification was carried out as described previously (Yamamoto et al., 2002). Primer extension analysis was performed using fluorescently labelled probes according to a published procedure (Yamada et al., 1998). In brief, 40 μg total RNA and 1 pmol 5′-FITC-labelled primer were mixed in 20 μl 10 mM Tris-HCl (pH 8.3 at 37 °C), 50 mM KCl, 5 mM MgCl2, 1 mM each of dATP, dTTP, dGTP and dCTP, and 20 U RNase inhibitor. The 5′-FITC-labelled primer used was 5′-GGTGACCTCATCCCCAATCA-3′. The primer extension reaction was initiated by the addition of 5 U avian myeloblastosis virus reverse transcriptase (TaKaRa). After incubation for 1 h at 50 °C, the reaction mixture was precipitated with ethanol and subjected to electrophoresis on a 6% polyacrylamide sequencing gel containing 8 M urea. Fluorescently labelled DNA in gels was detected using a slab gel DNA sequencing gel containing 8 M urea. Fluorescently labelled DNA in gels was detected using a slab gel DNA sequencing gel containing 8 M urea.

**Northern blotting analysis.** Northern blot analysis of allA, allD, csRA and gcl mRNAs was performed essentially as described previously (Shimada et al., 2007). In brief, total RNAs were prepared by the hot phenol method from both E. coli wild-type KP7600 and an allR-disrupted mutant, separated by agarose gel electrophoresis in the presence of formamide, and transferred to a positively charged nylon membrane (Roche). The DIG-labelled probes were prepared using a PCR DIG Probe Synthesis kit (Roche) with pairs of oligonucleotides: allA-F (5′-ATGAAATCCTGAGTACGTGTAATGTC-3′) and allA-R (5′-TCCGCCGCCTCGGTATATGTC-3′) for the allA probe; allD-F (5′-CCTTGCGGTGTTTATGTGC-3′) and allD-R (5′-TTATGGCGAAAGAGAGTTTGC-3′) for the allD probe; gcl-F (5′-CCGGCGTGTTTATGTGTC-3′) and gcl-R (5′-TTATGGCGAAAGAGAGTTTGC-3′) for the gcl probe; and csRA-F (5′-ATGCCTGATCCGTCGTTTGC-3′) and csRA-R (5′-TTAGTAACTGACTGTCGTTTGC-3′) for the csRA probe. For measurement of 16S and 23S rRNA used as reference controls, hybridization was carried out at 50 °C for 16 h in DIG Easy Hyb, as recommended by the supplier (Roche). Hybridized probes on membranes were detected by CDP-Star, Ready-to-Use (Roche).

**RESULTS AND DISCUSSION**

AllR is a member of the large IclR family of DNA-binding transcriptional regulators and shares 42% sequence identity with IclR (isocitrate lyase gene regulator), the founding member of this family (N. Fujita and A. Ishihama, unpublished data). The E. coli K-12 genome contains eight members of the IclR family, while a total of about 450 member genes have been identified in bacteria and archaea (Walker et al., 2006). The crystal structures of the C-terminal effector-binding domain of the AllR regulator and its complex with glyoxylate have been determined at the resolution of 1.7 and 1.8 Å (0.17 and 0.18 nm), respectively (cited in Walker et al., 2006). The active protomer of AllR is a dimer of dimers. Mutations in the tetramerization region, including the Cys135–Cys142 region, resulted in dissociation of the AllR tetramer to dimers in vitro and are functionally inactive in vivo.

**Isolation of AllR-binding sequences from the E. coli genome by genomic SELEX**

For the identification of DNA sequences that are recognized by E. coli AllR, we employed the genomic SELEX method (Shimada et al., 2005), in which a complete library of E. coli genome DNA fragments is used instead of synthetic oligonucleotides with all possible sequences used in the original SELEX method (Ellington & Szostak, 1990; Tuerk & Gold, 1990; Singer et al., 1997). First, we constructed the plasmid library, each carrying a size-fractionated DNA fragment (150–300 bp in length) from a pool of sonicated E. coli W3110 genome DNA. In each experiment, the DNA fragment mixture was regenerated after amplification of the inserted DNA fragments by PCR. Into the mixture of PAGE-purified DNA fragments, a fourfold molar excess of the purified His-tagged AllR protein was added, and the AllR–DNA complexes were affinity-purified. In the early stage of this genomic SELEX cycle, the AllR-bound DNA fragments gave smear bands on PAGE, as did the original genome fragment mixture. After two and three SELEX cycles, several discrete bands were identified, indicating that some DNA fragments with high affinity to AllR were enriched. These DNA fragments were recovered from the gel and cloned into pT7 Blue-T plasmid (Novagen) for sequencing.

A total of 163 clones were isolated from 65 regions [87 clones from 12 spacer regions (group A) and 76 clones from 53 different coding regions (group B)] of the E. coli genome (Table 1; and Supplementary Table S1). Table 1 includes the list of clones that were isolated at least twice from single loci. Since most regulatory proteins that affect transcription initiation generally bind upstream of their
Table 1. SELEX search for AllR-binding sites in the absence of allantoin

The genomic SELEX to search for AllR-binding sequences was performed in the absence of allantoin. A total of 163 DNA fragments were isolated, of which 87 contained sequences within spacer regions between two ORFs, while 76 fragments carried the sequences within protein-coding frames. This table shows the clones that were isolated at least twice independently. The SELEX fragments that were isolated only once are listed in Supplementary Table S1. The numbers on both sides of each SELEX ($) fragment indicate the boundaries in the revised E. coli genome map. The arrows indicate the direction of transcription of the neighbouring genes. Genes in bold type represent possible regulatory targets of AllR, because AllR binds to the upstream region.

<table>
<thead>
<tr>
<th>Left gene</th>
<th>SELEX fragment</th>
<th>Right gene</th>
<th>Number of fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group A AllR sites within spacer regions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>allR (→)</td>
<td>(0532575) S (0532845)</td>
<td>(→) gcl</td>
<td>48</td>
</tr>
<tr>
<td>allS (→)</td>
<td>(0531193) S (0531397)</td>
<td>(→) allA</td>
<td>24</td>
</tr>
<tr>
<td>lepA (→)</td>
<td>(2706603) S (2706770)</td>
<td>(→) rseC</td>
<td>3</td>
</tr>
<tr>
<td>nikR (→)</td>
<td>(3619099) S (3619361)</td>
<td>(→) rseB</td>
<td>3</td>
</tr>
<tr>
<td>ompC (→)</td>
<td>(2317853) S (2318097)</td>
<td>(→) yojN</td>
<td>2</td>
</tr>
<tr>
<td>btuB (→)</td>
<td>(4164841) S (4165061)</td>
<td>(→) murl</td>
<td>2</td>
</tr>
<tr>
<td>cycA (→)</td>
<td>(4430666) S (4430889)</td>
<td>(→) ytfE</td>
<td>2</td>
</tr>
</tbody>
</table>

**Group B AllR sites within ORFs**

cueO (→) | (0140176) S (gcl) (0140347) | (→) hpt | 4 |

ugpQ (→) | (3588140) S (ugpC) (3588367) | (→) ugpE | 3 |

surA (→) | (0055315) S (imp) (0055509) | (→) djIA | 2 |

ydiL (→) | (0145514) S (ydiE) (0145751) | (→) pantD | 2 |

lpxA (→) | (0203499) S (lpxB) (0203697) | (→) rhnB | 2 |

aupZ (→) | (0916729) S (ybjD) (0916939) | (→) ybiX | 2 |

cyD (→)  | (0931444) S (trxR) (0931676) | (→) lp | 2 |

hyA (→)  | (1034345) S (hyaB) (1034565) | (→) hyaC | 2 |

hns (→)  | (1295922) S (tdk) (1296142) | (→) ychG | 2 |

ynjI (→) | (1849084) S (topB) (1849297) | (→) selD | 2 |

wcaL (→) | (2120403) S (wcaK) (2120648) | (→) wzxA | 2 |

yfbF (→) | (2373855) S (yfbG) (2374052) | (→) yfbH | 2 |

mntH (→) | (2520137) S (mntC) (2520350) | (→) insL | 2 |

nadB (→) | (2711867) S (nfyC) (2712030) | (→) srmB | 2 |

recN (→) | (2753315) S (smpA) (2753303) | (→) yfI | 2 |

yjiV (→) | (3241854) S (uxaA) (3242039) | (→) uxaC | 2 |

glmS (→) | (3914110) S (glmU) (3914259) | (→) atpC | 2 |

ggpA (→) | (3964643) S (rhlB) (3964829) | (→) trxA | 2 |

cutA (→) | (4365619) S (dcaA) (4365880) | (→) aspA | 2 |

yjB (→)  | (4407779) S (rnr) (4407981) | (→) yjH | 2 |

| respective target genes or near their promoters, the regulation targets of group A binding sites are likely to be gcl, lepA, rhsB, murl and two divergently transcribed sets, allR and/or allS, and omcC and/or yojN. The most abundant 48 clones carried sequences from a spacer region that is located downstream of allR and upstream of gcl. In this case, AllR was considered to control the gcl promoter on the basis of gene orientation. Twenty-four clones carried sequences from a spacer region between allS and allA. The binding of AllR to the intergenic region of the divergently transcribed allS and allA operons indicates either autogenous regulation of allS itself or regulation of allS, which encodes another regulator that controls the second pathway for carbamoylphosphate synthesis. The products of the other candidate genes (LepA, RhsB, OmpC and Murl) possibly under the control of AllR are all membrane proteins, supposedly involved in transport and/or stress responses. In some known cases of transcription regulation in E. coli, transcription factors bind within ORFs of the upstream genes. Thus, AllR may also control group B genes such as hpt (hypoxanthine phosphoribosyltransferase) and ugpQ (glycerophosphodiester phosphodiesterase). |

Confirmation of AllR binding to SELEX fragments and search for effectors affecting AllR activity

To confirm the binding activity of AllR to the SELEX DNA fragments, we first examined two abundant SELEX fragments, allR-gcl and allS-allA spacer sequences, for their binding activity towards purified AllR protein. As shown in Fig. 1(a, b), both SELEX DNA fragments formed AllR complexes in a protein-dose-dependent manner. In order to identify the effectors affecting AllR function, we examined the influence of allantoin, glyoxylate and related
compounds on the binding of AllR to the target sequences [both allantoin and glyoxylate are the metabolites in the pathway of allantoin degradation (see Fig. 8a)].

In the presence of 60 mM allantoin, AllR–DNA interaction was enhanced for both SELEX fragment probes (Fig. 1a, b, compare lanes 1–4 with lanes 5–8). This finding indicates that the hitherto undescribed effector allantoin is a corepressor for enhancing the binding of the AllR repressor to its targets. On the other hand, when 40 mM glyoxylate was added in the simultaneous presence of allantoin, the level of AllR–probe complexes markedly decreased (Fig. 1c, d), confirming the notion that glyoxylate is an inducer for derepression of AllR action (Rintoul et al., 2002). Altering the residues Cys217, Ser234 and Ser236 is known to change AllR to a glyoxylate-independent repressive form (Walker et al., 2006). Structurally related compounds such as glyconate or D-lactate did not affect the DNA-binding activity of AllR (data not shown).

**Genomic SELEX screening in the presence of corepressor allantoin**

The initial genomic SELEX was performed using the purified AllR protein alone in the absence of effectors affecting AllR function. Since allantoin was identified as a corepressor of AllR, we repeated the genomic SELEX in the continuous presence of 90 mM allantoin. A total of 86 independent clones could be classified into two groups, 43 group A clones from 18 spacer regions and 43 group B clones from 35 coding sequences (Table 2, Supplementary Table S2). Again the allR-gcl and allS-allA spacer sequences were the most abundant in group A, both being isolated even in the absence of allantoin (see Table 1). In addition, several candidate novel targets were identified including artP, frdB, hisM, leuO (or leuL), yahB and ygdH (Table 2). Since the second genomic SELEX was performed in the presence of the corepressor allantoin, the sequences isolated may be more specific than those isolated in the initial genomic SELEX.

**Examination of AllR binding to novel SELEX fragments by gel mobility shift**

For confirmation of AllR binding to the sequences identified in the second genomic SELEX screening, we performed the gel mobility shift assay. Fig. 2 shows the gel pattern for five representative DNA sequences, leuL-leuO, yqcD-ygdH and csrA-alaS from group A, and deoD-[lyjj]-lplA and ycsF-[mfd]-ycfT from group B (see Table 2). All these SELEX fragments formed complexes with AllR (Fig. 2), even though the minimum concentrations of AllR needed to form these DNA complexes were lower than those required to form allS-allA and allR-gcl complexes (compare Figs 1 and 2; note that the experiments in Figs 1 and 2 were performed in parallel), indicating the higher affinity of the allS-allA and allR-gcl fragments for AllR than that of the other five probes. This finding also supports the prediction that a good correlation exists between the number of SELEX DNA sequences isolated and their affinities to the test DNA-binding protein (Ogasawara et al., 2007a, b; Shimada et al., 2005, 2007).

**Identification of AllR-binding sites within the target promoters**

To identify the DNA recognition sequence of AllR, we next carried out a DNase I footprinting assay for two group A probes, allS-allA and allR-gcl. After forming complexes in the presence of a fixed amount of the respective fluorescently labelled DNA probes and increasing amounts of AllR, DNase I treatment was carried out for a short period, and the partially digested DNA products were

---

**Fig. 1.** Gel mobility shift assay of AllR–target DNA complex formation and effects of allantoin and glyoxylate. Fluorescently labelled DNA probes, each containing a SELEX segment from the allR-gcl (a) or allS-allA (b) spacer region, were incubated at 37 °C for 30 min in the absence or presence of 60 mM allantoin, as indicated, with increasing amounts of AllR (lanes from left to right: 0, 0.18, 0.36 and 0.72 pmol). Binding of AllR to the same fluorescently labelled DNA probes, allR-gcl (c) and allS-allR (d), was also examined in the presence of 40 mM glyoxylate. The reaction mixtures were directly subjected to PAGE.
analysed by PAGE. For both probes examined, clear protection by AllR was found in the presence of high concentrations of AllR apparently at a single region (Fig. 3a, b). For the \textit{allS-allA} spacer region, AllR protected a 29 bp sequence between −18 and −46 of the \textit{allS} promoter from DNase I digestion (Fig. 3a). This protected segment

### Table 2. SELEX screening in the presence of allantoin

The genomic SELEX to search for AllR-binding sequences was performed in the presence of allantoin. A total of 86 DNA fragments were isolated, of which 43 contained sequences within spacer regions between two ORFs, while 43 fragments carried the sequences within the indicated coding frames. This table shows the clones that were isolated at least twice independently. The SELEX fragments that were isolated only once are listed in Supplementary Table S2. The numbers on both sides of each SELEX (S) fragment indicate the boundaries in the revised \textit{E. coli} genome map. The arrows indicate the direction of transcription of the neighbouring genes. Genes in bold type represent possible regulatory targets of AllR, because AllR binds to the upstream region.

<table>
<thead>
<tr>
<th>Left gene</th>
<th>SELEX fragment</th>
<th>Right gene</th>
<th>Number of fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{allS} (†)</td>
<td>(0531191) S (0531393)</td>
<td>\textit{allR}</td>
<td>16</td>
</tr>
<tr>
<td>\textit{allR} (†)</td>
<td>(0532645) S (0532820)</td>
<td>\textit{gcl}</td>
<td>5</td>
</tr>
<tr>
<td>\textit{artP} (†)</td>
<td>(0903829) S (0904028)</td>
<td>\textit{ybjP}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{leuL} (†)</td>
<td>(0083698) S (0083873)</td>
<td>\textit{leuO}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{ygcD} (†)</td>
<td>(2925614) S (2925794)</td>
<td>\textit{ygdH}</td>
<td>2</td>
</tr>
<tr>
<td>Group B AllR sites within ORFs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{ycsF} (†)</td>
<td>(1174932) S (1175100)</td>
<td>\textit{ycfT}</td>
<td>4</td>
</tr>
<tr>
<td>\textit{deoD} (†)</td>
<td>(4621241) S (4621421)</td>
<td>\textit{lplA}</td>
<td>3</td>
</tr>
<tr>
<td>\textit{entA} (†)</td>
<td>(0628273) S (0628457)</td>
<td>\textit{cstA}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{frdB} (†)</td>
<td>(4380243) S (4380411)</td>
<td>\textit{ponA}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{hisM} (†)</td>
<td>(2431860) S (2432038)</td>
<td>\textit{hisJ}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{nana} (†)</td>
<td>(3373372) S (3373577)</td>
<td>\textit{dcuA}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{rcsC} (†)</td>
<td>(2326385) S (2326591)</td>
<td>\textit{atoC}</td>
<td>2</td>
</tr>
<tr>
<td>\textit{yahB} (†)</td>
<td>(0335347) S (333684)</td>
<td>\textit{yahD}</td>
<td>2</td>
</tr>
</tbody>
</table>

**Fig. 2.** Gel mobility shift assay of AllR–SELEX fragment complexes. Fluorescently labelled DNA probes, each containing a SELEX segment from the \textit{leuL-leuO} (a), \textit{csrA-alaS} (b), \textit{ygcD-ydgH} (c), \textit{deoD-[yjjJ]-lplA} (d) and \textit{ycsF-[mfd]-ycfT} (e) regions, were incubated at 37 °C for 30 min in the absence or presence of 60 mM allantoin, as indicated, with increasing amounts of AllR (from left to right: 0, 0.18, 0.36 and 0.72 pmol). The reaction mixtures were directly subjected to PAGE.

http://mic.sgmjournals.org 3371

IP: 54.70.40.11
On: Fri, 15 Mar 2019 05:50:04
includes the proposed 17 bp AllR-binding palindromic sequence, TTGGAAAAAATTTC (boxed in Fig. 3a), hereafter designated the AllR box. The AllR box completely overlaps the allS promoter −35, in good agreement with the negative regulation of allS transcription by AllR (Rintoul et al., 2002).
The location of this AllR box, however, corresponds to the 17 bp sequence between positions −45 and −61 from the transcription initiation site (or between −107 and −123 from the initiation codon) of the divergently transcribed allA gene (Fig. 3a). allA gene transcription is also negatively regulated by AllR, even though AllR binds upstream of the allA promoter −35. Since only a single AllR-binding site was identified in the allS-allA spacer region, we concluded that a single, identical AllR molecule represses transcription from both allA and allS promoters, together directing divergent transcription. On the basis of AllR binding relative to the allA promoter, the mode of allR repression may be due to prevention of the promoter escape of RNA polymerase because of the tight association of AllR with RNA polymerase α subunit CTD. Transcription repression by an upstream-bound transcription factor has been reported in some cases (for example, see Yamamoto & Ishihama, 2003). The allS repression by the upstream-bound AllR may also take place through steric occlusion by the allA promoter complex, preventing αCTD binding to the allS promoter, or interference by negative supercoiling created by allA transcription.

In the case of the allR-gcl spacer region, AllR was found to protect a 24 bp sequence between positions −17 and −41 upstream of the gcl initiation codon, including the 17 bp conserved AllR box sequence between −21 and −37 (Fig. 3b). Although the transcription initiation site has not been identified for gcl, this AllR box must be located downstream from the gcl promoter, leading to tight repression of gcl operon transcription by AllR. In the presence of the inducer glyoxylate, the gcl operon might be derepressed to produce the enzymes for utilization of glyoxylate as an energy source.

**Induction in vivo of the gcl promoter by glyoxylate**

*In vitro* gel shift and DNase I footprinting assays indicated that the AllR binding to both allS-allR and gcl promoters is reduced in the presence of glyoxylate. To confirm the induction *in vivo* of the gcl promoter by glyoxylate, we employed the TFP promoter assay vector pGRH026, in which the expression of GFP is under the control of the gcl promoter while RFP expression is directed by the reference promoter lacUV5 (see Methods). In the continuous presence of 40 mM glyoxylate a significant level of GFP expression was detected (Fig. 4a).

We next carried out a time-course experiment for analysis of the induction kinetics of gcl expression by glyoxylate. Wild-type *E. coli* KP7600 was grown in M9–0.4% glucose medium, and at the cell density (OD600) of 0.3–0.5, glyoxylate was added to a final concentration of 40 mM. Induction of gcl expression was detected within 30 min and the GFP level reached a level higher than twofold at 2 h after glyoxylate addition (Fig. 4b).

**Roles of glyoxylate and allantoin in transcription in vivo of the AllR regulon genes**

Detailed and direct analysis of mRNA of the AllR regulon genes was then performed by Northern blot analysis.
mRNA levels of \( \text{allA} \) and \( \text{gcl} \), each being the first gene of their respective operons (see Fig. 8b), increased markedly after addition of the inducer glyoxylate to wild-type \( \text{E. coli} \) K7600 (Fig. 5a, lanes 1 and 2). Although the \( \text{allD} \) promoter is not under the direct control of \( \text{AllR} \), a significant increase was observed for \( \text{allD} \) mRNA resulting from derepression of the gene encoding \( \text{AllS} \), the activator of the \( \text{allD} \) operon. In good agreement with this finding, these
mRNAs were not identified in the allR-disrupted mutants even in the absence of glyoxylate, and the mRNA levels did not further increase after glyoxylate addition (Fig. 5a, lanes 3 and 4).

In the AllR–DNA complex formation assay in vitro, allantoin was identified as the co-repressor of AllR (see Fig. 1). We then analysed possible influences of allantoin on transcription in vivo of the target genes. Under the culture conditions employed, however, detectable levels of mRNA were not identified in the wild-type E. coli even in the absence of allantoin (Fig. 5b, lanes 1 and 2). In the allR mutant, the level of gcl mRNA was high, irrespective of the presence or absence of allantoin (Fig. 5b, lanes 3 and 4). The level of allID mRNA, however, significantly increased in the presence of allantoin, presumably due to the increase in the level of functional AllS (compare Fig. 5b, lanes 3 and 4).

### Novel targets of transcription regulation by AllIR

CsrA, an RNA-binding protein, is a pleiotropic global regulator of carbon source metabolism, affecting glycogen biosynthesis, gluconeogenesis, cell size and surface properties, and motility and flagellum biosynthesis (Liu et al., 1995; Dubey et al., 2003). In this study, we have identified the csrA promoter as one of the novel candidate promoters under the control of AllR. To identify the transcription initiation site in the csrA promoter, we carried out a primer extension assay. Both wild-type KP700 and allR mutant JW0494 were grown in LB to mid-exponential phase, then 10 mM glyoxylate was added for 30 min to half of each culture. At least two bands of csrA transcript were identified at 388 (P1) and 180 (P2) bp upstream from the translation initiation codon in both wild-type and allR mutant (Fig. 6). The AllR box-like sequence was located between –273 and –288 bp from the initiation codon, or downstream of the P1 and upstream of the P2 promoter, respectively. Neither allR disruption nor glyoxylate addition had a significant influence on the levels of csrA transcripts. This finding indicates that although AllR binds upstream of the csrA promoter, it does not affect transcription of the csrA gene.

We also performed Northern blot analysis of csrA RNA. The csrA gene appears to be highly expressed in both wild-type and allR-mutant E. coli under the culture conditions employed (see Fig. 5). No significant change was observed, however, for the level of csrA mRNA after the addition of glyoxylate or allantoin, in agreement with the primer extension analysis (see Fig. 6). Thus, the role of AllR binding to the csrA promoter region remains undetermined.

### Determination of the intracellular concentration of AllIR

Although AllR has been believed to be constitutively expressed, this notion has never been experimentally tested. We therefore determined the intracellular concentration of AllR in wild-type E. coli cell culture grown in LB under various growth conditions. For quantitative measurement, a standard curve of AllR protein was prepared using a mixture of purified AllR and a whole-cell extract of the allR-disrupted mutant. By using this quantitative immunoblot analysis (for details see Methods), the concentration of AllR protein in cells at various growth phases grown between 22 and 42 °C was determined (Fig. 7). The AllR level stayed relatively constant at the level of approximately 500 molecules per cell as compared with the level of RNA polymerase α subunit. From this experiment, we propose that the level of functional AllR is controlled by regulating its activity through interaction with two effectors, allantoin and glyoxylate.

### Control of AllIR activity by two effectors

Here, we identified that at least two effectors, allantoin and glyoxylate, are involved in the control of AllR activity. The gcl operon for the glyoxylate pathway (Fig. 8a) is repressed by the AllR–allantoin complex under anaerobic conditions or in the absence of oxygen. In the presence of oxygen, glyoxylate plays a role in derepression of the gcl operon for production of the enzymes involved in the glyoxylate pathway (Fig. 8b). The data described herewith support the model whereby AllR controls the allS gene which encodes another regulator of the allantoin regulon (Fig. 8b). The allS operon is under the control of the repressor AllIR, but

---

**Fig. 7.** Intracellular concentration of the AllR protein. Wild-type E. coli K-12 W3350 was grown in LB medium at 22, 27, 32, 37 and 42 °C under aerobic conditions. Cells were harvested at various times (lanes 1–4: OD_{600} 0.15, 0.3, 0.6 and 1.0), and subjected to quantitative Western blot analysis against anti-AllR (a) and anti-RpoA (b) antibodies. The intensity of immunostained bands was measured with an LAS-1000 Plus Lumino-image analyser and IMAGE GAUGE (Fuji Film).
once AllR is inactivated by glyoxylate, AllS is produced, which in turn activates the allDC operon for further enhancement of the initial stage of allantoin degradation for generation of carbamoyl phosphate and NH$_4^+$ for nitrogen assimilation (see Fig. 8a, b). Thus, the accumulation of glyoxylate, an intermediate product of the glyoxylate pathway of allantoin degradation, may induce switching to the alternative pathway for utilization of allantoin as a nitrogen source. In this regard, it is noteworthy that IclR, another member of the same regulator family, is also controlled by two antagonistic effectors (Lorca et al., 2007).

### Interplay between purine and pyrimidine degradation

RutR plays a key role in global regulation of the set of genes for synthesis of pyrimidine and arginine and degradation of both pyrimidines and purines (Shimada et al., 2007). By genomic SELEX screening of the RutR-binding sequences in the *E. coli* genome, we identified the RutR box within the gcl operon, which possibly affects its expression. The affinity of RutR decreases in the presence of pyrimidines, uracil or thymine (Shimada et al., 2007), while that of AllR increases in the presence of allantoin (Fig. 8). Thus, one interpretation is that the increase in pyrimidine pools

---

![Diagram](image_url)

**Fig. 8.** (a) Purine degradation pathway in *E. coli*; the enzymes (shown as gene product names) involved in each step of the pathways are described. (b) Transcriptional organization of the genes encoding the enzymes in the allantoin degradation pathways. The roles of transcription factors, AllR, AllS and RutR, and the effectors controlling their activities are described in the text.
activates the gcl operon for enhancement of allantoin degradation while the increase in the allantoin pool represses the gcl operon, thereby switching the catabolic pathway of allantoin to nitrogen assimilation. Taking these results together, we propose that the degradation of purines and pyrimidines is interrelated using the global regulators, AllR and RutR.

Conclusions

AllR is a global regulator of the genes for degradation of purines, downstream from allantoin, to yield glyoxylate for energy production or to yield an end product, NH$_3^+$, via carbamoylphosphate (the substrate for de novo synthesis of pyrimidines) for nitrogen reutilization. In this study, we identified a set of the regulation targets of AllR using genomic SELEX. RutR, the global regulator of pyrimidine degradation, was also suggested to control the genes of the purine degradation pathway. Together, these observations indicate a cross-talk in regulation between the genes for degradation of purines and pyrimidines and between the genes for nitrogen and carbon metabolism. Here, we determined, for the first time to our knowledge, the intracellular concentration of AllR and found that the AllR level stays constant during growth-phase transition from exponential to stationary phase and at different temperatures between 22 and 42 °C. Glyoxylate is known to be an inducer of AllR, but in addition we showed that allantoin functions as a co-repressor of AllR. Taking these findings together, we conclude that the level of functional AllR is controlled by regulating its activity through interaction with two effectors, the inducer glyoxylate and the co-repressor allantoin.

ACKNOWLEDGEMENTS

We thank Takenori Miki (Fukuoka Dental College) for the allR-disrupted mutant JW0494, and T. Shimada, J. Teramoto and K. Yamamoto (Hosei University) for discussions. This work was supported by Grants-in-Aid (17076016 and 18310133) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the Nano-Biology Project of the Micro–Nano Technology Research Center of Hosei University.

REFERENCES


Edited by: M. Schweizer