Structural and functional analyses of the repressor, RbsR, of the ribose operon of *Escherichia coli*

C.A. MAUZY AND M.A. HERMODSON

Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907

(RECEIVED November 18, 1991; REVISED MANUSCRIPT RECEIVED February 12, 1992)

Abstract

The DNA sequence encoding the *rbs* repressor protein, RbsR, has been determined. Amino acid sequence analyses of the product of an *rbsR-lacZ* fusion and of affinity-purified RbsR demonstrate that translation begins at an unusual codon, TTG, and that the initial amino acid is removed during maturation of the protein. DNA-binding assays indicate that RbsR binds to a region of perfect dyad symmetry spanning the *rbs* operon transcriptional start site and that the affinity for the *rbs* operator is reduced by addition of ribose, consistent with ribose being the inducer of the operon. RbsR is a member of a family of homologous repressor proteins having very similar DNA-binding sites and binding to similar operator sequences. [RbsR PIR accession number A41828.]

Keywords: *Escherichia coli*; repressors; ribose transport and utilization; transcription regulation

The *rbs* operon of *Escherichia coli* K12 encodes the high-affinity membrane transport system for ribose (Iida et al., 1984; Lopilato et al., 1984). Three open reading frames at the 5' end of the mRNA, designated *rbsD*, *rbsA*, and *rbsC* (Bell et al., 1986), are located in the region shown to be necessary for the membrane transport function (Iida et al., 1984; Lopilato et al., 1984). The coding region for the periplasmic ribose binding protein, *rbsB*, follows (Groarke et al., 1983), with that for ribokinase, *rbsK* (Hope et al., 1986), completing the set of genes that have been demonstrated to function as a single transcriptional unit (Lopilato et al., 1984).

The gene encoding the repressor for the *rbs* operon, *rbsR*, was shown to be located downstream from *rbsK*, and evidence indicated that it is on a separate transcriptional unit from *rbs* (Lopilato et al., 1984). During the sequence analysis, we found an open reading frame immediately downstream of *rbsK* that was clearly homologous to *lacI* and *galR* (Hope et al., 1986). We have now completed the sequence analysis of this open reading frame and have shown that it encodes RbsR, a member of a family of repressor proteins. Functional studies have confirmed that the protein is expressed and binds to DNA sequences upstream of *rbsD* that conform to the consensus operator sequences for this family of repressors.

These studies, together with the earlier genetic analyses (Lopilato et al., 1984), identify and locate the *rbsR* gene.

Results

Sequence of *rbsR*

The *rbsR* gene-coding sequence consists of 987 bp encoding a 329-amino acid protein of M_r 36,369 (Fig. 1). The Shine-Dalgarno site is located within the coding sequence of *rbsK*. The *rbsR* gene begins at the unusual start codon TTG located 3 bp downstream of the *rbsK* stop codon. The three additional amino acids from the predicted protein sequence produced with the TTG start codon are highly similar to other repressor sequences (Mauzy & Hermodson, 1992). An in-frame ATG codon is present 13 bp downstream of *rbsK*. However, the length of DNA between this codon and the predicted Shine-Dalgarno site is not optimal for ribosome binding (Gold et al., 1981). The *rbsR* gene sequence contains two strong dyad symmetries located 71 and 131 bp downstream from the TTG start, which have been implicated as transcriptional termination sites by nuclease S1 mapping (Hope, 1986).

In vitro transcription/translation of *rbsR*

To verify that the open reading frame encodes a protein of the predicted molecular weight, various plasmids con-

Reprint requests to: M.A. Hermodson, Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907.

Fig. 1. DNA and deduced protein sequences of the *rbsR* gene. Top numbers specify the DNA sequence position, whereas the lower numbers indicate protein sequence position. The first base of the *rbsR* sequence is located at +4,906 and the last base at +5,895 relative to the *rbs* operon transcriptional start site. The Shine-Dalgarno site for *rbsR* is boxed. The two regions of dyad symmetry are marked by horizontal arrows.

taining the proposed *rbsR* gene were used in an in vitro DNA-directed transcription/translation system incorporating [³⁵S]methionine into the newly synthesized protein. The plasmids pCMR150, pAO3, and pJG2 (Fig. 2), which contain the proposed *rbsR* gene, produced a protein of approximately *M_r* 36,000, consistent with that predicted for RbsR. This protein was not produced using pNR2, which does not contain the *rbsR* gene. The plasmid pCMR150, with *rbsR* under the control of the strong *lac* promoter, produced higher quantities of *rbsR* protein than either pAO3 or pJG2, which contain *rbsR* under the control of the presumably weaker *rbs* promoter.

Determination of the translational start codon of *rbsR*

To determine the start codon of the *rbsR* gene, a *rbsR-lacZ* fusion was constructed to produce a fusion protein containing the N-terminal 266 amino acids from RbsR

fused to 1,016 residues of β -galactosidase. The fused *rbsR-lacZ* gene was placed under transcriptional control of the *lac* promoter. This construct, called pRF3lacZ, was placed in the strain MC4100 (*lacZ*⁻) for production and immunoaffinity purification of the protein. Examination of the eluate from the affinity column by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) showed that the solution contained two distinct protein species and lower levels of various contaminants (data not shown). Both the higher and the lower molecular weight bands were submitted for protein sequencing. Six cycles of Edman degradation of the higher molecular weight species gave the sequence Ala-Thr-Met-Lys-Asp-Val, which are the first seven amino acids, minus the initiator methionine, of the predicted sequence of RbsR using the start codon TTG. Posttranslational removal of the N-terminal methionine is also seen in other repressor proteins (Rolfes & Zalkin, 1990). The lower molecular weight species, which represented approximately 80% of

Fig. 2. In vitro transcription/translation of the *rbsR* gene. Autoradiograph of the products of the in vitro transcription/translation. The gene products as encoded from the plasmid DNA are indicated at left. The gene locations of the constructs used for in vitro transcription/translation are indicated across the top. The plasmids pUC18 and pBR322 were added as controls for nonspecific band formation.

the eluted protein, gave the sequence Val-Gly-Val-Tyr-Gln-Ala-Leu. This sequence follows the internal Met-248. In this truncated RF3lacZ protein, both the first and second amino acids, Met and Ala, are posttranslationally removed if one assumes that initiation of translation occurred at Met-248. Alternatively and more likely, this species could be a proteolytic degradation product of the RbsR- β -galactosidase fusion protein.

Affinity purification of RbsR

Obtaining significant quantities of pure RbsR proved difficult, possibly due to the unusual start codon and the sites of partial transcription termination early in the coding sequence (see the Discussion). Late in this study enough RbsR to obtain an N-terminal sequence was purified by affinity chromatography using biotinylated *rbsO* DNA attached to biotinylated cellulose by means of a streptavidin bridge (Chodosh et al., 1986). Examination of the eluate from the *rbsO* affinity column (Fig. 3) revealed the presence of a protein band at the expected molecular weight of RbsR and a higher molecular weight protein, estimated to be less than 5% of the total protein. A similar contaminating protein was also seen in the purification of RhaR, the rhamnose repressor (Tobin & Schleif, 1990), and was shown to lack the ability to bind operator DNA. The major band from the SDS-PAGE (Fig. 3) was blotted onto a polyvinylidene difluoride membrane and subjected to N-terminal sequence analysis. It gave the same Ala-Thr-Met-Lys-Leu-Val sequence

as the larger of the two β -galactosidase fusion products (above).

Identification of RbsR binding sites

Three possible sites of transcriptional control of *rbsR* are at the *rbs* promoter, at a possible secondary promoter region between *rbsB* and *rbsK* (Hope et al., 1986), and at or near the start of the *rbsR* gene for possible autoregulation. Each of these regions of DNA sequence contains a sequence homologous to the consensus operator sequence (Fig. 4). The strongest candidate for an *rbsR* operator site, based on the sequence and location within the operon, was located at bp -2 to +18 relative to the *rbs* transcriptional start site (Bell et al., 1986).

To investigate the transcriptional control of the operon by RbsR, the *DdeI*-*BanI*, *Bss*HII-*Bst*EII, and *NheI*-*Nsp*(7524)V fragments from the *rbs* operon (Fig. 5) were examined for possible RbsR binding sites using the mobility shift DNA-binding assay. The *DdeI*-*BanI* fragment is located at bp -24 to +39, starting at the *rbs* promoter -10 site and ending 10 bp within the *rbsD* coding sequence. The *Bss*HII-*Bst*EII fragment is located at bp +3,626 to +4,070 of the *rbs* operon (Bell et al., 1986) and includes all of the intercistronic sequence between *rbsB* and *rbsK*. The *NheI*-*Nsp*(7524)V fragment (bp +4,664 to +5,159) includes the 3' end of *rbsK* and the first 254 bp of *rbsR*. This region contains several sites that are not strong consensus sequences. To increase the expression of RbsR, the *rbsR* gene was placed into both pBR322

Fig. 3. Purification of RbsR via an *rbs* operator affinity column. Biotinylated operator fragment from pCM01 was attached to biotin-cellulose via a streptavidin linkage. A crude protein extract prepared from pCMR100/MC4100 cells was passed three times over the *rbs* operator DNA affinity column equilibrated in 4 mM Tris, pH 7.9, 12 mM HEPES, pH 7.9, 60 mM KCl, 1 mM EDTA, and 1 mM DTT. The protein was eluted using 20 mM Tris, pH 7.9, 1 M KCl, 5 mM MgCl₂, 1 mM EDTA, 1 mM DTT, and 40 mM ribose and collected in 0.5-mL fractions. Thirty microliters of each fraction was analyzed using 15% SDS-PAGE. The proteins were stained with Coomassie blue.

(pCMR100) and into pUC18 (pCMR150). The strain MC4100 was transformed with these plasmids, and the crude extract produced was used for most of the binding studies.

RbsR does not bind to either the pCM03 fragment or the pCM04 fragment in the absence or presence of D-ribose (data not shown). However, it was shown that RbsR does bind to the pCM01 fragment (Fig. 6). Using the pCMR100/MC4100 extract, the interaction is specific; there was no binding artifact from the vector,

pBR322. The vector pUC18 did generate some background binding (pUC18/MC4100 lane). However, the addition of the *rbsR* gene to pUC18 increases binding at least 10-fold. Crude extracts made from MC4100 (*rbsR*⁻) or from the wild-type strain W3110 do not demonstrate binding, even though the genomic copy of *rbsR* in W3110 was expected to produce a slight band shift. To examine the expression and binding of RbsR in single copy, extract was prepared from MRi121 (characterized as *rbsR*⁺) grown in M9 medium containing either glucose, glycerol, D-ribose + glucose, or D-ribose + glycerol. No binding was seen from any of these extracts (data not shown).

The binding affinity of RbsR for the pCM01 fragment decreases upon the addition of D-ribose (Fig. 7). Other sugars such as ribose-5-phosphate, ribose-1-phosphate (Fig. 7), galactose, arabinose, L-ribose, and allose (not shown) do not affect RbsR binding to the proposed operator sequence found in pCM01. To demonstrate that this interaction was specific for the sequence found in the pCM01 fragment, a 146-bp fragment of DNA was isolated from pBR322 and used as a potential DNA-binding competitor. The addition of increasing amounts of unlabeled pBR322 fragment had no effect upon the binding to the pCM01 fragment (data not shown). Competition between unlabeled and labeled pCM01 fragment was demonstrated, indicative of the specific nature of the binding of RbsR to the pCM01 fragment.

Protection and interaction by RbsR to the operator region

To verify that the predicted operator sequence was *rbsO*, DNase footprinting was used to determine the region of DNA protected by RbsR (Fig. 8). To optimize the DNA fragment size for the ladder production, a slightly larger fragment containing *rbsO* from pCM01ESx was used.

Fig. 4. The *rbsO* sequence alignment with other known operator sequences. The operator sequences are shown and the respective control proteins are given in parentheses. The operator consensus sequence is indicated, and the regions of dyad symmetry are underlined. This figure is a modification of Rolfe and Zalkin (1988) with the addition of *rbsO*, *malO₁*, and *malO₂* (Reidl et al., 1989). The protected and partially protected guanines in *rbsO* are marked by arrows.

Fig. 5. DNA fragments used in the DNA mobility shift assays. All fragments are shown in their respective locations in the *rbs* operon.

The *Eco*RI-labeled fragment, the coding strand for the operon, demonstrates protection at bp +2 to +14 relative to the *rbs* operon transcriptional start site (bp 3–15 in the 5' direction of the dyad symmetry). Due to incomplete DNase digestion, it is unclear whether the second base of the dyad symmetry, a cytidine, forms part of the protected region. The *Hind*III fragment displayed a protected region from +5 to +19 bp (bp 1–15 in the 5' direction of the dyad symmetry). On this strand, the 16th base, a guanine, may be protected. When the protected bases are examined in a B-form DNA structure, the protected strands form 1.3 (*Eco*RI) and 1.5 (*Hind*III) helical turns. The protected double-stranded DNA forms a

staggered edge with the ends of the protected strands matching on one face of the DNA, each end on either side of the minor groove.

The specific binding contacts of RbsR to the operator fragment were also examined. In this assay, a methylated guanine that interferes with the repressor contact of the DNA binding site will not produce a repressor–operator complex, and the DNA fragment will remain in the unbound form. The site at which methylation interferes with binding will show up as a gap in the repressor-bound DNA ladder. Isolation and separation of the DNA products from both the RbsR bound and unbound fragments reveal that the second guanine from the 5' end of the operator sequence is missing from the sequence (Fig. 4) in the RbsR-bound form on both strands (Fig. 9A,B). The third guanine band is less intense than that found in the unbound form. When these two positions are examined in a DNA helical structure, one fully protected guanine from one strand and one partially protected guanine from the other strand define a binding region on one face of the DNA helix (Fig. 9C). The two sites are found on either side of the center of the dyad symmetry and are on the opposite face of the DNA relative to the ends of the DNase I-protected region.

Discussion

Early work (David & Weismeyer, 1970; Abou-Sabé et al., 1982) indicated that the control of ribose transport in *E. coli* was relatively complicated and strain dependent. An inducible system was demonstrated in *E. coli* K12 (strain X289) (David & Weismeyer, 1970), whereas evidence was obtained for both inducible and constitutive transport systems in *rbs*⁺ strains of *E. coli* B/r (Abou-Sabé et al., 1982). Later work by Lopilato et al. (1984) further defined the existence and role of a control protein for ribose high-affinity transport. A *lacZ* fusion to an *rbs* gene (most likely *rbsB*) was constructed in a lambda phage system. Studies using the *lac*–*rbs* fusion gene demonstrated that the *rbs* operon is inducible by D-ribose. All of the fusion experiments were performed in 13 mM ribose, however, and no attempt was made to determine the in vivo ribose concentration at which induction occurs. The in vitro band shift assays performed in this study examined both the location of the *rbs* operator site and the repressor binding in the presence of the probable inducer molecule, D-ribose. It was shown that the *rbs* repressor binds to a DNA sequence having perfect dyad symmetry located 6 bp downstream of the –10 promoter site and 12 bp upstream of the *rbsD* start. Control binding studies using a nonspecific pBR322 DNA fragment and operator fragment competition experiments verified that the band shift was due to a specific interaction with the operator fragment. The protein extract used in the binding studies was obtained from strains carrying the *rbsR* gene in either pBR322 or pUC18. We were unable to demonstrate any

Fig. 6. Mobility shift assay of the pCM01 fragment. Labeled fragment (0.5 ng) was incubated with the indicated cell extracts (5 µg of protein) for 15 min at 30 °C. Reduced mobility of the labeled fragment due to protein binding is observed in the right three lanes. See text for further interpretation of the results.

Fig. 7. Mobility shift assay of the pCM01 fragment in the presence of sugars. Assay conditions were the same as in Figure 6 for the pCMR100/MC4100 extract reaction.

Fig. 8. DNase footprinting of the *rbs* operator region. An *EcoRI*-*HindIII* DNA fragment from pCM01ESx was end-labeled at either the *EcoRI* or the *HindIII* sites using T4 polynucleotide kinase and selective restriction digestion. The labeled fragment was bound under the conditions determined by the mobility shift assay using an extract containing partially purified RbsR. The DNA-RbsR complex was subjected to a 1-min digestion at room temperature using 0.25 ng/ μ L of DNase I. The lane numbers indicate microliters of RbsR extract used per binding reaction. The standards are the Maxam-Gilbert G and AG sequencing reactions that run 1.5 base lengths faster than the equivalent DNase fragment. The perfect dyad symmetry regions are boxed in gray. The regions of protection are indicated by a box, whereas areas with unclear protection are marked with a stippled box. **Left:** DNase footprint of the *HindIII*-labeled operator fragment. **Right:** DNase footprint of the *EcoRI*-labeled operator fragment (coding strand).

Fig. 9. Methylation interference assay using *rbsO*. The DNA fragments generated from the methylation interference were separated using a 16% acrylamide urea gel. The Maxam-Gilbert standards of each strand are indicated as G, AG, and AC. The RbsR-bound DNA is designated B, whereas the unbound form is labeled UB. The partial and full protection positions are marked. **A:** Methylation interference data using labeled *HindIII* fragment. **B:** Methylation interference data using labeled *EcoRI* fragment (coding strand). **C:** Position of the full and partially protected guanines on B-form DNA. Large arrows indicate the center of the dyad symmetry of *rbsO*.

low-level binding using strains containing only single-copy *rbsR*. To determine if the lack of binding was due to an uncharacterized *rbsR* strain or to low RbsR expression, extracts were prepared from the strain MRi121 (demonstrated to be *rbsR*⁺ [Lopilato et al., 1984]) grown in sugars to control the induction by ribose or to control the catabolite repression. No binding was seen using any of the extracts; thus, binding could not be observed using genomically encoded *rbsR*.

The in vitro shift assays presented here using *rbsR*⁺ extract and D-ribose supported previous observations of the control by D-ribose upon ribose transport (Lopilato et al., 1984). Purified RbsR was shown to bind to the operator fragment (data not shown), providing evidence for the direct interaction of the repressor molecule with bind-

ing at the proposed operator site and indicating that DNA-binding activity of RbsR did not require a secondary protein.

Control of expression of RbsR is not well understood. The *rbsR* gene was previously proposed to form a separate transcriptional unit from the rest of the *rbs* operon (Lopilato et al., 1984). Analysis of the *rbsK* coding sequence revealed no obvious promoter sequence for *rbsR*. Nuclease S1 mapping of the 3' end of the *rbs* operon (Hope, 1986) revealed two transcriptional termination positions at hairpins 1 and 2 (Fig. 1) located in *rbsR*. A longer transcript was detectable, however, suggesting *rbsR* expression from the same transcript as *rbsK* with down-regulation at the hairpins.

The proximity of *rbsR* to *rbsK* and the placement of

the ribosome binding site within the *rbsK* gene also suggests that the *rbsR* gene is located on the same transcript as *rbsK*. Because RbsR did not bind to the region of DNA spanning the 3' end of *rbsK* and the first 260 bp of *rbsR* or to a potential secondary promoter region between *rbsB* and *rbsK*, regulated expression of *rbsR* from those regions appears unlikely, although constitutive expression from a secondary promoter upstream of *rbsR* has not been excluded. Another interesting feature of RbsR expression is the presence of a TTG start codon in the coding sequence, which was verified using the protein sequence from an RbsR- β -galactosidase fusion protein and from affinity-purified RbsR. Of all known start codons in the Genbank sequences, the TTG has a frequency of less than 2%. The unusual start codon for *rbsR* also suggests down-regulation of RbsR expression due to lower translational efficiency, as evidence indicates that there is a strong preference for the ATG start codon in the bacterial translational machinery (Munson et al., 1984; Schulz & Reznikoff, 1990).

Although a small amount of RbsR was purified by affinity chromatography late in this study (Fig. 3), we have had some difficulty producing significant amounts of the protein. The several sites at which down-regulation of expression may occur may be partly responsible for the low expression from the various plasmids. Toxicity of high levels of the repressor may also be a problem. Because we only had crude extracts available during most of this investigation, the DNA-binding studies were performed with them. The results, together with the control experiments run simultaneously, indicate that RbsR does bind *rbsO*, and of the sugars tested, only ribose was effective in disrupting that complex. No ancillary factors were needed to form the RbsR-*rbsO* complex when pure RbsR was tested. At least under the conditions used, no other potential RbsR binding sites in the *rbs* operon were found, although the use of crude extracts may have interfered with weak binding at other sites. We believe that to be less than likely on physiological grounds. If ribokinase is noncoordinately expressed relative to the *rbs* transport genes as early data suggest (David & Weismeyer, 1970), then a constitutive promoter within the operon is the most likely explanation. We have attempted to find evidence for transcription from such a promoter without success (Hope, 1986).

Analysis employing crude extracts prevents the determination of binding constants for the RbsR-*rbsO* complex and for the RbsR-sugar interaction. However, the footprinting assays should not be affected, since enormous excesses of competitor DNA are used in such assays.

By homology to LacI (Mauzy & Hermodson, 1992) the first 60 residues of RbsR form the DNA-binding domain. The highest degree of similarity between the members of this family of repressors is found in this domain, and the most similar region of all spans the first 22 residues of

RbsR. These residues form the first helix-turn-helix and contact the DNA (Kaptein et al., 1985; Gordon et al., 1988; Rolfes & Zalkin, 1988). The degree of similarity is so striking in both the protein sequences (Fig. 1; Mauzy & Hermodson, 1992) and the operator sequences (Fig. 6) that the question of how specificity is attained must be considered. Mutagenesis studies and analysis of amino acid sequence conservation within this family of proteins consistently indicate that the hydrophobic core residues that hold the bundle of helices together are very intolerant to change (Gordon et al., 1988; Rolfes & Zalkin, 1988). The hydrophobic core residues in RbsR would be M-4, V-7, A-8, V-17, V-20, I-21, A-38, I-39, L-42, and Y-44 by homology to LacI (Fig. 1; Kaptein et al., 1985; Gordon et al., 1988; Mauzy & Hermodson, 1992). These positions are either invariant between members of this protein family or show only conservative changes. This no doubt reflects the crucial roles these residues play in stabilizing the domain structure. Because the side chains of these residues project into the center of the domain, they would not be involved in DNA contacts. It has been suggested that other conserved residues in the helix-1-turn-helix-2 region could function as contacts to the conserved bases in the operators (Rolfes & Zalkin, 1988). These would include T-3, D-6, A-11, S-13, T-16, and S-18. That leaves few positions available to confer specificity, especially in helix-2, which presumably makes the principal DNA contacts (Kaptein et al., 1985; Gordon et al., 1988). T-14, S-15, H-19, and N-22 are the amino acids in RbsR at those positions, and an identical amino acid is found at each of those positions in at least one of the other repressors (RbsR and PurR differ only at position 15 in this region where the former has Ser and the latter Thr, yet the operator sequences for *rbs* and *pur* are about as different as any in this series [Fig. 6]). Thus, the explanation of the functional significance of these relationships must await more detailed structural analyses.

The proposed turn region between helix-1 and helix-2 in RbsR is truncated by one residue as compared to the other members of this repressor family (Fig. 1; Mauzy & Hermodson, 1992). Initial studies by Pabo and Sauer (1984) suggested that a glycine was critical at this position. However, later mutational studies (Gordon et al., 1988) and additional protein sequences of other members of this repressor family imply that there is much less sequence constraint at this position. Indeed, in the case of RbsR, the absence of one residue in this position indicates that this region does not require a three-residue turn. The residue following this deletion in RbsR is a Leu where it is Val in the other members of the protein family. The absence of a branch at the β -carbon of the Leu may confer the required additional flexibility to allow a shorter turn in RbsR, although the very short two-residue turn is bound to affect the conformations of the amino acids leading into the helices at that site.

Materials and methods

Materials

5'-[α - 35 S]Thio-dATP, 14 C-methylated protein molecular weight standards, and L-[35 S]methionine were obtained from Amersham. Immobilon was purchased from Promega. ProBlott membrane was obtained from Applied Biosystems. All restriction endonucleases and enzyme reagents used in the cloning procedures were obtained from commercial sources. *Eco*RI linkers were obtained from New England Biolabs, and the NuSieve GTG agarose was obtained from FMC BioProducts. The GeneClean DNA purification kit was purchased from BIO101, Inc. The Sequenase DNA sequencing kit was obtained from United States Biochemicals. Oligonucleotide purification cartridges (OPC) were obtained from Applied Biosystems, and the prokaryotic DNA-directed translation kit was obtained from Amersham. En³hance was obtained from New England Nuclear, and the BCA (bicinchoninic acid) assay used in the protein concentration determination was purchased from Pierce Chemical Company.

Strains and plasmids

Escherichia coli strains TG1, HB101, and JM101 (Sambrook et al., 1989) were obtained from laboratory stocks as was strain MC4100 (Silhavy et al., 1984). Wild-type strain W3110 (Bachman, 1987) was a gift from R. Somerville. The plasmid pJG2 and the strain MRi121 (Lopilato et al., 1984) were gifts from J. Beckwith, Harvard Medical School. The phagemid pUC118 was a gift from H. Zalkin. The *lacZ* fusion plasmid pMC1871 (available from Pharmacia) was a gift from K. Herrmann. The plasmids pUC18, pUC19, and pBR322 and the phage M13mp18 and M13mp19 were obtained from laboratory stocks.

Cell transformation, media, and growth conditions

Competent cells were prepared by the method of Hanahan (1983) using SOB and SOC media as described. The transformed cells were plated onto 2xYT plates (Miller, 1972) containing the appropriate antibiotic. Cells transformed with M13 phage were plated as described by Tabor and Richardson (1989). Unless otherwise noted, all bacterial cells in this study were grown in 2xYT medium at 37 °C.

DNA purification

Large and small plasmid preparations were done in accordance with Birnboim (1983). Plasmid DNA used in

the in vitro transcription/translation experiment was further purified using the GeneClean kit. M13 phage purification for both replicative and single-stranded forms is described by Tabor and Richardson (1989). All DNA fragment isolations were performed with agarose gels and the GeneClean kit unless otherwise noted.

Construction of plasmids

All ligations were performed overnight at 15 °C, and the ligation solution was used to transform competent TG1 or JM101 cells. The resulting transformants were screened by restriction analysis. All restriction digests were performed according to the manufacturer's instructions.

The *Nhe*I-*Sal*I fragment containing *rbsR* from pJG2 was ligated to the *Nhe*I-*Sal*I sites in pBR322 to create the plasmid pCMR100. The identical *Nhe*I-*Sal*I *rbsR* fragment was removed from purified pCMR100 and ligated to the *Xba*I-*Sal*I sites in pUC18 to create the plasmid pCMR150.

The 6,794-bp *Nde*I-*Sal*I fragment from pJG2 was isolated and ligated into the *Nde*I-*Sal*I sites of pUC19 to create the plasmid pAO1. The plasmid pAO3 was constructed as follows: pAO1 was digested with *Sal*I and blunt-ended by Klenow addition of dNTPs. *Eco*RI linkers were ligated to the blunt ends, excess linker removed by *Eco*RI recut, and the ends reclosed. Thus, pAO3 contains an *Eco*RI site in place of the *Sal*I site.

To create pNR2, the plasmid pAO3 was digested with *Nsp*(7524)V and the ends filled in using dNTPs and Klenow enzyme. *Eco*RI linkers were ligated to the blunted ends, and excess linker was removed by *Eco*RI digest. The resultant linear DNA was isolated and the *Eco*RI ends were ligated.

The 60-bp *Dde*I-*Ban*I fragment from pAO1 was isolated, blunt-ended by dNTP addition using Klenow enzyme, and ligated to a phosphatased *Sma*I site in pUC18 to create pCM01. The 444-bp *Bss*HII-*Bst*EII fragment of pNR2 was isolated, blunt-ended by Klenow, and ligated to a phosphatased *Sma*I site in pUC118 to create pCM03. The plasmid pCM04 contains the 496-bp *Nhe*I-*Nsp*(7524)V fragment inserted at the *Sma*I site in pUC18. All three constructs containing inserts used for the band shift studies were verified by plasmid sequencing (Kraft et al., 1988).

The *Sma*I-*Sal*I fragment of the *lacZ* fusion vector pMC1871 was isolated using agarose gel separation and GeneClean purification and was inserted into the *Eco*RV-*Sal*I sites of pCMR150 to create pRF3lacZ.

To create pCM01ESx, the *Eco*RI-*Sal*I fragment from pCM01 was isolated and blunt-ended using dNTPs and Klenow fragment. *Nco*I linkers were ligated to the fragment for 4 h at room temperature, and the modified fragment was cloned into the *Sma*I site of pUC19.

Construction of phage

The *NheI*–*SalI* fragment from pCMR100 was ligated to the *XbaI*–*SalI* sites in M13mp18 and M13mp19 to create pCMR2mp18 and pCMR2mp19, respectively.

Oligonucleotide synthesis and purification

All sequence-specific primers were synthesized using solid phase-based phosphoramidite chemistry performed on a 380A DNA Synthesizer from Applied Biosystems. The crude trityl-on oligomers were purified using OPC as described by the manufacturer.

DNA sequence analysis

DNA sequences were obtained by the dideoxy sequencing method (Sanger et al., 1977) with 5′-[α -³⁵S]thio-dATP using both M13 deletion constructs and sequence-specific primers. Both strands of the DNA were completely sequenced. The DNA was analyzed on 8% polyacrylamide urea gels poured with 0.2–1-mm wedge spacers.

In vitro transcription/translation

Approximately 5 μ g of each plasmid was analyzed using the DNA-directed transcription/translation system, which utilizes an S30 *E. coli* extract. Synthesized proteins were labeled during in vitro translation by the addition of L-[³⁵S]methionine. A 5-min cold methionine chase, per the manufacturer's directions, was performed. The reaction mix was diluted 1:1 with loading buffer and boiled 6 min to denature the proteins. Twelve microliters, or approximately 5×10^5 dpm, of each sample was analyzed on a 15% SDS polyacrylamide gel (Laemmli, 1970). After electrophoresis, the gel was fixed in 10% acetic acid/30% methanol for 1 h, soaked in En³hance for 1 h, and then rehydrated for 1 h in water. After drying, the gel was autoradiographed 9 h at room temperature.

Mobility shift DNA-binding assay

EcoRI–*SalI* digests of pCM01, pCM03, and pCM04 were phosphatased for 8 h at 37 °C. The DNA solutions were run separately over 6% acrylamide gels, and the respective insert bands were cut from the gel. The DNA was electroeluted from the gel slices and the recovered DNA solution was ethanol precipitated. The isolated insert DNA was labeled by T4 polynucleotide kinase using [γ -³²P]ATP (Ausubel et al., 1987–1989). To prepare protein extract, the appropriate strain was grown at 37 °C to Klett 150. The cells were harvested at low speed, washed in 20 mL cold 20 mM Tris, pH 7.5, and again harvested.

The resultant cell pellet was resuspended in 2.5 mL BS buffer (12 mM HEPES, pH 7.9, 4 mM Tris, pH 7.9, 60 mM KCl, 1 mM EDTA, 1 mM fresh dithiothreitol [DTT]). The cell suspension was lysed by two passes through a French press at 20,000 psi and centrifuged at $12,000 \times g$ for 30 min. The supernatant was used as the crude extract. A small example of the extract was analyzed for protein concentration using the BCA assay. Sterile glycerol was added to 20%, and the extract was divided into aliquots and frozen at –70 °C for later use.

To initiate binding, 5–10 μ g of protein extract, 1.5 μ g poly(dI–dC)·poly(dI–dC) carrier DNA, 1 μ g bovine serum albumin (BSA), and approximately 0.5 ng ³²P-labeled fragment were placed in BS buffer to a total of 10 μ L. The binding reactions were performed at 30 °C for 15 min and were loaded onto 4% low-ionic-strength polyacrylamide gels with recirculating buffer (Ausubel et al., 1987–1989). After electrophoresis, the gels were dried and autoradiographed 1–3 h at room temperature.

Methylation interference assay using *rbsO*

To obtain interference data on both operator strands, pCM01ESx was digested with either *EcoRI*–*PvuII* or *HindIII*–*PvuII*. Both digestion solutions were treated with alkaline phosphatase at 37 °C overnight, and the respective fragments were isolated using 7% acrylamide gels and electroelution. The precipitated DNA fragments were resuspended in kinase buffer (Ausubel et al., 1987–1989) and labeled 6 h with [γ -³²P]ATP using T4 polynucleotide kinase. After labeling, the solutions were heated at 75 °C for 15 min to remove the kinase activity. The *EcoRI*–*PvuII* fragment was digested with *HindIII* and the *HindIII*–*PvuII* fragment was digested with *EcoRI*. The labeled fragments containing the operator fragment were again isolated using 7% acrylamide gels and electroelution.

For RbsR–operator attachment studies, a crude protein extract was prepared from pCMR100 in the strain MC4100 as previously discussed, with the exception that the cell pellet was resuspended in cold ES buffer (10 mM Tris, pH 7.9, 1 mM EDTA). After centrifugation, the supernatant was subjected to a 30% ammonium sulfate precipitation and the resulting pellet was resuspended in cold ES buffer. To desalt the protein solution, the extract was loaded onto a G-25 Sephadex column (2.5 cm $\times$ 5.5 cm) equilibrated with ES buffer. The 600-nm absorbance of each fraction was determined, and the appropriate protein peaks were pooled. A small sample was saved and was subjected to a band shift assay to confirm operator binding abilities. The remaining solution was analyzed for protein concentration and the necessary components were added to change the ES buffer to BS buffer. Sterile glycerol was added to 10%, and the extract was placed in storage at –70 °C.

The methylation interference assay (Hendrickson & Schleif, 1985; Ausubel et al., 1987–1989) was modified slightly by employing 10 mM Tris, pH 8.0, 1.5 M NaCl, and 1 mM EDTA as the DEAE elution buffer. The elution buffer was added to the DEAE paper and incubated overnight at 37 °C to release the DNA. DNA sequence standards of both strands were prepared using the chemical sequencing method of Maxam and Gilbert (1980).

DNase footprinting of rbsO

DNase footprinting was done according to the published protocol (Ausubel et al., 1987–1989). Preparation of the labeled operator fragment and the protein extract was done in the same manner as for the methylation interference assay.

RbsR fusion protein purification and N-terminal sequencing

Escherichia coli strain MC4100 was transformed with pRF3lacZ, and the cells were grown to saturation in 2xYT medium. The cells were harvested and frozen overnight at –20 °C, then thawed on ice and resuspended in 10 mL of cold TEP buffer (100 mM Tris, pH 7.4, 10 mM EDTA, and 1 mM phenylmethylsulfonylfluoride [PMSF]). The cells were lysed by three passes through a French press, and the debris was pelleted by ultracentrifugation at 50,000 × *g* at 5 °C. The supernatant was removed, tested for protein concentration using the BCA protein assay, and diluted to 4 mg/mL with cold 50 mM Tris, pH 7.3. The protein solution was then purified using a ProtoSorb lacZ immunoaffinity column. Using this method, 1.3 mg of protein was eluted from the column.

To purify the protein further for sequencing, 94 µg of the eluted protein solution was run on a 7% SDS polyacrylamide gel. The gel was transferred to Immobilon membrane in CAPS (3-[cyclohexylamino]-1-propanesulfonic acid) buffer (Matsudaira, 1987). After electrophoresis at 70 V for 30 min, the membrane was removed, stained in fresh Coomassie blue stain, and destained in 50% methanol. The membrane was dried between two clean pieces of Whatman 3MM chromatography paper, and the appropriate bands were removed and submitted for protein sequencing. Protein sequencing was accomplished using a gas-phase Applied Biosystems 473 Protein Sequencer.

Purification of RbsR

The biotin-streptavidin *rbsO* affinity column (Ausubel et al., 1987–1989) was prepared using approximately 500 µg of the plasmid pCM01 digested with *Eco*RI and

*Hinc*II. The *Eco*RI end was blunt-ended using Klenow fragment and bio-11-dUTP, and the resulting biotinylated *rbsO* fragment was isolated using acrylamide gel electrophoresis and electroelution. The DNA pellet was resuspended in BS buffer. A crude protein extract was prepared using the plasmid pCMR100 in the strain MC4100. The pCMR100/MC4100 cells were grown to saturation in 500 mL of 2xYT medium containing ampicillin, harvested, and resuspended in 10 mL of BS buffer. The cell suspension was passed through a French press three times at 20,000 psi, and the solution was centrifuged at 12,000 × *g* for 30 min. The supernatant was removed and the protein concentration of the solution determined via BCA assay. The solution was diluted to 5 mg/mL using cold BS buffer. A binding reaction was performed using the isolated biotinylated *rbsO* fragment and the crude extract. After incubation at 30 °C for 30 min, a fivefold excess of streptavidin was added. The column was washed at room temperature using 30 mL of BS buffer, and the RbsR was eluted using 15 mL of BS buffer containing 40 mM D-ribose.

N-terminal sequencing of purified RbsR

A 300-µL sample of the peak fraction from the affinity column was loaded onto a 10% SDS-PAGE gel. The gel was washed and transferred in CAPS buffer using the same method as before with the exception that ProBlott membrane was used. After transfer, the membrane was prepared as before, and the appropriate band was removed and submitted for N-terminal sequencing.

Acknowledgments

This work was supported by PHS grant GM24602 and NSF grant DMB 9003012. Protein sequence analyses and oligonucleotide syntheses were performed by the Laboratory for Macromolecular Structure, supported by the Diabetes Research and Training Center (PHS grant P60DK20542). We thank Prof. Steven Broyles and Dr. Ronda Rolfes for assistance with several procedures. We thank Jo Crain for expert assistance with the assembly of the manuscript. This is journal paper number 13234 of the Purdue Agricultural Experiment Station.

References

- Abou-Sabé, M., Pilla, J., Hazuda, D., & Ninfa, A. (1982). Evolution of the D-ribose operon of *E. coli* B/r. *J. Bacteriol.* 150, 762–769.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., & Struhl, K., Eds. (1987–1989). *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley Interscience, New York.
- Bachman, B.J. (1987). Linkage map of *Escherichia coli* K-12. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, 7th Ed. (Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., & Umberger, H.E., Eds.), pp. 807–876. American Society for Microbiology, Washington, D.C.
- Bell, A.W., Buckel, S.D., Groarke, J.M., Hope, J.N., Kingsley, D.H.,

- & Hermodson, M.A. (1986). The nucleotide sequences of the *rbsD*, *rbsA*, and *rbsC* genes of *E. coli* K12. *J. Biol. Chem.* 61, 7652–7658.
- Birnboim, H.C. (1983). A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.* 100, 243–255.
- Chodosh, L.A., Carthew, R.W., & Sharp, P.A. (1986). A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* 6, 4723–4733.
- David, J. & Wiesmeyer, H. (1970). Regulation of ribose metabolism in *Escherichia coli*. I. The ribose catabolic pathway. *Biochim. Biophys. Acta* 208, 45–55.
- Gold, L., Pribnow, D., Schneider, T., Schinedling, S., Singer, B.S., & Stormo, G. (1981). Translational initiation in prokaryotes. *Annu. Rev. Microbiol.* 35, 365–403.
- Gordon, A.J.E., Burns, P.A., Fix, D.F., Yatagai, F., Allen, F.L., Horsfall, M.J., Halliday, J.A., Gray, J., Bernelot-Moens, C., & Glickman, B.W. (1988). Missense mutation in the *lacI* gene of *Escherichia coli* inferences on the structure of the repressor protein. *J. Mol. Biol.* 200, 239–251.
- Groarke, J.M., Mahoney, W.C., Hope, J.N., Furlong, C.E., Robb, F.T., Zalkin, H., & Hermodson, M.A. (1983). The amino acid sequence of D-ribose-binding protein from *Escherichia coli* K12. *J. Biol. Chem.* 258, 12952–12956.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557–580.
- Hendrickson, W. & Schleif, R. (1985). A dimer of AraC protein contacts three adjacent major groove regions of the *araI* DNA site. *Proc. Natl. Acad. Sci. USA* 82, 3129–3133.
- Hope, J.N. (1986). Primary structures of ribose binding protein and of ribokinase and regulation of the RBS operon of *Escherichia coli* K12. Ph.D. Thesis, Purdue University, West Lafayette, Indiana.
- Hope, J.N., Bell, A.W., Hermodson, M.A., & Groarke, J.M. (1986). Ribokinase from *Escherichia coli* K12. *J. Biol. Chem.* 261, 7663–7668.
- Iida, A., Harayama, S., Iino, T., & Hazelbauer, G.L. (1984). Molecular cloning and characterization of genes required for ribose transport and utilization in *E. coli* K-12. *J. Bacteriol.* 158, 674–682.
- Kaptein, R., Zuiderweg, E.R.P., Scheek, R.M., Boelens, R., & van Gunsteren, W.F. (1985). A protein structure from nuclear magnetic resonance data: *lac* repressor headpiece. *J. Mol. Biol.* 182, 179–182.
- Kraft, R., Tardiff, J., Krauter, K.S., & Leinwand, L.A. (1988). Using mini-prep plasmid DNA for sequencing double stranded template sequence. *BioTechniques* 6, 544–547.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- Lopilato, J.E., Garwin, J.L., Emr, S.D., Silhavy, T.J., & Beckwith, J.R. (1984). D-Ribose metabolism in *Escherichia coli* K-12: Genetics, regulation, and transport. *J. Bacteriol.* 158, 665–673.
- Matsudaira, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262, 10035–10038.
- Mauzy, C.A. & Hermodson, M.A. (1992). Structural homology between *rbs* repressor and ribose binding protein implies functional similarity. *Protein Sci.* 1, 843–849.
- Maxam, A. & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65, 499–560.
- Miller, J.H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Munson, L.M., Stormo, G.D., Niece, R.L., & Reznikoff, W.S. (1984). *lacZ* translation initiation mutations. *J. Mol. Biol.* 177, 663–683.
- Pabo, C.O. & Sauer, R.T. (1984). Protein-DNA recognition. *Annu. Rev. Biochem.* 53, 293–321.
- Reidl, J., Römisch, K., Ehrmann, M., & Boos, W. (1989). Mall, a novel protein involved in regulation of the maltose system of *Escherichia coli*, is highly homologous to the repressor proteins GalR, CytR, and LacI. *J. Bacteriol.* 171, 4888–4899.
- Rolfes, R.J. & Zalkin, H. (1988). *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. *J. Biol. Chem.* 263, 19653–19661.
- Rolfes, R.J. & Zalkin, H. (1990). Purification of the *Escherichia coli* purine regulon repressor and identification of corepressors. *J. Bacteriol.* 172, 5637–5642.
- Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Sanger, F., Nickeln, S., & Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Schulz, V.P. & Reznikoff, W.S. (1990). *In vitro* secondary structure analysis of mRNA from *lacZ* translation initiation mutants. *J. Mol. Biol.* 211, 427–445.
- Silhavy, T.J., Berman, M.L., & Enquist, L.W. (1984). *Experiments with Gene Fusions*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Tabor, S. & Richardson, C.C. (1989). Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis. *J. Biol. Chem.* 264, 6447–6458.
- Tobin, J.F. & Schleif, R.F. (1990). Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. *J. Mol. Biol.* 211, 75–89.