Functional Analysis of the *Escherichia coli* K-12 *cyn* Operon Transcriptional Regulation

ANNE-FRANCOISE J. LAMBLIN AND JAMES A. FUCHS*

Department of Biochemistry, University of Minnesota, St. Paul, Minnesota 55108

Received 13 January 1994/Accepted 25 August 1994

The *cynTSX* operon enables *Escherichia coli* K-12 to degrade and use cyanate as a sole nitrogen source. The promoter of this operon is positively regulated by cyanate and the CynR protein. CynR, a member of the LysR family of regulatory proteins, binds specifically to a 136-bp DNA fragment containing both the *cynR* and the *cynTSX* promoters. In this study, we report the results of DNase I digestion studies showing that CynR protects a 60-bp region on the *cynR* coding strand and a 56-bp sequence on the *cynTSX* coding strand. CynR binding was not affected by cyanate or its structural homolog azide, a gratuitous inducer of the operon. However, CynR-induced bending of two different DNA fragments was detected. The amount of bending was decreased by cyanate.

The observation that *Escherichia coli* strains harboring various *lac* deletions were not inhibited by cyanate led to the identification, isolation, and characterization of the *cyn* operon (31). The *cynTSX* operon, which encodes enzymes to degrade cyanate, is located adjacent to the *lac* operon. The *lac* mRNA and *cyn* mRNA overlap in their 3' ends by 98 nucleotides (29). Cyanate degradation requires carbonic anhydrase encoded by *cynT* and cyanate hydrolase encoded by *cynS*. CynS catalyzes the bicarbonate-dependent hydrolysis of cyanate into ammonia and two molecules of carbon dioxide (1). CynT recycles the carbon dioxide into bicarbonate (10). The third gene of the *cyn* operon, *cynX*, encodes a hydrophobic protein of undetermined function (10). The transcriptional activation of the *cynTSX* polycistronic message is under the positive control of the CynR protein and cyanate (30). The CynR gene is located upstream of the *cynTSX* operon and is transcribed in the opposite direction of the *cyn* operon. CynR negatively regulates its own synthesis independently of cyanate. The *cynR* promoter was found to overlap the *cynTSX* promoter on the opposite strand. The *cynR* gene encodes a 32-kDa protein identified as a member of the LysR family of regulatory proteins (11, 30). CynR recognizes a 136-bp DNA fragment encompassing the *cynR-cynTSX* intergenic region (15). In this study, DNase I footprinting was used to determine the CynR binding site for both promoters. Circular permutation assays indicated that CynR induces bending of the DNA upon binding and that this bending decreases in the presence of cyanate.

MATERIALS AND METHODS

Notation and abbreviations used. Codon numbers with an asterisk correspond to the position of the codon in the uninterrupted sequence *cynR-cynTSX* as it is found on the chromosome, with 1 corresponding to the 3' end of the *cynR* sequence. The bracketed numbers refer to the position of the codon in the *cynR* sequence as it appears in GenBank sequence M93053. The bracketed numbers with the symbol ¶

refer to the position of the codon in the *cynT-cynS* sequence as it appears in GenBank sequence M23219.

Bacterial strains and growth media. The strains used in this study are listed in Table 1. For β -galactosidase assays, cultures were grown in Luria broth medium (18). For plasmid DNA isolation, cultures were grown in terrific broth (32).

Plasmid constructions. Plasmids not described here are listed in Table 1. Plasmid pAL44 was constructed in two steps. (i) A *Bsp*HI restriction site at position 1292*[68] was introduced in plasmid pAL7 by oligonucleotide-directed mutagenesis (13) with the 18-mer oligonucleotide 5' TCTCATTTAT CATGAGTT 3'. (ii) The pAL7 mutagenized plasmid pAL41 was digested with the *Eco*RI and *Taq*I enzymes. The pAL41 *Eco*RI-*Taq*I fragment containing the *cynR-cynTSX* intergenic sequence was ligated into the *Eco*RI-*Acc*I sites of the vector pUC118 to give plasmid pAL44. Uracil-containing single-stranded pAL7 DNA was isolated from strain CJ236 (13). Plasmid pAL54Z was constructed by ligating a 1.2-kb *Bam*HI-*Cla*I fragment of pAL42Z into the pJEL170 vector (16) linearized with the same enzymes. Plasmids pAL55Z and pAL56Z were constructed by an identical procedure with DNA fragments from pAL45Z and pAL46Z, respectively. Plasmid pAL45Z required a three-step construction. (i) A 1.9-kb *Pst*I-*Acc*I fragment from plasmid pSJ104 was cloned into pUC119 linearized with the same enzymes to give plasmid pAL47. This 1.9-kb fragment contains the *cynR* gene, the promoter region, and two-thirds of the *cynT* gene. (ii) Uracil-containing single-stranded pAL47 DNA was isolated from strain CJ236 and used for site-directed mutagenesis. The *Bsp*HI restriction site was introduced at position 1292*[68] as described above to give plasmid pAL48. (iii) A 98-bp *Pvu*II-*Bsp*HI fragment was isolated from pAL43. The fragment was blunt ended at the *Bsp*HI end with mung bean nuclease. The blunt-end fragment containing only the -35 sequence of the *cynR* promoter and the complete *cynTSX* promoter (R2 region) with the N-terminal coding sequence of CynT was ligated into the *Sma*I site of pRS414 to give the CynT-LacZ translational fusion plasmid pAL45Z. Plasmid pAL46Z was obtained by ligating a 1.2-kb *Pvu*II fragment from pSJ104 into the *Sma*I site of pRS414. The 1.2-kb fragment contains the full *cynR* sequence, the promoter region, and the 5' end of the *cynT* gene. The 1-kb pAL45Z *Eco*RI-*Cla*I fragment and the 2-kb pAL46Z *Eco*RI-*Cla*I fragment were inserted into the pJEL170

* Corresponding author. Mailing address: Department of Biochemistry, University of Minnesota, 1479 Gortner Ave., St. Paul, MN 55108. Phone: (612) 624-1215. Fax: (612) 624-5780. Electronic mail address: james-f@molbio.cbs.umn.edu.

TABLE 1. Bacterial strains and plasmids used in this study

<i>Escherichia coli</i> strain or plasmid ^a	Relevant properties	Source or reference
Strains		
CJ236	<i>dut1 ung1 thi-1 relA1</i>	13
SJ100	<i>supE thi Δ(lac-proAB)</i> ; JM101 parental strain	31
JM101	<i>supE thi Δ(lac-proAB)/F' traD36 proA⁺B⁺ lacI^qZ ΔM15</i>	J. Messing
WAM105	<i>F' araD139 Δ(argF-lacU)U169 Δ(his-gnd) thu rpsL150 gltSo flbB5301 relA1 deoC1 rbsR rpoA341</i>	M. S. Thomas (33) G. C. Rowland (23)
Plasmids		
pAL42Z	CynR-LacZ translation fusion vector; pRS414 parental vector	This laboratory (15)
pJEL170	Single-copy plasmid; LacZ transcriptional fusion vector	J. E. Larsen (16)
pRS414	pBR origin vector; LacZ protein fusion vector	R. W. Simons (27)
pSJ104	2.6-kb <i>HincII-ClaI</i> <i>cynR-cynTS</i> fragment in pUC13	This laboratory (31)

^a Construction of all the other plasmids used in this study is described in Materials and Methods.

vector digested with the same enzymes to give, respectively, plasmids pAL55Z and pAL56Z. Plasmids pAL54Z, pAL55Z, and pAL56Z are *lacZ* translational fusion clones similar to plasmids pAL42Z, pAL45Z, and pAL46Z, respectively. Plasmid pJEL170, a *lacZ* transcription fusion vector, was converted to a *lacZ* translation fusion vector, pJEL170T, by replacing the *EcoRI-ClaI* fragment of pJEL170 with the *EcoRI-ClaI* fragment of pRS414. Because the pJEL170 origin of replication is under the control of the λ phage repressor cI857, the vector pJEL170 and its derivatives exist as a single copy at temperatures below 37°C. Shifting the culture to 42°C leads to uncontrolled replication of the vector. For studying CynR-induced bending, plasmids pAL51 and pAL52 were constructed. Vector pBend3 (40) with circularly permuted restriction sites was linearized at the unique *XbaI* site, and the ends were filled in by Klenow reaction. The *HindIII-EcoRI* ends of the 141-bp AL1 fragment were filled with Klenow enzyme, and the blunt-end fragment was ligated into the filled *XbaI* site of pBend3 to give plasmid pAL51. Plasmid pAL52 is a similar construct that contains only the R1 region of the CynR binding site. The R1 region sequence was isolated by digestion of plasmid pAL44 with *HindIII* and *BspHI* enzymes. The fragment was blunt-ended with Klenow enzyme, and the 64-bp R1 fragment was cloned into the filled *XbaI* site of pBend3 to give pAL52. The correct orientation of the insert in the various clones and in-frame ligations with the *lacZ* open reading frame was verified by dideoxy sequencing (24) with the Sequenase reagent kit (U.S. Biochemicals).

Determination of the CynR binding site by DNase I footprinting. DNase I footprinting of CynR on the *cynTSX* and *cynR* coding strands, both in the presence and the absence of cyanate, was conducted as described by Galas and Schmitz (9). A 134-bp *EcoRI-PstI* DNA fragment and a 132-bp *HindIII-KpnI* DNA fragment were isolated from plasmid pAL1. The *cynTSX* coding strand was labeled by using the 134-bp *EcoRI-PstI* DNA (AL1EP) and filling in the *EcoRI* end with [α -³²P]dATP and Klenow enzyme. The *cynR* coding strand was labeled by using the 132-bp *HindIII-KpnI* fragment (AL1HK) and filling in the *HindIII* end with [α -³²P]dATP, [α -³²P]dGTP, and Klenow enzyme. Protein-DNA complexes were allowed to form in a volume of 20 μ l. Each reaction mixture contained 1.23 ng of either AL1EP (136 bp or 6.8×10^{-10} M) or AL1HK (134 bp or 7.1×10^{-10} M) labeled DNA fragment, 15.5 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)-KOH (pH 7.4), 10 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 110 μ g of bovine serum albumin per ml, 5 μ g of poly(dI-dC) DNA per ml, and increasing amounts of purified

CynR protein (0, 6.5×10^{-10} , 3.25×10^{-9} , 6.5×10^{-9} , or 13×10^{-9} M). Where indicated, 1 mM potassium cyanate was added to the binding reaction mixture. After 15 min at 37°C, DNase I digestion was initiated by addition of 2 μ l of DNase mix (20 mM CaCl₂-MgCl₂, 0.5 U of Promega DNase QI). Digestion was conducted at 37°C for 10 to 40 s and was stopped with 100 μ l of phenol. Digested DNA was extracted and ethanol precipitated. Pellets were washed with 70% ethanol, dried, resuspended in 5 μ l of formamide loading buffer, and heat denatured. Samples were resolved on an 8% polyacrylamide-6 M urea-0.5 $\times$ Tris-borate-EDTA sequencing gel. The M13 sequence with the -40 direct primer provided the ladder to size the footprint. Gels were dried and exposed for autoradiography.

Effect of azide on DNA binding activity. The method of gel mobility shift described by Fried and Crothers (8) was used as follows. The 136-bp AL1EP radiolabeled fragment was incubated with the purified CynR protein at the concentration indicated in Fig. 3. DNA-CynR complexes, prepared as described in the DNase I footprinting section, were resolved on a 6% polyacrylamide native gel in 1 $\times$ Tris-acetate-EDTA buffer. When the stable and gratuitous inducer sodium azide (10 mM) was used in the binding reaction, azide was added to the gel and to the 1 $\times$ Tris-acetate-EDTA running buffer. The buffer was recirculated to keep azide constant in the protein environment. The dried gels were scanned with the Ambis4000 radioimager (Ambis, San Diego, Calif.) for 6 to 12 h. The radioactivity of the resolved CynR-DNA complexes and of free AL1EP DNA was quantitated with the Ambis analysis software. The percentage of DNA-protein complex formed was plotted as a function of the log₁₀ of the CynR molar concentration in the assay. The mean value of three experiments is presented in Fig. 3.

Subdivision of the CynR DNA binding site. The CynR DNA binding site was subdivided into two regions—R1 and R2—on the basis of the DNase I footprinting results. DNA binding assays were conducted with each of these sequences. R1 and R2 fragments were generated by PCR with the *XmnI* linearized plasmid pAL44 as template. The M13 -20 forward primer (10 pmol) and *BspHI* primer 5' TCTCATTTATCATGAGTT 3' (10 pmol) were used to amplify a 114-bp fragment encoding the R1 sequence. The M13 -20 reverse primer (10 pmol) and *BspHI*-T primer 5' AACTCATGATAAATGAGA 3' (10 pmol) were used to amplify a 132-bp fragment carrying the R2 region. Denatured samples were cycled 30 times (94°C for 1 min, 40°C for 1 min, ramping to 72°C in 1 min, and 72°C for 1 min) by using the PE9600 amplifier (Perkin-Elmer) and PCR

Taq polymerase (Perkin-Elmer). Amplified DNA was extracted with phenol and chloroform and precipitated with ethanol. The R1 and R2 fragments were both digested with *Bsp*HI enzyme. The R1 fragment was subsequently digested with *Hind*III, and the R2 fragment was subsequently digested with the *Eco*RI restriction enzyme. The *Bsp*HI-*Hind*III AL44 fragment (R1, or AL44BH) and *Bsp*HI-*Eco*RI AL44 fragment (R2, or AL44BE) were specifically labeled at the *Hind*III or *Eco*RI end by Klenow enzyme filling in with [α - 32 P]dATP. The final AL44BH fragment was 58 bp, and the AL44BE DNA fragment was 104 bp. The binding conditions were identical to the ones used in the footprinting experiments. AL1EP DNA was used as a control for DNA binding activity of CynR. All DNAs were used at the final assay concentration of 4.2×10^{-10} M. Purified CynR was used at 0.32×10^{-8} , 0.65×10^{-8} , and 1.3×10^{-8} M. DNA-CynR complexes and free DNA were resolved on a 6% polyacrylamide native gel in $0.5 \times$ Tris-borate-EDTA buffer. Gels were dried and autoradiogrammed.

CynR bending assay. A circular permutation assay (38, 40) was used to test the ability of CynR to induce bending upon binding of DNA fragments containing either the full intergenic region (both R1 and R2 sequences present) or only the R1 DNA sequence. Plasmid pAL51 was digested individually with the enzymes *Bgl*II, *Nhe*I, *Spe*I, *Xho*I, *Eco*RV, *Sma*I, *Nru*I, and *Bam*HI. The fragments generated by these enzymes were of equal length, with the binding sites located at different positions. The DNA fragments were purified from low-melting-point agarose. Restriction DNA fragments with blunt ends were dephosphorylated. The purified fragments were radiolabeled by either a Klenow- or kinase-catalyzed reaction with [α - 32 P]deoxynucleoside triphosphate or [γ - 32 P]ATP, depending on the 5' end generated. The labeled fragments were used for gel retardation analysis, and complexes were resolved by electrophoresis on a 6% native polyacrylamide gel in $0.5 \times$ Tris-borate-EDTA buffer. Where indicated, 1 mM cyanate was used in the binding reaction mixture and in the electrophoresis setup with $0.5 \times$ Tris-borate-EDTA recirculating buffer. The distance of migration was calculated by reference to the 298- or 396-bp fragment of the 1-kb phage DNA (Biochemical Research Laboratory) ladder labeled by Klenow enzyme with [α - 32 P]dATP. Gels were dried and autoradiogrammed. The extent of CynR-induced bending (α) was estimated with the equation $\mu^M/\mu^E = \cos(\alpha/2)$ (34), where μ^M corresponds to the lowest mobility of the complex when CynR is bound near the middle of the fragment and μ^E is the highest mobility of the complex when the protein is bound near the end of the fragment (38).

β -Galactosidase assay. The cells harboring the *lacZ* translation fusion plasmids pAL54Z, pAL55Z, pAL56Z, and pJEL170T were grown at 30°C in Luria broth medium supplemented with ampicillin (50 μ g/ml). When the culture reached an optical density at 660 nm of 0.5 to 0.6, potassium cyanate was added to the culture at a final concentration of 1 mM. The cultures were incubated at 30°C for 20 min and were then quickly chilled on ice. One-milliliter samples were frozen for protein and β -galactosidase assays. The production of 2-*ortho*-nitrophenol (ONP) from ONP- β -D-galactopyranoside (ONPG) upon addition of cellular extracts was monitored by the change of A_{412} with a kinetic microplate reader (Thermomax microplate reader; Molecular Devices). With the 96-well flat-bottom microtiter plate, the molar extinction coefficient of ONP was found to be 562 M $^{-1}$ for a 100- μ l assay (2). This value was used to calculate the β -galactosidase activity in cellular extracts as micromoles of ONP formed per minute. Samples of log cultures (30°C) of equivalent optical density at 660 nm were thawed on ice and sonicated, and the supernatant obtained

FIG. 1. CynR DNase I footprint on the *cynTSX* coding strand. In all lanes, the labeled 136-bp AL1EP DNA fragment was present at a 6.8×10^{-10} M final assay concentration. CynR protein was present in the assay at the following concentrations: no protein, lanes 1, 6, 7, and 12; 6.5×10^{-10} M, lanes 2 and 8; 3.25×10^{-9} M, lanes 3 and 9; 6.5×10^{-9} M, lanes 4 and 10; and 13×10^{-9} M, lanes 5 and 11. Where indicated, 1 mM potassium cyanate (KCNO) was added to the binding reaction mixture. Bands are labeled relative to the start of transcription of the *cynR* (R) and *cynTSX* (T) promoters. Bars on each side of the footprint indicate the limits of the protected region.

after centrifugation for 2 min (Eppendorf laptop centrifuge; $14,000 \times g$) was used for assays. Reactions were started by addition of 90 μ l of Z buffer (60 mM Na $_2$ HPO $_4$, 40 mM NaH $_2$ PO $_4$, 10 mM KCl, 1 mM MgSO $_4$, 50 mM β -mercaptoethanol, 900 mg of ONPG per ml) (20) to 10 μ l of cell extract. The total protein concentration was determined by the method of Bradford (3) by using the Bio-Rad protein assay kit and bovine gamma globulin as the protein standard.

RESULTS

DNase I footprinting analysis. We previously demonstrated that CynR is a DNA-binding protein. CynR binds specifically to a 136-bp DNA fragment containing both the *cynR* and *cynTSX* promoter sequences (15). To more precisely localize the site of interaction between CynR in the intergenic region, DNase I footprinting analysis was conducted. The site of interaction was examined on the *cynTSX* coding strand and on the *cynR* coding strand both in the presence and in the absence of 1 mM cyanate (Fig. 1 and 2, respectively). On the *cynTSX* coding strand, the CynR binding site was found to extend from -75 to inside the spacer region of the *cynTSX* promoter at -21 (positions are given relative to the *cynTSX* start site of transcription). On the *cynR* coding strand, the protection extends from +19 inside the coding region to -41 (positions are given relative to the *cynR* transcription start site). No

FIG. 2. CynR DNase I footprint on the *cynR* coding strand. In all lanes, the labeled 134-bp AL1KH DNA fragment was present at a 7.1×10^{-10} M final assay concentration. CynR protein was present in the assay at the following concentrations: no protein, lanes 1, 6, 7, and 12; 6.5×10^{-10} M, lanes 2 and 8; 3.25×10^{-9} M, lanes 3 and 9; 6.5×10^{-9} M, lanes 4 and 10; 13×10^{-9} M, lanes 5 and 11. When indicated, 1 mM potassium cyanate (KCNO) was added to the binding reaction mixture. Bands are labeled relative to the start of transcription of the *cynR* (R) and *cynTSX* (T) promoters. Bars on each side of the footprint indicate the limits of the protected region.

reproducible difference was observed in the DNase I footprinting pattern when the inducer was present in the reaction mixture. Strong DNase I hypersensitive sites were detected in the presence of the CynR protein. These hypersensitive sites are located on each side of the *cynR* -10 sequence (positions -11 and -12 and -20 to -22 on the *cynR* coding strand) and between the *cynR* -10 and *cynTSX* -35 sequences (positions -49 and -48 [-40, -39, and -38 on the *cynTSX* coding strand]) (Fig. 1 and 2). One interpretation of this result is that CynR binding changes the conformation of the DNA between the *cynR* -10 and *cynTSX* -35 sequences to a conformation more susceptible to DNase I endonuclease activity. An alternative explanation, which seems unlikely, is that nonspecific interactions of CynR and DNase I result in the hypersensitive sites. The *cynR* -10 region on the *cynR* coding strand is partially protected as is the complementary sequence on the *cynTSX* coding strand. The *cynTSX* -10 sequence is not protected by CynR. DNase I footprinting analysis was also conducted on larger fragments extending 205 bp into the CynR coding sequence (up to the *Bgl*II site at position 1053*[296]) and 73 bp into the CynT sequence (up to the *Pvu*II site at position 1441*[194]) on both coding strands (data not shown). No additional protection in either the presence or absence of cyanate was detected. Previous DNase I footprinting analyses conducted with the *cynTSX* coding strand, with a CynR protein

FIG. 3. Quantitation of CynR DNA binding to the *cynR-cynTSX* intergenic region. 32 P-labeled AL1EP DNA was used at a final assay concentration of 7.7×10^{-10} M. CynR protein was added to the binding reaction mixture at the following concentrations: 0, 1.25×10^{-11} , 2.5×10^{-11} , 1.25×10^{-10} , 1.87×10^{-10} , 2.5×10^{-10} , 6.25×10^{-10} , 1.25×10^{-9} , 1.75×10^{-9} , 2.25×10^{-9} , 2.5×10^{-9} , 6.25×10^{-9} , and 2.5×10^{-8} M. The fraction of the CynR bound to DNA was plotted as the logarithmic function of the CynR concentration in the presence (●) or absence (asterisks) of 10 mM azide in the assay.

purified with a procedure different from the one used in this study, suggested the presence of two binding regions, R1 and R2, with different affinities; R1 has the highest affinity (data not shown). The two regions were observed on the *cynTSX* coding strand (data not shown). DNA footprints were delimited by a nonprotected band in the -10 CynR promoter sequence. Although this difference in binding was not reproduced in the DNase I footprinting analysis presented here, we retained the assigned subdivision of CynR binding sites into R1 and R2 sites (Fig. 2).

CynR DNA binding affinity. Although the DNase I footprinting analysis did not reveal any reproducible effect of cyanate on CynR binding recognition, we investigated the potential effect of the inducer on the binding affinity of the protein. Because cyanate is not stable in solution for prolonged periods of time, we used sodium azide, a stable and nonmetabolizable analog of cyanate and a gratuitous inducer of the *cyn* operon. After gel shift assays, the electrophoretically resolved protein-DNA complexes were quantified with the Ambis4000 radioimager and software analysis package. The percentage of complex formed was plotted as the logarithmic function of the CynR protein concentration in the assay (Fig. 3). No difference was noted in the presence or the absence of azide. The CynR concentration that results in an equal amount of bound and free DNA is 3×10^{-9} to 4×10^{-9} M and does not change in the presence of 10 mM sodium azide in the reaction. Calculation of a true K_d was not possible from these data. The data indicate a very strong cooperative binding of CynR. A very sharp transition from 5% DNA complexed to more than 80% was accomplished with a protein concentration ranging from 2×10^{-9} to 5.5×10^{-9} M. The unimodal binding

FIG. 4. CynR DNA binding assay with the *cynR-cynTSX* restricted promoter. Three different DNA fragments were used: the 136-bp labeled AL1EP DNA fragment, which contains both the R1 and R2 sequences; the 58-bp AL44BH fragment, which contains only the R1 sequence; and the 104-bp AL44BE DNA fragment, which contains only the R2 sequence. In all reactions, the respective DNAs were used at the final concentration of 4×10^{-10} M. CynR protein was used in the assay at the following concentrations: no protein, lanes 1, 5, and 9; 1.3×10^{-8} M, lanes 2, 6, and 10; 0.65×10^{-8} M, lanes 3, 7, and 11; and 0.32×10^{-8} M, lanes 4, 8, and 12.

indicated by the shape of the curve would suggest that even if more than one binding site is present for CynR, as we suspect, CynR binding to multiple sites would occur simultaneously. In the binding assay, only one complex was observed at any protein concentration (Fig. 4, lanes 2 to 4) and with larger intergenic DNA fragments (data not shown). The CynR concentration that results in equal amounts of bound and free DNA agrees with the results presented in Fig. 1 and 2. A concentration of 3.25×10^{-9} M CynR protects approximately half of the DNA from digestion with DNase. DNase I protection is conducted in solution with an equilibrium of free and complexed DNA and thus is not susceptible to artifacts of gel retardation.

Subdivision of the CynR DNA binding site. On the basis of the DNase I protection results, we subdivided the CynR binding site into two regions, R1 and R2. We tested CynR binding to either R1 (Fig. 4, lanes 5 to 8) or R2 (Fig. 4, lanes 9 to 12). The AL1EP fragment was used as a positive control for CynR DNA binding activity (Fig. 4, lanes 1 to 4). While CynR formed a complex with the R1 fragment (Fig. 4, lanes 6 to 8), it did not bind the R2 sequence in the range of protein concentrations used in this assay (Fig. 4, lanes 10 to 12). Although the protein was used in large excess in all lanes, only a small fraction of the R1 fragment was retarded (Fig. 4, lanes 6 to 8). At the same protein concentration, the AL1EP fragment was fully complexed (Fig. 4, lanes 2 to 4).

CynR DNA bending assay. Because of the numerous DNase I hypersensitive sites observed upon CynR binding, the possi-

bility of a DNA conformational change due to bending was investigated. The ability of the CynR protein to induce DNA bending upon binding was examined by circular permutation analysis. The DNA fragments used in this assay were the same size (277 or 189 bp) but differed from each other by the position of the CynR DNA binding region within the fragment. Figures 5 and 6 show the positioning effect of the CynR binding site on the electrophoretic mobility of the various CynR-DNA complexes. Complexes in which the full intergenic region (Fig. 5) is located at or near the center of the fragment (Fig. 5, lane f) exhibited the lowest level of mobility, whereas complexes with the CynR binding sequence at the end of the fragment (Fig. 5, lane a) showed the highest level of mobility. The electrophoretic mobility of the different free DNA fragments did not vary, which suggests the lack of inherent curvature in the DNA sequence alone (data not shown). Similar observations were made with DNA fragments containing only the R1 sequence (Fig. 6). The *Bam*HI fragment (Fig. 6C, point h) had the highest level of mobility in the presence or absence of cyanate, but the fragment with the lowest mobility shifted from the *Xho*I DNA fragment in the absence of cyanate to the *Sma*I fragment in its presence (Fig. 6C, points d and c, respectively). The extent of bending was determined empirically with the formula $\mu^M/\mu^E = \cos(\alpha/2)$ (34). α represents the angle between the linear stretched DNA molecule and the bent DNA. μ^M was calculated from the *Sma*I fragment (Fig. 6C, point f), which had the slowest mobility with the CynR protein bound in the middle of the fragment. μ^E was calculated from the *Bgl*II fragment, which showed the highest mobility with the CynR protein bound at the end of the fragment (Fig. 6C, point a). Because the free DNA ran off the gel, the relative mobility was calculated by reference to the 396-bp DNA marker for the *cynR-cynTSX* intergenic sequence or to the 298-bp DNA marker when the R1 sequence was used. The bend induced by CynR to the *cynTSX-cynR* intergenic DNA fragment was reduced from an approximate 75°C to 65°C when cyanate was added to the reaction mixture (Fig. 6C). The bend was estimated to occur near position 1294*[56] located between the *cynR* -10 region and the *cynTSX* -35 region, which is AT rich. CynR protein was also found to induce bending of the R1 DNA fragment (Fig. 7), and the bend was mapped near position 1275*[75] in the absence of cyanate and to position 1281*[69] in its presence. A strong difference in electrophoretic mobility can be directly observed from the gel autoradiogram (Fig. 6A and B). With only the R1 sequence, the migration of the complex was strongly altered compared with that of the identical restriction digest fragment containing both the R1 and R2 sequences (Fig. 6C, points a and h, versus Fig. 5C, lanes a and h). For the R1 fragment, in the absence of cyanate, the CynR-induced angle was estimated to be approximately 45°C. The *Bam*HI DNA fragment (Fig. 6C, point h) was selected for the highest electrophoretic mobility, and the *Xho*I DNA fragment (Fig. 6C, point d) was selected for the lowest one. However, in the presence of 1 mM cyanate, the *Spe*I fragment (Fig. 6C, point c) had the most impaired migration (with the *Nhe*I fragment, Fig. 6C, point b). The absolute degree of bending may not be exact, since it could be affected by the length of DNA and base composition compared with that used by Thompson and Landy (34). Also, different electrophoresis conditions, and a different mass of protein to mass of DNA were used. The fact that there is strong bending and that it decreased in the presence of cyanate is important. A summary of the sites protected by DNase I, the hypersensitive sites, and the sites of bending in fragment AL1 and the bend with only R1 is given in Fig. 7.

A **C**

β -Galactosidase activity assays. The R2 DNA fragment contains both the -10 and -35 *cynTSX* promoter sequences. To test whether RNA polymerase was able to initiate transcription from this promoter, the R2 sequence followed by the 5' end of the *cynT* gene was translationally fused to the *lacZ* gene (pAL55Z). In the absence of cyanate and CynR, the *cynTSX* operon promoter in the R2 fragment (pAL55Z) directed transcription to the same extent as the *cynTSX* promoter in the nonaltered intergenic region (pAL56Z) (Table 2). These results thus indicate that CynR does not prevent the RNA polymerase from binding the *cynTSX* promoter and

B

FIG. 5. CynR-induced bending of the *cynR-cynTSX* DNA promoter region. A 152-bp DNA fragment containing the CynR binding site was cloned into the pBend3 vector. A set of eight fragments of equal length (277 bp) with the binding sites located at different positions within each fragment were generated with the following restriction enzymes: a, *Bgl*II; b, *Nhe*I; c, *Spe*I; d, *Xho*I; e, *Eco*RV; f, *Sma*I; g, *Nru*I; and h, *Bam*HI. These fragments were used in gel shift assays in the absence of cyanate (A) or in the presence of 1 mM cyanate (B). (C) Position of the bend within the 277-bp fragment. The relative mobility of the protein-DNA complexes was plotted as a function of the distance (nucleotides [nt]) between the pBend3 *Eco*RI site at the extreme left end of the polyclonal site and the center of the 277-bp DNA fragment.

from initiating transcription. The *cyn* operon promoter in the R2 fragment (pAL55Z) failed to respond to induction of transcriptional activation with both CynR, provided in *trans* (strain JM101), and cyanate present. With the intact intergenic sequence and the *cynR* gene provided in *cis* (pAL56Z), 66- and 48-fold increases in β -galactosidase activity driven from the *cynTSX* promoter were measured in the strains JM101 (*cyn*⁺) and SJ100 (*cyn* mutant), respectively (Table 2). It thus appears that the R1 sequence is required for the transcription activation of the *cyn* operon.

With CynR-GalK transcriptional fusion and galactokinase assays, Sung and Fuchs (30) showed that the CynR-negative autoregulation was observed when a copy on the *cynR* gene was provided in *trans*. In this study, the pAL54Z *cynR-lacZ* translational fusion reporter construct indicates that the negative autoregulation is constantly observed, both in a *cynR*⁺ (JM101) and a *cynR* mutant (SJ100) genetic background. The CynR-LacZ chimeric protein encoded in pAL54Z contains the CynR helix-turn-helix DNA binding domain and the LacZ multimerization domain. This combination might enable the chimeric CynR-LacZ protein to bind the *cynR-cynTSX* intergenic sequence and thus negatively regulate *cynR* transcription.

A **C**

DISCUSSION

Members of the LysR family of positive regulatory proteins all require a low-molecular-weight compound for activation of transcription. The DNase I protection observed with CynR was identical in the presence or absence of cyanate on either promoter. No general correlation has been established between the effect of the inducer and changes in the protein-DNA interactions for the members of the LysR family. DNase I protection assays indicated that TrpI (5) and CatR (21) protect additional sequences in the presence of their respective inducers, while NahR (25), *IlvY* (37), AmpR (17), and NodD (7) did not exhibit any changes in the protected region upon

B

FIG. 6. CynR-induced bending of the R1 DNA fragment. A 64-bp DNA fragment containing the CynR binding site was cloned into the pBend3 vector. A set of eight fragments of equal length (189 bp) with the binding sites located at different positions within each fragment were generated with the following restriction enzymes: a, *Bgl*II; b, *Nhe*I; c, *Spe*I; d, *Xho*I; e, *Eco*RV; f, *Sma*I; g, *Nru*I; and h, *Bam*HI. These fragments were used in gel shift assays in the absence of cyanate (A) or in the presence of 1 mM cyanate (B). (C) Position of the bend within the 189-bp fragment. The relative mobility of the protein-DNA complexes was plotted as a function of the distance (nucleotides [nt]) between the pBend3 *Eco*RI site at the extreme left end of the polyclonal site and the center of the 189-bp DNA fragment. ○, 1 mM cyanate; ■, no cyanate.

addition of the inducer. The OccR footprint was shown to decrease by 8 bp in the presence of octopine (36). However, in vivo methylation protection experiments with NahR (12) revealed some modifications in the DNA-NahR contacts in the presence of the inducer, which were not detected by in vitro footprinting.

The LysR family also varies in the ability of the inducer to alter regulatory protein binding strength. The CynR DNA binding affinity was not altered by sodium azide, a gratuitous inducer of the *cyn* operon (Fig. 3); similar results were obtained with Trp (5) and NahR (25) proteins and their respective inducers indoleglycerol phosphate and salicylate. CatR DNA binding activity is increased an estimated 20-fold by the presence of *cis,cis*-muconate (21), while the binding affinity of OccR is decreased by the presence of octopine in the assay (36). Thus, even though the members of the LysR family are all activator proteins sharing protein homology, the mechanisms of transcription activation appear to be different.

Like the binding sites of most transcriptional factors (activators or repressors) (6), which either overlap the σ^{70} RNA polymerase binding site or are located immediately upstream of this sequence, the binding sites of the LysR regulatory proteins are immediately adjacent to or overlap the sequences of the promoters they activate. MetR (4, 19), AmpR (17), and NahR (25) proteins bind immediately adjacent to the promoters they activate, while CatR (21), OccR (36), TrpI (5), RbcR

TABLE 2. Transcriptional efficiency of different sequences of the *cynR-cynTSX* intergenic region

<i>E. coli</i> strain	Plasmid	β -Galactosidase sp act (nmol of ONP/min/mg of protein) with:	
		No cyanate	0.5 mM cyanate
SJ100	pJEL170T	1	0.9
	pAL54Z	5.5	4.7
	pAL55Z	1.5	1.6
	pAL56Z	1.1	48
JM101	pJEL170T	1.7	0.5
	pAL54Z	4.3	3.2
	pAL55Z	1.4	0.7
	pAL56Z	0.9	66

(14), IlvY (37), and CynR (this study) bind to sites overlapping the RNA polymerase recognition sequence. CatR and TrpI binding sites are classified as overlapping even though in the absence of inducer the DNase I protected region is adjacent to and upstream of the positively regulated promoter. From these examples, a correlation emerges between the DNA binding site location, the DNA binding affinity, and the effect of the inducer on both. For DNA binding sites that extend over the activated promoter in the absence of the inducer, the binding affinity of the protein is not altered by the inducer. The DNA binding sites that map upstream of the activated promoter in the absence of the inducer exhibit increased binding in the presence of the inducer and extend over the promoter in the presence of inducer. The OccR system is a subdivision of the first class having a binding site that overlaps the activated promoter in the absence of inducer, but it has its binding site and binding affinity reduced in the presence of the effector (36).

The general hypothesis for transcriptional activation is that, upon inducer binding, the regulatory protein undergoes a conformational change that allows contact with the RNA polymerase. Alternatively, the protein conformational change could translate into a DNA conformational change such as bending. A change in DNA bending was reported with OccR (36) and, in this study, with CynR. The bend was located 12 bp upstream from the *cyn* operon -35 sequence. The OccR-induced bend was mapped 12 to 13 bp upstream from the *occQ* -35 sequence. In both cases, the DNA bend was partially released by the presence of the inducer (Fig. 6 and 7). It is reasonable to predict that, when examined, other members of the LysR family may exhibit the same behavior. A model of transcriptional activation in which DNA bending would act as a transcriptional switch has been proposed (22). DNA curvature, either inherent to a sequence or induced by bound protein, could facilitate promoter-RNA polymerase interaction when in phase with the promoter. Extrapolation of this model to the *cyn* operon gives a reasonable and attractive explanation for transcriptional activation of the *cyn* operon and transcriptional repression of *cynR*. Like catabolite activator protein, CAP, which introduces a 90° DNA bend (26), transcription activation of the *cyn* operon may be associated with bending. However, interprotein contact between RNA polymerase and the regulator and overcoming an activation energy barrier can also be critical factors in transcription activation. A possibility is that bending, interprotein contact, and free energy (ΔG) are all connected and contribute to the process. In their CAP study, Zinkel and Crothers (39) proposed that the bending energy might assist RNA polymerase in escaping the

FIG. 7. Summary of the DNase I protection and bending assays on the *cynR-cynTSX* intergenic region. The CynR (lower strand) and the CynTSX (upper strand) coding strands are represented. Sequences protected from DNase I digestion are framed. Arrowheads indicate DNase I hypersensitive sites. Asterisks indicate the partially protected and nonprotected sequences that are part of the CynR footprint. +1 *cynR*, *cynR* transcription start site; +1 *cynT*, *cynTSX* transcription start site. The -10 and -35 promoter sequences are indicated and thickly underlined. A potential CynR recognition sequence is indicated by the thin line outside the frame. The positions of the bend observed with the R1 fragment and of the AL1 fragment in the presence (+) and absence (−) of cyanate (CNO) are indicated by arrows.

promoter. Strong protein-protein and/or protein-DNA interaction could prevent the RNA polymerase from moving from the promoter.

Although CynR was found by gel filtration to be a dimer in solution (28), it has not been determined if CynR binds as a monomer or a dimer and then interacts to form either a dimer or a tetramer. We propose that CynR binds to R1 as a dimer and that simultaneously the second site of the dimer binds to R2. If correct, the binding must be strongly cooperative, since we do not observe, either by gel retardation or footprinting, that an intermediate species with only one site was occupied. Both R1 and R2 contain the sequence 5' ATAAGTAAA 3' at the edge of the protected region. If this sequence is the CynR recognition sequence, the two-base mismatch of TA versus GT could account for the difference in CynR binding affinity of the two sites when separated. Cooperativity of binding could explain the inability to observe protein at only one site when the two sites are adjacent. If, as hypothesized, the R1 sequence ATAAGTAAA is the CynR DNA-binding site, an exact copy of the former sequence in the R2 fragment would promote tighter CynR binding to the R2 sequence. This higher affinity might not allow enough flexibility for the DNA and the protein to enhance RNA polymerase transcription in the presence of the inducer. As suggested for many members of the LysR family, it is speculated that the inducer triggers a conformational change of the activator. However, no direct binding of an inducer to any member of the LysR family has yet been reported. We interpret the DNA bend upstream of the -35 *cynTSX* as a consequence of the interaction between the CynRs bound to the R1 and R2 regions. We propose that cyanate triggers a CynR conformational change that modifies the interaction between CynR and DNA and that the observed relaxation in bending (Fig. 6 and 7) is a consequence of this

protein structural change. We propose that the cyanate-induced conformational change of the CynR-DNA complex in the R2 region is optimal for the RNA polymerase to initiate transcription. This activated conformation could facilitate RNA polymerase escape as well as correct docking of the RNA polymerase on the promoter. The region with partial DNase I protection is framed by the DNase I hypersensitive sites on both strands, which agrees well with the estimated location of the bend and the binding of one CynR to R1 and a second to R2 with interaction between them. Forty-two base pairs or four turns of the DNA helix separate the centers of the proposed DNA recognition sequences ATAAGTAAA and ATAAGG TAA. This would position the two CynRs on the same face of the helix and facilitate protein-protein interaction. The change in bending of the R1 fragment is the result of only one CynR binding. The bending on the R1 fragment is located at the 3' end of the ATAAGNNAA sequence. However, the bend observed with the full promoter results from protein interactions between the CynR bound at distant sites on the DNA. This would be reminiscent of a mini loop structure.

A few alternative schemes for transcription activation and relevant features are presented above. The challenge remains to establish the cause-effect relationship between them. Some, such as the importance of the bending in the positioning of the activator to foster contact with the RNA polymerase, are structure related. Others, like the possibility of the RNA polymerase to recognize the promoter in the absence of the activator and to form a closed complex with enough time to allow a productive open complex are structure and kinetics related. Overall, to have a complete picture, one would have to consider the free energy changes in the various steps of the transcription activation process.

ACKNOWLEDGMENTS

We thank R. W. Simons for the gift of the pRS transcription and translation LacZ fusion vectors and R. E. Glass for providing us with the WAM105 bacterial strain.

This work was supported by Public Health Service grant GM33842-07 from the National Institutes of Health.

REFERENCES

- Anderson, P. 1980. Purification and properties of the inducible enzyme cyanase. *Biochemistry* **19**:2882-2888.
- Augustin, L. B. 1992. Regulation of ribonucleotide reductase expression in *Escherichia coli*. Ph.D. thesis. University of Minnesota, St. Paul.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
- Byerly, K. A., M. L. Urbanowski, and G. V. Stauffer. 1991. The MetR binding site in the *Salmonella typhimurium* *metH* gene: DNA sequence constraints on activation. *J. Bacteriol.* **173**:3547-3553.
- Chang, M., and I. P. Crawford. 1989. The roles of indoleglycerol phosphate and the TrpI protein in the expression of *trpBA* from *Pseudomonas aeruginosa*. *Nucleic Acids Res.* **18**:979-988.
- Collado-Vides, J., B. Magasanik, and J. D. Gralla. 1991. Control site location and transcription regulation in *Escherichia coli*. *Microbiol. Rev.* **55**:371-394.
- Fisher, R. F., and S. R. Long. 1989. DNA footprint analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible *nod* gene promoters. *J. Bacteriol.* **171**:5492-5502.
- Fried, M., and D. M. Crothers. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505-6525.
- Galas, D., and A. Schmitz. 1978. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **6**:302-314.
- Guillot, M. B., A. F. Lamblin, E. I. Kozliak, M. Gerami-Nejad, C. Tu, D. Silverman, P. M. Anderson, and J. A. Fuchs. 1993. A physiological role for cyanate-induced carbonic anhydrase in *Escherichia coli*. *J. Bacteriol.* **175**:1443-1451.
- Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace. 1988. A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA* **85**:6602-6604.
- Huang, J., and M. A. Schell. 1991. *In vivo* interactions of the NahR transcriptional activator with its target sequences. Inducer-mediated changes resulting in transcription activation. *J. Biol. Chem.* **266**:10830-10838.
- Kunkel, T. A., J. D. Roberts, and R. A. Zakoar. 1986. Rapid and efficient site-directed mutagenesis without phenotypic selection. *Methods Enzymol.* **154**:367-382.
- Kusano, T., and K. Sugawara. 1993. Specific binding of *Thiobacillus ferrooxidans* RbcR to the intergenic sequence between the *rbc* operon and the *rbcR* gene. *J. Bacteriol.* **175**:1019-1025.
- Lamblin, A.-F. J., and J. A. Fuchs. 1993. Expression and purification of the *cynR* regulatory gene product: CynR is a DNA-binding protein. *J. Bacteriol.* **175**:7990-7999.
- Larsen, J. E. L., B. Albrechtsen, and P. Valentin-Hansen. 1987. Analysis of the terminator region after the *deoCABD* operon of *Escherichia coli* K-12 using a new class of single copy number operon-fusion vectors. *Nucleic Acids Res.* **15**:5125-5140.
- Lindquist, S., F. Lindberg, and S. Normark. 1989. Binding of the *Citrobacter freundii* AmpR regulator to a single DNA site provides both autoregulation and activation of the inducible *ampC* β -lactamase gene. *J. Bacteriol.* **171**:3746-3753.
- Luria, S. E., and J. W. Burrous. 1957. Hybridization between *Escherichia coli* and *Shigella*. *J. Bacteriol.* **74**:461-476.
- Mares, R., M. L. Urbanowski, and G. V. Stauffer. 1992. Regulation of the *Salmonella typhimurium* *metA* gene by the MetR protein and homocysteine. *J. Bacteriol.* **174**:390-397.
- Miller, J. H. 1972. Experiments in molecular genetics, p. 431-432. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Parsek, M. R., D. L. Shinabarger, R. K. Rothmel, and A. M. Chakrabarty. 1992. Roles of CatR and *cis,cis*-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. *J. Bacteriol.* **174**:7798-7806.
- Pérez-Martin, J., and M. Espinosa. 1993. Protein-induced bending as a transcriptional switch. *Science* **260**:805-807.
- Rowland, G. C., P. M. Giffard, and I. R. Booth. 1984. Genetic studies of the *phs* locus of *Escherichia coli*, a mutation causing pleiotropic lesions in metabolism and pH homeostasis. *FEBS Lett.* **173**:395-400.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
- Schell, M. A., and E. Faris Poser. 1989. Demonstration, characterization, and mutational analysis of NahR protein binding to *nah* and *sal* promoters. *J. Bacteriol.* **171**:837-846.
- Schultz, S. C., G. C. Shields, and T. A. Steitz. 1991. Crystal structure of CAP DNA complex: the DNA is bent by 90°C. *Science* **253**:1001-1007.
- Simons, R. W., F. Houman, and N. Kleckner. 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* **53**:85-96.
- Sung, Y.-C. 1987. Identification and characterization of the *cyn* operon and its regulation in *Escherichia coli* K-12. Ph.D. thesis. University of Minnesota, St. Paul.
- Sung, Y.-C., and J. A. Fuchs. 1988. Characterization of the *cyn* operon in *Escherichia coli* K12. *J. Biol. Chem.* **263**:14769-14775.
- Sung, Y.-C., and J. A. Fuchs. 1992. The *Escherichia coli* K-12 *cyn* operon is positively regulated by a member of the *lysR* family. *J. Bacteriol.* **174**:3645-3650.
- Sung, Y.-C., D. Parsell, P. M. Anderson, and J. A. Fuchs. 1987. Identification, mapping, and cloning of the gene encoding cyanase in *Escherichia coli* K-12. *J. Bacteriol.* **169**:2639-2642.
- Tartof, K. D., and C. A. Hobbs. 1987. Improved media for growing plasmid and cosmid clones. *Bethesda Res. Lab. Focus* **9**:12.
- Thomas, M. S., and R. E. Glass. 1991. *Escherichia coli* *rpoA* mutation which impairs transcription of positively regulated sys-

- tems. *Mol. Microbiol.* **5**:2719–2725.
34. **Thompson, J. F., and A. Landy.** 1988. Empirical estimation of protein-induced DNA bending angles: applications to λ site-specific recombination complexes. *Nucleic Acids Res.* **16**:9687–9705.
35. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* **153**:3–11.
36. **Wang, L., J. D. Helmann, and S. C. Winans.** 1992. The *A. tumefaciens* transcriptional activator OccR causes a bend at a target promoter, which is partially relaxed by a plant tumor metabolite. *Cell* **69**:659–667.
37. **Wek, R. C., and G. W. Hatfield.** 1988. Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. *J. Mol. Biol.* **208**:643–663.
38. **Wu, H.-M., and D. M. Crothers.** 1984. The locus of sequence-directed and protein-induced DNA bending. *Nature (London)* **308**:509–513.
39. **Zinkel, S. S., and D. M. Crothers.** 1991. Catabolite activator protein-induced DNA bending in transcription initiation. *J. Mol. Biol.* **219**:201–215.
40. **Zwieb, C., and R. S. Brown.** 1990. Absence of substantial bending in *Xenopus laevis* transcription factor IIIA-DNA complexes. *Nucleic Acids Res.* **18**:583–587.