

Organization and Transcriptional Regulation of the *Escherichia coli* K-12 D-Serine Tolerance Locus

MADS NØRREGAARD-MADSEN,¹ ELIZABETH MCFALL,² AND POUL VALENTIN-HANSEN^{1*}

Department of Molecular Biology, Odense University, DK-5230 Odense M, Denmark,¹ and Department of Microbiology, New York University School of Medicine, New York, New York 10016²

Received 12 June 1995/Accepted 12 September 1995

We have reinvestigated the genetic organization and the transcription regulation of the *dsd* operon of *Escherichia coli*. By combining genetic and biochemical studies, it is demonstrated that the regulatory region of the operon and the gene encoding the specific regulator of D-serine tolerance (*dsdC*) had been misplaced in previous work on the *dsd* system. Also, the previous erroneous DNA sequence of the *dsdC* gene has been corrected. It turned out that an additional gene (*dsdX*) is present immediately upstream of *dsdA* (encoding D-serine deaminase) and that *dsdC* is located adjacent to *dsdX*. The *dsdXA* genes are cotranscribed from a common promoter region present in the *dsdX-dsdC* intercistronic region. The DsdC activator belongs to the LysR-type of transcriptional regulators and is absolutely required for *dsdA* expression. Additionally, the activity of the *dsdXA* promoter depends on the cyclic AMP receptor protein, and the two activators act in concert to synergistically activate transcription.

Here we report data that have brought forward a number of new aspects of the organization of the *Escherichia coli* D-serine tolerance locus. D-Serine is toxic to *E. coli* K-12 in minimal medium as a result of inhibition of L-serine and pantothenate synthesis (6). Thus there is a need for detoxification, which is accomplished by D-serine deaminase (DsdA). The enzyme degrades D-serine into pyruvate and ammonia in a multistep dehydratase reaction (8, 9) and enables *E. coli* to use this D-amino acid as sole carbon and nitrogen source. Genetic approaches have established that D-serine in complex with a specific activator protein enhances the synthesis of DsdA several-hundredfold, and the expression of *dsdA* is additionally subjected to catabolite control (12, 20).

The deaminase is encoded by the *dsdA* gene, while the specific activator designated DsdC is encoded by the *dsdC* gene. The two genes map at 50 min on the *E. coli* K-12 linkage map (19) and were believed to be transcribed divergently from an unusually large regulatory region, of about 800 bp.

We have investigated the proposed *cis*-acting nature of the D-serine activator (22) by means of low-copy *lacZ* fusions and find that the regulator in fact is *trans* acting. Furthermore, irregularities in the determined sequence of *dsdC* (29) have prompted us to reclone and resequence the region upstream of *dsdA*.

Our data show that the previously described organization of the D-serine tolerance locus in *E. coli* has to be adjusted. First, we have identified a new locus, immediately upstream of *dsdA*, that encodes a 445-amino-acid putative membrane protein homologous to the gluconate permease of *Bacillus subtilis* (10). We designate this gene *dsdX* and the corresponding protein DsdX. Second, the position and the DNA sequence of the gene for the regulatory protein DsdC have now been determined unambiguously. It is located adjacent to *dsdX*, and DsdC belongs to the LysR family of bacterial transcriptional regulators. The *dsdA* and *dsdX* genes constitute an operon, and the promoter region is present in the *dsdXA-dsdC* intercistronic re-

gion. Finally, we characterized the transcriptional regulation of the *dsdA* and *dsdX* genes by using *lacZ* transcriptional fusions.

MATERIALS AND METHODS

Bacterial strains. The *E. coli* strains used were NM1000 (*dsdA*⁺ *dsdC*⁺ *Δlac169 strA thi::Tn10*), NM1002 (HfrH *ΔdsdC-dsdC Δgal-uvrB relA Δlac169 strA thi::Tn10*), and TG-1 (*Δlac-pro supE thi hsdD5/F' traD36 proA*⁺ *proB*⁺ *lacI*^Δ *lacZΔM15*) (the latter was obtained from Amersham).

Plasmids. pJM241 is a 6.4-kb single-copy R1 runaway plasmid with an *EcoRI-BamHI* linker (18). To create pNM29, a 4-kb fragment stretching from *SalI* downstream of *dsdA* to *PvuII* (converted to an *EcoRI* site in pMM1) upstream in *dsdC* was isolated from pMM1 (25) and cloned in the *EcoRI-BamHI* sites of pJM241. To create pNM30, a 14-kb *EcoRI-BamHI* fragment was isolated from cosmid E145 of the *E. coli* K-12 W3110 cosmid collection (32) and cloned in the *EcoRI-BamHI* sites of pJM241. Strains carrying plasmid pNM30 were isolated by screening for the DsdA phenotype in strain NM1002 (*ΔdsdA*). To create pNM31, a 3-kb fragment stretching from *BfiI* downstream of *dsdA* to *PvuII* in *dsdC* was isolated from pNM30 and cloned in the *SmaI* site of pUC13. Subsequently, the *EcoRI-BamHI* fragment from this plasmid was cloned in the *EcoRI-BamHI* sites of pJM241. To create pNM44, a 4.2-kb fragment stretching from *BfiI* downstream of *dsdA* to *EcoRI* downstream of *dsdC* was isolated from pNM30 and cloned in the *EcoRI-SmaI* sites of pUC13. From this construct, the *EcoRI-BamHI* fragment was cloned in the *EcoRI-BamHI* sites of pJM241.

pJEL250 is a 12-kb single-copy runaway plasmid containing a promoterless *lacZ* gene preceded by an *EcoRI-SmaI-BamHI* linker and a *trpA* translation terminator (16, 34). To create pNM21, a fragment stretching from *SmaI* in *dsdA* to *PvuII* (*EcoRI*) in *dsdC* was isolated from pMM1 and cloned in the *SmaI-EcoRI* sites of pUC13. From this construct, an *EcoRI-BamHI* fragment was cloned to *EcoRI-BamHI* of pJEL250, giving a *dsdA-lacZ* transcriptional fusion. To create pNM22, a fragment stretching from *SphI* in *dsdX* to *SmaI* in *dsdC* was isolated from pNM30 and cloned in the *SmaI* site of pUC13. From this construct, an *EcoRI-BamHI* fragment was cloned in the *EcoRI-BamHI* sites of pJEL250, giving a *dsdA-lacZ* transcriptional fusion. To create pNM23, a fragment stretching from *MscI* in *dsdX* to *PvuII* in *dsdC* was isolated from pNM30 and cloned in the *SmaI* site in pUC13. The *EcoRI-BamHI* fragment from this construct was isolated and cloned in the *EcoRI-BamHI* sites of pJEL250, giving a *dsdX-lacZ* transcriptional fusion. To create pNM25, an *EcoRI-BamHI* fragment was isolated from pNM31 and cloned in the *EcoRI-BamHI* site of pJEL250, giving a *dsdA-lacZ* transcriptional fusion.

pGB2 is a low-copy-number plasmid containing an *EcoRI-SmaI-BamHI* linker. The plasmid confers resistance to spectinomycin and streptomycin (4). To create pNM26, an *EcoRI-StuI* fragment isolated from pNM21 was cloned in the *EcoRI-SmaI* sites of pGB2, giving a *dsdA-lacZ* transcriptional fusion.

Media. Cells were grown in LB medium (1) or AB minimal medium (5) supplemented as indicated.

Primer extension analysis. Cells were grown as described in the legend to Fig. 4. Total cellular RNA was isolated as described elsewhere (11). Primer extension was performed essentially as described by Moazed et al. (28). Approximately 15 μg of RNA was used per reaction. For analysis of the *dsdXA* promoter transcription start site, primer ODS-1 (5'-CCCCAGCGTTCGCC-3'; see Fig. 3) 5'-end labeled with ³²P was used.

* Corresponding author. Mailing address: Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark. Phone: 45 66158600. Fax: 45 65932781. Electronic mail address: kathe@pr-group.ou.dk.

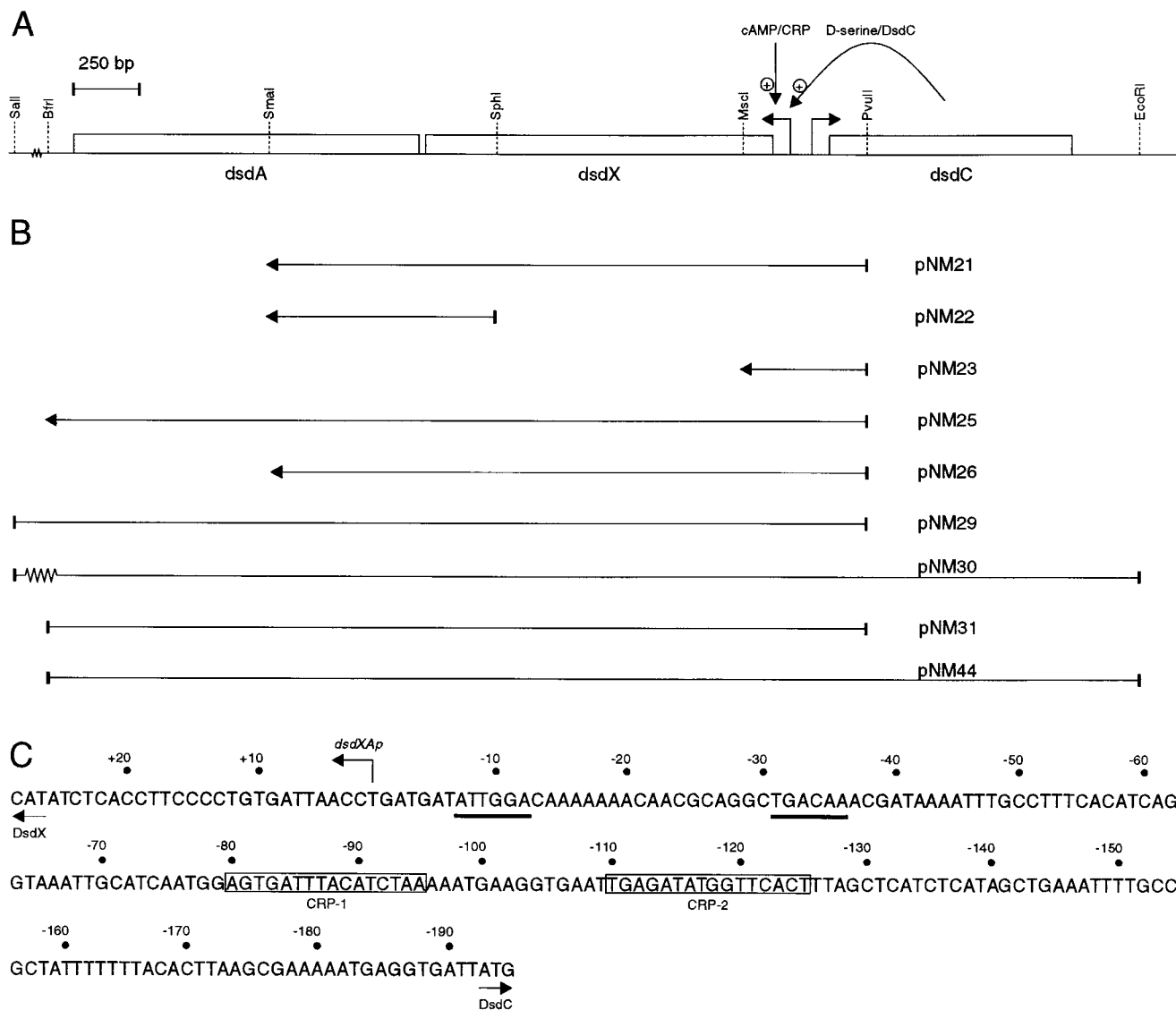


FIG. 1. (A) Model for the organization and regulation of the *dsd* locus. Restriction sites used for cloning are indicated. The kinked arrows indicate promoters. Regulatory sites for the D-serine-DsdC complex and cAMP-CRP complex are indicated by the straight and curved arrows. (B) Fragments used in this study to analyze the *dsd* locus. The waved line in pNM30 indicates that the cloned fragment stretches further 8 kb downstream of the *SalI* site; for the remaining constructs, arrows indicate the direction of transcription into *lacZ*. (C) Nucleotide sequence of the upper strand of the *dsdX-dsdC* intergenic region. +1 is at the start site of the *dsdXA* promoter transcription. -10 and -35 sequences of the *dsdXA* promoter are underlined. The boxed sequences indicate putative binding sites for the cAMP-CRP complex. The kinked arrow indicates the start of transcription, and the straight arrows indicate the start of translation.

Enzyme assays. (i) β -Galactosidase assay. Samples were taken for measurements of enzyme levels during exponential growth at 35°C in AB minimal medium or LB medium supplemented as indicated in the relevant tables. Samples were handled as described previously (27). Activities are expressed as optical density at 420 nm (OD_{420})/ OD_{450} /minute.

(ii) D-Serine deaminase assay. The assay was performed essentially as described by McFall (21). Samples were taken for enzyme measurements every hour for 7 h from cultures grown in AB minimal medium supplemented as indicated in the legend to Fig. 2. Activities are expressed as (OD_{520} / OD_{450} /20 min) \times 100.

DNA sequencing. DNA sequencing was performed on alkaline-denatured plasmid DNA preparations, using the Sequenase enzyme (U.S. Biochemical) as described by the supplier.

Nucleotide sequence accession number. The sequence of the *dsdX* and *dsdC* genes described here has been assigned EMBL accession number X91821.

RESULTS

The region upstream of *dsdA* does not encode a cis-acting regulator. Previous work on the *dsd* system led to a model in which the D-serine deaminase gene, *dsdA*, and the regulatory

gene, *dsdC*, were transcribed divergently from a common regulatory region of approximately 800 bp. DsdC, the product of the regulatory gene *dsdC*, and D-serine were found to be absolutely required for expression of *dsdA*, and several lines of evidence suggest that DsdC has a preference for acting in *cis* (22). To further examine the organization of the *dsd* operon and the proposed model for D-serine deaminase expression, we first constructed a *dsdA-lacZ* transcriptional fusion on the low-copy-number plasmid pGB2 (Fig. 1B). Plasmid pNM26 carries a 2.3-kb fragment stretching from the *SmaI* site in *dsdA* to *PvuII* downstream of the region that is believed to encode the *dsdC* gene. The regulation of this construct in response to D-serine was assessed by measuring the β -galactosidase activity in strain NM1002 (Δdsd) grown in LB medium (Table 1). Most surprisingly, this construct is completely inert and thus cannot support induction of *lacZ* expression. The most straightforward explanation for this finding is that there is no D-serine-

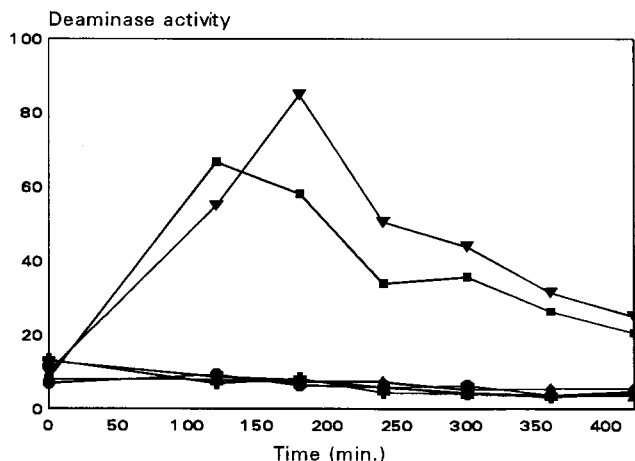


FIG. 2. Deaminase activity in response to D-serine. Symbols: \blacklozenge , NM1002; \bullet , NM1002/pJM241; \triangle , NM1002/pNM29; \blacktriangle , NM1002/pNM31; ∇ , NM1002/pNM30; \blacksquare , NM1002/pNM44. The cultures were grown in AB minimal medium with glucose as the carbon source, supplemented with 50 μg of D-serine per ml. The unit of activity is $(\text{OD}_{520}/\text{OD}_{450}/20 \text{ min}) \times 100$.

responsive regulated promoter on the cloned fragment or, alternatively, the region upstream of *dsdA* does not encode the DsdC protein.

Cloning of an intact chromosomal *dsd* locus. To define additional loci required for *dsd* regulation, we recloned the *dsd* region present on a 14-kb chromosomal *EcoRI*-*Bam*HI fragment of cosmid E145 (32) into the R1 low-copy-number vector pJM241 (generating pNM30). The restriction pattern of the cloned 14-kb fragment is in agreement with a previously determined restriction pattern of the *dsd* region (3). Moreover, the presence of pNM30 in the Δdsd strain NM1002 confers resistance to high concentrations of D-serine, suggesting that this plasmid bears a D-serine-responsive regulator. To test this idea, the *dsdA-lacZ* fusion plasmid pNM26 and plasmid pNM30 were cotransformed into strain NM1002. As shown in Table 1, *lacZ* expression activity directed from the *dsdA-lacZ* fusion is now activated (about 15-fold) by addition of D-serine. Taken together, these results show that (i) pNM30 must encode one or more *trans*-acting factors and (ii) a D-serine-responsive promoter for *dsdA* transcription is present on the *dsd-lacZ* fusion (pNM26). The region of pNM30 encoding the regulatory factor(s) was subsequently mapped by the construction of upstream and downstream deletion derivatives (pNM29, pNM31, and pNM44; Fig. 1B). The levels of D-serine deaminase enzyme expressed from these constructs were assessed in cultures grown in minimal medium with glucose as the carbon source (Fig. 2). The activity is recorded as specific activity against time of addition of D-serine (50 $\mu\text{g}/\text{ml}$). In this assay, we do not find enzyme activity expressed from pNM29 and pNM31 above the background level. However, in cells containing plasmids pNM30 and pNM44, the expression of *dsdA* is strongly induced; the deaminase activity peaks after 2 to 3 h, whereafter it declines because of depletion of D-serine in the medium. Thus, we conclude that only pNM30 and pNM44 encode an intact activator and, contrary to previous

TABLE 1. β -Galactosidase activities in strain NM1002 (Δdsd)^a

Plasmids	β -Galactosidase activity (U)		Fold activation
	-D-serine	+D-serine (1,000 $\mu\text{g}/\text{ml}$)	
pNM26 + pJM241	0.09	0.09	1.0
pNM26 + pNM30	0.70	8.70	14.5

^a Grown in LB containing ampicillin (50 $\mu\text{g}/\text{ml}$) and spectinomycin (100 $\mu\text{g}/\text{ml}$).

belief, that sequences upstream of the *Pvu*II site are absolutely required for transcription activation. These results prompted us to resequence the *dsdA* upstream region.

Identification of ORFs in the *dsdA* upstream region. First, we determined the sequence of the *dsdA* upstream region of pNM30 (Fig. 3). A search with the MAP program (7) for open reading frames (ORFs) in the *dsdA* upstream sequence detected one ORF of significant length (Fig. 3). This ORF, designated *orf445*, is 445 amino acids in length (molecular mass = 47 kDa), the stop codon being located 19 bp from the ATG start codon of DsdA. A FASTA (7) search in the SwissProt database (release 30.0) with the Orf445 polypeptide as the query sequence gave one hit of high homology (37% identity, 64% similarity) to the GntP protein of *B. subtilis*, which is a permease for gluconate (10). Compared with the sequence determined by Bornstein-Forst et al. (2), two errors were found in the 550-bp region adjacent to *dsdA*; two base pairs (GT) at position 857 (now deleted) and a CG base pair at position 960 (now inserted; note that this base pair is also present in the original sequence determined by McFall and Runkel [24]). The previously published sequence of the region further upstream of *dsdA*, however, contain numerous errors as well as insertions of nonrelated sequences (29).

Next the sequence of the region upstream of *orf445* was determined (Fig. 3). We find a single ORF of significant length. This ORF of 311 codons (molecular mass = 35 kDa) reveals by FASTA (7) analysis homology to the large family of LysR-like transcriptional regulators (13) and is closely linked to *orf445*, with a polarity opposite that of *orf445* and *dsdA*.

Taken together, our results show that (i) the regulator of the *dsd* operon, DsdC, is encoded by *orf311* (consequently, we designate this locus *dsdC*) and (ii) an additional gene, encoding an inner membrane-like protein, is situated between *dsdA* and *dsdC*. We designate this gene *dsdX* and the encoded protein DsdX, as its function is not yet known.

***dsdA*, *dsdX*, and *dsdC* are transcribed bidirectionally from a common promoter region.** To determine the transcriptional regulatory regions of *dsdA* and *dsdX*, we constructed transcriptional *lacZ* fusion plasmids with various DNA fragments encompassing the *dsd* locus (Fig. 1). *LacZ* expression from pNM21, pNM23, and pNM25 when present in strain NM1000 (*dsd*⁺; grown in LB with or without 1,000 μg of D-serine per ml) is activated in the presence of D-serine, while plasmid pNM22 is transcriptionally inactive (Table 2). Thus, it follows that the promoter responsible for *dsdA* transcription has to be located in the *dsdXA-dsdC* intercistronic region.

To detect transcripts initiating upstream of *dsdX*, we used

FIG. 3. Nucleotide sequence of *dsdX* (*orf445*), *dsdC*, and the intercistronic region. Below the nucleotide sequence are shown the translated peptide sequences of DsdX and DsdC. DsdX is homologous to gluconate permease (GntP) from *B. subtilis*. Homology between GntP and an 11-kDa hypothetical protein preceding *dsdA* has previously been noted by Reizer et al. (30). DsdC belong to the LysR family of proteins and is related most closely (50% similarity) to the *E. coli* glycine cleavage enzyme activator protein (GcvA) (35). The bold underlined sequence indicate the position of primer ODSD-1, used for primer extension analysis.

1 CATAGTACGCTCTTTTGGCGTTAGATGATAAATGACAGCAGGAATGTGCCGCCAGAGGGACGACTGAAGCGATAAATGTCGCTGTCGTATAGTATTTAA 100
GTAATCATGAGAAACCGGAATCTACTATTTACTGTCTGCTTACACGGGCGGCTCGCTGCTGACTTCGCTATTTACAGCGAACAGCATATCATAAATTT
M * I I F S L L F T G A L A V V S A I F T A T T Y Y K
DsdA start
101 ATGTTTCATTTAGCGTAGCGCCGCAATATTTGCTTACTAGCCAGAAAAGCGAGTCCGTAACGATAGTGCAGCCAAATGACCTGAACCGATAGCAATGCG 200
TACAAAGTAACCTCGATCGCGGCTTATAACGAAAGTATCGGCTTTTTCGCTAGGCATGCTATCAGCTCGGTTAACGTTGGACTTGGCTATGCTTAGG
F T E N L T A G C Y Q K V L W F L S D T V I T C G I A G S G I A I A
201 AATAATTTCCGGGCTGATGTGGGATACAGCGGCAGCATGGGTGCAACAATTTGCCGTTGCCCCCATCATTTGCCAGGTAGCGGAGCCCATCGCGCATGC 300
TTATTAAGGCCCGACTACAGCCCTATCTCGCGCTGTAACCCAGCTTGTAAACGGCAACGGGGTAGTAACGGTCCATCGCTCGGCTGAGCGGCTAGC
I I E P S I D P Y L P L M P A V I A T A G M M A V T A S G V A A H
301 AGAATAAGAGCCACTAACCCAGCCAGAAATCGGGTGCATATGTCATTTGGAGAGAATAAAGTCCAGCGGTATCAGCGAGACTGCTGCTTTTAAATG 400
L I L A V L W A L L I P H M H M N S L I V A L T D A L S S S K L I
401 CGTTGAATGCGCCTCCGGCCCGATAATCAGCAAAATATTAGCAATAGAACCAGGCAATTTCCGATGCTGAGCATGCTCCCATGCTCATATGCTG 500
GAACTTTACCGCGAGGCGCGGCTATTAGTCTGTTTATAATCTTATCTTGGCTTCGGTAAAGGCATAGACACTCGTAGCAGGGGTACGAGTATAGCAG
A N F A G G A G I I L L I N A I S G F G N E T H T L M T G M S M H Q
501 GCGTATACCCAAACATAATAGCGCAACACCGCGATAAACATGGCAGTATAGGGTTGCCAATAAATCAACCAGATATACAAACCACTCTCACCG 600
CGCATATGGGTTCTGCTATTTATCCGCTATTGTGCGCTATTGTACCGTCACTATCCCAAGGTTATTGAGTGGTTCATATGTTGGTGGAGTGGC
R I G L V Y Y A V F V A I F M A T I P N G I F E V L I Y L G S E R
601 GCCATATTCATTTCCGCAATCGTTTTAACCAACATCAGCGCAATGGTAGCAGTATGGTAACGTTGCCCTAATGACGGTACTGTTTTTCATCGC 700
CGGTATAAGTTAAGCGCTAGCAAAATGTTGTTAGTCCGCTACCCATCGCTATACCATTTATGCAACGGGATTAAGTCCATACAAATAAGTATGAGG
A M N L E A I T K V L M L A I P L L I T F L T A G L S P L T K E D
701 GAACCTTGAGATCTGCAAACTCTGTGGTACAGGTTTAAAGGCGAGTCTGTGACCCAGAAATTTAAGGAAAGTGGGCCACCGATCAGTATGCCATCAG 800
CTTGAACCTTAGACGTTGAGACACCCATGTCCAAATTTCCCGTCAGCAACTGGGCTTTAAATTCCTTTTACCAGGTTGGTACTCAGTACAGTATC
R V K L D A F E T P V P K F L R Q G L F K L F L P G G I L S A M L
801 CCCAACAGCAACCGTAGACATCAGCGAACCAGATATCTGCCGCCAGCTTATTGGCAACATATAAAGCAGCCGATGTTGGAGAACCCAGCAGTGCCT 900
GGTGTGGTCTGTTGCATCTGCTAGTGGCTTGGCTATAGACCGCGGTCGAAATACCGTTGTATMTTCCTGCGCCTACACCTCCTTGGTCCGTCACGTA
G V L L G Y V I V S G I D A G L K N A V Y L A A P H P P V V C H V
901 GCCATCAATCGCGTACATAGCGGAATGGCAAGCTTTAATTAAGGATTTGGTTTTTTTGGCAATGAAAAGCCAGAGGAATCAATGACGACGAGCCCA 1000
CGGTATGTTGCGGCTAGTATCGCCTTACCGTTTCGAAATTTACTCCATAACCAAAAAACCGTTAACTTTTTTGGTCTCCTTAGTATCGGCTGAGGCGG
A M L A T C L P I A L K L L S T N T K K A I S F A L P I L L V V G
1001 CTTCAACAAACAGCGTATGCCAAATCAGGCCAACAGCACCATAATGACATCAACTGAAAGCCAGCGGCAGCGTTGAAAGTGCAGACCAATTCITTC 1100
GAAGTGTGTTTTCGCGACTACCGTGTATTAGTCCGTTGGTCTGCTGTTACTGATGTTGACTTTTCGGTCCGCGCTCGCACTTCACAGTCTGGTTAAGAAAG
V E V F L T I G C I L G V L V M I V D V S L W R C R Q L T L G I R E
1101 TGGCGGCGGATACTTCCATCATTTTCCAGTATCGTCCAAAGCCGATAACCGCTGGGAGGAAACCCAGCGTCCCGCAATTCACCTTTCAATAGCA 1200
ACGCGGGCGCTATGAAGTGTAAAGGTCATAGCAGCGTCCCGCTATTGGCCAGCTCCTTGGGTCGCAAGCGGTTAAGGTGAAGATTACTG
A A G S V E M M K G L I T G L G I V A A L F G L T G G I G S E I A
1201 TTTACCATATCCAGTGGGCCCATGCCATCATCGTTCCACGAAAGCTGGCCAGCAACAGCGCCAGAAACCGGGTGAATTTGAATTCACGATGCTCA 1300
AAATGATATAGGTCACCGGGTACGGGTAGTAGCAAGGTTGCTTCTTCGACCGGCTGTTCTGCGGCTTTCGCCACCTTAACTTGAAGTGCATCCAGT
N V M D L P G M G M T G V F F S A L L L A L F P H F K V I T
1301 GTACAATTAACAGCATGCTGATAAGCAGCGTGTACAAACCCAGATTGAGAGTGCATATCTCACCTTCCCGTGTGATTAACCTGTGATTAACCTGATGA 1400
CATGTTAATTTGCTACGACTATTCGTGCGCAGGATGTTGGTCTAAACTCTCAGTATAGAGTGAAGGGGACACTAATTTGGACACTAATTTGGACTACT
L V I L V I S I L L T S V V W I Q S H M
DsdI start
1401 TATTGGACAAAAAACACCGAGCTGCAACAGATAAAATTTGCGCTTTCACATCAGGTAATTTGCATCAATGGAGTGAATTTACATCTAAAAATGAAGT 1500
ATAACCTGTTTTTTTGTGCTCGGACTGTTGCTATTTTAAACGGAAAGTGTAGTCCATTTAACGTAGTTACCTCACATAATGTAGATTTTTACTTCCA
1501 GAATTTGAGATATGCTCACTTTAGCTCATCTCATAGCTGAAATTTTGGCGCTATTTTTTTTACACTTAAAGGAAAAATGAGGTGATTTAGGAAACCCCTTCG 1600
CTTAACCTATACCAAGTGAATCGAGTAGATTCGACTTTTAAACGGCGATAAAAAATGTAATTCGCTTTTTTACTCCACTAATTAACCTTGGGGAAGC
M E P L R
DsdC start
1601 TGAATAAGAAATCGGCTGCTTAAACGGCTGGCAATATCAAAAATGCATACTTTTGAAGTGGCTGCCAGGCATCAGTCTTCCGCTGGCGGACAGAGGAA 1700
ACTTTATCTTAGCGGACGAAATGCGGACCGTTAATAGTTTTTACGATGAAACTTCACCGAGCGCTAGTTCAGGAAAGCGGACCGCGCTCCCTT
E I R N L L N G W Q L R S K M H T F E V A A R H Q S F A L A A E E
1701 TTGTGCTGAGCGCCAGTGGGTAAGTACCGTATCAATCAGCTGGAAGAATAATGGGCATTCAGTTGTTTGTGTTCCCATCGCAAGTGAATTA 1800
AACAGGACTCGGGCTACGCGCATTCAGTGCATAGTGTAGTGCCTTCTTCTTAAACCGTAAAGTCAACAACAAGCAAGGGTACGCTTACCTTAAT
L S L S P S A V S H R I N Q L E E E L G I Q L F V R S H R K V E L T
1801 CGCACGAGGGGAAACGTTTATTTGGGCGCTAAATTCGCTGGATACCTGAAACAGGAAATCTTGTATATAAAAAATCAGGATTTATCGGAAACGTT 1900
GGTCTCCCTTTGCAAAAATAACCGCGATTTTACAGCGCATATGGGACTTGGTCTTTAAGAACTATAGTTTTTTGTGCTCAATTAAGGCTTGCAAA
H E G K R V Y W A L K S S L D T L N Q E I L D I K N Q E L S G T L
1901 AACTCTGATTTCCCGGCCCTTATCGCCCAATGCTGTTGTTGCCCGCATTAGTGAATTTACACCGCGGTATCCGCTTATTTCCGCTACCGTCTCACT 2000
TTGAGACATAAAGGGCGGAGATAGCGGTTACGACCAACCAAGCGGCTAATCCACTGAANTGTGCGGCCATAGGCAGATAAAGCGAGTGGCAGAGTGA
T L Y S R P S I A Q C W L V P A L G D F T R R Y P S I S L T V L T
2001 GGTAAATGACAACGTAATCTGCAACGTTGCCGGAATCGATTGGCGATCACTTTGATGATCGGCGCTCAGCGCAACTCACTCACTTTCTGATGGATG 2100
CCATTTACTGTTGAGTGTAGACTTTGACAGCGCTTACGTAACCGCTAGATGAAACTACTAGCGGCGAGTGGCTTGTAGTGTAGTGAAGACTACCTAC
G N D N V N L Q R A G I D L A I Y F D D R P S A Q L T H H F L M D E
2101 AAGAAATCTGCCAGTATGAGCCGGAAATACGCTCAAAGACATCTTTAACCAACAGGTAATTAACCTGTGCTACTGTACGTTGCTCCATGACAGACA 2200
TTCTTTAGGACCGGTACATACCGGCTTATGCGAGTTTCTTACGAAATTTGGTGTGCCATTAATTTGACACAGTGCATGCAACGAGTACTGCTGT
E I L P V C S P E Y A Q H L L A L T N T V I N L C H C T L L H D R Q
2201 GGCATGGAGCAACGACTCCGGTACGGATGAATGGCATAGTTGGCGCAACATTTATCGGTTAATTTGCCGACATCTTCTGGAATTTGGCTTTGATCGTTCT 2300
CCGTACTCTGTTGCTGAGGCGCATGCTACTTACCCTATCAACCGCGTTGTAATACGCCAATTAACCGGCTGTAAGAACCTTAAACCGAACTAGCAGGA
A W S N D S G T D E W H S W A Q H Y A V N L P T S S G I G F D R S
2301 GATTTAGCTGTTATTTGCCGGATGAATCATATTTGGGGTGGGATGGGAAGAAAACGCCTGGTACAAAAAGGCTTCCAGTGGTGGCTGCTGCGCGCT 2400
CTAAATCGCAATAACCGCGCTACTTATGATAACCCACCGTACCCTTCTTTTGGGACCATGTTTTTCCGAAAGGTCACCACTCGAGCAGCGCGGCA
D L A V I A A M N H I G V A M G R K R L V Q K R L A S G E L V A P F
2401 TTGGCGATGAGCGTGAATGCTACGATTTTACATCACACATTTACCGGCGAGCGAGTGGCCAAAAATTTGAAGCGTTTTATTTCTGGCTAAGAGA 2500
AACCGCTACTGCTTACCTTACCGTACTGTAATAATGTTAGTGGTGAATTTGGCGCTCCGTCACCGGTTTTTAACTTGGCAATAATAGACCGATCTCT
G D M T V K C H Q H Y Y I T T L P G R Q W P K I E A F I I W L R E
2501 ACAGGTAACCAACCTTACATAATGAATAAACCAACAGGATTTCCGTGTGCTTATTCCGCTGCTAGTTGCGAATTTAATTTACCGTTTTTATGTTGGTTTT 2600
TGCTCAATTTTTGTTGAAATTAATTTTGGTCTCAAGGACCACAGGAAATAGCCGAGCAATCAAGCTTAAATTAATGCCAAAAATACACAAAA
Q V K T T S *

TABLE 2. β -Galactosidase activities in strain NM1000 (*dsd*⁺)^a

Plasmid	β -Galactosidase activity (U)		Fold activation
	-D-Serine	+D-Serine (1,000 μ g/ml)	
pJEL250	0.04	0.04	1.0
pNM21	0.33	2.60	7.9
pNM22	0.02	0.02	1.0
pNM23	1.40	16.10	11.5
pNM25 ^b	0.15	0.95	6.8

^a Grown in LB containing ampicillin (50 μ g/ml).

^b The low expression from pNM25 is due to the presence of a transcriptional terminator just downstream of *dsdA* (17).

primer ODS-D-1 (Fig. 3), which is complementary to sequences in the beginning of *dsdX*. The primer extensions were performed with total RNA isolated from strains NM1000 (*dsd*⁺), NM1002 (Δ *dsd*), and NM1002/pNM44, all grown in LB medium with or without 1,000 μ g of D-serine per ml. As shown in Fig. 4, the analyses revealed a single D-serine-induced transcript with RNA isolated from NM1000 or NM1002/pNM44 (compare lanes 1 and 2 with lanes 3 and 4). The 5' end of this transcript is located 26 nucleotides upstream of the first putative ATG start codon of *dsdX* and is preceded by a putative σ ⁷⁰ -10 box (-7 to -12) with the sequence TCCAAT as well as a putative σ ⁷⁰ -35 box (-31 to -36) with the sequence TTGTCA.

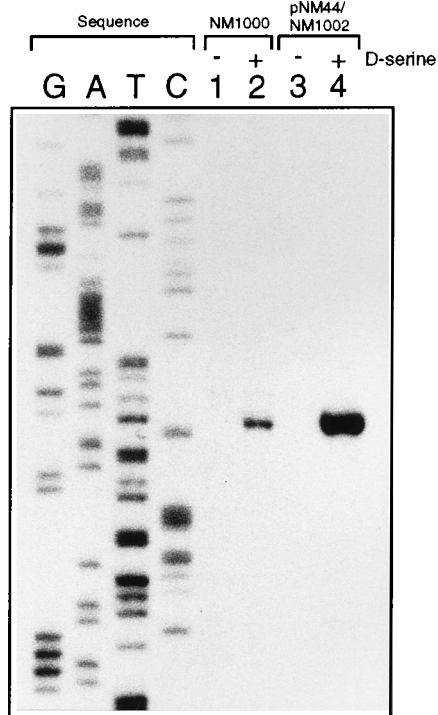


FIG. 4. Primer extension analysis of *dsdX*-specific transcripts. The G, A, T, and C sequence reactions were performed with plasmid pNM44, using the same 5'-end-labeled primer as was used for the primer extension reactions. Total RNA was isolated from NM1000 grown in LB at 35°C (lane 1), NM1000 grown in LB plus 500 μ g of D-serine per ml at 35°C (lane 2), NM1002/pNM44 grown in LB plus 50 μ g of ampicillin per ml at 35°C (lane 3), and NM1002/pNM44 grown in LB plus 50 μ g of ampicillin and 500 μ g of D-serine per ml at 35°C (lane 4). Note that before isolation of the RNA used in lanes 3 and 4, the copy number of pNM44 was amplified at 42°C for 15 min, which increases the ratio of specific to nonspecific mRNA.

DsdC and the cAMP-CRP complex are activators of *dsdXA* transcription. Previous reports (20) established that D-serine deaminase regulation is subjected to catabolite control. To clarify the role of the cyclic AMP (cAMP)-cAMP receptor protein (CRP) complex in the control of the *dsdXA* promoter, we measured the D-serine-induced transcriptional activity of plasmid pNM23 when present in three different genetic backgrounds: wild-type, Δ *dsd*, and Δ *crp* strains (Table 3). In the wild-type strain, the *dsdX* promoter is activated >200-fold by D-serine. In the Δ *crp* background, we find that pNM23 is still activated by D-serine but to a lower extent (31-fold). The basal level of β -galactosidase activity observed in each of the two strains is equal to the level seen in the Δ *dsd* strain.

Thus, it appears that DsdC can act as an independent activator of *dsdXA* transcription. The presence of the cAMP-CRP complex has a stimulatory effect on activation; however, CRP does not activate expression alone or directly.

DISCUSSION

Over the past decades, genetic and biochemical studies together with molecular cloning of specific genes have provided a rather comprehensive view of the organization and regulation of the *dsd* operon (for a review, see reference 23). However, a deeper understanding of the regulatory features of this system has been hampered by irregularities in the structure of some basic plasmids, in part of the sequencing work, and in mapping start sites for transcription (2). Moreover, these matters have led to a number of incorrect conclusions, e.g., that DsdC is a strictly *cis*-acting regulator (22).

The experiments presented here clearly establish that the organization of the *E. coli dsd* locus has to be adjusted as outlined in Fig. 1A. An additional gene, *dsdX*, is present immediately upstream of *dsdA*, and the specific regulator of D-serine tolerance (DsdC) is encoded by a gene located next to *dsdX*. The *dsdA* and *dsdX* genes are cotranscribed from a common promoter region present in the 219-bp-long *dsdX*-*dsdC* intercistronic region (Fig. 1C).

The DsdC protein belongs to the large and expanding family of LysR-like transcriptional regulators, and our regulatory studies show unambiguously that DsdC is a *trans*-acting regulator.

The role of the DsdX protein is uncertain. On the basis of homology arguments, we suggest that it is an inner membrane protein and may function in transport of D-serine. However, at present we cannot rule out the possibility that also DsdX is part of the *dsd*-regulatory apparatus.

Our regulatory studies show that expression of *dsdA* includes two activating elements, the specific activator-D-serine complex and the cAMP-CRP complex and thus confirm the previous proposed regulatory model. The DsdC activator can act independently and is absolutely required for transcription activation of the *dsdXA* promoter. Expression from this promoter can be further enhanced (about 7-fold) in the presence of the cAMP-CRP complex, leading to a more than 200-fold activation of transcription. The positive effect of the cAMP-CRP complex, however, is seen only in concert with DsdC. Thus, CRP cannot act on its own in this system. Accordingly, two putative targets for the cAMP-CRP complex, CRP-1 and CRP-2, with good match to the consensus CRP-binding site (TGTGAN₆TCACA [14]), are present in the *dsdXA* promoter region (Fig. 1C); CRP-1 is centered around -87.5, and CRP-2 is centered around -117.5, two positions at which the cAMP-CRP complex is unable to activate transcription as the sole activator (33). Clearly, the *dsdXA* promoter belongs to the more complex class III CRP-dependent promoters, in which a

TABLE 3. β -Galactosidase activities in minimal media

Plasmid	NM1000 (wild type) ^a			NM1002 (Δdsd) ^b			SØ2929 (Δcrp) ^c		
	β -Galactosidase activity (U)		Fold activation	β -Galactosidase activity (U)		Fold activation	β -Galactosidase activity (U)		Fold activation
	-D-Serine	+D-Serine		-D-Serine	+D-Serine		-D-Serine	+D-Serine	
pJEL250	0.09	0.09	1	0.08	0.08	1	0.04	0.04	1
pNM23	0.04	9.1	228	0.04	0.04	1	0.04	1.25	31

^a Grown in AB minimal medium containing 0.5% glycerol and thiamine (2 μ g/ml) with or without D-serine (500 μ g/ml).

^b Grown in AB minimal medium containing 0.2% glucose, thiamine (2 μ g/ml), and biotin (1 μ g/ml) with or without D-serine (5 μ g/ml).

^c Grown in AB minimal medium containing 0.2% glucose, thiamine (2 μ g/ml), 0.05% Casamino Acids, and D-serine (500 μ g/ml).

regulon-specific activator protein in addition to CRP is required for transcriptional activation. Examples of class III promoters include the *araBAD* (15), the *malE-malK* (31), and the *uhpT* (26) promoters. Interestingly, the roles of CRP in these types of promoters differ. It can be structural, as found in the *araBAD* promoter, in which case a CRP-induced DNA bend acts to disrupt a repressive DNA loop, or steric, as in the *malK* promoter, in which case CRP induces a repositioning of the MalT activator into an activating register. Finally, the action of CRP in the *uhpT* promoter seems to involve a mechanism different from those used in the *malK* and *araBAD* promoters.

On the basis of the present work, it should now be possible to dissect the regulatory features of the *dsd* system. We are currently attempting to isolate the DsdC protein in order to define how this activator in concert with the cAMP-CRP complex synergistically activate transcription from the *dsdXA* promoter.

ACKNOWLEDGMENTS

We thank S. Tabata and A. Higashitani for providing cosmids. This work was supported by the Danish Center for Microbiology.

REFERENCES

- Bertani, G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**:293-300.
- Bornstein-Forst, S. M., E. McFall, and S. Palchaudhuri. 1987. In vivo D-serine transcription start sites in wild-type *Escherichia coli* and in *dsdA* promoter mutants. *J. Bacteriol.* **169**:1056-1060.
- Carothers, A. M., E. McFall, and S. Palchaudhuri. 1980. Physical mapping of the *Escherichia coli* D-serine deaminase region: contiguity of the *dsd* structural and regulatory genes. *J. Bacteriol.* **142**:174-184.
- Churchward, G., D. Belin, and Y. Nagamine. 1984. A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene* **31**:165-171.
- Clark, D. J., and O. Maaløe. 1967. DNA replication and the cell cycle in *E. coli*. *J. Mol. Biol.* **23**:99-112.
- Cosloy, S. D., and E. McFall. 1973. Metabolism of D-serine in *Escherichia coli* K-12: mechanism of growth inhibition. *J. Bacteriol.* **114**:685-694.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387-395.
- Federiuk, C. S., R. Bayer, and J. A. Schafer. 1983. Characterization of the catalytic pathway for D-serine dehydratase: evidence for variation of the rate-determining step with substrate structure. *J. Biol. Chem.* **258**:5379-5385.
- Federiuk, C. S., and J. A. Schafer. 1983. A reaction pathway for transamination of the pyridoxal 5'-phosphate in D-serine dehydratase by amino acids. *J. Biol. Chem.* **258**:5372-5378.
- Fujita, Y., T. Fujita, Y. Miwa, J.-I. Nihahsi, and Y. Aratani. 1986. Organization and transcription of the gluconate operon, *gnt*, of *Bacillus subtilis*. *J. Biol. Chem.* **261**:13744-13753.
- Gerdes, K., T. Thisted, and J. Martinussen. 1990. Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. *Mol. Microbiol.* **4**:1807-1818.
- Heincz, M. C., and E. McFall. 1978. Role of the *dsdC* activator in regulation of D-serine deaminase synthesis. *J. Bacteriol.* **136**:96-103.
- 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**:6602-6606.
- Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya. 1993. Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* **62**:749-796.
- Lobell, R. B., and R. F. Schleif. 1991. AraC-DNA looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. *J. Mol. Biol.* **218**:45-54.
- Løve-Larsen, J. E., and P. Valentin-Hansen. Unpublished data.
- Marceau, M., E. McFall, S. D. Lewis, and J. A. Schafer. 1988. D-serine dehydratase from *Escherichia coli*. DNA sequence and identification of catalytically inactive glycine to aspartic acid variants. *J. Biol. Chem.* **263**:16926-16933.
- Martinussen, J. Unpublished data.
- McFall, E. 1967. Mapping of the D-serine deaminase region in *Escherichia coli* K-12. *Genetics* **55**:91-99.
- McFall, E. 1973. Role of adenosine 3',5'-cyclic monophosphate and its specific binding protein in the regulation of D-serine deaminase synthesis. *J. Bacteriol.* **113**:781-785.
- McFall, E. 1975. *Escherichia coli* K-12 mutant forming a temperature-sensitive D-serine deaminase. *J. Bacteriol.* **121**:1074-1077.
- McFall, E. 1986. *cis*-acting proteins. *J. Bacteriol.* **167**:429-432.
- McFall, E. 1987. The D-serine deaminase operon, p. 1520-1526. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umberger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
- McFall, E., and L. Runkel. 1983. DNA sequences of the D-serine deaminase control region and N-terminal portion of the structural gene. *J. Bacteriol.* **154**:1508-1512.
- McFall, E., S. S. Nikam, and S. Palchaudhuri. 1991. Effects of structural changes in the *dsdA-dsdC* intergenic region on D-serine deaminase synthesis. *J. Bacteriol.* **173**:1161-1167.
- Merkel, T. J., J. L. Dahl, R. H. Ebright, and R. J. Kadner. 1995. Transcription activation at the *Escherichia coli* *uhpT* promoter by the catabolite gene activator protein. *J. Bacteriol.* **177**:1712-1718.
- Miller, J. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Moazed, D., S. Stern, and H. F. Noller. 1986. Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. *J. Mol. Biol.* **187**:399-416.
- Palchaudhuri, S., V. Patel, and E. McFall. 1988. DNA sequence of the D-serine deaminase activator gene *dsdC*. *J. Bacteriol.* **170**:330-334.
- Reizer, A., J. Deutscher, M. H. Saier, Jr., and J. Reizer. 1991. Analysis of the gluconate (*gnt*) operon of *Bacillus subtilis*. *Mol. Microbiol.* **5**:1081-1089.
- Richet, E., D. Vidal-Ingigliardi, and O. Raibaud. 1991. A new mechanism for coactivation of transcription initiation: repositioning of an activator triggered by the binding of a second activator. *Cell* **66**:1185-1195.
- Tabata, S., A. Higashitani, M. Takanami, K. Akiyama, Y. Kohara, Y. Nishimura, A. Nishimura, S. Yasuda, and Y. Hirota. 1989. Construction of an ordered cosmid collection of the *Escherichia coli* K-12 W3110 chromosome. *J. Bacteriol.* **171**:1214-1218.
- Ushida, C., and H. Aiba. 1990. Helical phase dependent action of CRP: effect of the distance between the CRP site and the -35 region on promoter activity. *Nucleic Acids Res.* **18**:6325-6330.
- Valentin-Hansen, P., B. Albrechtsen, and J. E. Løve-Larsen. 1986. DNA-protein recognition: demonstration of three genetically separated operator elements that are required for repression of the *Escherichia coli* *deoCABD* promoters by the DeoR repressor. *EMBO J.* **5**:2015-2021.
- Wilson, R. L., and G. V. Stauffer. 1994. DNA sequence and characterization of GcvA, a LysR family regulatory protein for the *Escherichia coli* glycine cleavage enzyme system. *J. Bacteriol.* **176**:2862-2868.