Mol Gen Genet (1993) 240:395–402

MGG

TdcA, a transcriptional activator of the *tdcABC* operon of *Escherichia coli*, is a member of the LysR family of proteins

Yojana L. Ganduri, Srinivas R. Sadda*, Milton W. Datta, Rajalakshmi K. Jambukeswaran, Prasanta Datta**

Department of Biological Chemistry, The University of Michigan, Ann Arbor, MI 48109-0606, USA

Received: 21 December 1992 / Accepted: 8 March 1993

Abstract. The *tdcB* and *tdcC* genes of the *tdcABC* operon of *Escherichia coli* encode threonine dehydratase and a threonine-serine permease, respectively. These proteins are involved in transport and metabolism of threonine and serine during anaerobic growth. In this study, we functionally characterized *tdcA*, which encodes a 35 kDa polypeptide consisting of 312 amino acid residues. Nonpolar and partially polar mutations introduced into *tdcA* drastically reduced the expression of the genes downstream from *tdcA*. Complementation studies using single-copy chromosomal integrants of a *tdcB-lacZ* fusion harboring an in-frame deletion of *tdcA* with chromosomal or plasmid-borne *tdcA*$^+$ in *trans* showed complete restoration of *tdc* operon expression in vivo. The amino acid sequene at the amino-terminal end of TdcA revealed a significant homology to the helix-turn-helix motifs of typical DNA binding proteins. Sequence alignment of TdcA with LysR also showed considerable sequence similarity throughout their entire lengths. Our results suggest that TdcA is related to the LysR family of proteins by common ancestry and, based on its functional role in *tdc* expression, belongs to the LysR family of transcriptional activators.

Key words: *Escherichia coli* – *tdc* operon – *tdc* activator – Gene regulation – LysR family of proteins

Introduction

During anaerobic incubation in amino acid-rich medium, *Escherichia coli* synthesizes biodegradative (catabolic) threonine dehydratase (EC 4.2.1.16), which catalyzes the dehydration of L-threonine and L-serine to ammonia and their corresponding keto acids (Umbarger 1978). We have cloned *tdcB* the structural gene for threonine dehydratase, in a 6.3 kb *Eco*RI fragment (Goss and Datta 1985; Datta et al. 1987). A series of biochemical and genetic experiments including DNA sequence analysis, deletion studies, minicell expression and insertion mutagenesis revealed that *tdcB* is the second gene of a polycistronic operon, *tdcABC* (Goss and Datta 1985; Goss et al. 1988; Schweizer and Datta 1988). The *tdcC* gene encodes a membrane-associated L-threonine-L-serine permease (Sumantran et al. 1990), however, the function of the *tdcA* gene product still remains unknown. The operon maps at coordinate 3330 kb, or at 68.3 min, on the *E. coli* chromosome with clockwise direction of transcription (Schweizer and Datta 1988, 1991).

Recently, evidence has been accumulating that anaerobic expression of the *tdc* operon is a highly regulated system subject to global and operon-specific control at the level of transcription. For example, both integration host factor (IHF) and the cAMP-catabolite gene activator protein (CAP) complex bind to specific DNA sequences in the *tdc* promoter and act in concert to positively regulate the *tdc* genes (Wu and Datta 1992; Wu et al. 1992). The efficient in vivo expression of the *tdc* operon also requires the product of a regulatory gene, *tdcR*, located immediately upstream of *tdcABC* in the opposite transcriptional orientation (Schweizer and Datta 1989a). In this report we provide genetic evidence to suggest that the *tdcA* gene product acts as a *trans*-acting factor for positive regulation of the *tdc* operon.

Materials and methods

Bacterial strains and growth media. A list of *E. coli* strains used is given in Table 1. The *tdcB-lacZ* fusions were transferred as single copies to the TH274 chromosome to yield TH274λSH241 and TH274λYG2413 using the lambda vector λRZ5 and following the procedure described by Schweizer and Datta (1988).

Communicated by H. Hennecke

* *Present address*: School of Medicine, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA

** *Present address*: Biomedical Research and Development, The J.M. Holland Laboratory, American Red Cross, 15601 Crabbs Branch Way, Rockville, MD 20855, USA

Correspondence to: P. Datta

Table 1. Strains of *Escherichia coli* used

Strain	Genotype	Source or reference
K37	Str^r *galK*	Friedman et al. (1984)
DH5αF′	Ø 80 d *lacZ* Δ M15 Δ *(lacZYA-argF) U169 recA1 endA1 hsdR17* (r_k^- m_k^+) *supE44* λ*-thi-1 gyrA relA1*	Wu and Datta (1992)
MC4100	F-*araD139* Δ *(argF-lac) U169 rpsL150 deoC1 relA1 rbsR ptsF25 flbB5301*	Wu and Datta (1992)
TH16	MC4100 Δ *tdc-216 zgi*::Tn*10*	Schweizer and Datta (1988)
TH274	MC4100 Δ *tdc-274 zgi*::Tn*10*	Wu and Datta (1992)
TH274λSH241	TH274λ*tdcR*$^+$*tdcAB-lacZ*	This study
TH274λYG2413	TH274λ*tdcR*$^+$*tdcA*^{Δ}*B-lacZ*	This study

Bacteria were grown in Luria-Bertani (LB) medium (Miller 1972) supplemented with appropriate antibiotics, as needed, at the following concentrations: 100 μg/ml ampicillin, 50 μg/ml kanamycin, 15 μg/ml tetracycline, and 30 μg/ml chloramphenicol. For enzyme induction, cells were grown aerobically at 37° C for ~ 12 h in LB and then washed and resuspended in tryptone yeast extract (TYE) medium containing 2% tryptone and 1% yeast extract supplemented with salts and pyridoxine hydrochloride (Hobert and Datta 1983) or in the synthetic H4 medium of Hobert and Datta (1983), which contained four amino acids, threonine, serine, valine and isoleucine, plus cAMP and fumarate. After anaerobic incubation in still culture (without shaking) at 37° C for 8–12 h, cells were collected by centrifugation and washed in 0.1 M potassium phosphate buffer, pH 8.0, containing 3 mM AMP and 1 mM dithiothreitol, and then resuspended in the same buffer for enzyme assay.

Plasmids. The plasmids pEC61, pSH215 and pRS124 have been described previously (Goss and Datta 1985; Schweizer and Datta 1988; Goss et al. 1988). The extent of DNA carried by these and other plasmids employed in this study is shown in Fig. 1. pRS126, containing a frameshift mutation in *tdcB,* was constructed by digesting pRS124 with *Bgl*II, filling-in the recessed ends with Klenow DNA polymerase I, and subsequent ligation. Loss of the *Bgl*II site and lack of threonine dehydratase activity in strain TH16 (Δ*tdcABC*) were used to screen for the frameshift mutation. pRS127 harboring frameshift mutations in *tdcA* and *tdcB* was similarly constructed by digesting pRS126 with *Sal*I, and by filling-in and ligation as described above. The introduction of a frameshift at the unique *Sal*I site was confirmed by the loss of the *Sal*I site and generation of a *Pvu*I site. To construct pRS109 *(tdcP′ ABC),* the purified 3.4 kb *Hin*dIII fragment from pRS124 was ligated into *Hin*dIII-digested pTG107 (Goss and Datta 1985), which lacked the binding sites for IHF and CAP required for *tdc* expression. A 3.9 kb *Eco*RI-*Hin*dIII (partial) fragment from pRS109 was cloned into pUC18 to yield pYG111, which contained a truncated *tdcP* and the entire coding region of *tdcABC* downstream from *lacP,* which drives the expression of the *tdc* genes. pYG112 was obtained by inserting a 1.4 kb Kan^r *Eco*RI (blunt) fragment from pUC4K (Vieira and Messing 1982) at the vector *Aat*II site of pYG111. A 2.1 kb *Bal*I fragment deleted from pYG112 resulted in pYG113, which contained the *tdcA* coding sequence and the 5′ half of *tdcB.* Plasmid pSH234 was constructed (H.P. Schweizer and P. Datta, unpublished data) by isolating a 1.3 kb *Eco*RI-*Ssp*I fragment from pRS124 that contained the intact *tdcA* gene with its own promoter and the first 43 codons of *tdcB,* and then cloning it into *Eco*RI plus *Pvu*II-cleaved pACYC184. pYG234 and pYG184 were derived, respectively, from pSH234 and pACYC184 by introducing 1.4 kb Kan^r fragments from pUC4K. pYG18 was obtained by ligating the 1.4 kb Kan^r *Eco*RI fragment of pUC4K into pUC18.

pSH241 *(tdcR*$^+$ *tdcB-lacZ)* has been reported previously (Schweizer and Datta 1989a). pYW241 was obtained by ligating the 1.2 kb *Eco*RI-*Hin*dIII fragment of pSH241 into pUC19.

The plasmid pYG2413 harboring an in-frame deletion in *tdcA* was constructed as follows: a 0.34 kb *Hpa*I-*Sal*I fragment was excised from pYW241 and the *Sal*I end of the plasmid DNA was filled-in followed by self-ligation to yield pYG2412. A 0.84 kb *Eco*RI-*Hin*dIII fragment was purified from pYG2412 and ligated to the large (vector-containing) fragment of pSH241 previously digested with *Eco*RI and *Hin*dIII and treated with phosphatase. The presence of an in-frame deletion in *tdcA* in pYG2413 was confirmed by sequencing the fusion joint (data not shown) by the dideoxy DNA sequencing technique (Sanger et al. 1977).

Gel mobility shift assay. Gel mobility shift assays were carried out essentially as described by Hendrickson and Schlief (1984), with some minor modifications. For the binding reactions, two different DNA templates were used: the 354 bp *Dde*I-*Hpa*I fragment from pSH241 extending from codon 8 of *tdcR* to codon 32 of *tdcA* and harboring the wild-type *tdcP,* and an identical fragment from pYW75, which contained a mutation in the IHF binding site in *tdcP* (Wu and Datta 1992), to prevent the DNA mobility shift due to IHF binding. The fragments were end-labeled by filling-in with Klenow DNA polymerase I and [α-^{35}S]dATP. The labeled fragments were separately incubated with cell extracts of various *tdcA*$^+$ and *tdcA* strains in several different binding buffers including those used for IHF and CAP binding (Wu and Datta 1992; Wu et al. 1992). After incubation, the samples were electrophoresed through a 4% native polyacrylamide gel for 1–2 h at 350 V in several different

running buffers (Sambrook et al. 1989) including TBE (89 mM TRIS, 89 mM boric acid and 1 mM EDTA, pH 8.3). The gels were then dried under vacuum and exposed to Kodak X-AR5 film to visualize the bands by autoradiography.

Enzyme assay. Threonine dehydratase activity of toluene-treated cells was determined colorimetrically as previously described (Hobert and Datta 1983). Specific activity is expressed as nanomoles of 2-oxobutyrate formed per minute per milligram of protein. β-Galactosidase activity in sodium dodecyl sulfate-chloroform-permeabilized cells was measured according to Miller (1972), and specific activity is expressed in Miller units. All assays were repeated at least four times. Although the absolute values varied somewhat between experiments, the relative values (as shown in Table 3 in parentheses) were within 10% of each other.

Other methods. The general procedures for DNA manipulations, gel electrophoresis and end-labeling were as described by Sambrook et al. (1989). Restriction digestions of DNA were carried out following manufacturers' specifications. DNA sequencing was performed by the dideoxy procedure of Sanger et al. (1977) using the Sequenase Kit supplied by United States Biochemicals. Polymerase chain reaction (PCR; Saiki et al. 1988) was carried out using Taq polymerase according to the protocol supplied by Perkin-Elmer Cetus. The oligonucleotides used as primers for PCR were synthesized by the University of Michigan Biomedical Research Core Facility.

Results and discussion

Lack of dehydratase induction in strain K37

In a search for mutants lacking threonine dehydratase activity, Merberg and Datta (1982) found a number of strains of *E. coli,* obtained from several laboratories, that exhibited a wide variability in dehydratase activity when incubated anaerobically in TYE. A few strains with low enzyme levels produced 50- to 100-fold more enzyme when the TYE medium was supplemented with an electron acceptor such as fumarate or nitrate. A preliminary genetic analysis revealed that the dehydratase-negative phenotype was linked to several loci on the *E. coli* chromosome, including *pgi,* suggesting that mutations in multiple genes may influence the expression of the enzyme in vivo (Merberg and Datta 1982). Recently, we examined the strain K37, which produced very little or no enzyme activity when incubated anaerobically in TYE or in H4 medium containing 50 mM fumarate (Table 2). The wild-type strain MC4100 induced a high level of dehydratase under the same incubation conditions. That K37 cells might produce an inhibitory substance that somehow prevented enzyme induction or inactivated the enzyme was ruled out by assaying dehydratase activity in K37 transformed with the plasmid pEC61, which contains the cloned *tdc* operon (Goss and Datta 1985; Fig. 1). A large amount of dehydratase activity was found in K37 (pEC61) as compared with K37 harboring the vector or a control plasmid pSH215 *($tdcR^+$ Δ$tdcABC$)* (Table 2). We tentatively concluded from these data that the K37 cytoplasm appears to have no significant deleterious effect on the induction of plasmid-encoded TdcB or on the stability of the enzyme. Furthermore, because the DNA sequences flanking the *tdc* operon in pSH215 containing other open reading frames, including *tdcR,* (Goss et al. 1988; Schweizer and Datta 1989a) were not able to complement the function missing in K37, it appeared that lack of dehydratase synthesis in K37 might result from a mutation in its *tdc* DNA and/or at some other locus on the K37 chromosome. To test the validity of this notion we compared the dehydratase activities of K37 transformed with pRS126 *($tdcA^+B^-C^+$)* and K37 harboring pRS127 *($tdcA^-B^-C^+$)*. The results given in Table 2 show that K37 (pRS126) had high dehydratase activity, whereas K37 (pRS127) exhibited almost no enzyme activity. Both these plasmids lack functional *tdcB*, as seen from the control experiment with TH16 (Δ *tdcABC*) (pRS126); thus the enzyme activity found in K37 (pRS126) must have arisen from an intact K37 chromosomal *tdcB* gene. Furthermore, because pRS126 and not pRS127 contained $tdcA^+$, it is likely that expression of K37 *tdcB* was dependent on *tdcA* present in the multicopy plasmid. We concluded from this result that the K37 *tdcA* gene might harbor a mutation that prevented *tdcB* expression in this strain. These experiments

Table 2. Threonine dehydratase level in K37

Strain (plasmid)	Relevant genotype	TdcB specific activity in	
		TYE	H4
MC4100		927	1091
K37		71	0
K37 (pBR322)	(Vector)	71	0
K37 (pEC61)	($tdcA^+B^+C^+$)	1052	1058
K37 (pSH215)	(Δ *tdcABC*)	72	2
K37 (pUC19)	(Vector)	ND	5
K37 (pRS126)	($tdcA^+B^-C^+$)	ND	1224
K37 (pRS127)	($tdcA^-B^-C^+$)	ND	83
TH16 (pRS126)	Δ$tdcABC$ ($tdcA^+B^-C^+$)	ND	3

For culture conditions, enzyme assay and specific activity, see Materials and methods. ND, not determined

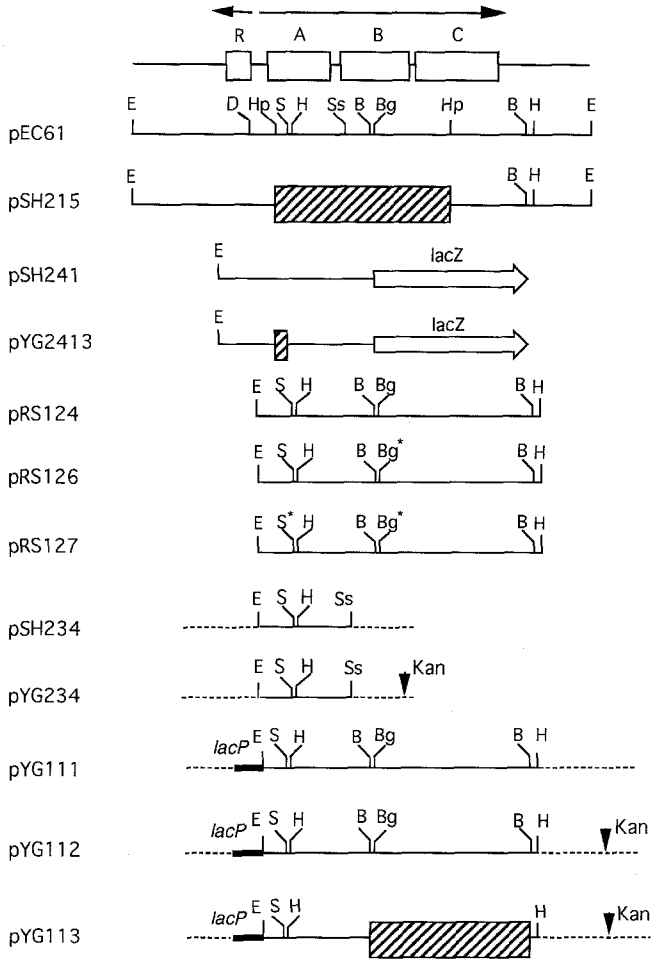

Fig. 1. Organization of the *tdc* operon and restriction maps of various plasmids. The *solid lines* represent *tdc* DNA, and the *broken lines* designate vector sequences. The *solid bars* in pYG111, pYG112 and pYG113 indicate the *lac* promoter sequence. The *hatched segments* in pSH215, pYG2413 and pYG113 indicate deleted regions, and the *open arrows* in pSH241 and pYG2413 indicate *lacZ* fusion. The *solid arrowheads* in pYG234, pYG112 and pYG113 indicate insertion of Kan^r cassette. The 5′ borders of the various plasmids are: pSH241 and pYG2413, nucleotide (nt) 1196 (see Schweizer and Datta 1989b), and include the *tdcR* reading frame; pRS124, pRS126, pRS127, pSH234 and pYG234, nt 1743, and harbor the *tdcP* region; pYG111, pYG112 and pYG113, nt 1848, and lack the integration host factor (IHF) and catabolite gene activator protein (CAP) binding sites in *tdcP* required for *tdc* expression. Abbreviations for restriction enzyme sites are: D, *Dde*I; E, *Eco*RI; Hp, *Hpa*I; S, *Sal*I; H, *Hin*dIII; Ss, *Ssp*I; B, *Bal*I; Bg, *Bgl*II. S* and Bg* denote, respectively, the location of frameshift mutations at the *Sal*I and *Bgl*II sites

provided for the first time an indication of a possible regulatory role of *tdcA* in *tdc* operon expression. A potential problem in this interpretation however is the mild polar effect of the frameshift mutation introduced in *tdcA* in reducing *tdcBC* expression. Indeed, we showed earlier that a frameshift mutation in *tdcA* partially reduced the TdcC-mediated threonine transport by about 50% (Sumantran et al. 1990), raising the concern that *tdcC* may also have some effect on *tdc* gene expression. The experiments described in a section below completely rule out this possibility.

Sequence of K37 tdc *DNA*

Does K37 *tdcA* harbor a mutation that influcences *tdc* operon expression in this strain? To answer this question directly we cloned the K37 *tdcA* gene and the *tdc* promoter region, employing the PCR technique using two sets of primers flanking the *tdcA* gene (Schweizer and Datta 1989b): (A), 5′ TGGTCTCAGCCCCTTTTTT-GTATTAACCAC 3′ corresponding to nucleotides (nt) 1641 to nt 1670 (forward primer) and 5′ TAGCAAC-CGGCAGATCGTATGTAATATGCA 3′ corresponding to nt 2970 to nt 2941 (reverse primer); and (B), 5′ CGCCTATGAGGGGAGAAGAA 3′ corresponding to nt 1350 to nt 1369 (forward primer) and the same reverse primer corresponding to nt 2970 to nt 2941. The standard PCR reactions with these primers and K37 genomic DNA were performed according to the manufacturer's specification, and the products were purified by gel electrophoresis. After digestion with *Dra*I, the 1086 bp fragment (from nt 1770 to nt 2855) was purified and cloned into the M13 cloning vector for DNA sequencing. The complete nucleotide sequence from both strands revealed that the DNA sequence of K37 *tdcA* (data not shown) was identical to the published nucleotide sequence of *tdcA* from strain W3110 (Goss and Datta 1985; Schweizer and Datta 1989b). Sequencing of the K37 *tdcP* region also revealed a nucleotide sequence identical to that found in strain W3110. In a separate experiment, we replaced the *tdcR-tdcPAB'* region of plasmid pSH241 by a homologous fragment from the cloned K37 *tdc* DNA and found wild-type LacZ expression in strain TH274 (data not shown). Thus, the K37 chromosome contains an intact *tdcA* gene and the wild-type *tdcP*, in addition to a functional *tdcB* gene as mentioned above. These findings raised an intriguing question: what genetic alteration(s) in K37 resulted in loss of *tdc* gene expression? One likely explanation is that a mutation in an unlinked regulatory locus on the K37 chromosome somehow abolished *tdc* expression in vivo; however, the exact mechanism by which a plasmid-borne *tdcA* gene (from pRS126) could rescue such a mutation (cf. Table 2) is not clear. An alternative possibility, alluded to earlier, is that *tdcC* may also play a part in *tdc* gene expression. Obviously, further analysis is needed to identify the mutation in K37 and to examine its involvment in *tdc* gene expression in vivo.

Involvement of tdcA *in* tdc *expression*

Despite the uncertain nature of the mutation in K37 that prevented *tdc* expression in this strain and the potentially complicating effect of polarity in the *tdcA* frameshift mutant in complementation studies depicted in Table 2, the results presented thus far suggested that *tdcA* was somehow involved in regulating the *tdc* operon. To analyze rigorously the in vivo role of *tdcA*, we redesigned the experimental protocol as follows, employing the strain MC4100, which shows the normal pattern of expression of the *tdc* genes. First, an in-frame deletion in *tdcA* was introduced by deleting a fragment between

Table 3. Requirement of *tdcA* for *tdcB-lacZ* expression

Expt	Strain[a] (plasmid)	Relevant genotype	LacZ activity in		TdcB activity in
			TYE	H4	TYE
I	TH274λYG2413 (pYG18)	*tdcR*$^+$*tdcA*$^\Delta$*B-lacZ* (vector)	155 (1.0)	267 (1.0)	361
	TH274λYG2413 (pYG113)	*tdcR*$^+$*tdcA*$^\Delta$*B-lacZ* (*tdcR lacPtdcA*$^+$*B'*)	744 (4.8)	1266 (4.7)	210
	TH274λYG2413 (pYG112)	*tdcR*$^+$*tdcA*$^\Delta$*B-lacZ* (*tdcR lacPtdcA*$^+$*B*$^+$*C*$^+$)	754 (4.9)	1348 (5.0)	36850
II	TH274λSH241 (pYG18)	*tdcR*$^+$*tdcAB-lacZ* (vector)	756 (1.0)	1053 (1.0)	334
	TH274λSH241 (pYG113)	*tdcR*$^+$*tdcAB-lacZ* (*tdcR lacPtdcA*$^+$*B'*)	775 (1.0)	704 (0.7)	190
	TH274λSH241 (pYG112)	*tdcR*$^+$*tdcAB-lacZ* (*tdcR lacPtdcA*$^+$*B*$^+$*C*$^+$)	839 (1.1)	1097 (1.0)	32050
III	TH274λYG2413 (pYG184)	*tdcR*$^+$*tdcA*$^\Delta$*B-lacZ* (vector)	210 (1.0)	237 (1.0)	ND
	TH274λYG2413 (pYG234)	*tdcR*$^+$*tdcA*$^\Delta$*B-lacZ* (*tdcR tdcPA*$^+$)	980 (4.7)	947 (4.0)	ND
IV	TH274λSH241 (pYG184)	*tdcR*$^+$*tdcAB-lacZ* (vector)	817 (1.0)	1084 (1.0)	ND
	TH274λSH241 (pYG234)	*tdcR*$^+$*tdcAB-lacZ* (*tdcR tdcPA*$^+$)	503 (0.6)	745 (0.7)	ND

For culture conditions, enzyme assays and specific activities, see Materials and methods. Relative values of LacZ expression as a function of vector control are shown in parentheses. ND, not determined

[a] Strain TH274 is Δ*lac* Δ(*tdcR-tdcABC*)

codons 33 and 144 within the *tdcA* reading frame in the *tdcB-lacZ* fusion plasmid pSH241 yielding pYG2413 (*tdcR*$^+$ *tdcPA*$^\Delta$*B-lacZ*), and then the fusion was transferred in a single copy to the TH274 Δ(*tdcR-tdcABC*) chromosome, a derivative of MC4100. Second, the new plasmid pYG112 was constructed by placing the *tdcABC* genes with a truncated *tdcP* downstream from the *lac* promoter to drive *tdc* expression from *lacP* (see Fig. 1); plasmid pYG113 was derived from pYG112 by deleting *tdcBC*. Third, all assays were performed in strain TH274. These changes allowed measurement of *tdc* gene expression in a known genetic background from single copy gene constructs, without polar effect, in the presence or absence of *tdcA* and *tdcC* driven by a strongly expressed independent promoter.

As shown in Table 3, experiment I, β-galactosidase activity in the *tdcA* in-frame deletion strain (TH274λYG2413) was low in both TYE and H4 medium, whereas, the same strain transformed with pYG112 (*tdcA*$^+$*B*$^+$*C*$^+$) or pYG113 (*tdcA*$^+$*B'*) exhibited five fold more enzyme activity. In absolute values, the levels of β-galactosidase seen in TH274λYG2413 harboring pYG112 or pYG113 were approximately the same as that found in the control strain TH274λSH241 (*tdcA*$^+$), with or without pYG112 or pYG113 (experiment II). Almost identical results were seen when the plasmid pYG234, harboring the cloned *tdcA*$^+$ gene with its own promoter, was used to complement the *tdcA*-minus TH274λYG2413 (experiment III and IV). It should be noted in this context that all constructs employed in these experiments had one copy of *tdcR*$^+$. These results clearly show that the nonpolar mutation introduced in *tdcA* drastically reduced the downstream *tdcB-lacZ* expression, and that the plasmid-encoded *tdcA* alone, driven by *tdcP* or *lacP*, fully complemented the *tdcA* mutation in the TH274λ lysogen.

Several additional pieces of information emerged from the data given in Table 3. First, as expected, expression of threonine dehydratase activity from the *lac* promoter in pYG112 was high in both TH274λ lysogens incubated anaerobically in TYE (cf. experiments I and II) or when grown aerobically in the same medium (data not shown), indicating normal functioning of *lacP*. Second, the high level expression of the *tdc* gene product(s) from the multicopy plasmids pYG112 and pYG113 had no significant effect on β-galactosidase activity from the wild-type lysogen with the intact *tdcA*$^+$ gene (cf. experiment II). And third, the presence or absence of *tdcC* in the complementing plasmids did not alter the levels of LacZ expression in *tdcA*$^+$ or *tdcA* lambda lysogens (experiments I and II). The last finding is reminiscent of earlier reports that when *tdcC* is truncated no significant reduction in *tdc* expression is seen from plasmid-borne *tdcB* (Goss and Datta 1985) or *tdcB-lacZ* fusions (Sumantran et al. 1990), although deletion of *tdcC* did reduce TdcC-mediated threonine uptake to a basal level of transport, presumably occurring via other amino acid transport systems in *E. coli* (Sumantran et al. 1990). In a separate experiment we found that the levels of LacZ expression in MC4100λSH241 and MC4100λYG2413 were almost the same (567 and 544 units of β-galactosidase, respectively), indicating that a single chromosomal copy of *tdcA* from the host was sufficient fully to restore *tdc* expression in the *tdcA*-minus lysogen. The cumulative data presented here provide genetic evidence that *tdcA*, and not *tdcC*, is involved in positive regulation of the *tdcABC* operon in vivo.

TdcA is a member of the LysR family of proteins

The *lysR* gene of *E. coli* encodes a protein of 311 amino acid residues, which activates transcription of *lysA*, the structural gene for diaminopimelate decarboxylase, involved in lysine biosynthesis (Stragier and Patte 1983). By searching DNA and protein databases, Henikoff et al. (1988) found extensive sequence similarities between LysR and several other bacterial proteins. Several members of this LysR family of proteins share two important structural and functional features: they are of similar size and have helix-turn-helix motifs typical of DNA binding proteins, and they are transcriptional activators of their adjacent operons. We previously reported that the TdcA polypeptide consists of 312 amino acid residues with a

```
LysR   MAAVNL---RHIEIFHAVMTAGSLTEAAHLLHTSQPTVSRELARFEKVIG   47
TdcA   MSTILLPKTQHLVVFQEVIRSGSIGSAAKELGLTQPAVSKIINDIEDYFG   50
CONS   MSMSMM-DLNHLKIFEAVMEEGSLTAAARALHLSQPAISRQIARLEQHLG   49

LysR   LKLFERVRGRLHPTVQGLRLFEEVQRSWYGLDRIVSAAESLREFRQGELS   97
TdcA   VELVVRKNTGVTLTPAGQLLLSRSESITREMKNMVNEISGMSSEAVVEVS  100
CONS   DQLFVRXGRGLRLTPAGEELLRXARQALXLIQRMLDAXDXXXPSESGRLF   99

LysR   IACLPVFSQSFLPQLLQPFLARYPDVSLNIVPQESPLLEEWLSAQRHDLG  147
TdcA   FGFPSLIGFTFMSGMINKFKEVFPKAQVSMYEAQLSSFLPAIRDGRLDFA  150
CONS   IACIGTFAXSVLPSLLENFRARYPHVSLXLTTHENXDPEEALRAGELDLA  149

LysR   LTETLHTPAGTERTELLSL-DEVCVLPPGHPLAVKKVLTPDDFQGENYIS  196
TdcA   I-GTLSAEMKLQDLHVEPLFESEFVLVASKSRTCTGTTTLESLKNEQWVL  199
CONS   I-S--XDPLHSPGTESXXLFEDXLVXVA---------------------L  175

LysR   LSRTDSYRQLLDQLFTEHQVKRRMIVETHSAASVCAMVRAGVGISVVNPL  246
TdcA   PQTNMGYYSELLTTLQRNGISIENIVKTDSVVTIYNLVLNADFLTVIPCD  249
CONS   PPDHPLAGKKXITXEEXXXHTLVSYXRTXSRRXLWWXLF-XXLQVXSRIX  224

LysR   TALDYAASGLVVRRFSIAVPFTVSLIRPLHRPSSALVQAFSGHLQAGLPK  296
TdcA   MTSPFGSNQFIT--IPVEETLPVAQYAAVWSKNYRIKKAASVLVE--LAK  295
CONS   VEATVAGSVXMV--VMLAAGVGIAALPLVLVXAXSXXVRVVXXXL--XQX  270

LysR   LVTSLDAILSSATTA    311
TdcA   EYSSYNGCRRRQLIEVG  312
CONS   XLXSIXXLRRPALAXR   286
```

Fig. 2. Alignment of TdcA sequence with LysR and with a consensus sequence of the LysR family of activator proteins. The *vertical bars* show identical amino acid residues and *colons* indicate conservative substitutions

calculated molecular weight of 35 544, and in minicells produced a polypeptide of expected length (Goss and Datta 1985; Goss et al. 1988). Figure 2 shows the sequence alignment of TdcA with LysR, as well as with a consensus sequence deduced from multiple alignment of nine LysR family members (Henikoff et al. 1988). The amino acid sequences of TdcA and LysR revealed 17% identity and an additional 23% similarity throughout their entire lengths although the proteins are more similar near their amino-terminal ends. A comparison of TdcA with the LysR family consensus sequence also shows a high degree of similarity. Thus, TdcA appears to be related to the LysR family of proteins by common ancestry, and based on its functional role in *tdc* expression, belongs to the LysR family of transcriptional activators.

Helix-turn-helix motif in TdcA

How does a *trans*-acting protein activate gene transcription? If TdcA were to function as a transcriptional regulatory factor for the *tdc* operon, it is likely to interact with the *tdc* DNA to modulate mRNA synthesis. Many prokaryotic transcription regulatory proteins appear to bind to their respective target DNAs by a helix-turn-helix motif (Pabo and Sauer 1984). The members of the LysR family of gene activator proteins also have helix-turn-helix motifs near their N-termini (Henikoff et al. 1988). Analysis of the TdcA sequence according to the method of Dodd and Egan (1987) revealed that the amino acid sequence between residues 24 and 43 shows a significant homology to the helix-turn-helix motifs of several DNA binding proteins (Fig. 3; Pabo and Sauer 1984). For example, TdcA has highly conserved residues at positions 28, 32, 38 and 39, hydrophobic residues at positions 31 and 33, and a basic residue at position 40. Thus, it appears reasonable that TdcA could function as a putative DNA binding protein to enhance *tdc* transcription.

To test the proposition that TdcA might bind to the *tdcP* DNA, a series of gel-shift assays was carried out using two DNA templates (with the wild-type *tdc* promoter sequence and a mutant promoter harboring a mutation in the IHF binding site) incubated at 4° or 22° C with extracts of several strains including TH274 [Δ*(tdcR-tdc)*] and the same strain transformed with

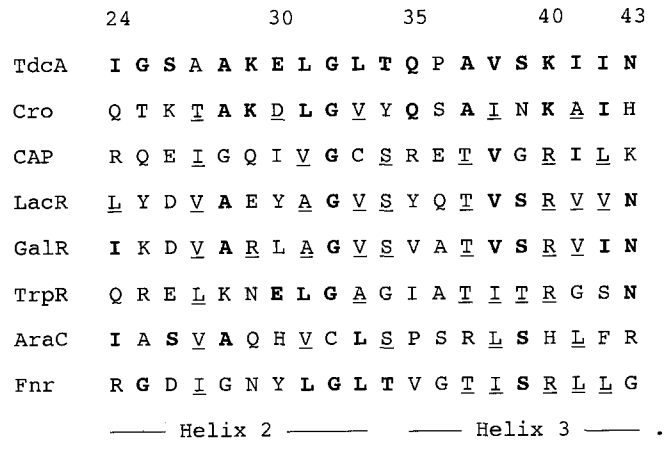

Fig. 3. Helix-turn-helix motif in TdcA polypeptide as compared with that of several known DNA binding proteins (taken from Pabo and Sauer 1984; Dodd and Egan 1987). The *numbers* on the top represent amino acid residues of TdcA polypeptide. *Bold letters* designate identical residues and *underlined letters* indicate homologous residues

pYG18 (vector), pYG112 *($tdcA^+B^+C^+$)* or pYG113 *($tdcA^+$)* in a number of different binding buffers with various concentrations of cations followed by electrophoresis at 4° C or at 22° C in TBE, TAE or sodium phosphate buffers as described in Materials and methods. No TdcA-specific shift in mobility of *tdcP* DNA was observed in a large number of trials. These results might tend to suggest that the TdcA-mediated regulation of the *tdc* operon does not appear to depend on its DNA binding activity to *tdcP*, however, we cannot completely rule out the possibility that some missing factor(s) might be involved in promoting *tdcP*-TdcA interaction. It may be recalled in this context that, despite numerous attempts, no DNA binding activity of the fumarate-nitrate reductase activator protein, Fnr, has yet been seen in gel-mobility shift assays using cell extracts or purified protein, although specific DNA binding to the consensus Fnr binding site has been recently reported in footprinting studies (Spiro and Guest 1990; Green et al. 1991). It is also possible to envisage that TdcA might interact with other gene(s) whose product(s) is needed for *tdcABC* expression.

Conclusion

By serendipity, the initial experiments in this project revealed that the *E. coli* strain K37 contained a mutation that prevented *tdc* expression in vivo and that a plasmid-borne *$tdcA^+$* gene was able to rescue this phenotype. Although the identity of the mutation in this strain still remains elusive, the follow up experiments with strain MC4100 summarized above led us to define for the first time the functional role of *tdcA*, the promoter-proximal gene of the *tdcABC* operon. Mutant analysis and complementation studies clearly indicated that the TdcA polypeptide acts as a *trans*-acting positive regulatory factor for *tdc* operon expression. Based on structural and functional similarities, TdcA appears to belong to the LysR family of bacterial activator proteins. We have recently reported that three other proteins, TdcR (Schweizer and Datta 1989a), IHF (Wu and Datta 1992), and CAP in combination with cAMP (Wu et al. 1992), are also required for positive transcriptional regulation of the *tdc* genes. Footprinting experiments revealed that purified IHF and CAP occupy their unique binding sites at around −100 and −40, respectively, on the *tdc* promoter DNA and together activate *tdc* transcription (Wu and Datta 1992; Wu et al. 1992). A preliminary deletion analysis indicated that a *cis*-acting sequence in the *tdcP* region is required for *tdcR* action (Schweizer and Datta 1989a). These regulatory proteins appear to facilitate *tdc* transcription by interacting with the *tdc* promoter. Further experiments are necessary to elucidate the biochemical mechanism of *tdc* operon expression by TdcA and its mutual interactions, if any, with the other *trans*-acting regulatory factors in transcription activation.

Acknowledgements. This work was supported by Public Health Service grant GM21436 from the National Institutes of Health. R.K. Jambukeswaran was supported in part by a Thurnau Postdoctoral Fellowship from the Center for Molecular Genetics at the University of Michigan. Computer facilities, supported by NIH grant M01 RR00042, were provided by the General Clinical Research Center, University of Michigan Medical School. We thank David Friedman for strain K37, Herbert Schweizer for strain TH274 and plasmid pSH234, and Byron Hagewood and Yifei Wu for stimulating discussions and their critical reading of the manuscript.

References

Datta P, Goss TJ, Omnaas JR, Patil RV (1987) Covalent structure of biodegradative threonine dehydratase of *Escherichia coli:* homology with other dehydratases. Proc Natl Acad Sci USA 84:393–397

Dodd IB, Egan JB (1987) Systematic method for the detection of potential λ cro-like DNA-binding regions in proteins. J Mol Biol 194:557–564

Friedman DI, Plantefaber LC, Olson EJ, Carver D, O'Dea MH, Gellert M (1984) Mutations in the DNA *gyrB* gene that are temperature sensitive for lambda site-specific recombination, Mu growth, and plasmid maintenance. J Bacteriol 157:490–497

Goss TJ, Datta P (1985) Molecular cloning and expression of the biodegradative threonine dehydratase gene *(tdc)* of *Escherichia coli.* Mol Gen Genet 201:308–314

Goss TJ, Schweizer HP, Datta P (1988) Molecular characterization of the *tdc* operon of *Escherichia coli* K-12. J Bacteriol 170:5352–5359

Green J, Tragesser M, Six S, Unden G, Guest JR (1991) Characterization of the FNR protein of *Escherichia coli,* an iron-binding transcriptional regulator. Proc R Soc Lond [Biol] 244:137–144

Hendrickson W, Schlief RF (1984) Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. J Mol Biol 174:611–628

Henikoff S, Haughn GW, Calvo JM, Wallace JC (1988) A large family of bacterial activator proteins. Proc Natl Acad Sci USA 85:6602–6608

Hobert EH, Datta P (1983) Synthesis of biodegradative threonine dehydratase of *Escherichia coli:* role of amino acids, electron acceptors, and certain intermediary metabolites. J Bacteriol 155:586–592

Merberg D, Datta P (1982) Altered expression of biodegradative threonine dehydratase in *Escherichia coli* mutants. J Bacteriol 150:52–59

Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Pabo CO, Sauer RT (1984) Protein DNA recognition. Annu Rev Biochem 53:293–321

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer directed enzymatic amplification of DNA with thermostable DNA polymerase. Science 239:487–491

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Schweizer HP, Datta, P (1988) Genetic analysis of the *tdcABC* operon of *Escherichia coli* K-12. J Bacteriol 170:5360–5363

Schweizer HP, Datta P (1989a) Identification and DNA sequence of *tdcR*, a positive regulatory gene of the *tdc* operon of *Escherichia coli*. Mol Gen Genet 218:516–522

Schweizer HP, Datta P (1989b) The complete nucleotide sequence of the *tdc* region of *Escherichia coli*. Nucleic Acids Res 17:3994

Schweizer HP, Datta P (1991) Physical linkage and transcriptional orientation of the *tdc* operon on the *Escherichia coli* chromosome. Mol Gen Genet 228:125–128

Spiro S, Guest JR (1990) FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol Rev 75:399–428

Stragier P, Patte J-C (1983) Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli* III Nucleotide sequence and regulation of the *LysR* gene. J Mol Biol 168:333–350

Sumantran VN, Schweizer HP, Datta P (1990) A novel membrane-associated threonine permease encoded by the *tdcC* gene of *Escherichia coli*. J Bacteriol 172:4288–4294

Umbarger HE (1978) Amino acid biosynthesis and its regulation. Annu Rev Biochem 47:533–606

Vieira J, Messing J (1982) The PUC plasmids, an M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259–268

Wu Y, Datta P (1992) Integration host factor is required for positive regulation of the *tdc* operon of *Escherichia coli*. J Bacteriol 174:233–240

Wu, Y, Patil RV, Datta P (1992) Catabolite gene activator protein and integration host factor act in concert to regulate *tdc* operon expression in *Escherichia coli*. J Bacteriol 174:6918–6927