Mol Gen Genet (1985) 201:308–314

MGG

© Springer-Verlag 1985

Molecular cloning and expression of the biodegradative threonine dehydratase gene (*tdc*) of *Escherichia coli* K12

Thomas J. Goss and Prasanta Datta

Department of Biological Chemistry, University of Michigan, Box 034, Ann Arbor, MI 48109-0010, USA

Summary. The biodegradative threonine dehydratase gene (*tdc*) of *Escherichia coli* was cloned by isolating a dehydratase-negative mutant after Tn5 mutagenesis, cloning the *tdc*::Tn5 DNA into pBR322 and then replacing the Tn5 element on the plasmid in vivo. Subcloning and nucleotide sequence data revealed two distinct procaryotic promoter-like elements each containing a potential CAP-binding site and AT-rich regions, and a Shine-Dalgarno sequence. One of these putative promoters, P_2, was located immediately upstream from the *tdc* coding region, and a second, P_1, was approximately 1 kilobase upstream from P_2. Deletion of the potential CAP-binding site from P_1 prevented *tdc* gene expression. However, removal of P_2 and a large segment of the upstream DNA had no discernible effect on dehydratase synthesis. A 936-base pair open reading frame was found between P_1 and the *tdc* coding region, which produced a polypeptide of about 32 kilodaltons. The data suggest that P_1, and not P_2, is necessary for *tdc* gene expression, and that the DNA sequences coding for the 32 KD polypeptide and threonine dehydratase are part of a single transcriptional unit.

Introduction

Although biodegradative threonine dehydratase (EC 4.2.1.16) of *Escherichia coli* is known to be induced under anaerobic conditions in tryptone-yeast extract (TYE) medium in the absence of fermentable carbohydrates (Umbarger 1978), two recent reports (Merberg and Datta 1982; Hobert and Datta 1983) indicated that the complex TYE medium contained various metabolites which suppressed enzyme synthesis, and mutations at different loci on the *E. coli* chromosome that altered the intracellular concentrations of some of these metabolites, influenced enzyme induction. Furthermore, a combination of only four amino acids, threonine, serine, valine and isoleucine, supplemented with cAMP and an electron acceptor (fumarate or nitrate), produced much higher levels of the dehydratase as compared to that in the TYE medium containing cAMP + fumarate. These observations are consistent with the notion that a complex regulatory system of intermediary metabolites controls enzyme synthesis in vivo. To gain an insight into the regulatory mechanism, we initiated a genetic analysis of this inducible system by isolating an *E. coli* mutant lacking functional threonine dehydratase after Tn5 transposon mutagenesis and cloning the DNA fragment with the integrated Tn5 element into the plasmid pBR322 (Goss and Datta 1984). Mapping experiments with the transposon element that caused insertional inactivation located the threonine dehydratase (*tdc*) gene at minute 67 on the *E. coli* chromosome. In this report we describe the isolation of the functional *tdc*$^+$ gene by replacement of the *tdc*::Tn5 DNA on the plasmid and provide evidence for the existence of a distal regulatory region on the cloned DNA, 1 kb upstream from the dehydratase coding region, which is essential for in vivo expression of the enzyme.

Offprint requests to: P. Datta

Materials and methods

Materials. Tryptone, yeast extract and other medium components were supplied by Difco Laboratories. Restriction enzymes, Bal31 nuclease, S1 nuclease, Klenow fragment of DNA polymerase, T4 DNA polymerase and T4 ligase were bought from either New England Biolabs, Inc. or Bethesda Research Laboratories. [^{35}S]-Methionine and reagents for DNA sequencing by the chain termination method of Sanger et al. (1977) were supplied by Amersham. The following items were bought from Sigma Chemical Company: L-amino acids, cAMP, various antibiotics, isopropyl-β-D-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-galactoside (X-gal) and ethidium bromide. Antibodies to the purified threonine dehydratase (Merberg and Datta 1982) were supplied by David Merberg. All other chemicals were of reagent grade.

Bacteria, plasmids and phages. E. coli strain JA194 (F$^-$ *leu thi gal-1,2 lac xyl ara ΔtrpE5 sup*$^{2+}$ *rec*$^+$ *tdc*$^+$ *hsdR*$^-$ *hsdM*$^+$) was obtained from Charles Yanofsky, Stanford University. Strain X 1411 was a gift from Dale Oxender of The University of Michigan. The pUC plasmids, their host strain JM103 (*Δ(lac−pro) thi* P1 lysogen *strA supE endA sbcB* F′ *traD36 proAB*$^+$ *lacI*^{Q} *lacZΔM15*) and M13 phages were supplied by Amersham. TG103 was a *tdc*::Tn5 transductant of JM103 from TG11. The sources of other bacterial strains and plasmids have been described (Goss and Datta 1984).

Plasmid construction and DNA preparation. The construction of plasmid pEC6, containing Tn5 insert in the *tdc* gene, has been described (Goss and Datta 1984). In general, the techniques outlined by Maniatis et al. (1982) were employed

for DNA isolation, purification of restriction fragments by electrophoresis through low melting agarose or polyacrylamide gels, cloning into plasmids and transformation of competent host strains. Restriction digests of DNA with appropriate enzymes were carried out according to manufacturers' specifications.

Deletion analysis. A nested set of deletions was generated by Bal31 nuclease digestion following the procedure described in the product profile of Bethesda Research Laboratories. The system was calibrated by incubating at 25° 1 µg of λ *Hin*dIII DNA fragments with 0.6 unit of Bal31 nuclease in a total volume of 24 µl of 20 mM Tris.Cl buffer, pH 8.1, containing 12 mM $MgCl_2$, 12 mM $CaCl_2$, 200 mM NaCl and 1 mM EDTA for various lengths of time; the course of digestion was analyzed by agarose gel electrophoresis. Purified DNA from pTG4B (see Fig. 3) was linearized with *Eco*RI, treated with Bal31 nuclease and fragments of desired lengths were pooled and purified by phenol extraction and ethanol precipitation. Typically, about 200 nucleotides were removed per minute per fragment end using the above conditions. The fragments were then end-repaired with T4 DNA polymerase in 33 mM Tris.acetate buffer, pH 7.6, containing 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT and 0.1 µg/µl of bovine serum albumin in two steps: first in the absence then in the presence of deoxynucleoside triphosphates (0.2 mM each). For subsequent cloning, the pooled DNA fragments were digested with *Pst*I, purified from vector DNA by agarose gel electrophoresis, and ligated separately to pUC13 previously digested with *Sma*I and *Pst*I. The ligation mixture was used to transform TG103 and Amp^r transformants with Lac^- phenotype were screened on X-gal/IPTG plates.

DNA sequencing. The dideoxy chain terminating method of Sanger et al. (1977), as outlined in the Amersham Handbook (1983), was used for DNA sequencing. Nucleotide sequence was obtained from both strands of each restriction fragment, cloned into M13 mp8 and mp9 in opposite orientations, and from overlapping regions of adjacent fragments.

Enzyme assay. Threonine dehydratase activity of individual cultures, grown anaerobically in still cultures for either 12 or 20 h at 37° C in tryptone yeast extract medium, was assayed either by the colorimetric procedure with toluene-treated cells or spectrophotometrically in sonic extracts (Merberg and Datta 1982).

Analysis of plasmid-encoded proteins in minicells. Minicells from *E. coli* strain X 1411 harboring various plasmids were purified and analyzed basically as described by Clark-Curtiss and Curtiss (1983) with a few minor changes. The growth and induction medium used was medium A containing lactose and all 19 amino acids except asparagine (Miller 1972) supplemented with 75 µg of ampicillin per ml. Strain X 1411 was grown at 37° C aerobically or anaerobically (with 10 mM KNO_3) and minicells were prepared by two rounds of differential rate sedimentation through 5%–30% sucrose gradient. Aliquots of minicells were labeled with 10 µCi of [^{35}S]-methionine and the labeled proteins were analyzed by SDS-polyacrylamide gel electrophoresis. For detection of [^{35}S]-labeled threonine dehydratase, the minicells were fractionated according to Braus et al. (1984), and the cytosolic fraction was incubated with anti-dehydratase antibodies, precipitated with Staph A cells and electrophoresed as above. The gels were dried and subjected to autoradiography.

Results

Cloning of the tdc gene

The plasmid pEC6 was obtained by cloning an 11.7 kb *tdc*::Tn*5* DNA fragment into the *Eco*RI site of pBR322 and screening for Kan^r transformants (Goss and Datta 1984). Figure 1 depicts the experimental strategy for isola-

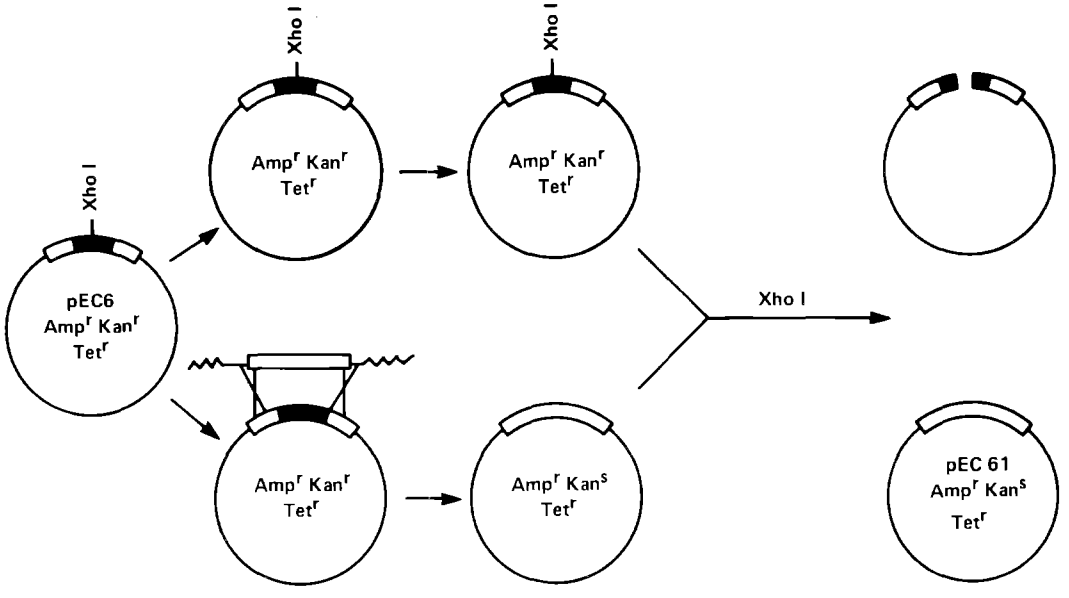

Fig. 1. Strategy for the isolation of kanamycin-sensitive transformants. *E. coli* strain JA194 harboring plasmid pEC6 (*amp tet kan*), grown in LB medium supplmented with ampicillin, yielded two classes of plasmid DNA molecules: a large number of parental molecules containing *amp tet* and *kan* genes, and relatively few with *amp* and *tet* genes. Restriction of this plasmid DNA pool with *Xho*I prior to retransformation selected against those molecules containing the *kan* gene thereby enriching for Amp^r, Tet^r, Kan^s transformants. For experimental detail, see text

Table 1. Threonine dehydratase activity in Kan^s transformants. Individual colonies were grown anaerobically in 5 ml of tryptone-yeast extract medium for 12 h at 37° C, and enzyme activity of toluene-treated cells was measured as detailed elsewhere (Merberg and Datta 1982). The specific activity is expressed as micromoles of α-ketobutyrate produced per minute per milligram of protein (Hobert and Datta 1983)

Clone designation	Phenotype	Specific activity
pEC61-C1	Kan^s	13.8
pEC61-C3	Kan^s	13.4
pEC61-C8	Kan^s	10.8
pEC61-C10	Kan^s	0
pEC61-C16	Kan^s	1.2
pEC61-C21	Kan^s	0
pEC61-C50	Kan^r	0
pEC61-C51	Kan^r	1.0
JA194	Kan^s	1.0–2.0

tion of kanamycin sensitive colonies of *E. coli* strain JA194 transformed with pEC6 DNA (*amp kan tet*). Plasmid DNA isolated from cells grown in LB medium with ampicillin contained two classes of molecules: a large number of parental molecules having the genes conferring resistance to ampicillin, kanamycin and tetracycline, and relatively few with genes for ampicillin and tetracycline; the latter class was generated either by rare recombinational events between homologous *E. coli* DNA inserts flanking the integrated Tn5 element in the plasmid and chromosomal DNA within and adjacent to the *tdc*$^+$ gene (as shown in Fig. 1), or transposition of Tn5 element from the plasmid. To select against the predominant class of DNA molecules with *amp kan tet* genes prior to retransformation of JA194, the plasmid DNA was digested with *Xho*I and treated with phosphatase; previous experiments indicated that neither the pBR322 vector DNA nor the *E. coli* DNA insert had recognition site for *Xho*I, whereas the Tn5 DNA was cleaved by the enzyme at 3 sites. From 24 μg of *Xho*I-restricted pEC6 DNA, a total of 500 Amp^r transformants were found. Upon replica plating on LB agar containing 50 μg/ml of kanamycin, 390 (or 78%) of these transformants were kanamycin sensitive. Thirty-two of the Kan^s colonies, picked at random and purified, were screened for threonine dehydratase activity. Some representative data are summarized in Table 1. Twenty-two out of 32 Kan^s clones (or 66%) had 5–15 times more enzyme activity compared to JA194 and appeared to harbor the cloned *tdc*$^+$ gene in high copy number. The remaining Kan^s clones had enzyme activity less than or equal to that found in JA194. It is likely that these Kan^s clones exhibiting little or no enzyme activity lost the *kan* gene by imprecise excision of Tn5, thus abolishing both *tdc*$^+$ and *kan* functions.

The cloning of the wild type *tdc* gene into plasmid pEC61 was supported by the DNA restriction profiles shown in Fig. 2. pEC6 DNA contained an 11.7-kb insert into pBR322 comprising of 5.5 kb Tn5 element integrated into a 6.2-kb *E. coli* DNA fragment that caused insertional inactivation of the *tdc* gene (Goss and Datta 1984). As predicted from the restriction map, digestion of pEC6 and pEC61 (clones 1, 3, 5, and 12, all Kan^s, dehydratase overproducers, cf. Table 1) DNAs with *Eco*RI generated two fragments in each case, 11.7 plus 4.3 kb (lane 10, top panel),

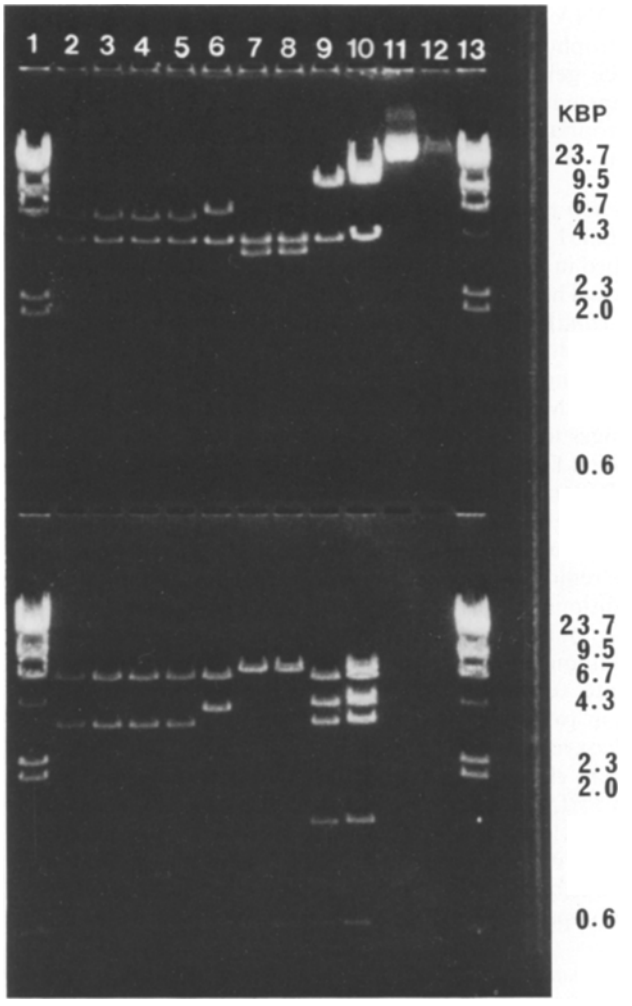

Fig. 2. Restriction profiles of plasmid DNAs isolated from various classes of Kan^s and Kan^r transformants treated with *Eco*RI (*top panel*) or *Hin*dIII (*bottom panel*). Lanes 1 and 13, *Hin*dIII-digested λ DNA; lanes 2–5, pEC61 clones 1, 3, 5 and 12, respectively (all Kan^s, dehydratase overproducers); lanes 6–8, pEC61 clones 16, 10, and 21, respectively (all Kan^s with no or basal level enzyme activity); lane 9, pEC61 clone 50 (Kan^r, no enzyme activity); lane 10, pEC6 DNA (Kan^r, no enzyme activity). Lanes 11 and 12, top panel, control DNAs from pEC6 and pEC61 clone 5, respectively, not digested with *Eco*RI. No DNAs were present in lanes 11 and 12 in bottom panel

and 6.2 plus 4.3 kb (lanes 2–5, top panel), respectively. Similarly, digestion with *Hin*dIII produced 5 expected fragments from pEC6 (approximately 6.5, 4.3, 3.3, 1.3 and 0.6 kb, lane 10, bottom panel) and 3 fragments from pEC61 (about 6.0, 3.5 and 0.6 kb, lanes 2–5, bottom panel). Control experiments with pEC61 clones (all Kan^s, having basal or no enzyme activity) showed altered restriction profiles due to extended deletions in clones 10 and 21 (lanes 7 and 8), and a somewhat smaller deletion in clone 16 (lane 6).

Evidence for an upstream sequence needed for tdc expression

The hybridization pattern of various restriction fragments of pEC6 DNA with a synthetic cDNA complementary to the DNA segment coding for a unique hexapeptide at the amino terminus end of the enzyme (amino acid residues 12–17) showed that the potential translation-start site for

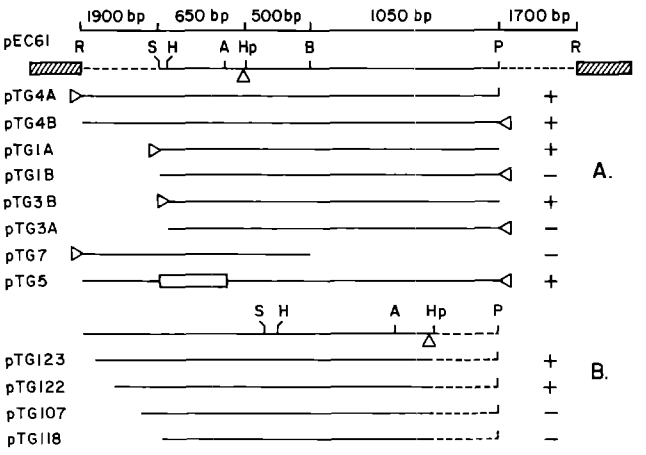

Fig. 3A, B. Subclones of pUC plasmids containing various DNA fragments and expression of threonine dehydratase activity. **A** pEC61 DNA was digested with appropriate restriction enzymes and the purified fragments were cloned into pUC8 and pUC13 in two different orientations with respect to the *lac* promoter as designated by the *arrowheads*. **B** A nested set of deletion mutants from the *Eco*RI end of pTG4B was constructed by Bal31 nuclease treatment as described in Materials and methods. Individual transformants of TG103 by these recombinant plasmids were purified and grown in TYE medium for enzyme assay using toluene-treated cells (Merberg and Datta 1982). + and − signs represent presence or absence of dehydratase activity, respectively. The *hatched areas* and the *horizontal line* represent pBR322 and *E. coli* DNA, respectively. The *open box* (in pTG5) shows the deleted segment. The sites for various restriction enzymes are designated as follows: R, *Eco*RI; S, *Sal*I; H, *Hin*dIII; A, *Acc*I, Hp, *Hpa*II; B, *Bgl*II; P, *Pst*I. The *open triangle* shows the translation-start site of the *tdc* gene. The direction of transcription is from *left* to *right*

the dehydratase protein was located within a 500-bp region from one end of Tn5 insertion site (Goss and Datta 1984). To define accurately the 5′ end of the *tdc* coding region and the direction of transcription, various restriction fragments of the 6.2 kb pEC61 DNA were subcloned into pUC plasmids in both orientations with respect to the *lac* promoter followed by transformation of strain TG103 (*tdc*::Tn*5*). Individual transformants were purified, grown in TYE medium anaerobically, and screened for dehydratase activity. The data presented in Fig. 3A show that a 2.2-kb *Sal*I/*Pst*I fragment in plasmid pTG1A (and a slightly smaller *Hin*dIII/*Pst*I fragment in pTG3B) contained the *tdc* coding region as judged by dehydratase synthesis; on the other hand, pTG7 harboring the *Eco*RI/*Bgl*II fragment showed no enzyme activity. The hybridization experiments with synthetic cDNA (Goss and Datta 1984) and the size of the dehydratase protein subunits of 36KD (Saeki et al. 1977; Kim and Datta 1982), predict that the protein coding region should span about 1000 bp from a location adjacent to the *Hpa*II site (also see below). Because the *Sal*I/*Pst*I fragment in pTG1A, extending only 650-bp leftward from *Hpa*II site, produced functional enzyme, it is likely that the direction of transcription is from left to right as depicted in Fig. 3 and the *Eco*RI/*Bgl*II fragment generated a truncated polypeptide with a missing C-terminal half of the molecule.

The experiments with several other subclones, shown in Fig. 3A, revealed an unexpected result: neither the 2.2-kb *Sal*I/*Pst*I fragment in pTG1B nor the smaller *Hin*dIII/*Pst*I fragment (pTG3A), cloned in the opposite orientation with respect to the *lac* promoter produced threonine dehydratase. Thus, the 2.2-kb fragment contains the dehydratase coding region but does not include its own promoter, although it extends about 650 bp upstream from the putative translation-start site of the *tdc* gene.

In order to locate the promoter region in the cloned DNA, a nested set of deletions was generated from the *Eco*RI end of the insert DNA in plasmid pTG4B as described in Materials and methods. By direct enzyme assay, deletions extending about 500-bp to the left of the *Sal*I site (in pTG123 and pTG122) showed dehydratase activity; whereas longer deletions of DNA sequence, as in pTG107 and pTG118, resulted in no enzyme production (Fig. 3B). These data strongly suggest that a nucleotide sequence between the end-points of pTG107 and pTG122, about 1 kb upstream from the *tdc* coding region, is required for the *tdc* gene expression.

Nucleotide sequence of the upstream region

Figure 4 shows the nucleotide sequence of DNA between the end-point of pTG122 and the *Hpa*II site which includes the putative promoter element and the translation start site of the *tdc* gene (cf. Fig. 3). It is evident that pTG122 harbors a recognizable procaryotic promoter-like sequence with a potential CAP-binding site (nucleotides 81–98) and an AT-rich region, and a ribosome binding site (Shine and Dalgarno 1975) between nucleotides 135–144 followed by an ATG codon 15 bp downstream. In contrast, pTG107 lacks the putative CAP site although it contains the other features seen in pTG122. Because cAMP is known to be required for anaerobic induction of threonine dehydratase (Shizuta and Hayaishi 1970; Phillips et al. 1978; Hobert and Datta 1983), and that a CAP-binding site on DNA is necessary for the cAMP-mediated activation of mRNA transcription of procaryotic genes (cf. de Crombrugghe et al. 1984; Ebright et al. 1984) it is likely that the deletion of the potential CAP site in pTG107 prevented transcription of the *tdc* gene.

Goss and Datta (1984) showed that lack of functional dehydratase in strain TG425 was due to insertional inactivation of the *tdc* gene by Tn*5*, and the integration site of the transposon element was within 500 nucleotides from the 5′ end of the coding region. The exact site of Tn*5* integration was located between nucleotide 785–786 (see Fig. 4) by determining the nucleotide sequence of the Tn*5* – *tdc* DNA junction in plasmid pEC6 and comparing it with the sequence obtained from the tdc^+ DNA. This result also supports the conclusion that a distal promoter is required for *tdc* gene expression because termination of transcription caused by integration of the transposon element between the promoter and the *tdc* coding region abolished dehydratase synthesis in strain TG425 (*tdc*::Tn*5*).

From the known amino acid sequence of the first 25 residues at the N-terminus of threonine dehydratase (Saeki et al. 1977; Kim and Datta 1982), we can align the translation-start site of the *tdc* gene at the ATG codon at nucleotides 1197–1199 (Fig. 4). Furthermore, by comparing with the published sequences of various bacterial operons the following features immediately upstream from the *tdc* coding region can be recognized: (a) a purine-rich region (nucleotides 1177–1187), 9 bp upstream from the ATG codon, resembling a Shine-Dalgarno sequence, (b) an 18 bp se-

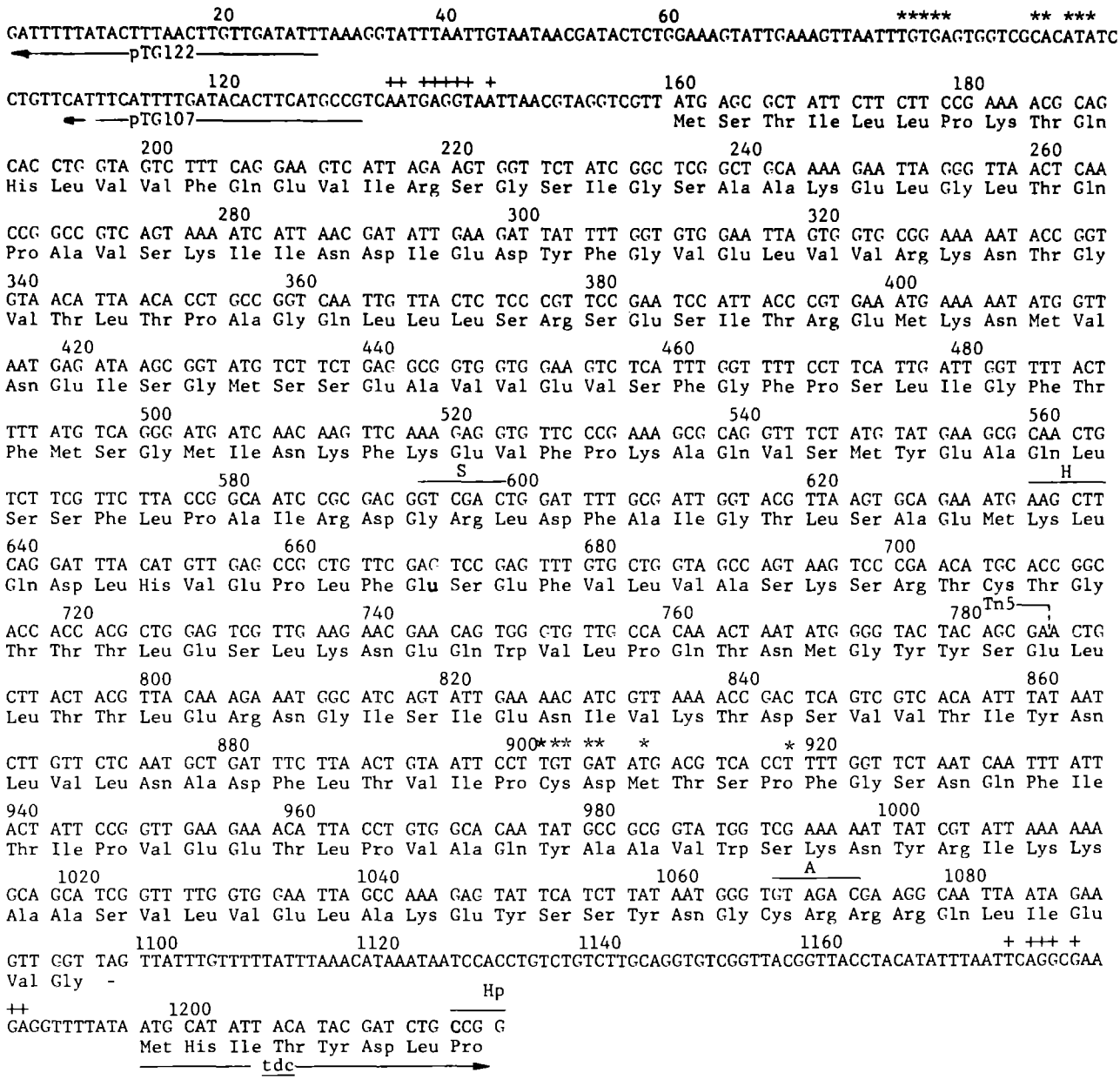

Fig. 4. Nucleotide sequence of the 5′ flanking region of the *tdc* gene. The nucleotide number is assigned arbitrarily beginning with the end-point of pTG122. The end-point of pTG107 is at nucleotide 105. The *asterisks* and the + signs above the nucleotide sequence represent homology with the consensus CAP-binding site (de Crombrugghe et al. 1984; Ebright et al. 1984) and the Shine-Dalgarno sequence, respectively. A *horizontal line* over a sequence of nucleotides designates a restriction enzyme site; the letters S, H, A and Hp denote *Sal*I, *Hin*dIII, *Acc*I and *Hpa*II, respectively. The insertion site for Tn*5* is shown between nucleotides 785–786. The translation-start site of the *tdc* gene as determined from the N-terminal amino acid sequence (Saeki et al. 1977; Kim and Datta 1982), is aligned with the ATG codon at nucleotides 1197–1199. The deduced amino acid sequence of the open reading frame is also shown

quence (nucleotides 901–918) with a diad symmetry having a good homology with the consensus cAMP:CAP binding site (de Crombrugghe et al. 1984; Ebright et al. 1984), and (c) several AT-rich regions that may serve as potential RNA polymerase binding sites. Although these features imply that this region of DNA may serve as a putative promoter, its role in *tdc* gene expression remains to be deciphered (see Discussion).

Plasmid-directed protein synthesis

The nucleotide sequence in Fig. 4 also revealed a 936-bp open reading frame starting from the ATG codon at nucleotides 160–162. To detect the presence of this putative polypeptide with 312 amino acid residues, in vivo protein synthesis was carried out in *E. coli* minicells harboring various plasmids. The data presented in Fig. 5 show that pTG4B (in lanes 1 and 2) produced, among others, two polypeptides of about 32 KD and 36 KD, both of which were absent in the vector pUC13 (lane 6). The 36 KD polypeptide was identified as threonine dehydratase by its predicted size (Saeki et al. 1977; Kim and Datta 1982) and immunological reactivity with antibodies raised against the purified protein (lane 7). Control experiments indicated that pTG7A (same as pTG7 but with inverted orientation with respect to the *lac* promotor) harboring the RI/BglII fragment produced

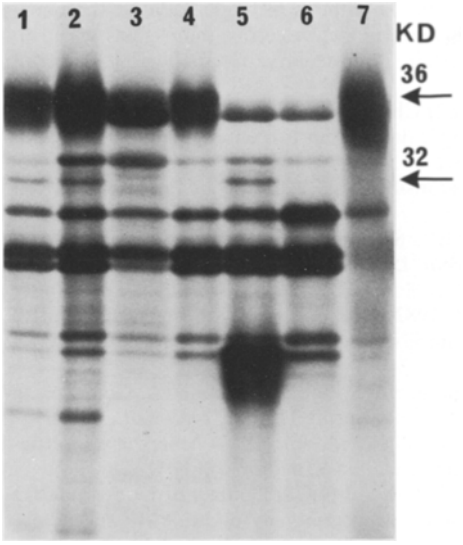

Fig. 5. Autoradiograph of [^{35}S]-methionine-labeled polypeptides synthesized in minicells harboring various plasmids. Lanes 1, 2 and 7, pTG4B; 3, pTG5; 4, pTG8; 5, pTG7A; 6, pUC13. Minicells were grown aerobically with the exception of pTG4B in Lane 1 where cells were grown anaerobically. In Lane 7, the cytosolic fraction of labeled minicells harboring pTG4B was precipitated with antibodies against purified threonine dehydratase (see Materials and methods). The positions of 36 KD and 32 KD polypeptides, based on migration of molecular weight markers, are indicated by *arrows*

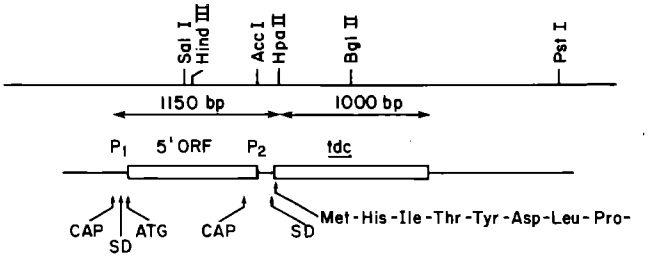

Fig. 6. A schematic representation of the *tdc* gene and its 5′ upstream region. For a discussion, see text

only the 32 KD polypeptide (lane 5), whereas pTG5 with the SalI/AccI deletion, which removed almost half of the 5′ open reading frame, synthesized the dehydratase polypeptide only (lane 3). Similarly, the plasmid pTG8 containing a frameshift in the 5′ open reading frame (constructed by digesting pTG4B with *Sal*I, blunt-ended with S1 nuclease and ligated under blunt-end conditions) also produced the dehydratase polypeptide and not the 32 KD polypeptide (lane 4). These data suggest that the 936-bp open reading frame indeed codes for a polypeptide chain, and the presence or absence of this polypeptide does not have any effect on the expression of the *tdc* gene.

Discussion

Cloning of a gene and its regulatory region is an essential first step to study the structure and regulation of gene expression at the molecular level. For a long time, the isolation of the *tdc* gene of *E. coli* presented a practical problem because of the difficulty in selecting mutants due to lack of phenotype and the absolute requirement of four amino acids, an electron acceptor and cAMP for enzyme induction in an anaerobic environment. To overcome this problem, we isolated a threonine dehydratase-negative mutant by insertional inactivation with Tn5 and cloned the DNA containing the integrated Tn5 into pBR322 (Goss and Datta 1984). The experiments reported herein describe a simple, rapid procedure for isolating the functional *tdc*$^+$ gene by in vivo replacement of the Tn5 element on the plasmid. The exact mechanism for the acquisition of Kan^s phenotype and *tdc*$^+$ gene however remains unclear. The recovery of Kan^s phenotype from Amp^r plasmid DNA was approximately 10^{-5}, a value significantly higher than that observed (10^{-6} to 10^{-10}) for precise excision of Tn5 (Kleckner 1977). This result tends to suggest that exchange-recombination of plasmid DNA with chromosomal DNA rather than transposition resulted in the replacement of *tdc*::Tn*5* on the plasmid.

The method outlined here may have wider applicability for cloning *E. coli* genes that lack easily scorable phenotype or require complex culture conditions for gene expression. A crucial element in the procedure was to find a restriction enzyme that cleaved the integrated Tn5 DNA but not the plasmid vector or *E. coli* insert DNA thereby facilitating the selection of Kan^s transformants. A computer search of 52 kb nucleotide sequences of *E. coli* DNA from EMBL Nucleotide Data Base revealed only 4 *Xho*I target sites (i.e. on average one site/13 kb) as contrasted with the random distribution of one site/4 kb predicted for a hexanucleotide target. Although the number of nucleotide sequences examined represents only about 1% of the total *E. coli* genome, it would appear that a large number of genes may have no *Xho*I site within their coding sequences and thus can be cloned in vivo without the use of a genomic library.

The organization of the *tdc* gene and its 5′ flanking sequence is schematically depicted in Fig. 6. The nucleotide sequence data revealed a promoter-like sequence, designated P_1, approximately 1 kb upstream from the *tdc* coding region. Results of the subcloning experiments and Bal31 deletion analysis showed that P_1 is needed for *tdc* gene expression. This conclusion was strengthened by two other observations: (a) deletion of the *Sal*I/*Acc*I region (in pTG5, see Figs. 3A and 4), which removed a large part of the upstream DNA, had no discernible effect on dehydratase synthesis, and (b) insertion of Tn5 at a site about 400-bp upstream from the *tdc* coding region abolished dehydratase synthesis. These cumulative findings also suggest that the DNA segment between P_1 and the end of the *tdc* gene is transcribed as a single transcriptional unit, and that the nucleotide sequence immediately upstream of the *tdc* coding region may not function as a *cis*-dominant regulatory sequence for *tdc* gene expression in TYE medium.

Although the nucleotide sequence immediately upstream from the *tdc* coding region appears to contain a putative CAP site and an AT-rich region (designated P_2 in Fig. 6), the significance of this region as a potential transcription start site remains to be deciphered, especially when transcription of the *tdc* gene presumably occurs by readthrough from P_1. It is possible to envisage that the region around P_2 may have a regulatory role in influencing transcription from P_1 under specific physiological conditions.

The expression in minicells of various recombinant plasmid DNAs revealed a 32 KD polypeptide from the 936-bp open reading frame upstream from the *tdc* gene. The pres-

ence or absence of this polypeptide did not influence dehydratase synthesis suggesting that it does not play any regulatory part in *tdc* gene expression. Because the in vivo function of threonine dehydratase is still not understood, the physiological significane of the 32 KD polypeptide in anaerobic threonine metabolism in *E. coli* also remains to be established. A computer search of *Newat* date base (Doolittle 1981) has failed to detect homology with a known protein (Doolittle personal communication).

Acknowledgments. A part of this work was carried out by P.D. in the laboratory of Charles Yanofsky at Stanford University to whom we are grateful. We thank R.F. Doolittle of the University of California, San Diego, D.L. Oxender and L. Grossman of the University of Michigan for help in computer search. This work was supported by U.S. Public Health Service grant GM 21436 from the National Institutes of Health. P.D. was recipient of a Senior Fellowship 1 F33 GM 08957 from the National Institute of General Medical Sciences.

References

Braus G, Argast M, Beck CF (1984) Identification of additional genes on transposon Tn10: *tetC* and *tetD*. J Bacteriol 160:504–509

Clark-Curtiss JE, Curtiss R (1983) Analysis of recombinant DNA using *Escherichia coli* minicells. Meth Enzymol 101:347–362

de Crombrugghe B, Busby S, Buc H (1984) Cyclic AMP receptor protein: role in transcription activation. Science 224:831–838

Doolittle RF (1981) Similar amino acid sequences: change or common ancestry? Science 214:149–159

Ebright RH, Cossart P, Gicquel-Sanzey B, Beckwith J (1984) Mutations that alter DNA sequence specificity of the catabolite gene activator protein of *E. coli*. Nature 311:232–235

Goss TJ, Datta P (1984) *Escherichia coli* mutation that inactivates biodegradative threonine dehydratase by transposon Tn5 insertion. J Bacteriol 158:826–831

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

Kim SS, Datta P (1982) Chemical characterization of biodegradative threonine dehydratases from two enteric bacteria. Biochim Biophys Acta 706:27–35

Kleckner N (1977) Translocatable elements in procaryotes. Cell 11:11–23

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 1–545

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, Cold Spring Harbor, New York, p432

M13 cloning and sequencing handbook (1983) Amersham International plc, Amersham, United Kingdom, pp 1–41

Phillips AT, Egan RM, Lewis B (1978) Control of biodegradative threonine dehydratase inducibility by cyclic AMP in energy-restricted *Escherichia coli*. J Bacteriol 135:828–840

Saeki Y, Ito S, Shizuta Y, Hayaishi O, Kagamiyama H, Wada H (1977) Subunit structure of biodegradative threonine deaminase. J Biol Chem 252:2206–2208

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

Shine J, Dalgarno L (1975) Determinant of cistron specificity in bacterial ribosomes. Nature 254:34–38

Shizuta Y, Hayaishi O (1970) Regulation of biodegradative threonine dehydratase synthesis in *Escherichia coli* by cyclic adenosine 3′,5′-monophosphate. J Biol Chem 245:5416–5423

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

Communicated by E.K.F. Bautz

Received June 5 / August 15, 1985

Note added in proof

We have found two errors in the nucleotide sequence shown in Fig. 4. The corrected sequence should read A at nucleotide 166 and G at nucleotide 852