Cloning and Nucleotide Sequence of the *Escherichia coli* Cytidine Deaminase (*ccd*) Gene†,‡

Charles Yang,§,⊥ Dean Carlow,§ Richard Wolfenden,§ and Steven A. Short*,∥

Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina 27599, and Molecular Genetics and Microbiology, Wellcome Research Laboratory, Research Triangle Partk, North Carolina 27709

Received November 25, 1991

ABSTRACT: The structural gene that encodes cytidine deaminase (*cdd*) in *Escherichia coli* was cloned from Kohara phage λ365 (7F1), and its nucleotide sequence was determined. Plasmids harboring the gene complemented chromosomal *cdd* mutations, enhanced cytidine deaminase activity in cell extracts, and directed the synthesis of a protein identical in mass and N-terminal amino acid sequence with cytidine deaminase purified from wild-type bacteria. Metal analysis of the purified, plasmid-encoded deaminase indicated a single atom of tightly bound zinc per subunit. Earlier work has shown that bacterial cytidine deaminase and mammalian adenosine deaminase are remarkably alike in their mechanisms of action, in their free energies of interaction with analogue inhibitors resembling tetrahedral intermediates in nucleophilic substitution, and in their ability to discriminate between analogue inhibitors differing by a single hydroxyl group. In contrast to these shared catalytic similarities, the deduced amino acid sequence of *E. coli* cytidine deaminase (monomer MW 31 540) differs markedly from the mammalian adenosine deaminase sequence suggesting major differences in their tertiary structures. Nevertheless, cytidine deaminase and mammalian plus bacterial adenosine deaminases share a single region (TVHA) of sequence identity that is tentatively identified as part of the cytidine deaminase active site.

Bacterial cytidine deaminase and mammalian adenosine deaminase are excellent catalysts, enhancing the rates of hydrolytic deamination of nucleosides by factors in the neighborhood of 10^{12} (Frick et al., 1987). These enzymes share a susceptibility to strong inhibition by compounds that resemble unstable intermediates in direct water attack on their substrates. In addition, rare hydrated forms of inhibitory substrate analogues, in which hydrogen replaces the leaving group ammonia, are stabilized at the active sites of both cytidine and adenosine deaminases by factors of approximately 10^{10} (Kati & Wolfenden, 1989a,b; Frick et al., 1989). In both cases, a single ligand hydroxyl group appears to contribute roughly ~10 kcal/mol to the binding affinity of the hydrated inhibitor, the highest level of enzyme binding discrimination that appears to have been recorded for a single hydroxyl substituent.

To uncover the structural basis of these remarkable powers of binding discrimination and to assist the solution of the crystal structure of bacterial cytidine deaminase, it would be desirable to learn the amino acid sequence of bacterial cytidine deaminase. The amino acid sequence and three-dimensional structure of mouse adenosine deaminase have been reported recently (Chang et al., 1991; Wilson et al., 1991), and crystals of bacterial cytidine deaminase have been prepared in the presence and absence of the transition-state analogue inhibitor 5-fluoropyrimidin-2-one ribonucleoside (Betts et al., 1989). This paper describes the cloning of the *cdd* gene that encodes cytidine deaminase in *Escherichia coli* K-12 and the determination of its DNA sequence. The similarities in amino acid sequence between cytidine deaminase and adenosine deaminase, which are very limited, point to several mechanistic inferences.

Table I: Bacterial Strains and Plasmids

strain or plasmid	characteristics	source
E. coli		
N99	F⁻ *gal rpsL*	M. Gottesman
N4956	F⁻ *thi-1 leuB6 thr-1 pro lacY1 tonA21 supE44 hsdR* λ⁻	M. Gottesman
JM103	Δ(*lac-pro*) *thi rpsL supE endA sbcB15*/F′ *traD36 proAB lacI*ᵠ *lacZ*Δ*M15*	J. Messing
Sϕ441	F⁻ *cdd-5 upp-11 relA1 metB1 rpsL*	B. Bachmann
plasmids		
pCY1	Ap^r plus a 13-kb *Bam*HI/*Hin*dIII-ended DNA insert from Kohara phage λ365	this study
pCY2	Ap^r, *Hpa*I deletion derivative of pCY1	this study
pCY3	Ap^r, *Eco*RV deletion derivative of pCY1	this study
pCY4	Ap^r, *Cla*I deletion derivative of pCY1	this study
pCY5	Ap^r, *Bam*HI/*Mlu*I deletion derivative of pCY4	this study
pCY6	Ap^r, *Bam*HI/*Hpa*I deletion derivative of pCY4	this study
pCY9	Ap^r, *Bgl*II⁻ derivative of pCY4	this study
pCY10	Ap^r, *Ssp*I deletion derivative of pCY5	this study
pCY11	Ap^r, pCY10 having a *Bam*HI linker (GGTGGATCCACC) inserted at the unique *Ssp*I site	this study

Materials and Methods

Bacterial Strains, Plasmids, and Bacteriophage. The *E. coli* strains used in this study are listed in Table I. The *E. coli* strains N99 and JM103 were used for the propagation of λ and M13 phages, respectively. Plasmids constructed during the cloning of *cdd* are derivatives of pBR322 and are described in Table I. Stocks of the Kohara λ phage set encompassing the 46-min region of the *E. coli* chromosome were kindly provided by K. Rudd (Kohara et al., 1987).

Media and Growth Conditions. All *E. coli* strains were grown at 37 °C in a complete minimal medium as described by Vogel and Bonner (1956). The carbon sources for liquid and solid media were 0.2% glucose, 0.4% glycerol, or 0.2%

† This work was supported, in part, by Research Grant GM-18325 from the National Institutes of Health (to R.W.).

‡ The *cdd* nucleotide sequence given in this paper has been assigned GenBank Accession Number M60916.

* Corresponding author.

§ University of North Carolina.

⊥ Present address: Lineberger Cancer Center, University of North Carolina, Chapel Hill, NC 27599.

∥ Wellcome Research Laboratory.

0006-2960/92/0431-4168$03.00/0

cytidine. Amino acids required to satisfy auxotrophic requirements and ampicillin were added to the media at final concentrations of 50 μg/mL and 125 μg/mL, respectively. Other media were prepared as described by Miller (1972).

Recombinant DNA Procedures. Plasmid DNA and DNA extracted from the Kohara λ phages were isolated and digested, and their restriction digests were analyzed as described previously (Barbier & Short, 1985; Maniatis et al., 1982; Short & Singer, 1984). Specific DNA fragments used in *cdd* subcloning experiments were isolated from preparative agarose gels by electroelution or by binding to GeneClean glassmilk (Bio101, La Jolla, CA). DNA ligations using T4 DNA ligase (New England BioLabs, Inc., Beverly, MA) were carried out in TA buffer prepared as described by O'Farrell et al. (O'Farrell et al., 1980). Competent *E. coli* were prepared and transformed as described by Maniatis et al. (1982). Kohara λ phage harboring the *cdd* structural gene were identified by in situ hybridization using a ^{32}P-labeled synthetic oligonucleotide (Sambrook et al., 1989). The 19-base oligonucleotide probe was designed by back-translation of an N-terminal cytidine deaminase heptapeptide and was a degenerate pool composed of 48 oligonucleotides.

For determination of the *cdd* nucleotide sequence, M13mp19 derivatives having the *cdd* structural gene cloned in opposite orientations were constructed. Ordered, overlapping deletions within the DNA insert of each CDD$^+$ M13mp19 derivative were prepared using T4 DNA polymerase as described by Dale et al. (1985). Dideoxy chain termination sequencing (Sanger et al., 1977) was performed using [α-^{35}S]dATPαS and Sequenase Version 2.0 (United States Biochemical Corp., Cleveland, OH). Compressions were resolved by replacing dGTP sequencing reactions with reactions containing either 7-deaza-dGTP or dITP. Electrophoresis of the DNA sequencing reactions employed 8% polyacrylamide/8 M urea exponential wedge gels (spacers from CBS Scientific, Delmar, CA) maintained at 55 °C using a thermostatic plate. Computer analysis of the nucleotide sequences obtained was performed with the GCG sequence software package (Devereux et al., 1984).

Nucleotide Sequence Accession Number. The *cdd* nucleotide sequence has been assigned GenBank Accession Number M60916.

Analysis of Cytidine Deaminase in Cell Extracts. Bacterial strains to be assayed for cytidine deaminase activity were grown for 16–18 h at 37 °C in complete minimal medium containing glycerol as the carbon source. The final A_{600} for these cultures ranged from 2.9 to 3.2. Cells collected from 25 mL of each culture were washed once with an equal volume of cold, 100 mM potassium phosphate (pH 7.0) and then resuspended in 1 mL of the same buffer containing 1 μg/mL lysozyme. The cell suspension was subjected to two freeze–thaw cycles and then sonicated for 30 s in 10-s bursts followed by 50-s periods of cooling on ice. Cell debris and intact cells were removed by centrifugation, and the cell-free extracts were stored at −20 °C without loss of activity. Cytidine deaminase activity present in each cell-free extract was determined at 30 °C by measuring the decrease in absorbance at 282 nm due to the conversion of cytidine to uridine [$\Delta\epsilon$ (mM) = −3.60] in a reaction mixture containing 50 mM Tris-HCl (pH 7.5) and 400 μM cytidine. One cytidine deaminase unit catalyzes the deamination of 1 μmol of cytidine to uridine per minute per milliliter.

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of cell-free extracts was performed by using mini-slab gels (0.75 mm × 7 cm) and the running buffer described by Thomas and Kornberg (1975). The slab gels consisted of a 6-cm 16.5% separating gel (200:1 acrylamide:bisacrylamide) overlaid with a 1-cm 4.5% stacking gel (36:1 acrylamide:bisacrylamide). Electrophoresis of the proteins through the stacking and separating gels was carried out at a constant voltage of 50 V and 200 V, respectively. The protein concentration of cell-free extracts was measured by the method of Lowry (1951) using bovine serum albumin as the protein standard.

Purification and Characterization of Cytidine Deaminase Expressed from CDD$^+$ Recombinant Plasmids. Cytidine deaminase was purified from *E. coli* S*φ441*(pCY4). The bacterial culture was grown at 37 °C in Luria broth (Miller, 1972) supplemented with 0.2% glycerol and 150 μg/mL ampicillin to late log phase; the cells were collected and washed with cold 50 mM Tris-HCl (pH 7.5) containing 50 mM NaCl, and the enzyme was purified as described previously (Ashley & Bartlett, 1984; Frick et al., 1989) except that all chromatography steps were carried out at 4 °C using FPLC columns. This procedure yielded enzyme having a purity >98% with an overall yield of ~30%. The amino acid composition and N-terminal amino acid sequence were determined as described by Matsudaira (1987). The molar absorptivity for cytidine deaminase was calculated from spectroscopic data obtained with solutions of known enzyme concentrations. The protein concentrations of these cytidine deaminase solutions were determined by amino acid analysis using a Beckman System 6300 amino acid analyzer. The metal content of dialyzed, purified cytidine deaminase was measured using an ARL Model 34000 inductively coupled plasma optical emission spectrometer, through the courtesy of Dr. J. W. Olesik (Department of Chemistry, UNC—Chapel Hill). For this measurement, enzyme (5 mg) was dissolved in 0.01 M imidazole hydrochloride buffer (5 mL) and dialyzed against the same buffer (1000 mL) with three changes over a period of 72 h.

Production of Anti-Cytidine Deaminase Antibody. Cytidine deaminase purified to greater than 98% homogeneity was resuspended in Freund's complete adjuvant and used to immunize New Zealand rabbits. Two weeks after the initial immunization, the rabbits were again injected with purified enzyme emulsified in Freund's complete adjuvant, followed two weeks later by a second booster injection of enzyme suspended in incomplete Freund's adjuvant. The rabbits were bled prior to the onset of the immunization regime (preimmune serum) and at 7, 14, and 21 days following injection of enzyme in incomplete Freund's adjuvant. Pre- and postimmune sera were tested against purified cytidine deaminase using the double-immunodiffusion technique. Preimmune serum produced no specific cytidine deaminase precipitin bands. The different postimmune serum samples showing cytidine deaminase specific precipitin bands were pooled, and the polyvalent antibody was purified by ammonium sulfate precipitation. The final ammonium sulfate precipitated γ-globulin fraction was dissolved in 10 mM potassium phosphate buffer (pH 7.6) and dialyzed extensively against the same buffer. Aliquots of the antibody were frozen at −70 °C.

Results

Cloning the cdd Structural Gene. A collection of Kohara λ phage spanning the 46-min region of the *E. coli* chromosome was screened for the *cdd* structural gene by in situ hybridization using a mixed oligonucleotide probe having the sequence 5′-TGCA(C/T)CC(G/A/T)CG(C/T)TT(C/T)CA-(G/A)AC-3′. The nucleotide sequence of this oligonucleotide probe was derived following back-translation of the first seven cytidine deaminase amino-terminal amino acids (MHPRFQT)

Table II: Cytidine Deaminase Specific Activities Determined for Sϕ441 Clones Harboring pCY1 and Its Derivative Plasmids

strain	sp act.[a] (units/ mg of protein)	strain	sp act.[a] (units/ mg of protein)
Sϕ441 (pBR322)	<0.003	Sϕ441 (pCY5)	3.76
Sϕ441 (pCY1)	15.0	Sϕ441 (pCY6)	<0.005
Sϕ441 (pCY2)	<0.002	Sϕ441 (pCY9)	<0.007
Sϕ441 (pCY3)	<0.006	Sϕ441 (pCY10)	3.86
Sϕ441 (pCY4)	14.2		

[a] The cytidine deaminase specific activity was determined for cell-free extracts prepared from each clone grown in complete minimal medium containing 0.4% glycerol and 150 μg/mL ampicillin as described in Materials and Methods. One cytidine deaminase unit is defined as 1 μmol of cytidine deaminated/(min·mL).

using codons found predominantly in highly expressed *E. coli* proteins (Aota et al., 1988). Of the phage tested, only plaques formed by Kohara phage λ365 (Kohara no. 7F1) hybridized with the oligonucleotide probe back-translated from the cytidine deaminase amino-terminal hexapeptide.

DNA fragments of Kohara phage λ365 that resulted from single- or double-restriction endonuclease digestions using *Eco*RI, *Bam*HI, *Eco*RV, and *Hin*dIII were cloned into appropriately digested pBR322. Each potential *cdd* plasmid gene bank was used to transform *E. coli* N4956 and a pooled population of transformants selected in liquid media containing ampicillin. For each Ap^r N4956 transformant pool, minilysate plasmid DNA was prepared and used to transform *E. coli* S*ϕ441* (*cdd*). Since a functional cytidine deaminase is required for growth of S*ϕ441* on cytidine as the sole carbon and energy source, CDD$^+$ S*ϕ441* transformants were selected for by requiring growth on minimal agar media containing 0.2% cytidine and ampicillin. Of the various gene banks prepared from Kohara phage λ365, only pBR322 derivatives harboring *Bam*HI/*Hin*dIII-ended λ365 DNA inserts yielded CDD$^+$ S*ϕ441* transformants. The pBR322 hybrid isolated from one CDD$^+$ transformant was designated pCY1. Analysis of the DNA insert carried by pCY1 indicated that this plasmid contained a single 13-kb DNA insert bounded by single *Bam*HI and *Hin*dIII cleavage sites (Figure 1). The data presented in Table II confirm that the CDD$^+$ growth phenotype of S*ϕ441*(pCY1) results directly from the fact that pCY1 carries a functional *cdd* gene. The cytidine deaminase specific activity determined for the crude cell-free extract prepared from *E. coli* S*ϕ441*(pCY1) was approximately equal to 6% of the specific activity determined for the purified enzyme (Frick et al., 1989) and is at least 5000-fold higher than the specific activity measured for CDD$^-$ strain S*ϕ441*(pBR322).

Localization of the cdd Gene within the pCY1 DNA Insert. Deletion derivatives of pCY1 were constructed in order to map the location of the *cdd* gene within the pCY1 DNA insert and to prepare a subclone containing the *cdd* gene on a small DNA fragment suitable for nucleotide sequence determination. Deletion of pCY1 insert DNA bounded by *Hpa*I or *Eco*RV endonuclease cleavage sites yielded plasmids pCY2 and pCY3, respectively (Figure 1), both of which lacked the ability to encode a functional cytidine deaminase (Table II). The construction of *Cla*I deletion derivative pCY4 yielded CDD$^+$ S*ϕ441* transformants on cytidine minimal agar medium, yielded S*ϕ441*(pCY4) cell extracts having cytidine deaminase specific activity equal to that determined for S*ϕ441*(pCY1) extracts (Table II), and thus localized the *cdd* gene to the *Bam*HI proximal, 4-kb DNA segment of pCY1 (Figure 1). Plasmid pCY5 was constructed by removal of pCY4 DNA bounded by *Bam*HI and *Mlu*I endonuclease cleavage sites (Figure 1). In the generation of pCY5, the *Mlu*I site was abolished and the *Bam*HI site was repositioned adjacent to the original *Mlu*I site of pCY4. All subsequent plasmid constructs derived from pCY5 possess this repositioned *Bam*HI site. Plasmid pCY5 conferred a CDD$^+$ phenotype upon all S*ϕ441* transformants. Determination of the cytidine deaminase specific activity for S*ϕ441*(pCY5) extracts demonstrated that this plasmid encoded a functional cytidine deaminase but that the level of enzyme synthesis was reduced approximately 4-fold relative to the cytidine deaminase level recorded for S*ϕ441* clones harboring plasmids pCY1 or pCY4 (Table II).

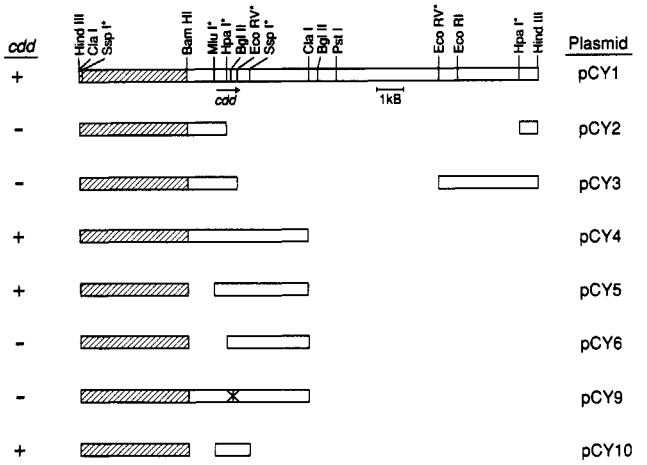

FIGURE 1: Plasmid maps for CDD$^+$ pCY1 and its derivatives generated in the cloning of the *cdd* gene. Relevant restriction sites used in the characterization of pCY1 and for the construction of deletion derivative plasmids are indicated. The hatched segment of each plasmid map denotes pBR322 DNA and the open segment denotes cloned DNA. Deletions are represented by gaps. The *cdd* genotype for each construct is indicated on the left of the drawing. Restriction sites marked with an asterisk (*) indicate that additional endonuclease cleavage sites are present in pCY1 but were not mapped in this study. In pCY9, the cross through the line representing the *Bgl*II site signifies the destruction of this endonuclease recognition site.

Deletion of pCY4 DNA between the unique *Bam*HI and *Hpa*I sites (see Figure 1) or destruction of the unique *Bgl*II cleavage site of pCY4 resulted in plasmids pCY6 and pCY9, respectively (Figure 1), both of which were incapable of directing the synthesis of functional cytidine deaminase (Table II). These results, in conjunction with a knowledge of the subunit size determined for purified cytidine deaminase (Ashley & Barlett, 1984; Frick et al., 1989), indicated that deletion of DNA delimited by the *Ssp*I sites located in the *cdd*-containing insert and vector regions of pCY5 would generate a plasmid that should maintain a functional *cdd* gene. That this was the case is shown by data obtained with plasmid pCY10 (Figure 1; Table II). From pCY10, CDD$^+$ plasmid pCY11 was constructed by insertion of a *Bam*HI linker (GGTGGATCCACC) into the unique *Ssp*I site, yielding a plasmid having the *cdd* gene on a 1.3-kb insert bounded by *Bam*HI endonuclease cleavage sites.

Identification of the Plasmid-Encoded Cytidine Deaminase Polypeptide. When cell extracts prepared from selected CDD$^+$ and CDD$^-$ S*ϕ441* clones were analyzed by SDS–polyacrylamide gel electrophoresis, the protein product coded for by the *cdd* structural gene was readily identified (Figure 2). Comparison of the protein profiles obtained with extracts from CDD$^+$ clones S*ϕ441*(pCY1) or S*ϕ441*(pCY4) with that from CDD$^-$ clone S*ϕ441*(pCY2) demonstrated that the CDD$^+$ plasmids encode a protein having a subunit molecular weight of approximately 32 000 that is absent from the protein profile obtained with the extract from a CDD$^-$ clone. The polypeptide identified in Figure 2 as the *cdd* gene product had the same

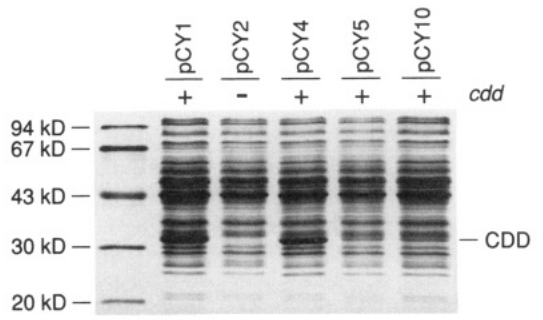

FIGURE 2: Protein profiles obtained following polyacrylamide electrophoresis of cell extracts prepared from Sϕ*441* clones harboring CDD$^{+/-}$ plasmids. Plasmid *cdd* genotype is indicated below each name, and the cytidine deaminase polypeptide ($M_r \sim 32\,000$) is indicated by CDD at the right of the figure. The protein profiles shown in this figure were obtained by electrophoresis of cell extract samples containing 25 μg of protein.

subunit size as that reported for chromosomally encoded cytidine deaminase purified from other *E. coli* strains (Ashley & Bartlett, 1984; Frick et al., 1989) and reacted with polyvalent antibody raised against purified cytidine deaminase. Extracts prepared from Sϕ*441* harboring plasmids pCY5 and pCY10 also contained the cytidine deaminase polypeptide identified in extracts from pCY1/pCY4 clones, but the level of cytidine deaminase subunit polypeptide in the pCY5/pCY10 extracts was reduced relative to the level of cytidine deaminase polypeptide found in the pCY1/pCY4 extracts. The difference in cytidine deaminase polypeptide abundance found for the two plasmid sets is consistent with and predicted by the differences in cytidine deaminase specific activities measured for extracts prepared from the pCY5/pCY10 clones relative to the cytidine deaminase activity observed for the pCY1/pCY4 clones (Table II).

Nucleotide Sequence of the cdd Gene. The DNA sequence for the *cdd* structural gene, given in Figure 3, was determined from both DNA strands using a complete and overlapping set of M13mp19 derivatives carrying the 1.3-kb insert isolated from pCY11. Computer analysis of the sequence identified a single 294-codon open reading frame (ORF) beginning with an ATG which when translated would correspond to a polypeptide having a subunit molecular weight of 31 540. A survey of the nucleotide sequence 5′ to the cytidine deaminase ORF revealed (i) a sequence GAGG (nucleotides 81–84) which when translated into mRNA would be complementary to the 3′ end of 16S rRNA and could function as a ribosome binding site (Shine & Dalgarno, 1974); (ii) the sequence 5′-TAATGAGATTCAGATCACATAT starting at nucleotide 14 in Figure 3 which has 78% identity with the 22-bp consensus CAP binding site (Zhang & Ebright, 1990) and has been identified as one of the two CAP sites reported for the *cdd* promoter region (Valentin-Hansen et al., 1989); and (iii) a sequence identified as a potential −10 region of the *cdd* promoter located at bases 54–59. It should also be noted that the *cdd* promoter lacks a −35 region but contains a 22-nucleotide sequence having homology with the canonical CAP site upstream of the potential −10 region of the *cdd* promoter as would be expected for a gene whose transcription is CAP activated (Munch-Petersen et al., 1972).

Properties of Cytidine Deaminase Encoded by pCY4. The enzyme purified from *E. coli* expressing the *cdd* gene from a multicopy plasmid was found to have physical and catalytic parameters that were similar to those reported previously for this deaminase (Ashley & Bartlett, 1984; Cohen & Wolfenden, 1971; Vita et al., 1985). Purified cytidine deaminase exhibited a K_m value for cytidine of 120–140 μM and a specific activity of 600–700 μmol of cytidine deaminated/(min·mg of protein) ($k_{cat} = 299\ s^{-1}$), a value which is approximately 3-fold higher than that reported previously for this deaminase. The extinction coefficient at 280 nm for purified *E. coli* cytidine deaminase was determined to be 39 000 $M^{-1}\ cm^{-1}$. When analyzed for zinc by inductively coupled plasma optical emission spectroscopy, the results of four experiments indicated that tightly bound zinc was present at a concentration equivalent to 1.01 ± 0.05 g-atoms/31 540 g of protein, which is equivalent to 1 zinc per deaminase subunit. No other metal, including Ni, Co, Cu, Mn, or Fe, was present in significant quantities.

Cytidine deaminase was stored as a precipitate under 70% ammonium sulfate at 4 °C for >6 months without a loss of activity. The native, dimeric form of the enzyme appeared to be very stable; individual subunits were not detected following gel filtration chromatography of the crude or purified enzyme or by centrifugation of the purified enzyme through sucrose gradients. When purified cytidine deaminase was analyzed for amino acid composition and N-terminal sequence, the data obtained were in complete agreement with the amino acid composition and N-terminal sequence predicted from the *cdd* DNA sequence. This latter result demonstrates that mature, cytosolic cytidine deaminase does not undergo post-translational methionine excision. This finding agrees with the specificity reported for *E. coli* methionyl-aminopeptidase (Hirel et al., 1989).

Discussion

Several lines of evidence suggest that the open reading frame encoding cytidine deaminase has been correctly identified. Specifically, endonuclease cleavage sites, identified as critical for the synthesis of functional enzyme, are located within the DNA identified as the *cdd* structural gene (Figures 1 and 3). The subunit molecular weight calculated from the deduced amino acid sequence (M_r 31 540) agrees with the value (32 000) determined for the purified enzyme (Ashley & Bartlett, 1984; Frick et al., 1989). In cell extracts prepared from the CDD$^+$ clones, the level of cytidine deaminase activity and the immunoreactivity with polyvalent antibody raised against pure cytidine deaminase were found to match the relative abundance of a protein having a subunit molecular weight of ~32 000 (Table II and Figure 2). Finally, the N-terminal amino acid sequence, deduced from the nucleotide sequence, agrees with the N-terminal 22-residue sequence observed for the purified enzyme (Frick et al., 1989).

The *cdd* structural gene has previously been mapped to the 46–47-min region of the *E. coli* chromosome (Josephsen et al., 1983). If the restriction map prepared for the DNA inserts of the CDD$^+$ pCY plasmids constructed in this study are aligned with the map for Kohara phage λ365, then the *cdd* structural gene can be positioned to the 2241–2242-kb segment of the *E. coli* chromosome in such a way that *cdd* transcription proceeds in a clockwise direction (Kohara et al., 1987; Rudd et al., 1990). Comparison of the restriction maps for CDD$^+$ pCY plasmids and the sequence data for *cdd* furnishes a plausible explanation of the higher levels of enzyme activity that were observed for plasmids pCY1 and pCY4 than those observed for plasmids pCY5 and pCY10 (Table II). A recent study of transcriptional regulation at the *cdd* promoter, using *cdd–lac* gene fusions, identified tandem CAP sites located at positions 41 and 93 nucleotides 5′ to the site of initiation of *cdd* transcription (Valentin-Hansen et al., 1989). Plasmids pCY1 and pCY4 contain an intact *cdd* promoter region 5′ of the structural gene, whereas plasmids pCY5 and pCY10 possess only one (the *cdd* proximal site) of the two CAP

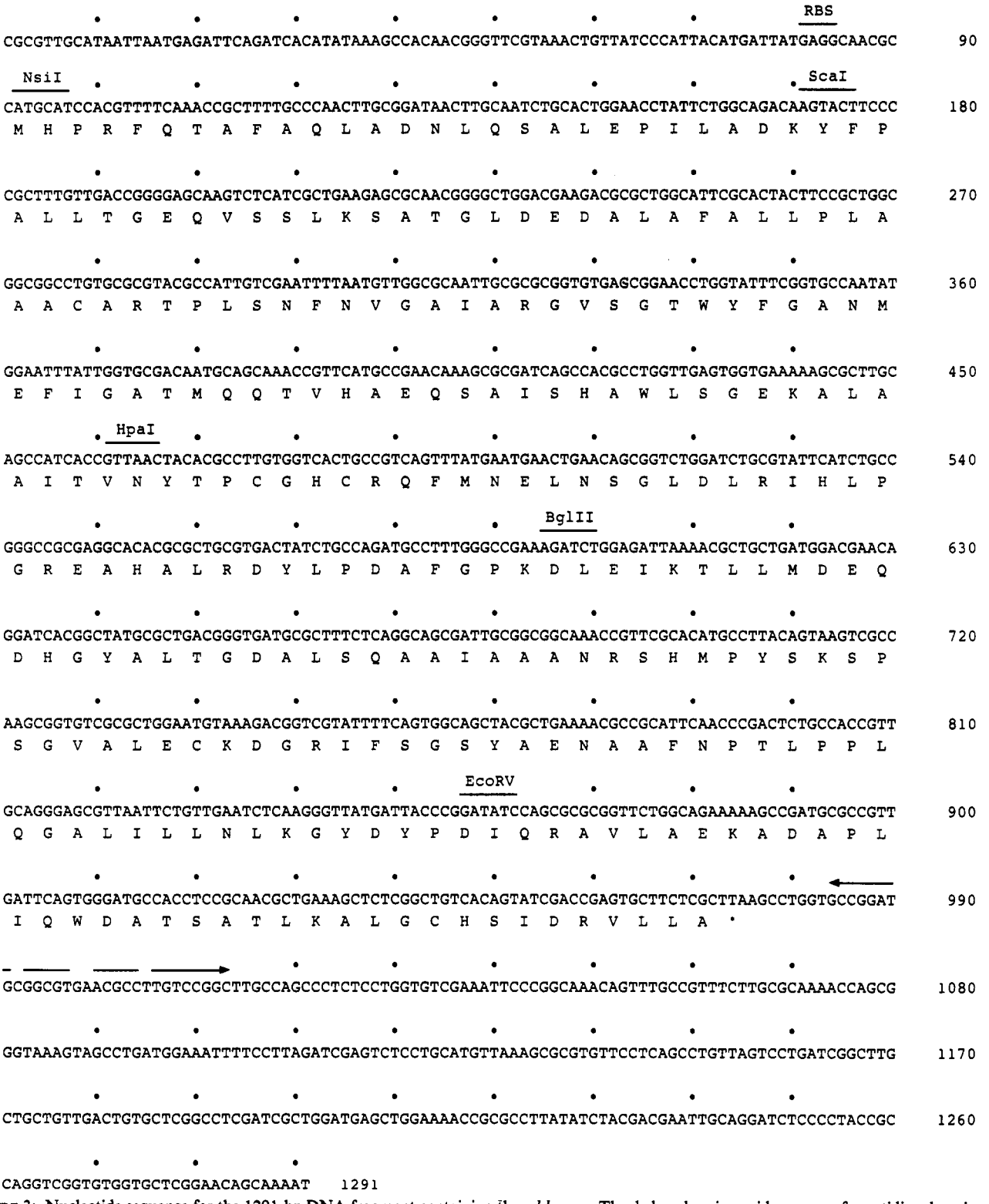

FIGURE 3: Nucleotide sequence for the 1291-bp DNA fragment containing the *cdd* gene. The deduced amino acid sequence for cytidine deaminase is shown. The sequence corresponding to a potential ribosome binding region (RBS) is indicated. The arrows 3′ to the cytidine deaminase termination codon indicate an inverted repeat that could function as a ρ-dependent transcription terminator. The sequence CGCGT corresponds to the last five bases of the *Mlu*I site shown in Figure 1 which were retained during subcloning of the *cdd* gene for sequence determination.

binding sites identified in the DNA sequence of the wild-type *cdd* promoter region. It remains to be determined whether the reduction in enzyme expression from plasmids pCY5 and pCY10 results from a decrease in the efficiency of mRNA synthesis from a truncated and weakened *cdd* promoter, which may be now only marginally CAP-activated or from mRNA initiated at a plasmid-borne promoter, such as pBR-P4 (Queen & Rosenberg, 1981), following inactivation of the *cdd* promoter during construction of plasmids pCY5 and pCY10. In either case, the level of functional *cdd* mRNA would be reduced relative to mRNA synthesized from the wild-type *cdd* promoter.

Comparison of the amino acid sequence of bacterial cytidine deaminase with the sequences for deoxy-CMP deaminases of yeast (McIntosh & Haynes, 1986) and bacteriophages T2 and T4 (Maley et al., 1983, 1990) and with human, mouse, and

E. coli cytidine deaminase	qTVHAE.qsAl sagekalaal ..tvnYt... PCghC
E. coli adenosine deaminase	lTVHAg.eaAg pesiwqaire .....
human adenosine deaminase	rTVHAg.evgs pevvReavdl .....
mouse adenosine deaminase	rTVHAg.evgs pevvReavdl .....
S.cerevisiae dCMP deaminase	lclHAEen.Al leagRdrvgq n.atlYcdtc PClLC
bacteriophage T2 dCMP deaminase	e.iHAEln.Al lfaaRngssl egatmYvtls PCpdC
bacteriophage T4 dCMP deaminase	e.iHAEln.Al lfaaRngssl egatmYvtls PCpdC
SEQUENCE SIMILARITY	-TVHAE---A- ----R----l -----Y---- PC--C

FIGURE 4: Regional amino acid sequence similarity between *E. coli* cytidine deaminase and other purine and pyrimidine nucleoside/nucleotide deaminases. Line 1: cytidine deaminase amino acid sequence beginning at residue 99. Line 2: *E. coli* adenosine deaminase beginning at residue 193 (Chang et al., 1991). Line 3: human adenosine deaminase beginning at residue 211 (Wiginton et al., 1984). Line 4: mouse adenosine deaminase beginning at residue 211 (Yeung et al., 1985). Line 5: deoxycytidylate deaminase from *Saccharomyces cerevisiae* beginning at residue 230 (McIntosh & Haynes, 1986). Lines 6 and 7 show bacteriophage T2 (line 6) and T4 (line 7) deoxycytidylate deaminase, each beginning at amino acid residue 102 (Maley et al., 1983, 1990). In line 8, which represents the sequence similarity, amino acids found in 4 of the 7 protein sequences are shown.

bacterial adenosine deaminases (Wiginton et al., 1984; Yeung et al., 1985; Chang et al., 1991) identified only a single region of amino acid sequence similarity in each deaminase (Figure 4). The region of amino acid similarity for each protein was used to construct a derivative profile, which was then compared with all protein sequences in the NBRF data base using the program PROFILESEARCH as described by Gribskov et al. (1990). This search uncovered only the proteins composing the original profile, and iterative reconstruction, in which each of these enzymes was omitted in turn, always found the enzyme that had been omitted intentionally.

A short but striking region of sequence identity, TVHA, was observed between bacterial cytidine deaminase and, bacterial and mammalian, adenosine deaminases. These enzymes resemble each other in the general mechanisms by which they appear to bring about substrate hydrolysis (as indicated by earlier studies involving inhibitors resembling intermediates in direct water attack) and also in the magnitudes of the large numerical contributions made by single hydroxyl groups to the free energies of binding of 3,4-dihydrouridine by cytidine deaminase (Frick et al., 1989) and 6-hydroxy-1,6-dihydropurine ribonucleoside by adenosine deaminase (Kati & Wolfenden, 1989a). In addition, both adenosine deaminase (Wilson et al., 1991) and cytidine deaminase contain an essential zinc atom at the active site. In view of these functional similarities, it seems hardly surprising that cytidine and adenosine deaminases should share a region (TVHA) of sequence identity, at positions 100–103 in cytidine deaminase and at positions 212–215 and 194–197 in mammalian and bacterial adenosine deaminase, respectively. This span of four contiguous amino acids is their only region of amino acid sequence identity, rendering it probable that this common sequence plays a role, direct or indirect, in catalyzing the deamination of both cytosine and adenine derivatives. In the crystal structure of adenosine deaminase complexed with 6-hydroxy-1,6-dihydropurine ribonucleoside, the sequence TVHAGE (residues 212–217) contains a histidine coordinated to zinc and a glutamic acid that interacts with N-1 of the purine ring (Wilson et al., 1991). In the crystal structure of cytidine deaminase complexed with 5-fluoropyrimidin-2-one ribonucleoside (L. Betts and C. W. Carter, Jr., personal communication), the sequence TVHAE (residues 100–104) contains a histidine that is coordinated to zinc.

Aside from this short region of similarity, the amino acid sequences of cytidine and adenosine deaminases are strikingly different. Inspection of the cytidine deaminase amino acid sequence suggests that it is unlikely that the active site zinc atom is coordinated by three histidine ligands as in adenosine deaminase. Instead, the most probable zinc ligands appear to include, in addition to His-102, the cysteine residues at positions 129 and 132. Separated by two residues, these cysteines might form a reverse β-turn that could serve as part of a high-affinity site for chelation of zinc (Arnold & Haymore, 1991) and could also serve as targets for 5-chloromercuricytidine, an irreversible inhibitor described by Ashley and Bartlett (1984). The amino acid sequence homologies shared by *E. coli* cytidine deaminase and the dCMP deaminases also suggest, by inference, that these bacteriophage and yeast nucleotide deaminases are zinc metalloproteins and that the histidyl and cystidyl residues identified in Figure 4 are involved in metal coordination. Answers to questions regarding the function of specific amino acids await solution of the cytidine deaminase crystal structure at high resolution. The functional significances of the TVHAE sequence, of the Cys-X_2-Cys sequence, and of other amino acid residues that may participate in the catalytic process are now being explored by site-directed and proofreading mutagenesis experiments.

ACKNOWLEDGMENTS

We thank M. L. Powers-Hemphill for her help in the production of the antibody directed against cytidine deaminase and B. Merrill and W. Chestnut for performing the analytical amino acid analysis of cytidine deaminase. We also thank L. Betts and C. W. Carter, Jr. for sharing information about the crystal structure of cytidine deaminase in advance of publication. We thank Dr. J. W. Olesik and S. E. Hobbs for performing elemental analysis by inductively coupled plasma optical emission spectroscopy.

REFERENCES

Aota, S., Gojobori, T., Ishibashi, F., Maruyama, T., & Ikemura, T. (1988) *Nucleic Acids Res. 16* (Suppl.), r315–r402.

Arnold, F. H., & Haymore, B. L. (1991) *Science 252*, 1796–1797.

Ashley, G. W., & Bartlett, P. A. (1984) *J. Biol. Chem. 259*, 13615–13620.

Barbier, C. S., & Short, S. A. (1985) *Gene 36*, 37–44.

Betts, L., Frick, L., Wolfenden, R., & Carter, C. W., Jr. (1989) *J. Biol. Chem. 264*, 6737–6740.

Chang, Z. Y., Nygaard, P., Chinault, A. C., & Kellems, R. E. (1991) *Biochemistry 30*, 2273–2280.

Cohen, R. M., & Wolfenden, R. (1971) *J. Biol. Chem. 246*, 7561–7565.

Dale, R. M., McClure, B. A., & Houchins, J. P. (1985) *Plasmid 13*, 31–40.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Frick, L., MacNeela, J. P., & Wolfenden, R. (1987) *Bioorg. Chem. 15*, 100-108.
Frick, L., Yang, C., Marquez, V. E., & Wolfenden, R. (1989) *Biochemistry 28*, 9423-9430.
Gribskov, M., Luthy, R., & Eisenberg, D. (1990) *Methods Enzymol. 183*, 146-159.
Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., & Blanquet, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8247-8251.
Josephsen, J., Hammer Jespersen, K., & Hansen, T. D. (1983) *J. Bacteriol. 154*, 72-75.
Kati, W. M., & Wolfenden, R. (1989a) *Biochemistry 28*, 7919-7927.
Kati, W. M., & Wolfenden, R. (1989b) *Science 243*, 1591-1593.
Kohara, Y., Akiyama, K., & Isono, K. (1987) *Cell 50*, 495-508.
Lowry, O. H., Rosebrough, A. L., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265-275.
Maley, G. F., Guarino, D. U., & Maley, F. (1983) *J. Biol. Chem. 258*, 8290-8297.
Maley, G. F., Duceman, B. W., Wang, A. M., Martinez, J., & Maley, F. (1990) *J. Biol. Chem. 265*, 47-51.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular cloning, a laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035-10038.
McIntosh, E. M. & Haynes, R. H. (1986) *Mol. Cell Biol. 6*, 1711-1721.
Miller, J. H. (1972) in *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Munch-Petersen, A., Nygaard, P., Hammer-Jespersen, K., & Fiil, N. (1972) *Eur. J. Biochem. 27*, 208-215.
O'Farrell, P. H., Kutter, E., & Nakanishi, M. (1980) *Mol. Gen. Genet. 179*, 421-435.
Queen, C., & Rosenberg, M. (1981) *Nucleic Acids Res. 9*, 3365-3377.
Rudd, K. E., Miller, W., Ostell, J., & Benson, D. A. (1990) *Nucleic Acids Res. 18*, 313-321.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) in *Molecular cloning, a laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342-1346.
Short, S. A., & Singer, J. T. (1984) *Gene 31*, 205-211.
Thomas, J. O., & Kornberg, R. R. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2626-2630.
Valentin-Hansen, P., Holst, B., Josephsen, J., Hammer, K., & Albrechtsen, B. (1989) *Mol. Microbiol. 3*, 1385-1390.
Vita, A., Amici, A., Cacciamani, T., Lanciotti, M., & Magni, G. (1985) *Biochemistry 24*, 6020-6024.
Vogel, H. J., & Bonner, D. M. (1956) *J. Biol. Chem. 218*, 97-106.
Wiginton, D. A., Adrian, G. S., & Hutton, J. J. (1984) *Nucleic Acids* 4Res. *12*, 2439-2446.
Wilson, D. K., Rudolph, F. B., & Quiocho, F. A. (1991) *Science 252*, 1278-1284.
Yeung, C. Y., Ingolia, D. E., Roth, D. B., Shoemaker, C., Al Ubaidi, M. R., Yen, J. Y., Ching, C., Bobonis, C., Kaufman, R. J., & Kellems, R. E. (1985) *J. Biol. Chem. 260*, 10299-10307.
Zhang, X. P., & Ebright, R. H. (1990) *J. Biol. Chem. 265*, 12400-12403.

An Inquiry into the Source of Stereospecificity of Lactate Dehydrogenase Using Substrate Analogues and Molecular Modeling†

Richard D. LaReau‡ and Vernon E. Anderson*
Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912
Received September 5, 1991; Revised Manuscript Received December 31, 1991

ABSTRACT: Lactate dehydrogenase catalyzes the stereospecific hydride transfer to and from the *re* face of the nicotinamide coenzyme. The demonstrated probability of transfer to the *si* face of less than 2×10^{-8} indicates that the free energy of any diastereotopic transition state leading to a *si* transfer must be over 10 kcal/mol greater than the free energy for transfer to or from the *re* face. The general notion of closed, desolvated active sites suggests the a priori hypothesis that steric hindrance prevents the nicotinamide ring from assuming a conformation that would lead to transfer of the *pro-S* hydrogen. In this paper we report that the probability of transfer of the *pro-S* proton is less than 9×10^{-7} with 3-pyridinealdehyde adenine dinucleotide as coenzyme and less than 4×10^{-7} during the lactate dehydrogenase catalyzed disproportionation of glyoxylate. Examination of the crystal structure of lactate dehydrogenase further suggests that steric exclusion does not enforce the extreme stereospecificity of the reaction. An electrostatic interaction with the macrodipole associated with the α2F helix is suggested as a potential molecular source of the stereospecificity.

The pyridine nucleotide dependent dehydrogenases catalyze the reversible hydride transfer to and from the C-4 carbon of the nicotinamide moiety of $NAD(P)^+$. It has been well established that this reaction is stereospecific, such that a given dehydrogenase transfers either the *pro*-4*R* (A-side) or *pro*-4*S* (B-side) hydrogen of the coenzyme (Fisher et al., 1953; Loewus et al., 1953; Jarabak & Talalay, 1960; Cornforth & Ryback, 1962; Arnold et al., 1976; You, 1985). In our original experiments, the reaction

† This work was supported by Grant GM 36562 from the National Institute of General Medical Sciences. V.E.A. is a Sloan Foundation Fellow.
‡ Present address: Department of Biochemistry, University of Rochester, Rochester, NY 14642.

0006-2960/92/0431-4174$03.00/0