Random mutagenesis and selection of *Escherichia coli* cytosine deaminase for cancer gene therapy

Sheri D.Mahan\(^1\), Greg C.Ireton\(^2\), Catherine Knoeber\(^1\), Barry L.Stoddard\(^2\) and Margaret E.Black\(^1,3\)

\(^1\)Department of Pharmaceutical Sciences and the School of Molecular Biosciences, PO Box 646534, Washington State University, Pullman, WA 99164-6534 and \(^2\)Fred Hutchinson Cancer Research Center, and the Graduate Program in Molecular and Cell Biology, University of Washington, 1100 Fairview Avenue North A3-023, Seattle, WA 98109, USA

\(^3\)To whom correspondence should be addressed. E-mail: blackm@mail.wsu.edu

**Cytosine deaminase (CD)** is currently being used as a suicide gene for cancer gene therapy. The premise of this therapy is the preferential deamination of 5-fluorocytosine (5FC) to 5-fluorouracil by cancer cells expressing cytosine deaminase. However, a lack of efficient gene transfer to tumors combined with inefficient 5FC turnover currently limits the clinical applications of this gene therapy approach. We have used random mutagenesis to create novel bacterial cytosine deaminases that demonstrate an increased preference for 5FC over cytosine. Among the 15 mutants isolated, one conferred sensitivity to *Escherichia coli* in a negative selection system at a concentration of 5FC that was 10-fold lower than a sublethral dose for wild-type CD. Evaluation of individual substitutions found in this double mutant (Q102R, D314G) demonstrated that the substitution at residue D314 was solely responsible for the observed increase in sensitivity to 5FC. Additional mutagenesis at D314 resulted in the identification of two more substitutions with the ability to confer enhanced 5FC sensitivity to *E.coli*. Structure determinations of the three CD variants in the presence and absence of a transition state 5FC analogue provide insights to the determinants of substrate binding specificity at the 5\(^{\prime}\) position of the pyrimidine ring. CD mutant D314A is a promising candidate for further gene therapy studies.

**Keywords**: cytosine deaminase/gene therapy/5-fluorocytosine/mutagenesis

**Introduction**

Cytosine deaminase (CD; EC 3.5.4.1) plays an important role in pyrimidine salvage by catalyzing the deamination of cytosine to uracil and ammonia. The cytosine deaminase enzyme from *Escherichia coli* (bCD) is a hexamer with an approximate molecular mass for the native form of around 300 kDa and is dependent upon Fe\(^{2+}\) for maximal activity. While bacteria and yeast encode cytosine deaminase, mammalian cells do not (Nishiyama et al., 1985; Katsuragi et al., 1987). This feature and the ability of cytosine deaminase to deaminate 5-fluorocytosine (5FC) and produce the potent antimetabolite, 5-fluorouracil (5FU), make cytosine deaminase particularly attractive for use in gene therapy (Austin and Huber, 1992; Mullen et al., 1992; Huber et al., 1993; Hirschowitz et al., 1995; Kievit et al., 1999). In suicide gene therapy applications, the gene for cytosine deaminase is delivered to cancer cells and the individual is subsequently treated systemically with the non-toxic 5FC prodrug. Once deaminated by CD, the analogue is activated to 5-FdUMP by cellular enzymes and acts as an irreversible inhibitor of thymidylate synthase, thereby preventing DNA replication by blocking dTTP synthesis. The cytosine deaminase from *E.coli* (bCD) has been used successfully in gene therapy for a variety of animal tumor models and is currently under investigation for the treatment of human cancers (Mullen et al., 1992; Ohwada et al., 1996; Cunningham and Nemunaitis, 2001; Freytag et al., 2002).

In addition to cytosine deaminase/5-fluorouracil combined therapy, a number of additional enzyme prodrug combinations are under intense scrutiny for suicide gene therapy (Greco and Dachs, 2001). Perhaps the most advanced and widely investigated of these systems is the combination of Herpes Simplex Virus type 1 (HSV) thymidine kinase (TK) and ganciclovir, in which the prodrug is phosphorylated within the transfected cell by TK and subsequently converted into a triphosphorylated nucleoside analogue (Miller and Miller, 1980; Moolten, 1986). Although transfection efficiencies for this and other prodrug conversion enzymes are low, and the product molecules cannot cross cell membranes (relying instead on gap junctions and/or active transport between neighboring cells), complete tumor regression is often observed in a variety of experimental systems (Moolten, 1986; Culver et al., 1992; Bi et al., 1993; Freeman et al., 1993; Colombo et al., 1995; Dilber et al., 1997; Kuriyama et al., 1998). Now known as the bystander effect, this general toxicity of suicide gene therapy to non-transfected neighboring and distal tumor cells is a key component of efficacy.

The CD/5FC system also relies on the bystander effect for tumor ablation; however this system displays the advantage that the prodrug product is freely diffusible across cell membranes, allowing localized general toxicity to be independent of gap junctions or active transport in the tumor bed (Domin et al., 1993; Kuriyama et al., 1998). Additionally, CD/5FC therapy is amenable to delivery using antibody directed enzyme prodrug therapy (ADEPT) (Senter et al., 1991). In situations where HSVTK/GCV combined therapy is not successful in eradicating the tumor, or where an immune response to HSVTK would preclude its use in a second round of gene therapy, the CD/5FC system offers a viable alternative, and may in fact be the preferred first gene therapy course of treatment.

The limiting factors for successful anti-tumor gene therapy are transfection efficiency and the ability of the enzyme to turn over the prodrug, which is an analogue of its natural substrate. From a kinetic perspective, 5FC is a poor substrate for bCD (K\(_m\) = 3.5 mM) compared to its native substrate, cytosine (K\(_m\) = 0.2 mM). Recent studies suggest the CD from *Saccharomyces cerevisiae* (yCD) displays a kinetic advantage towards 5FC over bCD (Kievit et al., 1999). However, yCD is considerably less thermostable than bCD, a characteristic that may make...
the bacterial enzyme a preferable catalyst system for gene therapy.

The crystal structure bCD has been determined in the presence and absence of a mechanism-based inhibitor. The bCD enzyme fold corresponds to an 8-stranded α/β barrel, commonly termed a ‘TIM’ fold (Ireton et al., 2002). The enzyme forms an enzymatically active hexamer, in which the N- and C-terminal regions (residues 1–52 and 363–407) form an additional domain-swapped fold that is critical for packing and stability of the enzyme hexamer. Despite its complex oligomeric structure, the enzyme does not display kinetic cooperativity and all six enzyme subunits are catalytically active and independent of one another.

We sought to create novel, drug-specific bCD variants as a means to enhance tumor cell killing without increased toxicity. In the present study we have used molecular evolution to introduce mutations randomly throughout the entire reading frame to create thousands of bCD variants. We have established positive and negative genetic complementation in E.coli to identify mutants that confer increased sensitivity to 5FC. Crystal structure determinations of three key bCD mutants provide novel insight to the molecular mechanism of bCD and may be important for future engineering of the enzyme.

Materials and methods

Materials

Restriction endonucleases and T4 DNA ligase were purchased from Gibco BRL (Rockville, MD) or New England Biolabs (Beverly, MA). Oligonucleotides used for polymerase chain reaction (PCR), site-directed mutagenesis and DNA sequencing were obtained from Operon (San Pablo, CA) or Genset (Boulder, CO). Nickel affinity chromatography agarose (Ni-NTA Agarose) used to purify mutant and wild-type bCDs was purchased from Qiagen (Valencia, CA). Enzyme assay reagents and other chemicals were purchased from Sigma (St Louis, MO) except where designated otherwise.

Bacterial strains

The cytosine deaminase deficient E.coli strain GIA38 strain (thr- dad B3 leuA21 thrA21 codA1 lacY1 tsx-95 glnV44(AS) λ- pyrF101 his-108 argG6 ilvA63 thi-1 deoC1 gli-15) was obtained from the E.coli Genetic Stock Center (CGSC #5594). Escherichia coli GIA39 was lysogenized with DE3 according to the instructions from Novagen (Madison, WI). The derived strain GIA39(DE3) was used in the genetic complementation assays for cytosine deaminase activity. Escherichia coli strain NM522 [F' lacAΔ(lacZ)-M15proAB-/supE thiA(lac-proAB)Δ(hsdMS-mcrB)S(rk mbr McrBC)] was used as a recipient for certain cloning procedures. Escherichia coli strain CI236 (F' lacIΔmcrB - ung-1, relA1, dut-1, spoT1, thi-1) was used to produce single stranded DNA for site directed mutagenesis procedures. Escherichia coli BL21(DE3) tdk- [F' ompT[tol] hsdSb (rB mB)] gal dcm met (DE3)] was used for expression of wild-type and mutant bCD proteins.

Vectors

Initially, the bCD gene was amplified from pCD2 (a gift from Mike Blaese) using oligos (MB204, 5’ GTTATTCGCTAGGCTAGCTGG 3’ and MB201, 5’ ACTAGCTCTGTAACCC 3’ ) that introduced an NcoI site at the start codon and a HindIII site just 3’ to the stop codon. The unc7 1.6 kb fragment was cloned into pCR2.1TOPO. The bCD gene was excised from pCR2.1TOPO:bCD as an NcoI/HindIII fragment and cloned into the bacterial expression vector, pETHT (Brady et al., 1996), restricted with the same enzymes. The resulting plasmid was designated pETHT:bCD.

Library construction

The template for the error prone PCR (epPCR) was pETHT:bCD. Two primers used to amplify the bCD gene hybridize to vector sequences flanking the open reading frame (T7 promoter primer, 5’ TAATACGACTCACTATAGGG 3’ and T7 terminator primer, 5’ GCTAGTTATGCTACGGG 3’). The 50 µl reaction mix contained the following reagents: 1× PCR buffer (Gibco BRL), 0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP, 1 mM dTTP, 7 mM MgCl2, 0.5 mM MnCl2, 5 U Taq polymerase, 0.5 pmol of T7 promoter primer, 0.5 pmol of T7 terminator primer and 0.16 ng of pETHT:bCD. Amplification was carried out in an Eppendorf thermocycler as follows: 94°C for 3 min followed by 30 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 40 s. At the end of the 30 cycles the reaction was incubated at 72°C for 3 min and then chilled to 4°C.

The resulting PCR fragments were restricted with Ncol and HindIII and ligated to pETHT digested with Ncol and HindIII. Ligation mixtures were used to transform GIA39(DE3) competent cells by electroporation. Following electroporation, a small fraction of cells was plated onto uracil-containing plates to determine the total number of transformants and the remainder of the transformation was plated onto cytosine-containing plates to select functional clones. The plates were incubated at 37°C for 36 h prior to scoring. Colonies that grew on cytosine plates were picked and streaked onto fresh cytosine plates to confirm the phenotype. Positive selection was based on the genetic complementation of a functional bCD mutant in the codA GIA39(DE3) strain. The CD selection medium contains 1.96 g of yeast synthetic dropout without uracil, 0.1 mM CaCl2, 0.24 mM (or 0.267 mg/ml) cytosine, 111 ml of M9 cocktail [100 ml of 10×M9 salts (30 g of KH2PO4, 67.8 g of Na2HPO4, 5 g of NaCl, 10 g of NH4Cl per liter), 1 ml of 1 M MgSO4, 5 ml of 20% glucose per liter] and 50 mg carbenicillin per liter. For uracil-containing medium (nonselective), cytosine was omitted and the yeast synthetic dropout minus uracil was replaced by 0.78 g yeast synthetic dropout minus leucine and supplemented with 2 ml of 2% leucine per liter. For plates, 17 g of Bactoagar per liter was included in the media described above.

Selection of 5FC-sensitive mutants

Functional clones were subjected to a secondary selection to determine 5FC sensitivity. Secondary selection media contained 5FC concentrations ranging from 2 to 20 µg/ml. The 5FC selection medium is the same as CD selection medium with the addition of 5FC.

Real-time PCR

Real-time PCR was used to establish the cloning frequency of the error prone bCD fragments. The SYBR Green PCR Core Reagent kit was used (PE Applied Biosystems, Foster City, CA) with oligonucleotides MB264, 5’ CGGCACACACACCAGTATATC 3’ and MB265, 5’ CAATCTAGCTTGTCCTTTCGGG 3’ as follows. For each 25 µl reaction, 2 mM MgCl2, 1× SYBR Green Buffer, 1 mM dNTPs, 0.25 pmol of MB264, 0.25 pmol of MB265, 0.625 U ATaq Gold Polymerase, 0.25 µM ERASE UNG, 3 µl of template and H2O to
25 µl were mixed. Eighty-seven colonies from uracil plates were individually resuspended in 250 µl of H2O and 3 µl used as template. In addition, colonies of pETHT and pETHT:bCD were resuspended in the same fashion for use as negative and positive controls, respectively. The starting temperature for the PCR was 60°C followed by 55 cycles of 55°C for 30 s and 72°C for 30 s. The PCR and analysis were performed using a PE Applied Biosystems GeneAmp 5700 Sequence Detection System Version 3.1.

**Sequencing**
Plasmid DNA was isolated from cultures of individual colonies using the Wizard mini-prep kit according to the manufacturer’s instructions (Promega, Madison, WI). Sequencing reactions were performed at the core sequencing facility at Washington State University.

**Site-directed mutagenesis**
Site-directed mutagenesis (Kunkel, 1985) was used to individually introduce the following substitutions: Q102R, D314G, D314H, D314R and D314K. The oligonucleotide MB286 (5’-CCATGGCAGGCTTTCACATC-3’) was used to introduce the Q102R mutations as well as to introduce a new BssHII restriction site. The following oligonucleotides were used to introduce the D314 codon changes as well as a silent mutation to introduce a MscI restriction site to facilitate identification of mutated sequences. The following oligonucleotides were used for site-directed mutagenesis: for D314G, MB287, 5’-GAAAGACACATCGTGCCAAAG-3’; D314K, MB288, 5’-GAAACATTTATCGTGCCAAAG-3’; D314R, MB289, GAAACACATCGTGCCAAAG-3’; D314H, MB290, 5’-GAAACATGATCGTGCCAAAG-3’; D314S, MB291, 5’-GAAACACTATCGTGCCAAAG-3’; and D314A, MB292, 5’-GAAACAGCATCGTGCCAAAG-3’. After transformation of E.coli NM522, plasmid DNA was isolated from individual clones, restricted with BssHII or MscI and the fragments separated by agarose gel electrophoresis. DNA sequencing was used to confirm the mutations. Plasmid DNA of each site-directed mutant was then used to transform E.coli GIA39(DE3). All seven mutants and the wild-type bCD construct (pETHT:bCD) were then tested for genetic complementation on cytosine-containing selection plates. Clones scored as positive for cytosine deaminase activity were streaked onto plates containing various concentrations of 5FC to evaluate the degree of prodrug sensitivity.

**Enzyme purification**
Expression of the wild-type and mutant cytosine deaminases was performed in BL21 (DE3) tdk− cells without induction as described previously (Mahan et al., 2004). Fractions containing a single band corresponding to the molecular mass of bCD were pooled and dialyzed against a large volume of dialysis buffer (50 mM NaCl and 50 mM Tris, pH 7.5) at 4°C. After dialysis, the samples were collected and stored at −20°C. The enzyme concentration of the dialyzed sample was determined using a spectrophotometer based assay (Pharmacia Biotech Ultrospec 2000). A stock of 10–20 mM cytosine was made in 50 mM Tris–HCl, pH 7.5 and diluted 1:1 with enzyme. OD280 readings were taken immediately upon mixing every 10 s for a total of 6.5 min (390 s). This was repeated using the same amount of enzyme for every cytosine concentration within the optimal range. Double reciprocal plots were used to determine Km values for each mutant and wild-type bCD enzyme. The turnover number (kcat = Vmax/[E], where [E] is the total enzyme amount) was also determined for each mutant and wild-type bCD. Assays were performed five to seven times for each substrate concentration.

**Enzyme assays with 5FC as the substrate**
Kinetic values for the wild-type and D314 mutants (G, A and S) were obtained using a protocol similar to that described above (Hayden et al., 1998) with 5FC substituted for cytosine as the substrate. A stock of 30 mM 5FC was made in 50 mM Tris–HCl, pH 7.5. A 1 ml reaction was containing the desired concentration of 5FC, enzyme and 50 mM Tris–HCl, pH 7.5 was placed at 37°C. Aliquots (50 µl) of the reaction were taken every 90 s over a 15 min time period and immediately quenched in 0.1 N HCl. Readings were then taken at OD290 (OD at which 5FC absorbs) and OD255 (OD at which 5FU absorbs). Assays were repeated five to eight times. Kinetic parameters were determined as above.

**Determination of mutant enzyme thermostability**
The stabilities of the wild-type enzyme and the three point mutants described in this work were measured by chemical denaturation using circular dichroism. Circular dichroism spectra were collected on an Aviv 62A DS spectrometer. Far UV circular dichroism wavelength scans (200–260 nm) at varying protein concentrations (5–20 µM) were collected in a 1 mm pathlength cuvette. Protein denaturation due to thermal melting was followed by change in ellipticity at 220 nm.

**Crystallography**
The expression, purification and crystallization of bCD and bCD variants were carried out as previously described (Ireton et al., 2002). X-ray diffraction data were collected at beamlines 5.0.1 and 5.0.2 at the ALS (Advanced Light Source, Lawrence Berkeley Laboratory, Berkeley, CA) using a four panel ADSC CCD area detector. Data were processed and scaled using the DENZO/SCALEPACK program packages (Otwinowski, 1993; Otwinowski and Minor, 1997). The structures were refined using CNS with a random 5% of the data excluded for the purpose of cross-validation (Brunger, 1993). Crystals containing the enzyme in complex with a mechanism based inhibitor [5-fluoro-4-(S)-hydroxyl-3,4-dihydropyrimidine or 5FDHP] were prepared by soaking in cryo-buffer containing 5–10 mM 5-fluoro-2-hydroxypyrimidine (Aldrich, Milwaukee, WI). Unbiased (Fobs 2FDHP − Fcalc apo)2σa maps were examined and used for rebuilding. All structures were refined using CNS, as described above for the uncomplexed enzymes.

**Results and discussion**

**Library construction and selection of mutants with increased 5FC sensitivity**
CD plays an essential role in activation of 5FC and the yeast and bacterial CD are both being explored for use as suicide
genes for cancer gene therapy. Despite the ability of these two enzymes to catalyze the same reaction, there is no identifiable primary sequence homology or conserved motifs that might reveal important regions (Ireton et al., 2003). At the time this study was initiated, the crystal structure of bCD was not yet available to model potential active site amino acid interactions with substrates. In the absence of a clear region to target, we used epPCR to introduce random nucleotide changes throughout the bCD coding region in combination with both positive and negative selections to create and identify novel bCD mutants that enhance 5FC sensitivity to cells (Figure 1).

In our mutant bCD library, we estimate the number of transformants containing a bCD gene at around $1.8 \times 10^4$. Approximately 1400 transformants or 7.7% grew on the cytosine selection plates. Sequence analysis of bCD genes from non-selected transformants recovered from uracil plates revealed an overall mutation frequency of 0.73%. This is similar to frequencies from error prone mutagenesis studies reported by others (Cadwell and Joyce, 1992; Stemmer, 1994; Kawate et al., 2002).

All 1400 CD-positive clones were picked and streaked onto plates containing 5FC at a concentration determined to be sublethal to wild-type bCD (20 µg/ml 5FC). Starting at 20 µg/ml 5FC, where 70 clones or 0.05% were more sensitive than wild-type bCD, the concentration was progressively reduced until a single remaining mutant, #550 was determined to be most sensitive to the lowest experimental concentration of 5FC (2 µg/ml). Fourteen additional clones displayed sensitivity at a slightly higher concentration of 5FC (10 µg/ml) and were also included for subsequent analyses.

Plasmid DNA from mutant 550 and other mutants with sensitivity to 10 µg/ml 5FC was isolated and the entire bCD open reading frame sequenced. Sequence results show that the 5FC-sensitive clones contain at least two and up to eight amino acid substitutions. The spectrum of amino acid substitutions that lead to increased 5FC sensitivity show a uniform distribution of mutations throughout the open reading frame as shown in Figure 2. One mutant contained four point mutations and a mutation at the stop codon that resulted in the fusion of an additional 23 amino acids at the C-terminus. Four amino acid positions were observed to contain two different amino acid substitutions, e.g. T42 was substituted by A and S in two independently isolated clones. The same substitution was found independently twice for two different positions (T248A and Q339L). Of particular note is the occurrence of D314G in three of the mutant sequences, including mutant 550 (Figure 2). The mutation at codon D314 in all three clones occurs at nucleotide 944 and reflects an A to G transition. Because it is the only nucleotide alteration in common between mutant 550 (three changes), 1456 (three changes) and 1491 (eight changes), it is possible that the mutation at nucleotide 944 occurred early during the PCR cycle and all other mutations observed in these clones occurred after the initial mutational event.

Because mutant 550 displayed the greatest sensitivity to 5FC we assessed the role of its individual amino acid substitutions (Q102R and D314G) using site-directed mutagenesis. Both Q102R and D314G were found to display active cytosine deaminase activity by complementation but only the D314G substitution exhibited sensitivity to 5FC at 2 µg/ml. Examination of the bCD molecular structure indicates that residue Q102 is distal to the active site with the side chain extending towards the periphery of the subunit. In contrast, D314 is located within the active site with the side chain directed towards the fluorine atom of 5FC.

To further investigate the role of the aspartic acid at position 314, several other substitutions were made at that position: D314K, D314R, D314H, D314S and D314A. The basic side chain substitutions (K, R and H) yielded inactive enzymes as assessed by complementation. However, D314S and D314A were both functional towards cytosine and conferred sensitivity to 5FC at 4 and 2.5 µg/ml, respectively. In a separate study, alanine scanning mutagenesis was used to substitute residues 310–320 with alanine. Surprisingly, only D314A was shown to confer increased 5FC sensitivity to GIA39 (DE3) cells (Mahan et al., 2004).

**Enzyme behavior and kinetics**

Purification of wild-type bCD, D314G, D314A and D314S enzymes for kinetic analysis was achieved using batch nickel affinity chromatography. All enzymes displayed non-cooperative, saturable Michaelis–Menten profiles. The kinetic parameters determined for the wild-type bCD and the three D314 mutants are shown in Table I. The kinetic values

---

**Fig. 1.** Outline of library construction and PCR gel. Lanes 1 and 8 are molecular weight markers. Lanes 2–6 show amplified PCR product at ~1.6 kb, lane 7 is a no template control PCR.
determined for the wild-type bCD correspond to values previously reported in the literature by Porter (2000). All three enzyme point mutants displayed thermal melting points (Tm) that were unchanged from the wild-type enzyme (85 ± 2°C).

With respect to turnover of cytosine, the D314A mutant is most strongly affected, with a kcat/Km value that is 6% of the wild-type enzyme. This reduction in activity corresponds to a significant reduction in affinity (9% of wild-type) and a more modest reduction in turnover rate, as shown in Table I. The other two point mutations (D314G and D314S) also display impaired catalytic efficiencies, with kcat/Km values for cytosine that are 15 and 34% of wild-type, respectively. For the D314S mutant, the reduction in catalytic efficiency is primarily caused by reduced substrate binding affinity, while for D314G, the effect is predominantly due to a reduced turnover rate.

Fluorocytosine is a relatively poor substrate for bCD, primarily because it binds the wild-type enzyme with a Km increased by 16.5-fold relative to cytosine (Table I). This effect is presumed to be caused by electrostatic repulsion of the fluorine atom by the side chain carboxylate oxygens of D314. Evaluation of the 5FC kinetic profiles for point mutants at residue 314 indicates that the mutations have minor effects on the binding of 5FC (ranging from a 15% reduction in Km for D314A to a 3-fold increase for D314G) and consistent increases in kcat (as much as 3-fold higher for D314S). As a result, the catalytic efficiency of the mutants towards 5FC is slightly decreased for one mutant (D314G) but improved for the remaining two mutants (D314S and D314A). Taken together, the 5FC kinetics, with a kcat/Km value twice that of wild-type bCD indicates that D314A displays the most significant shift in specificity towards 5FC.

From a comparison of wild-type and D314A kinetics for 5FC, one might not predict that D314A would provide a significantly enhanced prodrug sensitivity to cells. However, because endogenous cytosine within the cell competes with 5FC for the active site, it is important to consider the ratio
of specificity constants for the prodrug and cytosine. We used the equation \( \frac{[k_{cat}/K_m \text{ (5FC)}]}{[k_{cat}/K_m \text{ (cytosine)}]} + \frac{[k_{cat}/K_m \text{ (5FC)}]}{[k_{cat}/K_m \text{ (cytosine)}]} \) to take this into account (Table I). D314G and D314S display moderate shifts in their relative specificities when they are compared to the wild-type bCD at 2.5- and 4.2-fold, respectively. In comparison, the D314A mutant is shifted 19-fold in favor of 5FC.

**Structural studies**

To determine the effect of each amino acid substitution on the bCD structure and substrate binding, each mutant (D314G, D314A and D314S) was crystallized in the absence and presence of a fluorinated mechanism-based inhibitor, 5-fluoro-2-hydroxypyrimidine (5FDHP). The six structures (D314G, D314A, D314S, with and without 5FDHP) were all refined between 1.6 and 1.1 Å resolution. Crystallographic statistics are listed in Table II.

Previously, we had shown that bCD is a hexameric protein of 300 kDa molecular mass with a single catalytic iron coordinated by three histidine residues and residue D313 in each monomeric unit (Ireton et al., 2002). Residue 314 lies on a flexible loop structure that undergoes a conformational shift upon substrate binding so that several residues contact the pyrimidine ring. Substitution of D314 does not cause any overall change in the bCD fold but appears to cause small structural alterations of the enzyme active site (Figure 3). In the absence of bound compound, there is a small shift in D313. In the wild-type enzyme, this residue is a direct iron (II) ligand, with Oδ2–Fe2+ distance to 4.0 Å. In all unbound mutant structures, the catalytic metal assumes an irregular tetrahedron coordinated by Oδ2 of H61 (2.1 Å), H63 (2.1 Å), H214 (2.2 Å) and a bound water molecule (1.9 Å).

**Mutants complexed with 5FDHP**

Unbiased Fourier difference maps calculated with diffraction data from bCD mutant crystals soaked with the mechanism-based inhibitor 5-fluoro-2-hydroxypyrimidine display clear electron density for the bound compound adjacent to the active site metal ion (Figure 4). Occupancy refinement on the bound inhibitor leads to a values of \( \sim70-90\% \) for the different mutant data sets.

When the inhibitor is bound, all of the mutant active site structures and coordination geometries return to much closer structural agreement and superposition with the wild-type enzyme bound to an analogous, unfluorinated inhibitor dihydropyrimidine (DHP). As observed previously for wild-type bCD, the D314 mutants in complex with the fluorinated inhibitor display a significant conformational change relative to the apo-enzyme, burying over 750 Å² of accessible surface within the binding cleft and increasing the hydrophobic environment around the bound pyrimidine ring. The conformational change involves a peptide lid (residues 68–96) that borders the entrance to the barrel pore and the active site. The α-helix in this region moves approximately 5 Å inward to close off the mouth of the barrel in the presence of bound compound. The conformational change is induced by direct contacts between protein side chains and the bound pyrimidine ring, including the movement of L81 and I85 ~5 Å, and W319 ~3 Å, to create a hydrophobic pocket below the bound compound that blocks access to solvent.

### Table II. Crystallographic statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>D314G</th>
<th>D314G-5FDHP</th>
<th>D314A</th>
<th>D314A-5FDHP</th>
<th>D314S</th>
<th>D314S-5FDHP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit cell dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a = b )</td>
<td>109.04</td>
<td>109.26</td>
<td>109.57</td>
<td>109.34</td>
<td>109.79</td>
<td>109.48</td>
</tr>
<tr>
<td>( c )</td>
<td>240.19</td>
<td>240.84</td>
<td>204.09</td>
<td>240.64</td>
<td>240.70</td>
<td>241.09</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20–1.58(1.63–1.58)</td>
<td>20–1.10(1.14–1.10)</td>
<td>50–1.55(1.60–1.55)</td>
<td>50–1.35(1.40–1.35)</td>
<td>20–1.32(1.36–1.32)</td>
<td>20–1.32(1.36–1.32)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>76,994</td>
<td>211,957</td>
<td>80,401</td>
<td>117,634</td>
<td>130,224</td>
<td>126,518</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.3(98.5)</td>
<td>95.3(67.1)</td>
<td>100(100)</td>
<td>100(100)</td>
<td>94.4(89.3)</td>
<td>97.6(95.7)</td>
</tr>
<tr>
<td>Rmerge (Å)</td>
<td>0.068(0.338)</td>
<td>0.041(0.223)</td>
<td>0.075(0.319)</td>
<td>0.098(0.395)</td>
<td>0.060(0.254)</td>
<td>0.076(0.339)</td>
</tr>
<tr>
<td>Average I/σ(I)</td>
<td>18.8(3.9)</td>
<td>36.3(3.0)</td>
<td>26.4(5.3)</td>
<td>20.4(3.5)</td>
<td>22.9(3.9)</td>
<td>18.8(3.5)</td>
</tr>
<tr>
<td>Refinement statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20–1.58</td>
<td>20–1.12</td>
<td>20–1.57</td>
<td>20–1.40</td>
<td>20–1.32</td>
<td>20–1.32</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>72,793</td>
<td>198,719</td>
<td>74,561</td>
<td>107,622</td>
<td>125,729</td>
<td>124,047</td>
</tr>
<tr>
<td>Test set (%)</td>
<td>3682</td>
<td>9889</td>
<td>3806</td>
<td>5367</td>
<td>6311</td>
<td>6221</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>18.37</td>
<td>17.16</td>
<td>18.53</td>
<td>18.59</td>
<td>17.34</td>
<td>18.02</td>
</tr>
<tr>
<td>r.m.s.d.</td>
<td>0.0055</td>
<td>0.0072</td>
<td>0.0064</td>
<td>0.0068</td>
<td>0.0075</td>
<td>0.0075</td>
</tr>
<tr>
<td>Angles (Å)</td>
<td>1.336</td>
<td>1.414</td>
<td>1.356</td>
<td>1.411</td>
<td>1.399</td>
<td>1.413</td>
</tr>
<tr>
<td>Protein atoms</td>
<td>3321</td>
<td>3321</td>
<td>3321</td>
<td>3321</td>
<td>3321</td>
<td>3321</td>
</tr>
<tr>
<td>Waters</td>
<td>472</td>
<td>466</td>
<td>473</td>
<td>432</td>
<td>476</td>
<td>434</td>
</tr>
<tr>
<td>Ligand atoms</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Average B (Å²)</td>
<td>12.69</td>
<td>8.65</td>
<td>13.02</td>
<td>14.51</td>
<td>11.25</td>
<td>10.56</td>
</tr>
<tr>
<td>Protein</td>
<td>28.58</td>
<td>22.74</td>
<td>27.79</td>
<td>27.2</td>
<td>25.24</td>
<td>24.54</td>
</tr>
<tr>
<td>Waters</td>
<td>10.70</td>
<td>14.63</td>
<td>11.02</td>
<td>11.02</td>
<td>11.02</td>
<td>11.02</td>
</tr>
</tbody>
</table>
contacts are made to the bound compound by Oδ1 and Ne2 of Q156, that make a pair of bifurcated contacts to its N1 and O2 atoms. In addition, the N3 nitrogen of cytosine contacts the e2 side chain oxygen of E217. Finally, the compound is also directly contacted by several polar residues that, in combination with the coordination geometry of the catalytic Fe²⁺ ion, determine the stereochemistry of the reaction. The hydroxyl group at position 4 of the pyrimidine ring, an analog of the nucleophilic hydroxyl in the reaction transition state, is directly coordinated to the catalytic metal ion (2.1 Å) and is within contact distance of H246 Ne2. In the mutant structures, the side chain of residue 314 is pointing directly at the 5' fluorine on the pyrimidine ring. The distance of the terminal side chain atom of residue 314 to the fluorine are as follows: D314G backbone alpha carbon, 5.1 Å; D314A methyl carbon, 4.2 Å; and D314S hydroxyl oxygen, 3.0 Å. Previous attempts to co-crystallize or soak the fluorinated compound into wild-type bCD were unsuccessful. When 5FDHP is modeled into the wild-type bCD active site, the distance between the fluorine and the D314 terminal oxygen is only 2.3 Å.

**Effects of mutations at residue 314 on binding and deamination of cytosine and 5FC**

A comparison of high-resolution crystallographic analyses of wild-type enzyme bound to a DHP (a transition state analogue of cytosine) and of the mutants bound to 5FDHP (a corresponding 5' fluorinated version of the same compound) indicates that binding interactions of these compounds to the active sites are similar with the exception of the distance from the 5' position of the pyrimidine ring to residue 314 and the chemical structure of the side chain at that position. The primary effect of the mutations that cause increased cellular sensitivity towards 5FC is not simply a significant improvement in the affinity of the prodrug, but rather a mixture of effects on relative substrate binding affinities, transition-state contacts and the resulting turnover rate for cytosine and 5FC. While two of the mutants (D314G, D314S) display decreased affinity towards 5FC (3.3- and 2-fold higher $K_m$, respectively), this effect is less than the corresponding decrease in affinity towards cytosine. In contrast, the point mutant with the most significant shift in specificity in favor of 5FC
(D314A), displays a slightly improved affinity towards 5FC and at the same time also displays the largest decrease in affinity towards cytosine (11-fold higher $K_m$). This point mutant is the only one with roughly equivalent $K_m$ values towards the two substrates.

The observed trends in 5FC binding affinity for the three mutants appear to be well correlated with their crystal structures bound to the fluorinated transition-state analog. The two mutants with poor 5FC $K_m$ values display inappropriate van der Waals contact distances to the fluorine atom: either too long for effective favorable interactions (D314G; 5.1 Å) or somewhat short, leading to unfavorable crowding of electronegative oxygen and fluorine atoms (D314S, 3.0 Å). Only the D314A methyl group displays a distance to the prodrug fluorine atom that is clearly within a favorable van der Waals contact range (4.2 Å).

Other trends in the kinetic behavior of the substitutions at residue 314 are more difficult to correlate confidently with the crystallographic structures. For example, one mutant (D314S) displays an improved turnover rate ($k_{cat}$), relative to wild-type enzyme, for both cytosine (up 2-fold) and 5FC (up 3-fold, the largest increase of the mutants). In contrast, the other two mutants (D314G and D314A) display reduced turnover rates towards cytosine, but slightly elevated turnover rates with 5FC. Structural analyses of the mutants in the absence of bound inhibitor indicate that the D314S mutant displays the largest shift of D313 away from the active site iron atom, as described above (Figure 4). It is possible that the movement of active site atoms in the mutants, particularly for D314S, alters either the electrostatic potential of the bound iron atom and the $pK_a$ of its bound nucleophilic water, and/or the dynamics of lid movement across the active site during product binding and release.

In summary, we have used an approach of directed evolution, combining random mutagenesis and genetic complementation, to create novel cytosine deaminase variants that impart enhanced sensitivity to 5-fluorocytosine in transformed cells. While we have made substantial improvements to HSVTK with respect to prodrug activation using a regiospecific randomized approach (Black et al., 1996, 2001; Kokoris et al., 1999, 2002), this is the first demonstration that the substrate preference of cytosine deaminase can be significantly improved for prodrug gene therapy using a randomized approach. Surprisingly, an alanine scanning mutagenesis study carried out at the same time to investigate the role of residues 310–320 in the structure to function relationship of bCD also revealed the importance of D314 in 5FC activity (Mahan et al., 2004).

The study presented here demonstrates a direct correlation between the observed changes in kinetic profiles for cytosine and 5-fluorocytosine and the effect on cell viability in the presence of the prodrug. Despite the thermolability of yCD,
others have suggested that yCD is superior to bCD in gene therapy settings due to a 23-fold relative substrate preference for 5FC displayed by yCD (Kievit et al., 1999). However, given the thermostability of bCD and the 19-fold relative substrate preference the bCD mutant D314A displays towards 5FC, bCD D314A may be a superior suicide gene to yCD. The results indicate that the best bCD mutant described here (D314A) is an excellent candidate for subsequent preclinical comparisons with wild-type bCD and yCD, and is the obvious starting construct for further mutagenesis and selection of variants with larger shifts in substrate preference towards 5FC.

Acknowledgements

This work was supported by National Institutes of Health grants CA85939 (to M.E.B.), GM49857 (to B.L.S.) CA97380 (to B.L.S. and M.E.B.) and an NCI training grant T32 CA09437 (G.C.I.). We thank R. Michael Blaese for the kind gift of pCD2.

References


Received June 10, 2004; revised August 31, 2004; accepted September 6, 2004

Edited by Dan Tawfik