THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 21, Issue of May 23, pp. 19111–19117, 2003

Printed in U.S.A.

Crystal Structure of Yeast Cytosine Deaminase

INSIGHTS INTO ENZYME MECHANISM AND EVOLUTION*

Received for publication, January 27, 2003, and in revised form, March 11, 2003
Published, JBC Papers in Press, March 13, 2003, DOI 10.1074/jbc.M300874200

Tzu-Ping Ko‡§, Jing-Jer Lin§¶, Chih-Yung Hu‖, Yi-Hsin Hsu¶, Andrew H.-J. Wang‡, and Shwu-Huey Liaw‖‡‡§§**

*From the ‡Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan, ¶Institute of Biopharmaceutical Science, ‖Institute of Biochemistry, and **Department of Life Science, National Yang-Ming University, Taipei 11221, Taiwan, and ‡‡Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei 11217, Taiwan*

Yeast cytosine deaminase is an attractive candidate for anticancer gene therapy because it catalyzes the deamination of the prodrug 5-fluorocytosine to form 5-fluorouracil. We report here the crystal structure of the enzyme in complex with the inhibitor 2-hydroxypyrimidine at 1.6-Å resolution. The protein forms a tightly packed dimer with an extensive interface of 1450 Å^2 per monomer. The inhibitor was converted into a hydrated adduct as a transition-state analog. The essential zinc ion is ligated by the 4-hydroxyl group of the inhibitor together with His^{62}, Cys^{91}, and Cys^{94} from the protein. The enzyme shares similar active-site architecture to cytidine deaminases and an unusually high structural homology to 5-aminoimidazole-4-carboxamide-ribonucleotide transformylase and thereby may define a new superfamily. The unique C-terminal tail is involved in substrate specificity and also functions as a gate controlling access to the active site. The complex structure reveals a closed conformation, suggesting that substrate binding seals the active-site entrance so that the catalytic groups are sequestered from solvent. A comparison of the crystal structures of the bacterial and fungal cytosine deaminases provides an elegant example of convergent evolution, where starting from unrelated ancestral proteins, the same metal-assisted deamination is achieved through opposite chiral intermediates within distinctly different active sites.

Cytosine deaminase (CD,[1] EC 3.5.4.1) catalyzes the deamination of cytosine to uracil and 5-methylcytosine to thymine. The enzyme has been found in bacteria and fungi, where it plays an important role in pyrimidine salvage. However, it is not present in mammalian cells, which utilize cytidine deaminase (CDA) instead (1). The bacterial and fungal CDs are distinct from each other and have evolved separately. The 426-residue hexameric *Escherichia coli* enzyme like the murine adenosine deaminase belongs to the $(\beta/\alpha)_8$-barrel amidohydrolase superfamily, in which four histidines and one aspartate located at similar spatial positions are conserved for metal coordination and enzyme catalysis (2–4). On the other hand, the 158-residue dimeric yeast counterpart may share two conserved signature sequences, H*X*E and C*XX*C, with a variety of deaminases, and thus has been grouped into the cytidine and deoxycytidylate deaminase family in the Pfam protein family data base (5, 6). The crystal structure of *E. coli* CDA reveals that the signature sequences contain a zinc binding motif, with histidine and two cysteines acting as zinc ligands while the glutamate serves as a proton shuttle (7).

The antimetabolite 5-fluorouracil (5-FU) is one of the most active chemotherapeutic agents for the treatment of colorectal cancer, but it has limited efficacy due to gastrointestinal and hematological toxicities (8). Because of its ability to convert the relatively nontoxic 5-fluorocytosine (5-FC) into 5-FU and its absence in mammalian cells, CD has become an attractive candidate for the reduction of 5-FU toxicity toward normal distal tissues in enzyme-prodrug gene therapy (9). The combination of 5-FC with CD has been shown to effectively control tumor growth in animals and is currently being evaluated in several human clinical trials (10, 11). Previous studies also demonstrate that the use of yeast CD instead of the *E. coli* enzyme significantly improves treatment efficacy, perhaps due to the higher conversion efficiency of the 5-FC into 5-FU by the yeast CD (12, 13). Yeast CD, therefore, appears to be a better candidate for gene therapy. To provide insights into the catalytic mechanism, evolution, and gene therapy applications of the yeast CD, we have refined the enzyme structure to 1.6-Å resolution.

EXPERIMENTAL PROCEDURES

The protein crystallization, diffraction data collection, phase determination, and building of an initial dimeric model with 218 alanine residues have been described in our preliminary report (14). However, subsequent inclusion of additional 52 residues and most side chains as well as the inhibitor and zinc ion into the model did not yield satisfactory results upon further refinement. Subsequent modification, analysis, and computational refinement of the model including simulated annealing were carried out using O (15) and CNS (16). Five percent of randomly selected reflections were set aside for calculating R_{free} to monitor the progress of the refinement. An *R*-value of 0.45 at 1.6 Å resolution was obtained after rigid-body refinement of the model, and a difference Fourier map of $2F_o - F_c$ was computed. In most regions of this map, the polypeptide backbones and side chains can be clearly seen. The protein model was modified accordingly. However, in both

* This study was supported by Program for Promoting Academic Excellence of Universities Grant 89-B-FA22-2-4 (to J.-J. L.), an Academia Sinica Main Theme research grant (to A. H.-J. W.), Taipei Veteran General Hospital Grant VGH 91-376-6, and National Science Council Grant NSC 91-3112-B-010-012 (to S.-H. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1UAQ) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ These authors contributed equally to the work.

§§ To whom correspondence should be addressed: Dept. of Life Science, National Yang-Ming University, Taipei, Taiwan 11221. Tel.: 886-2-2826-7278; Fax: 886-2-2820-2449; E-mail: shliaw@ym.edu.tw.

[1] The abbreviations used are: CD, cytosine deaminase; 5-FC, 5-fluorocytosine; 5-FU, 5-fluorouracil; CDA, cytidine deaminase; AICAR transformylase, 5-aminoimidazole-4-carboxamide ribonucleotide transformylase.

TABLE I
Statistics of data collection and structural refinement
The values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P2_1$
Unit cell (Å)	47.11, 53.68, 68.08, β = 105.3°
Resolution range (Å)	50–1.6 (1.66–1.60)
Total observations	277,542 (18,894)
Unique reflections	42,187 (3,945)
Completeness (%)	96.7 (91.2)
$I/\sigma\langle I\rangle$	36.1 (2.9)
R_{merge} (%)	6.6 (33.8)
Refinement	
Resolution range (Å)	50–1.6 (1.66–1.60)
Reflections ($F > 0\ \sigma_F$)	38,222 (3,157)
R_{cryst} (%) for 95% data	17.7 (26.8)
R_{free} (%) for 5% data	21.5 (31.3)
Estimated coordinate error by Luzzati plot (Å)	0.18
Estimated coordinate error by SigmaA plot (Å)	0.20
Root Mean Square deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.6
Average *B*-factors (Å^2)	
1208 Main-chain atoms	22.9
1148 Side-chain atoms	26.5
484 Water molecules	42.1
16 Inhibitor atoms	17.6
2 Zinc ions	17.7

subunits the region corresponding to amino acid residues 48–62 showed a degree of disordered peptide modeling that did not match the electron densities. Non-crystallographic symmetry averaging and density modification of the map did not make much improvement. In the end a model was reconstructed manually with alternating reference to the densities of both subunits. The model was then refined with strong non-crystallographic symmetry restraints, yielding R and R_{free} values of 0.285 and 0.317, respectively. The resulting electron density map was much better. After the addition of water molecules and extension of the terminal polypeptides in the model, the refinement proceeded with release of the non-crystallographic symmetry restraints.

Statistics for the refined model are shown in Table I. More than 92% of the residues are in the most favored regions of the Ramachandran plot, with the remaining in the additional allowed regions. An anomalous data set at the absorption edge of zinc was collected to 1.9-Å resolution for completeness and R_{merge} values of 98.8 and 5.3%. The anomalous difference Fourier map for zinc was calculated using phase angles (−90°) from the "native" model. With almost identical unit cell dimensions, the initial R was 0.183, and it decreased immediately to 0.178 after rigid-body refinement at 1.9 Å. Two unique peaks of more than 50 σ levels were seen in the map, which correspond to locations of the zinc ions. Figs. 1, 3, *B* and *C*, and 4*A* were generated by MolScript (17) and Raster3D (18), Fig. 2*B* was generated by Insight II (Molecular Simulations Inc.), Fig. 3*A* was generated by BobScript (19), and Fig. 4, *B* and *C* was generated by Grasp (20).

RESULTS AND DISCUSSION

The Overall Structure—The current model contains residues 8–158 with clear electron density. The protein structure is composed of a central five-stranded β-sheet (β1-β5) with the strand order 2, 1, 3, 4, 5 and with β1 running antiparallel to other strands. The β-sheet is sandwiched by two α-helices (αA and αE) on one side and by three α-helices (αB-αD) on the other side (Fig. 1*A*). The enzyme forms a tightly packed dimer in the crystallographic asymmetric unit (Fig. 1*B*), without significant differences between the two subunits (root mean square deviations of 0.67 Å for all protein atoms). The dimer interface, constituted mainly by the helical layer (αB-αD) and the C-terminal tail, is extensive and buries 1450 Å^2 of the 7450-Å^2 molecular surface area per monomer. At least 26 residues from each subunit are involved in dimer formation. There are 18 direct hydrogen bonds between the protein atoms across the dimer interface, including four salt bridges between Arg^{73} and Glu^{154} and Asp^{92} and Arg^{125}.

In the dimer center, two small hydrophobic patches are formed by Met^{93}, Ala^{97}, and Met^{100} and a buried water molecule is surrounded by Thr^{60}, Ile^{65}, and Met^{93}. The anti-parallel disposition of the two dyad-related αD helices also contributes to the interface by dipole-dipole interactions. Interestingly, the side chains of Tyr^{121} from the two subunits stack very well, with a distance of 3.6 Å between the aromatic rings, and enclose four water molecules in a cavity that has a 62-Å^3 volume between helices αC and αD and is flanked by the two salt bridges between Asp^{92} and Arg^{125}.

A New Structural Superfamily—A structural homology search by DALI (21) reveals that despite the apparent lack of sequence similarity, yeast CD displays significant structural similarity to chicken ATIC (22) and to *Bacillus subtilis* and *E. coli* CDAs (7, 23), with Z-scores of 11.3, 9.0, and 8.2, respectively. ATIC is a 64-kDa bienzyme, responsible for the final steps of *de novo* purine biosynthesis. It contains an N-terminal inosine 5′-monophosphate cyclohydrolase domain and a C-terminal 5-amino-4-imidazolecarboxamide ribonucleotide (AICAR) transformylase domain. The AICAR transformylase domain is made up of subdomains 2 and 4, with a large insertion (residues 469–532, subdomain 3) in subdomain 4. The insertion does not appear to be directly involved in the enzyme activity (24), being only present in the eukaryotic ATICs but not in the prokaryotic enzymes. On the other hand, two types of CDAs have been identified; they are a dimeric form such as *E. coli* CDA that contains a catalytic domain with a non-catalytic C-terminal domain and a tetrameric form such as *B. subtilis* CDA that only has a catalytic domain. The dimeric CDA may have arisen by a gene duplication of tetrameric CDA because there is pronounced structural similarity between the catalytic and the C-terminal domains (7, 23).

Detailed structural comparisons reveal that yeast CD displays a similar fold to those involving the following five segments: subdomain 2 of the AICAR transformylase domain, subdomain 4 of the AICAR transformylase domain, the catalytic domain of *E. coli* CDA, the C-terminal domain of *E. coli* CDA, and the single catalytic domain of *B. subtilis* CDA. The main-chain atoms of the 60–62 structurally equivalent residues overlay with root mean square deviations of 1.11, 1.72, 1.34, 1.21, and 1.00 Å and with 7, 5, 17, 15, and 22% sequence identity, respectively (Fig. 2). Surprisingly, even though yeast CD shares a similar active site and a high sequence identity with the CDAs, its closest structural match is the subdomain 2 of the AICAR transformylase domain. This is perhaps due to the common αD helix and the same direction of β5 strand. Yeast CD and subdomains 2 and 4 of the AICAR transformylase domain contain common four helices αA-αD, whereas the CDAs only have three helices, αA-αC. Only the strand β1 in yeast CD and AICAR transformylase run antiparallel to the others, whereas both β1 and β5 are antiparallel in CDAs. There are only four β-stands (β1-β4) in the C-terminal domain of *E. coli* CDA.

The strong conservation of the tertiary structures of these domains suggests that they are descended from a single ancestral gene and thereby define a new superfamily. This gene may have been duplicated and fused to form *E. coli* CDA and the AICAR transformylase domain, with subsequent evolution causing significant divergence. For example, yeast CD and CDAs have developed a zinc binding motif for their metal-assisted catalysis. The orientations between helices αB and αC in the C-terminal domain of *E. coli* CDA and in the subdomains 2 and 4 in the AICAR transformylase domain are different from those of yeast CD and CDAs, perhaps due to the lack of a zinc binding motif. The missing αD helix in CDAs has perhaps caused the β5 strand to evolve to be antiparallel to β4 and the

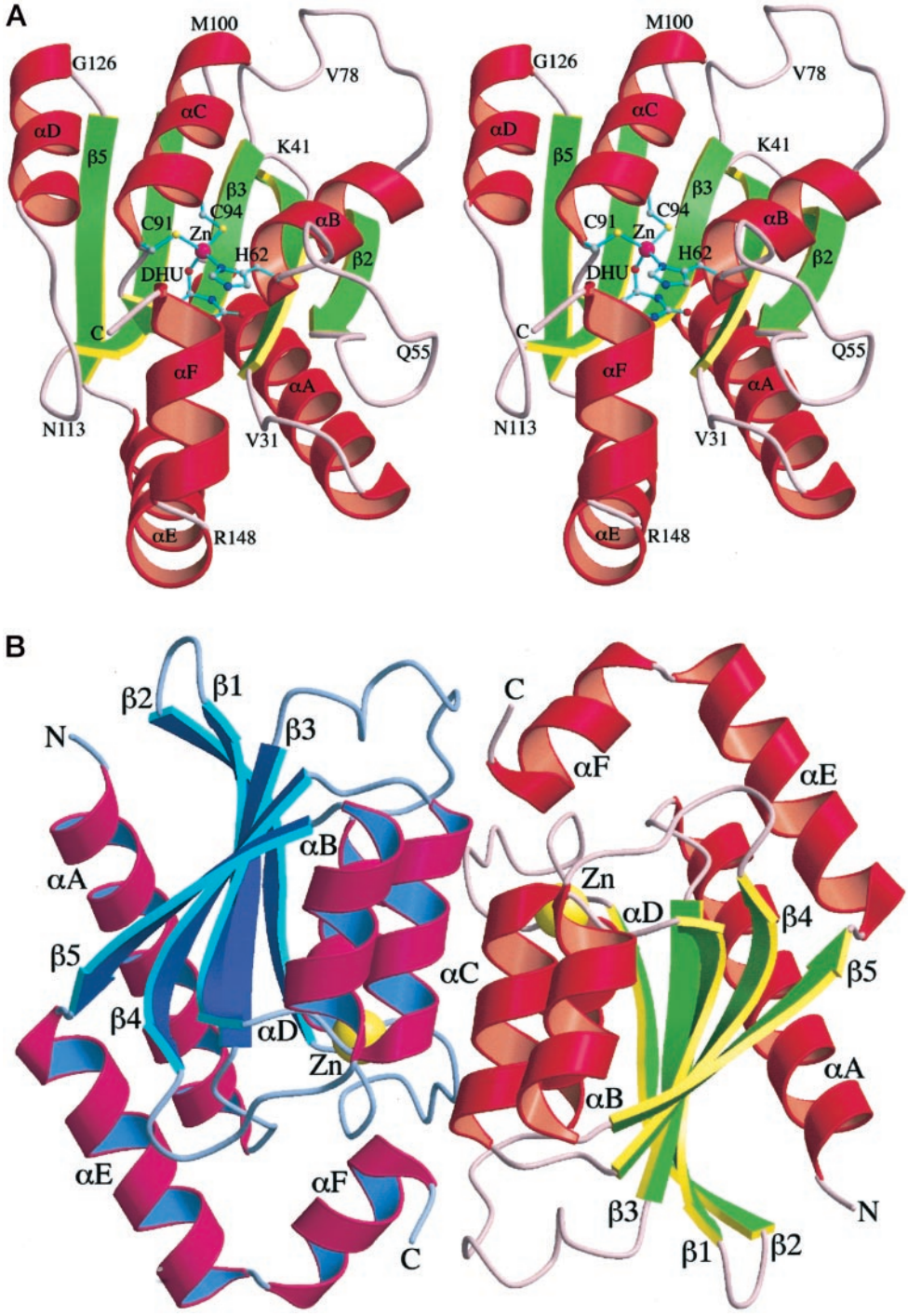


FIG. 1. **Structure of yeast CD.** *A*, stereo view of the monomer, which is a three-layered $\alpha/\beta/\alpha$ structure with a central β-sheet sandwiched on either side by α-helices. The tightly bound zinc ion is shown as a *magenta sphere* with its ligands and the inhibitor (3,4-dihydrouracil (*DHU*)) as *ball-and-stick* representations. *B*, the dimer has one monomer colored in *red* (helices) and *green* (strands), whereas the other is colored in *magenta* (helices) and *blue* (strands). The zinc ions are shown as *yellow spheres*.

C terminus to be in an opposite direction (Fig. 2*B*). The C-terminal copy in *E. coli* CDA retains its structural integrity but without any catalytic functionality. On the other hand, subdomains 2 and 4 of the AICAR transformylase have developed another enzyme mechanism (24).

The Active-site Architecture and Substrate Binding—The active site contains one tightly bound metal ion even though no metal ions were added either during protein purification or crystallization. The metal ion was identified by a very strong spherical electron density in the Fourier map and was assigned as zinc based on the zinc anomalous data (Fig. 3*A*). The zinc ion is tetrahedrally coordinated by His^{62} $N^{\delta 1}$ (2.06 Å), Cys^{91} S^{γ} (2.31 Å), Cys^{94} S^{γ} (2.25 Å), and the hydroxyl group of the bound inhibitor (2.08 Å) (Fig. 3*B*). The protein ligands are located on the N termini of the αB and αC helices, and the helix dipoles may partially neutralize the negative charges of the cysteinate ligands to lower the pK_a value of the nucleophilic water molecule (23).

2-Hydroxypyrimidine has been shown to be converted by *E. coli* CD into 4-(*S*)-hydroxyl-3,4-dihydropyrimidine (or 3,4-dihydrouracil), a hydrated adduct, and this acts as a transition-state intermediate analog (2). Similar inhibitors have been

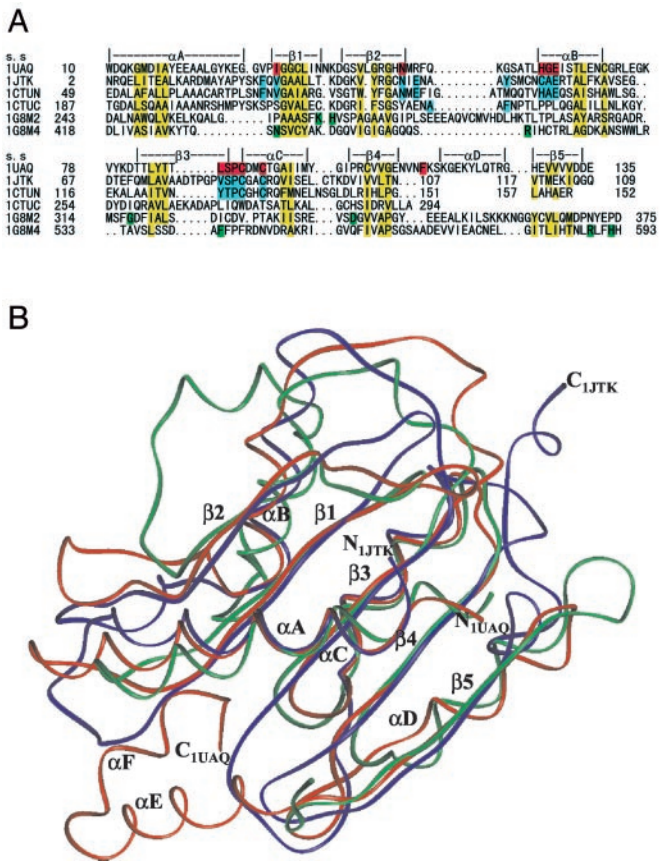


FIG. 2. **A new structural superfamily.** *A*, structure-based sequence alignment of the yeast CD (Protein Data Bank code 1UAQ), the *B. subtilis* CDA (Protein Data Bank code 1JTK), the catalytic and the non-catalytic C-terminal domains of *E. coli* CDA (1CTUN and 1CTUC from Protein Data Bank code 1CTU), and the subdomains 2 and 4 of the chicken AICAR transformylase domain (1G8M2 and 1G8M4 from Protein Data Bank code 1G8M). The active-site residues for yeast CD, CDAs, and AICAR transformylase are shaded in *red*, *cyan*, and *green*, respectively. The common secondary structure elements (αA-αD, β1-β5) for yeast CD are labeled (*s.s*), and the residues for the conserved hydrophobic core are shaded in *yellow*. *B*, structural superposition of 1UAQ (*red*), 1JTK (*blue*), and 1G8M2 (*green*). The letters N_{1UAQ}, C_{1UAQ}, N_{1JTK}, and C_{1JTK} indicate the N and C termini of yeast CD and *B. subtilis* CDA, respectively. Structural comparison reveals that the common β-strands and α-helices are highly correlated, whereas the surface-exposed loops diverge significantly.

used in previous crystallographic studies of murine adenosine deaminase (3, 25) and *E. coli* CDA (7, 26). In the structure presented here, the inhibitor 2-hydroxypyrimidine is firmly embedded in the active site as the hydrated form and is identified by its strong electron density that correlates well with the shape of a pyrimidine ring. The orientation of the pyrimidine ring is unambiguous because of clear electron density bulges for the 2-keto oxygen (O^2) and the hydroxyl group at position 4 (OH^4) (Fig. 3*A*) and because of its extensive hydrogen bond interactions and hydrophobic contacts (Fig. 3*B*).

The OH^4 group of the pyrimidine ring, which is an analog to the nucleophilic hydroxyl in the reaction transition state, directly coordinates to the catalytic zinc ion and interacts with Glu^{64} $O^{\epsilon 2}$ (2.45 Å) and Cys^{91} N (3.10 Å). Additionally, the N^3 atom interacts with Glu^{64} $O^{\epsilon 1}$ (2.81 Å), the O^2 atom forms hydrogen bonds with Gly^{63} N (2.79 Å) and Asn^{51} $N^{\delta 2}$ (2.95 Å), and the N^1 atom contacts with Asp^{155} $O^{\delta 1}$ (2.70 Å). In addition to the extensive hydrogen bond network described above, the imidazole ring of His^{62} stacks on the pyrimidine ring with an interplanar distance of 3.3 Å. Also, the C^2, C^5, and C^6 atoms of the pyrimidine ring make close contacts with Ile^{33} $C^{\delta 1}$ (3.20 Å), Lue^{88} $C^{\delta 2}$ (4.36 Å), Phe^{114} $C^{\epsilon 2}$ (4.08 Å), Trp^{152} C^{h2} (3.74 Å), and Ile^{156} $C^{\gamma 2}$ (3.98 Å). Therefore, the structure of the complex presented here reveals that yeast CD converts the inhibitor 2-hydroxypyrimidine into 4-(*R*)-hydroxyl-3,4-dihydropyrimidine, which is enantiomeric to the configuration observed in *E. coli* CD; this will be discussed later (Fig. 3*C*).

A Comparison of Yeast CD and CDAs—The active-site architectures of yeast CD and *E. coli* and *B. subtilis* CDAs share striking similarity (Fig. 4*A*). Superposition of their active sites reveals a very similar interaction network between the attacking water molecule, the zinc ion, the zinc ligands, and the proton shuttle. Thus, the four highly conserved residues in the signatures, H*X*E (or C*X*E), and C*XX*C, of the cytidine and deoxycytidylate deaminase family (5, 6), are responsible for generation of the attacking hydroxide ion and for enhancement of the attacked substrate susceptibility during the enzyme catalysis, suggesting a similar zinc-assisted deamination mechanism in the family.

On the other hand, each member in the family should contain its unique substrate recognition residues, for example, the C-terminal tail in yeast CD. Cytosine has been shown to be a slow substrate for *E. coli* CDA, whereas cytidine is not a substrate for yeast CD (27, 28). The position of the pyrimidine ring of the inhibitor in yeast CD is structurally equivalent to that in the CDAs (Fig. 4). However, several interactions between the inhibitor and enzyme in yeast CD are different from those in the CDAs. The conserved asparagine forms a hydrogen bond with the $OH^{3'}$ of the ribose in the CDAs, whereas it interacts with the pyrimidine 2-keto oxygen in the yeast CD. The imidazole ring of the conserved histidine for zinc ligation, His^{102} in the *E. coli* CDA, does not stack on the inhibitor pyrimidine ring. The ribosyl sugar of the cytidine cannot be accommodated within the active site of the yeast CD, because Asp^{155} occupies part of the ribose-binding site (Fig. 4).

Another major difference between yeast CD and the CDAs is a change in the mechanism of the active-site entrance. The structure reported here, the yeast CD, appears to have a closed conformation, because the C-terminal tail forms a "flap" across the entrance and, thus, caps the narrow opening of the active-site cavity (Fig. 4*B*). This flap may function as a gate that controls access to the active site and, thus, regulates both substrate binding and product release. The closed conformation described here may be due to direct interactions between the inhibitor and the Trp^{152}, Asp^{155}, and Ile^{156} residues of the C terminus, allowing sealing of the entrance. Unlike the active site of yeast CD, which is complete within the monomer itself, the active site in one monomer of the tetrameric *B. subtilis* CDA is made up of three subunits, and its entrance appears to be controlled by the C-terminal helix from the adjacent subunit (Fig. 4*C*). However, the flexibility of the C-terminal tail in the presence of the inhibitor is high, perhaps due to a lack of significant interaction with the inhibitor. This suggests that the C-terminal tail can easily swing out and allow substrate entry or product release (23). On the other hand, the active-site entrance switch in the dimeric *E. coli* CDA is more complicated and still unclear. In a single monomer complex, there are solvent-accessible areas of the inhibitor, but these are covered by three loops from the second monomer in the dimer, which complete the active site in the first monomer and form a lid on the active site (7).

A Proposed Catalytic Mechanism—On the basis of the structural studies on yeast CD and CDAs (7, 26, 29, 30), we propose a deamination mechanism for yeast CD as outlined in Scheme 1. First, the zinc ion activates the attacking water molecule to form a hydroxide ion with the assistance of Glu^{64}. The substrate binds to the active site and induces a closed conforma-

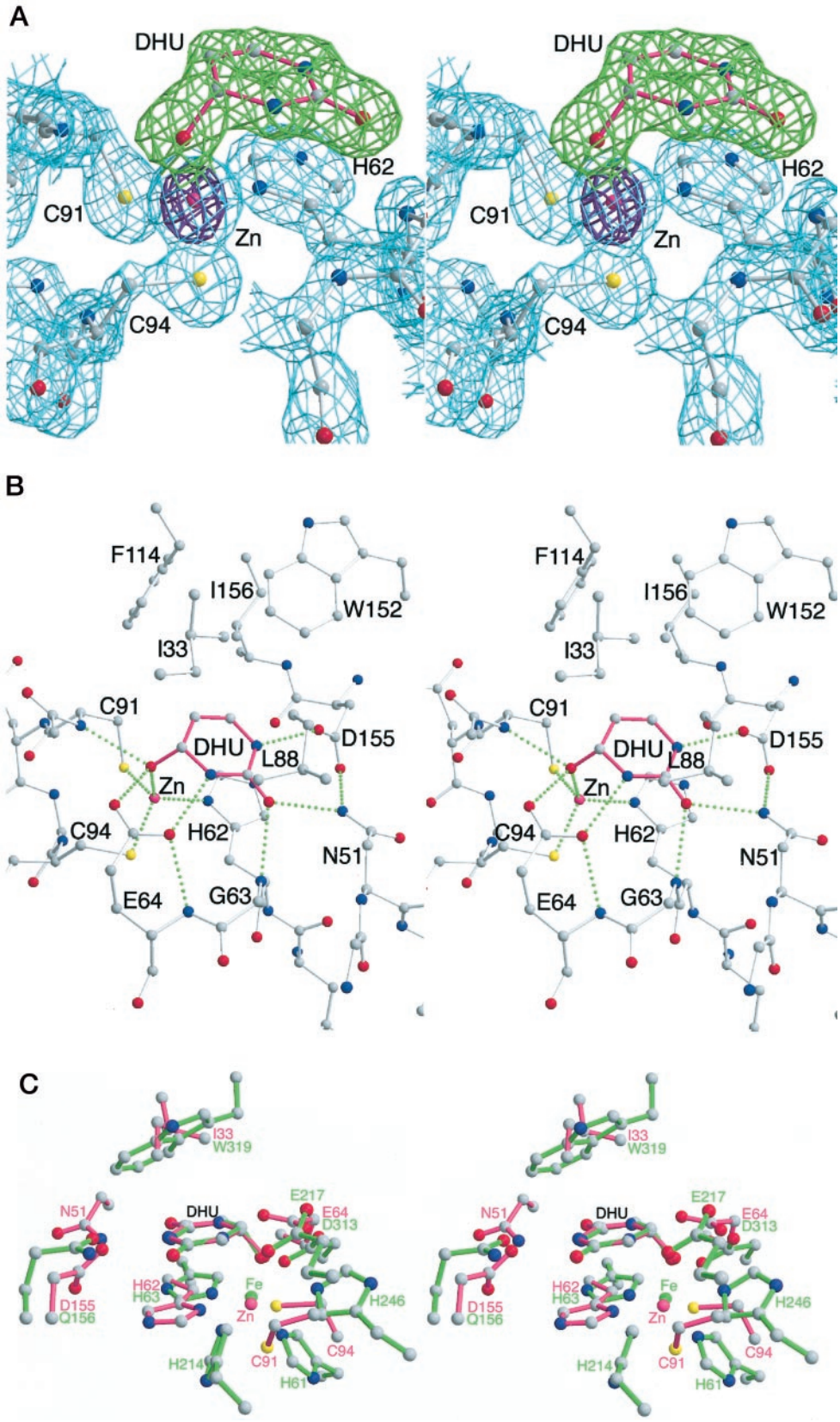


FIG. 3. **The active site.** *A*, stereo view of the $2F_o - F_c$ electron density map of the active site contoured at 1.5 σ level and shown in *cyan* and the difference anomalous map for the zinc ion contoured at 30 σ level and shown in *purple*. The densities of the inhibitor molecule are highlighted in *green*. The active-site residues and the inhibitor (3,4-dihydrouracil (*DHU*)) are shown as *ball-and-stick* representations, and the zinc ion is shown as a *magenta sphere*. *B*, stereo view of the interaction networks in the active site. There are six direct hydrogen bonds between the protein molecule and the inhibitor (see *The Active Site Architecture and Substrate Binding*). *C*, stereo view of a comparison of the active sites of the yeast (*magenta*) and *E. coli* cytosine deaminase (Protein Data Bank code 1K70, *green*) based on superposition of the inhibitor 3,4-dihydrouracil. The residue numbering is labeled in the same color for each protein. The bacterial enzymes utilize three histidines and one aspartate for iron ligation (His^{61}, His^{63}, His^{214}, and Asp^{313}). Asp^{313} also forms a hydrogen bond with the attacking water molecule. Glu^{217} interacts with the N^3 atom of the pyrimidine ring, and Gln^{156} interacts with the O^2 and N^1 atoms.

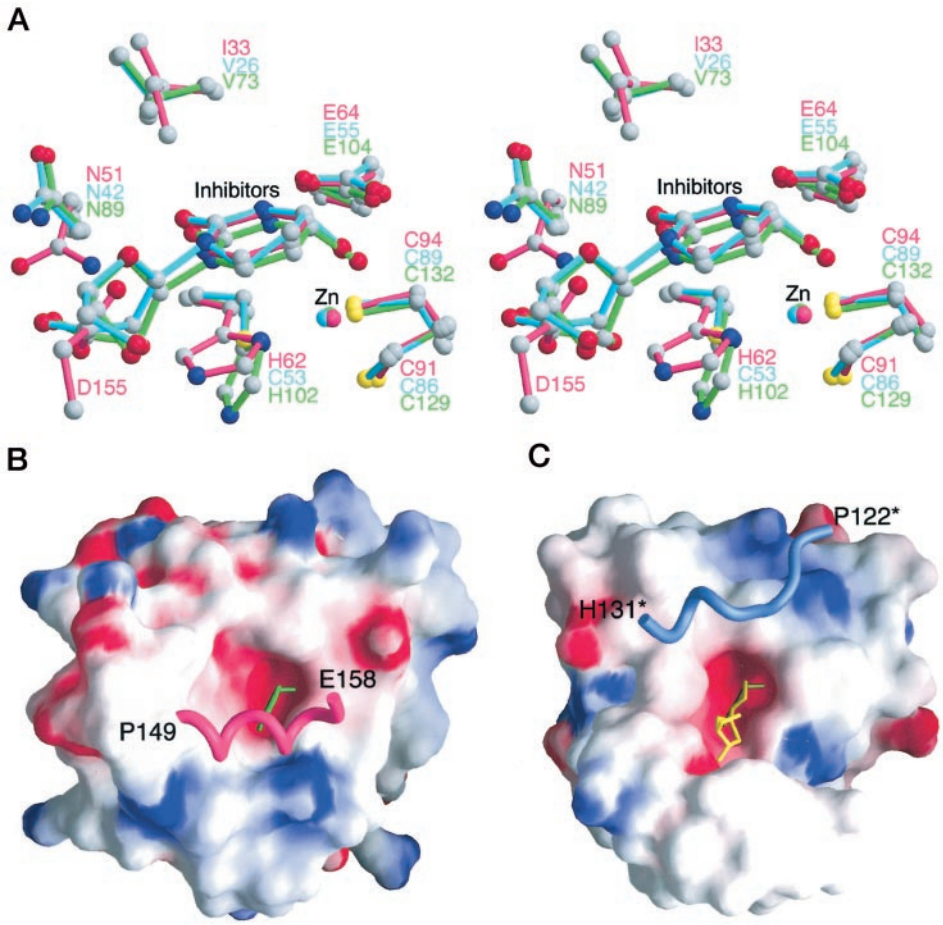
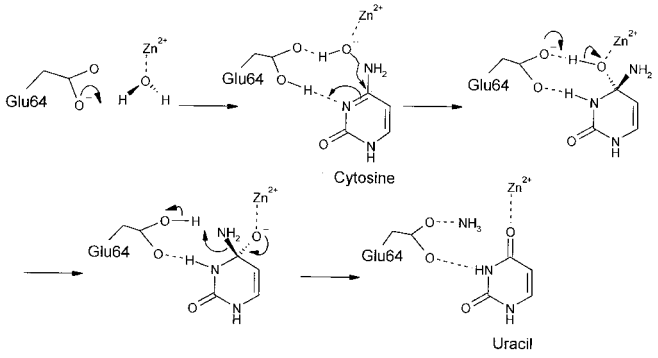


FIG. 4. **A comparison of yeast CD and CDAs.** *A*, superposition of the active sites of the yeast CD (1UAQ, *magenta*), and *B. subtilis* (1JKT, *cyan*) and *E. coli* (1CTU, *green*) CDAs based on 61-residue common structure (β1-β4, αA-αC). Two classes of CDAs have been identified: a dimeric form such as *E. coli* CDA that utilizes one histidine and two cysteines for zinc ion coordination and a tetrameric form such as *B. subtilis* CDA that uses three cysteines instead. Molecular surfaces of yeast CD (*B*) and *B. subtilis* CDA (*C*) are colored for electrostatic potential from -10 k_BT (*red*) to 10 k_BT (*blue*). The C-terminal tail (residues 149–158) of yeast CD is not included in surface construction but is shown explicitly as worms as is the C-terminal residues (residues 122*-131*) from one adjacent subunit in *B. subtilis* CDA. The substrate analogues for cytosine and cytidine are colored *green* and *yellow*, respectively. The zinc ion is embedded at its the deepest part, whereas the inhibitor lies near the cavity opening. The CDAs have space to accommodate the ribosyl sugar of cytidine, where yeast CD does not, due to the blocking by the C-terminal tail.



SCHEME 1. **The proposed catalytic mechanism for yeast CD.**

tion to sequester the reaction from solvent. Glu^{64} serves as a proton shuttle, abstracting a proton from the zinc-bound water on one hand and subsequently protonating the N^3 of cytosine on the other hand, thereby reducing the N^3 to C^4 double bond character and making the C^4 atom more susceptible to the hydroxide ion for formation of the tetrahedral intermediate. Cleavage of the carbon-nitrogen bond is assisted by transferring the proton from OH^4 to $NH_2{}^4$ via Glu^{64} and a pyrimidinolate (O^-) intermediate. The newly formed uracil may move toward the zinc ion for ligation, and this would weaken its interaction with the C-terminal tail, allowing its release from the active site.

An Example for Convergent Evolution—The crystal structures of bacterial and fungal CDs provide an elegant example of convergent evolution in that they have evolved from unrelated ancestral proteins but have achieved the same deamination reaction (Fig. 3*C*). The two enzymes have completely different tertiary structures and distinct active sites, and they even catalyze the reaction through opposite chiral intermediates. However, both enzymes utilize a divalent cation to activate the water molecule and protein carboxylate group(s) for proton abstraction and donation. In yeast CD, the zinc ion is tetrahedrally coordinated, and there is only one carboxylate group (Glu^{64}) to facilitate proton shuttling during catalysis. On the other hand, in *E. coli* CD, the iron ion is penta-coordinated, and the proton transfer is carried out by at least two different residues, Asp^{313} for proton abstraction from the nucleophile water and Glu^{217} for proton donation to the N^3 atom (2). The different interaction networks between the two enzymes and the substrates lead to opposite enantiomeric configurations of their transition-state intermediates.

Previous studies demonstrate that yeast CD has a greater therapeutic potential as part of an enzyme-prodrug strategy in cancer patients compared with the bacterial enzymes. This is perhaps due to its more efficient conversion of 5-FC into 5-FU

(12, 13). The K_m value of *E. coli* CD for 5-FC is 8-fold higher than that for the natural substrate cytosine, whereas the K_m of yeast CD for 5-FC is only one-fifth of that for cytosine (12). The structure presented here shows that the C^5 atom of the pyrimidine ring is adjacent to the hydrophobic cluster, which includes Phe^{114}, Trp^{152}, and Ile^{156} (Fig. 3*B*). In particular, the C^5 atom faces toward Phe^{114}, with a shortest distance of 4.1 Å. A fluorine atom is attached to an aromatic ring will make it very hydrophobic. Therefore, the hydrophobic cluster in yeast CD should enhance the binding of 5-FC. On the other hand, the C^5 atom faces toward Asp^{314} in *E. coli* CD, with a distance between C^5 and Asp^{314} $O^{\delta 1}$ of 3.6 Å. Close contact with such a polar residue would be unfavorable to 5-FC binding. These structural observations are consistent with the enzyme kinetic measurements.

Conclusion—In summary, the yeast CD structure shares high similarity to the AICAR transformylase domain, which may define a new superfamily. The enzyme structure also reveals a common zinc-assisted deamination for the cytidine and deoxycytidylate deaminase family. The unique C-terminal tail in yeast CD is involved in dimerization, substrate specificity, and as an active-site entrance switch. On the basis of the structural information, some protein engineering trials to obtain new functions through site-directed mutation and deletion of the C-terminal tail have been initiated.

Acknowledgments—The synchrotron radiation experiments were performed at the National Synchrotron Radiation Research Center, Hsinchu, Taiwan, at the Photon Factory, Tsukuba, Japan, and at the SPring-8, Sayo, Japan. Thanks to Professor Ralph Kirby for help in preparing this manuscript.

REFERENCES

1. Nishiyama, T., Kawamura, Y., Kawamoto, K., Matsumura, H., Yamamoto, N., Ito, T,. Ohyama, A., Katsuragi, T., and Sakai, T. (1985) *Cancer Res.* **45,** 1753–1761
2. Ireton, G. C., McDermott, G., Black, M. E., and Stoddard, B. L. (2002) *J. Mol. Biol.* **315,** 687–697
3. Wilson, D. K., Rudolph, F. B., and Quiocho, F. A. (1991) *Science* **252,** 1278–1284
4. Liaw, S.-H., Chen, S.-J., Ko, T.-P., Hsu, C.-S., Chen, C.-J., Wang, A. H.-J., and Tsai, Y.-C. (2003) *J. Biol. Chem.* **278,** 4957–4962
5. Reizer J., Buskirk, S., Bairoch, A., Reizer, A., and Saier, M. H., Jr. (1994) *Protein Sci.* **3,** 853–856
6. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) *Nucleic Acids Res.* **30,** 276–280
7. Betts, L., Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W., Jr. (1994) *J. Mol. Biol.* **235,** 635–656
8. Vauthey, J. N., Marsh Rde, W., Cendan, J. C., Chu, N. M., and Copeland, E. M. (1996) *Br. J. Surg.* **83,** 447–455
9. Greco, O., and Dachs, G. U. (2001) *J. Cell. Physiol.* **187,** 22–36
10. Nyati, M. K., Symon, Z., Kievit, E., Dornfeld, K. J., Rynkiewicz, S. D., Ross, B. D., Rehemtulla, A., and Lawrence, T. S. (2002) *Gene Ther.* **9,** 844–849
11. Miller, C. R., Williams, C. R., Buchsbaum, D. J., and Gillespie, G. Y. (2002) *Cancer Res.* **62,** 773–780
12. Kievit, E., Bershad, E., Ng, E., Sethna, P., Dev, I., Lawrence, T. S., and Rehemtulla, A. (1999) *Cancer Res.* **59,** 1417–1421
13. Kievit, E., Nyati, M. K., Ng, E., Stegman, L. D., Parsels, J., Ross, B. D., Rehemtulla, A., and Lawrence, T. S. (2000) *Cancer Res.* **60,** 6649–6655
14. Hu, C.-H., Hsu, Y.-H., Lin, J.-J., and Liaw, S.-H. (2003) *Acta Crystallogr. Sec. D* **59,** 950–952
15. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sec. D* **47,** 110–119
16. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sec. D* **54,** 905–921
17. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
18. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
19. Esnouf, R. M. (1999) *Acta Crystallogr. Sec. D* **55,** 938–940
20. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
21. Dietmann, S., Park, J., Notredame, C., Heger, A., Lappe, M., and Holm L. (2001) *Nucleic Acids Res.* **29,** 55–57
22. Greasley, S. E., Horton, P., Ramcharan, J., Beardsley, G. P., Benkovic, S. J., and Wilson, I. A. (2001) *Nat. Struct. Biol.* **5,** 402–406
23. Johansson, E., Mejlhede, N., Neuhard, J., and Larsen, S. (2002) *Biochemistry* **41,** 2563–2570
24. Wolan, D. W., Greasly, S. E., Beardsley, G. P., and Wilson, I. A. (2002) *Biochemistry* **41,** 15505–15513
25. Wang, Z., and Quiocho, F. A. (1998) *Biochemistry* **37,** 8314–8324
26. Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W., Jr. (1995) *Biochemistry* **34,** 4516–4523
27. Carlow, D., and Wolfenden, R. (1998) *Biochemistry* **37,** 11873–11878
28. Ipata, P. L., and Cercignani, G. (1978) *Methods Enzymol.* **51,** 394–401
29. Xiang S., Short, S. A., Wolfenden, R., and Carter, C. W., Jr. (1996) *Biochemistry* **35,** 1335–1341
30. Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W., Jr. (1997) *Biochemistry* **36,** 4768–4774

Crystal Structure of Yeast Cytosine Deaminase: INSIGHTS INTO ENZYME MECHANISM AND EVOLUTION

Tzu-Ping Ko, Jing-Jer Lin, Chih-Yung Hu, Yi-Hsin Hsu, Andrew H.-J. Wang and Shwu-Huey Liaw

J. Biol. Chem. 2003, 278:19111-19117.
doi: 10.1074/jbc.M300874200 originally published online March 13, 2003
