X-ray Structure of *Paramecium bursaria* Chlorella Virus Arginine Decarboxylase: Insight into the Structural Basis for Substrate Specificity†

Rahul Shah,‡ Radha Akella,§ Elizabeth J. Goldsmith,§ and Margaret A. Phillips*,‡

Departments of Pharmacology and Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9041

Received November 13, 2006; Revised Manuscript Received December 26, 2006

ABSTRACT: The group IV pyridoxal-5′-phosphate (PLP)-dependent decarboxylases belong to the β/α barrel structural family, and include enzymes with substrate specificity for a range of basic amino acids. A unique homolog of this family, the *Paramecium bursaria* Chlorella virus arginine decarboxylase (*cv*ADC), shares about 40% amino acid sequence identity with the eukaryotic ornithine decarboxylases (ODCs). The X-ray structure of *cv*ADC has been solved to 1.95 and 1.8 Å resolution for the free and agmatine (product)-bound enzymes. The global structural differences between *cv*ADC and eukaryotic ODC are minimal (rmsd of 1.2−1.4 Å); however, the active site has significant structural rearrangements. The key "specificity element," is identified as the 3_{10}-helix that contains and positions substrate-binding residues such as E296 *cv*ADC (D332 in *T. brucei* ODC). In comparison to the ODC structures, the 3_{10}-helix in *cv*ADC is shifted over 2 Å away from the PLP cofactor, thus accommodating the larger arginine substrate. Within the context of this conserved fold, the protein is designed to be flexible in the positioning and amino acid sequence of the 3_{10}-helix, providing a mechanism to evolve different substrate preferences within the family without large structural rearrangements. Also, in the structure, the "K148-loop" (homologous to the "K169-loop" of ODC) is observed in a closed, substrate-bound conformation for the first time. Apparently the K148 loop is a mobile loop, analogous to those observed in triose phosphate isomerase and tryptophan synthetase. In conjunction with prior structural studies these data predict that this loop adopts different conformations throughout the catalytic cycle, and that loop movement may be kinetically linked to the rate-limiting step of product release.

Amino acid decarboxylation is an essential reaction for the formation of many important cellular metabolites, including polyamines, amino acids, and neurotransmitters (*1*−*3*). Because of the roles these metabolites play in cellular proliferation, enzymes catalyzing amino acid decarboxylation have been of interest as drug targets for the prevention and treatment of a number of proliferative diseases, such as cancer and parasitic infections (*4*−*7*). A well-studied inhibitor of polyamine biosynthesis, α-difluoromethylornithine (DFMO), a suicide inhibitor of ornithine decarboxylase (ODC), is an approved treatment for African sleeping sickness. Two distinct structural families of pyridoxal-5′phosphate dependent basic amino acid decarboxylases have been described. The group III decarboxylases are structurally homologous to aspartate amino transferase and the group IV decarboxylases are homologs of alanine racemase, assuming a β/α barrel fold (*8*, *9*). Representatives of both structural classes have substrate specificity for ornithine, lysine and arginine; although, the group IV decarboxylases have expanded substrate specificity that includes the dicarboxylic acid diaminopimelate (*8*, *9*). Enzymes in the Group IV family with differing specificities share low sequence identity (i.e., < 20%) and are distantly related. A novel exception is the *Paramecium bursaria* Chlorella virus arginine decarboxylase (*cv*ADC),[11] which shares 40% sequence identity with eukaryotic ODCs, typical of the level of identity shared among homologs with similar specificity (*10*, *11*). It is the only example of a group IV decarboxylase with different substrate specificity found to group with the ODCs, thus providing a unique opportunity to study the evolution of substrate specificity.

The X-ray structures of several eukaryotic ODCs (*12*−*14*), and bacterial diaminopimelate decarboxylases (DAPDC) have been solved (*15*, *16*). The enzymes are obligate homodimers and two active sites are formed at the subunit interface between the N-terminal β/α barrel domain, which houses PLP, and the two C-terminal β-barrel domains. Ligand-bound *Trypanosoma brucei* ODC (*tb*ODC) structures have been solved with putrescine, DFMO and D-ornithine (*14*, *17*, *18*). The ligand is in an extended conformation at

† This work was supported by National Institutes of Health grants (R01 AI34432) (to MAP), and Welch Foundation grants, I-1257 (to M.A.P.) and I-1128 (to E.J.G.).

* Author to whom all correspondence should be addressed. Telephone: (214) 645-6164. Fax: (214) 645-6166. E-mail: margaret.phillips@UTSouthwestern.edu.

‡ Department of Pharmacology, University of Texas Southwestern Medical Center.

§ Department of Biochemistry, University of Texas Southwestern Medical Center.

[1] Abbreviations: *cv*ADC, *Paramecium bursaria* Chlorella virus arginine decarboxylase; PLP, pyridoxal-5′-phosphate; ODC, ornithine decarboxylase; *tb*ODC, *Trypanosoma brucei* ornithine decarboxylase; *h*ODC, human ornithine decarboxylase; DAPDC, diaminopimelate decarboxylase; *mt*DAPDC, *Mycobacterium tuberculosis* DAPDC; *ec*DAPDC, *Escherichia coli* DAPDC, and *mj*DAPDC, *Methanococcus jannaschii* DAPDC.

10.1021/bi6023447 CCC: $37.00
Published on Web 02/17/2007

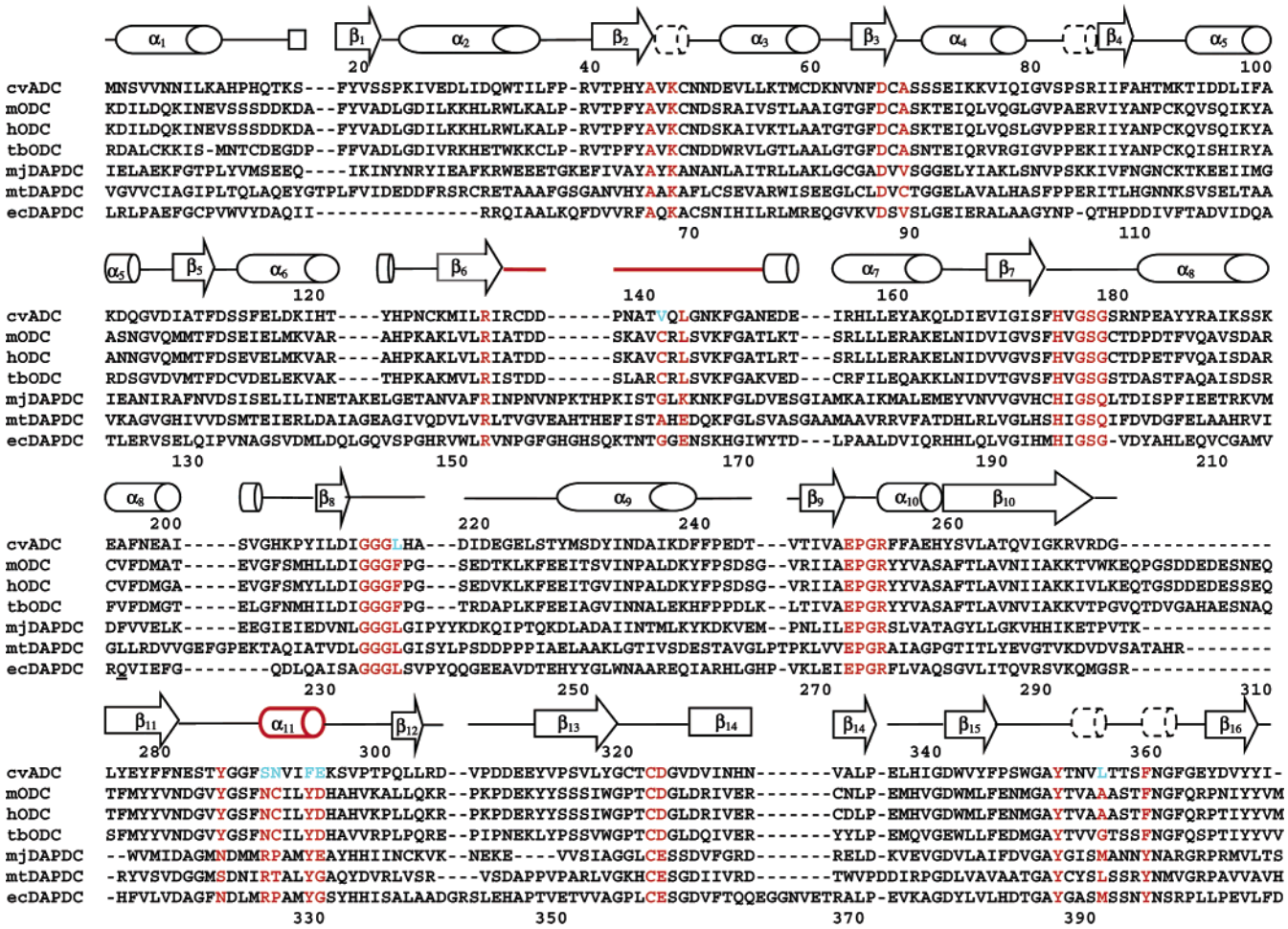

FIGURE 1: Sequence alignment of Group IV decarboxylases: The alignment was produced using the DALI alignment server and includes only sequences for which structural information is available: ODC from mouse, human and *T. brucei*, *cv*ADC, and DAPDC from *M. jannaschii*, *M. tuberculosis*, and *E. coli*. The secondary structure of *cv*ADC is illustrated above the sequences: cylinders (α-helices), arrows (*β*-strands) and lines (loops). Cylinders in dashed-lines represent short, 3-residue 3_{10} helices. The active site lid is shown as a red-highlighted line between β_6 and α_7 and the "specificity element" is the 3_{10}-helix α_{11} shown as a red-highlighted cylinder. Highlighted in red in the multiple sequence alignment are residues in the active site, which is defined as a 5 Å shell around the agmatine product. Highlighted in blue are active site residues that are variable between the ODC sequences and *cv*ADC.

the subunit interface and each ligand makes similar contacts. The α-amino group forms a Schiff base with PLP and the *δ*-amino group interacts with two aspartic acid residues, D332 *tb*ODC from the C-terminal *β*-barrel domain of one monomer and D361 tbODC from the C-terminal *β*-barrel domain of the opposite monomer. The DAPDC active site is arranged similarly although the sequence identity between ODCs and DAPDCs is less than 20%. Based on the *Methanococcus jannaschii* DAPDC structure, E373 *mj*DAPDC is thought to play a similar role as D361 *tb*ODC and the carboxylate of the bound lysine product makes contacts with R297 DAPDC (R277 *tb*ODC) from the α/*β* barrel domain. However, overlays of the *mj*DAPDC structure onto the *tb*ODC structure show numerous changes making it unclear which are involved in specificity determination.

P. bursaria Chlorella virus encodes enzymes to synthesize homospermidine from L-arginine, one of which is *cv*ADC (*10*, *19*). In previous studies this enzyme was demonstrated to prefer L-arginine over L-ornithine as a substrate suggesting that it had undergone rapid evolution to change function (*11*). Comparison of the active sites of *cv*ADC and ODC suggested that a relatively limited number of substitutions might account for the specificity difference (Figure 1). An obvious difference was E296 of *cv*ADC replaces D332 of ODC. However, both *T. brucei* ODC (D332E) and *cv*ADC (E296D) are inactive enzymes (*11*).

In order to study the structural basis of specificity and learn more about the catalytic mechanism in the group IV decarboxylases, the X-ray structures of *cv*ADC in both free and product-bound states have been determined at 1.95 and 1.8 Å resolution, respectively. The overall monomeric structure of *cv*ADC is similar to eukaryotic ODC (rmsd of 1.2−1.4 Å). However, the origin of the arginine specificity of *cv*ADC is apparent. It arises from different positioning of a 3_{10}-helix (residues S291-E296 *cv*ADC) in the C-terminal domain that includes the substrate-binding residue E296 *cv*ADC (D332 *tb*ODC). Further, a conserved loop near the active site, residues D138−V143 (K148 *cv*ADC-loop) and analogous to D159−C164 in ODC (K169 ODC loop), is observed in a closed conformation for the first time. In comparison with human ODC, where the homologous loop adopts an open conformation (*13*), it appears that the loop serves the function of a mobile lid allowing substrate access, and sequestering substrates from solvent.

Table 1. Data Collection, Processing and Refinement Statistics

	Se-Met native *cv*ADC	*cv*ADC-agmatine
	(a) Data Collection and Processing	
wavelength (Å)	0.98066	0.97945
resolution (Å)	50−1.95	50−1.8
space group	$P2_12_12_1$	$P2_12_12_1$
cell dimensions	$a = 116.1$ Å, $b = 116.9$ Å, $c = 269.4$ Å, $\alpha = \beta = \gamma = 90°$	$a = 116.2$ Å, $b = 117.3$ Å, $c = 269.7$ Å, $\alpha = \beta = \gamma = 90°$
tot. no. of reflections	1,722,504	1,619,455
no. of unique reflections	235,861	336,913
redundancy (last shell)	4.9 (2.8)	4.7 (4.2)
completeness (%) (last shell)	88.7 (50.1)	99.5 (97.5)
I/σ (last shell)	9.8 (2.8)	14.7 (2.2)
R_{merge} (last shell)	0.073 (0.21)	0.118 (0.48)
	(b) Refinement	
R/R_{free}	21.1/24.2	21.9/23.8
no. of non-hydrogen atoms	23550	23676
no. of H_2O atoms	1182	1467
rmsd for bond lengths (Å)	0.02	0.01
rmsd for bond angles (deg)	1.41	1.19
average *B*-value (Å^2)	12.5	14.1
B rmsd for bonded main-chain (side chain)	0.85 (2.02)	0.48 (1.42)
Ramachandran plot		
% in most favored region	89.1	89.1
% in additional allowed region	10.9	10.9
% generously allowed or disallowed	0	0

EXPERIMENTAL PROCEDURES

Materials. Methyl-pentanediol (MPD) was purchased from Hampton Research, and all other chemicals were purchased from Sigma.

Methods. cvADC Expression and Purification. The expression and purification of His_6-tagged *cv*ADC from pODCTM9 and pODCTM9 T142A was described previously (*11*). Briefly, protein was produced heterologously in *Escherichia coli* (BL21 DE3 strain) under an inducible promoter and purified by Ni^{2+}-agarose column chromatography followed by gel filtration column chromatography. Purified protein was concentrated via ultracentrifugation. The T142A mutation arose during prior cloning of the gene (*10*), however the two constructs produce enzymes with equivalent activity, and they have both been used in these studies (*11*). Protein derived from pODCTM9 T142A was used for the crystallization of free *cv*ADC protein. Site-directed mutagenesis (Quickchange, Stratagene, La Jolla, CA) was employed to regenerate the wild-type pODCTM9 plasmid. The protein was used to crystallize *cv*ADC bound to agmatine.

A Se-Met derivative of *cv*ADC was produced in *E. coli*. An overnight culture of *E. coli* transfected with the pODCTM9 T142A was used to inoculate 1 L of M9 minimal media with ampicillin. At log-phase growth ($OD_{600} = 0.7$), IPTG (0.5 mM), 60 mg of selenomethionine, 100 mg each of threonine, lysine, and phenylalanine, and 50 mg each of leucine, isoleucine, and valine were added. After 13 h of induction, cells were harvested by centrifugation and lysed. Purification of the Se-Met derivative was performed as described above. Mass spectrometry analysis (Protein Chemistry Technology Center, UT-Southwestern Med. Ctr.) revealed that there were 5 Se-Met residues per monomer in *cv*ADC produced from the minimal media (data not shown).

Crystallization of cvADC. Approximate crystallization conditions were found with the Wizard I screen (Emerald Biostructures, BainBridge Island, WA). Free (Se− labeled) and agmatine-bound *cv*ADC crystallized in similar conditions. The optimized crystallization conditions are as follows: 5 μL of 20 mg/mL protein in buffer (10 mM HEPES pH 7.2, 50 mM NaCl, 1 mM DTT, 0.5 mM EDTA pH 8.0, 0.03% (v/v) Brij-35) is mixed with an equal volume of the crystallization solution (6% PEG-8000, 100 mM imidazole (pH 7.5), 200 mM calcium acetate) and set up in sitting drops at 16 °C. Agmatine-bound *cv*ADC crystals were produced by incubating 20 mg/mL protein in buffer with 5 mM agmatine at 4 °C for 1 h prior to setup of the crystallization drop. Crystals generally grew within 24 h. The final diffraction quality crystals were produced by streak-seeding procedures. Cryo-loops (Hampton Research) were used to store crystals, which were cryo-protected with a solution containing the mother liquor and 25−30% 2-methyl-2,4-pentanediol (MPD) and subsequently frozen in liquid propane.

Data Collection and Processing. Diffraction data with anomalous dispersion were collected at the Se absorption edge wavelength ($\lambda = 0.98066$) at beamline 19-BM (SBC-CAT) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). Data for agmatine-bound *cv*ADC crystals were collected at the 19-ID beamline at the same facility. The data were indexed, integrated, and scaled using HKL2000 (*20*). Data processing statistics are summarized in Table 1.

Structure Determination and Refinement. Both the free (Se-Met labeled) and agmatine-bound *cv*ADC structures crystallized in the $P2_12_12_1$ space group and similar unit cell parameters were observed for the free ($a = 116.1$ Å, $b = 116.9$ Å, $c = 269.4$ Å, $\alpha = \beta = \gamma = 90°$) and agmatine-bound ($a = 116.2$ Å, $b = 117.3$ Å, $c = 268.7$ Å, $\alpha = \beta = \gamma = 90°$) structures. The data set collected at the Se absorption edge was used for phasing by the single-wavelength anomalous dispersion method (SAD) (*21*). Selenium sites were found by direct methods in *ShelX (22)*- and 40 sites with correlation coefficients of 0.42 or better were identified. Noncrystallographic symmetry (NCS) operators were obtained with *RESOLVE* (*23*). Phases were

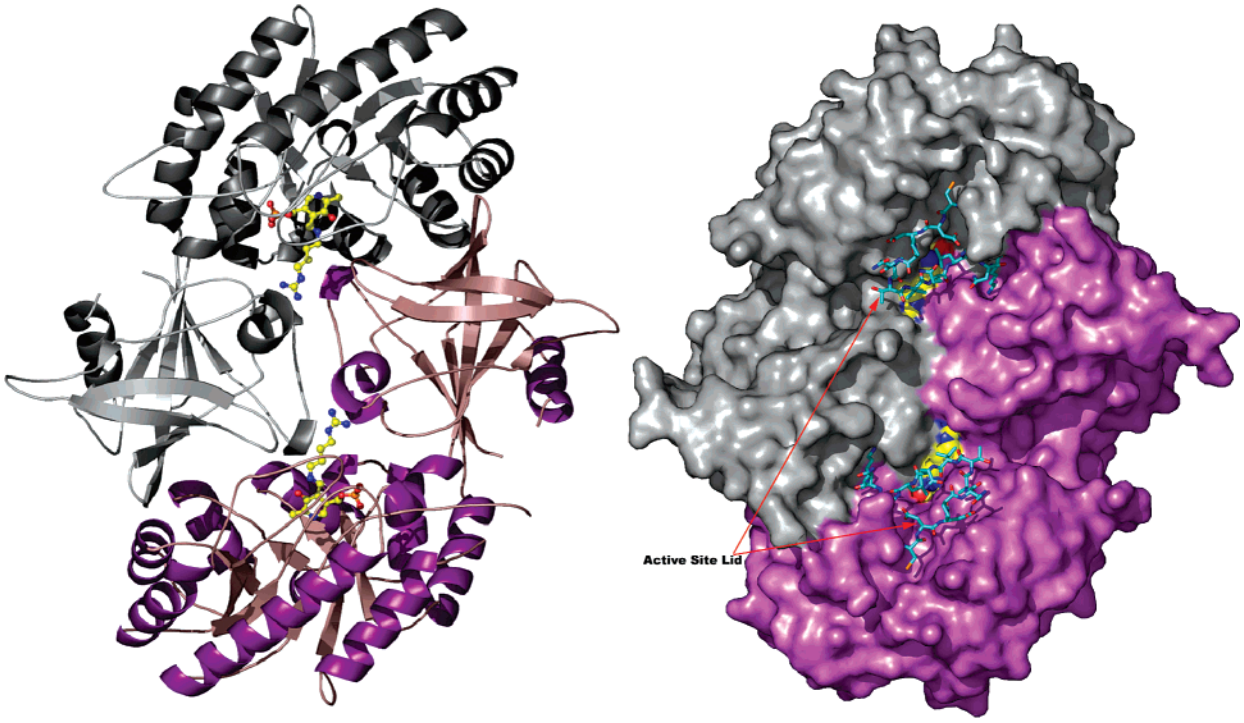

FIGURE 2: Dimeric structure of *cv*ADC. (A) The *cv*ADC structure is displayed as a ribbon (monomer A (purple), and monomer B (gray)) with the bound agmatine and the PLP cofactor displayed in yellow as ball and stick. Nitrogen atoms are displayed in blue, oxygen in red, phosphate in orange, and sulfur in yellow. (B) Surface representation of the *cv*ADC homodimer. The PLP-agmatine complex is shown in spheres and the K148 loop (RCDDPNATVQLGNK) is shown in stick representation and colored cyan.

calculated in *MLPHARE* and density modification was performed in *DM* (*24*, *25*). With the NCS operators and phases, a density-modified map was calculated. An initial model, containing 2300 of 2900 expected residues in the asymmetric unit, was built using the automated model-building program *ARP/wARP* (*26*). The agmatine-bound *cv*ADC structure was determined by using the free structure as the initial phasing model for molecular replacement with the CNS package (*27*). Both structures were rebuilt with the graphics program *Coot* (*28*) and the refinement package REFMAC5 (*29*, *30*). For the agmatine-bound structure, the Fourier difference map ($F_o - F_c$) showed readily interpretable density for the agmatine product bound to the PLP cofactor. The agmatine product was modeled into the structure and further refined. NCS restraints were released in the final rounds of refinement and waters were added as difference Fourier peaks above 3.0σ for both structures.

The free structure of *cv*ADC was refined with data to 1.95 Å and the agmatine-bound structure included data to 1.8 Å. The final models contain eight monomers in the asymmetric unit (labeled A − H), in the form of two tetramers, containing a total of 2901 residues and eight PLP molecules (or eight PLP-agmatine complexes). Dimer pairs are formed by chains A and B, C and D, E and F, and G and H. The density for the majority of residues and the PLP cofactors was easily interpretable. Residues omitted from the molecular structure of free-*cv*ADC are as follows: chain A, 229; chain C, 222−225, 229; chain D, 13−15, 223−224, 227, 229, 245−246; chain E, 14, 229; chain F, 222−223, 310; chain G, 10−11, 13−14, 143, 221−225; chain H, 222−224, 229, 311. Residues omitted from *cv*ADC-agmatine are as follows: chain A, 223−225; chain B, 14, 223, 311; chain C, 14, 222−224, 229; chain D, 223; chain E, 14, 223; chain F, 14, 222−223, 229; chain G, 14, 229, 239, 245; chain H, 10−14, 222−223, 229. The final refined models contained no peptide torsion angles in disallowed regions of the Ramachandran plot, and 90% of non-glycine and non-proline residues were within the most favored regions of the plot, as confirmed by PROCHECK (*31*). The four dimers of the agmatine-bound *cv*ADC overlay with an rmsd of 0.2 Å, although there were small differences in interpretable density. The A/B dimer is the best representative, and will be described below.

Molecular Modeling. Structures were displayed using the graphics program PyMol (*40*). All rmsd calculations were based on structural alignment by PyMol and calculated within the program. The solvent accessible surface was calculated using AREAIMOL (*29*).

RESULTS

X-ray Structure Determination of Free and Agmatine-Bound cvADC. The overall structures of *cv*ADC in both free and agmatine-bound states are similar to the previously determined eukaryotic ODC structures. The *cv*ADC structure consists of an N-terminal β/α-barrel from residues 23−261 and a C-terminal β-barrel from residues 262−372 (Figure 2). In addition, a short α-helix (residues 2−11) and β-strand (residues 18−22) are observed at the N-terminus of both the native and agmatine-bound structures. The N-terminal β-strand is part of an extended β-sheet forming interactions with two strands from the C-terminal domain. The N-terminal region of this structure has been observed in a number of other ODC structures and forms approximately the same secondary elements observed in the *cv*ADC structure (*12*, *13*, *17*, *18*).

As with all other members of this fold-type, two identical active sites are formed at the dimer interface (Figure 2A). Residues from the β/α barrel and β-barrel of one monomer as well as the β-barrel from the opposite monomer contribute

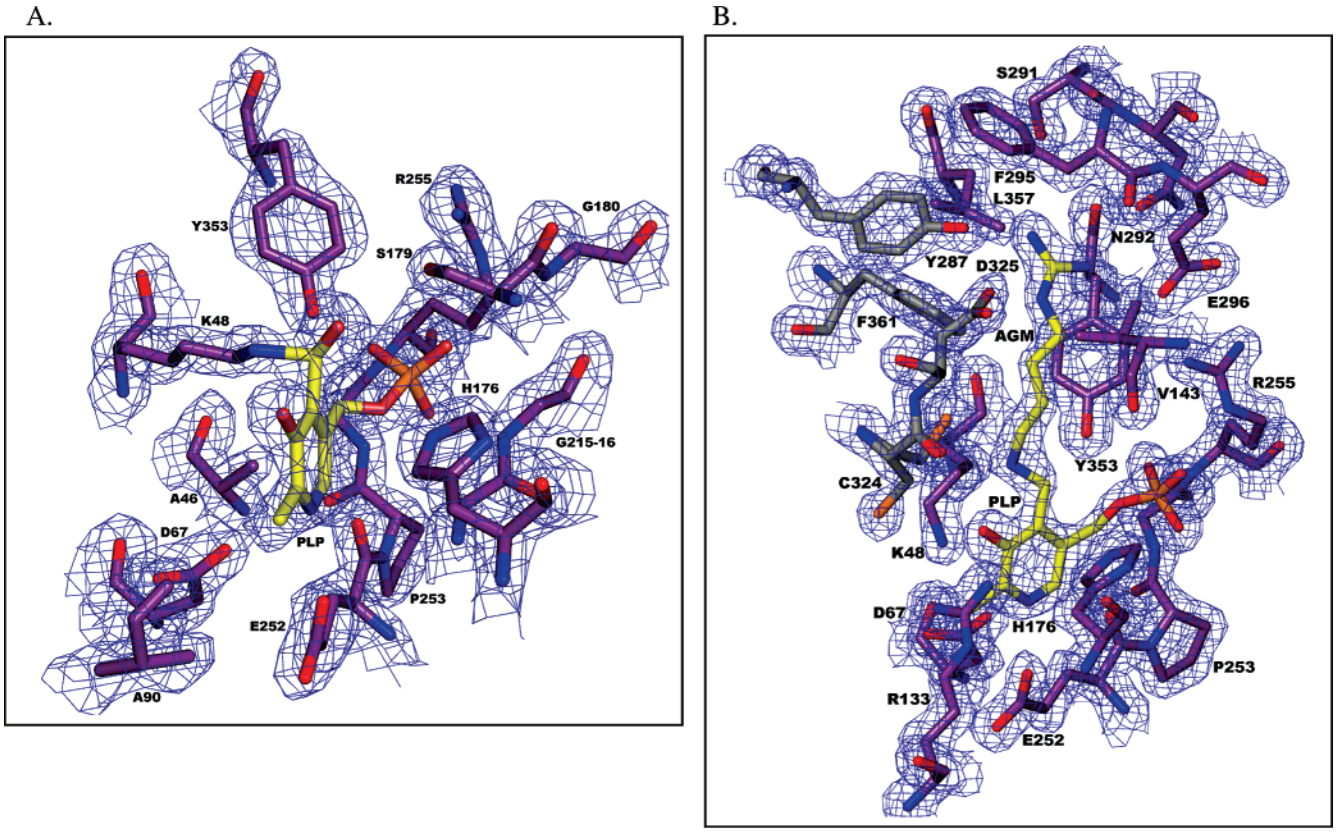

FIGURE 3: Electron density maps of the active sites of free and ligand-bound *cv*ADC: (A) Electron density map of the carbinolamine intermediate observed in the Free *cv*ADC structure. The intermediate is formed between the NZ atom of K48, the C4′ of PLP and a hydroxyl group. (B) Electron density observed in the active site of the agmatine (AGM)-bound *cv*ADC structure. Residues that form H-bond contacts (N292, E296, D325′, Y287′) and van der Waals packing (L145, Y353) with the agmatine molecule are shown. Maps ($2F_o - F_c$) are contoured at 1.0σ. The PLP molecule and agmatine are colored yellow, while protein residues are colored purple (chain A) or gray (chain B).

to the formation of the active sites. The loop consisting of residues 135−152 is observed in a down conformation and appears to function as an active site lid (Figure 2B). Two residues in the loop, *cv*ADC V143 and L145 are within about 5 Å of the bound agmatine ligand.

The Active Site of cvADC. The electron density maps of the active site showed good, interpretable density for both the native and agmatine-bound structures (Table 1 and Figure 3). In the native structure the PLP cofactor forms an interaction with K48 *cv*ADC (K69 ODC). This residue forms a Schiff's base with PLP and has been demonstrated to be a key catalytic residue accelerating the rates of Schiff base formation with substrate and decarboxylation in *tb*ODC (*32*). However, unlike previous ligand-free ODC structures (*12*−*14*), PLP is not bound to K48 *cv*ADC through a Schiff's base. Instead the density around the C4′ of the cofactor suggests a tetrahedral configuration consistent with a carbinolamine species (Figure 3A). The model for the native enzyme was built with a carbinolamine intermediate in all eight active sites of the asymmetric unit and no residual density was found in the recalculated difference Fourier map. Density for the sulfhydryl group of the active site cysteine residue C324 *cv*ADC (C360 in eukaryotic ODC), which had been proposed to function as a general base in the reaction (*18*), indicates that the cysteine side-chain exists in dual conformations within each of the eight active sites in both the native and agmatine-bound structures. Thus, we modeled C324 *cv*ADC as dual conformations with the sulfhydryls positioned 145° apart. The homologous residue in ODCs has been observed to occupy both of these conformations (*12*−*14*, *17*, *18*, *33*).

In the active site of the agmatine-bound structure good density was observed for the agmatine in all four dimers in the asymmetric unit (Figure 3B). The side-chain of residue K48 *cv*ADC has shifted 5 Å away from the cofactor to form an interaction with D67 *cv*ADC, thus making room for the bound product (Figure 4). The electron density is consistent with a Schiff's base structure between N1 of agmatine and C4′of PLP. The aliphatic portion of agmatine packs against Y353, F361 and V143, while the guanidinium moiety is within H-bonding distance of the side-chains of E296, N292, D325, and Y287.

The K148 cvADC (K169 ODC) loop functions as an active site lid. To investigate the structural differences between free *cv*ADC and eukaryotic ODC, the Cα atoms of the N-terminal β/α-barrel domain of *cv*ADC (residues 23−261), *T. brucei* ODC (1QU4, residues 43−283), and Human ODC (1D7K, residues 45−283) were aligned. The overall rmsd between all three structures at the level of the monomer was small ranging from 1.2 (*tb*ODC) to 1.4 (*h*ODC) Å (Figure 5 and 6). However, regions of the structure displayed rmsd values above 2.5 Å. Notably significant changes are observed in the K148 *cv*ADC (K169 ODC) loop which is formed by residues 135 − 152 *cv*ADC at the end of the β_6-strand (Figure 5). With the exception of the human ODC structure

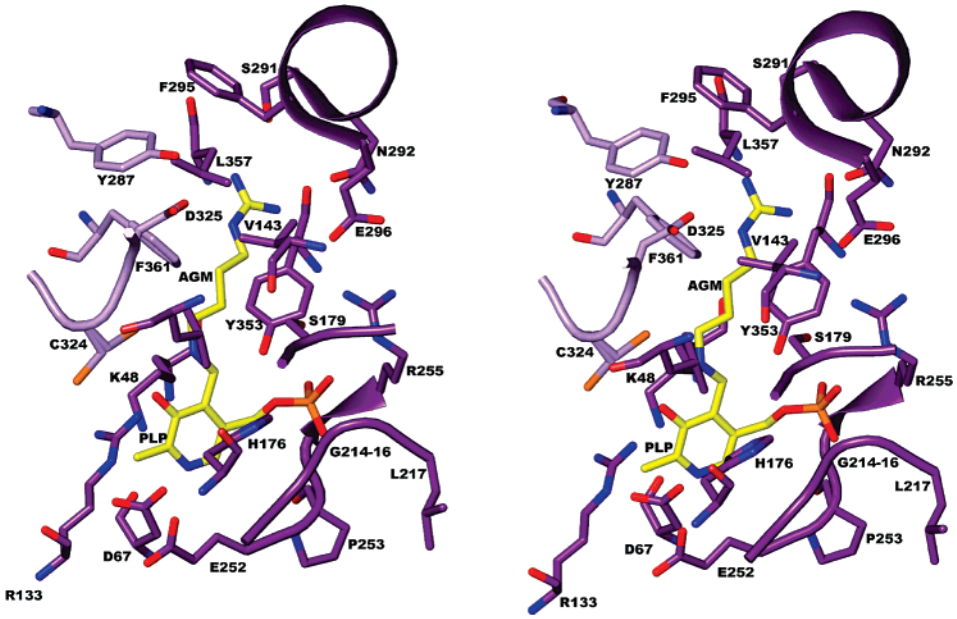

FIGURE 4: Stereopair of the structure of the ligand-bound active site of *cv*ADC. The 5 Å shell of residues surrounding the PLP-agmatine complex is shown in stick representation. The cofactor-product complex is shown in yellow and residues are shown in purple (chain A) or light purple (chain B).

(*13*), portions of this loop (residues D158−R165 ODC) have typically been disordered in the eukaryotic ODC structures that have been solved (e.g., the mouse ODC and the *T. brucei* ODC structures (*12*, *14*, *17*, *18*, *33*)). Density for the entire loop was observed in the human ODC structure where it assumes an open, solvent accessible, conformation that points away from the actives site (Figure 5). In contrast, in both the native and agmatine-bound structures of *cv*ADC the loop has swung 9 Å toward the active site, pivoting around residues P139 and N147, and it is observed for the first time in a closed conformation. In this conformation, the bound ligands are sequestered from solvent (Figure 2B and 5). In the agmatine-bound *cv*ADC structure there is no solvent-accessible surface on agmatine, while 2.0 $Å^2$ is accessible on the adjacent PLP molecule. In contrast, the putrescine molecule in the ligand-bound *tb*ODC structure has 21 $Å^2$ of solvent exposed surface compared to 1.3 $Å^2$ for its PLP molecule. Both conformations of the loop are likely to be available to both enzymes, since there appears to be no route for substrate entry in *cv*ADC. Thus, these data suggest that this loop functions as a mobile active site lid in this family of enzymes, controlling access to the substrate binding site during the catalytic cycle. In addition to this change the Cα atoms of residues S179−N183 *cv*ADC, which cover the phosphate of PLP, have shifted 3 Å toward the phosphate moiety in *cv*ADC. This movement allows the formation of a new H-bond interaction between the backbone of S179 and V143 in *cv*ADC (Figure 4). Additional backbone H-bond interactions between residues G178−G180 and T142−V143 *cv*ADC seem to promote the shift of the S179−N183 *cv*ADC region and add to the stabilization of the K148 *cv*ADC loop.

The 3_{10}-helix is a key determinant of substrate specificity.

In order to delineate the determinants of substrate specificity in *cv*ADC, agmatine-bound *cv*ADC and putrescine-bound *tb*ODC (1F3T) were compared (Figure 6). The subunits aligned by their N-terminal β/α barrel domains overlaid with an rmsd of 1.2 Å for the monomers. Very little or no domain rotation is observed. Residues at the bottom of the pocket that interact with the PLP cofactor superimpose precisely between the two structures, including the positions of D67, K48, and E252-R255 in *cv*ADC. The β-barrel C-terminal domains of the opposite subunit are also closely aligned, including the catalytic base and substrate recognition residues, C324′ and D325′ *cv*ADC. Part of the C-terminal β-barrels of the primary subunits of *cv*ADC and *tb*ODC also superimpose well, including the substrate binding residues Y287, Y353, and F360 of *cv*ADC. However, significant differences in the structures are apparent, providing new insight into the structural basis of specificity in this family of enzymes. In both the *cv*ADC and ODC structures residues S291-E296 *cv*ADC (S325−D332 ODC) form a short 3_{10}-helix at the distal portion of the substrate-binding pocket. The helix in the *tb*ODC structure begins 2 residues prior to that observed in *cv*ADC, however, the helix ends at the same position (E296/D332) in both (Figure 1). This residue, E296 *cv*ADC (D332 ODC) is positioned from the helix to form hydrogen bond contacts with the guanidinium of agmatine in *cv*ADC (the δ-amino of putrescine in ODC). In the *cv*ADC structure the helix is translated 2.1 Å away from the cofactor when compared to *tb*ODC (as measured between C4′ of PLP and the Cα of E296/D332), effectively enlarging the pocket to accommodate the arginine substrate (Figure 6). The shortening of the helix may contribute to this shift. However despite this translation the longer aliphatic portion of E296 *cv*ADC accommodates this more distal position of the helix without changing the final position of the carboxylate portion of the side chain, which remains similarly placed to that of D332 ODC. Thus E296 *cv*ADC is able to retain key contacts with the remainder of the active site, including a H-bond interaction with R255 *cv*ADC (R277 ODC). This residue forms interactions with the phosphate moiety of PLP, thus its precise positioning in the active site is necessary for the function of the enzyme.

In addition to the repositioning of the 3_{10}-helix the structural alignment identifies several other amino acid substitutions in the distal region of the active site that may

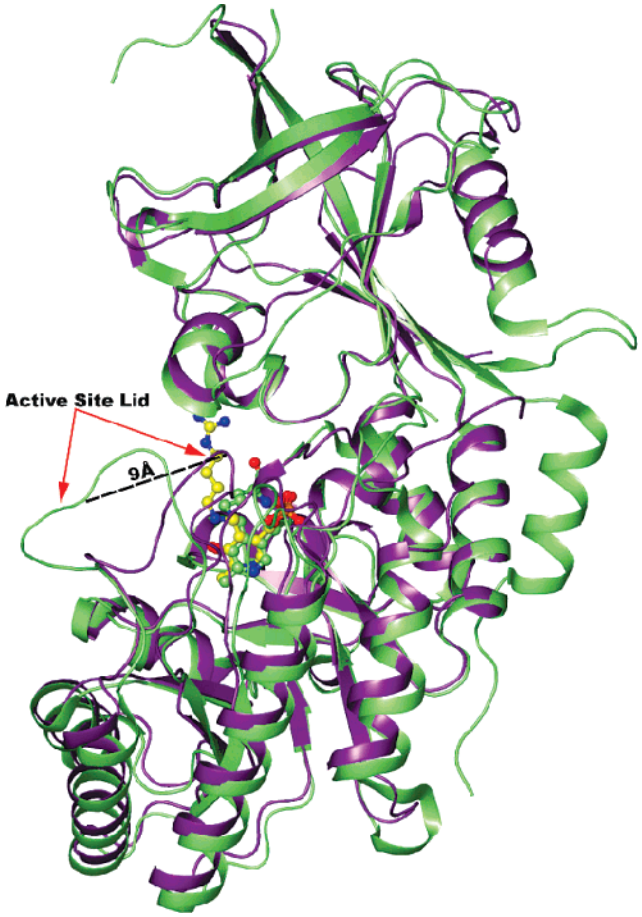

FIGURE 5: Multiple conformations of the K148 *cv*ADC loop. Alignment of the structures of native *cv*ADC (purple) and human ODC (1D7K; light green). Structures were aligned using the N-terminal residues of *cv*ADC (residues 23−261) and human ODC (45−283) in the β/α-barrel domain. The overall rmsd between the structures is 1.4 Å as calculated for the monomer. The secondary structure is represented in ribbon while the PLP molecules are colored yellow for both structures. The residues at the hinge positions in the K148 *cv*ADC (K169 *h*ODC) loops are P139 and N147.

contribute to the specificity change (Table 2). The more noteworthy substitutions in the region include C328 ODC (N292 *cv*ADC) and G393 TBODC (L357 *cv*ADC). The side-chain of C328 ODC is turned out of the active site and forms an H-bond interaction with the backbone carbonyl of S325. In contrast, the side-chain of N292 *cv*ADC is pointed toward the active side and the carbonyl oxygen forms a H-bond with the guanidinium moiety of agmatine (3.3 Å). L357′ *cv*ADC (contributed to the active site from the opposite subunit) does not form a direct interaction with the agmatine ligand (distance 5.67 Å), however the increased bulk of the side chain in comparison to ODC (Gly or Ala at this position), may play a role in positioning of the 3_{10}-helix. The side-chain of L357′ is within van der Waals contact of the side-chain of S291 *cv*ADC of the 3_{10}-helix. This contact may help stabilize the position of the 3_{10}-helix in order to maintain the proper volume of the active site to accommodate arginine.

Structural Comparison of cvADC and DAPDCs. To further probe the structural basis of substrate specificity the active site of *cv*ADC was compared to the three available structures of DAPDC. The N-terminal β/α barrel domain of *cv*ADC with *Mycobacterium tuberculosis* DAPDC (residues 48−305, 1HKV), *E. coli* DAPDC (residues 33−271, 1KO0) and *M. jannaschii* DAPDC (residues 51−312, 1TWI) were superimposed, and the rmsd for the monomer ranged from 2.3 to 2.5 Å (Figure 7). The significantly greater divergence in the structures is consistent with the fact that DAPDC and *cv*ADC share less than 20% sequence identity. The differences between *cv*ADC and the DAPDC structures arise from differences distributed throughout both the N- and C-terminal domains. Unlike for the comparison to alanine racemase (*14*) no apparent domain rotation has taken place between DAPDC and *cv*ADC. Despite the higher rmsd values for the overall structures, the active site of DAPDC is similar to *cv*ADC (or ODC). As with the *cv*ADC and *tb*ODC comparison, the position of the 3_{10}-helix is substantially different between *cv*ADC and the DAPDC structures, and indeed it assumes a different conformation in all three structures. The distance between C4′of PLP and the helix (Cα of D332 *tb*ODC, E296 *cv*ADC and E348 *mj*DAPDC) is 10.0 Å for *tb*ODC, 12.1 Å for *cv*ADC, 15.3 Å for *mj*DAPDC.

Additionally amino acid substitutions in the helix provide a diversity of interactions for the various substrates that are accommodated in the different structures. The key interactions with ligand for ODC and *cv*ADC include one common interaction D332 *tb*ODC (E296 *cv*ADC), while agmatine in *cv*ADC makes an additional interaction with N292 *cv*ADC positioned at the start of the 3_{10}-helix. For DAPDC the residue at the start of the helix has been substituted with R343 *mj*DAPDC, which forms a salt bridge with the carboxylate of the bound lysine. Additionally the hydroxyl group of Y347 *mj*DAPDC forms H-bond interactions with the amine group of lysine, while the residue equivalent to D332 *tb*ODC faces out of the pocket. These data support the conclusion that the 3_{10}-helix is a major determinant defining substrate specificity in the β/α-barrel fold decarboxylase family. Within the context of this conserved fold, the protein is designed to be flexible in the positioning and amino acid sequence of the 3_{10}-helix, providing a mechanism to evolve different substrate preferences without large structural changes.

DISCUSSION

The β/α-fold type decarboxylases are composed of a large, diverse family with a broad range of substrate preferences for basic amino acids (*8*, *9*). The discovery of *cv*ADC, an arginine decarboxylase closely related to the eukaryotic ODCs, provided a unique opportunity to gain insight into the evolution of enzyme specificity in this family (*11*). The structure of *cv*ADC identifies the E296 *cv*ADC (D332 ODC) 3_{10}-helix as the key determinant of substrate specificity. Comparison of *cv*ADC with both eukaryotic ODC and bacterial DAPDC structures shows that the 3_{10}-helix assumes a different orientation in each structure (Figures 5−7). Previous structures of ODC and DAPDC had demonstrated that residues within this helix formed key interactions with substrate analogs bound to the active sites of these enzymes, and suggested that the distance between these residues and the PLP cofactor served as a molecular ruler guiding substrate preference (*14*−*16*). However given the divergence of these two structures the importance of the position of the helix was unclear. In contrast the *cv*ADC and ODC structures are very similar. Most of the active site, including the catalytic residues, and those that interact with PLP superimpose

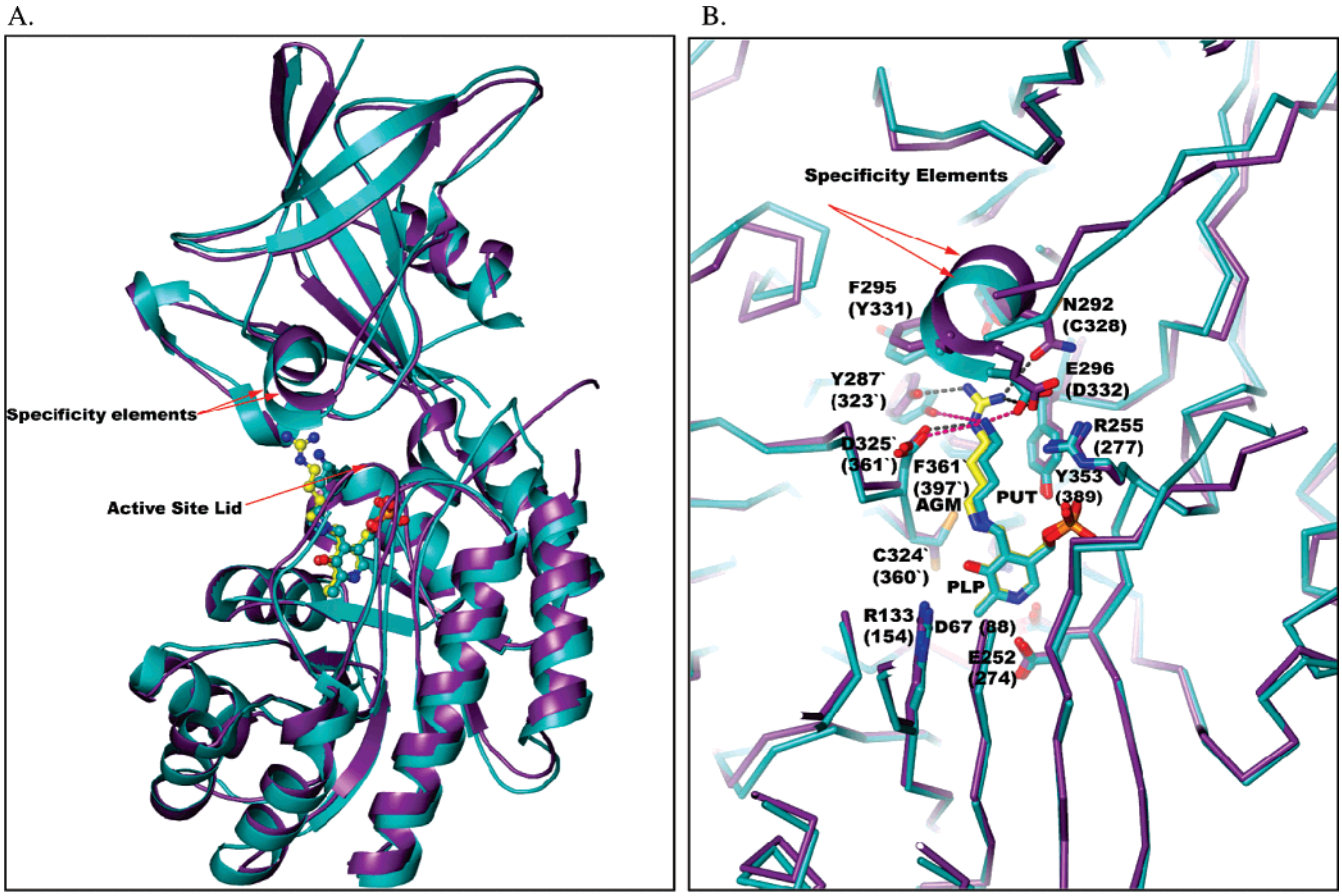

FIGURE 6: Comparison of product-bound structures of *cv*ADC to *T. brucei* ODC: **A**. Ribbon diagram of the superimposed monomers of *cv*ADC (purple) bound to agmatine and the *tb*ODC bound with putrescine (1F3T; teal). The structures were aligned using the N-terminal residues in the β/α-barrel domain (*cv*ADC 23−261 and *tb*ODC 43−283). The overall rmsd between the structures is 1.2 Å for the monomer. **B.** Structural alignment of the active sites of *cv*ADC and *tb*ODC. Residues are numbered according to the *cv*ADC sequence with the equivalent ODC residue in parentheses. Residue numbers with a prime notation come from the opposite monomer. The PLP cofactor/agmatine complex is shown in yellow for *cv*ADC and PLP/putrescine is displayed in teal for tbODC.

Table 2. Amino Acid Variations Observed within 5 Å of Product in Active Site of *cv*ADC Relative to *tb*ODC and *mj*DAPDC and Where the Primary Substrate for Each Enzyme Is Also Listed[a]

*cv*ADC L-arginine	*tb*ODC L-ornithine	*mj*DAPDC L-diaminopimelate
V143	C164	G188
L217	F238	L265
S291	N327	R343
N292	C328	P344
F295	Y331	Y347
E296	D332	E348
L357	G393	M405

[a] The coordinates have been deposited in the protein data base: unliganded *cv*ADC, 2NV9; *cv*ADC complexed to agmatine, 2NVA.

exactly, thus the repositioning of the 3_{10}-helix within this context provides a clear correlation between the helix conformation and the altered substrate preference.

Flexibility in the position of the 3_{10}-helix provides a mechanism to adjust the volume in the pocket to accommodate different substrates, however changes in amino acid composition are required to build the differently structured active sites (Table 2 and Figure 1). In *cv*ADC E296 has replaced D332 ODC, and the longer side chain arm allows the helix to be positioned further back into the pocket, while still placing the carboxylate portion of E296 in the same position as D332 ODC. This allows contact with the guanidinium portion of the arginine substrate, while importantly preserving a salt bridge with R255 *cv*ADC (R277 ODC). This interaction is likely to be important for orienting this residue, which was previously shown in *tb*ODC to be essential for high affinity PLP binding through interaction with the phosphate (*34*). The structure also explains why the D332E *tb*ODC and E296D *cv*ADC mutants were inactive (*11*). This single point mutation is not sufficient to encode the repositioning of the 3_{10}-helix, nor the required substrate interactions. In addition to the amino acid change at this position, C328 *tb*ODC is replaced by N292 *cv*ADC. The side chain of this residue is reoriented toward the active site and makes a hydrogen bond with the bound agmatine, suggesting that this interaction is also an important component of the reorganization of the pocket to recognize the arginine substrate. Additionally, the structure of *cv*ADC suggests two mechanisms that may contribute to differences in the positioning of the helix. First the length of the helix differs between homologs in the family. Second, substitutions of second shell residues that interact with the helix may also contribute to the different conformations. L357′ *cv*ADC packs against the side chain of S291 *cv*ADC in the 3_{10}-helix. This contrasts with ODC where either Ala or Gly-393 are found at this position and form an interaction with the side-chain of N327 ODC in the 3_{10}-helix. Thus the sequence of the 3_{10}-helix clearly coevolves multiple changes that play a

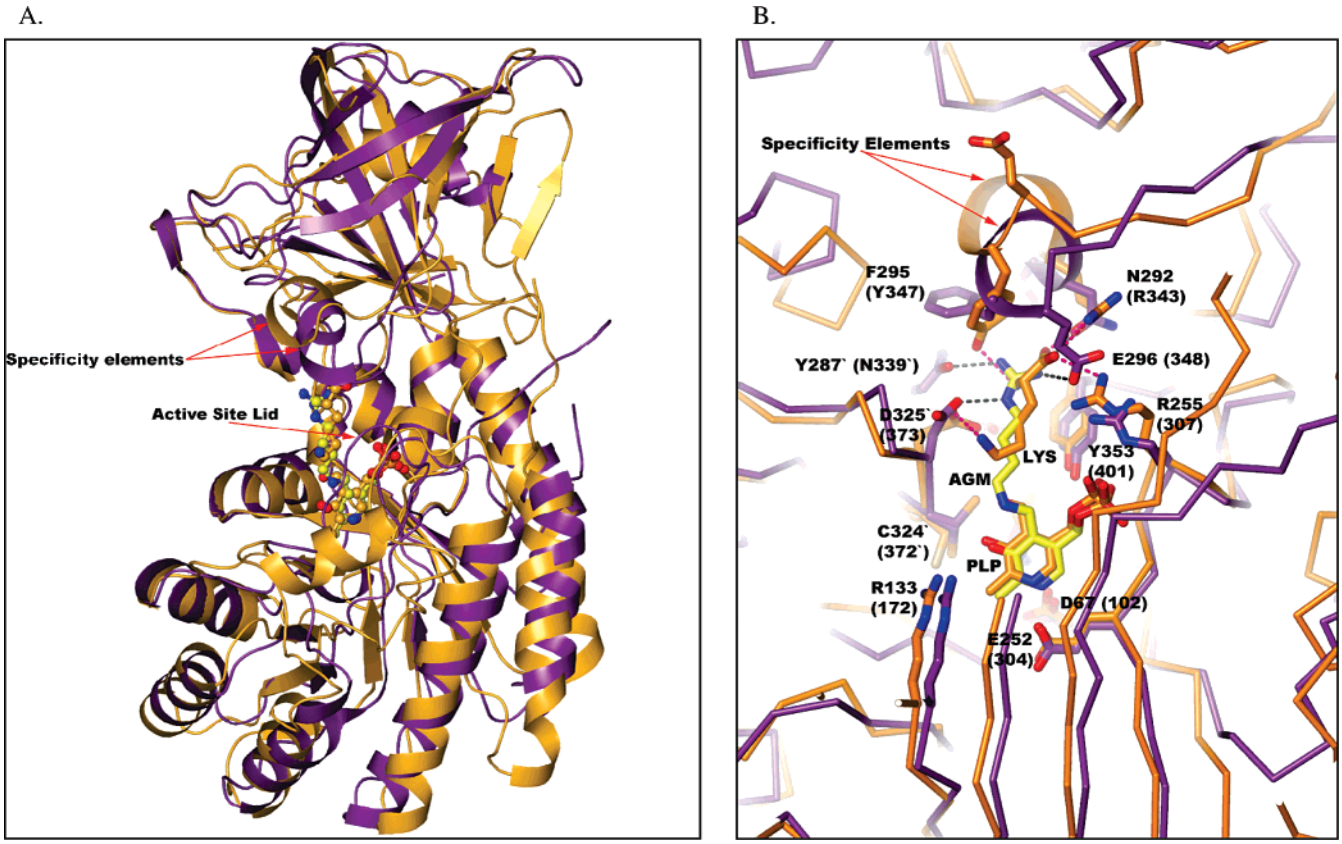

FIGURE 7: Comparison of the product-bound structures of *cv*ADC and *M. jannaschii* DAPDC. (A) Ribbon diagram of the monomeric structures of *cv*ADC (purple) and *mj*DAPDC (1TWI; orange) for the alignment of the N-terminal domain of *cv*ADC (residues 23−261) and *mj*DAPDC (residues 51−312). (B) Comparison of the active sites of *cv*ADC and *mj*DAPDC. The overall rmsd between the structures is 2.4 Å for the monomer. The PLP cofactor and agmatine from *cv*ADC are shown in yellow sticks and the PLP and lysine molecule from *mj*DAPDC are shown in orange sticks.

role in the final orientation and chemistry of the helix in the structure.

The substrate preference for DAPDC can also be understood in the context of changes in the amino acid composition of the 3_{10}-helix. Accommodation of the dicarboxylate substrate diaminopimelate requires the pocket to accommodate interactions with a carboxylate on the ligand as well as the amino group. R343 *mj*DAPDC is positioned in the active site of DAPDC and forms interactions with the carboxylate of azelic acid (1TUF) or lysine (1TWI) bound in the structure. This position is a neutral residue in the ODC and *cv*ADC structures, and the side chain points away from the active site in both. Thus this interaction in the DAPDC structure, which was not previously recognized (*15*, *16*), is likely to play an important role in substrate recognition. Likewise residue Y347 mjDAPDC is positioned differently than the equivalent residue in ODC, and interacts with the ϵ-amino group of ligand. Y331 ODC is not oriented into the active site but instead is at the 2-fold axis of the dimer interface. Thus, evolutionary changes that alter both the position of the helix, and its amino acid sequence composition, provide the flexibility to accommodate different substrates into this structural fold.

Finally, the finding that the active site is sequestered by a mobile-loop (K148 *cv*ADC-loop) in the *cv*ADC structure has provided novel insight into the reaction mechanism of the β/α fold-type decarboxylases. Kinetic and structural studies of mutations introduced at residues in the dimer interface of *tb*ODC previously suggested that dynamic motion of the analogous loop (K169 *tb*ODC loop) played a role in the catalytic cycle (*33*). However, multiple conformations of the loop have never previously been observed in a structure. Indeed the K169 ODC loop is disordered in all ODC structures except that of human ODC, where the loop was observed in an open conformation, distant from the active site (Figure 5). In contrast for the *cv*ADC structure the K148 *cv*ADC (K169 ODC) loop is observed in the down or closed conformation, positioned as a lid over the active site. In this conformation it makes contacts with key active site residues (e.g., S179 *cv*ADC), and it effectively sequesters the active site from solvent. The conformation of a second active site loop formed by S179−N183 *cv*ADC is also repositioned in the closed conformation structure, shortening the H-bond distance between S179 and the phosphate of PLP.

The sequestering of the substrate from solvent in the "lid down" conformation implies that the loop must be in the open position for substrate to access the active site and for product to be released. Previous investigations have demonstrated that product release is the rate-limiting step in *tb*ODC catalysis (*35*), suggesting that the rate of lid movement may in fact be the rate-limiting step. Additionally, the ligand-free *cv*ADC reveals a tetrahedral carbinolamine adduct between the C4′ of PLP and K48 *cv*ADC, rather than a Schiff's base. Although the carbinolamine has been observed in mutant ODC structures, it was unclear if this species could form in the wild-type enzyme (*33*). The observation of the carbinolamine intermediate in *cv*ADC suggests that the closed or loop down conformation may promote formation

of the tetrahedral structure, perhaps because it mimics transition state structures along the reaction coordinate. In support of a catalytic role for the loop, mutation of K169 *tb*ODC significantly reduces the catalytic efficiency of the enzyme (*36*).

The data taken together suggest that both closed and open conformations are used in all Group IV decarboxylases to accommodate sequestering of the substrate from solvent, substrate entry, and product release. Dynamic loops have previously been reported for enzymes in this α/β fold (e.g., triosephosphate isomerase (TIM) and tryptophan synthase (TS), inosine monophosphate dehydrogenase (IMPDH)) (*37*−*39*). It is interesting that the mobile loop in *cv*ADC occupies the same position in the TIM barrel as those observed in TIM and TS. Thus, apparently a mobile loop that controls substrate access appears to be a conserved structural feature of the β/α barrel fold-type enzymes.

ACKNOWLEDGMENT

We thank Jeffrey Baldwin for help in purification of *cv*ADC and site-directed mutagenesis, Mischa Machius and Chad Brautigam for aid in collection of data sets from the synchrotron, Andrej Joachimiak, Norma Duke, and members of the SBC-CAT for assistance with data collection, and Dominika Borek for assistance with scaling of the data and phasing procedures.

REFERENCES

1. Pegg, A. E. (2006) Regulation of ornithine decarboxylase, *J. Biol. Chem. 281*, 14529−32.
2. Hutton, C. A., Southwood, T. J., and Turner, J. J. (2003) Inhibitors of lysine biosynthesis as antibacterial agents, *Mini Rev. Med. Chem. 3*, 115−27.
3. Schousboe, A., and Waagepetersen, H. S. (2006) Glial modulation of GABAergic and glutamat ergic neurotransmission, *Curr. Top Med. Chem. 6*, 929−34.
4. Fries, D., and Fairlamb, A. (2003) Antiprotozoal agents, in *Burger's Medicinal Chemistry and Drug Discovery* (Abraham, D., Ed.) pp 1033−1087, John Wiley & Sons, Inc., New York.
5. Gerner, E., and Meyskens, F. (2004) Polyamines and cancer: old molecules, new understanding, *Nature Rev. Cancer 4*, 781−792.
6. Marton, L., and Pegg, A. (1995) Polyamines as targets for therapeutic intervention, *Annu. Rev. Pharmacol. Toxicol. 35*, 55−91.
7. Wang, C. (1995) Molecular mechanisms and therapeutic approaches to the treatment of African Trypanosomiasis, *Annu. Rev. Pharmacol. Toxicol. 35*, 93−127.
8. Grishin, N., and Phillips, M. (1994) The subunit interfaces of oligomeric enzymes are conserved to a similar extent to the overall protein sequences, *Protein Sci. 3*, 2455−2458.
9. Sandmeier, E., Hale, T., and Christen, P. (1994) Multiple evolutionary origin of pyridoxal-5′-phosphate-dependent amino acid decarboxylases, *Eur. J. Biochem. 221*, 997−1002.
10. Morehead, T. A., Gurnon, J. R., Adams, B., Nickerson, K. W., Fitzgerald, L. A., and Van Etten, J. L. (2002) Ornithine decarboxylase encoded by chlorella virus PBCV-1, *Virology 301*, 165−75.
11. Shah, R., Coleman, C. S., Mir, K., Baldwin, J., Van Etten, J. L., Grishin, N. V., Pegg, A. E., Stanley, B. A., and Phillips, M. A. (2004) Paramecium bursaria chlorella virus-1 encodes an unusual arginine decarboxylase that is a close homolog of eukaryotic ornithine decarboxylases, *J. Biol. Chem. 279*, 35760−7.
12. Kern, A., Oliveira, M., Coffino, P., and Hackert, M. (1999) Structure of mammalian ornithine decarboxylase at 1.6 Å resolution: Stereochemical implications of PLP-dependent amino acid decarboxylase, *Structure 7*, 567−581.
13. Almrud, J., Oliveira, M., Grishin, N., Phillips, M., and Hackert, M. (2000) Crystal structure of human ornithine decarboxylase at 2.1 Å resolution: structural perspectives of antizyme binding, *J. Mol. Biol. 295*, 7−16.
14. Grishin, N., Osterman, A., Brooks, H., Phillips, M., and Goldsmith, E. (1999) The X-ray structure of ornithine decarboxylase from *Trypanosoma brucei*: the native structure and the structure in complex with α-difluoromethylornithine, *Biochemistry 38*, 15174−15184.
15. Gokulan, K., Rupp, B., Pavelka, M. S., Jr., Jacobs, W. R., Jr., and Sacchettini, J. C. (2003) Crystal structure of Mycobacterium tuberculosis diaminopimelate decarboxylase, an essential enzyme in bacterial lysine biosynthesis, *J. Biol. Chem. 278*, 18588−96.
16. Ray, S. S., Bonanno, J. B., Rajashankar, K. R., Pinho, M. G., He, G., De Lencastre, H., Tomasz, A., and Burley, S. K. (2002) Cocrystal structures of diaminopimelate decarboxylase: mechanism, evolution, and inhibition of an antibiotic resistance accessory factor, *Structure 10*, 1499−508.
17. Jackson, L., Goldsmith, E., and Phillips, M. (2003) X-ray structure determination of T. brucei ornithine decarboxylase bound to D-ornithine and to G418: Insights into substrate binding and ODC conformational flexibility, *J. Biol. Chem. 278*, 22037−22043.
18. Jackson, L., Brooks, H., Osterman, A., Goldsmith, E., and Phillips, M. (2000) Altering the Reaction Specificity of Eukaryotic Ornithine Decarboxylase, *Biochemistry 39*, 11247−11257.
19. Kaiser, A., Vollmert, M., Tholl, D., Graves, M., Gurnon, J., Xing, W., Lisec, A., Nicerson, K., and Van Etten, J. (1999) Chlorella virus PBCV-1 encodes a functional homospermidine synthase, *Virology 263*, 254−262.
20. Otwinoski, A., and Minor, W. (1997) *Meth. Enzymol. 276*, 307−326.
21. Dodson, E. (2003) Is it jolly SAD? *Acta Crystallogr. D, Biol. Crystallogr. 59*, 1958−65.
22. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD, *Acta Crystallogr. D, Biol. Crystallogr. 58*, 1772−9.
23. Terwilliger, T. C. (2000) Maximum-likelihood density modification, *Acta Crystallogr. D, Biol. Crystallogr. 56*, 965−72.
24. Cowtan, K. D., and Zhang, K. Y. (1999) Density modification for macromolecular phase improvement, *Prog. Biophys. Mol. Biol. 72*, 245−70.
25. Otwinoski, Z. (1991) Maximum-likelihood density modification, *Acta Crystallogr. D, Biol. Crystallogr. 56*, 965−972.
26. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) ARP/wARP and automatic interpretation of protein electron density maps, *Meth. Enzymol. 374*, 229−44.
27. Brunger, 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) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D, Biol. Crystallogr. 54*, 905−21.
28. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. D, Biol. Crystallogr. 60*, 2126−32.
29. Collaborative Computing Project, N. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
30. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D, Biol. Crystallogr. 53*, 240−55.
31. Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993) Procheck: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
32. Osterman, A., Brooks, H., Jackson, L., Abbott, J., and Phillips, M. (1999) Lys-69 plays a key role in catalysis by *T. brucei* ornithine decarboxylase through acceleration of the substrate binding, decarboxylation and product release steps, *Biochemistry 38*, 11814−11826.
33. Jackson, L. K., Baldwin, J., Akella, R., Goldsmith, E. J., and Phillips, M. A. (2004) Multiple active site conformations revealed by distant site mutation in ornithine decarboxylase, *Biochemistry 43*, 12990−9.
34. Osterman, A., Brooks, H., Rizo, J., and Phillips, M. (1997) The role of Arg-277 in the binding of pyridoxal 5′-phosphate to *Trypanosoma brucei* ornithine decarboxlyase, *Biochemistry 36*, 4558−4567.
35. Brooks, H., and Phillips, M. (1997) Characterization of the reaction mechanism of Trypanosoma brucei ornithine decarboxylase by

multiwavelength stopped-flow spectroscopy, *Biochemistry 36*, 15147−15155.
36. Myers, D., Jackson, L., Ipe, V., Murphy, G., and Phillips, M. (2001) Long-range interactions in the dimer interface of ornithine decarboxylase are important for enzyme function, *Biochemistry 40*, 13230−13236.
37. Hedstrom, L., and Gan, L. (2006) IMP dehydrogenase: structural schizophrenia and an unusual base, *Curr. Opin. Chem. Biol. 10*, 520−5.
38. Sampson, N. S., and Knowles, J. R. (1992) Segmental motion in catalysis: investigation of a hydrogen bond critical for loop closure in the reaction of triosephosphate isomerase, *Biochemistry 31*, 8488−94.
39. Brzovic, P. S., Hyde, C. C., Miles, E. W., and Dunn, M. F. (1993) Characterization of the functional role of a flexible loop in the alpha-subunit of tryptophan synthase from Salmonella typhimurium by rapid-scanning, stopped-flow spectroscopy and site-directed mutagenesis, *Biochemistry 32*, 10404−13.
40. DeLano, W. L. The PyMOL Molecular Graphics System (2002); http://www.pymol.org/.

BI6023447