

STRUCTURE AND ACTIVITY OF THE PARAMECIUM BURSARIA CHLORELLA  
VIRUS ARGININE DECARBOXYLASE

APPROVED BY SUPERVISORY COMMITTEE

---

Margaret A. Phillips, Ph.D. ( Supervising Professor )

---

Nikolai V. Grishin, Ph.D. (Committee Chair)

---

Elizabeth J. Goldsmith, Ph.D.

---

Richard Auchus, M.D., Ph.D.

This thesis is dedicated to my family.

STRUCTURE AND ACTIVITY OF THE PARAMECIUM BURSARIA CHLORELLA  
VIRUS ARGININE DECARBOXYLASE

by

RAHUL HARSHAD SHAH

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences  
The University of Texas Southwestern Medical Center at Dallas  
In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY  
The University of Texas Southwestern Medical Center at Dallas  
Dallas, Texas  
December, 2007

Copyright  
by  
RAHUL HARSHAD SHAH, 2007  
All Rights Reserved



## **Acknowledgements**

There are a number of people to thank for making my graduate school experience so rewarding. First, I would like to express my deepest gratitude to my mentor, Dr. Margaret A. Phillips. Her vast knowledge of enzymology and her infectious passion for science was an indispensable asset that helped me get through my graduate career. I could not imagine a more supportive and encouraging mentor from whom I could learn to think critically about science.

I would also like to thank all past and current members of the Phillips laboratory. Their friendliness and enthusiasm for science truly encouraged me to become a better scientist. In particular, I would like to thank Dr. Jeffrey Baldwin for teaching me how to purify proteins and conduct enzyme assays. I would also like to give a special thanks to Dr. Laurie Jackson for the tremendous amount of knowledge she communicated to me regarding protein crystallization and structure determination. Dr. Radha Akella deserves special mention for aiding me with structure determination and teaching me a great deal about the theoretical and practical aspects of x-ray crystallography.

I would like to thank my dissertation committee members Dr. Nick Grishin, Dr. Elizabeth Goldsmith and Dr. Richard Auchus for their advice, support and encouragement over the course of my graduate studies.

Finally, I would like to thank my family and friends for all their support and encouragement both before and throughout my graduate career. It is certain that without my parent's, Harshad and Raksha Shah, positive and supportive attitude I could not have completed my graduate school education. I would also like to thank my wife, Shilpa, whose love and support are especially important to me. Finally, I would like to thank my sisters, Roshni and Reshma Shah, for providing encouraging words and support.

STRUCTURE AND ACTIVITY OF THE PARAMECIUM BURSARIA CHLORELLA  
VIRUS ARGININE DECARBOXYLASE

RAHUL HARSHAD SHAH, Ph.D.  
The University of Texas Southwestern Medical Center at Dallas, 2007

Supervising Professor: Margaret A. Phillips, Ph.D.

The substrate specificity of enzymes has been studied with keen interest for many years. An understanding of the structural basis of specificity may help to explain how enzymes have evolved such enormous rates of catalysis above uncatalyzed reactions. By understanding how enzymes coordinate residues within and outside of the active site, enzyme engineering efforts may be aided. Finally, in the case of enzymes as drug targets, the structural basis of enzyme-substrate interactions may facilitate medicinal chemistry efforts to modulate the activity of a targeted enzyme.

The investigations presented here focus on a homolog of ornithine decarboxylase, a proven drug target in the treatment of parasitic infections. The *Paramecium bursaria Chlorella* virus arginine decarboxylase is a member of the Group IV pyridoxal-5'-phosphate-dependent decarboxylase family. The enzyme

is a close homolog of eukaryotic ornithine decarboxylases and is only distantly related to bacterial arginine decarboxylases. The goals of my dissertation project were: 1) to determine the substrate preference of the *Paramecium bursaria* Chlorella virus arginine decarboxylase, 2) to determine the structural basis of specificity of this enzyme, 3) to attempt to define the amino acid determinants of specificity for both ornithine and arginine decarboxylases, 4) to provide a more thorough understanding of the catalytic cycle of these decarboxylases. To accomplish these goals of my thesis I employed biochemical and biophysical techniques ranging from HPLC-based analysis of reaction products to x-ray crystallography to fluorescence spectroscopy.

The results of these efforts have produced a number of unique and intriguing observations. First, I have demonstrated that the Chlorella virus arginine decarboxylase prefers arginine by over 600-fold compared ornithine or lysine, and, therefore, represents a new activity within the ODC clade of Group IV decarboxylases. Second, I have shown that the structural basis of specificity of Group IV decarboxylases is a short helix which functions as a molecular ruler that selects substrates based on chain-length. Mutants of both arginine and ornithine decarboxylase that I created demonstrate the importance of the sequence identity and precise positioning of this helical ruler in determining substrate specificity.

## TABLE OF CONTENTS

TITLE-FLY .....	i
DEDICATION .....	ii
TITLE PAGE.....	iii
COPYRIGHT.....	iv
ACKNOWLEDGEMENTS.....	v
ABSTRACT.....	vi
TABLE OF CONTENTS.....	viii
PRIOR PUBLICATIONS .....	xii
LIST OF FIGURES .....	xiii
LIST OF TABLES.....	xvi
LIST OF ABBREVIATIONS.....	xvii
I. INTRODUCTION.....	1
A. POLYAMINE BIOSYNTHESIS.....	1
B. DISCOVERY OF THE <i>PARAMECIUM BURSARIA</i> CHLORELLA VIRUS	
“ODC” .....	2
C. EVOLUTION OF PLP DECARBOXYLASES.....	3
D. MECHANISM OF PLP-DEPEDENT DECARBOXYLATION.....	4
E. GROUP IV DECARBOXYLASES AS DRUG TARGETS.....	5
F. SUBSTRATE SPECIFICITY OF GROUP IV ENZYMES.....	6
II. SPECIFICITY DETERMINATION OF PBCV-1 ODC HOMOLOG.....	
A. INTRODUCTION.....	20
B. MATERIALS .....	21

C. EXPERIMENTAL PROCEDURES.....	22
D. RESULTS.....	26
1. ANALYSIS OF PBCV-1 ODC HOMOLOG SUBSTRATE PREFERENCE AND ACTIVITY.....	26
2. INHIBITION OF PBCV-1 ODC HOMOLOG PROTEIN WITH DFMO AND DFMA.....	28
E. DISCUSSION .....	31
III. X-RAY STRUCTURE DETERMINATION OF THE CHLORELLA VIRUS ARGININE DECARBOXYLASE .....	
A. INTRODUCTION .....	42
B. MATERIALS .....	44
C. EXPERIMENTAL PROCEDURES.....	44
D. RESULTS.....	49
1. X-RAY STRUCTURE DETERMINATION OF FREE AND AGMATINE-BOUND CVADC.....	49
2. THE ACTIVE SITE OF CVADC .....	50
3. THE K148 CVADC (K169 ODC) LOOP FUNCTIONS AS AN ACTIVE SITE LID.....	51
4. THE $3_{10}$ -HELIX IS A KEY DETERMINANT OF SUBSTRATE SPECIFICITY.....	52
5. STRUCTURAL COMPARISON OF CVADC AND DAPDC.....	54
6. STRUCTURAL COMPARISON OF CVADC AND AR.....	55

7. STRUCTURAL COMPARISON OF CVADC AND PYRUVOYL- DEPENDENT ADC.....	56
E. DISCUSSION.....	57
IV DETERMINANTS OF SUBSTRATE SPECIFICITY IN GROUP IV DECARBOXYLASES .....	
A. INTRODUCTION .....	87
B. MATERIALS .....	91
C. EXPERIMENTAL PROCEDURES.....	91
D. RESULTS.....	93
1. MODELS OF CVADC AND <i>T. BRUCEI</i> ODC MUTANTS.....	93
2. INHIBITION OF WILD-TYPE CVADC AND <i>T. BRUCEI</i> ODC WITH SUBSTRATE ANALOGS .....	94
3. STEADY-STATE ANALYSIS OF WILD-TYPE AND MUTANT CVADC AND <i>TBODC</i> .....	94
E. DISCUSSION .....	97
V. LOOP DYNAMICS IN GROUP IV FOLD ENZYMES.....	108
A. INTRODUCTION .....	108
B. MATERIALS .....	109
C. EXPERIMENTAL PROCEDURES.....	110
D. RESULTS.....	110
1. STEADY-STATE ANALYSIS OF WILD-TYPE AND MUTANT ENZYMES.....	110

2. FLUORESCENCE SPECTRA OF WILD-TYPE AND MUTANT	
CVADC ENZYMES.....	111
E. DISCUSSION.....	111
VI. CONCLUSIONS .....	116
REFERENCES .....	122

#### PRIOR PUBLICATIONS

**Shah, R.**, Coleman, C.S., Mir, K., Baldwin, J., Van Etten, J.L., Grishin, N.V., Pegg, A.E., Stanley, B.A., 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 (34), 35760-35767

**Shah, R.**, Akella, R., Goldsmith, E.J., Phillips, M.A. (2007) X-ray Structure of *Paramecium bursaria* Chlorella Virus Arginine Decarboxylase: Insight into the Structural Basis for Substrate Specificity. *Biochemistry*, 46, 2831-2841.



## LIST OF FIGURES

FIGURE 1-1.....	9
FIGURE 1-2.....	10
FIGURE 1-3.....	11
FIGURE 1-4.....	12
FIGURE 1-5.....	13
FIGURE 1-6.....	14
FIGURE 1-7.....	15
FIGURE 1-8.....	16
FIGURE 1-9.....	17
FIGURE 1-10.....	18
FIGURE 1-11.....	19
FIGURE 2-1 .....	35
FIGURE 2-2 .....	36
FIGURE 2-3 .....	37
FIGURE 2-4 .....	39
FIGURE 2-5 .....	40
FIGURE 2-6 .....	41
FIGURE 3-1 .....	63
FIGURE 3-2.....	64
FIGURE 3-3 .....	65
FIGURE 3-4 .....	66
FIGURE 3-5 .....	69

FIGURE 3-6 .....	70
FIGURE 3-7 .....	71
FIGURE 3-8 .....	72
FIGURE 3-9 .....	73
FIGURE 3-10 .....	74
FIGURE 3-11 .....	75
FIGURE 3-12 .....	76
FIGURE 3-13 .....	77
FIGURE 3-14 .....	78
FIGURE 3-15 .....	79
FIGURE 3-16 .....	81
FIGURE 3-17 .....	82
FIGURE 3-18 .....	83
FIGURE 3-19 .....	84
FIGURE 3-20 .....	85
FIGURE 3-21 .....	86
FIGURE 4-1.....	100
FIGURE 4-2.....	101
FIGURE 4-3.....	102
FIGURE 4-4 .....	103
FIGURE 4-5 .....	104
FIGURE 4-6 .....	107

FIGURE 5-1 .....	114
FIGURE 5-2 .....	115
FIGURE 6-1 .....	121

## LIST OF TABLES

TABLE I .....	38
TABLE II.....	67
TABLE III .....	68
TABLE IV .....	80
TABLE V .....	105
TABLE VI .....	106

## LIST OF ABBREVIATIONS

AATase	Aspartate aminotransferase
ADC	Arginine decarboxylase
AIH	Agmatine iminohydrolase
APAO	Amino polyoyl amine oxidase
AR	Alanine racemase
a.s.u	asymmetric unit
$\beta$ -ME	beta-mercaptoethanol
CAPSO	3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic Acid Sodium Salt
CP	Carbamoyl putrescine
CPA	Carbamoyl putrescine amidohydrolase
CNS	Crystallography and NMR system
<i>cv</i> ADC	Chlorella virus arginine decarboxylase
DAPDC	Diaminopimelate decarboxylase
DFMA	Difluoromethylarginine
DFMO	Difluoromethylornithine
DNase	Deoxyribonuclease
DTT	Dithiothreitol
eIF	eukaryotic translation initiation factor
EPMR	Evolutionary program for molecular replacement
ES <sup>‡</sup>	Enzyme-substrate transition state
FPP	Farnesyl pyrophosphate
FTase	Farnesyl transferase
GGPP	Geranylgeranyl pyrophosphate
GGTase	Geranylgeranyl transferase
HPLC	High-performance liquid chromatography
IMPDH	Inosine monophosphate dehydrogenase
$k_{\text{cat}}$	enzyme turnover rate
$K_{\text{m}}$	Michaelis equilibrium constant
LDH	Lactate dehydrogenase

LPS	Lipopolysaccharide
MALDI-TOF	Matrix-assisted laser desorption/ionization time-of-flight
MDH	Malate dehydrogenase
μM	micromolar
mM	millimolar
<i>mj</i> DAPDC	<i>Methanococcus jannaschii</i> diaminopimelate decarboxylase
MS	Mass spectrometry
ODC	Ornithine decarboxylase
NCS	Non-crystallographic symmetry
PBCV	<i>Paramecium bursaria</i> Chlorella virus
PLP	Pyridoxal-5'-phosphate
RMSD	Root-mean-square deviation
RT	Retention time
SAD	Single-wavelength anomalous dispersion
SMO	Spermine oxidase
SSAT	Spermine N-acetyltransferase
TATase	Tyrosine aminotransferase
<i>tb</i> ODC	<i>Trypanosoma brucei</i> ornithine decarboxylase
TIM	Triosephosphate isomerase
TS	Tryptophan synthase



# CHAPTER ONE

## INTRODUCTION

### **Polyamine Biosynthesis**

Polyamines are small, cationic molecules found in all species of life. The most prominent polyamines, putrescine, spermidine and spermine (Figure 1-1), have demonstrated roles in cell survival, replication and differentiation (5). The metabolic pathway common to most, but not all, eukaryotic cells is shown in Figure 1-2. Because of the role polyamines play in the survival and growth of cells, enzymes in the polyamine metabolic pathway are attractive targets for anti-parasitic and anti-neoplastic therapies. A number of polyamine biosynthetic enzymes have been targeted for various diseases, some with clinical success (6, 7). The first committed step in the biosynthesis of polyamines is the formation of putrescine from ornithine, a reaction catalyzed by ornithine decarboxylase (ODC). Spermidine synthase then links a molecule of putrescine to the aminopropyl group of decarboxylated S-adenosylmethionine to form spermidine. A subsequent reaction, similar to that carried out by spermidine synthase, is catalyzed by spermine synthase to form spermine. Spermidine and spermine support cell survival and growth through a number of mechanisms including association with nucleic acids (8), gene regulation (9), ion-channel regulation (10) and the formation of eukaryotic translation initiation factor 5A (eIF5A) (11).

Variations of the polyamine metabolism depicted in Figure 1-2 are observed in the various uni- and multi-cellular species. In mammals, ornithine is produced by the catabolism of arginine by the enzyme arginase (12). Alternatively, putrescine can be formed from arginine, commencing with the decarboxylation of arginine to form agmatine (1-(4-aminobutyl)guanidine). Putrescine can be formed from agmatine by the enzyme agmatinase ((13) Figure



1-3). This enzyme has been characterized in plants (14) and mammals (15). Alternatively, putrescine production can proceed through formation of carbamoyl putrescine (CP) by agmatine iminohydrolase (AIH) and subsequent removal of the carbamoyl moiety of CP by carbamoyl putrescine amidohydrolase (CPA) ((16), Figure 1-3). Thus far the ADC-based pathway for putrescine production has not been found in mammals (16); however, this pathway has been identified both in plants and some prokaryotes (17, 18). Enzymes for the ADC-based pathway for putrescine production have also been characterized from the *Paramecium bursaria* Chlorella Virus ((2), Figure 1-4). The Chlorella Virus arginine decarboxylase (*cvADC*) is the focus of my dissertation. In the following chapters I will discuss efforts I have undertaken to determine the enzymatic activity and three-dimensional structure of *cvADC*. Results of those studies will be discussed with respect to the larger family of homologs to which *cvADC* belongs. Furthermore, I will present attempts to isolate the determinants of substrate specificity and the role of loop dynamics in the catalytic cycle of *cvADC* and its homologs.

### **Discovery of the *Paramecium bursaria* Chlorella Virus “ODC”**

The *Paramecium bursaria* Chlorella virus is a large, icosahedral, double-stranded DNA virus that is distantly related to the Herpes simplex virus type 2 (HSV-2) (19). The virus infects certain green algae, which exhibit a symbiotic relationship with the *Paramecium bursaria*. In Figure 1-5, the Chlorella virus is shown infecting and lysing a chlorella algae. Van Etten and colleagues, with the assistance of the Institute for Genomic Research, sequenced the genome of the Chlorella virus. It has an unusually large genome, of 330 kilo-bases, that contains genes coding for proteins involved in the biosynthesis of such molecules as fucose, hyaluronic acid and polyamines (19). Examples of polyamine biosynthetic enzymes coded by the Chlorella virus genome are agmatine iminohydrolase, N-

carbamoylputrescine amidohydrolase and homospermidine synthase. Additionally, an open reading frame, a207r, which codes for a 372 amino acid protein with approximately 40% amino acid sequence identity to eukaryotic ornithine decarboxylase (ODC) was discovered (20), which displayed a low level of ornithine decarboxylase activity. The recombinant protein was found to catalyze the decarboxylation of L-ornithine; however, the catalytic efficiency was significantly lower than that observed for the enzymes from mouse and *T. brucei* (21). The authors noted their investigation comprised the first observation of a virally-encoded polyamine biosynthetic pathway and classified the gene product of a207r as an ODC-like protein.

### **Evolution of PLP decarboxylases**

Ornithine decarboxylase belongs to a large group of enzymes that utilize the cofactor vitamin B<sub>6</sub> or pyridoxal-5'-phosphate. In the role as an enzyme cofactor, the aldehyde species of vitamin B<sub>6</sub> is converted to an imine by the formation of a Schiff's base with the  $\epsilon$ -amino group of one of the enzyme's lysine residues (Figure 1-6). PLP is one of the most versatile cofactors because of its ability to catalyze reactions such as racemization, transamination and  $\alpha$  or  $\beta$  decarboxylation. More than 140 distinct enzymatic activities, mainly with amino acid substrates, have been attributed to PLP-dependent enzymes by the Enzyme Commission (EC) (22). These enzymes have been categorized based on their evolutionary relationships. Sandmeier et al. classified PLP-dependent enzymes based on profile analysis of multiple amino-acid-sequence alignments (23). This study placed PLP enzymes into four groups. Group III and IV contained the decarboxylases of basic amino acids. Group III is related to PLP-dependent aminotransferases and, Group IV is homologous to alanine racemase (AR). Soon after, Grishin et al classified PLP enzymes based on their three-dimensional structures (24). The authors of this study placed enzymes catalyzing

aminotransferase reactions, such as aspartate aminotransferase (AATase), and prokaryotic decarboxylases into Fold-type I. On the other hand, Fold-type III contained eukaryotic decarboxylases, alanine racemase and some prokaryotic diaminopimelate and arginine decarboxylases.

PLP enzymes involved in the biosynthesis of polyamines have been found to belong to both Group III and IV. For example, PLP-dependent ODCs of both Group III (bacterial) or Group IV (eukaryotic) families have been described (Figure 1-7). The structure of the Group III ODC from *Lactobacillus* 30a consists of five domains. The N-terminal domain folds into a five-stranded  $\beta$ -sheet and is followed by a linker domain consisting of  $\alpha$ -helices. The third domain folds into  $\alpha/\beta$  domain that is similar to the PLP-binding domain of the aspartate aminotransferase. The fourth domain also shares similarities with the small domain of aspartate aminotransferase. The fifth domain contains both anti-parallel loops and seven  $\alpha$ -helices (25). The group IV eukaryotic ODCs are composed of an N-terminal  $(\beta/\alpha)_8$ -barrel and a C-terminal  $\beta$ -barrel (26-32). The biologically active oligomer of eukaryotic ODC is a dimer and two active sites are formed at the subunit interface. The suicide inhibitor  $\alpha$ -difluoromethyl-ornithine (DFMO) targets only the Group IV ODC (33). In Figure 1-8 the structure of *T. brucei* ODC complexed with DFMO is shown. The inhibitor is bound to PLP through a Schiff's base and to the catalytically relevant Cys-360 (28).

### **Mechanism of PLP-dependent decarboxylation**

The minimal reaction mechanism of ODC-catalyzed decarboxylation of L-ornithine was characterized by multi-wavelength stopped-flow absorbance spectroscopy (34). The reaction proceeds from an internal aldimine (with Lys-69) to an external aldimine with substrate via a gem-diamine intermediate (Figure 1-9). Subsequent decarboxylation was followed by a structurally rigid quinoid

intermediate which decomposed to a Schiff's base with product. Product release entailed the exchanging of the external aldimine with product to one with Lys-69 via gem-diamine intermediate.

The investigation suggested that the formation of a Schiff base intermediate formed rapidly (i.e. during the dead time of the stopped-flow instrument) and was followed by formation of a quinoid intermediate with a rate constant of  $21 \text{ s}^{-1}$  (34). Protonation at  $\text{C}\alpha$ , as opposed to the  $\text{C4}\alpha$  of PLP, is thought to be promoted by the favoring of a quinoid tautomer with a characteristic absorbance maxima at 350 nm. The quinoid intermediate is thought to decay in two steps to a Schiff base with putrescine. Finally, the authors concluded that the rate of product release is the rate-limiting step.

#### **Group IV decarboxylases as drug targets**

*T. brucei* ODC is the most successful example of a Group IV decarboxylase as a drug target. Di-fluoromethylornithine (DFMO) is a clinically relevant therapy for the treatment of African sleeping sickness (Figure 1-8). Other Group IV enzymes have been targeted or are suggested to be targets for the treatment of various human diseases. Human ODC has been targeted by DFMO in the context of some cancers (6). Alanine racemase (AR) is the target of cycloserine, which is used in the treatment of bacterial infections (35). Similar to DFMO cycloserine acts as an irreversible inhibitor of AR. Inhibition of AR limits the available pool of D-alanine, a component of bacterial cell walls, and ultimately kills the targeted bacterium (36). Diaminopimelate decarboxylase, which decarboxylates meso-diaminopimelate (Figure 1-10) to form lysine, is considered to be a potential target for various microbial infections. Currently, only low affinity inhibitors of DAPDC have been synthesized. However, the potential for DAPDC as a drug target is significant because lysine synthesis is necessary for both protein production and formation of the bacterial

peptidoglycan cell wall (37). It is conceivable, then, that a potent inhibitor of DAPDC could also have potent anti-bacterial properties. Arginine decarboxylase (ADC) from various human pathogens, including the causative agent of bubonic plague, *Yersinia pestis*, has also been contemplated as a drug target (38). The product of arginine decarboxylation, agmatine, is an inhibitor of nitric oxide synthase and a suppressor of apoptosis, so its production would be advantageous for an intracellular bacterial parasite (39).

In light of past and current drug development efforts, the Group IV PLP-dependent decarboxylases represent a class of drug targets with proven, and potentially greater, clinical relevance (40).

### **Substrate specificity in Group IV enzymes**

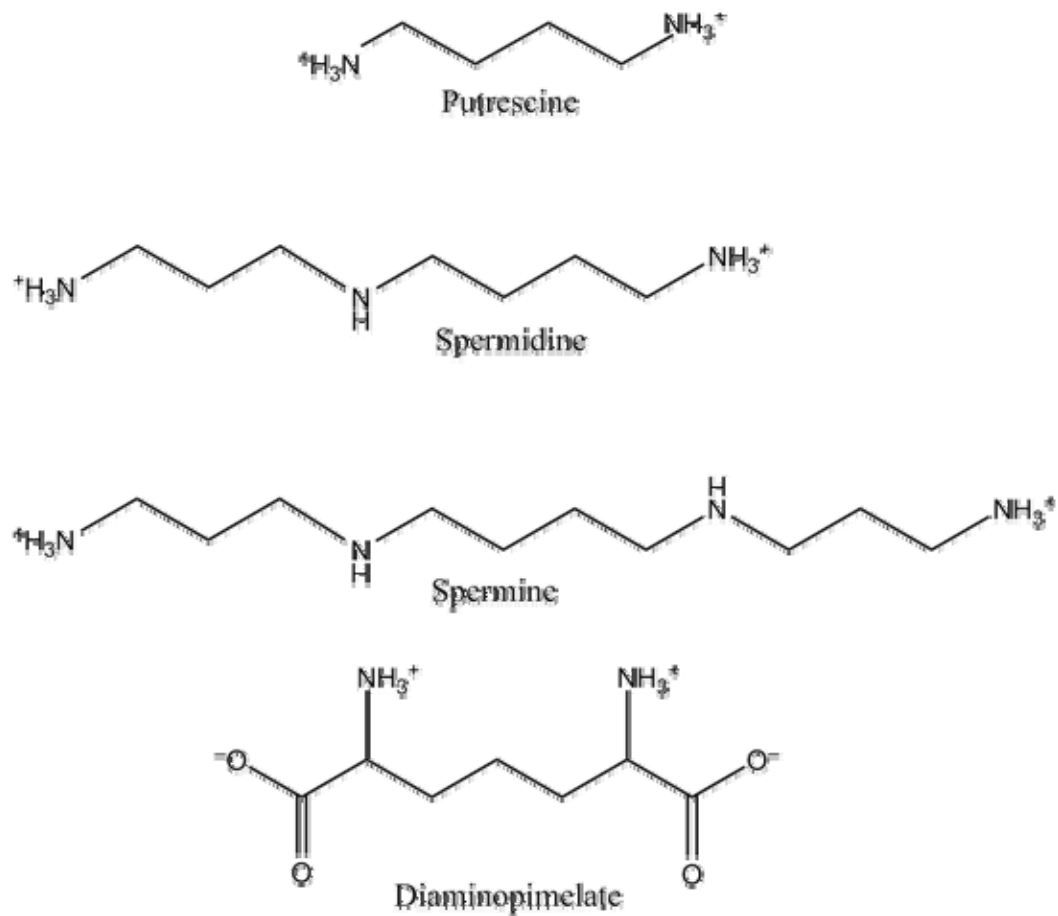
The elucidation of the determinants of substrate specificity in Group IV decarboxylases will yield a better understanding of the evolution of this family and aid in further development of therapies targeting these enzymes. Phylogenetic analysis suggests that Group IV decarboxylases have undergone a considerable amount of evolution in order to develop activity towards a specific substrate. That is, Group IV decarboxylases with different substrate specificities share low amino acid sequence identity ( $\leq 20\%$ ). Decarboxylases with the same substrate preference share considerably higher sequence identity ( $\geq 40\%$ ) (Figure 1-10). Previous data on activity of ODC with a range of substrates suggests that these enzymes evolved to minimize activity on non-cognate substrates (21). Substrate discrimination may, therefore, be an integral part of the evolutionary history of these enzymes, a feature which would be reflected in the structure of these enzymes.

A comparison of the monomeric structures of *T. brucei* ODC and *Mycobacterium tuberculosis* DAPDC reveals that the large sequence divergence (i.e.  $\sim 20\%$  sequence identity) translates into multiple structural changes,

especially around the PLP binding site (Figure 1-11). Although the overall RMSD between the two structures is not exceedingly high (2.5 Å), there are numerous structural changes in the active site. Multiple loops and secondary elements that comprise the PLP and substrate binding pocket have shifted in *M. tuberculosis* DAPDC relative to *T. brucei* ODC. Hence, determining which of the many structural differences are relevant to discriminating between substrates is problematic.

The discovery of a novel ODC-like protein from the Chlorella virus presents a unique opportunity to isolate the determinants of substrate specificity in Group IV decarboxylases. An analysis of the amino acid content of Chlorella algae, the Chlorella virus host, revealed that arginine was the most abundant while ornithine was amongst the least abundant amino acid (41). Coupling this knowledge to the discovery of genes coding for agmatine imidohydrolase and N-carbamoylputrescine amidohydrolase in the Chlorella virus genome leads to the hypothesis that the Chlorella virus utilizes the ADC-based pathway for polyamine production. As mentioned before, the putative ODC from the Chlorella virus genome shares about 40% sequence identity with eukaryotic ODC. This is the minimal level of identity shared amongst the various eukaryotic ODCs. For example, yeast ODC and mouse ODC share approximately 44% identity. Despite this level of sequence conservation, there is at least one active site change (D332E) that suggests the binding pocket of this ODC-like protein is altered with respect to eukaryotic ODC (see Chapter II). Thus, the distinct possibility exists for a scenario in which the Chlorella virus evolved the specificity of an ODC gene from a host to establish its own ADC-based pathway for polyamine synthesis. I will show in the next chapter that the Chlorella virus ODC-like protein's specificity differs significantly from *bona fide* ODCs. In chapter III I will demonstrate that its structure elucidates the basis of specificity in the larger family of Group IV PLP-dependent decarboxylases. Finally, in chapter IV I used

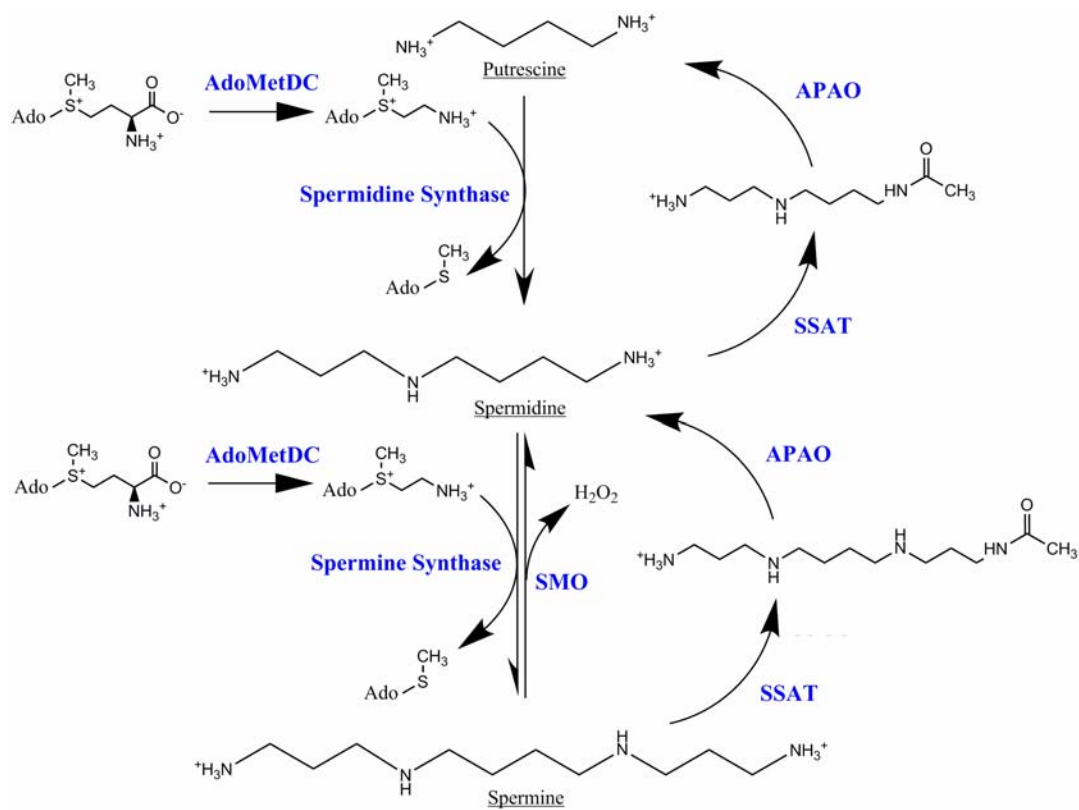
structural insights to guide the design of mutagenesis experiments which attempted to switch specificities of the ODC-like protein and eukaryotic ODC. Finally, I discuss the initial attempts to understand the role of loop dynamics in the catalytic cycle of Group IV decarboxylases in chapter V.



**FIGURE 1-1**

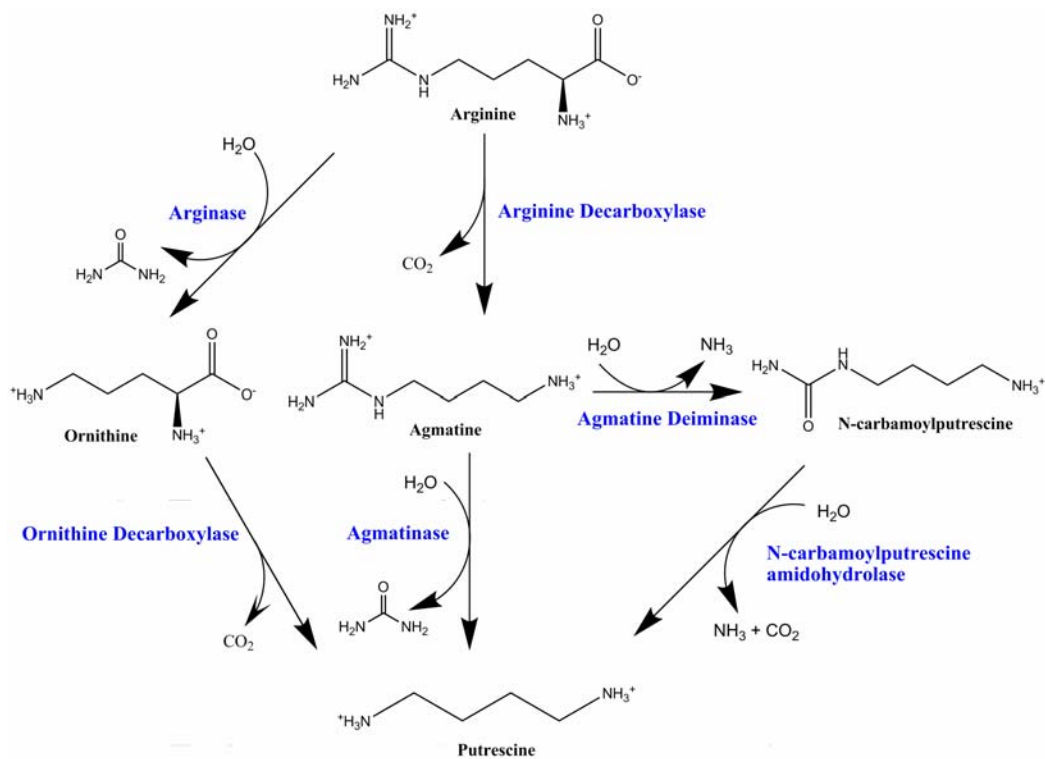
Putrescine, spermidine and spermine are the most prevalent polyamines. Diaminopimelate is a component of bacterial cell walls and also serves as a source of lysine.





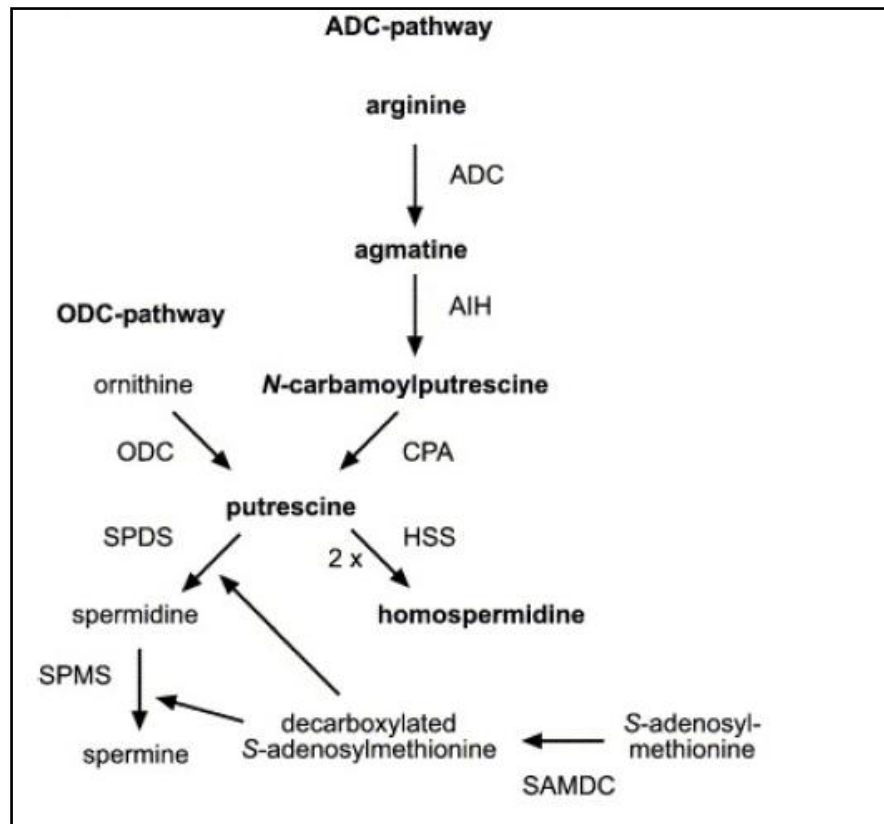
**FIGURE 1-2**

The polyamine metabolic pathway of mammalian cells. Enzymes of the pathway are shown in blue text. AdoMetDC is short for S-adenosylmethionine decarboxylase. SMO is spermine oxidase, SSAT is spermine N-acetyltransferase, APAO is amino polyamine oxidase.



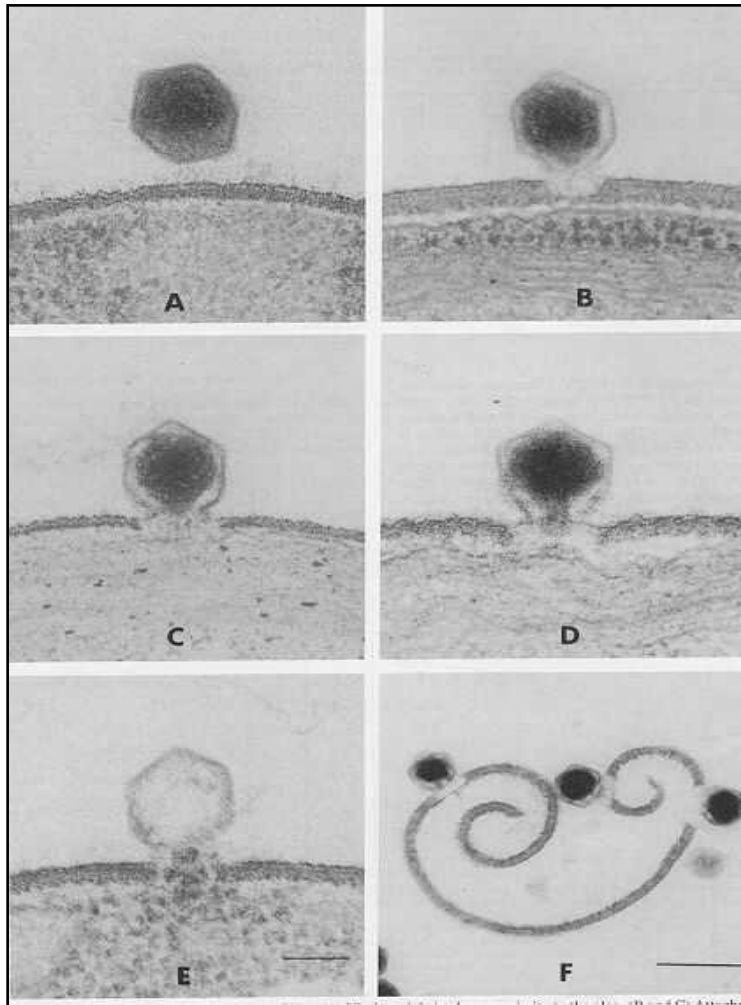
**FIGURE 1-3**

Alternative pathways for the conversion of arginine to putrescine.



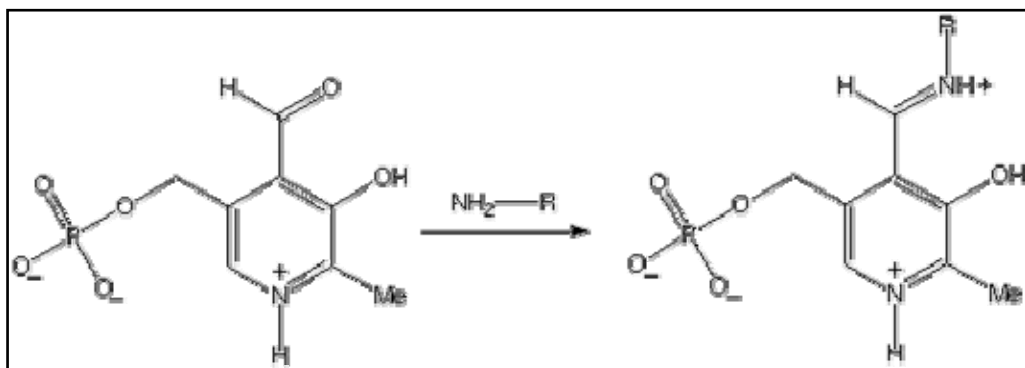
**FIGURE 1-4**

The demonstrated polyamine pathway of the *Paramecium bursaria* Chlorella virus. This figure was adapted from (2).



**FIGURE 1-5**

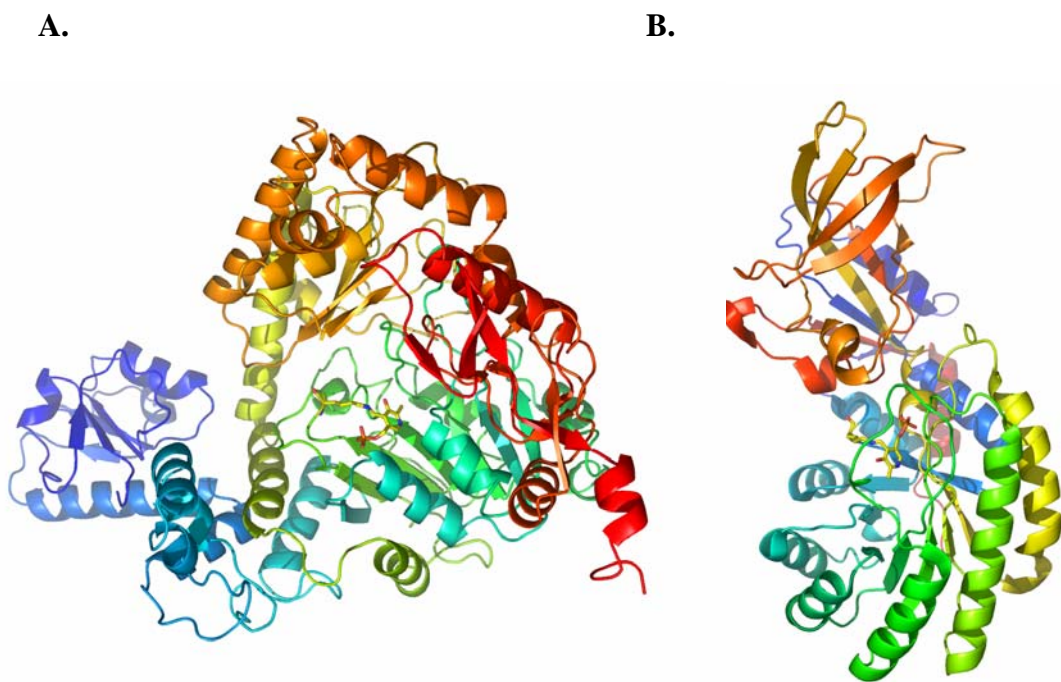
The Chlorella virus approaches the membrane of its host (A), fuses to the membrane (B), lyses the membrane (C), inserts its dsDNA genome (D, E) and subsequently lyses the algal host. These images were obtained from [HTTP://PLANTPATH.UNL.EDU/FACILITIES/VIROLOGY/](http://plantpath.unl.edu/facilities/virology/).



**FIGURE 1-6**

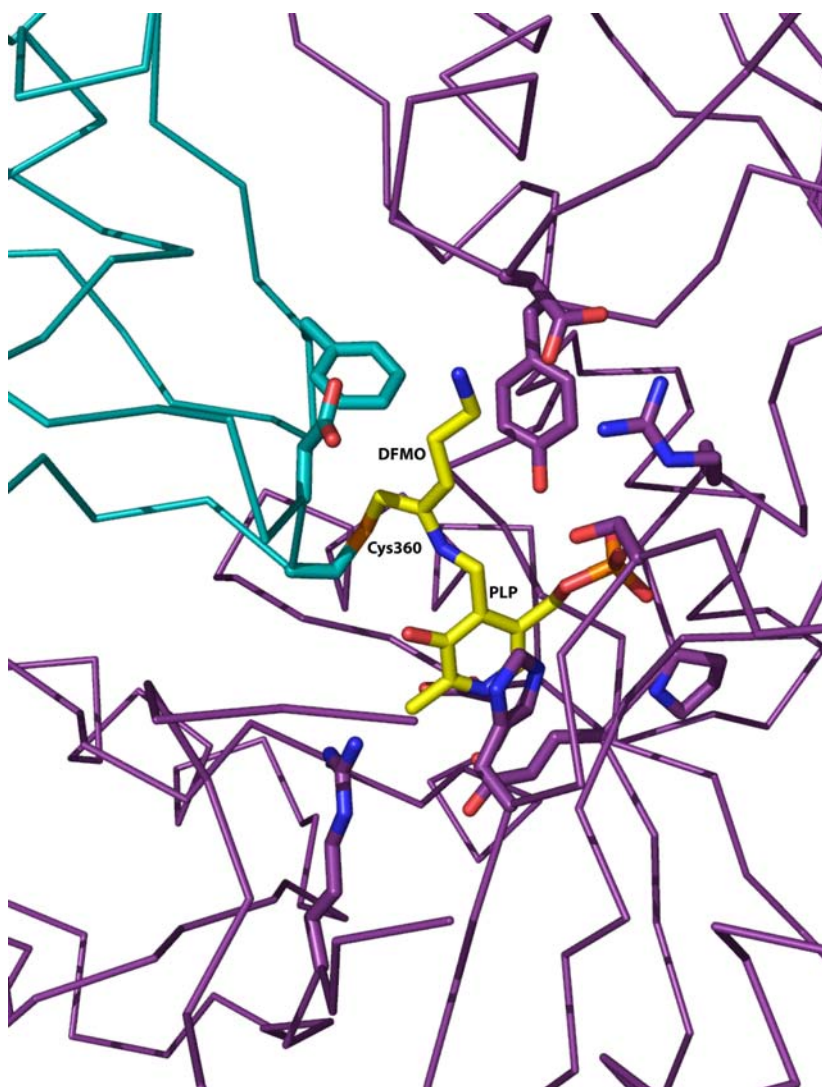
Vitamin B6, pyridoxal-5'-phosphate (PLP), is converted from an aldehyde to an imine by forming a Schiff's base with a lysine residue with which PLP functions as a cofactor.

[TONGA.USIP.EDU/GMOYNA/BIOCHEM341/LECTURE27.HTML](http://TONGA.USIP.EDU/GMOYNA/BIOCHEM341/LECTURE27.HTML)



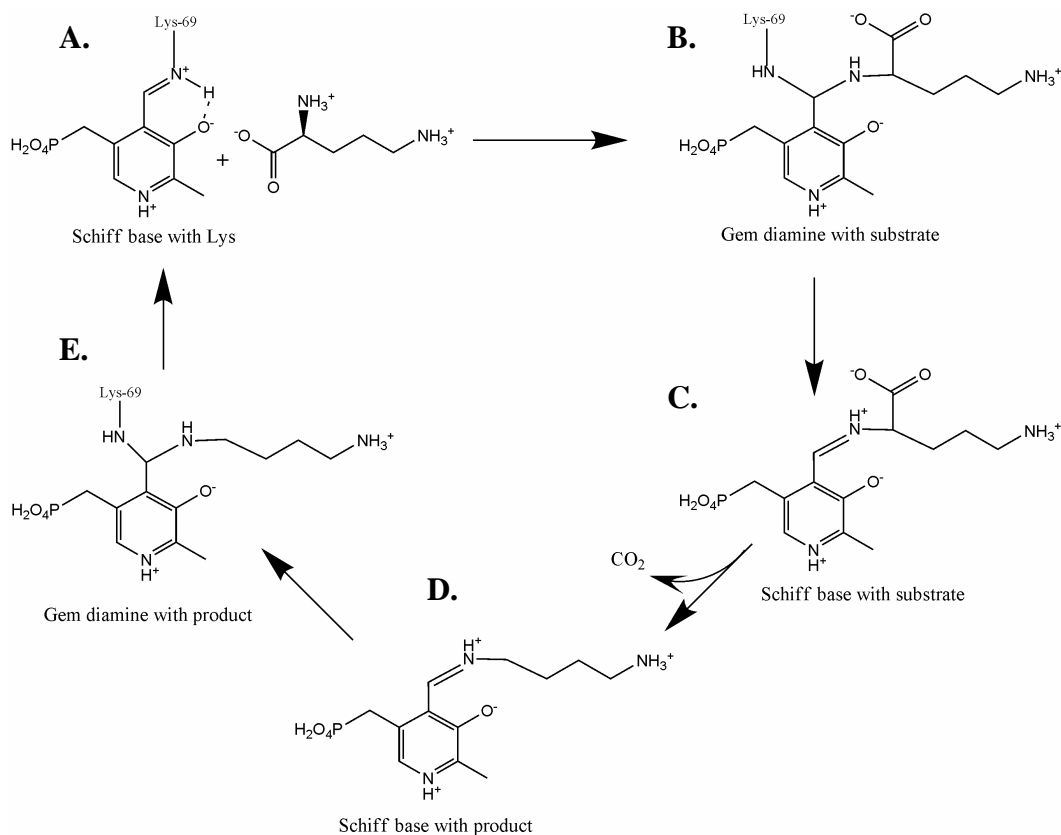
**FIGURE 1-7**

**A.** The Group III PLP-dependent ODC contains multiple domains which are structurally similar to the domains observed in aspartate aminotransferase. **B.** The Group IV PLP-dependent ODC is composed of an N-terminal TIM ( $\beta/\alpha$ )<sub>8</sub> barrel and an N-terminal  $\beta$ -barrel.



**FIGURE 1-8**

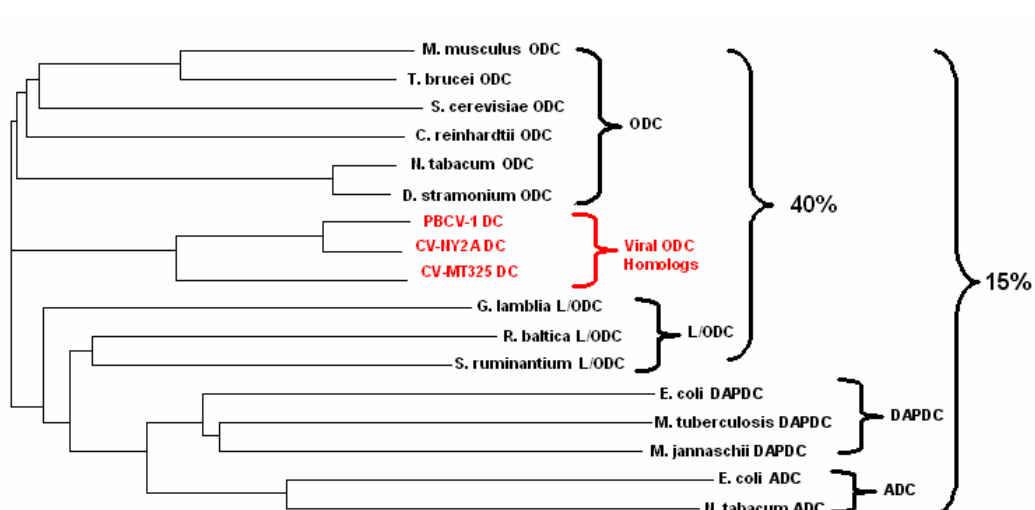
The structure of *T. brucei* ODC bound to DFMO, an approved therapy for African sleeping sickness. DFMO (yellow) is covalently bound to PLP (yellow) and Cys-360 in the active site of ODC. The active site is formed at the dimer interface and is composed of residues from both chains A (purple) and B (cyan). Atoms other than carbon are colored as follows: Nitrogen atoms are blue, oxygen atoms are red, sulfur atoms are yellow, and phosphate atoms are orange.



**FIGURE 1-9**

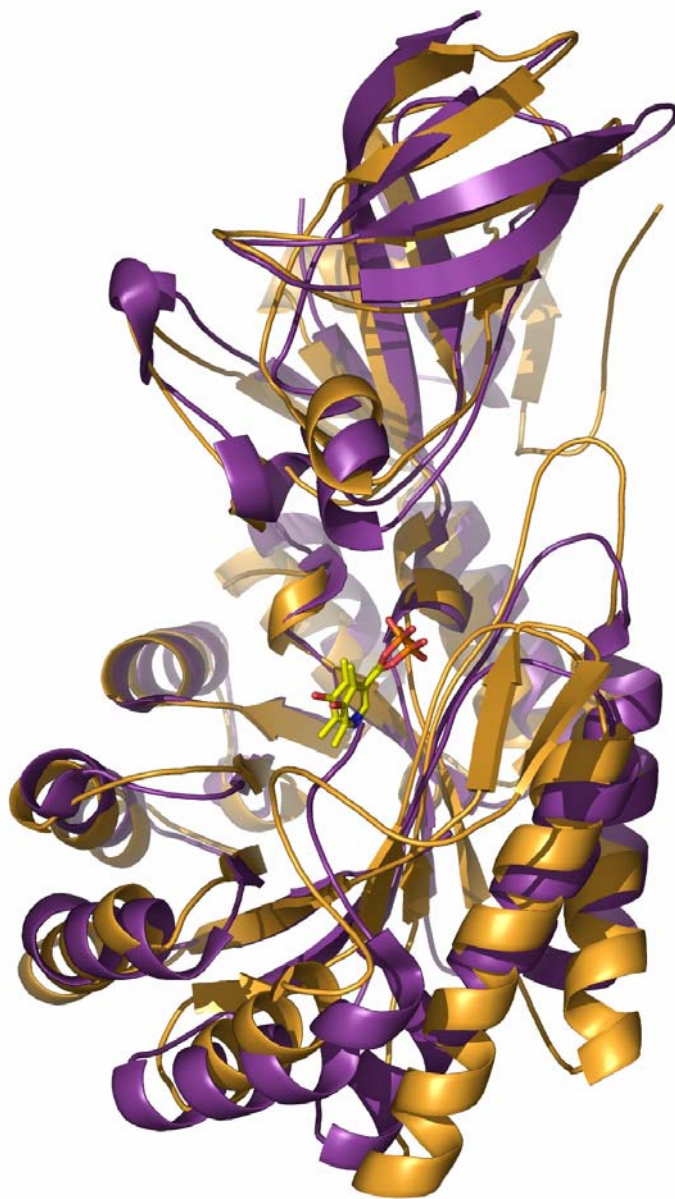
Minimal mechanism of ODC-catalyzed decarboxylation of L-ornithine. The internal aldimine (**A**) with PLP and Lys proceeds through a gem-diamine intermediate (**B**) with ornithine to a Schiff's base with substrate (**C**). Upon decarboxylation a Schiff's base with product (**D**) returns, through a gem-diamine intermediate (**E**), to the internal aldimine.





**FIGURE 1-10**

Evolutionary relationships of Group IV PLP-dependent enzymes. Chlorella virus ODC-like protein (labeled PBCV-1 DC) as well as the equivalent gene from two other Chlorella virus isolates (CV-NY2A and CV-MT325) are shown in red and referred to as Viral ODC homologs. Decarboxylases with the same substrate preference share high sequence identity with one another but low identity with enzymes with different preferences.



**FIGURE 1-11**

A structural comparison of *T. brucei* ODC and *M. tuberculosis* DAPDC. The overall RMSD between the two structures is 2.5 Å. Nonetheless, the deviations of loops and secondary elements in and around the active site are plentiful. This fact precludes a definitive determination of the structural basis of specificity in Group IV decarboxylases. The structures were aligned by C $\alpha$  atoms using the DALI structural alignment server.

## CHAPTER TWO

### SUBSTRATE SPECIFICITY DETERMINATION OF A PBCV-1 ODC HOMOLOG

#### Introduction

The 372 codon open-reading frame (a207r) in the *Paramecium bursaria* Chlorella virus genome encodes a homolog of eukaryotic ODC that shares 40% sequence identity with eukaryotic ODC (20). This ODC homolog is expected to fold similarly to other Group IV PLP-dependent decarboxylases. That is, the N-terminal is characterized by a  $(\beta/\alpha)_8$  barrel and the C-terminal as a  $\beta$ -barrel structure. In the structures of such enzymes, the active sites are formed at the dimer interface between the N-terminal domain of one monomer and the C-terminal domain of the other (Figure 2-1). While all enzymes in this family share a number of essential active site residues, eukaryotic ODCs are very distantly related to bacterial ADCs and DAPDCs (Figure 2-2 (3, 4, 21, 28, 42)). In general, eukaryotic ODCs share about 20% overall sequence identity, based on pair-wise comparisons, with the bacterial ADCs and DAPDCs. In contrast, a bacterial enzyme from this fold type has been described that has activity on both lysine and ornithine (Lys/OrnDC) and shares extensive sequence similarity (~35%) with eukaryotic ODCs (43).

Phylogenetic analysis indicates that the ODC-like protein from the Chlorella virus branches with the eukaryotic ODCs and with the bacterial Lys/OrnDC enzymes (Figure 1-9). Alignment of the primary amino acid sequence of this enzyme with the X-ray structures for mammalian and trypanosome ODCs reveals a key amino acid substitution in the substrate-binding pocket that is predicted to alter the substrate specificity (Figure 2-3 A & B). The structures of *T. brucei* ODC in complex with several substrate and product analogs demonstrate

that the  $\delta$ -nitrogen of L-ornithine forms salt-bridge interactions with Asp-361 from the C-terminal domain of the opposite monomer, and with Asp-332 from the C-terminal domain across the subunit boundary (28, 30, 32). These two residues are conserved in most functional ODCs that have been described. In the Chlorella virus ODC-like protein, Asp-332 has been substituted with a glutamic acid (Glu-296) (Figure 2-3 B). Though it is relatively conservative, this substitution occurs within the confines of an active site of limited volume. Thus, the magnitude and nature of the interaction between the  $\delta$ -nitrogen of putrescine and the carboxylate moiety of Glu-296 in the Chlorella virus ODC-like protein must be altered with respect to that observed in *T. brucei* ODC. Taken together with the presence of agmatine processing enzymes in the Chlorella virus, these data suggested the possibility that the substrate specificity of the Chlorella virus ODC-like protein differs from that of the eukaryotic enzyme.

To investigate the substrate preference of Chlorella virus ODC-like protein, I determined the decarboxylation activity of the enzyme on a range of basic amino acids. I will show in this chapter that, while this enzyme contains detectable activity with L-ornithine, its activity with L-arginine is significantly higher. The fact that it is more strongly related to the enzymes with specificity for L-ornithine and L-lysine, than to the ADCs from bacteria and plants, suggests that the Chlorella virus ODC-like protein represents a new activity within the clad of the ornithine specific enzymes. This observation raises the possibility that a minimal number of substitutions in the Chlorella virus ODC-like protein, relative to ODC, have occurred to confer L-arginine specificity. Thus, the basis of specificity, both in this enzyme and in Group IV decarboxylases, may be clarified by the analysis of the primary and tertiary structure of this enzyme.

## Materials

Amino acids, polyamines, and the carbon dioxide kit were purchased from Sigma (St. Louis, MO). The AccQ-Fluor Reagent Kit for labeling amino acids was purchased from Waters (Milford, MA). L-[1-<sup>14</sup>C] and L-[2, 3, <sup>3</sup>H]ornithine (47.7 mCi/ mmol) were obtained from New England Nuclear (Boston, MA) and American Radiochemicals, Inc. (St. Louis, MO), and L-[U-<sup>14</sup>C] arginine (310 mCi/mmol) was purchased from Amersham Pharmacia Biotech (Piscataway, NJ). Sequencing grade modified trypsin was from Promega (Madison, MI) and endopeptidase Glu-C was from Roche Applied Science (Indianapolis, IN). DFMO was obtained from ILEX Oncology (San Antonio, TX) and DFMA, originally obtained from Dr. A. Bitonti (Merrell Dow Research Institute, Cincinnati, OH), was kindly provided by Dr. Patrick M. Woster (Wayne State University, Detroit , MI)

## Experimental Methods

*Modeling and sequence analysis*- The program PyMol (Delano Scientific, South San Francisco, CA) was used to display the active site of *T. brucei* ODC for the wild-type enzyme in complex with putrescine [PDB ID 1f3t; (32)]. The distance tree was built using the Treeview (44) package from the alignment of the displayed sequences constructed with the program EXPRESSO (45). Sequence analysis of polymorphisms in the ODC family was performed using Gen Bank version 2.2.8. Blast analysis was done to identify eukaryotic ODC homologs and to generate an alignment. In total 53 eukaryotic sequences that had published sequence data were included, hypothetical proteins were excluded from the analysis.

*Protein Expression and Purification*- the Chlorella virus ODC-like protein was produced in *E. coli* from the expression plasmid for PBCV-1 DC, termed pODCTM9 which was described in (20). This vector directs the expression of an N-terminal His<sub>6</sub>-tagged PBCV-1 DC from a pET-15b plasmid that contains the

cDNA corresponding to the open reading frame a207r present in the PBCV-1 genome. Two sources of PBCV-1 ODC were used in the studies, 1) a protein with the wild-type sequence was used in the  $^{14}\text{CO}_2$  release assay (Table I) and for DFMO and DFMA inactivation studies (Fig. 2-6 A and 4B), 2) Protein used for the coupled-enzyme assays contained the substitution of Ala for Thr at position 142 (Figure 2-4). This change was generated during cloning but has no effect on the activity or substrate specificity of the enzyme. Recombinant Chlorella virus ODC-like protein was expressed and purified as described previously for *T. brucei* ODC (46). After induction of protein production via the IPTG-inducible lacZ promoter on the plasmid, the *E. coli* cells were harvested by centrifugation at 4° C, 1500 rpm for 30 minutes. The harvested cells were frozen in liquid nitrogen. Cells were lysed by stirring in buffer (200 mM NaCl, 50 mM Tris, pH 8.0, 5% glycerol, 7.1 mM  $\beta$ -ME, 200  $\mu$ M PLP, 2 mM PMSF and 500 to 1000 mgs of leupeptin, antipain, benzamidine, pepstatin and chymostatin) with approximately 0.1 g/L lysozyme at 4°C for 1.5 hours. The lysate was sonicated and DNase from bovine pancreas was added to the lysate and mixed thoroughly. The sonicated lysate was centrifuged at 100,000 X g at 4° C for one hour. The supernatant was loaded onto a  $\text{Ni}^{2+}$ -agarose column and washed with a low-salt (100 mM NaCl, 50 mM Tris, pH 8, 0.03% Brij-35, 2.2 mM  $\beta$ -ME) and high-salt buffer (1.1 M NaCl, 50 mM Tris, pH 8, 0.03% Brij-35, 2.2 mM  $\beta$ -ME). The protein was eluted with buffer containing imidazole (100 mM NaCl, 50 mM HEPES, pH 7, 2.2 mM  $\beta$ -ME, 0.15 % Brij-35, 200 mM imidazole, pH 7.1) using a step gradient. The majority of protein bound to the nickel-resin eluted with 100 mM imidazole. Fractions were resolved via SDS-PAGE and those containing the desired protein product (~ 42 kDa) were concentrated and purified further via gel filtration techniques. Purified protein was concentrated via ultracentrifugation using a

Centriprep (10 kDa molecular weight cut-off) ultrafiltration column (Millipore, Bedford, MA).

*Radiolabel Release-based Enzymatic Assays-* ODC and ADC activities were determined by measuring the release of  $^{14}\text{CO}_2$  from the enzymatic decarboxylation of L-[1- $^{14}\text{C}$ ] ornithine or L-[U- $^{14}\text{C}$ ] arginine as previously described (47). The  $^{14}\text{CO}_2$  released was trapped in a center well containing hyamine hydroxide and counted by liquid scintillation spectrometry as described previously. Kinetic constants were measured under two conditions using either 25 mM Tris HCl buffer pH 8.2 at 37°C or 50 mM CAPSO buffer pH 9 at 42 °C. All reaction mixes contained 0.04 mM pyridoxal 5'- phosphate and 2.5 mM dithiothreitol. The concentration of L-ornithine was varied from 5 -250 mM and L-arginine from 0.05 - 5 mM. Protein was determined by the Bradford assay using bovine serum albumin as a standard.

The molar ratios of  $\text{CO}_2$  formed from the labeled substrate were confirmed to be equivalent to the product by analyzing the products by reverse phase HPLC using a standard system for separation of polyamines (48) coupled to a Canberra Packard A140 radiomatic detector as previously described(16). L-[2, 3,- $^3\text{H}$ ] ornithine was used to detect the formation of putrescine and L-[U- $^{14}\text{C}$ ] arginine was used for the formation of agmatine. Authentic U-[ $^{14}\text{C}$ ]agmatine and [2, 3,- $^3\text{H}$ ]putrescine were used as markers.

*Coupled Enzymatic Assays-* Steady-state kinetics of the decarboxylation of L-ornithine (1 to 25 mM), L-arginine (0.5 to 16 mM), or L-lysine (1-50 mM) by the Chlorella virus ODC-like protein was measured spectrophotometrically at 37° C using Sigma Diagnostics carbon dioxide detection kit. Assays were conducted with 100  $\mu\text{M}$  PLP. The Sigma kit couples decarboxylation of substrate to the oxidation of NADH ( $\lambda_{\text{max}} = 340\text{nm}$ ) using phosphoenolpyruvate carboxylase and malate dehydrogenase (49).

*HPLC Analysis of Reaction Products-* The products of an enzymatic reaction with L-arginine and the Chlorella virus ODC-like protein were analyzed by high-performance liquid chromatography (HPLC) using the AccQ\_TAG kit (Waters, Milford, MA) in buffer (5% sodium tetraborate) and fluorescent labeling reagent (6-aminoquinolyl-n-hydrozysuccinimidyl in acetonitrile) (28). The Chlorella virus ODC-like protein was incubated with L-arginine (20 mM) in buffer (15 mM KPO<sub>4</sub>, 1 mM DTT, 0.15 mM PLP) at 37°C for 2.5, 5.0, and 10.0 minutes. Enzyme concentrations of 475 and 950 nM were used in the reactions. Enzymatic reactions were terminated by the addition of trichloroacetic acid (TCA) to a final concentration of 6%. Labeled samples (5 µL) were injected onto an AccQ\_TAG column using previously published buffers and gradient (28). The column was calibrated with known amounts of the following derivatized reagents: Agmatine (RT = 27.8 min), cadaverine (RT = 40.1 min), L-ornithine (RT = 34.8 min), putrescine (RT = 44.5 min), L-arginine (RT = 21.1 min).

*Inactivation of PBCV-1 DC with DFMO or DFMA-* The His-tagged Chlorella virus ODC-like protein in 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8, 300 mM NaCl, 250 mM imidazole, 0.04 mM PLP, 2.5 mM DTT, 0.5 mg/ml BSA (20) was dialyzed to remove DTT and imidazole before passing over a Talon<sup>R</sup> metal affinity column (Clontech, Palo Alto, CA) to remove BSA. After elution and further dialysis an aliquot of protein [104 µg in 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8, 0.04 mM PLP, 0.5 mM DTT] was incubated in the absence or presence of 10 mM DFMO or 1 mM DFMA for 2 h at 37°C in a total volume of 250 µl. Aliquots (5 µl) of each reaction were removed at the times indicated and diluted accordingly to monitor inactivation of the protein.

*Preparation of samples of PBCV-1 DC protein for MALDI-TOF analysis-* Reactions containing 4.2 µg (~100 nmol) of untreated, DFMO or DFMA inactivated PBCV-1 DC protein in a total volume of 100 µl were subjected to tryptic or endoprotease Glu-C digestion using a protease:protein ratio over the



range 1:84 to 1:21 (w/w) as indicated. Samples were digested with modified trypsin in 50 mM  $\text{NH}_4\text{HCO}_3$  buffer pH 8 containing 10 % v/v acetonitrile for 16 h, stopped by the addition of 4  $\mu\text{l}$  of glacial acetic acid and stored frozen until analyzed by MALDI-TOF. Samples were digested with Glu-C in 25 mM  $\text{NH}_4\text{HCO}_3$  buffer pH 7.8 containing 10 % v/v acetonitrile for 2-16 h at 25°C, stopped with glacial acetic acid and stored frozen until analysis by MALDI-TOF using an Applied Biosystems 4700 Proteomics analyzer. Digested samples were evaporated and resuspended in 200  $\mu\text{l}$  of deionized water three times to remove volatile digestion buffers ( $\text{NH}_4\text{HCO}_3$ ) which can interfere with subsequent binding to strong cation exchange resins. The final resuspension was evaporated to ~10  $\mu\text{l}$ , then 1/9th volume of 1.0% TFA was added to bring the final TFA concentration to 0.1%. Using a BioHit Multipipettor, ZipTip SCX tips were equilibrated with three times 10  $\mu\text{l}$  of 0.1 % TFA, then the sample aliquots were pipetted up and down across the resin 15 times to bind and concentrate peptides. The bound peptides were washed with 5 X 10  $\mu\text{l}$  of 0.1% TFA to remove salts, followed by elution of the bound peptides by carefully pipeting 2  $\mu\text{l}$  of freshly prepared 5%  $\text{NH}_4\text{OH}$ /30% MeOH up and down 4-5 times. The droplet containing the eluted peptides was deposited onto a polished stainless steel MALDI target plate (Applied Biosystems) and allowed to dry. After drying, each spot was overlaid with 0.6  $\mu\text{l}$  of freshly prepared CHCA matrix (5 mg per ml recrystallized  $\alpha$ -cyano-hydroxycinnamic acid/ 2 mg/ml  $\text{NH}_4\text{H}_2\text{PO}_4$ /50% acetonitrile/0.1% TFA).

## Results

### *Analysis of PBCV-1 ODC-like Protein Substrate Preference and Activity*

The decarboxylation of L-ornithine or L-arginine by *cvADC* was measured by following  $^{14}\text{CO}_2$  release from L-[1- $^{14}\text{C}$ ]ornithine or L-[U- $^{14}\text{C}$ ]arginine (Table I). These studies were carried out in the laboratory of A.E.

Pegg (Department of Cellular and Molecular Physiology, Pennsylvania State University College of Medicine, Hershey, PA). In studies carried out in Tris-HCl buffer, pH 8.2 at 37°C, the protein was highly active on L-arginine with a  $K_m$  of 0.45 mM and a  $k_{cat}$  of 15 s<sup>-1</sup>. It also had detectable activity on L-ornithine; however, while the  $k_{cat}$  was only slightly lower, the  $K_m$  was about 400 times higher ( $K_m$  of 180 mM; Table I), giving an overall difference in substrate preference of 550-fold as measured by  $k_{cat}/K_m$ . These results contrast with a previous report of a  $K_m$  of 0.78 mM for L-ornithine (20). The previously reported assays were conducted in CAPSO buffer pH 9 at 42°C. The assays were therefore repeated using these conditions and the ability to decarboxylate L-ornithine was slightly improved ( $K_m$  of 46 mM; Table I). However, even under these conditions, the protein still preferred L-arginine as a substrate with a  $K_m$  two orders of magnitude lower than that for L-ornithine. The products of the reaction were identified by HPLC (16, 48) after the protein had acted upon either L-[U-<sup>14</sup>C]arginine or L-[2, 3-<sup>3</sup>H]ornithine. As expected, [<sup>14</sup>C] agmatine was found in stoichiometric amounts with <sup>14</sup>CO<sub>2</sub> when L-[U-<sup>14</sup>C]L-arginine was decarboxylated, and [<sup>3</sup>H]putrescine was formed in equivalent amounts when L-[2, 3-<sup>3</sup>H]ornithine was the substrate.

I measured the substrate preference of the Chlorella virus ODC-like protein using a NADH-oxidation coupled assay which allows for continuous determination of decarboxylation via the decrease in absorbance at 340 nm. Steady-state decarboxylation was measured for L-arginine, L-ornithine and L-lysine over a wide range of concentrations (Figure 2-4). The relative substrate preferences for L-arginine over L-ornithine were similar to those observed by the <sup>14</sup>CO<sub>2</sub>-release assay, however the  $k_{cat}/K_m$  on both substrates was 8-fold lower under the conditions of the spectrophotometric assay (Table I). L-Lysine was also a substrate for the reaction and was decarboxylated with similar efficiency to L-ornithine (Figure 2-4). This result is in contrast to the ODCs from both mouse

and *T. brucei* for which the  $k_{\text{cat}}/K_m$  for L-lysine is 300-fold lower than for L-ornithine (21, 47). The similar catalytic efficiency of the Chlorella virus ODC-like protein on L-ornithine and L-lysine, however is reminiscent of the bacterial Lys/OrnDC from *Selenomonas ruminantium* (43).

A further study of the activity of the Chlorella virus ODC-like protein with L-arginine was conducted by measuring the rate of product (agmatine) formation via HPLC. For this method, substrate and products were labeled with a fluorophore that forms covalent bonds with primary and secondary amines (See experimental procedures). Based on the rate of agmatine formation, the  $k_{\text{cat}}$  of the Chlorella virus ODC-like protein with L-arginine was calculated to be  $15 \text{ s}^{-1}$ , a value that agrees well with the previously determined rate constants (Figure 2-5 A). I also determined the substrate dependence of the Chlorella virus ODC-like protein with L-arginine via this HPLC method. The calculated  $K_m$  based on the HPLC data was 1.5 mM (Figure 2-5 B). It has been previously observed for amino-acid decarboxylases, that decarboxylation-dependent transamination occurs as a rare side reaction (less than 1 in  $10^4$  events) for physiological substrates while it occurs more frequently for non-physiological substrates due to improper protonation at the C4' of PLP (50). Transamination is an exceedingly rare event, as is the case with ODC and its cognate substrate, L-ornithine (32), as the turnover rate agrees with the rate determined by other methods. Furthermore, it has been shown that commercially available L-arginine has up to 1% contamination with L-ornithine (3). Thus, the turnover rate of the Chlorella virus ODC-like protein with L-arginine observed with the coupled-enzyme assay could reflect the formation of putrescine from L-ornithine. The HPLC profile of the reaction products of L-arginine with the Chlorella virus ODC-like protein demonstrates that the observed rate constant is, indeed, that for the decarboxylation of L-arginine, as the formation of putrescine is undetectable.

*Inhibition of Chlorella virus ODC-like protein with DFMO and DFMA*

Morehead et al. previously reported that DFMA was a more potent inhibitor of CVADC than was DFMO (20). To confirm this finding, members of the Pegg laboratory undertook similar inhibition experiments with DFMA and DFMO (Figure 2-6 A). The Chlorella virus ODC-like protein was irreversibly inactivated by incubation with either DFMO or DFMA. Consistent with previous reports, DFMA inactivated the enzyme more rapidly than DFMO; after a two hour incubation with 1 mM DFMA only 0.34% of the activity remained, while it required 10 mM DFMO to reduce the enzyme activity to 4.4% over the same incubation time.

Mouse ODC is inactivated by DFMO with the formation of a covalent S-[(2-(1-pyrroline)]methyl adduct at Cys-324 (major product. 90%) and a minor product (c. 10%) at Lys-48 (51, 52). Although previous studies have shown that DFMA is an irreversible inhibitor of bacterial and plant ADCs (53), the adduct formation site has not been identified. However, based on the inactivation of mouse ODC, the likely sites of interaction in the native Chlorella virus ODC-like protein are residues Cys-347 and Lys-71, which correspond to Cys-360 and Lys-69 in mouse ODC.

Endoproteinase Glu-C digestions of untreated Chlorella virus ODC-like protein or of protein inactivated by DFMO or DFMA were analyzed by linear MALDI-TOF MS in positive ion mode, using Chlorella virus ODC-like protein and Glu-C self-digested peaks for internal calibration of the spectra. The spectra of the inactivated enzyme had new peptide fragments with masses corresponding to adducts on the Lys<sup>321</sup>-Glu<sup>361</sup> peptide (KSVPTPQLLRDVPDDEEYVPSVLYGCTCDGVDVINHNVALPE) containing Cys347 (Figure 2-6 B). The presence of this peptide in the Glu-C digest and the absence of shorter fragments cleaved at the internal DDEEY sequence is probably due to the digestions being carried out in NH<sub>4</sub>CO<sub>3</sub> buffer which increases the

specificity of Glu-C but restricts activity and the known inactivity of Glu-C to cut at clusters of acidic residues (54). As seen in Figure 2-5B, these masses ( $m/z$  4681.0 and 4740.8) were not observed in the control digests, while all three digests contained a peak slightly smaller than the theoretical value (4599.2 average mass) of the equivalent unmodified fragment. (It should be noted that within the limited resolution of these linear spectra, the potential contributions of the unmodified peptide and either of two similar mass Glu-C self-digest peptides to the observed peak cannot be separated definitively).

The new peptide which appears after DFMO inactivation ( $m/z$  4681.0) is 81.8 Da heavier than the unmodified peptide Lys<sup>321</sup>-Glu<sup>361</sup>, which within experimental error agrees with the theoretical 81.1 Da difference previously observed for the S-[(2-(1-pyrroline))]methyl-cysteine adduct formed by DFMO at Cys360 of mouse ODC (52). In the digest of the protein inactivated by DFMA, the new peptide observed at  $m/z$  4740.8 is 141.6 Da larger than the unmodified peptide Lys<sup>321</sup>-Glu<sup>361</sup>. This value is within experimental error of the theoretical adduct mass of 140.2 Da which would occur with the analogous linear DFMA adduct to a Cys in the peptide. This result is consistent with DFMA inactivation occurring similarly to DFMO inactivation of ODC, with decarboxylation of the DFMA-PLP Schiff base leading to the loss of a fluoride ion to generate a reactive electrophilic conjugated imine that binds covalently to Cys-347. Subsequent elimination of the second fluoride anion followed by an internal transaldimination reaction with Lys-71 would generate the adduct. This adduct cannot cyclize in the same way as the adduct derived from DFMO because of the replacement of the terminal amino group with a guanidino group. The Lys<sup>321</sup>-Glu<sup>361</sup> peptide does contain another Cys residue at position 345. Since we were unable to produce interpretable  $ms/ms$  spectra of the modified peptides, addition to this site cannot be ruled out but, based on the existing data on the structure and inactivation of

mammalian and trypanosomal ODC by DFMO (26, 28, 30, 52), Cys-347 is the likely site of attachment.

Assuming that the peaks in the DFMO- and DFMA-inactivated protein spectra at  $m/z$  4597.9 and 4598.6 respectively do represent unmodified peptide, then not all of the Chlorella virus ODC-like protein was modified at this site. Although MALDI analysis is not quantitative, the relative sizes of the modified and parent peaks are similar, which suggests that only about half the protein was altered at Cys-347, in contrast to the almost total observed loss of enzyme activity, with >99% inhibition (with DFMA) and >90% (with DFMO). One explanation for this discrepancy is that interaction with another site such as Lys71 occurs to a larger extent than with the inactivation of mouse ODC with DFMO. However, we were unable by MS to identify a putative DFMO or DFMA adduct to peptides containing Lys71 after digestion by Glu-C or by trypsin (where an adduct at Lys71 would prevent tryptic cleavage at that site). Alternative explanations could be that: (a) only one of the two protein subunits making up the homodimer need to be modified to cause loss of catalytic activity; (b) that analytical workup resulted in the preferential loss of some of the modified peptide; (c) that, although we did not find any other peptides that were significantly different between control and DFMO- or DFMA-inactivated Chlorella virus ODC-like protein in the analysis of either Glu-C or trypsin digests, modification of an additional residue at the active site occurs. The sequence coverage was about 50% in the Glu-C digests, so modifications of the unrepresented proteolytic fragments would not have been observed; (d) that some proportion of the recombinant protein extract may be enzymatically inactive prior to the start of the reaction, and thus unable to react with DFMO or DFMA. It is also possible that a significant portion of the “unmodified” peak represents Glu-C self-digestion fragments.

## Discussion

The PLP-dependent decarboxylases that belong to the  $\beta/\alpha$ -barrel fold are from both bacterial and eukaryotic origin, and they include enzymes capable of decarboxylating a range of basic amino acid substrates (24). The eukaryotic ODCs, and a group of bacterial enzymes with dual specificity on L-ornithine and L-lysine [e.g. *S. ruminantium* Lys/OrnDC ] (Figure 1-9) (43), share high sequence similarity, while the enzymes with specificity for L-arginine and other basic amino acids share very low sequence similarity with the ornithine specific enzymes (Fig. 2). Thus, the Chlorella virus ODC-like protein is an anomaly because it shares greater sequence identity with the ODCs, while it strongly prefers L-arginine as a substrate. Subsequent to the publication of my characterization of this enzyme's specificity profile, it has appropriately been reclassified as the Chlorella virus arginine decarboxylase (*cvADC*). The preference for L-arginine over L-ornithine and L-lysine is reflected in the 550-fold higher  $k_{cat}/K_m$  observed for L-arginine. DFMA is a more potent inhibitor of the *cvADC* than DFMO, although both compounds form adducts at the same site. Thus, the data clearly indicate that *cvADC* is a legitimate ADC and not an ODC, as a previous study claimed (20).

The PBCV-1 genome is the first viral genome known to encode polyamine biosynthetic enzymes (19). In addition to encoding *cvADC*, the virus contains the enzymes necessary to produce putrescine from agmatine (agmatine iminohydrolase and N-carbamoylputrescine amidohydrolase) (Figure 1-4), and it also contains homospermidine synthase, but not spermidine synthase, suggesting that the end product of the pathway in the virus is homospermidine (55). It is interesting to speculate on the evolutionary driving force for the substrate specificity of *cvADC* for L-arginine. In algae, including chlorella isolates, amino acid analyses demonstrated that L-arginine is one of the most abundant amino acids, while L-ornithine is one of the least abundant (41). Thus the lack of L-

ornithine in these cells suggests that the switch in substrate specificity of PBCV-1 DC to L-arginine was required for efficient production of putrescine.

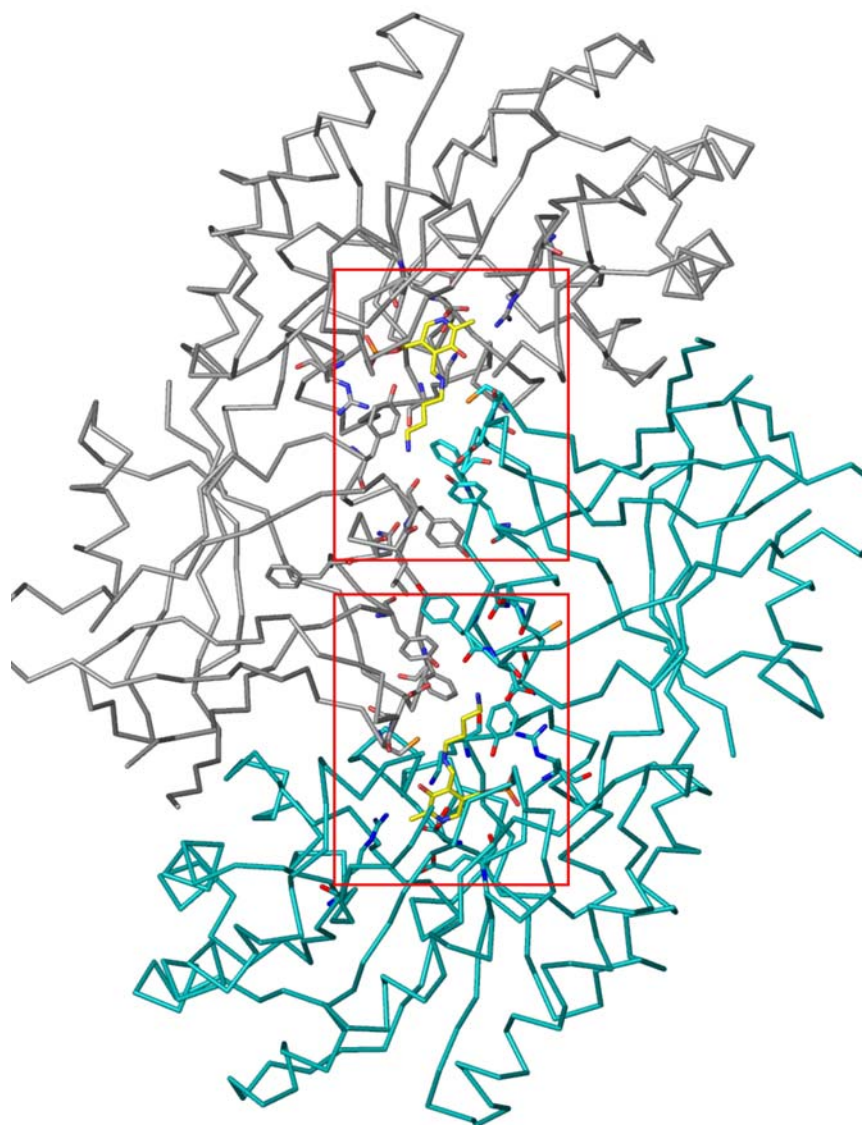
The high amino acid sequence identity (40%) between *cvADC* and eukaryotic ODC suggests that the *cvADC* is more closely related to enzymes with ornithine specificity than to bacterial ADCs, despite their common substrate preference. A potentially important link in this evolution is the observation that a group of bacterial enzymes which also share strong sequence similarity to the eukaryotic ODCs, have dual specificity on L-lysine and L-ornithine (43). *cvADC* groups most closely to these bacterial enzymes in a phylogenetic analysis, and these sequences form a bridge between the eukaryotic ODCs and the more distantly related ADCs from bacteria and plants (Fig. 2). Interestingly, *cvADC*, like the bacterial enzymes, has equal activity on both L-lysine and L-ornithine, though it is distinct in preferring L-arginine. These observations suggest that *cvADC* may have been acquired from a bacterial source that contained a dual specificity Lys/OrnDC, and that it has evolved a minimum set of amino acid substitutions to switch specificity without adversely affecting activity. Interestingly, the *cvADC* homospermidine synthase is also more closely related to bacterial enzymes than to plant enzymes (55), suggesting the entire polyamine biosynthetic pathway in the virus may have been acquired from bacteria, or from a common ancestor.

The role of putrescine and homospermidine in viral pathogenesis remains unclear, since the host cell also makes both polyamines. However, the virus is known to inhibit protein translation in the host (19, 56), and altered levels of polyamines could play a role in this inhibition. Eukaryotic translation initiation factor 5A (eIF-5A) plays an essential role in translation of selective messages required for cell division and proliferation (57). This factor is synthesized as an inactive precursor that is modified to its active form by the covalent attachment of hypusine via two enzymatic steps. The first step is catalyzed by deoxyhypusine



synthase and utilizes spermidine as the donor substrate. Recently it was reported that this same enzyme can catalyze the reverse reaction using putrescine as an acceptor producing homospermidine and unmodified eIF-5A (58). Putrescine levels increase significantly in chlorella cells after infection with PBCV-1 (55), and this increase could potentially alter the levels of hypusine modified eIF-5A thereby affecting host cell translation.

Regardless of the purpose of *cvADC* in the life cycle of the Chlorella virus, the conclusion that *cvADC* is a legitimate ADC with relatively high sequence identity to eukaryotic ODC presents an opportunity to isolate the specificity determinants of the Group IV decarboxylases. As mentioned in Chapter I, the previously known structures of Group IV decarboxylases, ODC and DAPDC, demonstrated that large sequence divergence is correlated with numerous structural differences in the active sites between these enzymes of differing specificity (Figure 1-10). Having established *cvADC*'s preference, the high sequence identity between ODC and *cvADC* can now be harnessed to understand at the primary and tertiary sequence level what positions and residues are necessary for determining substrate specificity. In the next chapter, I present my efforts to determine the three-dimensional structure of *cvADC* and offer insights into the structural basis of specificity of Group IV decarboxylases.



**FIGURE 2-1**

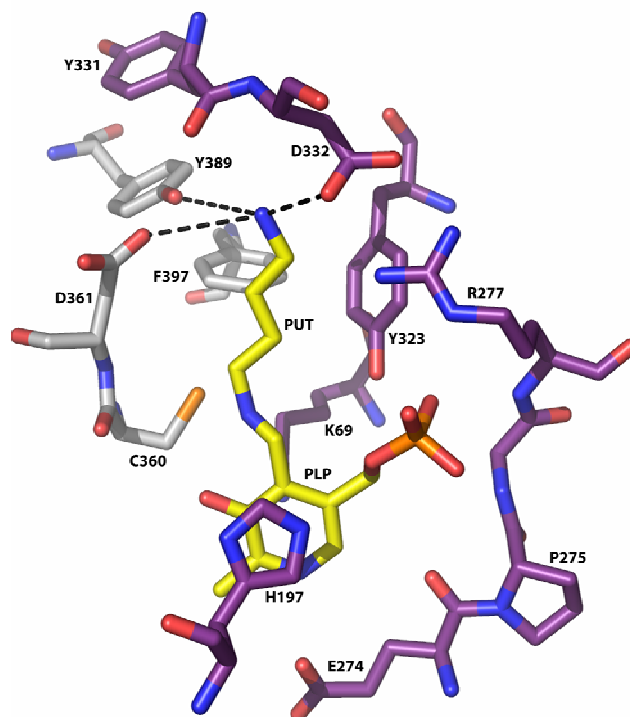
The two active sites of eukaryotic ODC (in red boxes) are formed at the dimer interface and include residues from the N-terminal of monomer A (cyan) and the C-terminal of monomer B (gray). The PLP-putrescine complex is shown as yellow sticks.

	70	90	110	130	150	Id										
cvdc	HYAVKCNNDVLLKTMCDKN--VNFD	CASSSEIKKVIQIGVSP----	SRIIFAHTMKTIDDLIFAKDQG----	VDIATFDSFELDKIHTYHF----	NCKMILRIR	37%										
hODC	FYAVKCNDSKAIKTLAATG--TGFD	CASKTEIQLVQSLGVFP----	ERIYANPCQVQSIKYAANNG----	VQMMTFDSSELMKVARAHP----	KAKVLRIA	92%										
tbODC	FYAVKCNDDWRVLGTLAALG--TGFD	CASNTEIQVRGIGVFP----	EKIYANPCQISHIRYARDSG----	VDVMTFDCVDELEKVAKTHP----	KAKMVLRIA	63%										
srLODC	FYAMKANPTPEILSLLAGLG--SHFD	VASAGEMEILHELGVDPG----	SQMIYANPVKDARGLKAADYN----	VRRFTFDDPSEIDKMAKAVP----	GADVLVRIA	36%										
scODC	FYAVKCNPDTKVLSLLAELG--VNFD	CASKVEIDRVLSMNISF----	DRIVYANPCQVASFIRYAASN----	VMKSTFDNVEELHKIKKFHP----	ESQLLRIA	45%										
mtDAPDC	HYAAKAFLCSEVARWISSEEG--LCLD	VTGGELAVALHASFPF----	ERITLHGNNKSVSELTAAVKAG----	VGHIVVDSMTIERLDAIAGEAGIVQDVLVRLT		21%										
ecDAPDC	RFAQKACSNIHILRLRMREQG--VKVD	SVLSGIERALAAAGYNPQTHEDD	IVFTADVADQATLERVSELQ----	I-PVNAGSVMDLQGLQGVSP----	GHRVWLRVN	21%										
mjDAPDC	AYAYKANANLAITRLAKLG--CGAD	VVSGGELYIAKLSNVFS----	KKIVFNGCNCKTEEIMGIEAN----	IRAFNVDSISELILINETAKELGETANVAFRIN		24%										
ecADC	VYPILKVNQHRRVIESLIHSGEPLGL	EAGSKAELMAVLAHAGMT----	RSVIVCNGYKDREYIRLALIGEKG	MHKVYLVIEK4SEIAIVLEAE----	RLNVVFLRG	22%										
ypADC	VYPILKVNQHRRVIESLVNSGEFLGL	EAGSKAEMMAVLAHAGMT----	RSVIVCNGYKDREYIRLALIGEKG	MHKVYLVIEK4SEIKMVEAE----	RLNVVFLRG	22%										
	170	190	210	230												
cvdc	CDD-----FNATVQLGNKFGAN-ED--	EIRHLEYAKQLD-I-EVIGISFHV	GSGRNPEAYYRAIKSSKEAFNEA----	ISVGH-KPYILDI	GGGLHADI											
hODC	TDD-----SKAVCRLSVKFGAT-LR--	TSRLLLERAKELN-I-DVVGVSFHV	GSCTDPETFVQAISDARCVDMG----	AEVGF-SMYLLDI	GGGFPQSE											
tbODC	TDD-----SLARCRLSVKFGAK-VE--	DCRFLEQAKKLN-I-DVTGVSFHV	GSSTDAFTFAQISDSRFVDMG----	TELG-FNMHILDI	GGGFPQTR											
srLODC	VRN-----NKALVDLNTKFGAP-VE--	EALDLLKAAQDAG-L-HAMGICFHV	GSQSLSTAAYEEALLVARRLFDEA----	EEMGM-HLTDLDI	GGGFPVDP											
scODC	TDD-----STAQCRSLTKYGCE-ME--	NVDVLLKAIKELG-L-NLAGVSFHV	GSASDFTSLYKAVRDARTVFDKAA----	NEYGLFPLKILDV	GGGFQFE											
mtDAPDC	VGVE---AHTEFISTAHEDQKFGLS-	VASGAAMAARRVFATDHL-RLVGLHS	HIGSQIFDVDFGFEAAHVRVIGLLRDV	VGEFGEKTA-QIATVDI	GGGLGISY											
ecDAPDC	PGPG---HGH5QKTITGGENSKHGIW-	YTDLPAALDVIQRHH-L-QLVGIHMH	HIGSGVDYAHLEQVCGAMVRQVIEF----	GQ-DLQATSA	GGGLSVPY											
mjDAPDC	PNNV---PKTHPKISTGLKNKFGLD-	VESGIAMKMALEMEYV-NVVGVC	HIGSQLTDISPFIEETRKVMDFVEL----	KEEGI-EIEDVNL	GGGLGIPY											
ecADC	VRARLASQSGKQWQSSGGSEKSKFGLA-	AT--QVLQVETLREAGRLDLSQLLHF	HIGSQMANIRDIATGVRESARFYVEL----	HKLGV-NIQCFDV	GGGLGVDI											
ypADC	VRARLASQSGKQWQSSGGSEKSKFGLS-	AT--QVLQVDMLEANSLESLLQLLHF	HIGSQLSNIRDISGVRESARFYVEL----	HKLGV-NIQCFDV	GGGLGVDI											
	250	270	290	310												
cvdc	D---EGELSTYMSDYINDAIKDFFF--	ED---TVTIVA	EPGRFAEHYSVLATQVIGKVRV----	-----DGLYE												
hODC	D---VKLFEEITGVINPALDKYFF--	S---DS---GVRI	IAEPGRYYVASAFTLAVNI	IAKKIVLKEQTSDDDED-----	ESSEQTFM											
tbODC	D---APLKFEIAGVINNALEKHF--	F---DL---KLTIVA	EPGRYYVASAFTLAVNI	IAKKIVPGVQTDVGAHA-----	ESNAQTFM											
srLODC	AKG-LNVDLAAMEAINKQIDRLFF--	-----DTAVMT	EPGRYMGCTAVNLVTSVIGTKTR----	-----GEQPF												
scODC	-----SFKESTAVLRLALEEFF--	V---GC---GVDI	IAEPGRYVATAFTLASHVIAKRL----	-----SENEAM												
mtDAPDC	LPSSDPPIAELAAKLGTIVSDEST--	AV---GLP---TPKL	VVPEPRAIAGFGTITLYEVGTVKDVSAT----	-----AHRRY												
ecDAPDC	QGGEEAVDTHEHYGLWNAAREQIAR--	HL---GH---FVKLEI	EPGRFLVAGSGVLITQVRSVKMG-----	-----SRHF												
mjDAPDC	YKKQIPTQKDLADAIINTMLKYKD--	K---VE---MPLIL	EPGRSLVATAGYLLGKVHHIKETP----	-----VTKW												
ecADC	EGTR---SQSDCSVNYGLNEYANNI	IWAIGDACEENGLPHTVIT	ESGRVATAHTVLVSNIGVERNEYVPT	TAPAEADAPRALQSMWETQEMHFGTRSLR												
ypADC	EGTR---SQSDCSVNYGLNEYANNV	IWIGDADCEHGLPHTVIT	ESGRVATAHTVLVSNIGVERNEFCEP	QPEAGAPRALLESWDTQEMQEFENRRSLR												
	330															
cvdc	YFFNES-----	-----TYGGF-SNVIF	EKSVFPT-----	-----												
hODC	YYVNDG-----	-----VYGSF-NCILY	DHAAVKP-----	-----												
tbODC	YYVNDG-----	-----VYGSF-NCILY	DHAVVRP-----	-----												
srLODC	YILDEG-----	-----IYGEF-SGIMY	DHWTY-F-----	-----												
scODC	IYTNDG-----	-----VYGNM-NCILE	DHQEPHRTLYHNLEFHYDDFES													
mtDAPDC	VSDVGG-----	-----MSD-N-IR	TALYGAQYDV-----	-----												
ecDAPDC	VLDVAG-----	-----FND-L-MRP	AMYGSYHHI-----	-----												
mjDAPDC	VMDIAG-----	-----MNDM-MRP	AMYEAHHI-----	-----												
ecADC	EWLHDSQMDLHDHIGYSSGTFSLQERAWAE	QLYLSMCHEVQKQLDPQNAHRFP	IIDELQERMADKMYVNFSLFQSMFPAWGI	DQLFPV-----												
ypADC	EWLHDSQMDLHDHVTQAHAGMLDLTHRAWAE	QLYLSICNEIQKQLDFSNRAHRFP	IIDELQERMADKLYVNFSLFQSMFPAWGI	DQLFPV-----												
	350	370	390													
cvdc	---QLLR---DVPDDEEYVPSVLYGCT	CDGVDVIN-----	HNVALPE-LH--IGDWVYF	FPFGAYTNVLTTSFNGFGE-YDVYII-	-----											
hODC	---LLQK---RPKPDERYSSSIWGP	CDGLDRIV-----	ERCDLPE-MH--VGDWMLFENM	GAYTAAASTFNGFQR-PTIYYVM	SGFAWGLMQ--	F										
tbODC	---LPQR---EPINKEKLYPSSVWGP	CDGLDQIV-----	ERYYLPE-MQ--VGEWLLFEDM	GAYTVGTSSFNGFQS-PTIYYV	VSGLFDHVRE--	I										
srLODC	---LHCF---GKGNKPFSTFGGFS	CDGIDVLY-----	RDFMAPE-LK--IGDKVLVTEM	GSYTSVSATRFNGFYL-APTII	FEDRPMWQLMKQ--	L										
scODC	TTAVLDSI--NKTRSEYFYKSVI	WGP	CDGLDIA-----	KEYYMKHDI--VGDWFF	FPALGAYTSSAATQFNGFEQ	TADIVYDSQPEYAA	RLTDEDD--									
mtDAPDC	---RLVS---RVSD-APPV	PARLVGKH	CSGDIIV-----	RTVVPDDIR--PGDLVAV	AATGAYCYLSSRYNMVGR-P	AVVAVHAGNARLVLR--	E									
ecDAPDC	---SALA	ADGRSL	EHAFTVETVAGP	CESGDVFTQ	QEGGNV--ETRALPE-VK--	AGDYLVLHDTGAYG	ASMSSNYSRPL-L	PELVFDNGQARLIRRR--	Q							
mjDAPDC	---INCK---VK---NEKE	VVSIA	AGGL	CESSDVFG-----	RDRDLK-VE--VGDV	LAI	FDV	GAYGISMANNYNARGR-PR	MVLTSSKKGVFLIRER--	E						
ecADC	---LP	LEGLDQV	PERRAVLLDIT	CDSDG	AI	DHYIDG	GIAT	TMMPPE-YDPEN	PMLGFFMV	GAYQ	EILGNMHNLF	FGD-TEAVDV	FV-----			
ypADC	---LP	LEGLDKP	PERRAVLLDIT	CDSDG	IT	DHYIDG	GIAT	TMMPPE-YDPEN	PMLGFFMV	GAYQ	EILGNMHNLF	FGD-TA	AVDVYV	FPDGS	VEVELS	DEGD

**FIGURE 2-2**

An alignment of 10 Group IV decarboxylase sequences. The sequences in the alignment are Chlorella virus ODC-like protein (cvDC), human ODC (hODC), *T. brucei* ODC (tbODC), *S. ruminantium* L/ODC (srLODC), *S. cerevisiae* ODC (scODC), *M. tuberculosis* DAPDC (mtDAPDC), *E. coli* DAPDC (ecDAPDC), *M. jannaschii* DAPDC (mjDAPDC), *E. coli* ADC (ecADC), and *Y. pestis* ADC (ypADC). The sequence identity shared between these sequences and mouse ODC is shown to the right of the sequences. Residue numbering is based on the mouse ODC sequence. Residues within 5 Å of PLP are highlighted in red.

B.



	332		361	
<b>MODC</b>	GSFNCILY <b>D</b> HAHVKA	LLQKRPKPDE-----	KYYSSSIWGPT <b>CD</b> GLDRIVERCNLP-EMHV	377
<b>tbODC</b>	GSFNCILY <b>D</b> HAVVRP	LPQREPIPN-----	KLYPSSVWGPT <b>CD</b> GLDQIVERYYLP-EMQV	375
<b>ChODC</b>	GSFNCILY <b>D</b> GQNPGY	KVVRSPLMADSTDS-	RTFLSTLWGPT <b>CD</b> SADCVYKDVTLP-VLRN	343
<b>cvDC</b>	GGFSNVIF <b>E</b> KSVPTP	QLLRDVPDDE-----	EYVPSVLYGCT <b>CD</b> GVVDVINHNVALP-ELHI	341
<b>ntODC</b>	GSMNCVLY <b>D</b> HATVNA	TPLAVLSNRSNVTCG	KTFTPTTVFGPT <b>CD</b> ALDTVLRDYQLP-ELQV	395
<b>scODC</b>	GNMNCILF <b>D</b> HQEPHP	RTLYHNLEFHYDDFE	STTAVLDSINKTRSE	429
			YPYKVS IWGPT <b>CD</b> GLDCIAKEYYMKHDVIV	

### FIGURE 2-3.

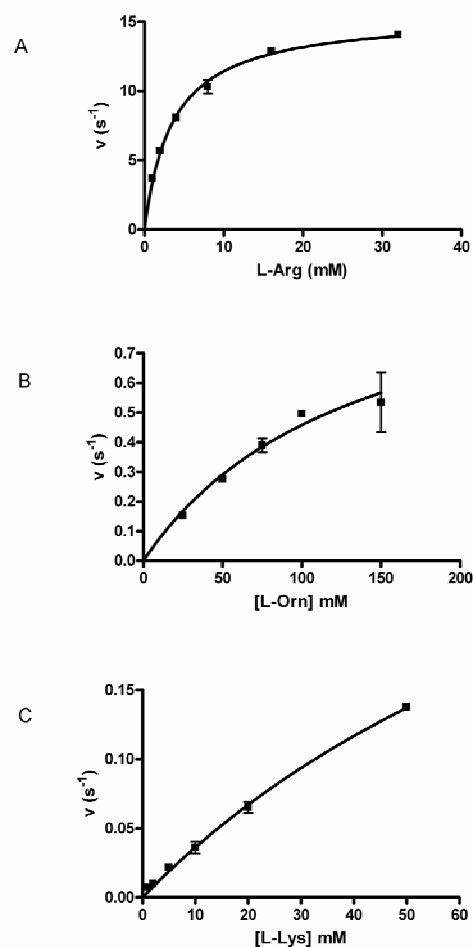
**A.** Structure of the active-site of *T. brucei* ODC complexed with putrescine (PDB ID 1F3T). Residues from the N-terminal domain of monomer A (purple) and the C-terminal domain of monomer B (gray) that are within 4.5 Å of putrescine are displayed.

**B.** Partial sequence alignment for a representative set of eukaryotic ODCs and *cvDC*. Sequences accession numbers are as follows: *cvDC*, NP\_048554; *Chlamydomonas reinhardtii* ODC, CAE46409; *tbODC*, *Trypanosoma brucei* ODC, 1NJJ\_A; Mouse ODC, P00860; *Nicotiana tabacum* ODC, AAQ14852; *Saccharomyces cerevisiae*, NP\_012737.

<b>Substrate</b>	<b>K<sub>m</sub> (mM)</b>	<b>k<sub>cat</sub> (s<sup>-1</sup>)</b>	<b>k<sub>cat</sub>/ K<sub>m</sub> (s<sup>-1</sup>M<sup>-1</sup>)</b>	<b>Buffer</b>	<b>pH/ °C</b>
<b>L-Orn</b>	<b>180</b>	<b>10</b>	<b>60</b>	<b>Tris</b>	<b>8.2/37</b>
<b>L-Orn</b>	<b>46</b>	<b>14</b>	<b>3.0 x 10<sup>2</sup></b>	<b>CAPSO</b>	<b>9.0/42</b>
<b>L-Arg</b>	<b>0.45</b>	<b>15</b>	<b>3.3 x 10<sup>4</sup></b>	<b>Tris</b>	<b>8.2/37</b>
<b>L-Arg</b>	<b>0.48</b>	<b>24</b>	<b>5.0 x 10<sup>4</sup></b>	<b>CAPSO</b>	<b>9.0/42</b>

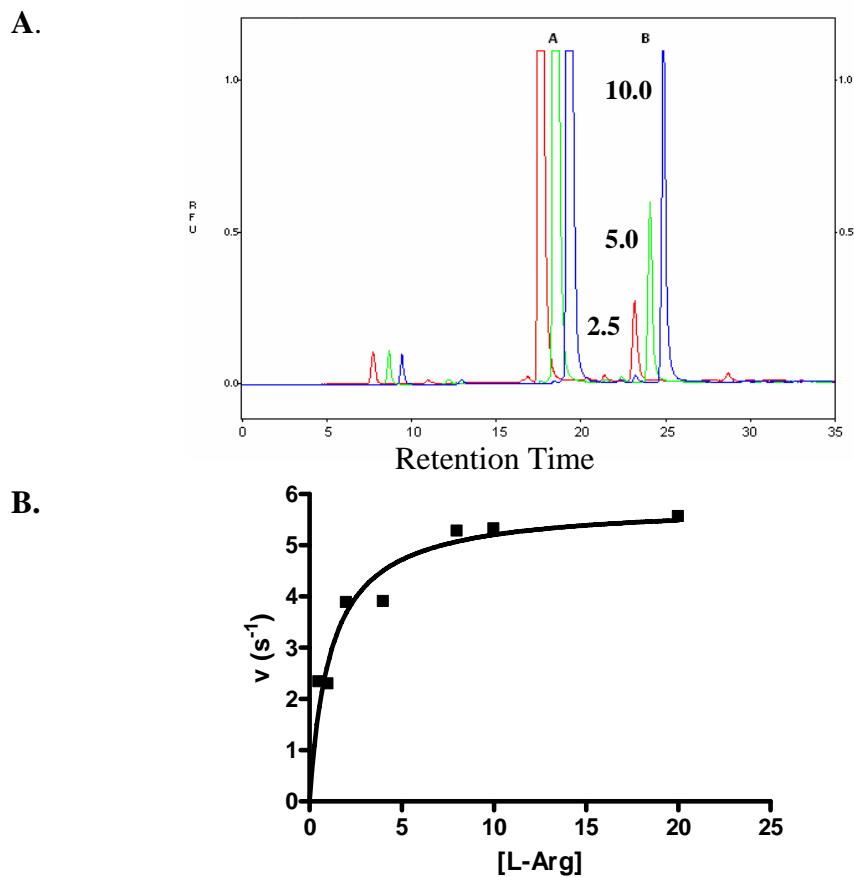
**TABLE I.**

The steady-state kinetic parameters of the Chlorella virus ODC-like protein on L-ornithine and L-arginine were obtained in at different pHs and temperatures. Assays were performed using the <sup>14</sup>CO<sub>2</sub>-trapping method by members of the Pegg laboratory (Pennsylvania State University College of Medicine, Hershey, PA).



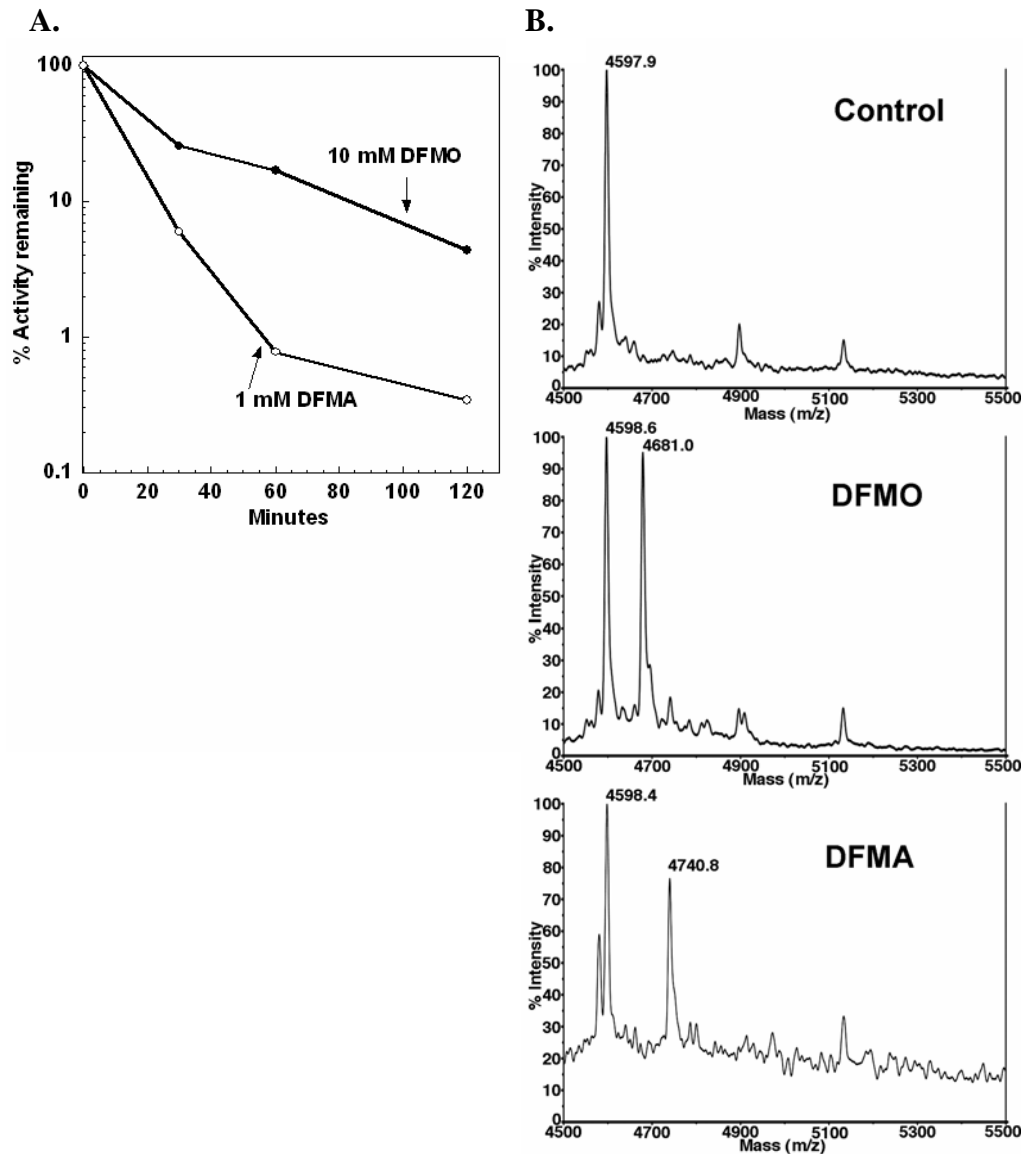
**FIGURE 2-4**

The Chlorella virus ODC-like protein prefers L-arginine (**A**) over L-ornithine (**B**) or L-lysine (**C**). Decarboxylation was coupled to NADH oxidation (see Experimental Procedures). Data points (filled squares) are the average of at least two experiments and error bars represent standard error of the mean. Data was fit by non-linear regression to the Michaelis-Menten equation and the fitted parameters are: L-arginine ( $K_m = 3.5 \pm 0.2$  mM,  $k_{cat} = 15 \pm 0.3$   $\text{s}^{-1}$ ), L-ornithine ( $K_m = 136 \pm 44$  mM,  $k_{cat} = 1.0 \pm 0.2$   $\text{s}^{-1}$ ) and L-lysine ( $K_m = 115 \pm 27$  mM,  $k_{cat} = 0.45 \pm 0.08$   $\text{s}^{-1}$ ). Data was analyzed with GraphPad Prism version 4 using the Michaelis-Menten equation.



**FIGURE 2-5**

HPLC analysis of the reaction products of Chlorella virus ODC-like protein with L-arginine. **A.** L-arginine (peak A) was incubated with enzyme for 2.5 (red profile), 5.0 (green) and 10 minutes (blue). The rate formation of agmatine (peak B) was calculated to be  $15 \text{ s}^{-1}$ . **B.** The substrate dependence curve for enzyme with L-arginine demonstrates that the  $K_m$  is 1.5 mM.



**FIGURE 2-6**

Inactivation of Chlorella virus ODC-like protein by DFMA and DFMO. **A.** The loss of decarboxylase activity of enzyme after incubation with 1 mM DFMA or 10 mM DFMO as indicated. **B.** Part of the linear MALDI-TOF MS in positive ion mode of Glu-C digested Chlorella virus ODC-like protein (upper plot), protein inactivated with 10 mM DFMO (middle plot) and protein inactivated with 1 mM DFMA (lower plot).



## CHAPTER THREE

### X-RAY STRUCTURE DETERMINATION OF THE CHLORELLA VIRUS ARGININE DECARBOXYLASE

#### Introduction

The observation that *cvADC* is closely related to eukaryotic ODC while preferring L-arginine over 600-fold over ornithine provides a unique opportunity to study the structural basis of substrate specificity in Group IV decarboxylases. As discussed in Chapter I, a comparison of *T. brucei* ODC (PDB ID: 1F3T) and *M. jannaschii* DAPDC (PDB ID: 1TWI) reveals that low sequence conservation (~ 20% sequence identity) yields a relatively large divergence in the structure (Figure 1-10). The RMSD of C $\alpha$  atoms (for 350 aligned residues) between ODC and DAPDC is 2.5 Å. Despite the difference in structure, these enzymes share many structural features that are characteristic of this fold. All Group IV PLP-dependent decarboxylases are obligate homodimers with two active sites forming at the subunit interface between the N-terminal  $\beta/\alpha$  barrel domain, which houses PLP, and the two C-terminal  $\beta$ -barrel domains (Figure 2-1) (23, 59). One of the most structurally well-characterized Group IV enzymes is the *Trypanosoma brucei* ODC. Ligand-bound structures of *tbODC* have been solved with putrescine, DFMO and D-ornithine (28, 30, 32). The ligand is in an extended conformation at the subunit interface and each ligand makes similar contacts (Figure 1-8). The  $\alpha$ -amino group forms a Schiff base with PLP and the  $\delta$ -amino group interacts with two aspartic acid residues, D332 from the C-terminal  $\beta$ -barrel domain of one monomer and D361 *tbODC* from the C-terminal  $\beta$ -barrel domain of the opposite monomer. The DAPDC active site is arranged similarly. Based on the *Methanococcus jannaschii* DAPDC structure, E373 *mjDAPDC* is thought to play a similar role as D361 *tbODC* and the carboxylate of the bound lysine

product makes contacts with R307 DAPDC (R277 *tb*ODC) from the  $\alpha/\beta$  barrel domain (Figure 3-1). However, overlays of the *mj*DAPDC structure onto the *tb*ODC structure show numerous changes both within and distant from the active site (Figure 1-10). A plot of RMSD values of 350 aligned C $\alpha$  atoms residues reveals that nearly 100 residues deviate by more than 2.5 Å between *tb*ODC and *mj*DAPDC (Figure 3-2). Thus, isolating the determinants of substrate specificity in the Group IV PLP-dependent decarboxylase fold is difficult.

*cv*ADC presents a unique opportunity to determine the residues and positions in Group IV decarboxylases that play a dominant role in substrate specificity. In the previous chapter, I established that this enzyme preferred L-arginine over L-ornithine as a substrate. Coupled to the fact that *cv*ADC shares high sequence identity with eukaryotic ODC, it seems that *cv*ADC had undergone rapid evolution to change function from an ODC to an ADC. Comparison of the active site residues of *cv*ADC and ODC, based on the primary sequences, suggested that a relatively limited number of substitutions might account for the specificity difference (Figure 2-2 and 2-3). In order to study the structural basis of specificity and to gain a better understanding of the catalytic mechanism of the Group IV decarboxylases, I solved the X-ray structures of *cv*ADC in both free and product-bound states at 1.95 and 1.8 Å resolution, respectively. The overall monomeric structure of *cv*ADC is similar to *T. brucei* and human ODC (RMSD of 1.2 and 1.4 Å, respectively). However, the origin of the arginine specificity of *cv*ADC is apparent. It arises from different positioning of a  $3_{10}$ -helix (*cv*ADC residues S291-E296) in the C-terminal domain that includes the substrate-binding residue E296 (D332 in *Tb*ODC). Furthermore, 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 (26), it appears that the loop serves the function of a mobile lid

allowing substrate access, and sequestering substrates from solvent.

## Materials

Methyl-pentanediol (MPD) was purchased from Hampton Research, and all other chemicals were purchased from Sigma. Crystallization screens were purchased from Emerald Biostructures.

## Experimental Methods

*cvADC Expression and Purification* The expression and purification of His<sub>6</sub>-tagged *cvADC* from pODCTM9 and pODCTM9 T142A were described in Chapter II. Protein was produced heterologously in *E. coli* (BL21 DE3 strain) under an IPTG-inducible promoter. Once the *E. coli* cells had reached an O.D.<sub>600</sub> of 0.5-0.7, IPTG (200  $\mu$ M) was added to each liter of media. Cells were harvested after 12 hours of induction by centrifugation at 1500 rpm, 4° C for 30 minutes. The cells were then lysed, sonicated and centrifuged at 100,000 x g for 1 hr. The supernatant was loaded onto a column containing Ni<sup>2+</sup>-agarose resin. After washing protein was eluted with the addition of imidazole. The T142A mutation arose during prior cloning of the gene, however the both wild-type and *cvADC* T142A display equivalent activity, and they have both been used in these studies. Protein derived from pODCTM9 T142A was used for the crystallization of Seleno-methionine labeled *cvADC* protein. Site-directed mutagenesis (Quickchange<sup>TM</sup>, Stratagene, La Jolla, CA) was employed to regenerate the wild-type pODCTM9 plasmid. The protein was used to crystallize *cvADC* bound to agmatine.

A selenomethionine derivative of *cvADC* 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<sub>600</sub> = 0.7), IPTG (0.5 mM), 60 mg Seleno-methionine, 100 mg each of threonine, lysine and phenylalanine, and 50 mg each of leucine, isoleucine and valine were

added. After 13 hours of induction, cells were harvested by centrifugation and lysed. Purification of the Se-Met derivative was performed as described in Chapter II. Mass spectrometry analysis (Protein Chemistry Technology Center, UTSW) revealed that there were 5 Se-Met residues per monomer of *cvADC* produced from the minimal media (Figure 3-3).

*Crystallization of cvADC* Approximate crystallization conditions were found with the Wizard I screen (Emerald Biostructures, BainBridge Island, WA) (Figure 3-4A). Free (Se - labeled) and agmatine-bound *cvADC* crystallized in similar conditions. The optimized crystallization conditions are as follows: 5  $\mu$ l 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 (Figure 3-4B). Agmatine-bound *cvADC* crystals were produced by incubating 20 mg/ml protein in buffer with 5 mM agmatine at 4° C for one hour prior to setup of the crystallization drop. Crystals generally grew within 24 hours. 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.

After the initial crystallization conditions were found, I attempted to re-crystallize *cvADC* in another condition. I undertook this task because the diffraction data for the original crystals revealed that the unit cell parameters ( $a = 116.1 \text{ \AA}$ ,  $b = 116.9 \text{ \AA}$ ,  $c = 269.4 \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ ) were large and that the space group was  $P2_12_12_1$ . Assuming a crystal solvent content of 50%, the Matthews' coefficient ( $V_M$ ) serves as an empirically-derived guide to the probability that the asymmetric unit contains X number of monomers (where  $V_M = \text{volume of unit cell} / (\text{MW of protein} * \text{number of asymmetric units} * X)$ ) (60). The orthorhombic

crystal I produced is predicted to contain eight molecules of the *cvADC* monomer in each asymmetric unit. Another crystal form with smaller cell dimensions and/or higher symmetry might allow for facile structure determination. I re-conducted crystallization trials with reagents from the Wizard screens (series I and II) and the Hampton Crystal Screens (series I and II). Despite this re-crystallization attempt, no new conditions were found. Curiously, I found another crystal morphology by simply increasing the pH of the original crystallization condition from pH 7.5 to pH 8.3 (Figure 3-4C). This new crystal's unit cell parameters ( $a=58\text{ \AA}$ ,  $b=153\text{ \AA}$ ,  $c=174\text{ \AA}$ ,  $\alpha=\gamma=90^\circ$ ,  $\beta=99^\circ$ ) and space group (P2) suggests that there are 12 molecules in the asymmetric unit. In addition to containing 4 more molecules per asymmetric unit than the original crystal form I found, the new monoclinic crystals diffracted x-rays no farther than  $3.0\text{ \AA}$ , displayed higher mosaicity and proved to be too fragile for cryo-protection techniques. It became clear that the orthorhombic crystals I had originally discovered were the superior choice for structure determination of *cvADC*.

#### *Data Collection and Processing*

I collected a diffraction dataset for a ligand-free, orthorhombic crystal of *cvADC* at a home-source x-ray beam ( $\lambda = 1.5418\text{ \AA}$ ) and R-Axis IV detector (Rigaku, The Woodlands, TX). I indexed, integrated and scaled the data with HKL2000 (61). Data processing statistics are summarized in Table II.

Diffraction data with anomalous dispersion for a Se-Met-labeled *cvADC* crystal 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). This data was kindly collected by Mischa Machius (UTSW Structural Biology Laboratory). I collected data for an agmatine-bound *cvADC* crystals at the 19-ID beamline at the same facility. The data from the Se-Met-labeled and agmatine-bound crystals were indexed, integrated and scaled

using HKL2000 (61). Data processing statistics for both the Se-Met labeled and agmatine-bound datasets are summarized in Table III.

#### *Structure Determination and Refinement*

Both the ligand-free (Se-Met labeled) and agmatine-bound *cvADC* structures crystallized in the  $P2_12_12_1$  space group and similar unit cell parameters were observed for the free ( $a = 116.1 \text{ \AA}$ ,  $b = 116.9 \text{ \AA}$ ,  $c = 269.4 \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ ) and agmatine-bound ( $a = 116.2 \text{ \AA}$ ,  $b = 117.3 \text{ \AA}$ ,  $c = 268.7 \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ ) structures. I attempted to solve the structure of ligand-free *cvADC* using the molecular replacement method. For this method I believed that the structure of *T. brucei* ODC would be an appropriate phasing model. I used a number of programs to try to find an initial model of *cvADC*. Despite using programs such as CNS (62), *MolRep* (63), *AMoRe* (64), *EPMR* (65) and *Phaser* (66), I was unable to generate a solution for the *cvADC* structure using the *T. brucei* ODC model and the data I collected at the home-source detector. The failure to find a solution is most likely due to the large unit cell of the crystal. The Matthew's coefficient for the orthorhombic crystal indicates that each asymmetric unit should contain 8 molecules of the *cvADC* monomer. Thus, the numerous molecular replacement trials I attempted were not successful because the programs were not capable of determining the proper rotation and translation function to account for the large number of scattering factors in the asymmetric unit, absent any constraints such as non-crystallographic symmetry operators. As a result of these unproductive trials, I produced, purified and crystallized a Seleno-methionine derivative of *cvADC* in order obtain experimental phases.

The data set collected at the Se absorption edge was used for phasing by the Single-wavelength Anomalous Dispersion method (SAD) (67). The Selenium substructure was determined by direct methods in *ShelX* (68). The top solution included all 40 Se sites expected in the asymmetric unit (5 Se-Met residues/monomer X 8 monomers/ a.s.u.) with correlation coefficients higher than

0.48. Based on the Se substructure Non-Crystallographic Symmetry (NCS) operators were obtained with the program *RESOLVE* (69). Phases were calculated in *MLPHARE* (70) and an overall figure-of-merit of 0.29. An inherent problem with the SAD phasing method is the ambiguity between the two hands of a phase solution. Density modification has been used to break this ambiguity. For my structure I used the program *DM* (71) to determine the hand of the correct solution. With the NCS operators and phases, a density-modified map was calculated. This map was inputted into the automated model-building program *ARP/wARP* (72) and an initial model, containing 2300 of 2900 expected residues in the asymmetric unit, was built. The agmatine-bound *cvADC* structure was determined by using the free structure as the initial phasing model for molecular replacement with the CNS package (62). Both structures were rebuilt with the graphics program *Coot* (73) and the refinement package *REFMAC5* (74). For the agmatine-bound structure, the Fourier difference map ( $F_o - F_c$ ) showed readily interpretable density for the agmatine molecule bound to the PLP cofactor. The agmatine molecule 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  $3.0 \sigma$  for both structures.

The free structure of *cvADC* was refined with data to  $1.95 \text{ \AA}$  and the agmatine-bound structure included data to  $1.8 \text{ \AA}$ . The final models contain eight monomers in the asymmetric unit (labeled A – H), in the form of two tetramers (Figure 3-5), 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 *cvADC* 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 *cvADC*-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 (75). The four dimers of the agmatine-bound *cvADC* 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 (DeLano, W.L. The PyMOL Molecular Graphics System (2002) on World Wide Web <http://www.pymol.org>). All RMSD calculations were based on structural alignment by PyMol and calculated within the program. Solvent accessible surface was calculated using AREAIMOL (74).

## **Results**

*X-ray structure determination of free and agmatine-bound cvADC* The overall structures of *cvADC* in both free and agmatine-bound states are similar to the previously determined eukaryotic ODC structures. The *cvADC* structure consists of an N-terminal  $\beta/\alpha$ -barrel from residues 23-261 and a C-terminal  $\beta$ -barrel from residues 262-372 (Figure 3-6). In addition, a short  $\alpha$ -helix (residues 2-11) and  $\beta$ -strand (residues 18-22) form at the N-terminus of both the native and agmatine-bound structures. The N-terminal  $\beta$ -strand is part of an extended  $\beta$ -sheet forming interactions with two strands from the C-terminal domain. The N-terminal region preceding the  $\beta/\alpha$ -barrel has been observed in a number of other ODC structures and forms approximately the same secondary elements observed in the *cvADC* structure (26, 27, 30, 32).



As with all other Group IV decarboxylases, two identical active sites are formed at the dimer interface (Figure 3-7). Residues from the  $\beta/\alpha$  barrel and  $\beta$ -barrel of one monomer as well as the  $\beta$ -barrel from the opposite monomer contribute 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 3-8). Two residues in the loop, *cvADC* 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. In the native structure the PLP cofactor forms an interaction with Lys-48 *cvADC* (Lys-69 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 *tbODC* (3). However, unlike previous ligand-free ODC structures (26-28), PLP is not bound to K48 *cvADC* through a Schiff's base. Instead the density around the C4' of the cofactor suggests a tetrahedral configuration consistent with a carbinolamine species (Figure 3-9). The model for the ligand-free enzyme was built with a carbinolamine intermediate in all eight active sites of the asymmetric unit and no residual density was found in the re-calculated difference Fourier map. Density for the sulfhydryl group of the active site cysteine residue Cys-324 *cvADC* (C360 in eukaryotic ODC), which had been proposed to function as a general base in the reaction (32), 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, I modeled C324 *cvADC* in a dual conformation with the sulfhydryls positioned 145° apart (Figure 3-9). The homologous residue in ODCs has been observed to occupy both of these conformations (26-28, 30-32).

In the active site of the ligand-bound structure, readily interpretable density was observed for agmatine in all eight monomers in the asymmetric unit

(Figure 3-10). The side-chain of residue Lys-48 *cvADC* has shifted 5 Å away from the cofactor to form an interaction with Asp-67 *cvADC*, thus making room for the bound product. 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 Tyr-353, Phe-361 and Val-143, while the guanidinium moiety is within H-bonding distance of the side-chains of Glu-296, Asn-292, Asp-325, and Tyr-287.

*The Lys-148 cvADC (Lys-169 ODC) loop functions as an active site lid.* To investigate the structural differences between free *cvADC* and eukaryotic ODC, the Cα atoms of the N-terminal β/α-barrel domain of *cvADC* (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 at between 1.1 (*tbODC*) – 1.2 (hODC) Å (Figure 3-11 A and B). The RMSD values for dimer-to-dimer comparison of *cvADC* were 1.2 and 1.5 Å for *tbODC* and hODC, respectively. However, regions of the structure displayed RMSD values above 2.5 Å. Notably significant changes are observed in the Lys-148 *cvADC* (Lys-169 ODC) loop which is formed by residues 135 – 152 *cvADC* at the end of the β<sub>6</sub> strand. With the exception of the human ODC structure (26), portions of this loop (residues Asp-158 – Arg-165 ODC) have typically been disordered in the eukaryotic ODC structures that have previously been solved (e.g. the mouse ODC and the *T. brucei* ODC structures (27, 28, 30-32)). Density for the entire loop was observed in the human ODC structure where it assumes an open conformation allowing the active site to be solvent-exposed. In contrast, in both the native and agmatine-bound structures of *cvADC* a seven residue segment (Asn-140 to Gly-146) of the loop has swung towards the active site, pivoting around residues Pro-139 and Asn-147 (Figure 3-12). This is the first observation of this loop in a closed conformation with respect to the active sites. Some of the residues in the mobile region of the loop have moved over 10 Å

towards the  $\beta/\alpha$ -barrel. In this conformation, the bound ligands are sequestered from solvent. In the agmatine-bound *cvADC* structure there is no solvent-accessible surface on agmatine, while 2.0 Å<sup>2</sup> is accessible on the adjacent PLP molecule. In contrast, the putrescine molecule in the ligand-bound *tbODC* structure has 21 Å<sup>2</sup> of solvent exposed surface compared to 1.3 Å<sup>2</sup> 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 *cvADC*. 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 C $\alpha$  atoms of residues Ser179-Asn183 *cvADC*, which cover the phosphate of PLP, have shifted 3 Å towards the phosphate moiety in *cvADC*. This movement allows the formation of a new H-bond interaction between the backbone of Ser-179 and Val-143 in *cvADC* (Figure 3-13). Additional backbone H-bond interactions between residues Gly-178 – Gly-180 and Thr-142 – Val-143 *cvADC* seem to promote the shift of the Ser-179-Asn-183 *cvADC* region and add to the stabilization of the Lys-148 *cvADC* loop.

*The 3<sub>10</sub> –helix is a key determinant of substrate specificity.* In order to delineate the determinants of substrate specificity in *cvADC*, agmatine-bound *cvADC* and putrescine-bound *tbODC* (1F3T) were compared (Figure 3-14). The subunits aligned by their N-terminal  $\beta/\alpha$  barrel domains overlayed 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 Asp-67, Lys-48, and Glu-252 – Arg-255 in *cvADC*. The  $\beta$ -barrel C-terminal domain of the opposite subunit are also closely aligned, including the catalytic base and substrate recognition residues, Cys-324' and Asp-325' *cvADC*. Part of the C-terminal  $\beta$ -barrels of the primary subunits of *cvADC* and *tbODC* also superimpose well, including the

substrate binding residues Tyr-287, Tyr-353 and Phe-360 of *cvADC*. A plot of RMSD values of 368 aligned C $\alpha$  atoms reveals that only about 30 residues deviate by more than 2.5 Å between *cvADC* and ODC (Figure 3-15). This result stands in stark contrast to the comparison of ODC and DAPDC in which nearly 100 positions were found to deviate by more than 2.5 Å between the structures (Figure 3-2). However, significant differences in the active site are apparent, providing new insight into the structural basis of specificity in this family of enzymes. In both the *cvADC* and ODC structures residues Ser-291-Glu-296 *cvADC* (Ser-325 – Asp-332 ODC) form a short  $3_{10}$ -helix at the distal portion of the substrate-binding pocket. The helix in the *tbODC* structure begins 2 residues prior to that observed in *cvADC*, however, the helix ends at the same position (Glu-296/Asp-332) in both (Figure 3-14). This residue, Glu-296 *cvADC* (Asp-332 ODC) is positioned from the helix to form hydrogen bond contacts with the guanidinium of agmatine in *cvADC* (the  $\delta$ -amino of putrescine in ODC). In the *cvADC* structure the helix is translated 2.1 Å away from the cofactor when compared to *tbODC* (as measured between C4' of PLP and the C $\alpha$  of Glu-296/Asp-332), effectively enlarging the pocket to accommodate the arginine substrate. The shortening of the helix may contribute to this shift. However despite this translation the longer aliphatic portion of Glu-296 *cvADC* 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 Asp-332 ODC. Thus Glu-296 *cvADC* is able to retain key contacts with the remainder of the active site, including an H-bond interaction with Arg255 *cvADC* (Arg-277 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 (4).

As reported in Chapter II, a primary sequence alignment suggested that

only one substitution in the active site of *cvADC* (E296D) had occurred relative to eukaryotic ODC. The structural alignment of *cvADC* and *tbODC* identifies several other amino acid substitutions in the distal region of the active site that may contribute to the specificity change (Table IV). The more noteworthy substitutions in the region include Cys-328 ODC (Asn-292 *cvADC*) and Gly-393 *tbODC* (Leu-357 *cvADC*). The side-chain of Cys-328 ODC is turned out of the active site and forms an H-bond interaction with the backbone carbonyl of Ser-325. In contrast, the side-chain of Asn-292 *cvADC* is pointed towards the active site and the carbamoyl oxygen forms a H-bond with the guanidinium moiety of agmatine (3.3 Å). Leu357' *cvADC* (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 eukaryotic ODC (Gly or Ala at this position), may play a role in positioning of the  $3_{10}$ -helix. The side-chain of Leu-357' is within van der Waals contact of the side-chain of Ser-291 *cvADC* of the  $3_{10}$ -helix (Figure 3-10). 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 *cvADC* was compared to the three available structures of DAPDC. The N-terminal  $\beta/\alpha$  barrel domain of *cvADC* with *M. 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 – 2.5 Å (Figure 3-16). The significantly greater divergence in the structures is consistent with the fact that DAPDC and *cvADC* share about 20% sequence identity. The differences between *cvADC* and the DAPDC structures arise from differences distributed throughout both the N- and C-terminal domains. Unlike for the comparison to alanine racemase (28) no apparent domain rotation has taken place between

DAPDC and *cvADC*. Despite the higher RMSD values for the overall structures, the active site of DAPDC is similar to *cvADC*. As with the *cvADC* and *tbODC* comparison, the position of the 3<sub>10</sub>-helix is substantially different between *cvADC* and the DAPDC structures, and indeed it assumes a different conformation in all three structures. The distance between C4' of PLP and the helix (C $\alpha$  of Asp-332 *tbODC*, Glu-296 *cvADC* and Glu-348 *mjDAPDC*) is 10.0 Å for *tbODC*, 12.1 Å for *cvADC*, 15.3 Å for *mjDAPDC*.

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 *cvADC* include one common interaction Asp-332 *tbODC* (Glu-296 *cvADC*), while agmatine in *cvADC* makes an additional interaction with Asn-292 *cvADC* positioned at the N-terminal of the 3<sub>10</sub>-helix. For DAPDC the residue at the start of the helix has been substituted with Arg-343 *mjDAPDC*, which forms a salt bridge with the carboxylate of the bound lysine (Figure 3-17). Additionally the hydroxyl group of Tyr-347 *mjDAPDC* forms H-bond interactions with the amine group of lysine, while the residue equivalent to Asp-332 *tbODC* faces out of the pocket. These data support the conclusion that the 3<sub>10</sub>-helix is a major determinant defining substrate specificity of Group IV decarboxylases. Within the context of this conserved fold, the protein is designed to be flexible in the positioning and amino acid sequence of the 3<sub>10</sub>-helix, providing a mechanism to evolve different substrate preferences without large structural changes.

*Structural Comparison of cvADC and AR* The structures of *Bacillus stearothermophilus* alanine racemase has been determined in various ligand-bound states (76-80) and from the human pathogens *Pseudomonas aeruginosa* (81) and from *Mycobacterium tuberculosis* (82). Like all other Group IV enzymes, these enzymes fold into an N-terminal  $\alpha/\beta$  barrel and C-terminal  $\beta$ -barrel. However, the sequence identity between *cvADC* and AR from these

species is very low (~ 10%). Accordingly, the RMSD for C $\alpha$  atoms is on the order of 4 Å (e.g. *cvADC* and *M. tuberculosis* AR). As in the comparison of *T. brucei* ODC and *B. stearotherophilus* AR (28), a rotation of the C-terminal  $\beta$ -barrel in AR has occurred relative to *cvADC* (Figure 3-18 A and B). Furthermore, the topology of the C-terminal  $\beta$ -barrel of AR diverges from that of *cvADC* (Figure 3-18 C and D). Despite the differences between the C-terminal domains of these structures, the conserved catalytic Cys-324 (Cys-317 in *mtAR*) is positioned roughly similarly in the active sites of both structures. Because AR is active on a smaller substrate, L-alanine versus L-arginine, the substrate binding pocket of AR is compressed, in part by Met-312 (ref), relative to the *cvADC* active site.

*Structural Comparison of cvADC and Pyruvoyl-dependent ADC* To date there are three structurally-defined ADC classes (83). As mentioned before, Group III PLP-dependent decarboxylases, which are structurally related to the AATase family, include enzymes with activity for L-arginine. *cvADC* is the first Group IV ADC to be structurally characterized. Finally, an unrelated ADC from the archeon *M. jannaschii* has been cloned and characterized (84). The *M. jannaschii* ADC uses a pyruvoyl group as a cofactor and shares no sequence or structural similarity with Group III or Group IV ADCs. Furthermore, the oligomeric state differs between each of the three classes of ADC. Group III ADCs are thought to form dodecamers similar to *Lactobacillus* 30a ODC (25). The active form Group IV ADC is a homodimer (Figure 3-7). The oligomeric state of *M. jannaschii* ADC is a homotrimer (Figure 3-19) (85). The protomer fold is a four-layer  $\alpha\beta\beta\alpha$  sandwich with topology similar to pyruvoyl-dependent histidine decarboxylase. Despite the lack of structural similarity, there are a number of active site features common amongst these ADCs. In terms of the chemistry of decarboxylation, both enzymes proceed through a Schiff base between either a PLP or a pyruvoyl group and the

amino group of the substrate (86). In both Group IV and *M. jannaschii* ADC, the predicted placement of the CO<sub>2</sub> group of L-arginine is in a pocket surrounded by the carboxylate group of an acidic residue (Asp-361 for Group IV ADC or Glu-109 for *M. jannaschii* ADC) and a variety of hydrophobic side chains (Phe-397, Tyr-389 for Group IV ADC or Leu-31 and Phe-34 for *M. jannaschii* ADC) (29, 85). Interestingly, the torsion angles of the agmatine molecules in *cvADC* and *M. jannaschii* ADC are different. In *cvADC* the agmatine molecule is in an extended conformation (Figure 3-6). On the other hand, the agmatine molecule in the structure of *M. jannaschii* ADC is extended except for the C $\gamma$ -C $\delta$  torsion angle, which is in a gauche conformation (Figure 3-19).

## Discussion

The Group IV PLP-dependent decarboxylases are composed of a large, diverse family with a broad range of substrate preferences for basic amino acids (23, 59). The discovery of *cvADC*, 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 (87). The structure of *cvADC* identifies the E296 *cvADC* (D332 ODC) 3<sub>10</sub>-helix as the key determinant of substrate specificity. Comparison of *cvADC* with both eukaryotic ODC and bacterial DAPDC structures shows that the 3<sub>10</sub>-helix assumes a different orientation in each structure (Figures 3-14, -17, -20). 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 (28, 88, 89). However given the divergence of these two structures the importance of the position of the helix was unclear. In contrast the *cvADC* and ODC structures are very similar. Most of the active site, including the catalytic residues, and those that interact with PLP



superimpose 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 IV and Figure 3-14, -17). In *cvADC* Glu-296 has replaced Asp-332 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 Glu-296 in the same position as Asp-332 ODC. This allows contact with the guanidinium portion of the arginine substrate, while importantly preserving a salt bridge with Arg-255 *cvADC* (Arg-277 ODC). This interaction is likely to be important for orienting this residue, which was previously shown in *tbODC* to be essential for high affinity PLP binding through interaction with the phosphate (4). The structure also explains why the D332E *tbODC* and E296D *cvADC* mutants were inactive (87). 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, Cys-328 *tbODC* is replaced by Asn-292 *cvADC*. The side chain of this residue is reoriented towards 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 *cvADC* 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. Secondly, substitutions of second shell residues that interact with the helix may also contribute to the different conformations. Leu357' *cvADC* packs against the side chain of Ser-291 *cvADC* in the  $3_{10}$ -helix. This contrasts to ODC where either Ala or Gly-393 are found at this position, and form an interaction with the side-chain of Asn-327 ODC in the

3<sub>10</sub>-helix. Thus the sequence of the 3<sub>10</sub>-helix clearly coevolves multiple changes that play a 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<sub>10</sub>-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. Arg-343 *mj*DAPDC is positioned in the active site of DAPDC and forms interactions with the carboxylate of azelic acid (PDB ID 1TUF) or lysine (PDB ID 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 (88, 89), is likely to play an important role in substrate recognition. Likewise residue Tyr-347 *mj*DAPDC is positioned differently than the equivalent residue in ODC, and interacts with the  $\epsilon$ -amino group of ligand. Y331 ODC is not oriented into the active site but instead is at the two-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.

Additional examples of molecular rulers within the structures of enzymes have been previously documented. A comparison of the structures of farnesyl transferase (FTase) and geranylgeranyl transferases (GGTase) reveals a similar method of substrate discrimination as that observed within the Group IV decarboxylase family (90). FTase transfers a 15-carbon isoprenoid unit while GGTase transfers a 20-carbon unit to various protein targets (91). Both enzymes transfer an isoprenoid group to the cysteine residue of a C-A-A-X motif on the targeted protein. The basis for isoprenoid chain-length discrimination seems to be mediated by the identity of residues on helices 5 and 17 of the  $\beta$  chain (Figure 3-21 A). In the FTase context, bulky residues (Trp-102 and Tyr-365) may clash

with the longer geranylgeranyl pyrophosphate (GGPP) molecule, but not with farnesyl pyrophosphate (FPP), explaining FTase's low transfer efficiency of GGPP relative to FPP (92). In comparison the bulky residues in FTase have been substituted with a Thr and Phe residues in the GTase structure, allowing for the binding of its cognate substrate (Figure 3-21 B). The hypothesis that these residues function as the length-discriminating molecular rulers was tested by site-directed mutagenesis (1). An FTase mutated at position 102 from Trp to Thr displayed a strong preference for GGPP over FPP. Although the identities of the residues on these helices are important for chain-length discrimination, the positioning of the helices seems to be less important for these enzymes.

PagP, an *E. coli* enzyme catalyzing the transfer of a palmitate chain from a phospholipid molecule to the lipid A component of lipopolysaccharide (LPS) (93), also contains within its structure a molecular ruler. PagP is capable of discriminating acyl chains differing by only a single carbon atom (94). PagP folds into an eight-stranded antiparallel  $\beta$ -barrel (Figure 3-21 C). Previous studies have shown that the identity of position 88 at the floor of the lipid-binding pocket selects for acyl-chain length (Figure 3-21 D) (94). The Gly-88 residue in wild-type PagP allows for the binding of a 16-carbon acyl chain, while mutants containing either Ala or Met at position 88 preferentially catalyze the transfer of 15- or 12-carbon acyl chains, respectively (94). In contrast to these lipid-metabolizing enzymes, the Group IV decarboxylases use both the identity and positioning of its molecular ruler to select for substrates based on length. Thus far, this finding seems to be a unique structural characteristic of Group IV PLP-dependent enzymes.

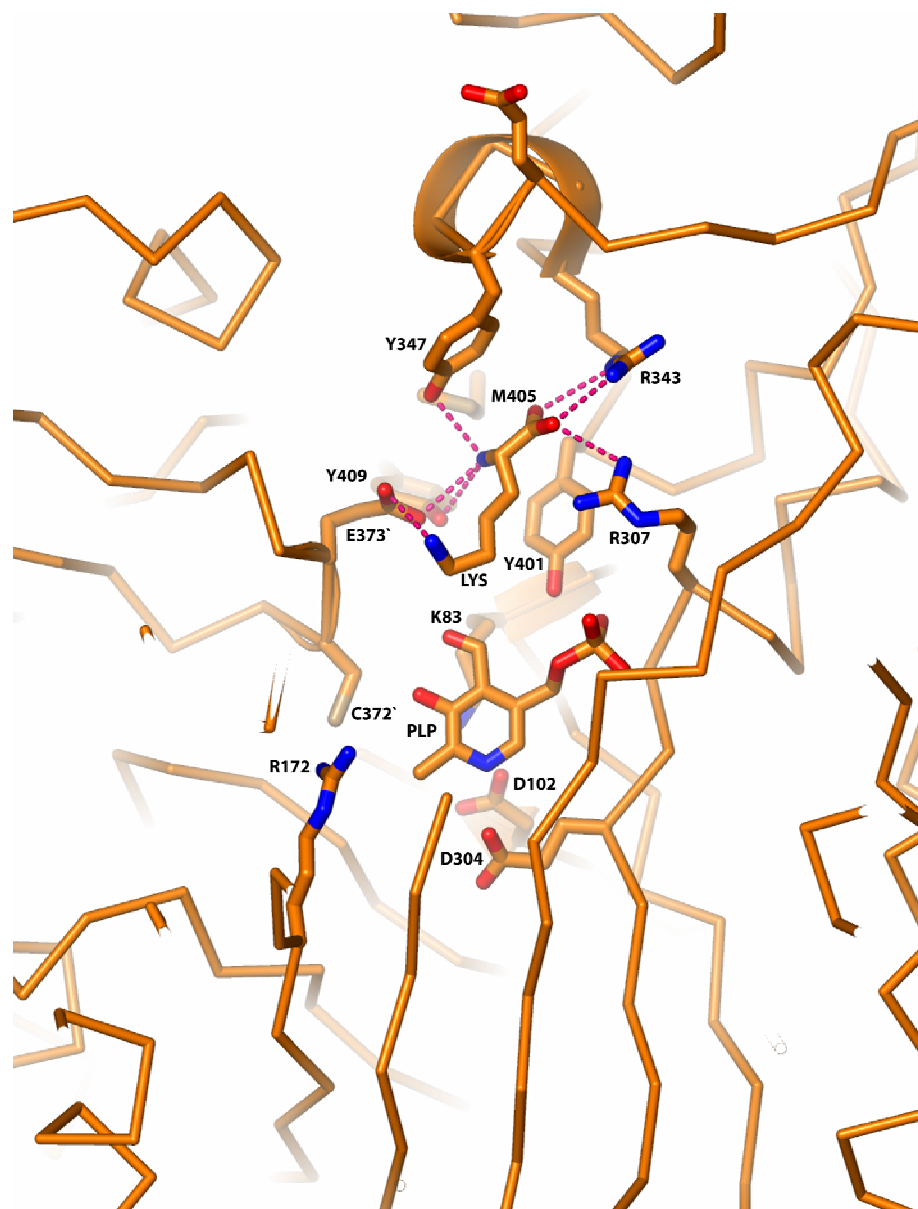
Finally, the finding that the active site is sequestered by a mobile-loop (Lys-148 *cvADC*-loop) in the *cvADC* structure has provided novel insight into the reaction mechanism of the  $\beta/\alpha$  fold-type decarboxylases. Kinetic and structural studies of mutations introduced at residues in the dimer interface of *tbODC*

previously suggested that dynamic motion of the analogous loop (Lys-169 *tb*ODC loop) played a role in the catalytic cycle (31). However, multiple conformations of the loop have never previously been observed in a structure. Indeed the Lys-169 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 3-12). In contrast for the *cv*ADC structure the Lys-148 *cv*ADC (Lys-169 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. Ser-179 *cv*ADC), and it effectively sequesters the active site from solvent. The conformation of a second active site loop formed by Ser-179-Asn-183 *cv*ADC is also repositioned in the closed conformation structure, shortening the H-bond distance between Ser-179 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 (34), 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 Lys-48 *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 (31). 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 Lys-169 *tb*ODC significantly reduces the catalytic efficiency of the enzyme (95).

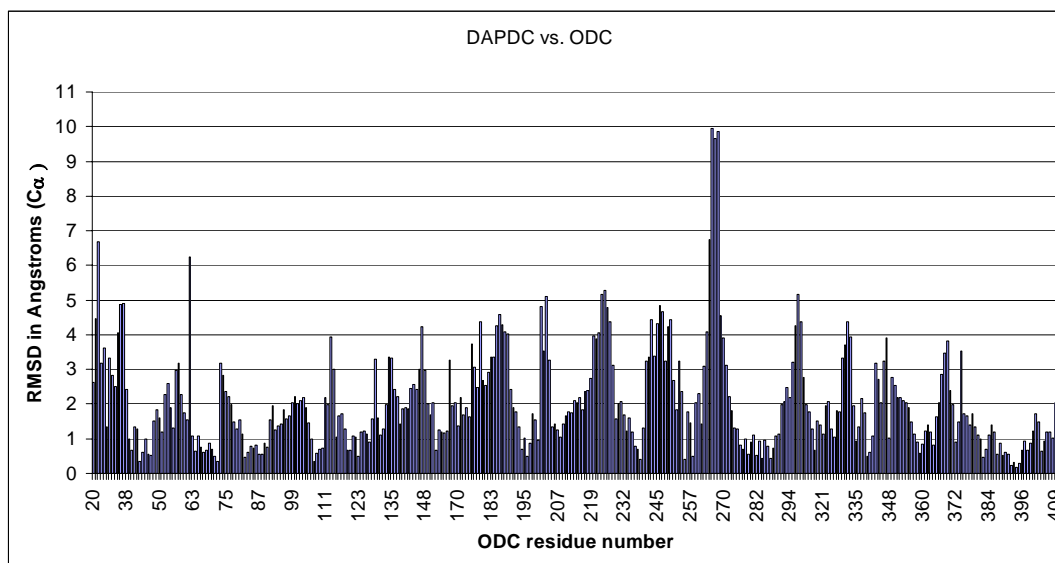
The data taken together suggest that both closed and open conformations are used in all group IV decarboxylases to accommodate sequestering of the

substrate from solvent, substrate entry, and product release. Dynamic loops have previously been reported for enzymes in this  $\alpha/\beta$  fold (e.g. triosephosphate isomerase (TIM) and tryptophan synthase (TS), inosine monophosphate dehydrogenase (IMPDH)) (96-98). It is interesting that the mobile loop in *cvADC* 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  $\beta/\alpha$  barrel fold-type enzymes.



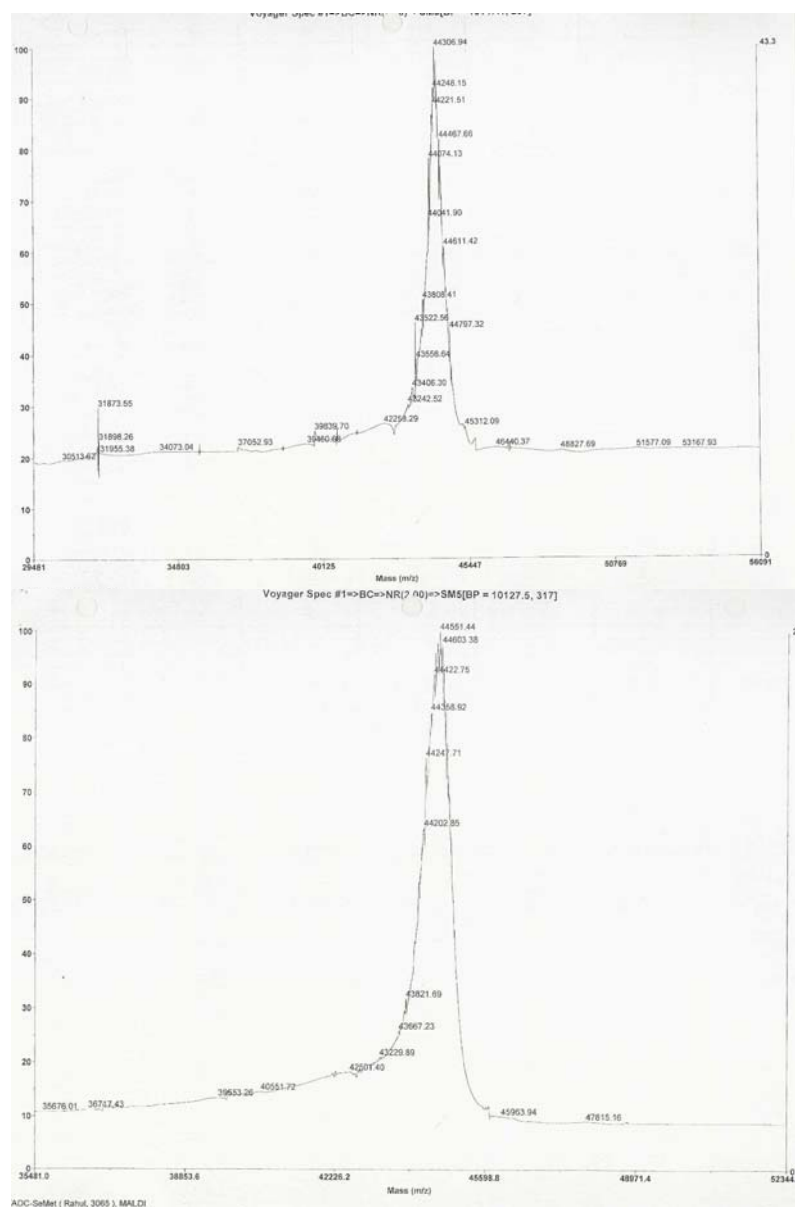
**FIGURE 3-1**

The active site of *M. jannaschii* DAPDC. The bound product (LYS) interacts with numerous residues from the N-terminal and C-terminal of the same monomer and the C-terminal of the opposite monomer (e.g. C372').



**FIGURE 3-2**

RMSD plot from the structural alignment of *M. jannashciii* DAPDC and *T. brucei* ODC. The overall RMSD between the structures is 2.5 Å (See FIGURE 1-10). However, there almost 100 residues for which the C $\alpha$  RMSD value is equal to or greater than 2.5 Å.



**FIGURE 3-3**

Mass spectrometer analysis of native (A) and Se-Met (B) *cvADC*. The mass of native *cvADC* was found to be 44306 Da and Se-Met *cvADC* has a mass of 44551. The increase in the mass of Se-Met *cvADC* (247 Da) is in accordance with the addition of 5 Se atoms, on average, in each monomer (MW = 47 Da).



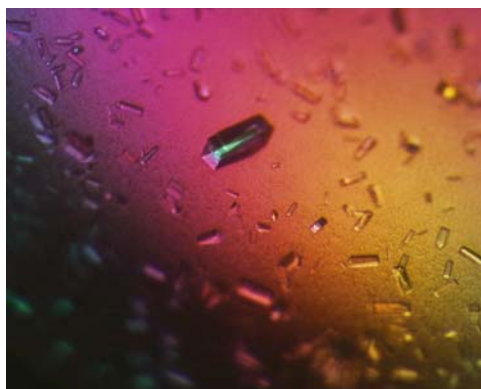
**A.**



**B.**



**C.**



**FIGURE 3-4**

Crystals of *cvADC*. The lead condition produced clustered crystals lacking well-defined faces (**A**). I refined the lead condition to produce orthorhombic crystals (**B**) that were separate and had very well-defined faces. Curiously, with an increase in the pH of the mother liquor to 8.3 a monoclinic crystal form was produced (**C**).

	Native <i>cvADC</i>
<b>Data collection and processing</b>	
Wavelength (Å)	1.514
Resolution (Å)	50 – 2.2
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions	$a = 115.8 \text{ Å}$ , $b = 117.3 \text{ Å}$ , $c = 269.0 \text{ Å}$ , $\alpha=\beta=\gamma=90^\circ$
Total number of reflections	646,432
Number of unique reflections	172,284
Redundancy (last shell)	3.9 (3.9)
Completeness (%) (last shell)	89.0 (90.9)
I/ $\sigma$ (last shell)	10.1 (2.1)
R <sub>merge</sub> (last shell)	0.1 (0.536)

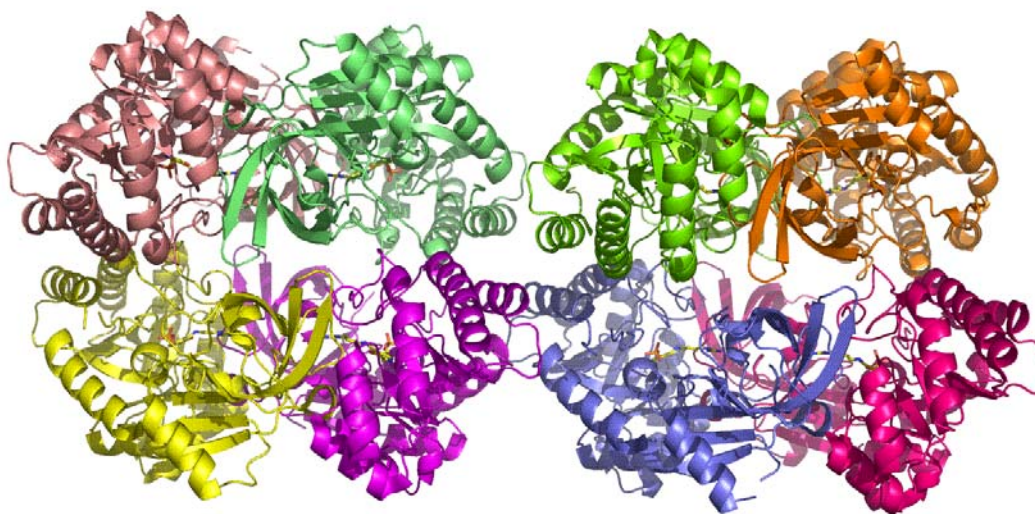
**TABLE II**

Data collection and processing from a *cvADC* crystal at the home-source R-AXIS IV x-ray detector.

	Se-Met native <i>CVADC</i>	<i>CVADC</i> -agmatine
<b>a) Data collection and processing</b>		
Wavelength (Å)	0.98066	0.97945
Resolution (Å)	50 - 1.95	50 - 1.8
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions	$a = 116.1 \text{ Å}, b = 116.9 \text{ Å}, c = 269.4 \text{ Å}, \alpha = \beta = \gamma = 90^\circ$	$a = 116.2 \text{ Å}, b = 117.3 \text{ Å}, c = 269.7 \text{ Å}, \alpha = \beta = \gamma = 90^\circ$
Total number of reflections	1,722,504	1,619,455
Number 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 <sub>merge</sub> (last shell)	0.073 (0.212)	0.118 (0.476)
<b>b) Refinement</b>		
R/R <sub>free</sub>	21.1/24.2	21.9/23.8
Number of non-hydrogen atoms	23550	23676
Number of H <sub>2</sub> O 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 <sup>2</sup> )	12.49	14.14
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

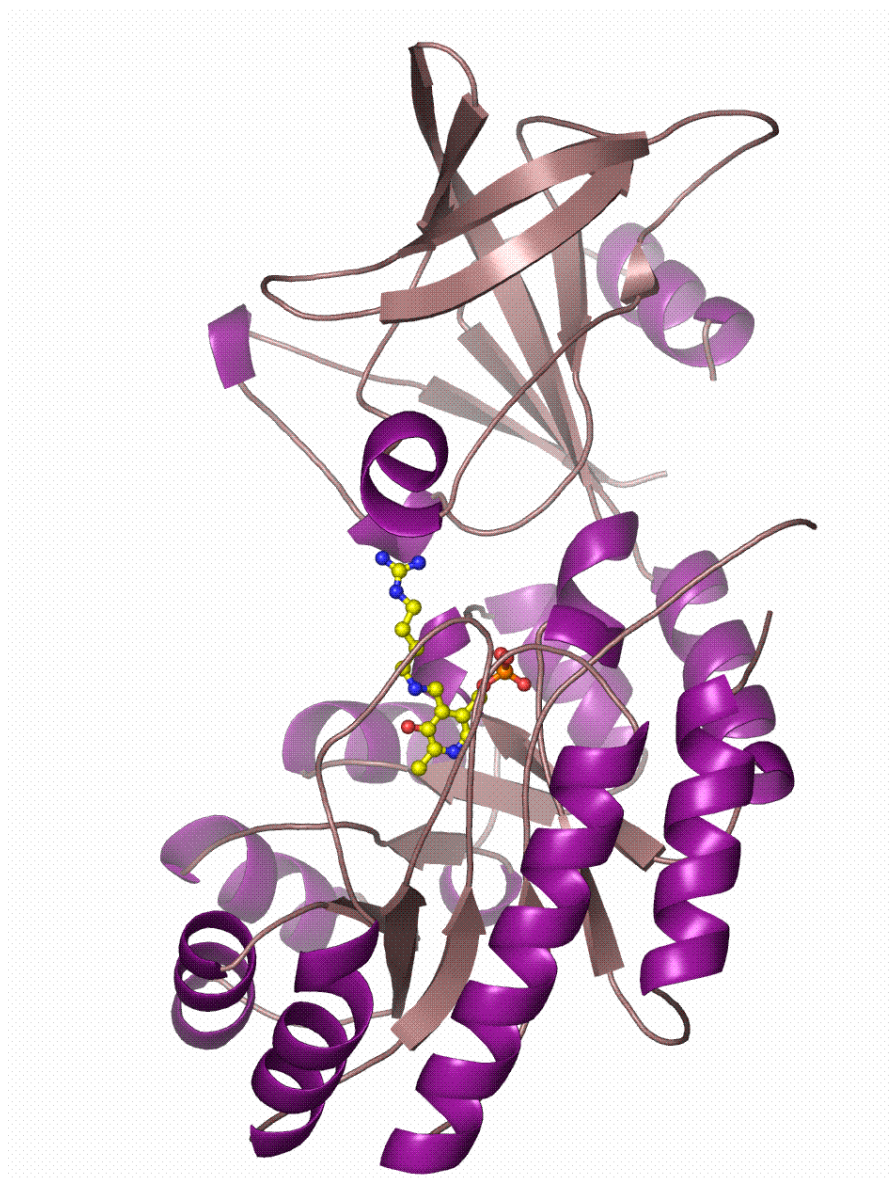
**TABLE III:**

The data collection, processing and refinement statistics for both the Se-Met labeled and agmatine-bound *cvADC* crystals from the APS beamlines BM-19 and ID-19, respectively are summarized.



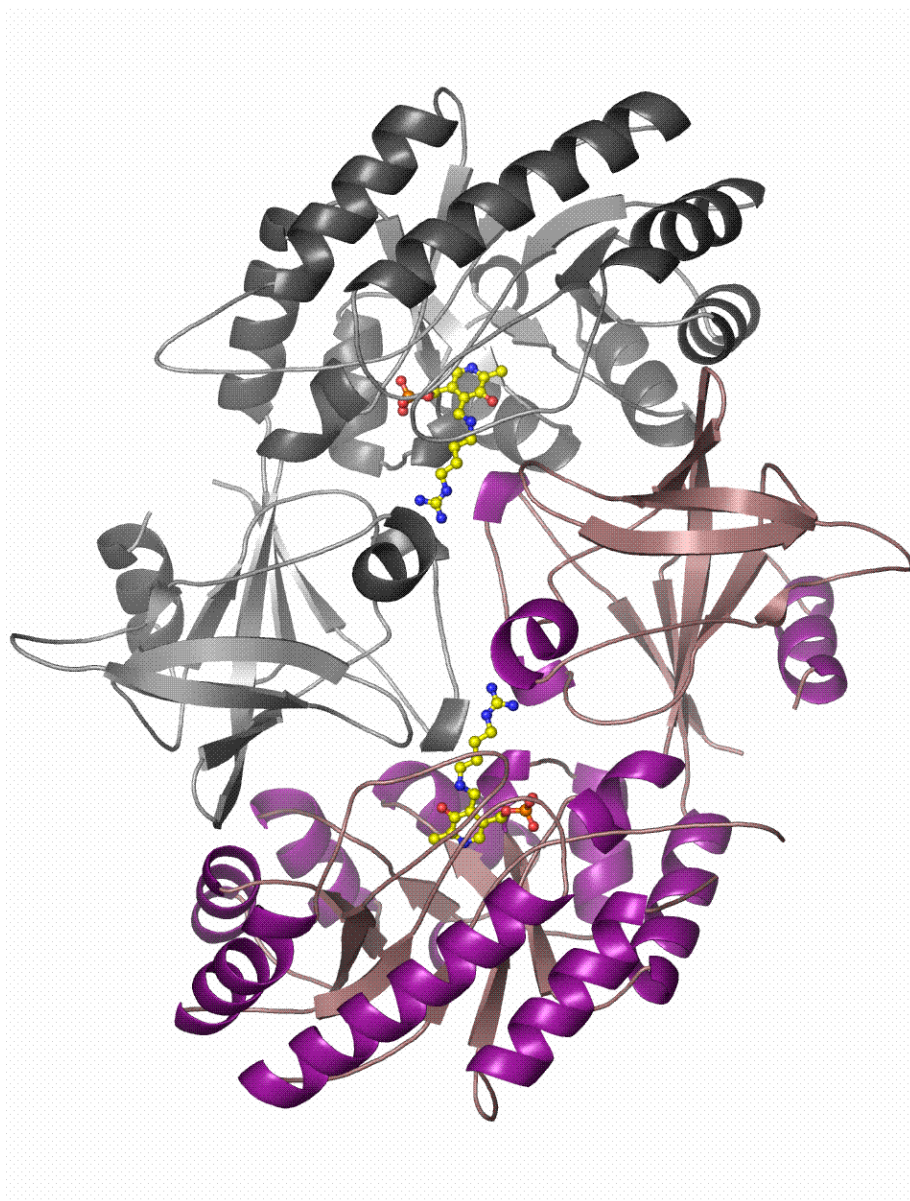
**FIGURE 3-5**

The contents of the asymmetric unit of the orthorhombic crystals include eight monomers of *cvADC* packed as a dimer of tetramers.



**FIGURE 3-6**

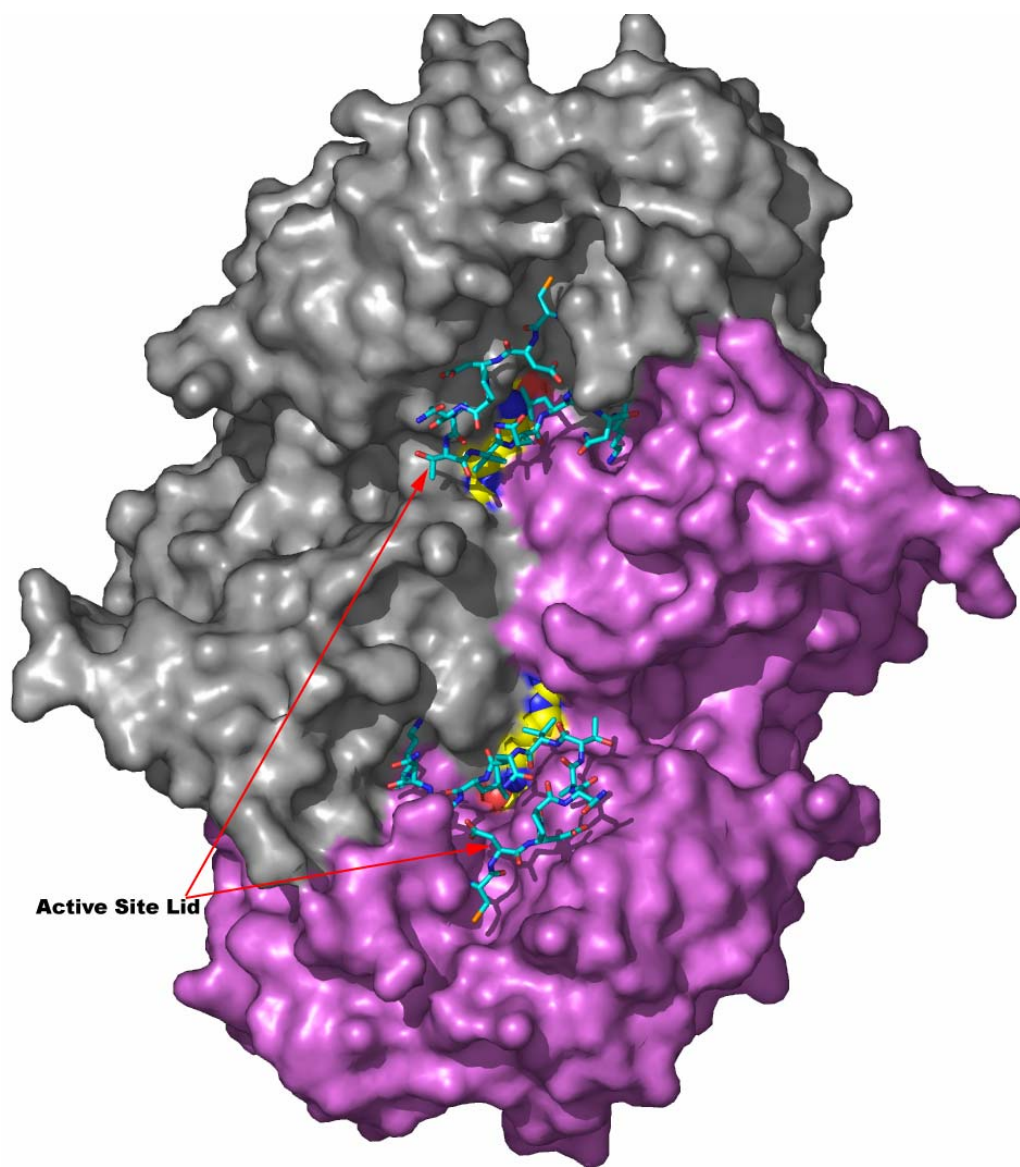
Monomeric structure of *cvADC*. The *cvADC* monomer is displayed as a ribbon with the PLP cofactor bound to agmatine (via a Schiff's base) displayed in yellow as ball and stick. Nitrogen atoms are displayed in blue, oxygen in red, phosphate in orange, and sulfur in yellow.



**FIGURE 3-7**

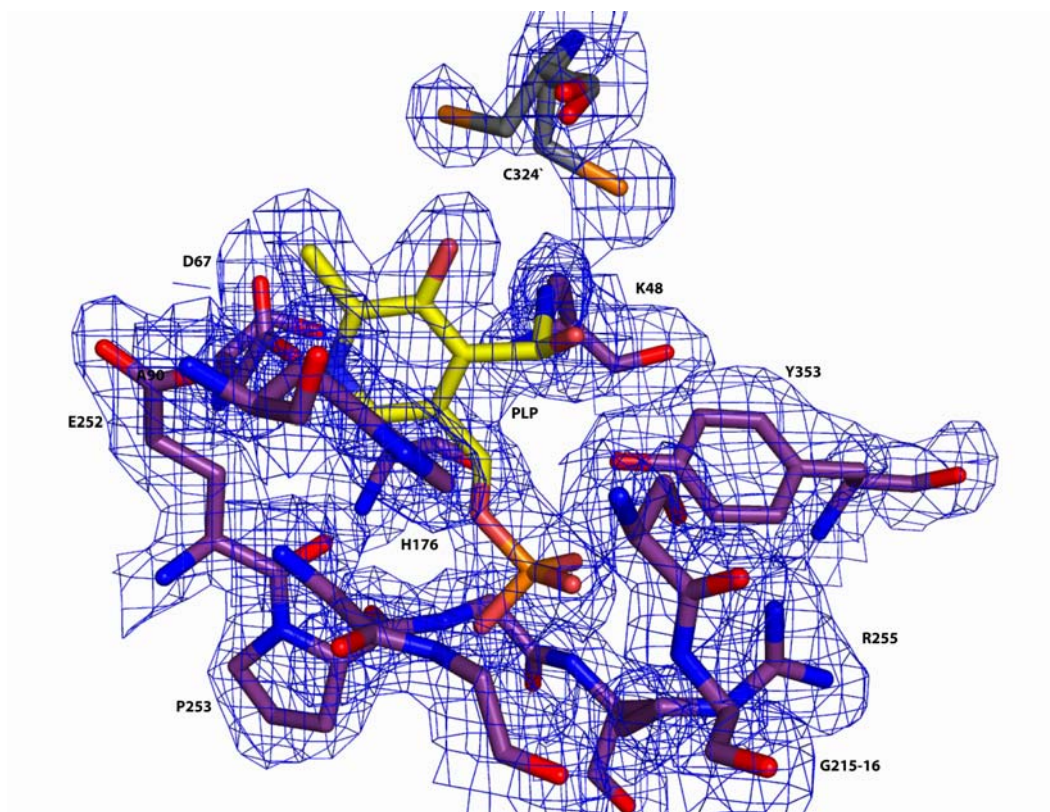
Dimeric structure of *cvADC*. The *cvADC* structure is displayed as a ribbon (monomer A (purple), and monomer B (grey)) with the bound agmatine and the PLP cofactor displayed in yellow as ball and stick. Nitrogen atoms are displayed in blue, oxygen in red, phosphate in orange, and sulfur in yellow.





**FIGURE 3-8**

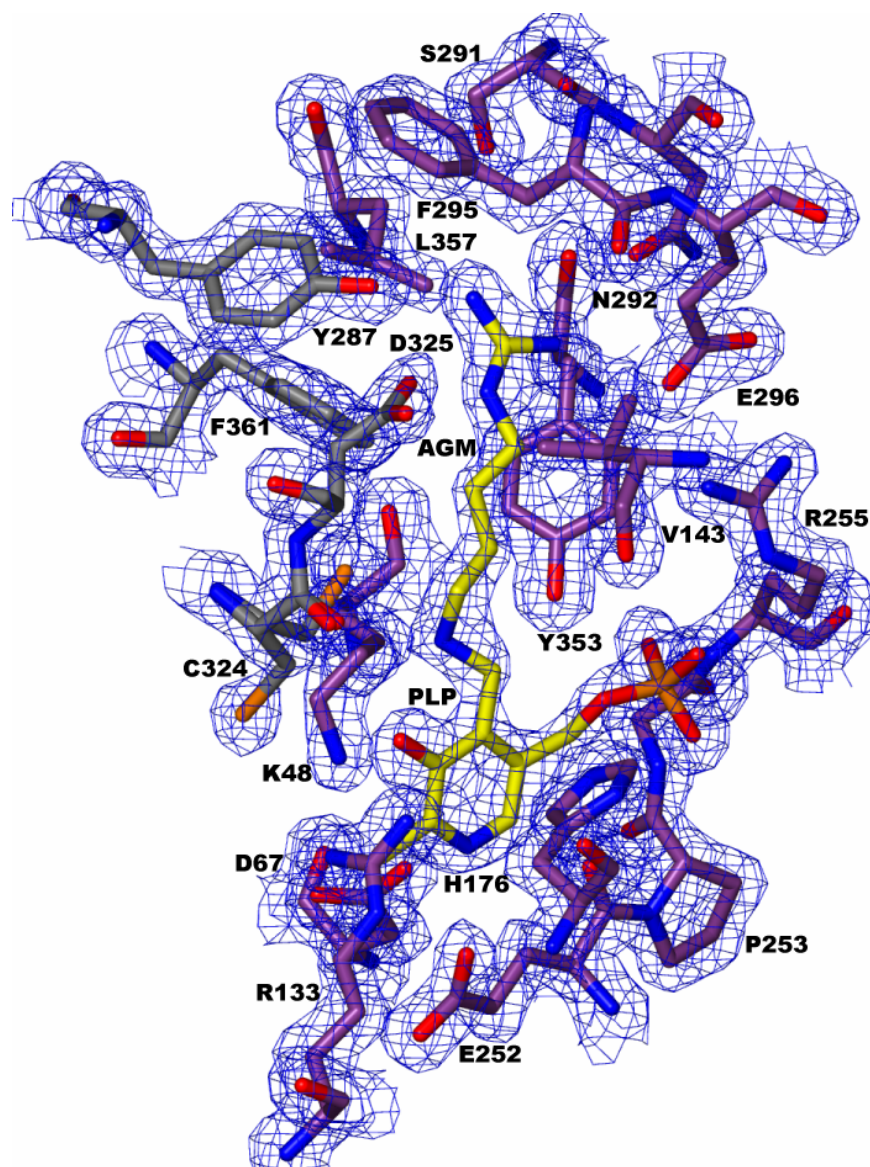
Surface representation of the *cvADC* homodimer (monomer A colored purple and monomer B in colored grey). The PLP-agmatine complex is shown in spheres and the K148 loop (RCDDPNATVQLGNK) is shown in stick representation and colored cyan.



**FIGURE 3-9**

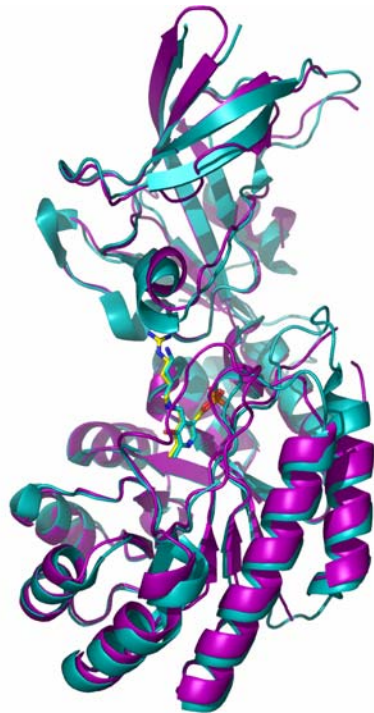
The electron density map of the carbinolamine intermediate observed in the ligand-free *cvADC* structure. The intermediate is formed between the NZ atom of K48, the C4' of PLP and a hydroxyl group. The map ( $2F_o - F_c$ ) is contoured at  $1.0\sigma$ . The PLP molecule is colored yellow, while protein residues are colored purple (chain A) or gray (chain B).





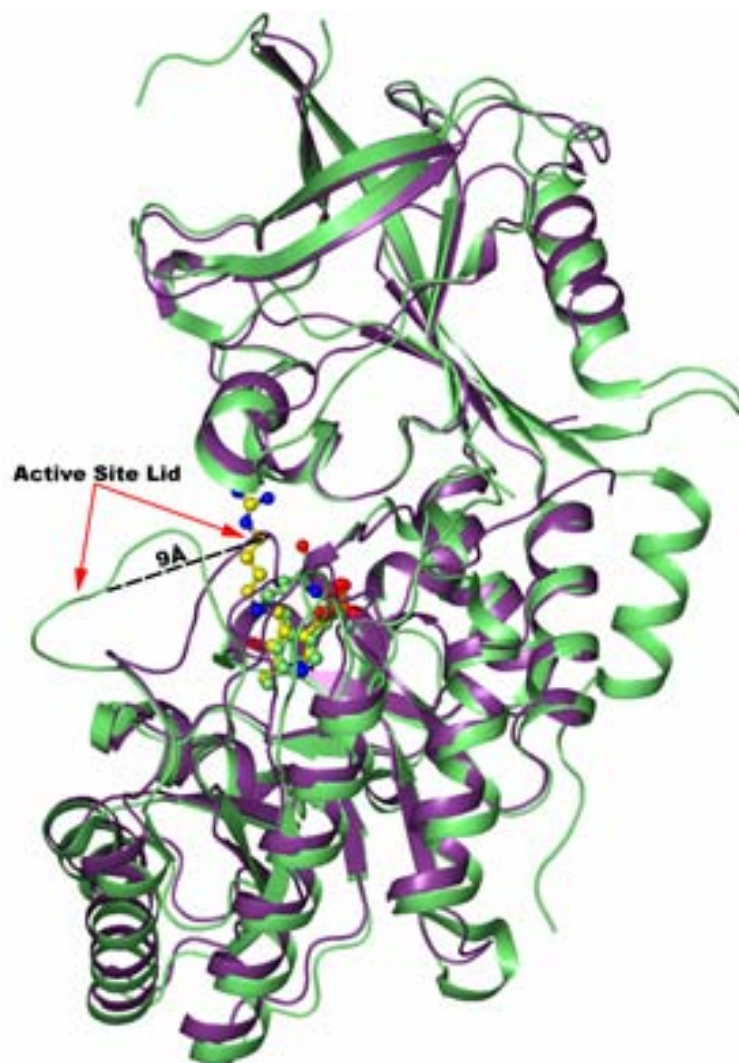
**FIGURE 3-10**

Electron density observed in the active site of the agmatine (AGM)-bound *cvADC* 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  $1.0\sigma$ . The PLP molecule and agmatine are colored yellow, while protein residues are colored purple (chain A) or gray (chain B).

**A****B**

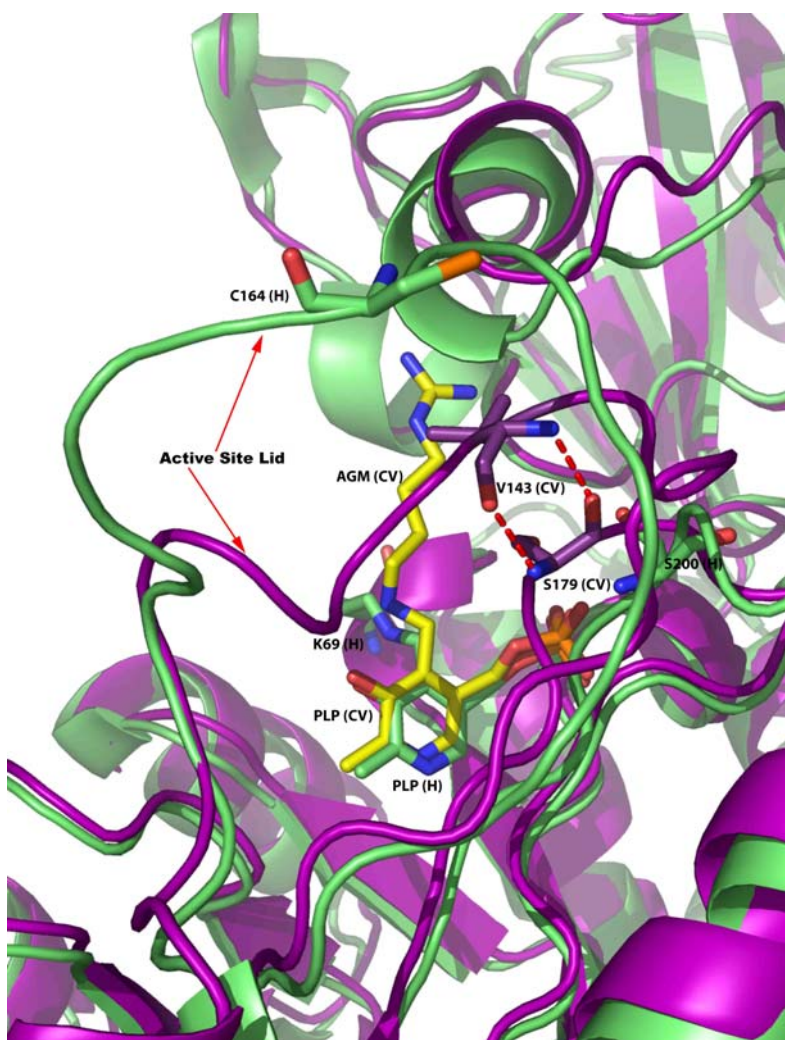
**FIGURE 3-11**

Structural alignment of *cvADC* with *T. brucei* ODC (**A**) and human ODC (**B**). The overall RMSD between *cvADC* and the eukaryotic ODCs is on the order of 1.0 Å, commensurate with the high sequence identity shared between *cvADC* and ODC.



**FIGURE 3-12**

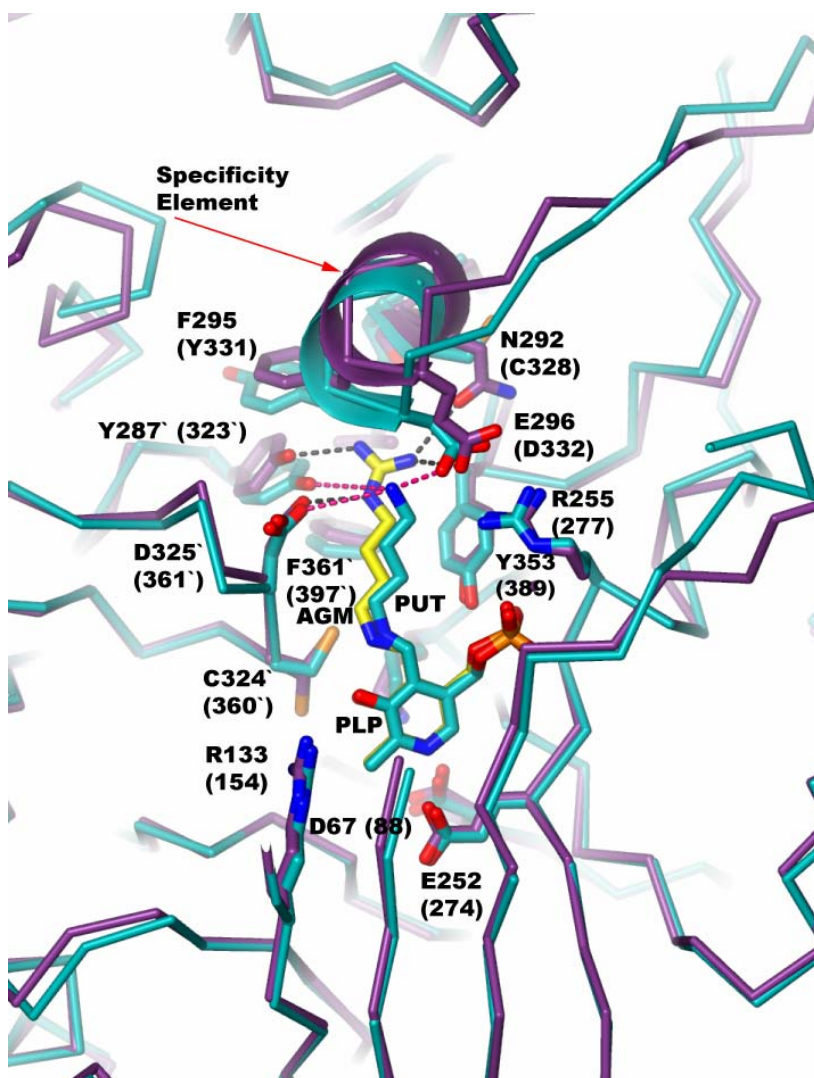
Multiple conformations of the K148 *cvADC* loop. Alignment of the structures of native *cvADC* (purple) and human ODC (1D7K ; light green). Structures were aligned using the N-terminal residues of *cvADC* (residues 23-261) and human ODC (45-283) in the  $\beta/\alpha$ -barrel domain. The overall rmsd 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 *cvADC* (K169 *hODC*) loops are P139 and N147.



**FIGURE 3-13**

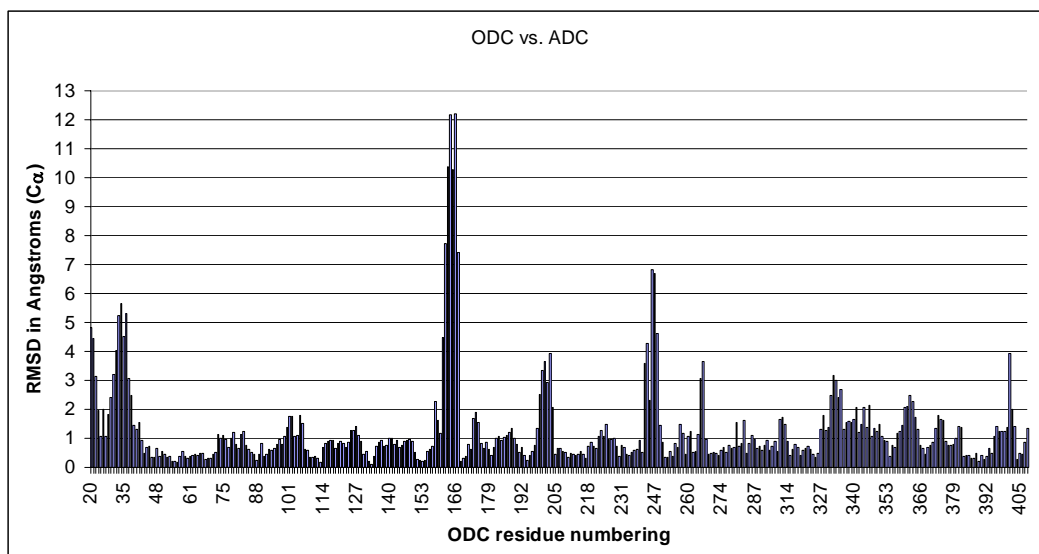
Mobile loops in Group IV decarboxylases. In the open conformation (human ODC structure in lime) the active site lid residues make no contact with a conserved Ser residue at position 200 (ODC numbering). Upon binding of ligand, the active site lid closes in toward the  $\beta/\alpha$  barrel and backbone atoms of lid residues (e.g. V143 in *cvADC*) form interactions with the backbone atoms of the conserved Ser (S179 in *cvADC*). These interactions may stabilize the loops lining the active site and promote catalysis by sequestering substrate from solvent. Labeled residues and molecules from *cvADC* are denoted by a (CV) and those of human ODC by (H).





**FIGURE 3-14**

Comparison of product-bound structures of *cvADC* to *T. brucei* ODC: Structural alignment of the active sites of *cvADC* (purple) and *T. brucei* ODC (1F3T; teal). Residues are numbered according to the *cvADC* 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 *cvADC* and PLP-putrescine is displayed in teal for *T. brucei* ODC.



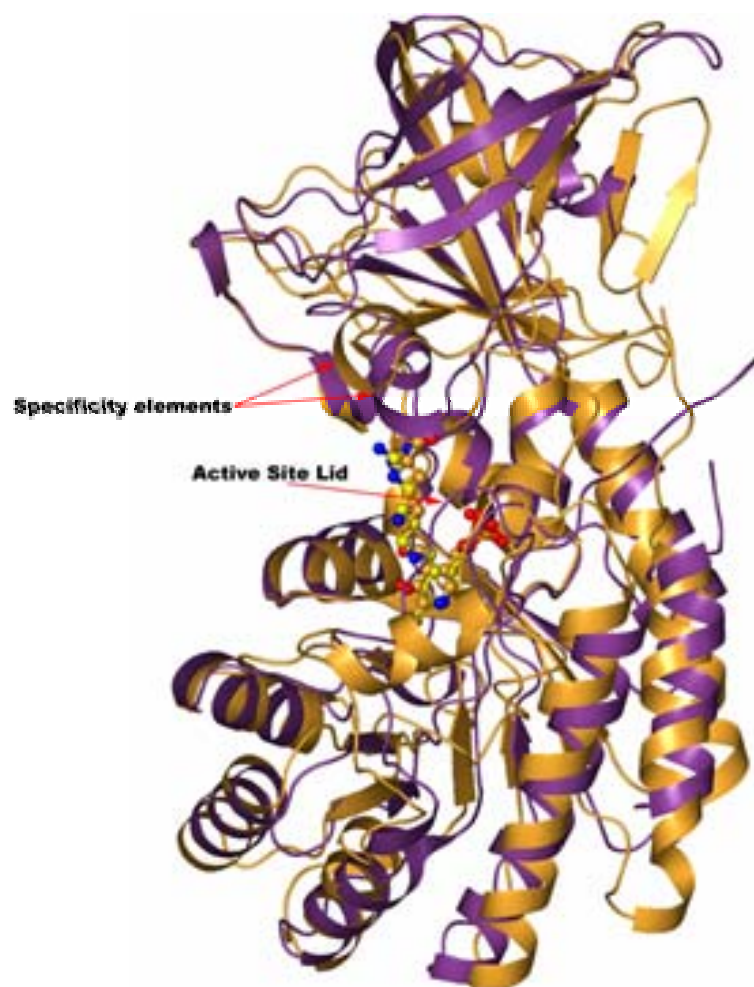
**FIGURE 3-15**

RMSD plot from the structural alignment of *cv*ADC and *T. brucei* ODC. The overall RMSD between the structures is 1.2 Å (See FIGURE 3-11). There are 33 residues in *cv*ADC that deviate by more than 2.5 Å in comparison to the ODC structure.

Chlorella virus ADC L-arginine	<i>T. brucei</i> ODC L-ornithine	<i>M. jannaschii</i> DAPDC L-diaminopimelate
V143	C164	G188
L217	F238	L265
S291	N327	R343
N292	C328	P344
F295	Y331	Y347
E296	D332	E348
L357	G393	M405

**TABLE IV**

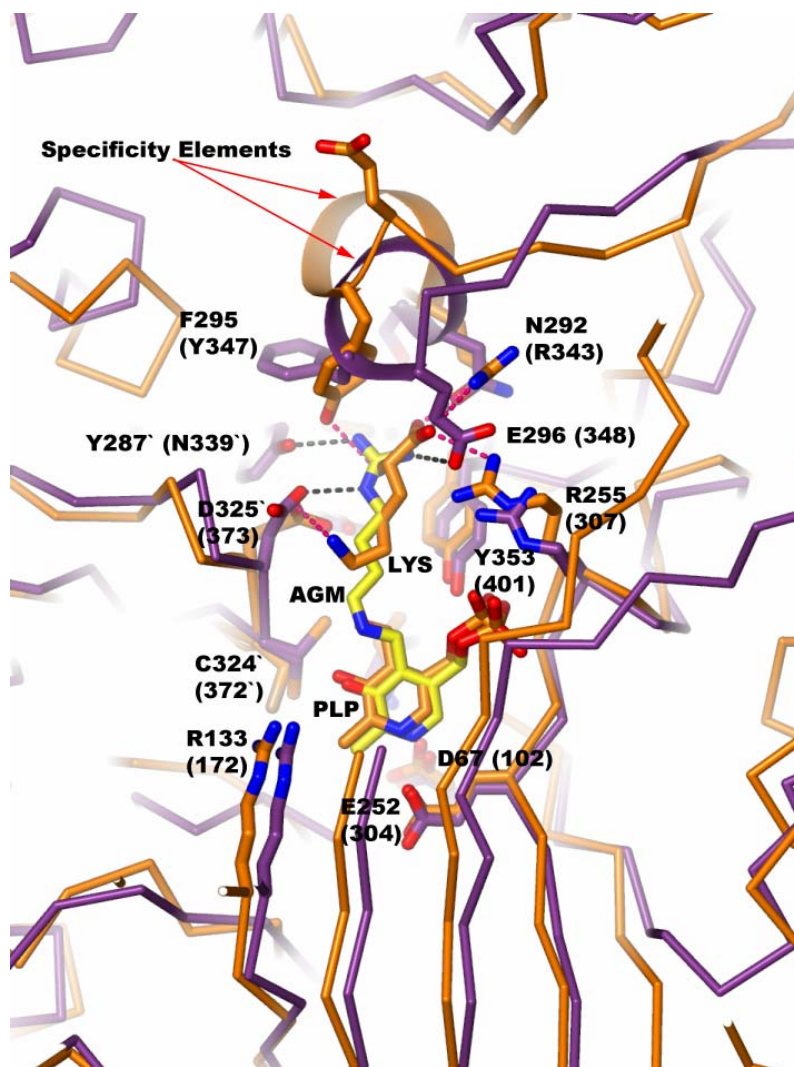
Amino acid variations observed within 5 Å of product in the active site of *cvADC* relative to *tbODC* and *mjDAPDC*. The primary substrate is listed below the enzyme name.



**FIGURE 3-16**

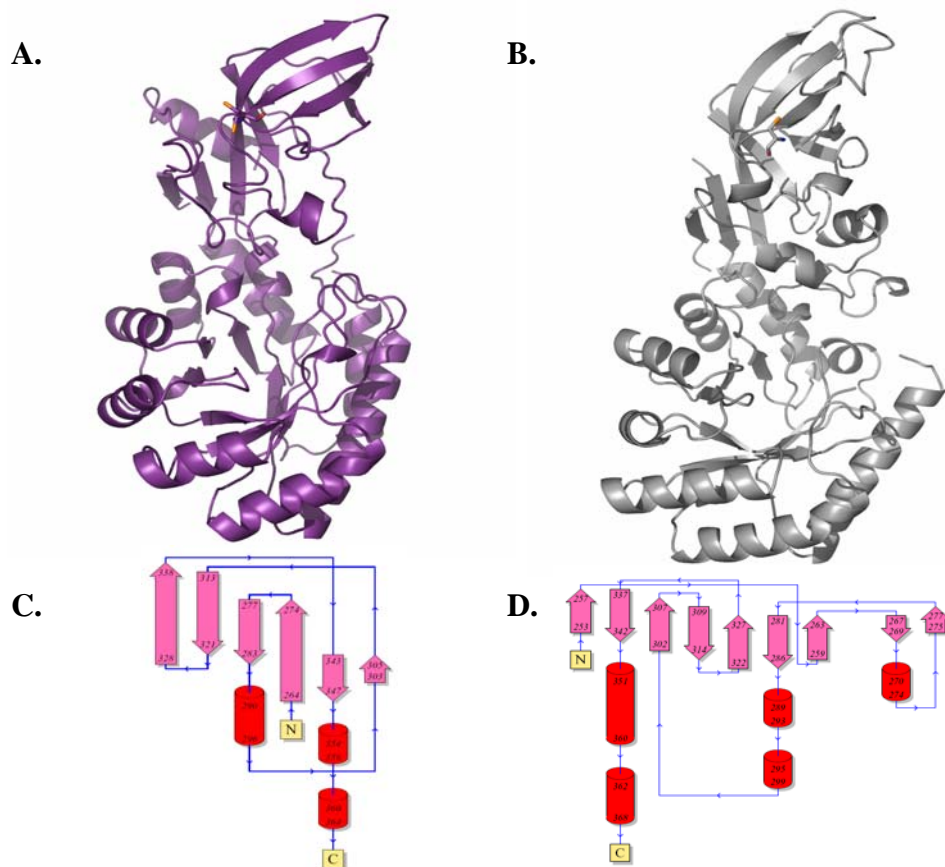
Comparison of the product-bound structures of *cvADC* and *M. jannaschii* DAPDC: A ribbon diagram of the monomeric structures of *cvADC* (purple) and *M. jannaschii* DAPDC (1TWI ; orange) for the alignment of the N-terminal domain of *cvADC* (residues 23-261) and *M. jannaschii* DAPDC (residues 51-312). The overall RMSD between the structures is 2.4 Å for the monomer.





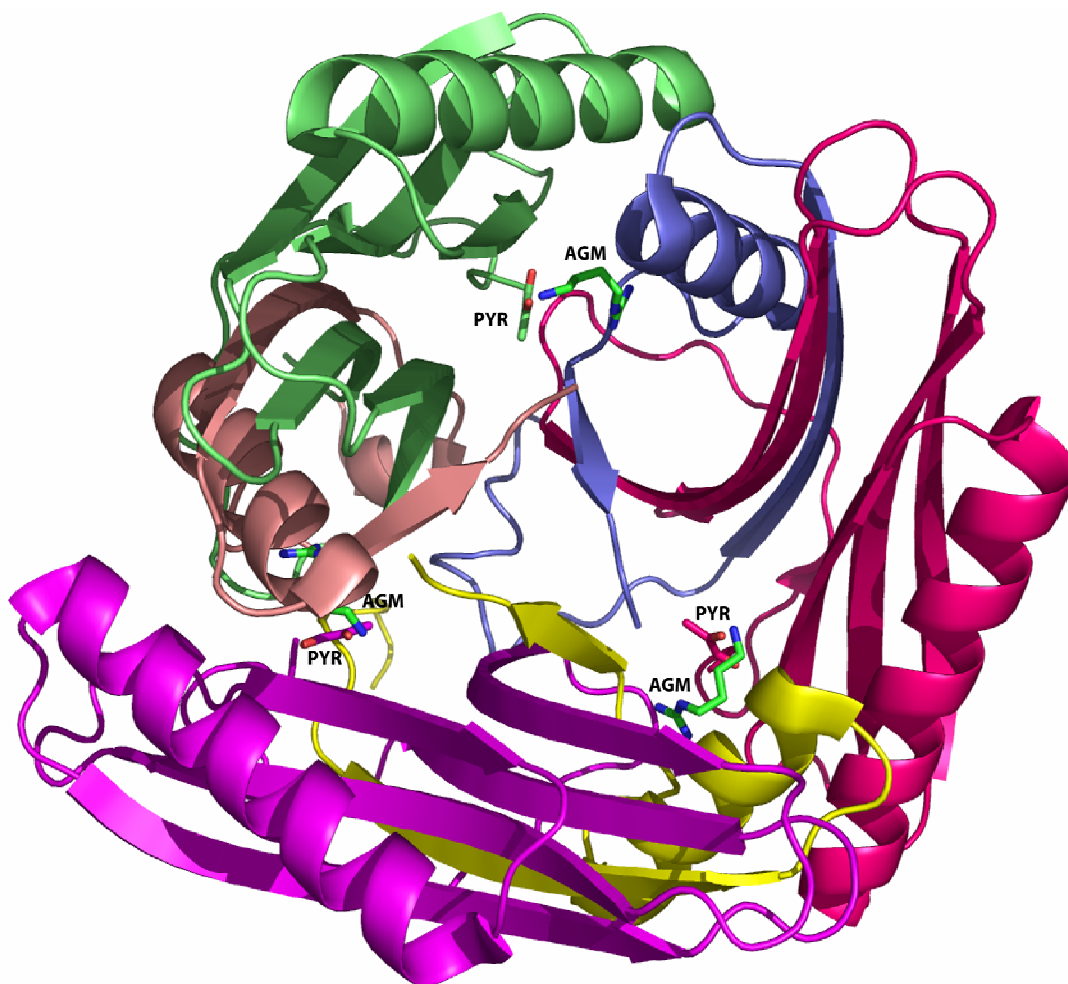
**FIGURE 3-17**

Comparison of the active-sites of product-bound *cvADC* and *M. jannaschii* DAPDC. *cvADC* is shown in purple and *M. jannaschii* DAPDC in orange. Residues are numbered according to the *cvADC* sequence with the equivalent DAPDC residue in parentheses. The PLP cofactor and agmatine from *cvADC* are shown in yellow sticks and the PLP and lysine molecule from *mjDAPDC* are shown in orange sticks.



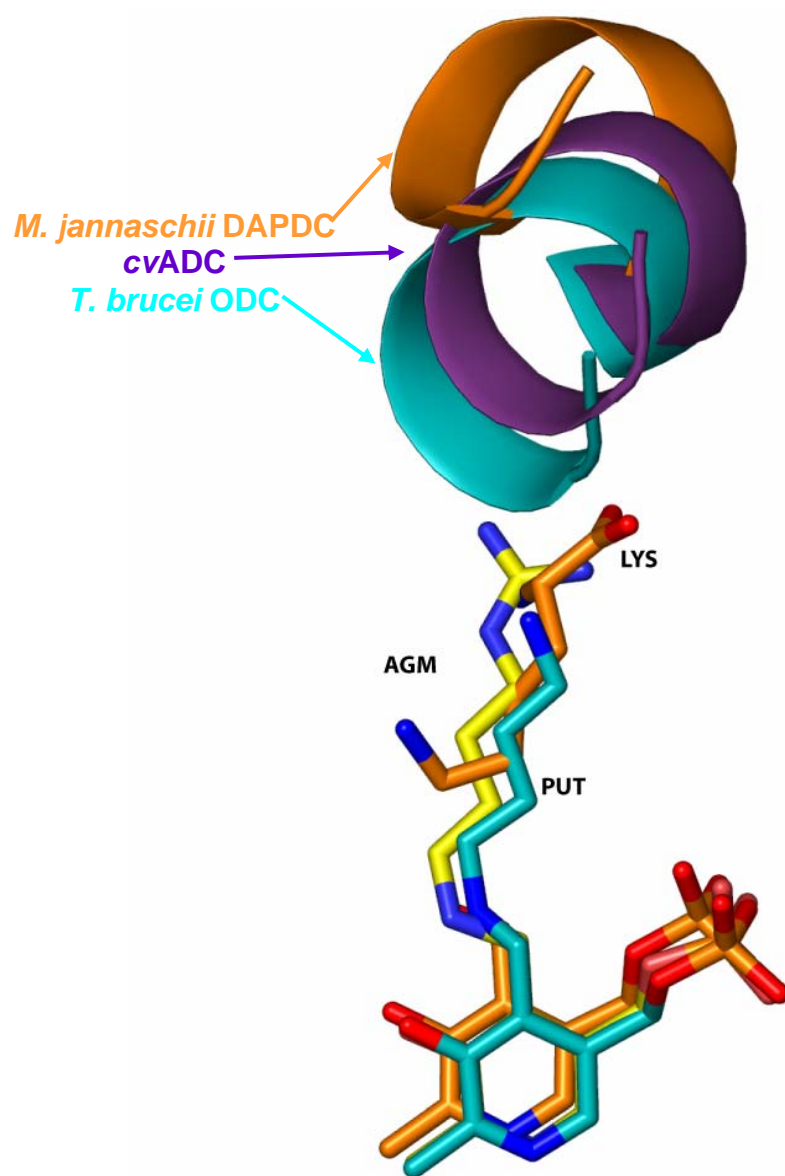
**FIGURE 3-18**

Comparison of *cvADC* and *M. tuberculosis* AR monomers. Both *cvADC* (A) and *mtAR* (B) fold into the characteristic Group IV structures. However, the C-terminal b-barrel of *mtAR* has rotated towards to N-terminal a/b barrel relative to *cvADC*. The topologies of the C-terminal barrel domains of *cvADC* (C) and *mtAR* (D) are shown. The N- and C-terminals of b-barrel are indicated.



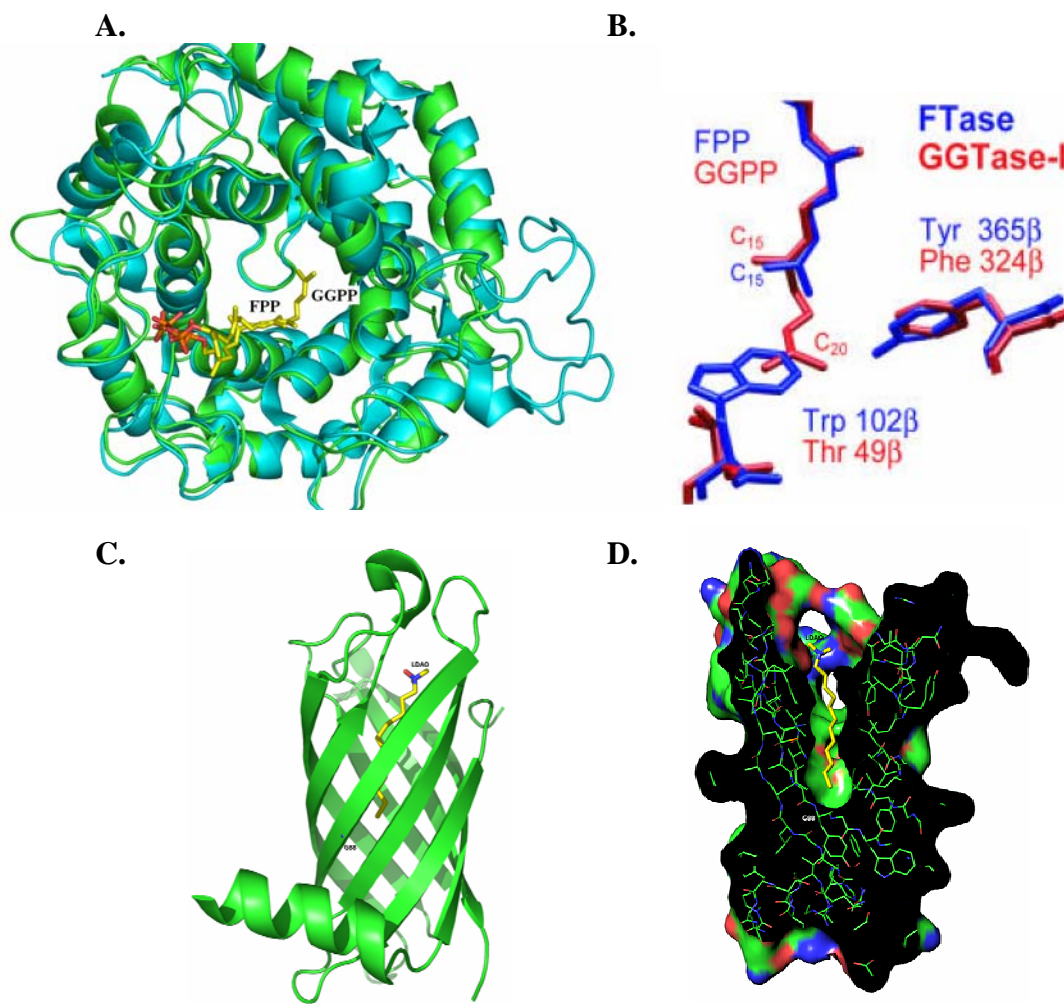
**FIGURE 3-19**

Structure of the pyruvoyl-dependent arginine decarboxylase from *M. jannaschii*. The biologically active oligomer is a trimer. Each trimer contains a pyruvoyl cofactor (labeled PYR) created by an auto-catalytic serinolysis reaction. The bound agmatine (AGM) is observed in each of three active sites of the oligomer adjacent to the cofactor.



**FIGURE 3-20**

Molecular rulers in Group IV decarboxylases. Group IV decarboxylases have evolved their specificities by increasing the distance between the C4' atom of PLP and the C $\alpha$  atoms of the specificity element.



**FIGURE 3-21**

Further examples of enzymes with molecular rulers (**A**) The structures of FTase and GGTase are nearly superimposable. (**B**) The basis of isoprenoid substrate chain-length selectivity is the identity of residues on helix 5 and 12 (Trp-102/Tyr-365 in FTase and Thr-49/Phe-324 in GGTase). This figure was adopted from (90). (**C**) The structure of *E. coli* PagP with a bound detergent molecule. (**D**) Cut-away of PagP demonstrating that Gly-88 serves as a molecular ruler for substrate acyl chain-length.

## CHAPTER FOUR

### DETERMINANTS OF SUBSTRATE SPECIFICITY IN GROUP IV DECARBOXYLASES

#### Introduction

Though the global structural differences between eukaryotic ODC and *cvADC* are minimal, the identity and positioning of the specificity element may explain the basis of specificity in Group IV decarboxylases. In the comparison of *mjDAPDC* and *tbODC*, it was observed that nearly 100 residues deviated by more than 2.5 Å between the structures (Figure 3-2). Comparison of *cvADC* and ODC reveals that only about 30 residues share a similar level of deviation (Figure 3-15). This raises the possibility that a limited number of structural changes, presumably conferred by a limited number of substitutions, could determine substrate specificity in the Group IV fold. In order to understand the sequence-structure-specificity relationship in Group IV decarboxylases, I undertook the task of switching specificities between ODC and ADC. The goal of this project was to isolate the positions or residues that contribute significantly to the enzymes' substrate preference. By mutating residues at specific sites in ADC to the corresponding residues in ODC, and vice versa, I hoped to increase the activity ( $k_{cat}/K_m$ ) of ADC for L-ornithine and the activity of ODC for L-arginine. To predictably manipulate an enzyme's substrate specificity indicates that the basis of specificity is understood to some level.

Defining determinants of specificity by swapping residues between structurally similar enzymes with differing preferences has been attempted with varying degrees of success. One notable achievement was the altering of lactate dehydrogenase specificity (LDH) to that of malate dehydrogenase (MDH) (99). The authors of this study found that by swapping a single residue (Gln-102) in

LDH to the corresponding Arg observed in MDH, the specificity for LDH could be effectively changed from the cognate substrate, pyruvate, to oxaloacetate. The guanidinium group of Arg in the LDH Q102R mutant was predicted to directly contact the 3-carboxyl moiety of oxaloacetate. The introduction of this new salt-bridge interaction was thought to provide the extra binding energy needed to preferentially catalyze hydride transfer from the cofactor, NADH, to oxaloacetate. On the other hand, with pyruvate as substrate, the guanidinium of Arg-102 in LDH Q102R would be unpaired, leading to an active site unfavorable to the proper arrangement of residues to catalyze hydride transfer.

A now classic case of swapping specificities was the attempt by Hedstrom et al. to convert trypsin to chymotrypsin (*100*). Trypsin prefers, on the order of nearly one-million fold, substrates with Arg or Lys at the P1 position. Chymotrypsin prefers substrates with large hydrophobic residues at the substrate P1 position. Substituting the S1 binding site residue D189 of Trypsin with the Ser observed in chymotrypsin conferred a change in ester hydrolysis specificity (*101*). However, for a change in amide hydrolysis specificity, substitutions of surface loops distal from the active site were necessary. Even with these additional substitutions, the mutant Trypsin (Trypsin [S1+L1+L2]) did not exhibit the same level of proficiency ( $k_{\text{cat}}/K_{\text{m}}$  for Trypsin [S1+L1+L2] of  $2.8 \times 10^3$  for P1-Phe substrates vs.  $k_{\text{cat}}/K_{\text{m}}$  of  $1.6 \times 10^6$  for wt Chymotrypsin with P1-Phe substrates). It became clear from this study that substitutions both proximal and distal to the active site play a significant role in determining the conformation of amino acids in the active site and/or modulating the level of flexibility necessary for efficient catalysis (*102*).

Swapping of specificities of a PLP-dependent enzyme has also been attempted. Kirsch and colleagues have put much effort into determining the basis of specificity in *E. coli* aspartate aminotransferase (AATase) and tyrosine aminotransferase (TATase). Through numerous investigations, they have



determined that a set of six residues, which they refer to as Hex, plays a dominant role in substrate preference in this fold (103, 104). In the case of converting AATase to TATase, the best case (as measured by  $k_{cat}/K_m$ ) was the six-mutant “Hex” AATase which displayed nearly equal preference for aspartate and phenylalanine, whereas wild-type AATase prefers aspartate to phenylalanine by nearly 1000-fold. In the opposite direction, the investigators were less than successful. The six-mutant “retroHex” TATase, showed reduced activity for both aspartate and phenylalanine as substrates.

One can derive a number of conclusions from these investigations. First, only in rare cases can a single mutation substantially alter the specificity of an enzyme. Second, residues not directly contacting the substrate (i.e. second-shell residues) play an important role in substrate preference. Finally, substitutions are highly context-dependent. That is, although substitution of residues in one direction (e.g. AATase  $\rightarrow$  TATase) can result in a meaningful change in specificity, the substitutions in the opposite direction (e.g. TATase  $\rightarrow$  AATase) may not necessarily produce a reciprocal result.

Determining the residues that play a dominant role in specificity in the Group IV decarboxylases is a worthwhile pursuit for a number of reasons. As mentioned in Chapter I, several Group IV decarboxylases are known drug targets while others may yet become clinically relevant. Furthermore, because PLP is such a versatile cofactor, many groups have taken on the task of engineering both the reaction and substrate specificity of PLP-dependent enzymes for commercial purposes (105, 106). Jackson et al. have previously demonstrated that the reaction specificity of ODC can be altered via a single active site residue substitution (32). In this chapter I will discuss my efforts to alter the specificity of ODC and ADC. Previously (Chapter II) I mentioned that a single active site mutation may account for the altered specificity of *cvADC* with respect to ODC (Figure 2-7). Asp-332 is highly conserved in the ODC family and plays an important role in substrate



binding and catalysis [Figure 2-3B; (95)]. The equivalent position in the *cvADC* is residue 296, which contains a Glu, suggesting that this substitution is a key determinant in the change in substrate specificity observed for this enzyme (Figure 2-3B). Here I will test the validity of this hypothesis. Additionally, having solved the structure of ADC, additional active site differences were identified that may play a role in determining specificity. One important revelation from the agmatine-bound *cvADC* structure was the substitution of Cys-328 ODC with Asn-292 *cvADC*. As mentioned in Chapter III, the carbamoyl oxygen of Asn-292 in *cvADC* is with H-bonding distance (3.3 Å) of the N1 atom of agmatine. Though the interaction of Asn-292 with substrate was not predicted by the alignment of *cvADC* and *tbODC* (Figure 2-3), an alignment of various ODCs and ADCs reveals that an Asn at this position seems to be well conserved (Figure 4-1). Thus, it may be that ADCs have evolved this new interaction to specify arginine over other basic amino acids. In terms of residues that position the  $3_{10}$  helix, an interesting substitution is a Leu-357 in *cvADC* (Figure 3-10), which is a Gly or Ala in ODC (Figure 4-1). A Leu at this position seems to be well conserved in ADCs throughout evolution and a small Ala or Gly seems to predominate at this position in ODC. In Chapter III I discussed the proximity of this residue to the  $3_{10}$ -helix-residue Ser-291, which influences the position of the helix. In this Chapter I will discuss my attempt to convert the specificity of *cvADC* from L-arginine to L-ornithine, and vice versa for *tbODC*. In addition to creating a series of single-mutant *cvADC*s and *tbODC*s, I created triple-mutants (i.e. *cvADC* N292C/E296D/L357G and *tbODC* C328N/D332E/G393L) with the belief that the combination of a select number of substitutions affecting both the nature of substrate binding and overall volume in the active site would yield an appreciable change in substrate preference.

In addition to altering the specificity of ODC and ADC, I also determined the inhibition constants of various substrate analogs (D-ornithine, D-lysine and D-

arginine) and products (putrescine, cadaverine and agmatine) for ODC and ADC. The goal of this study is to understand the contribution of ground-state binding to specificity. Previous steady-state kinetic data from *tb*ODC demonstrated that the  $K_i$  value for putrescine (350  $\mu$ M) was similar to the  $K_m$  for ornithine (46). Furthermore, specificity for L-ornithine by *tb*ODC over L-lysine or L-arginine was manifest mainly in a difference in the Michealis constant ( $K_m$ ) rather than the rate of catalysis (21). The authors of this study also demonstrated that a mutation of Asp-361 to Ala in *tb*ODC yielded a 2000-fold increase in  $K_m$  for Orn decarboxylation without significantly lowering the turnover rate. Furthermore, the binding affinity of D361E mutant of *tb*ODC for putrescine is reduced to the same extent as for Orn ( $K_m = K_i = 30$  mM) suggesting that  $K_m$  is a reflection of the true dissociation constant (21). If  $K_m$  and  $K_i$  can be considered reasonable estimates of  $K_S$  than it may be possible to extrapolate the contribution of ground-state binding to specificity from inhibition studies of *cv*ADC and *tb*ODC.

## Materials

The Infinity™ carbon dioxide detection kit was purchased from Thermo Scientific. All other chemicals and reagents were purchased from Sigma.

## Experimental Methods

*Protein Structure Modelling* Mutants of *cv*ADC and *tb*ODC were created with the *SWISS-MODEL* homology modeling program (107). The homology models for *cv*ADC N292C, *cv*ADC E296D, *tb*ODC G393L and *tb*ODC C328N/D332E/G393L were created using the structure of agmatine-bound *cv*ADC (PDB ID 2NVA). The homology models for *cv*ADC L357G, *cv*ADC N292C/E296D/L357G, *tb*ODC C328N and *tb*ODC E296D were created using the structure putrescine-bound *tb*ODC (PDB ID 1F3T). The homology models were created by the ProMod package (108) and energy minimization of all models was carried out by GROMOS (109).

*Site-directed mutagenesis* Single-mutants ODC C328N, D332E, G393L, the triple-mutant ODC C328N/D332E/G393L from *T. brucei*, single-mutants *cvADC* N292C, E296D, L357G and the triple mutant N292C/E296D/L357G were produced using the QuikChange<sup>TM</sup> site-directed mutagenesis kit from Stratagene (La Jolla, CA). Mutagenic primers were purchased from Integrated DNA Technologies (Coralville, IA). The sequences for all plasmids were determined (UTSW DNA Sequencing Center).

*Production and purification of CVADC and T. brucei ODC mutant* The *T. brucei* ODC mutant proteins were produced as described in Chapter II with the following changes. Protein production was induced at 23° C for 6 hours. *CVADC* mutants were produced exactly as described in Chapter II. All mutants were produced as His<sub>6</sub>-tagged proteins and were purified as described in Chapter II.

*Inhibition of wild-type cvADC and T. brucei ODC with substrate analogs* The inhibition constants ( $K_i$ ) were determined for D-arginine, agmatine, D-ornithine, putrescine, D-lysine and cadaverine with both *cvADC* and *T. brucei* ODC. Assays were conducted with substrates (0.1-50mM), either L-arginine for *cvADC* or L-ornithine for *T. brucei* ODC. Enzyme activity was measured for at least four concentrations on each inhibitor. For each inhibitor concentration, the apparent  $K_m$  ( $K_{m,app}$ ) values were calculated by the Michaelis-Menten equation.

$$v = V_{max} * [S] / ([S] + K_{m,app})$$

A plot  $K_{m,app}$  values (y-axis) versus the concentration of inhibitor [I] (x-axis) was then produced for each inhibitor. The value for  $K_i$  was obtained from the intercept of the fitted line on the x-axis (i.e.  $y=0$ ,  $[I] = -K_i$ ).

*Steady-state analysis of wild-type and mutant enzymes* Steady-state parameters ( $k_{cat}$  and  $K_m$ ) were obtained using the carbon dioxide detection kit from Thermo Electron Corporation (See Chapter II). Wild-type and mutant enzymes were assayed in buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 500  $\mu$ M EDTA, 0.03% Brij-35, 1mM DTT and 20  $\mu$ M PLP) at concentration of 60 to 120

nM. Substrate (L-Arg, L-Orn, L-Lys) concentrations ranged from 0.1 to 50 mM. All assays were conducted at 37° C.

## Results

*Models of cvADC and T. brucei ODC mutants* The mutants produced by the SWISS-MODEL homology modeling program are shown in Figures 4-2 and 4-3 (*cvADC* mutants) and Figures 4-4 and 4-5 (*tbODC* mutants). For the *cvADC* N292C mutant (Figure 4-2A), the model suggests that this substitution removes the H-bond interaction between the Asn-292 and the guanidinium group of agmatine. In the model, both putrescine and agmatine make equal number of H-bond interactions with active site residues. This is also the case in the *cvADC* E296D model (Figure 4-2B). In the *cvADC* L357G model (Figure 4-3 A), the shortening of the distance between the 3<sub>10</sub> helix and PLP has pushed Glu-296 out of the binding pocket. Regardless, there are still more H-bond interactions between agmatine in comparison to putrescine. The model of *cvADC* N292C/E296D/L357G (Figure 4-3 B) illustrates that the triple-mutant's active site is smaller in volume, similar to the *cvADC* L357G model, and that it would form more H-bond interactions with agmatine in comparison to putrescine.

The *tbODC* C328N (Figure 4-4 A) model illustrates the notion that this mutant may form more H-bond interaction with agmatine, in comparison to ornithine, but that steric hindrance may reduce the affinity of this mutant of L-arginine. The *tbODC* D332E (Figure 4-4 B) mutant reveals that an equal number of H-bonds could be formed between the mutant enzyme and agmatine and putrescine, however, the side-chain of Glu-332 has swung out of the binding pocket and does not form interactions with substrate. The model of *tbODC* G393L (Figure 4-5 A) suggests that a larger number of H-bonds could be formed between the enzyme and putrescine in comparison to agmatine. However, the substitution could increase the overall volume of the pocket, facilitating the

binding of larger ligands. The *tbODC* C328N/D332E/G393L model (Figure 4-5 B) suggests that preferential catalysis of L-arginine over L-ornithine by the larger number of H-bond interaction between the mutant enzyme and agmatine and an increase in the volume of the active site.

*Inhibition of wild-type cvADC and T. brucei ODC with substrate analogs* The inhibition constants ( $K_i$ ) of the products and substrate analogs for *cvADC* and *tbODC* are listed in Table V. It is clear to see that the inhibitors most closely resembling the cognate substrate for each enzyme display the lowest  $K_i$  values. For *cvADC* agmatine and D-arginine are at least 20- and 10-fold more potent, respectively, than the next most potent inhibitor (cadaverine). For *tbODC* putrescine and D-ornithine are approximately 10- and 3-fold more potent than the next most potent inhibitor, agmatine. Curiously the pattern of inhibition potency does not strictly follow that of substrate preference. For instance, *cvADC* displays a 600-fold preference for L-arginine over L-ornithine and a 1000-fold preference for L-arginine over L-lysine (Figure 2-4). In contrast cadaverine is only about 20-fold weaker at inhibiting *cvADC* than agmatine, while putrescine is 50-fold weaker than agmatine. *tbODC* displays a 400-fold preference for L-ornithine over L-lysine and at least a 1000-fold preference for L-ornithine over L-arginine (46). On the other hand, agmatine is only a 10-fold weaker inhibitor compared to putrescine and cadaverine is a 30-fold weaker inhibitor than putrescine.

*Steady-state analysis of wild-type and mutant cvADC and tbODC* The steady-state kinetic parameters of wild-type and mutant *cvADC* and *tbODC* are listed in Table VI. The activity of single mutants varied from inactive (e.g. *cvADC* E296D) to nearly as active as wild-type enzymes (e.g. *cvADC* L357G). Both triple-mutant enzymes (*cvADC* N292C/E296D/L357G and *tbODC* C328N/D332E/G393L) were almost as active as wild-type enzymes. On the other hand, at least one mutant enzyme, *cvADC* E296D, displayed no detectable activity, within the context of the coupled-enzyme assay, with certain L-lysine.

With the exception of *cvADC* E296D, all the *cvADC* mutants displayed broadened substrate specificities with respect to wild-type *cvADC* (Figure 4-2). Prior to solving the structure of *cvADC*, the only residue in direct contact with the substrate predicted to differ between ODC and *cvADC* was Asp-332, which is a Glu in *cvADC*. As shown in Table VI and Figure 4-6, *cvADC* E296D prefers L-arginine 20-fold more (as measured by  $k_{\text{cat}}/K_m$ ) than L-ornithine, a much decreased level of discrimination in comparison to the 600 fold preference displayed by wild-type *cvADC*. A recent investigation conducted by Jeongmi Lee in the Phillips laboratory on the activity of a number of ODC homologs from various microbial species has revealed that the identity of position 332 is not predictive of L-ornithine specificity (110). Although the residue at this position is important to overall activity, it seems that it is not the dominant residue in determining specificity. *cvADC* N292C displayed nearly equal activity for all substrates (Table VI and Figure 4-6). Its activity was reduced with respect to L-arginine mainly by a greater than 10-fold increase in  $K_m$ . The gain in catalytic efficiency of *cvADC* N292C for L-ornithine and L-lysine was largely a result in a greater than 20-fold decrease in  $K_m$  for both substrates. *cvADC* L357G displayed the highest level of activity of any of the *cvADC* single mutants. This mutant also maintained a nearly 10-fold preference for L-arginine, though its  $K_m$  for L-arginine was increased by a factor of five. Its activity with L-ornithine and L-lysine was improved primarily by a 20-fold decrease in  $K_m$  while  $k_{\text{cat}}$  was marginally increased (Table VI and Figure 4-6). The triple mutant *cvADC* N292C/E296D/L357G displayed nearly equal activity towards L-arginine and L-lysine and a slightly lower activity towards L-ornithine (Table VI and Figure 4-6). The overall proficiency of the *cvADC* triple mutant was due mainly to the sub-millimolar  $K_m$  values for L-arginine or L-lysine, with the  $k_{\text{cat}}$  values below one per second.

All the *tb*ODC mutants displayed some level of preference of L-ornithine, though activity on non-cognate substrates was increased in certain contexts. *tb*ODC D332E maintained a 10-fold preference for L-ornithine over L-lysine and a 1000-fold preference over L-arginine. *tb*ODC C328N exhibited a 100-fold preference for L-ornithine over L-lysine or L-arginine (Table VI and Figure 4-6). *tb*ODC G393L preferred L-ornithine and L-lysine nearly equally and 10-fold discrimination against L-arginine. This stands in stark contrast to the kinetic data for wild-type *tb*ODC, which shows a 1000-fold preference L-ornithine over L-arginine (Figure 4-6). The improved activity of *tb*ODC G393L for L-arginine was due both to a 20-fold increase in  $k_{cat}$  and a 3-fold decrease in  $K_m$ . The *tb*ODC triple mutant's specificity profile diverged from those of the single-mutants. Specifically, the triple mutant preferred L-lysine nearly 10-fold more than L-ornithine or L-arginine. Most of the *tb*ODC triple-mutant's preference is reflected in the sub-millimolar  $K_m$  value for L-lysine in comparison to millimolar  $K_m$  values for L-ornithine or L-arginine.

Several studies have shown that amino acid residues that are distant from the active site, and do not contact ligand directly, are often important in the change of function between homologous proteins [e.g. Refs (100, 111-113) ]. In the case of ODC, previous studies have demonstrated that amino acid residues in the dimer interface distant from the active site are important for enzyme activity (95). The results of my mutagenesis investigations demonstrate that the positions I chose to mutate have significant impacts on substrate preference in both the *cv*ADC and *tb*ODC context. The observation that many of the mutants displayed increased activity for non-cognate substrates suggests that the three positions I targeted for the mutagenesis have profound impacts on the formation of the active site such that non-cognate substrates may be decarboxylated at rates comparable to wild-type values for cognate substrates.

## Discussion

Engineering a new specificity into an existing enzyme construct is a challenging pursuit at best. Many investigators have attempted to build in new specificities by swapping residues between two structurally homologous but functionally different enzymes (99, 100, 103). The previous attempts at fully emulating the specificity profile and overall proficiency of naturally observed enzymes via rationally-designed mutants have been short of success. This is so because identifying all the residues playing dominant roles in specificity is difficult. I believe that the increase in sequence data and the rapidly-developing tools of bioinformatics will ultimately facilitate an understanding of the determinants of specificity such that *de novo* enzymes can be built.

The three previous examples of specificity-engineering investigations mentioned above offer some valuable lessons. The single-site mutation that effectively converted an LDH to an MDH demonstrated that it is possible that the determinants of enzyme specificity could be isolated (99). The authors of this study benefited from detailed kinetic, mechanistic and structural data for the enzyme. Thus, they were able to deduce not only which position/residues plays a dominant role in specificity but also how preference is carried out. Specifically, the authors rationalized that induced-fit is utilized to confer discrimination between similar substrates. In the attempt to convert Trypsin specificity to that of Chymotrypsin, two important aspects of enzyme-substrate interactions were elucidated. First, specificity is overwhelmingly conferred in the rate of chemistry (i.e. acylation for proteases) and not in ground-state substrate binding. Second, residues, and more specifically loops, distant from the active site play a large role in binding the binding of the transition-state structure of the substrate. Therefore, positions distant from the site of chemistry determine specificity (100). Finally, the attempt by Kirsch and colleagues to swap the specificities of AATase and TATase indicate that substrate interactions with enzyme active-site residues are






highly context-dependent. An appreciable shift, though broadening seems to be the apt term, in specificity was observed for AATase mutants, which displayed a greater than 300-fold improvement for Phe compared to wild-type (104). However, none of the TATase mutants exhibited legitimate improvement with Asp as substrate in comparison to the wild-type TATase activity with Asp. Thus, success in converting enzyme specificities can sometimes be achieved only in one direction.

For my study of determinants of specificity in Group IV decarboxylases, I chose to mutate residues at the distal portion of the active site. Specifically, the residues I chose to mutate form interactions with the fragment substrate not undergoing chemistry. Although the mutants I produced were designed to “bind” substrates differently, it has been clear for some time that ground-state discrimination does not fully account for specificity (114). However, residues thought to bind substrates do impart a fair level in specificity in that the energy derived from the interactions of either the ground-state or transition-state conformation of a substrate is channeled into enhancing the rate of catalysis (115). Previous studies on the mechanism of the ODC revealed that the decarboxylation rate, following Schiff base formation, is the dominant kinetic step for substrate discrimination (116). The ideal “swapped” mutant would be one in which the new energy of interactions between a non-cognate substrate and mutated residues would be harnessed more efficiently than cognate substrate.

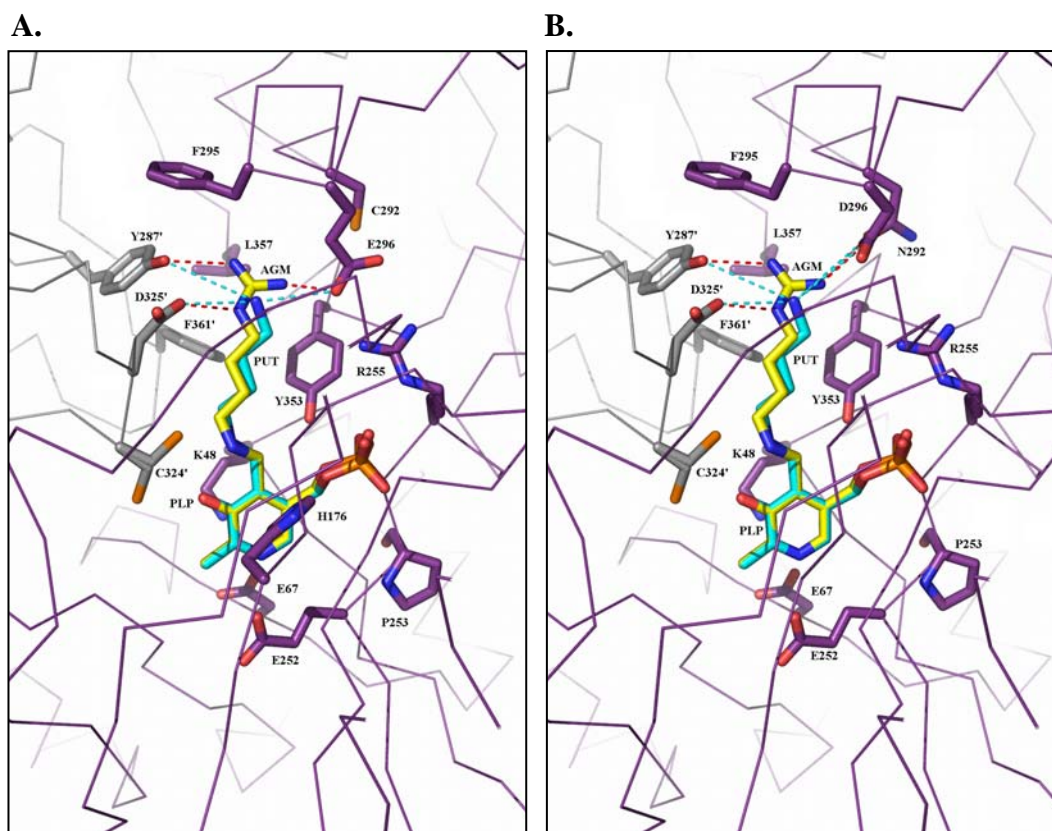
A plot of the activity ( $\log k_{\text{cat}}/K_m$ ) of wild-type and mutant *cvADC* and *tbODC* enzymes for L-arginine, L-ornithine or L-lysine is shown in Figure 4-2. One common theme that emerges from this plot is the broadened specificity of the various *cvADC* and *tbODC* mutants. Most striking is the specificity profile of *cvADC* N292C mutant which exhibits a slight preference for L-ornithine and L-lysine over L-arginine. This result demonstrates the importance of the carbamoyl moiety of Asn-292 of *cvADC* in distinguishing between L-arginine and L-

ornithine or L-lysine. The corresponding mutant *tb*ODC C328N, which displays a preference of L-ornithine over L-lysine and no activity with L-arginine, suggests that the identity of this position in Group IV decarboxylases is highly context-dependent. The specificity profile of *cv*ADC L357G suggests that the amino acid identity at this position also plays an important role in determining specificity. Mutation of the bulky leucine for a glycine in *cv*ADC has likely modified the positioning of the  $3_{10}$  helix at the distal end of the active site. In the context of the *cv*ADC L357G mutant, the  $3_{10}$  helix may be afforded a larger range of motions. On the other hand, all the *tb*ODC mutants displayed preference for L-ornithine over L-arginine. The specificity swap attempted here shares a number of similarities to the previous specificity-engineering investigations. Consistent with serine protease literature, it is abundantly evident that residues outside of the immediate substrate binding site likely play significant roles in determining specificity. Furthermore, interactions between substrate and enzyme active site residues are context dependent, a phenomenon observed in the aminotransferase investigations. Specifically, I have shown that while the single-mutant *cv*ADC N292C displays a slight preference for L-ornithine over L-arginine, the corresponding *tb*ODC C328N mutant plainly does not display altered specificity in comparison to wild-type *tb*ODC.

	Active site loop		3 <sub>10</sub> helix	
				
	143	217	292 296	337
cvADC	RCD-DP---NATVQLGNKFGAN	LDIGGGLHAD	YGGFSNVI FEKSVPTP	WGAYTNVL TTSFNG---FGE 366
NY2A	RCD-DP---NATVQLGNKFGAN	LDIGGGLHAD	YAGFSNVI FEKSVPIP	WGAYTNVL VTSFNG---FGE 366
MT325	RCD-DP---HALVKL-EKYGAL	IDIGGGMYAD	YGGFSNVI YEKAVPEP	WGAYTNCL VTSFNG---FGE 366
ecADC	ASQ-GSGKMQSSGGEKSKFGLA	FDVGGGLGVD	DKMYWNFS LFQSMPPDA	VGAYQEILGNMHNLF--GD 581
ypADC	ASQ-GSGKMQSSGGEKSKFGLS	FDVGGGLGVD	DKLWNFS LFQSMPPDA	VGAYQEILGNMHNLF--GD 582
crODC	RAD-DP---DARWPLGLKYGAE	LDIGGGFTGH	YGSFNCILYDGGQNPY	AGAYTVAGACDFNGIEFTT 370
tbODC	STD-DS---LARCLSVKFGAK	LDIGGGFPGT	YGSFNCILYDHAVVRP	MGAYTVVGTSSFNG---FQS 400
ngODC	KPMFDG---NARCPMGPKYGAL	LDVGGGFTSG	YGSMNCVL YDHATVNA	MGAYTKAAGSNFNG---FNT 419
srLDC	AVR-NN---KALVDLNTKFGAP	LDIGGGFPVP	YGCFSGIMYDHWY-P	MGSYTSVSA TRFNG---FYL 365
hODC	ATD-DS---KAVCLSVKFGAT	LDIGGGFPGS	YGSFNCILYDHAHVKP	MGAYTVAAASTFNG---FQR 402
mODC	ATD-DS---KAVCLSVKFGAT	LDIGGGFPGS	YGSFNCILYDHAHVKA	MGAYTVAAASTFNG---FQR 402
dmODC	KSE-AK---EAQCPLGDKFGCD	LDIGGGFPGS	YGSFNCILYDHQVIA	MGAYTMPIASAFNG---FEV 385
ceODC	AVS-DP---TATCPNLNLFKAD	IDMGGGFPGA	YGSFNCILFDHAHPIG	MGAYTLAAATTFNG---FSK 401
scODC	ATD-DS---TAQCRLSTKYGCE	LDVGGGFQFE	YGNMNCILFDHQEHPH	LGAYTSSAATQFNG---FEQ 454
ldODC	KTN-DS---KAQCSFSTKFGAP	LDIGGGFPGT	YHSFNCILFDHAHPPL	MGSYTTAAAGFFNG---FAT 676
ntODC	KPMLDG---NARSAMGPKYGAL	LDVGGGFTSG	YGSMNCVL YDHATVNA	MGAYTKAAGSNFNG---FNT 420
	164	238	328 332	393

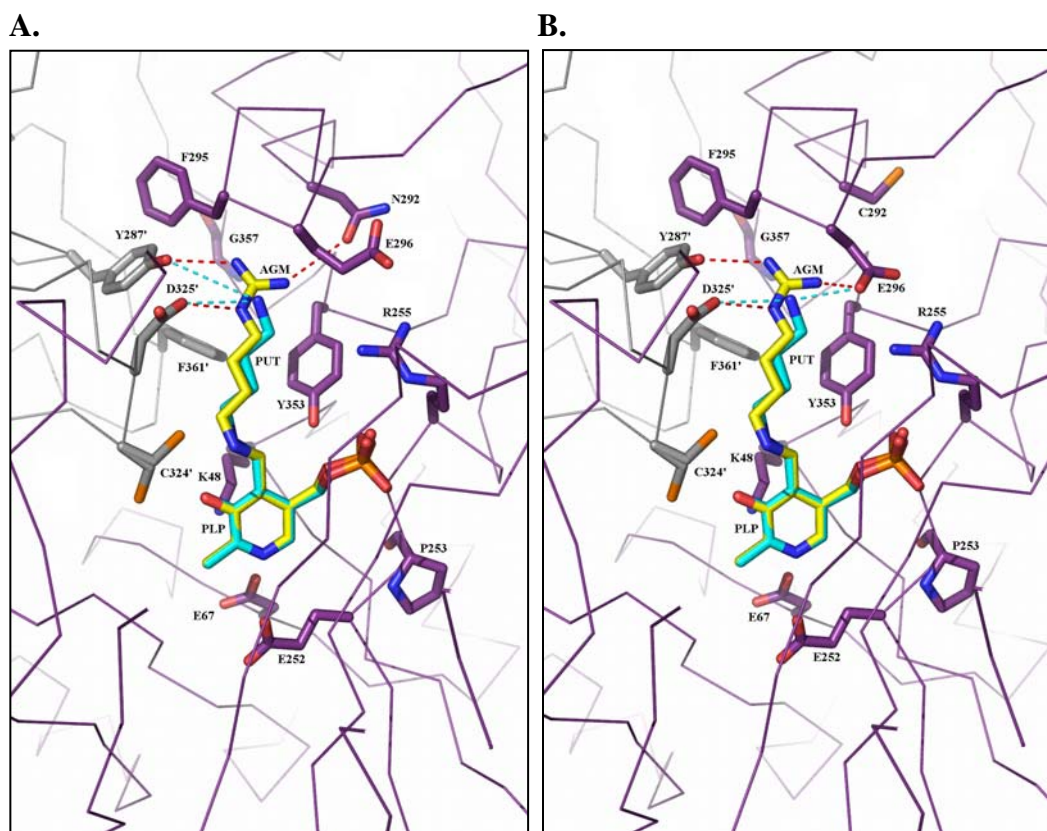
**FIGURE 4-1**

A partial alignment of 16 Group IV decarboxylases. The sequences in the alignment are *Paramecium bursaria* Chlorella virus Arginine decarboxylase (cvADC), two cvADC-related sequences (NY2A and MT325) from alternate Chlorella virus isolates, *E. coli* ADC (ecADC), *Y. pestis* ADC (ypADC), *Chlamydomonas reinhardtii* (crODC), *T. brucei* ODC (tbODC), *Nicotiana glutinosa* (ngODC), *S. ruminantium* L/ODC (srLDC), human ODC (hODC), mouse ODC (mODC), *Drosophila melanogaster* ODC (dmODC), *Caenorhabditis elegans* ODC (ceODC), *Saccharomyces cerevisiae* ODC (scODC), *Leishmania donovani* ODC (ldODC) and *Nicotiana tabacum* ODC (ntODC). Residues within the active site are highlighted red.



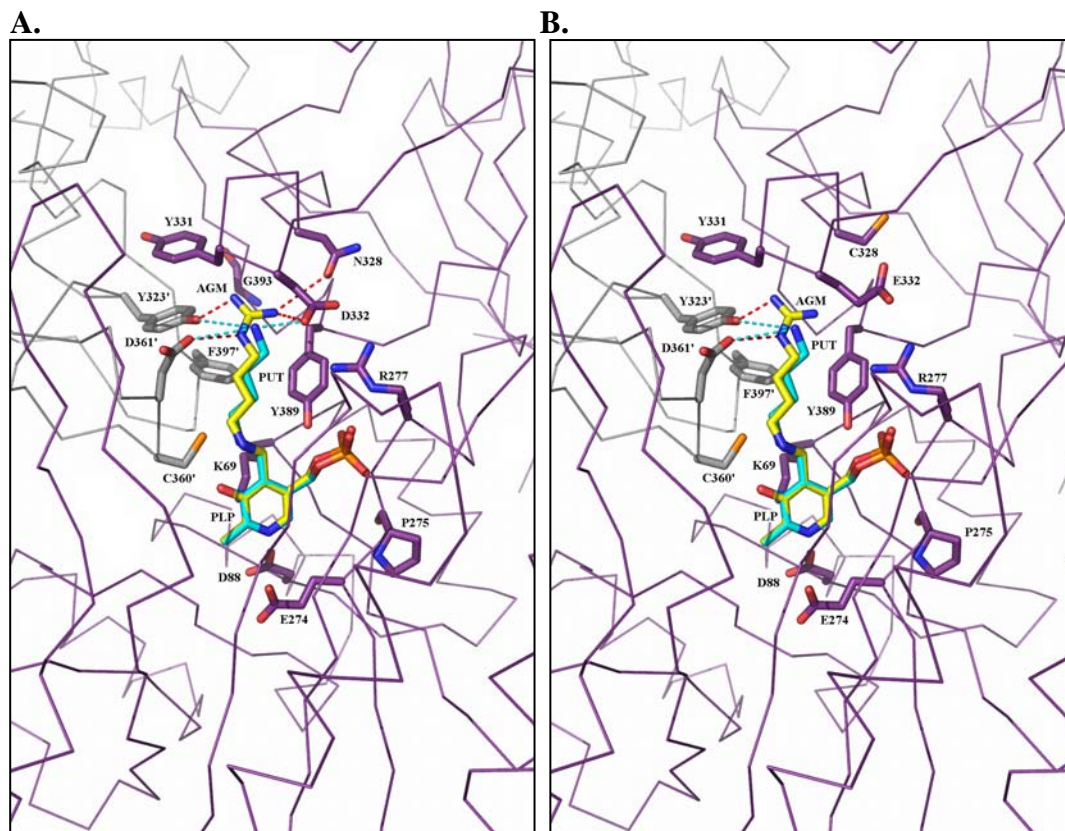
**FIGURE 4-2**

Active sites of mutant enzymes *cvADC* N292C *cvADC* E296D. The active sites of *cvADC* N292C (**A**) and *cvADC* E296D (**B**) are shown with both PLP-agmatine and PLP-ornithine complexes. Hydrogen bonds between agmatine (AGM) and enzyme are shown as red dashes and H-bonds between putrescine (PUT) are shown in cyan dashes.



**FIGURE 4-3**

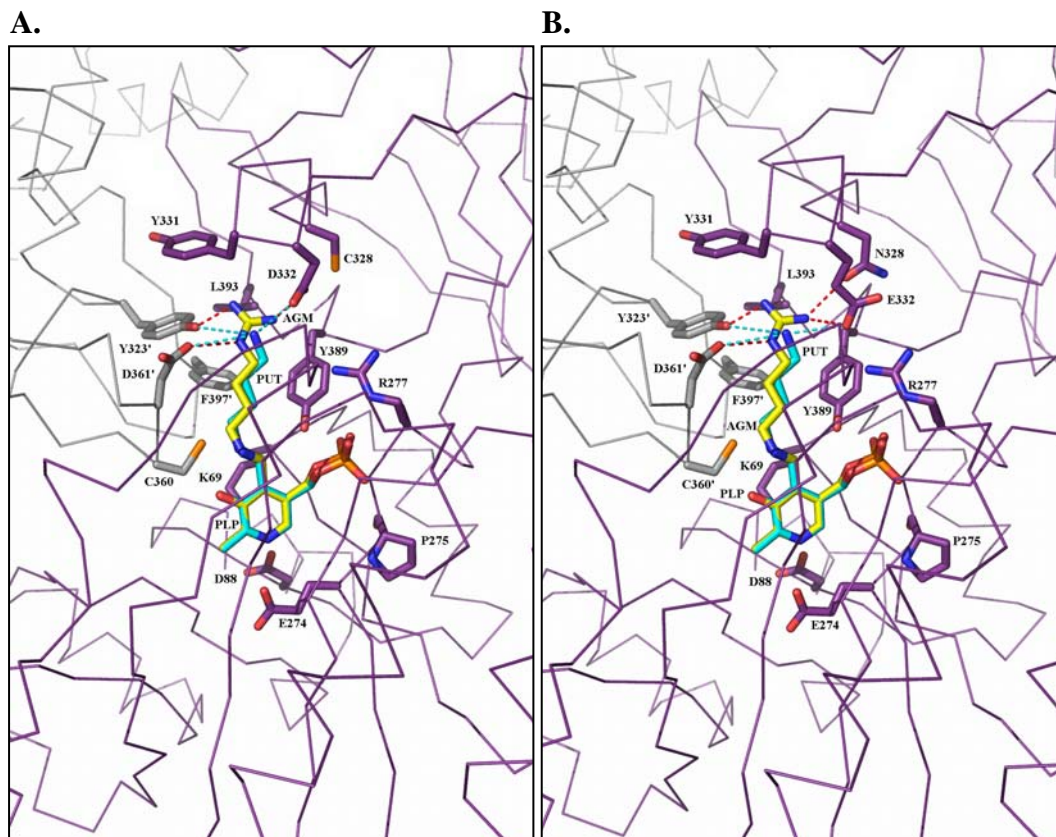
Active sites of mutant enzymes *cvADC* L357G and *cvADC* N292C/E296D/L357G. The active sites of *cvADC* L357G (**A**) and *cvADC* N292C/E296D/L357G (**B**) are shown with both PLP-agmatine and PLP-ornithine complexes. Hydrogen bonds between agmatine (AGM) and enzyme are shown as red dashes and H-bonds between putrescine (PUT) are shown in cyan dashes.



**FIGURE 4-4**

Active sites of mutant enzymes *tbODC* C328N and *tbODC* D332E. The active sites of *tbODC* C328N (**A**) and *tbODC* D332E (**B**) are shown with both PLP-agmatine and PLP-ornithine complexes. Hydrogen bonds between agmatine (AGM) and enzyme are shown as red dashes and H-bonds between putrescine (PUT) are shown in cyan dashes.





**FIGURE 4-5**

Active sites of mutant enzymes *tbODC* G393L and *tbODC* C328N/D332E/L357G. The active sites of *tbODC* G393L (**C**) and *tbODC* C328N/D332E/L357G (**D**) are shown with both PLP-argmatine and PLP-ornithine complexes. Hydrogen bonds between argmatine (AGM) and enzyme are shown as red dashes and H-bonds between putrescine (PUT) are shown in cyan dashes.

K <sub>i</sub> (mM)	Enzyme	
	<i>cv</i> ADC	<i>tb</i> ODC
D-arginine	1.8 ± 0.3	> 50
Agmatine	0.7 ± 0.1	3.1 ± 1.1
D-Ornithine	> 50	1.3 ± 0.4
Putrescine	32.7 ± 10.7	0.37 ± 0.1
D-Lysine	> 50	20.3 ± 5.4
Cadaverine	12.4 ± 2.4	12.3 ± 3.2

**TABLE V**

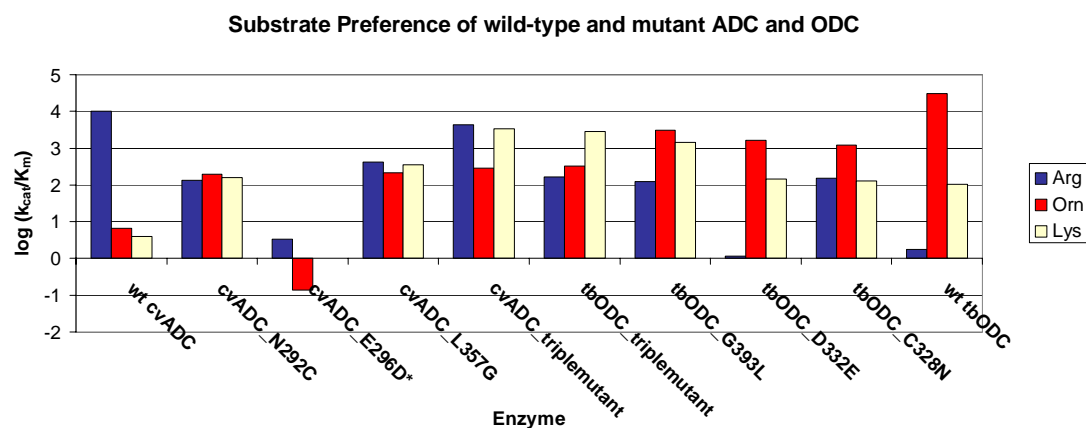
Inhibition constants of substrate analogs and products for *cv*ADC and *tb*ODC. Assays were performed using the coupled-enzyme system. All measurements are the mean (± standard error) of at least two independent trials.



Substrate	cvADC								
	wild-type			N292C			E296D		
	$k_{cat}$	$K_m$	$k_{cat}/K_m$	$k_{cat}$	$K_m$	$k_{cat}/K_m$	$k_{cat}$	$K_m$	$k_{cat}/K_m$
L-Arg	15 $\pm$ 0.3	3.5 $\pm$ 0.2	4.2 x 10 <sup>3</sup>	5.8 $\pm$ 1.0	42.9 $\pm$ 12.5	1.4X 10 <sup>2</sup>	0.14 $\pm$ 0.00 4	42 $\pm$ 2.7	3.3
L-Orn	0.9 $\pm$ 0.2	136 $\pm$ 44	7.4	0.72 $\pm$ 0.13	3.9 $\pm$ 2.2	1.8 X 10 <sup>2</sup>	0.02 $\pm$ 0.00 3	145 $\pm$ 30	0.14
L-Lys	0.4 5 $\pm$ 0.0 8	115 $\pm$ 27	3.9	0.87 $\pm$ 0.27	5.58 $\pm$ 3.68	1.56 X 10 <sup>2</sup>	no detectable activity		
	L357G			triple-mutant					
	$k_{cat}$	$K_m$	$k_{cat}/K_m$	$k_{cat}$	$K_m$	$k_{cat}/K_m$			
	L-Arg	6.08 $\pm$ 0.38	14. 26 $\pm$ 2.4 7	4.26 x 10 <sup>3</sup>	0.72 $\pm$ 0.15	0.16 $\pm$ 0.05	4.28 x 10 <sup>3</sup>		
	L-Orn	1.30 $\pm$ 0.21	6.0 2 $\pm$ 3.5 4	2.15 X 10 <sup>2</sup>	0.63 $\pm$ 0.24	2.19 $\pm$ 0.92	2.87 x 10 <sup>2</sup>		
	L-Lys	0.97 $\pm$ 0.17	2.7 0 $\pm$ 1.6 6	3.58 X 10 <sup>2</sup>	0.58 $\pm$ 0.12	0.18 $\pm$ 0.05	3.33 X 10 <sup>3</sup>		
	tbODC								
	wild-type			C328N			D332E		
	$k_{cat}$	$K_m$	$k_{cat}/K_m$	$k_{cat}$	$K_m$	$k_{cat}/K_m$	$k_{cat}$	$K_m$	$k_{cat}/K_m$
L-Arg	0.03	14.0	1.79	0.89 $\pm$ 0.25	6.0 $\pm$ 2.4	1.5 x 10 <sup>2</sup>	0.04 $\pm$ 0.00 5 <sup>a</sup>	34 $\pm$ 10 <sup>a</sup>	1.2
L-Orn	12	0.4	3.0 X 10 <sup>4</sup>	3.01 $\pm$ 0.05	0.25 $\pm$ 0.03	1.23 X 10 <sup>4</sup>	10 $\pm$ 0.6	6.1 $\pm$ 1.2	1.6 X 10 <sup>3</sup>
L-Lys	2.2	21	1.05 X 10 <sup>2</sup>	0.72 $\pm$ 0.12	5.72 $\pm$ 2.54	1.27 X 10 <sup>2</sup>	0.98 $\pm$ 0.19	6.63 $\pm$ 2.45	1.48 X 10 <sup>2</sup>
	G393L			triple-mutant					
	$k_{cat}$	$K_m$	$k_{cat}/K_m$	$k_{cat}$	$K_m$	$k_{cat}/K_m$			
	L-Arg	0.49 $\pm$ 0.10	3.94 $\pm$ 1.37	1.24 X 10 <sup>2</sup>	0.68 $\pm$ 0.11	4.11 $\pm$ 2.10	1.64 X 10 <sup>2</sup>		
	L-Orn	0.37 $\pm$ 0.03	0.12 $\pm$ 0.05	3.07 X 10 <sup>3</sup>	0.45 $\pm$ 0.06	1.39 $\pm$ 0.77	3.24 X 10 <sup>2</sup>		
	L-Lys	0.30 $\pm$ 0.03	0.2 $\pm$ 0.05	1.43 X 10 <sup>3</sup>	0.35 $\pm$ 0.04	0.12 $\pm$ 0.05	2.87 X 10 <sup>3</sup>		

**TABLE VI**

Steady-state kinetic parameters of wild-type and mutant cvADC and tbODC enzymes with L-arginine, L-ornithine and L-lysine. All measurements are the mean of at least two independent trials ( $\pm$  standard error). Kinetic assays were performed using the spectrophotometric method that measures the oxidation of NADH in a coupled assay. <sup>a</sup>Data for these enzymes were also collected by HPLC analysis of agmatine formation using AccQ\_tag labeling as described in the Experimental Procedures, and similar results were obtained. <sup>b</sup>Data were taken from (3, 4).



**FIGURE 4-6**

Comparison of activity of wild-type and mutant *cvADC* and *tbODC* for L-arginine, L-ornithine and L-lysine. The log ( $k_{cat}/K_m$ ) value for L-arginine (blue), L-ornithine (red) or L-lysine (beige) are plotted (from left to right) for wild-type *cvADC*, *cvADC* N292C, *cvADC* E296D, *cvADC* L357G, *cvADC* N292C/E296D/L357G (triple-mutant), *tbODC* C328N/D332E/G393L (triple-mutant), *tbODC* G393L, *tbODC* D332E and wild-type *tbODC*. Data are displayed in log-scale. \**cvADC* E296D exhibited no detectable activity towards L-lysine.

## CHAPTER FIVE

### LOOP DYNAMICS IN GROUP IV FOLD ENZYMES

#### Introduction

The importance of protein dynamics in enzyme function has been known for some time (117). Kern and colleagues have characterized the dynamics of the prolyl cis-trans isomerase cyclophilin A (CypA) (118). The authors of this study proposed that the intrinsic dynamics of an enzyme form the basis of catalytic rate enhancements. One of the most well-studied enzymes, with respect to conformational dynamics, is dihydrofolate reductase (DHFR), which catalyzes the reduction of dihydrofolate to tetrahydrofolate (119). The dynamics of DHFR, and its relationship to the catalytic cycle, has been studied by fluorescence spectroscopy, nuclear magnetic resonance and molecular dynamics simulation. Of particular interest in the DHFR investigations is the motion of a Met20 loop which undergoes conformational changes on the timescale of product-release, the rate-limiting step in the catalytic cycle. A similar phenomenon is observed in the enzyme triose phosphate isomerase (TIM). Previous structural data led to the belief that loop dynamics had an intimate role in the catalytic cycle (120). Specifically, the loop was thought to close down upon the active site upon substrate binding and open prior to product release. The kinetics of the active site loop 6 in TIM have been measured by NMR and fluorescence spectroscopy techniques and in each case the loop's movements were found to occur on the same timescale as product release (121-123).

Previous studies on ODC have shown that a protease-sensitive loop played a role in positioning active site residues (124). Subsequent structural studies indicated that this loop, termed the Lys-169 loop, forms interactions with active site residues and, therefore, mediates the efficient decarboxylation of substrate

(31). Specifically, in the structure of a K294A mutant of *tb*ODC, Leu-166 is positioned within 4.5 Å of D-ornithine. This residue is predicted to form interactions with the  $\alpha$ -carboxylate group of L-ornithine in the gem-diamine species but not the Schiff base complex with PLP. Furthermore, the reduced catalytic efficiency of the K294A *tb*ODC mutant was thought to be mediated by changes in the stability of the Lys-169 loop, which is then propagated to the active site loop that contains the catalytically relevant residues Asp-361 and Cys-360.

The structures of *cv*ADC presented in Chapter III, in comparison with the *h*ODC structure (Figure 3-12 and 3-13) suggests that the Lys-169 loop (Lys-148 in *cv*ADC) plays a similar role in Group IV decarboxylases as loop 6 in TIM. As mentioned in Chapter III, both the Lys-169 loop in Group IV decarboxylases and loop 6 in TIM exist in the same topological position within the b/a barrel, indicating that a catalytic loop is a conserved structural feature of TIM barrel proteins. The structure I presented in Chapter III leads to the testable hypothesis that the function of this loop is to control the rate of product release, which is known to be the rate limiting step in the catalytic cycle (34). Measuring loop dynamics via fluorescence spectroscopy techniques in stopped-flow apparatus will help establish role of this loop in catalysis (125).

I have mutated two residues (Val-143 and Thr-142) on the *cv*ADC Lys-148 loop to Trp (Figure 5-1 A and B). Ideally, the fluorescent property of a Trp on the loop would be correlated to the dynamics of loop movement upon the addition of ligands.

## Materials

The Infinity™ carbon dioxide detection kit was purchased from Thermo Scientific. The CO<sub>2</sub> kit was purchased from Thermo Electron. All other chemicals and reagents were purchased from Sigma.

## Experimental Methods

*Site-directed mutagenesis* Mutants human ODC K161W, hODC A162W, hODC V163W, hODC C164W, *cvADC* A141W, *cvADC* A142W, *cvADC* V143W, *cvADC* Q144W and *cvADC* L145W were produced using the QuikChange™ site-directed mutagenesis kit from Stratagene (La Jolla, CA).

*Production and purification of mutant and wild-type cvADC* Mutant and wild-type *cvADC* enzymes were produced exactly as described in Chapter II. All enzymes were produced as His<sub>6</sub>-tagged proteins and were purified as described in Chapter II.

*Steady-state analysis of wild-type and mutant enzymes* Steady-state parameters ( $k_{\text{cat}}$  and  $K_{\text{m}}$ ) were obtained using the CO2 kit from Thermo Electron Corporation (See Chapter II). Enzyme was assayed in buffer (20 mM sodium phosphate, 2 mM reduced glutathione) at concentration of 60 nM. Substrate (L-arginine) concentrations ranged from 0.1 to 50 mM. All assays were conducted at 37° C.

*Fluorescence spectra of wild-type and mutant CVADC enzymes* The fluorescence spectra of wild-type *cvADC* (5 – 50  $\mu\text{M}$ ), *cvADC* A142W (5 – 50  $\mu\text{M}$ ) and *cvADC* V143W (5 – 50  $\mu\text{M}$ ) were collected on a Photon Technology International fluorescence spectrometer. All spectra were collected at 25° C.

## Results

*Steady-state analysis of wild-type and mutant enzymes* The substrate dependence of wild-type *cvADC*, *cvADC* A142W and *cvADC* V143W are shown in Figure 5-2. Unlike assays presented in previous chapters, I did not add excess PLP to the reaction conditions. This may have contributed to lower overall activity of the enzymes in the phosphate buffer. The activity of the Trp *cvADC* mutants with L-arginine is not severely attenuated compared to wild-type *cvADC*,

suggesting that the substitution on the active site loop does not significantly alter the catalytic mechanism of *cvADC*.

*Fluorescence spectra of wild-type and mutant cvADC enzymes* The fluorescence spectra of wild-type *cvADC*, *cvADC* V143W and *cvADC* A142W were measured over three excitation/emission regimes. Enzymes were excited at 280 nm and fluorescence was observed from 290-540 nm. A second regime used 330 nm excitation with emission followed from 340-540 nm. Finally, 420 nm light was used to excite the enzymes and emission was followed from 430-540 nm. Preliminary fluorescence spectra were acquired under all three regimes. A difference in fluorescence intensity was observed upon agmatine-binding for both *cvADC* A142W and *cvADC* V143W. However, this initial observation was not repeated in further investigations with agmatine or either D- or L-arginine.

## Discussion

The preliminary, steady-state fluorescence measurements mentioned above have yet to be validated and robustly correlated with the nature of enzyme-ligand interactions. Thus, much more investigation is necessary before I can draw meaningful conclusions regarding the role of the active site loop in the catalytic cycle of Group IV decarboxylases. Nonetheless, the plethora of data from previous studies on TIM and other enzymes that utilize loops in their catalytic cycles suggests that the active site loop in Group IV decarboxylases is intimately linked to the reaction cycle. Furthermore, previous kinetic and structural data on ODC has established the importance of this loop in this enzyme's function (31, 121). The current hypothesis holds that it is the opening of this loop (i.e. away from the active site) that facilitates product release. It has been established that product-release is the rate-limiting step in the reaction cycle and, therefore, it is likely that the motion of the loop ultimately dictates the rate of release of the decarboxylated product. Conducting future experiments harnessing a stopped-

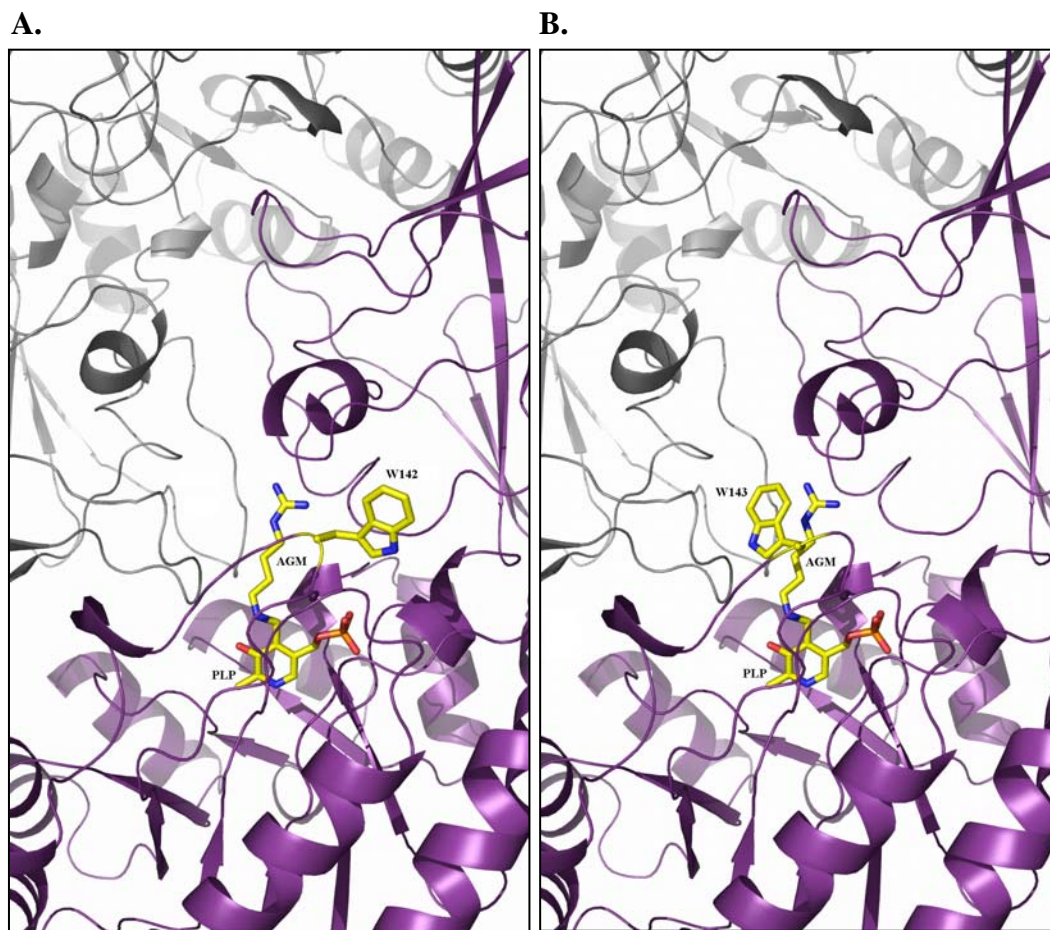
flow apparatus will aid in validating this hypothesis (125). Further testing of other loop mutants with a Trp substituted at different sites on the loop is necessary to find a fluorescence signal that is both robust and titratable with respect to enzyme-ligand interactions.

Some of the possible results from stopped-flow measurements of loop dynamics via fluorescence quenching may be considered. The ideal, simple case would yield measurements of loop dynamics on the time-scale of product-release. This would indicate that loop movement is the dominant rate-limiting step. Because the internal and external aldimine species are thought to be energetically equivalent, at least for favorable substrates (3), the breakdown of the PLP-product Schiff's base is likely to be faster than the overall turnover rate (34). On the other hand, if the rate of loop movement is found to occur at a rate much higher than product-release then at least two conclusions may be drawn. It may be that the active site loop's movement is completely decoupled from the catalytic cycle. Alternatively, the loop's dynamics may have a more intimate relationship with catalysis than product-release. Gutteridge et al. have catalogued the significance of dynamics, including loop movement, on the catalytic abilities of enzymes (126). In certain enzymes (e.g. LDH, DHFR, and IMPDH) a loop adjacent to the active site does not merely close in on the ES complex but actively promotes chemistry by introducing catalytic residues into the active site. Though such a notion for Group IV decarboxylases is mere speculation, it is interesting to consider if residues on the active site loop could play a role in the promotion of decarboxylation of substrates in the active site. For Group IV decarboxylases it may be that residues on the active site loop contact substrate (31) at a critical point on the reaction pathway, possibly enhancing the rate of decarboxylation. A pathway leading from substrate binding residues (e.g. Asp-361) to residues on the active site loop (K169) has been postulated. Furthermore, it has been shown that

unfavorable dynamics of the active site loop yields a higher propensity for non-productive binding of substrate to the active site (31).

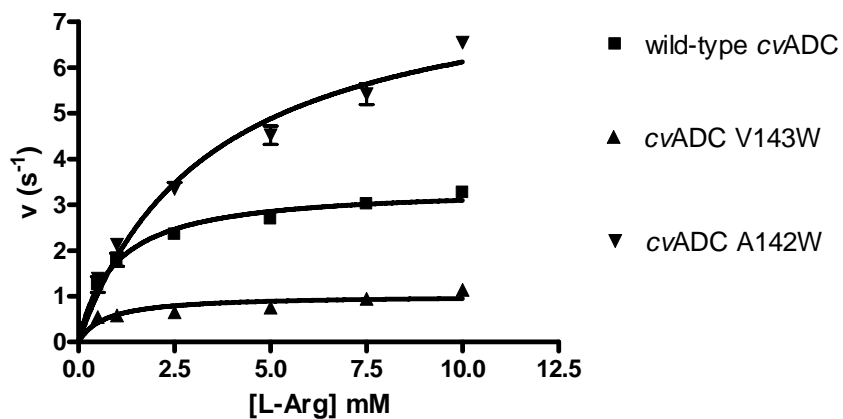
Finally, an even more intriguing case to consider is the role of dynamics in controlling substrate specificity. It has been stated for some time that the basis of specificity is due to favorable enzyme-transition-state ( $ES^\ddagger$ ) binding with cognate substrate versus non-cognate substrates (117). Recent investigations on the fidelity of DNA polymerases and DNA methyltransferases have provided experimental evidence implicating the crucial role of dynamics in catalyzing reactions for the correct substrates (127, 128). One of the authors of these studies has suggested enzymes have evolved such that incorrect substrates actively disorient catalytic residues, thereby lowering overall proficiency (127). With an awareness of a pathway of residues leading from substrate binding residues (Cys-360, Asp-361) to active site loop residues, one can imagine a case in which non-cognate substrates interact unfavorably to active site residues and, consequently, an altered dynamic property of the loop. The altered dynamic may be a case in which the loop does not fully close down on the substrate and promote efficient catalysis. Establishing the role of dynamics in specificity of Group IV decarboxylases is non-trivial and will involve much more kinetic and structural data than currently available, creating new avenues of research for this increasingly valued class of enzymes.





**FIGURE 5-1**

Models of the mutants *cvADC* V143W (**A**) and *cvADC* A142W (**B**) used in the measurements of loop dynamics. The putative position of the Trp substitutions and the PLP-arginine complex is shown in yellow sticks.



**FIGURE 5-2**

Substrate dependence of wild-type and mutant *cvADCs* used for fluorescence assays. Data points are the average of at least two experiments and error bars represent standard error of the mean. Data was fit by non-linear regression to the Michaelis-Menten equation and the fitted parameters are: wild-type *cvADC* ( $K_m = 0.94 \pm 0.14$  mM,  $k_{cat} = 3.39 \pm 0.12$  s<sup>-1</sup>), *cvADC* V143W ( $K_m = 0.72 \pm 0.27$  mM,  $k_{cat} = 1.0 \pm 0.08$  s<sup>-1</sup>) and *cvADC* A142W ( $K_m = 3.41 \pm 0.63$  mM,  $k_{cat} = 8.21 \pm 0.63$  s<sup>-1</sup>). Data was analyzed with GraphPad Prism version 4 using the Michaelis-Menten equation.

## CHAPTER SIX

### CONCLUSIONS

#### **Rapid evolution in Group IV PLP-dependent decarboxylases**

The previous substrate specificity paradigm for Group IV PLP-dependent decarboxylases held that evolving activity towards a new substrate (e.g. ODC to ADC) required a considerable degree of evolution. This notion seems apparent by observation of sequence alignments and phylogenetic analyses of Group IV decarboxylases of different specificities. It was believed that for an ODC to evolve into an ADC or DAPDC, or vice versa, many amino acid substitutions would need to occur. The data I presented in Chapter II demonstrates that this paradigm is in need of updating. The *cvADC* enzyme shares 40% identity with eukaryotic ODC and only about 20% identity with various ADCs (e.g. *E. coli*, *Y. pestis*). Thus, an observer of the Group IV decarboxylase literature would be inclined to first classify this enzyme as an ODC. The kinetic data I collected clearly demonstrates that *cvADC* is a *bona fide* ADC and, in fact, has subsequently been reclassified as such. The conclusion that can be drawn from this study is that, though rare, it is possible that a limited number of substitutions would be necessary to evolve a new specificity into an existing Group IV decarboxylase construct. The precise number of substitutions necessary to facilitate a change in preference remains to be determined.

The mutagenesis studies I presented in Chapter IV add some credence to the hypothesis that limited substitutions can produce a change in preference. It was certainly clear from the mutant data that single and even triple-mutant constructs were incapable of recapitulating the specificity profile of wild-type *cvADC* or *tbODC*. Nonetheless, I did demonstrate that residues on the specificity element (Cys-328 and Asp-332) as well as at least one residue believed to

position the specificity element (Gly-393) play important roles in substrate selection. With the advent of genome sequencing and the subsequent deposition of numerous, putative Group IV decarboxylase sequences, it is possible to harness bioinformatics techniques to determine more thoroughly the positions in the sequences of these decarboxylases that play dominant roles in the substrate specificity.

### **Structural Basis of Specificity of Group IV PLP-dependent decarboxylases**

The determination of the crystal structure of *cvADC* has helped illuminate the structural basis of specificity of Group IV PLP-dependent decarboxylases. Previous kinetic studies demonstrated that eukaryotic ODC was able to discriminate between substrates by chain-length (116). Based on the structure of *T. brucei* ODC complexed to putrescine, Grishin et al. proposed that a possible basis of specificity was a molecular ruler which measures the distance between the Schiff base nitrogen and the Asp-361/Asp-332 substrate-binding pair (28). Comparison of structures of ODC, DAPDC and *cvADC* demonstrate quite nicely that the  $3_{10}$  helix at the distal end of the active site is a ruler of amino acid side chain length (Figure 3-20). Further conformation of the molecular ruler hypothesis is supported by the recent structure determination of a dual-specificity L-ornithine/lysine decarboxylase (110). Kinetic data from  $3_{10}$  helix residue mutants demonstrate that the identity of residues on this specificity element is important for substrate selection, though as-yet-unidentified positions seem to be necessary to confer a particular specificity profile to a given Group IV decarboxylase. The considerable change in amino acid side-chain bulk at position 393 in ODC (357 in *cvADC*) may play some role in the positioning of the specificity element. It may be worthwhile to determine the structure of either *cvADC* L357G or *tbODC* G393L to ascertain the importance of this substitution in positioning of the helix.

### **A (more) complete reaction cycle**

The minimal scheme for decarboxylation of amino acids by Group IV decarboxylases is shown in Figure 1-9. A further benefit of the structure determination of *cvADC* is a more complete picture of the reaction cycle of these decarboxylases. Along with the numerous other structures of ODCs solved previously in the Phillips lab, it is now possible to follow the reaction with respect both to chemical and structural phenomenon. The resting state of the enzyme (internal aldimine) is captured by the structure of the ligand-free human ODC structure (PDB ID 1D7K) in which the active site loop allows for substrate entrance and PLP is bound to K69. A K294A *tbODC* mutant offers a view of the gem diamine intermediate (PDB ID 1SZR) which forms prior to the Schiff base structure with substrate, in this case D-ornithine from (PDB ID 1NJJ). The PLP-product complex (PDB ID 1F3T) remains intact as it awaits the movement of the active site loop (PDB ID 2NVA). Once the active site loop has flipped away from the active site, product is released and the resting state of the enzyme is regenerated.

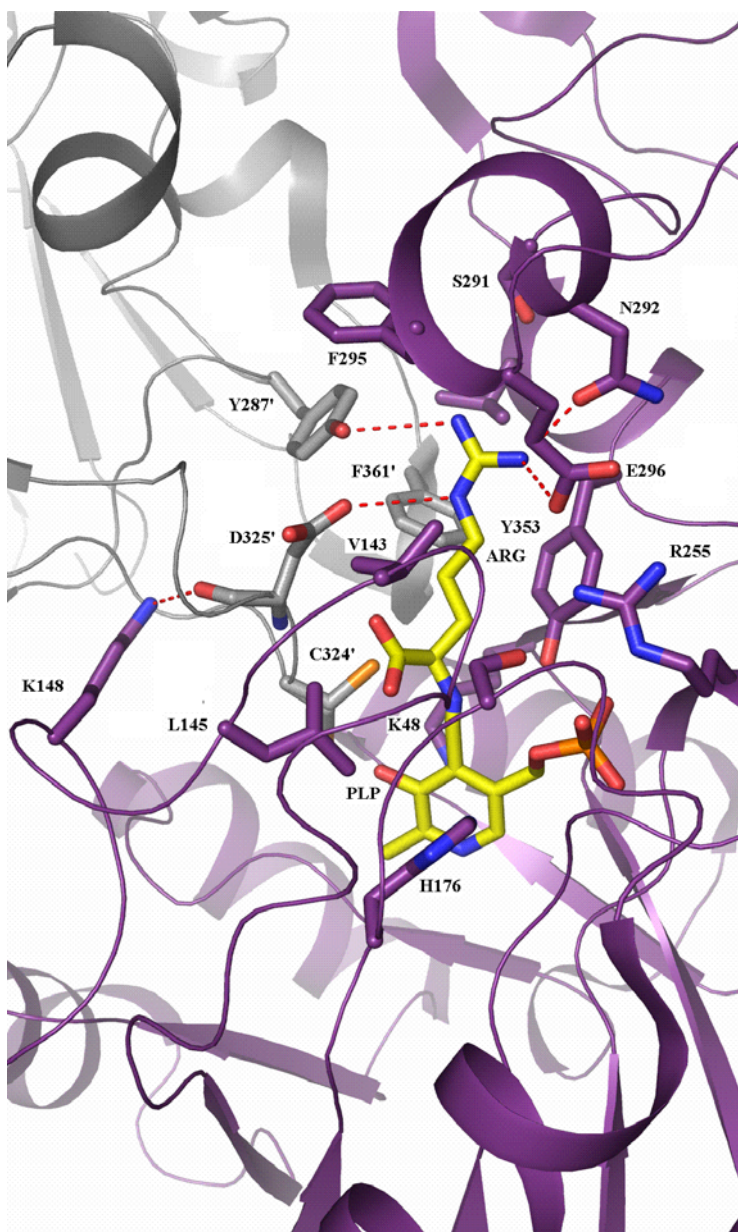
I have created a molecular movie of the chemical and structural transformations occurring in the reaction cycle of eukaryotic ODC. In this movie the rate of loop closure and opening is coupled to the decarboxylation reaction. With a more complete picture of the reaction cycle, it is worthwhile to speculate on the means by which substrate binding enhances the rate of chemical catalysis. In Figure 6-1 the active site of *cvADC* is shown with a molecule of L-arginine modeled in place of the agmatine molecule. I placed the carboxyl group of based on the position of the experimentally-derived position of the agmatine molecule in my structure. One can glean from this figure a pathway leading from the substrate-binding residue Asp-325 (from the opposite monomer) to the active site loop. The guanidine nitrogen of agmatine forms H-bond interactions with Asp-

325. Lys-148 from the active site loop form H-bond interactions with the backbone carbonyl of Asp-325. From the model of L-arginine bound to *cvADC* one can see that the active site loop residues Val-43 and Leu-145 form unfavorable van der Waals contacts with the carboxylate of the substrate. Previous studies have shown that a similar unfavorable interaction between the carboxylate of L-ornithine and ODC is necessary for the enhancement of the rate of decarboxylation (29). Thus, one can surmise that the binding energy gained by the interaction between Asp-325 and substrate is propagated to the “chemical” end of the active site via Lys-148. It seems, then, that the binding energy is used to stabilize the active site loop such that the hydrophobic residues, Val-143 and Leu-145, promote decarboxylation by creating an unfavorable environment for the carboxylate of the substrate.

Precedent for the involvement of an active site loop in the reaction pathway of an enzyme is plentiful. Previous data from investigations of TIM and other enzymes that utilize flexible loops in their catalytic cycles suggests that an active site loop in the Group IV decarboxylases is intimately linked to the reaction cycle. Furthermore, previous kinetic and structural data on ODC has established the importance of this loop in this enzyme’s function (31, 121). The current hypothesis holds that it is the opening of this loop (i.e. swinging away from the active site) that facilitates product release. It has been established that product-release is the rate-limiting step in the reaction cycle and, therefore, it is likely that the motion of the loop ultimately dictates the rate of release of the decarboxylated product.

In light of the importance of the active site loop for catalysis in Group IV decarboxylases, it is interesting to consider the role of this loop in substrate specificity. It has been stated for some time that the basis of specificity is due to favorable enzyme-transition-state ( $ES^\ddagger$ ) binding with cognate substrate versus non-cognate substrates (117). Recent investigations on the fidelity of DNA

polymerases and DNA methyltransferases have provided experimental evidence implicating the crucial role of dynamics in catalyzing reactions for the correct substrates (127, 128). One of the authors of these studies has suggested enzymes have evolved such that incorrect substrates actively disorient catalytic residues, thereby lowering overall proficiency (127). With an awareness of a pathway of residues leading from substrate binding residues (i.e. Asp-325) to active site loop residues, one can imagine a case in which non-cognate substrates interact unfavorably to active site residues and, consequently, an altered dynamic property of the loop. The altered dynamic may be a case in which the loop does not fully close down on the substrate and promote efficient catalysis. In the case of *cvADC*, it may be that the amino groups of L-ornithine and L-lysine simply do not form H-bond interactions of the same magnitude, energetically, as with the guanidine nitrogen of L-arginine. It may be that further substrate binding residues on the specificity element of *cvADC* orient L-arginine such that the guanidine nitrogen forms the optimal interaction with Asp-325, realizing the full amount of energy required to stabilize the active site loop and promote decarboxylation. It may be that the non-cognate substrates, L-ornithine and L-lysine, do not form this optimal interaction and, as a consequence, spend much more time in the active site of *cvADC* in a non-productive conformation, with respect to decarboxylation.



**FIGURE 6-1**

Model of *cvADC* bound to L-arginine. A model of L-arginine in the active site of *cvADC* was built based on the *cvADC* structure bound to agmatine. The model illustrates the idea that the carboxylate group of the substrate is placed in a hydrophobic pocket with unfavorable interactions occurring between the carboxylate and Leu-145 and Val-143. A pathway leading from the active site to the active site loop is mediated through Lys-148.



## REFERENCES

- (1) Taylor, J. S., Reid, T. S., Terry, K. L., Casey, P. J., and Beese, L. S. (2003) Structure of mammalian protein geranylgeranyltransferase type-I. *Embo J* 22, 5963-74.
- (2) Baumann, S., Sander, A., Gurnon, J. R., Yanai-Balser, G. M., Van Etten, J. L., and Piotrowski, M. (2007) Chlorella viruses contain genes encoding a complete polyamine biosynthetic pathway. *Virology* 360, 209-17.
- (3) 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.
- (4) 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 decarboxylase. *Biochemistry* 36, 4558-4567.
- (5) Tabor, C., and Tabor, H. (1984) Polyamines. *Ann Rev Biochem* 53, 749-790.
- (6) Casero, R. A., Jr., and Marton, L. J. (2007) Targeting polyamine metabolism and function in cancer and other hyperproliferative diseases. *Nat Rev Drug Discov* 6, 373-90.
- (7) Bacchi, C. J., and Yarlett, N. (2002) Polyamine metabolism as chemotherapeutic target in protozoan parasites. *Mini Rev Med Chem* 2, 553-63.
- (8) Feuerstein, B. G., Williams, L. D., Basu, H. S., and Marton, L. J. (1991) Implications and concepts of polyamine-nucleic acid interactions. *J Cell Biochem* 46, 37-47.
- (9) Hobbs, C. A. G., S.K. (2006) *Polyamine Cell Signaling: Physiology, Pharmacology, and Cancer Research*, Humana Press, Totowa.
- (10) Kurata, H. T., Marton, L. J., and Nichols, C. G. (2006) The polyamine binding site in inward rectifier K<sup>+</sup> channels. *J Gen Physiol* 127, 467-80.
- (11) Park, M. H. (2006) The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). *J Biochem (Tokyo)* 139, 161-9.
- (12) Meijer, A. J., Lamers, W. H., and Chamuleau, R. A. (1990) Nitrogen metabolism and ornithine cycle function. *Physiol Rev* 70, 701-48.
- (13) Yu, H., Yoo, P. K., Aguirre, C. C., Tsoa, R. W., Kern, R. M., Grody, W. W., Cederbaum, S. D., and Iyer, R. K. (2003) Widespread expression of arginase I in mouse tissues. Biochemical and physiological implications. *J Histochem Cytochem* 51, 1151-60.
- (14) Klein, R. D., Geary, T. G., Gibson, A. S., Favreau, M. A., Winterrowd, C. A., Upton, S. J., Keithly, J. S., Zhu, G., Malmberg, R. L., Martinez, M. P.,

- and Yarlett, N. (1999) Reconstitution of a bacterial/plant polyamine biosynthesis pathway in *Saccharomyces cerevisiae*. *Microbiology* 145 ( Pt 2), 301-7.
- (15) Mistry, S. K., Burwell, T. J., Chambers, R. M., Rudolph-Owen, L., Spaltmann, F., Cook, W. J., and Morris, S. M., Jr. (2002) Cloning of human agmatinase. An alternate path for polyamine synthesis induced in liver by hepatitis B virus. *Am J Physiol Gastrointest Liver Physiol* 282, G375-81.
  - (16) Coleman, C. S., Hu, G., and Pegg, A. E. (2004) Putrescine biosynthesis in mammalian tissues. *Biochem J* 379, 849-55.
  - (17) Janowitz, T., Kneifel, H., and Piotrowski, M. (2003) Identification and characterization of plant agmatine iminohydrolase, the last missing link in polyamine biosynthesis of plants. *FEBS Lett* 544, 258-61.
  - (18) Mercenier, A., Simon, J. P., Haas, D., and Stalon, V. (1980) Catabolism of L-arginine by *Pseudomonas aeruginosa*. *J Gen Microbiol* 116, 381-9.
  - (19) Van Etten, J. L. (2003) Unusual Life Style of Giant Chlorella Viruses. *Annu. Rev. Genet.* 37, 153-95.
  - (20) 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.
  - (21) Osterman, A., Kinch, L., Grishin, N., and Phillips, M. (1995) Acidic residues important for substrate binding and cofactor reactivity in eukaryotic ornithine decarboxylase identified by alanine scanning mutagenesis. *J. Biol. Chem.* 270, 11797-11802.
  - (22) Percudani, R., and Peracchi, A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep* 4, 850-4.
  - (23) 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.
  - (24) Grishin, N., Phillips, M., and Goldsmith, E. (1995) Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Science* 4, 1291-1304.
  - (25) Momany, C., Ernst, S., Ghosh, R., Chang, N., and Hackert, M. (1995) Crytsallographic structure of a PLP-dependent ornithine decarboxylase from *Lactobacillus* 30a to 3.0Å resolution. *J. Mol. Biol.* 6, 643-655.
  - (26) 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.
  - (27) 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.
- (28) 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  $\alpha$ -difluoromethylornithine. *Biochemistry* 38, 15174-15184.
  - (29) Jackson, L., Brooks, H., Myers, D., and Phillips, M. (2003) Ornithine decarboxylase promotes catalysis by binding the carboxylate in a buried pocket containing Phe-397. *Biochemistry in press*.
  - (30) 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. submitted*.
  - (31) 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.
  - (32) Jackson, L. K., Brooks, H. B., Osterman, A. L., Goldsmith, E. J., and Phillips, M. A. (2000) Altering the Reaction Specificity of Eukaryotic Ornithine Decarboxylase. *Biochemistry* 39, 11247-11257.
  - (33) Tabor, C. W. T., H. (1989) *Ornithine decarboxylase in microorganisms*, Pergamon Press Inc., New York.
  - (34) 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.
  - (35) Feng, Z., and Barletta, R. G. (2003) Roles of *Mycobacterium smegmatis* D-alanine:D-alanine ligase and D-alanine racemase in the mechanisms of action of and resistance to the peptidoglycan inhibitor D-cycloserine. *Antimicrob Agents Chemother* 47, 283-91.
  - (36) Wang, E., and Walsh, C. (1978) Suicide substrates for the alanine racemase of *Escherichia coli* B. *Biochemistry* 17, 1313-21.
  - (37) 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.
  - (38) Counts, K. G., Wong, I., and Oliveira, M. A. (2007) Investigating the geminal diamine intermediate of *Yersinia pestis* arginine decarboxylase with substrate, product, and inhibitors using single wavelength stopped-flow spectroscopy. *Biochemistry* 46, 379-86.
  - (39) Piacenza, L., Peluffo, G., and Radi, R. (2001) L-arginine-dependent suppression of apoptosis in *Trypanosoma cruzi*: contribution of the nitric oxide and polyamine pathways. *Proc Natl Acad Sci U S A* 98, 7301-6.

- (40) Amadasi, A., Bertoldi, M., Contestabile, R., Bettati, S., Cellini, B., di Salvo, M. L., Borri-Voltattorni, C., Bossa, F., and Mozzarelli, A. (2007) Pyridoxal 5'-phosphate enzymes as targets for therapeutic agents. *Curr Med Chem* 14, 1291-324.
- (41) Brown, M. R., Jeffrey, S.W. (1992) Biochemical composition of microalgae from the green algal classes Chlorophyceae and Prasinophyceae. 1. Amino acids, sugars and pigments. *Journal of Experimental Marine Biology and Ecology* 161, 91-113.
- (42) Coleman, C., Stanley, B., and Pegg, A. (1993) Effect of mutations at active site residues on the activity of ornithine decarboxylase and its inhibition by active site-directed irreversible inhibitors. *J Biol Chem* 268, 24572-24579.
- (43) Takatsuka, Y., Yamaguchi, Y., Ono, M., and Kamio, Y. (2000) Gene cloning and molecular characterization of lysine decarboxylase from *Selenomonas ruminantium* delineate its evolutionary relationship to ornithine decarboxylases from eukaryotes. *J Bacteriol* 182, 6732-41.
- (44) Page, R. D. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12, 357-8.
- (45) Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Espresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res* 34, W604-8.
- (46) Osterman, A., Grishin, N., Kinch, L., and Phillips, M. (1994) Formation of functional cross-species heterodimers of ornithine decarboxylase. *Biochemistry* 33, 13662-13667.
- (47) Pegg, A., and McGill, S. (1979) Decarboxylation of ornithine and lysine in rat tissues. *Biochimica et Biophysica Acta*. 568, 416-427.
- (48) Pegg, A. E., Wechter, R., Poulin, R., Woster, P. M., and Coward, J. K. (1989) Effect of S-adenosyl-1,12-diamino-3-thio-9-azadodecane, a multisubstrate adduct inhibitor of spermine synthase, on polyamine metabolism in mammalian cells. *Biochemistry* 28, 8446-53.
- (49) Burns, D. H., and Aberhart, D. J. (1988) A general coupled spectrophotometric assay for decarboxylases. *Anal Biochem* 171, 339-45.
- (50) Choi, S., and Churchich, J. (1986) Glutamate decarboxylase side reactions catalyzed by the enzyme. *Eur. J. Biochem.* 160, 515-520.
- (51) Metcalf, B., Bey, P., Danzin, C., Jung, M., Casara, P., and Vever, J. (1978) Catalytic irreversible inhibition of mammalian ornithine decarboxylase by substrate and product analogues. *J. Amer. Chem. Soc.* 100, 2551-2553.
- (52) Poulin, R., Lu, L., Ackermann, B., Bey, P., and Pegg, A. (1992) Mechanism of the Irreversible Inactivation of Mouse Ornithine

- Decarboxylase by  $\alpha$ -difluoromethylornithine: Characterization of Sequences at the Inhibitor and Coenzyme Binding Sites. *J Biol Chem* 267, 150-158.
- (53) Bitonti, A. J., Casara, P. J., McCann, P. P., and Bey, P. (1987) Catalytic irreversible inhibition of bacterial and plant arginine decarboxylase activities by novel substrate and product analogues. *Biochem J* 242, 69-74.
  - (54) Coligan, J., Dunn, B., Ploegh, H. L., Speicher, D., and Wingfield, P. (1995) *Current protocols in protein science.*, John Wiley & Sons Inc.
  - (55) Kaiser, A., Vollmert, M., Tholl, D., Graves, M. V., Gurnon, J. R., Xing, W., Lisec, A. D., Nickerson, K. W., and Van Etten, J. L. (1999) Chlorella virus PBCV-1 encodes a functional homospermidine synthase. *Virology* 263, 254-62.
  - (56) Van Etten, J. L., and Meints, R. H. (1999) Giant viruses infecting algae. *Annu Rev Microbiol* 53, 447-94.
  - (57) Tome, M. E., and Gerner, E. W. (1997) Cellular eukaryotic initiation factor 5A content as a mediator of polyamine effects on growth and apoptosis. *Biol Signals* 6, 150-6.
  - (58) Park, J. H., Wolff, E. C., Folk, J. E., and Park, M. H. (2003) Reversal of the deoxyhypusine synthesis reaction. Generation of spermidine or homospermidine from deoxyhypusine by deoxyhypusine synthase. *J Biol Chem* 278, 32683-91.
  - (59) Grishin, N., and Phillips, M. (1994) The subunit interfaces of oligomeric enzymes are conserved to a similar extent to the overall protein sequences. *Protein Science* 3, 2455-2458.
  - (60) Matthews, B. W. (1968) Solvent content of protein crystals. *J Mol Biol* 33, 491-7.
  - (61) Otwinoski, A., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology* 276, 307-326.
  - (62) Brunger, A., Adams, P., Clre, G., Delano, W., Gros, P., Grosse-Kunstleve, R., Jiang, J., Kuszewski, J., Nilges, N., Pannu, N., Read, R., Rice, L., Simonson, T., and Warren, G. (1998) Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Cryst D* 54, 905-921.
  - (63) Vagin, A., and Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol Crystallogr* 56, 1622-4.
  - (64) Navaza, J. (1994) AMoRe:an automated package for molecular replacement. *Acta Cryst. a* 50, 157-163.
  - (65) Kissinger, C. R., Gehlhaar, D. K., Smith, B. A., and Bouzida, D. (2001) Molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr* 57, 1474-9.

- (66) Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 60, 432-8.
- (67) Dodson, E. (2003) Is it jolly SAD? *Acta Crystallogr D Biol Crystallogr* 59, 1958-65.
- (68) Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* 58, 1772-9.
- (69) Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* 56 ( Pt 8), 965-72.
- (70) Otwinoski, Z. (1991) Maximum likelihood refinement of heavy-atom parameters in isomorphous replacement and anomalous scattering. *Proceedings of the CCP4 Study Weekend, SERC Daresbury Laboratory, Warrington, UK*, 80-86.
- (71) Cowtan, K. D., and Zhang, K. Y. (1999) Density modification for macromolecular phase improvement. *Prog Biophys Mol Biol* 72, 245-70.
- (72) Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol* 374, 229-44.
- (73) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-32.
- (74) Collaborative Computing Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D50*, 760-763.
- (75) 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.
- (76) Shaw, J., Petsko, G., and Ringe, D. (1997) Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-Å resolution. *Biochemistry* 36, 1329-1342.
- (77) Stamper, C., Morollo, A., and Ringe, D. (1998) Reaction of alanine racemase with 1-aminoethylphosphonic acid forms a stable external aldimine. *Biochemistry* 37, 10438-10445.
- (78) Morollo, A., Petsko, G., and Ringe, D. (1999) Structure of a Michaelis complex analogue: propionate binds in the substrate carboxylate site of alanine racemase. *Biochemistry* 38, 3293-3301.
- (79) Watanabe, A., Yoshimura, T., Mikami, B., Hayashi, H., Kagamiyama, H., and Esaki, N. (2002) Reaction mechanism of alanine racemase from *Bacillus stearothermophilus*: x-ray crystallographic studies of the enzyme bound with N-(5'-phosphopyridoxyl)alanine. *J Biol Chem* 277, 19166-72.
- (80) Fenn, T. D., Holyoak, T., Stamper, G. F., and Ringe, D. (2005) Effect of a Y265F mutant on the transamination-based cycloserine inactivation of alanine racemase. *Biochemistry* 44, 5317-27.

- (81) LeMagueres, P., Im, H., Dvorak, A., Strych, U., Benedik, M., and Krause, K. L. (2003) Crystal structure at 1.45 Å resolution of alanine racemase from a pathogenic bacterium, *Pseudomonas aeruginosa*, contains both internal and external aldimine forms. *Biochemistry* 42, 14752-61.
- (82) LeMagueres, P., Im, H., Ebalunode, J., Strych, U., Benedik, M. J., Briggs, J. M., Kohn, H., and Krause, K. L. (2005) The 1.9 Å crystal structure of alanine racemase from *Mycobacterium tuberculosis* contains a conserved entryway into the active site. *Biochemistry* 44, 1471-81.
- (83) Mehta, P. K., and Christen, P. (2000) The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Adv Enzymol Relat Areas Mol Biol* 74, 129-84.
- (84) Graham, D. E., Xu, H., and White, R. H. (2002) *Methanococcus jannaschii* uses a pyruvoyl-dependent arginine decarboxylase in polyamine biosynthesis. *J Biol Chem* 277, 23500-7.
- (85) Tolbert, W. D., Graham, D. E., White, R. H., and Ealick, S. E. (2003) Pyruvoyl-dependent arginine decarboxylase from *Methanococcus jannaschii*: crystal structures of the self-cleaved and S53A proenzyme forms. *Structure* 11, 285-94.
- (86) Recsei, P. A., and Snell, E. E. (1970) Histidine decarboxylase of *Lactobacillus* 30a. VI. Mechanism of action and kinetic properties. *Biochemistry* 9, 1492-7.
- (87) 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* chloroella virus-1 encodes an unusual arginine decarboxylase that is a close homolog of eukaryotic ornithine decarboxylases. *J Biol Chem* 279, 35760-7.
- (88) 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.
- (89) Ray, S. S., Bonanno, J. B., Rajashankar, K. R., Pinho, M. G., He, G., De Lencastre, H., Tomasz, A., and Burley, S. K. (2002) Cocystal structures of diaminopimelate decarboxylase: mechanism, evolution, and inhibition of an antibiotic resistance accessory factor. *Structure* 10, 1499-508.
- (90) Lane, K. T., and Beese, L. S. (2006) Thematic review series: lipid posttranslational modifications. Structural biology of protein farnesyltransferase and geranylgeranyltransferase type I. *J Lipid Res* 47, 681-99.
- (91) Caplin, B. E., Hettich, L. A., and Marshall, M. S. (1994) Substrate characterization of the *Saccharomyces cerevisiae* protein

- farnesyltransferase and type-I protein geranylgeranyltransferase. *Biochim Biophys Acta* 1205, 39-48.
- (92) Yokoyama, K., McGeady, P., and Gelb, M. H. (1995) Mammalian protein geranylgeranyltransferase-I: substrate specificity, kinetic mechanism, metal requirements, and affinity labeling. *Biochemistry* 34, 1344-54.
  - (93) Guo, L., Lim, K. B., Poduje, C. M., Daniel, M., Gunn, J. S., Hackett, M., and Miller, S. I. (1998) Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. *Cell* 95, 189-98.
  - (94) Ahn, V. E., Lo, E. I., Engel, C. K., Chen, L., Hwang, P. M., Kay, L. E., Bishop, R. E., and Prive, G. G. (2004) A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *Embo J* 23, 2931-41.
  - (95) 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.
  - (96) Hedstrom, L., and Gan, L. (2006) IMP dehydrogenase: structural schizophrenia and an unusual base. *Curr Opin Chem Biol* 10, 520-5.
  - (97) 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.
  - (98) 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.
  - (99) Wilks, H., Hart, K., Feeney, R., Dunn, C., Muirhead, H., Chia, W., Barstow, D., Atkinson, T., Clarke, A., and Holbrook, J. (1988) A specific highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. *Science* 242, 1541-1544.
  - (100) Hedstrom, L., Sylagi, L., and Rutter, W. (1992) Converting Trypsin to Chymotrypsin: The role of surface loops. *Science* 255, 1249-1253.
  - (101) Graf, L., Craik, C., Patthy, A., Rocznik, S., Fletterick, R., and Rutter, W. (1987) Selective alteration of substrate specificity by replacement of Asp-189 with Lys in the binding pocket of trypsin. *Biochemistry* 26, 2616-2623.
  - (102) Perona, J., L. H., Rutter, W., and Fletterick, R. (1995) Structural origins of substrate discrimination in trypsin and chymotrypsin. *Biochemistry* 34, 1489-1499.
  - (103) Onuffer, J. J., and Kirsch, J. F. (1995) Redesign of the substrate specificity of *Escherichia coli* aspartate aminotransferase to that of *Escherichia coli* tyrosine aminotransferase by homology modeling and site-directed mutagenesis. *Protein Sci* 4, 1750-7.



- (104) Shaffer, W. A., Luong, T. N., Rothman, S. C., and Kirsch, J. F. (2002) Quantitative chimeric analysis of six specificity determinants that differentiate *Escherichia coli* aspartate from tyrosine aminotransferase. *Protein Sci* 11, 2848-59.
- (105) Toney, M. D. (2005) Reaction specificity in pyridoxal phosphate enzymes. *Arch Biochem Biophys* 433, 279-87.
- (106) Toscano, M. D., Muller, M. M., and Hilvert, D. (2007) Enhancing activity and controlling stereoselectivity in a designed PLP-dependent aldolase. *Angew Chem Int Ed Engl* 46, 4468-70.
- (107) Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 31, 3381-5.
- (108) Peitsch, M. C., Wilkins, M. R., Tonella, L., Sanchez, J. C., Appel, R. D., and Hochstrasser, D. F. (1997) Large-scale protein modelling and integration with the SWISS-PROT and SWISS-2DPAGE databases: the example of *Escherichia coli*. *Electrophoresis* 18, 498-501.
- (109) Christen, M., Hunenberger, P. H., Bakowies, D., Baron, R., Burgi, R., Geerke, D. P., Heinz, T. N., Kastenholz, M. A., Krautler, V., Oostenbrink, C., Peter, C., Trzesniak, D., and van Gunsteren, W. F. (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem* 26, 1719-51.
- (110) Lee, J., Michael, A. J., Martynowski, D., Goldsmith, E. J., and Phillips, M. A. (2007) Phylogenetic diversity and the structural basis of substrate specificity in the beta /alpha -barrel fold basic amino acid decarboxylases. *J Biol Chem*.
- (111) Oue, S., Okamoto, A., Yano, T., and Kagamiyama, H. (1999) Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues. *J. Biol. Chem.* 274, 2344-2349.
- (112) Rajagopalan, P. T., Lutz, S., and Benkovic, S. J. (2002) Coupling interactions of distal residues enhance dihydrofolate reductase catalysis: mutational effects on hydride transfer rates. *Biochemistry* 41, 12618-28.
- (113) Jeffrey, C., Gloss, L., Petsko, G., and Ringe, D. (2000) The role of residues outside the active site: structural basis for function of C191 mutants of *Escherichia coli* aspartate aminotransferase. *Protein Engineering* 13, 105-112.
- (114) Hedstrom, L., Farr-Jones, S., Kettner, C., and Rutter, W. (1994) Converting trypsin to chymotrypsin: ground-state binding does not determine substrate specificity. *Biochemistry* 33, 8764-8769.
- (115) Kraut, D. A., Carroll, K. S., and Herschlag, D. (2003) Challenges in enzyme mechanism and energetics. *Annu Rev Biochem* 72, 517-71.

- (116) Swanson, T., Brooks, H., Osterman, A., O'Leary, M., and Phillips, M. (1998) Carbon-13 isotope effect studies of *Trypanosoma brucei* ornithine decarboxylase. *Biochemistry* 37, 14943-14947.
- (117) Fersht, A. (1984) Enzyme structure and mechanism, in *Enzyme structure and mechanism* (Fersht, A., Ed.) pp 311-342, WH Freeman and Co., New York.
- (118) Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E., and Kern, D. (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438, 117-21.
- (119) Schnell, J. R., Dyson, H. J., and Wright, P. E. (2004) Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct* 33, 119-40.
- (120) Lolis, E., Alber, T., Davenport, R., Rose, D., Hartman, F., and Petsko, G. (1990) Structure of yeast triosephosphate isomerase at 1.9Å resolution. *Biochemistry* 29, 6609-6618.
- (121) Sampson, N. S., and Knowles, J. R. (1992) Segmental movement: definition of the structural requirements for loop closure in catalysis by triosephosphate isomerase. *Biochemistry* 31, 8482-7.
- (122) Kursula, I., Salin, M., Sun, J., Norledge, B. V., Haapalainen, A. M., Sampson, N. S., and Wierenga, R. K. (2004) Understanding protein lids: structural analysis of active hinge mutants in triosephosphate isomerase. *Protein Eng Des Sel* 17, 375-82.
- (123) Desamero, R., Rozovsky, S., Zhadin, N., McDermott, A., and Callender, R. (2003) Active site loop motion in triosephosphate isomerase: T-jump relaxation spectroscopy of thermal activation. *Biochemistry* 42, 2941-51.
- (124) Osterman, A., Lueder, D., Quick, M., Myers, D., Canagarajah, B., and Phillips, M. (1995) Domain organization and a protease-sensitive loop in eukaryotic ornithine decarboxylase. *Biochemistry* 34, 13431-13436.
- (125) Zheng, Y., Mamdani, F., Tootygin, D., Brand, L., Stivers, J. T., and Cole, P. A. (2005) Fluorescence analysis of a dynamic loop in the PCAF/GCN5 histone acetyltransferase. *Biochemistry* 44, 10501-9.
- (126) Gutteridge, A., and Thornton, J. (2004) Conformational change in substrate binding, catalysis and product release: an open and shut case? *FEBS Lett* 567, 67-73.
- (127) Tsai, Y. C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry* 45, 9675-87.
- (128) Youngblood, B., and Reich, N. O. (2006) Conformational transitions as determinants of specificity for the DNA methyltransferase EcoRI. *J Biol Chem* 281, 26821-31.

## **VITAE**

Rahul H. Shah was born in Los Angeles, California on August 14, 1980, the son of Harshad and Raksha Shah. After completing secondary education at Servite High School in Anaheim, California, he attended the University of Southern California in Los Angeles, California. As a student at the University of Southern California he conducted research in the fields of pharmacology and molecular biology. He received a degree of Bachelor of Science with a major in biomedical engineering from the School of Engineering at the University of Southern California in 2002. Upon graduation from college, he entered the Graduate School of Biomedical Sciences at the University of Texas Southwestern Medical Center at Dallas. He was awarded the degree of Doctor of Philosophy in December, 2007.

Permanent Address: 2817 Wagon Train Lane  
Diamond Bar, CA 91765