

STRUCTURAL, KINETIC, AND BIOPHYSICAL CHARACTERIZATION OF AN α -
CARBONIC ANHYDRASE EXPRESSED IN *THIOMICROSPIRA CRUNOGENA*
GAMMAPROTEOBACTERIUM

By

NATALIA ADRIANA DIAZ TORRES

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2012

© 2012 Natalia Adriana Díaz Torres

To my amazing parents,
José A. Díaz Escribano and Virgen M. Torres Díaz

ACKNOWLEDGMENTS

I am eternally grateful for working with a most wonderful mentor, Dr. Robert McKenna, whose guidance and advice, as well as his sense of humor, helped me to achieve my best when I was feeling my worst. I thank him greatly for his understanding and his willingness to listen and help in any way he can. I doubt that I will ever be able to work with a superior as great as him.

I extend my thanks to all those who helped me complete my project. Dr. David Silverman was very helpful and also very attentive to my progress. Dr. Chingkuang Tu helped me conduct some experiments which were crucial to the completion of my project. Dr. Linda Bloom was a part of my defense committee and helped me understand what was necessary to complete the requirements to obtain my graduate degree. Without them, I would not have been able to get where I am today.

I would also like to thank all of my laboratory research fellows (lab mates) for all their advice. Particularly, I would like to thank Mayank Aggarwal, Chris Boone, and Dr. Balasubramanian Venkatakrishnan, for having the patience to deal with me on a daily basis and help me out in any way they could. I would also like to thank all the folks at the virus lab, Dr. Mavis Abgandje-McKenna included, for all of their help and advice at the joint group meetings which helped me grow as a researcher and a presenter.

I would like to thank my parents José A. Díaz Escribano and Virgen M. Torres Díaz, and my brother Ricardito, for their unconditional love and support. It was great to have them at my back, cheering me on no matter the circumstances. And lastly, I would like to thank my fiancé, Juan Carlos Castillo, without whom I would not have been able to carry forward with this project. My time at the University of Florida has been filled with many great memories that I will cherish wherever I go.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	6
LIST OF FIGURES.....	7
LIST OF ABBREVIATIONS.....	9
ABSTRACT.....	14
CHAPTER	
1 INTRODUCTION.....	16
Classification of CAs.....	16
Catalytic Mechanism of CAs.....	19
Chemolithoautotrophs and <i>Thiomicrospira crunogena</i>	21
The <i>T. cru</i> α -CA.....	23
2 MATERIALS AND METHODS.....	32
Expression and Purification of <i>T. crunogena</i> α -CA.....	32
Gel Filtration Chromatography.....	33
Crystallization and X-ray Data Collection of <i>T. cru</i> α -CA.....	33
Structure Determination of <i>T. cru</i> α -CA.....	34
Oxygen-18 Exchange Kinetic Analysis.....	35
Differential Scanning Calorimetry (DSC).....	37
3 RESULTS.....	38
4 DISCUSSION.....	59
Crystal Structure of <i>T. cru</i> α -CA.....	59
The N-terminus of <i>T. cru</i> α -CA.....	64
The Disulfide Bond.....	65
The Dimeric Interface.....	68
5 CONCLUSIONS.....	95
LIST OF REFERENCES.....	98
BIOGRAPHICAL SKETCH.....	108

LIST OF TABLES

<u>Table</u>		<u>page</u>
3-1	Data collection and refinement statistics for the crystallographic study of <i>T. cru</i> α -CA.	45
3-2	Comparison of maximal (pH independent) catalytic parameters for <i>T. cru</i> α -CA and hCA II.....	46
3-3	Comparison of inhibition constants for <i>T. cru</i> α -CA and hCA II.	47
3-4	Thermodynamic parameters of unfolding of <i>T. cru</i> α -CA at different pH values.	48
4-1	Possible interfaces observed in the crystal structure of <i>T. cru</i> α -CA as determined by PISA (102)	74
4-2	Hydrogen bond interactions at the <i>T. cru</i> α -CA dimer interface.....	75
4-3	Comparison of the interfaces of <i>T. cru</i> α -CA and other dimeric α -CAs.....	76
4-4	Comparison of dimer interfaces by secondary structure matching (SSM) superposition of <i>T. cru</i> α -CA using Coot (64)	77
4-5	List of interfacing residues of dimeric α -CAs	78

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Cartoon representation of hCA II.....	25
1-2 Cartoon representation of hCA II active site residues.....	26
1-3 Cartoon representation of the β -CA expressed in Haemophilus influenzae bacterium.....	27
1-4 Cartoon representation of Cam, the Y-CA expressed in the thermophile Methanosarcina thermophila.	28
1-5 Hydrogen bonding network involved in the zinc-hydroxide mechanism of hCA II.	29
1-6 Transmission electron micrograph of Thiomicrospira crunogena XCL-2 cells.	30
1-7 Putative model for the interaction of four carbonic anhydrases in the <i>T. cru</i> gammaproteobacterium.....	31
3-1 Crystals of <i>T. cru</i> α -CA.	49
3-2 Crystal structure of <i>T. cru</i> α -CA.	50
3-3 Superposition of the <i>T. cru</i> α -CA monomer and the hCA II crystal structure.	51
3-4 Close-up stick representation of superimposed active site residues in <i>T. cru</i> α -CA and hCA II.	52
3-5 The pH profiles for R_1 for the hydration of CO_2 catalyzed by <i>T. cru</i> α -CA and hCA II	53
3-6 The pH profiles for $k_{\text{cat}}^{\text{exch}}/K_{\text{eff}}^{\text{CO}_2}$ ($\text{M}^{-1}\text{s}^{-1}$) for the hydration of CO_2 catalyzed by <i>T. cru</i> α -CA and hCA II.	54
3-7 The pH profile for $R_{\text{H}_2\text{O}}/[\text{E}]$ (μs^{-1}) for proton transfer in the dehydration direction catalyzed by <i>T. cru</i> α -CA and hCA II.....	55
3-8 The inhibition profiles of <i>T. cru</i> α -CA by iodide, chloride, and bromide.....	56
3-9 The pH profile of the melting temperature of the dissociation of dimeric <i>T. cru</i> α -CA..	57
3-10 The pH profile of the melting temperature for the unfolding of dimeric <i>T. cru</i> α -CA.	58
4-1 Representation of the crystal packing of <i>T. cru</i> α -CA.	79

4-2	Ramachandran plot of <i>T. cru</i> α -CA residues obtained from <i>PROCHECK</i>	80
4-3	Cartoon representation of <i>T. cru</i> α -CA.....	81
4-4	Close-up stick representation of the <i>T. cru</i> α -CA active site.	82
4-5	Comparison of <i>T. cru</i> α -CA and hCA II.	83
4-6	Sequence alignment of hCA II and <i>T. cru</i> α -CA determined by ClustalW.....	84
4-7	Disorder probability plot obtained from PONDR-FIT.....	85
4-8	Cartoon representation of the dimer formed by chains A and B of <i>T. cru</i> α -CA.	86
4-9	Possible interfaces present at the crystal structure of <i>T. cru</i> α -CA.	87
4-10	Buried lysines at the <i>T. cru</i> α -CA dimer interface.	88
4-11	Superposition of the dimeric α -CAs <i>T. cru</i> α -CA and hCA VI.....	89
4-12	Superimposition of <i>T. cru</i> α -CA and hCA IX.	90
4-13	Superimposition of <i>T. cru</i> α -CA and hCA XII.	91
4-14	Superposition of <i>T. cru</i> α -CA and <i>Cr</i> - α CA1.	92
4-15	Superimposition of <i>T. cru</i> α -CA and <i>Ao</i> CA.	93
4-16	Superimposition of all dimeric α -CAs on <i>T. cru</i> α -CA.	94

LIST OF ABBREVIATIONS

Å	Angstrom
ACAP	<i>Rhodopseudomonas palustris</i> α -CA
AoCA	<i>Aspergillus oryzae</i> carbonic anhydrase
ARG	Arginine
ASN	Asparagine
ASP	Aspartate
AU	Arbitrary unit
AU	Asymmetric unit
B	Base
B-FACTOR	Temperature factor
C	Celsius
CA	Carbonic anhydrase
CA ²⁺	Calcium
CCD	Charge-coupled detector
CCM	Carbon concentrating mechanism
CM	centimeter
CO ₂	carbon dioxide
CR- α CA1	<i>Chlamydomonas reinhardtii</i> α -carbonic anhydrase
CsoCA	Carboxysomal β -carbonic anhydrase
CSS	Complexation significance score
CV	Column volume
CYS	Cysteine
3-D	Three-dimensional
DA	Dalton

DIC	Dissolved inorganic carbon
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
E	Enzyme
<i>E. COLI</i>	<i>Escherichia coli</i>
F _c	Calculated structure factor
F _o	Observed structure factor
FPLC	Fast protein liquid chromatography
ΔG	Gibbs' free energy
GDN-HCL	Guanidine hydrochloride
GLN	Glutamine
GLU	Glutamate
GLY	Glycine
G3P	Glyceraldehyde-3-phosphate
GPI	Glycophosphatidylinositol
H	Hour
ΔH	enthalpy
ΔH _v	van't Hoff enthalpy
HCA II	Human carbonic anhydrase isoform II
HCA IV	Human carbonic anhydrase isoform IV
HCA VI	Human carbonic anhydrase isoform VI
HCA IX	Human carbonic anhydrase isoform IX
HCA XII	Human carbonic anhydrase isoform XII
HCL	Hydrochloric acid
HCO ₃ ⁻	Bicarbonate

HEPES	4-(2-hydroxyethyl)-1-piperazinesulfonic acid
HIS	Histidine
H ₂ O	Water
HS ⁻	Hydrogen sulfide
ILE	Isoleucine
IPTG	Isopropyl-β-D-galactopyranoside
K	Potassium
K _B	Catalytic turnover for proton transfer
K _{CAT}	Catalytic turnover rate
K _{CAT} /K _M	Catalytic efficiency
KDA	Kilodalton
K _{EFF} ^{CO₂}	Apparent affinity constant of CO ₂ to carbonic anhydrase
LEU	leucine
Li ⁺	Lithium
LYS	Lysine
M	Molar
M ²⁺	Bivalent metal ion
MET	Methionine
MG	Milligram
Mg ²⁺	Magnesium
MIN	Minute
MM	Millimeter
MM	Millimolar
MR	Molecular Replacement
<i>M. THERMOPHILA</i>	<i>Methanosarcina thermophila</i>

μG	Microgram
μS	Microsecond
N_2	Nitrogen
NaCl	Sodium chloride
Na_2SO_4	Sodium sulfate
NCBI	National Center for Biotechnology Information
NGCA	<i>Neisseria gonorrhoea</i> carbonic anhydrase
NM	Nanometer
^{18}O	Oxygen-18
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
P-AMBS	p-(aminomethyl)benzenesulfonamide
PDB	Protein Data Bank
PEG	Polyethylene glycol
PGA	Phosphoglyceric acid
PH	Negative decimal logarithm of the hydrogen ion activity in a solution
PHE	Phenylalanine
PISA	Protein Interfaces, Surfaces, and Assemblies
PK_A	Acid dissociation constant
PRO	Proline
PSR	Proton shuttle residue
R	Correlation coefficient
RuBisCO	Ribulose-1,5-bisphosphate carboxylase oxygenase
RMSD	Root-mean-square deviation
<i>R. PALUSTRIS</i>	<i>Rhodospseudomonas palustris</i>

RPM	Revolutions per minute
S	Second
SDS	Sodium dodecyl sulfate
SER	Serine
SR ²⁺	Strontium
SUMO	Small ubiquitin-related modifier
TCA	tricarboxylic acid
TCEP	tris(2-carboxyethyl)phosphine
<i>T. CRU</i>	<i>Thiomicrospira crunogena</i>
THR	Threonine
T _M	Melting temperature
TRIS	Tris(hydroxymethyl)aminomethane
TRP	Tryptophan
TYR	Tyrosine
UV	Ultraviolet region of electromagnetic spectrum
VAL	Valine
VIS	Visible region of electromagnetic spectrum
V _M	Matthews' coefficient
2XYT	2X Yeast extract and tryptone
ZN ²⁺	Zinc

Abstract Of Thesis Presented To The Graduate School
Of The University Of Florida In Partial Fulfillment Of The
Requirements For The Degree Of Master Of Science

STRUCTURAL, KINETIC, AND BIOPHYSICAL CHARACTERIZATION OF AN α -
CARBONIC ANHYDRASE EXPRESSED IN THIOMICROSPIRA CRUNOGENA
GAMMAPROTEOBACTERIUM

By

Natalia Adriana Díaz Torres

August 2012

Chair: Robert McKenna

Major: Biochemistry and Molecular Biology

Carbonic anhydrases (CA) are mainly zinc metalloenzymes that catalyze the interconversion of water and carbon dioxide to bicarbonate and protons. As such, CAs are involved in many important physiological processes, such as respiration and pH regulation in tissues, making them essential to most living organisms. *Thiomicrospira crunogena* XCL-2 (*T. cru*) is a novel sulfur-oxidizing chemolithoautotroph that plays a significant role in the sustainability of deep-sea hydrothermal vent communities. The recently discovered deep-sea vent gammaproteobacterium, *Thiomicrospira crunogena* XCL-2, encodes and expresses four different CAs from all evolutionarily and structurally distinct families of carbonic anhydrases: a α -CA, two β -CAs, and a γ -CA [1]. In order to characterize and elucidate physiological roles of these CAs, structural, kinetic, and biophysical studies have been performed on the α -CA expressed by the *T. cru* gammaproteobacterium. The structure and catalytic properties of the α -CA were shown to be similar to hCA II, which could imply that the *T. cru* α -CA could exert similar physiological roles to hCA II. However, significant differences regarding the quaternary structure and the thermal stability were observed. *T. cru* α -CA forms a homodimer in

solution, which is rare among the mostly monomeric α -class CAs. Also, the thermal stability of the *T. cru* α -CA was greatly enhanced as compared to hCA II, which could be the result of the presence of a disulfide bond between residues Cys99 and Cys256. The results obtained in this project are used to provide insights to the function of the *T. cru* α -CA in the biological context.

CHAPTER 1 INTRODUCTION

Carbonic anhydrases (CAs) are mainly zinc metalloenzymes that catalyze the interconversion of carbon dioxide and water to bicarbonate and protons (1-3). Carbonic anhydrase was initially discovered as the enzyme responsible for catalyzing the conversion from bicarbonate to carbon dioxide necessary for rapid transit from the red blood cells to the lung capillaries in mammals (4). However, CAs are found in virtually all living organisms and have been shown to be involved in many important and complex physiological processes, such as respiration, pH regulation in tissues, renal tubular acidification, bone resorption, and secretory processes including the formation of cerebrospinal fluid and aqueous humor in animals (3, 5 -7). CAs have also been found to play important roles in physiological processes of more primitive organisms, including photosynthesis, CO₂ fixation and transport in plants and bacteria, proving they're essential in most living organisms (8, 9).

Classification of CAs

Three evolutionary and structurally distinct classes of CAs have been established according to sequence similarities and the overall folds, both of which indicate different phylogenetic and physiological pathways for each class (10, 11). The first and best-characterized class is the α -CA class, which mostly comprises mammalian CAs. There have been at least 16 different α -CA isoforms isolated from mammals. These vary in their cellular expression and localization, as well as catalytic activity. The α -CAs are mainly cytosolic (CA I, II, III, VI, XIII), transmembrane or GPI-anchored (CA IV, IX, XII, XIV), secretory (CA VI), and mitochondrial (CA VA, VB) (3). However, some α -CAs, which up until recently were thought to be exclusively expressed by eukaryotic

organisms, have also been found in the periplasmic space of certain prokaryotes, such as the Gram-negative bacterium *Neisseria gonorrhoeae* (12), the nitrogen-fixing alphaproteobacterium *Mesorhizobium loti* (13) and the purple non-sulfur bacterium *Rhodopseudomonas palustris* (14). Furthermore, most α -CAs are monomeric in nature, but there are a few that are dimeric, such as hCA VI, hCA IX, hCA XII, AcaP (α -CA in *R. palustris*) and α -CA I from green alga *Chlamydomonas reinhardtii*, among others. Most α -class CAs are roughly spherical enzymes with a highly conserved active site formed by a conical cavity that is approximately 15 Å deep, as can be observed in hCA II (Figure 1-1) (7). At the bottom of the active site lies the zinc ion tetrahedrally coordinated by 3 conserved histidines and a water molecule/hydroxide ion. The active site cavity is split into two very different environments, where one side of the zinc believed to be where the CO₂ binds is lined with mostly hydrophobic residues, and the other side which leads out of the active site into the bulk solvent is lined with hydrophilic amino acids (Figure 1-2). The dual nature of the active site in α -CAs allows the rapid and sustained catalytic conversion of CO₂ to bicarbonate (2).

A second, less-understood class of carbonic anhydrases is comprised by β -class CAs, predominantly found in plants, algae, and prokaryotic organisms belonging to both Archaea and Bacteria domains (10, 15, 16). The first crystal structure for β -CAs was reported in 2000 (17), confirming that the understanding of these enzymes is far behind compared to that of α -CAs. There are many differences between the β -CAs and the other two classes. Structurally, it has been demonstrated that β -CAs form dimers, tetramers, hexamers, and octamers, suggesting that dimers are the basic building units for β -CAs, unlike monomeric α -CAs and trimeric γ -CAs (18). An example of a crystal

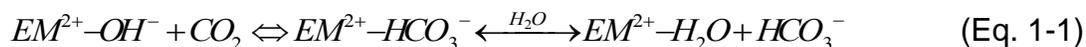
structure of a β -CA from *Haemophilus influenzae* (PDB ID: 2A8C) is presented in Figure 1-3 (19). Furthermore, though β -CAs are much more diverse than α - and γ -CAs, it can be established that β -CAs have a higher α -helical content in their secondary structures than the other two classes of CAs (15). Not much information is available on the different isoforms comprising the β -class CAs, as β -CAs are much more phylogenetically diverse than the α -class or γ -class CAs (20). Sequence analyses demonstrated that only 5 residues are conserved amongst β -CAs, which are the three zinc ligands, an arginine and an aspartic acid (21). In contrast to α -CAs, the zinc is tetrahedrally coordinated by two cysteines and a histidine plus a fourth ligand, which can be a water molecule seen in the crystal structure of the “cab” type β -CA, denoted Cab, from *Methanobacterium thermoautotrophicum* (16) an acetate molecule observed in the β -CA from *Pisum sativum* crystallized in acetate (18) or even the conserved aspartate of β -CAs, as seen in the *Escherichia coli* and *Porphyridium purpureum* β -CA crystal structures (15, 17). Despite the differences in the active site to α - and γ -CAs, the β -CAs are believed to catalyze the interconversion of CO₂ and bicarbonate through a zinc-hydroxide mechanism as well, explained in detail further below.

The third class of CAs is the γ -CA class, initially thought to be solely expressed in methane-producing bacteria. The first γ -CA reported was Cam, isolated from *Methanosarcina thermophila* in 1994 (22) and crystallized in 1996 (Figure 1-4) (PDB ID: 1QRG; 23, 24). Only a handful of γ -CAs have been crystallized thus far, but they all share the same structural characteristics. γ -CAs crystallized to date are homotrimeric structures whose monomers share a distinctive left-handed parallel β -helix fold that is predicted by a unique sequence motif also observed in the superfamily of

acyltransferases (23, 25). As in the α -CA active site, the γ -CA zinc is tetrahedrally coordinated by 3 histidines and a water molecule or hydroxide ion, but the location of the active site is not buried deep within the enzyme. The active site of γ -CA is located at the interface between two monomers, where two histidine ligands are donated by one monomer, and the other histidine comes from the adjacent monomer forming the active site interface (10, 26). Cam was initially thought to be a zinc metalloenzyme, as high-level expression in *E. coli* yields a zinc enzyme. However, substitution of zinc with other transition metals, such as iron and cobalt, yielded higher rates of CO₂ hydration for Cam, implying that native Cam expressed in *M. thermophila* might function with a different transition metal, with iron appearing to be the physiologically relevant metal (27-29). Independent of the transition metal occupying the active site, the γ -CAs that are active exhibit a metal hydroxide catalytic mechanism with proton transfer as the rate-limiting step, as observed for α - and β -CAs (10).

Catalytic Mechanism of CAs

The mechanism underlying the catalytic activity of CAs has been extensively studied for half a century. Studies have shown that α -, β -, and γ -class CAs share the same overall metal (Zn)-hydroxide ping-pong mechanism composed of two independent catalytic steps, as shown below in Eq. 1 and 2, where E is the enzyme and M²⁺ is a bivalent ion, typically Zn²⁺ in CAs (26, 30-32).



Equation 1-1 shows the first step of catalysis, the hydration of CO₂ to bicarbonate, where a pair of electrons on the metal-bound hydroxide (in most CAs, zinc-bound

hydroxide) nucleophilically attacks the substrate CO_2 to form zinc-bound bicarbonate (30, 31). A water molecule subsequently displaces the zinc-bound bicarbonate, leaving a zinc-bound water at the active site. In α -CAs, it has been shown that CO_2 binds in a conserved hydrophobic region of the active site before the nucleophilic attack (2). A more detailed look into the active site of one of the fastest CAs known, hCA II, shows a hydrogen bond network essential for the hydration of CO_2 , formed by residues Thr199 and Glu106. Briefly, hydrogen bonding between the Thr199 hydroxyl and the Glu106 carboxylate oxygen allows the Thr199 hydroxyl group to act as a hydrogen-bond acceptor to the zinc-bound hydroxide, optimizing the orientation of the electron pair of the hydroxide ion for nucleophilic attack on CO_2 (33). Furthermore, the backbone amide of Thr199 hydrogen bonds with CO_2 , slightly polarizing the carbon on CO_2 for nucleophilic attack and providing an environment which promotes the dissociation of bicarbonate for more efficient product removal from the active site (Figure 1-5) (34-36). Interestingly, β - and γ -CAs do not have corresponding residues to Thr199 and Glu106, but they do have residues that are proposed to have the same catalytic function. Structural and kinetic studies have shown that the active site residue Gln158 in the β -CA from *Arabidopsis thaliana*, and equivalent Gln151 in *P. sativum* β -CA, might provide the same function as the backbone amide of Thr199 (18, 37, 38). Similarly, the residues Asn202 and Gln75 in the γ -CA Cam have been demonstrated to assume the catalytic function of Thr199, where Asn202 is responsible for polarizing the CO_2 molecules and Gln75 primes the hydroxide ion for nucleophilic attack on CO_2 (32).

During a second, independent step shown in Eq. 1-2, a proton is transferred from the zinc-bound water to the bulk solvent, regenerating the zinc-bound hydroxide ion.

This proton-transfer stage is the rate-limiting step of catalysis, which for hCA II is in the order of 10^6 s^{-1} (31, 39). In Eq. 1-2, B represents a proton donor/acceptor, which can be a buffer molecule in solution for intermolecular proton transfer, or a proton shuttle residue (PSR) in the enzyme itself. The rate for proton transfer directly from the zinc-bound water with a $\text{pK}_a \sim 7$ to the bulk solvent (k_{cat}) is approximately 10^4 s^{-1} . Enzymes exhibiting a $k_{\text{cat}} > 10^4 \text{ s}^{-1}$ (reflective of the proton-transfer rate) have an intermediate PSR responsible for transferring the proton from the zinc-bound water to the bulk solvent (2, 10, 26). In hCA II, a network of highly ordered water molecules at the active site shuttle the proton to His64, which acts as a PSR, transferring the proton out of the active site to the bulk solvent (40-42). PSRs have also been observed in other CAs, such as Tyr131 in the α -class CA-VA isozyme, Asp34 in the β -CA Cab from *M. thermoautotrophicum* (which is the conserved aspartate in the “cab” subtype of β -CAs), and Glu84 in the γ -CA Cam from *M. thermophila*, which further underscores the importance of PSRs in the regeneration of the active site in CAs (10, 24, 43, 44).

Chemolithoautotrophs and *Thiomicrospira crunogena*

Genomic studies of the novel deep-sea vent chemolithoautotroph *Thiomicrospira crunogena* XCL-2 (*T. cru*) revealed the coding and expression of four different CAs, which span the three evolutionarily and structurally distinct α -, β -, and γ -CA families (45, 46). Chemolithoautotrophs are defined as bacterial organisms that obtain energy from the oxidation of inorganic compounds to fuel carbon fixation using CO_2 as their primary source of carbon (45, 46). Some of these bacteria, including *T. cru*, live in deep-sea hydrothermal vents, a harsh habitat where nutrients, such as dissolved inorganic carbon (DIC) and oxygen, are temporally and spatially limited. Although the hydrothermal fluid chemical composition might vary geographically, hydrothermal fluid usually contains

millimolar concentrations of alkali metal cations Li^+ and K^+ in solution, as well as alkaline earth cations Ca^{2+} , Mg^{2+} , and Sr^{2+} , which generally form deposits upon reactions with the reduced sulfate anions present in the fluid. Vent fluid also contains high concentrations of some trace minerals, such as manganese, iron, and zinc, as well as high concentrations of reduced sulfur anions, chloride, and low concentrations of bromide (47). The environment at hydrothermal vents is constantly changing, as dilute hydrothermal fluid (warm, anoxic, highly reduced, pH 5-8, [DIC] = 2-7 mM) interacts with bottom water (2°C, oxic, neutral pH, [DIC] = 20µM), leading to oscillatory habitat chemistry dominance (45). Interestingly enough, many of these bacteria have developed a wide array of adaptations, such as carbon-concentrating mechanisms (CCM), symbiotic relationships with invertebrates, and formation of microbial mats or veils to produce a stable microenvironment, which allow them to grow and thrive in these nutrient-limited habitats (8).

T. cru is a sulfur-oxidizing gammaproteobacterium that plays a significant role in the sustainability of deep-sea hydrothermal vent communities. It is a spiral-shaped, colorless, obligate aerobe isolated from the East Pacific Rise deep-sea hydrothermal vents that uses hydrogen sulfide and other reduced sulfur species to fuel cell maintenance and carbon fixation (Figure 1-6). *T. cru* exhibits optimum growth in an environment with a pH range of 5.0 to 8.5, NaCl concentration > 45 mM, and temperature range from 4 to 38.5 °C (48, 49). *T. cru* encodes a carbon-concentrating mechanism which allows it to rapidly grow under conditions of bicarbonate and CO_2 scarcity due to an increase in the cells' affinity for these carbon compounds, a critical adaptation to survive in the harsh environment found at the hydrothermal vents (45).

Previous results have shown that active transport of bicarbonate or CO₂ plays a role in the CCM. Therefore, the expression of four different CAs suggests a role in bicarbonate and carbon dioxide transport, as well as an indirect role in CO₂ fixation, as part of the CCM (46). Genomic and expression studies have shown that *T. cru* expresses a periplasmic α-CA, a cytoplasmic β-CA, a carboxysomal β-CA, and a γ-CA like protein. Figure 1-7 shows a putative model of the molecular interaction of the CAs as part of the CCM in *T. cru*. In the figure, the function of the periplasmic α-CA could involve CO₂ or bicarbonate sequestration to prevent diffusion out of the cell and facilitate CO₂ diffusion into the cytoplasm, as is observed for the periplasmic α-CA expressed in *R. palustris* (14). Another possible function of the α-CA could be pH homeostasis, similarly to the α-CA in *Helicobacter pylori* (50). As CO₂ diffuses into the cytoplasm, the cytoplasmic β-CA could be associated with maintaining a CO₂-bicarbonate pool near-chemical equilibrium in order to facilitate the use of CO₂ by form II ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO II), or CO₂ entering the carboxysome. The carboxysomal β-CA (CsoCA) is involved in carbon fixation, possibly sequestering CO₂ in the carboxysome in order to be fixated into phosphoglyceric acid (PGA) by RuBisCO I, which has a low affinity for CO₂ and low turnover rates (51). Subsequently, PGA enters the Calvin-Benson-Bassham cycle and is converted to glyceraldehyde 3-phosphate (G3P), which then enters a reductive tricarboxylic acid cycle (TCA) to form starch, pyruvate, and other carbon skeletons required for normal cellular function (45).

The *T. cru* α-CA

To gain insight into the physiological relevance of these CAs in the carbon-concentrating mechanism of *T. cru*, we have focused on the characterization of the α-CA. The *T. cru* α-CA gene encodes a full-length protein consisting of 315 amino acid

residues. The periplasmic location of some α -CAs in other previously mentioned prokaryotes prompted the analysis of the amino acid sequence by the SignalP software (52, 53), which predicted a signal peptide -NVAAP- at the N-terminus. Therefore, the α -CA expressed in *T. cru* contains 292 amino acid residues, with a calculated molecular weight of 33 kDa. Also, its expression in the periplasm of a hydrothermal vent gammaproteobacterium suggests that the structure could exhibit features related to increased thermal stability, which could potentially be applied to engineer more thermostable CA isoforms that could be used in industrial applications. Attempts to obtain a *T. cru* α -CA knockout mutant have been unsuccessful (unpublished data by the Scott laboratory), which suggests that *T. cru* α -CA has an essential role in the *T. cru* gammaproteobacterium's growth and survival.

We hypothesized that *T. cru* α -CA has a similar structure and activity to hCA II, and is possibly involved in CO₂ entrapment and carbon fixation. To complete this project, we performed kinetic studies of the *T. cru* α -CA using oxygen-18 (¹⁸O) exchange at chemical equilibrium measured by membrane inlet mass spectrometry, which is further explained in Chapter 2, used differential scanning calorimetry at different pH values to study its thermal stability, and obtained a crystal structure of the *T. cru* α -CA by X-ray crystallography. The results of these studies are presented in Chapter 3, and discussed in further detail in Chapter 4.

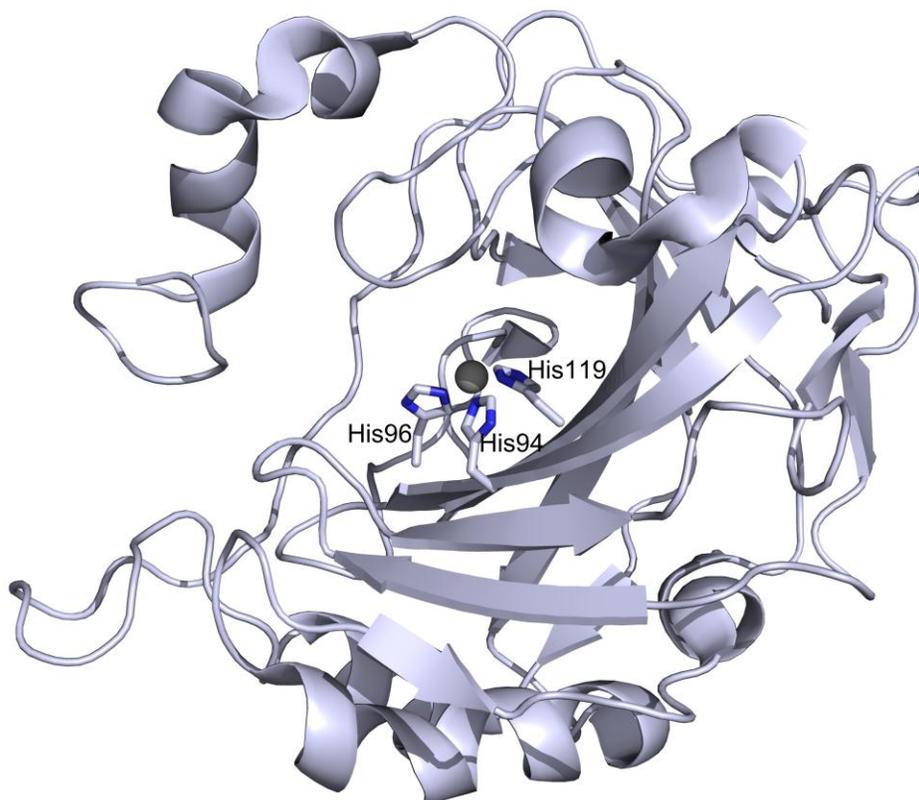


Figure 1-1. Cartoon representation of hCA II. The zinc-coordinating histidines are represented as sticks, and the zinc ion is shown as a gray sphere. Nitrogen atoms are shown in blue. (PDB ID: 3KS3; 54).

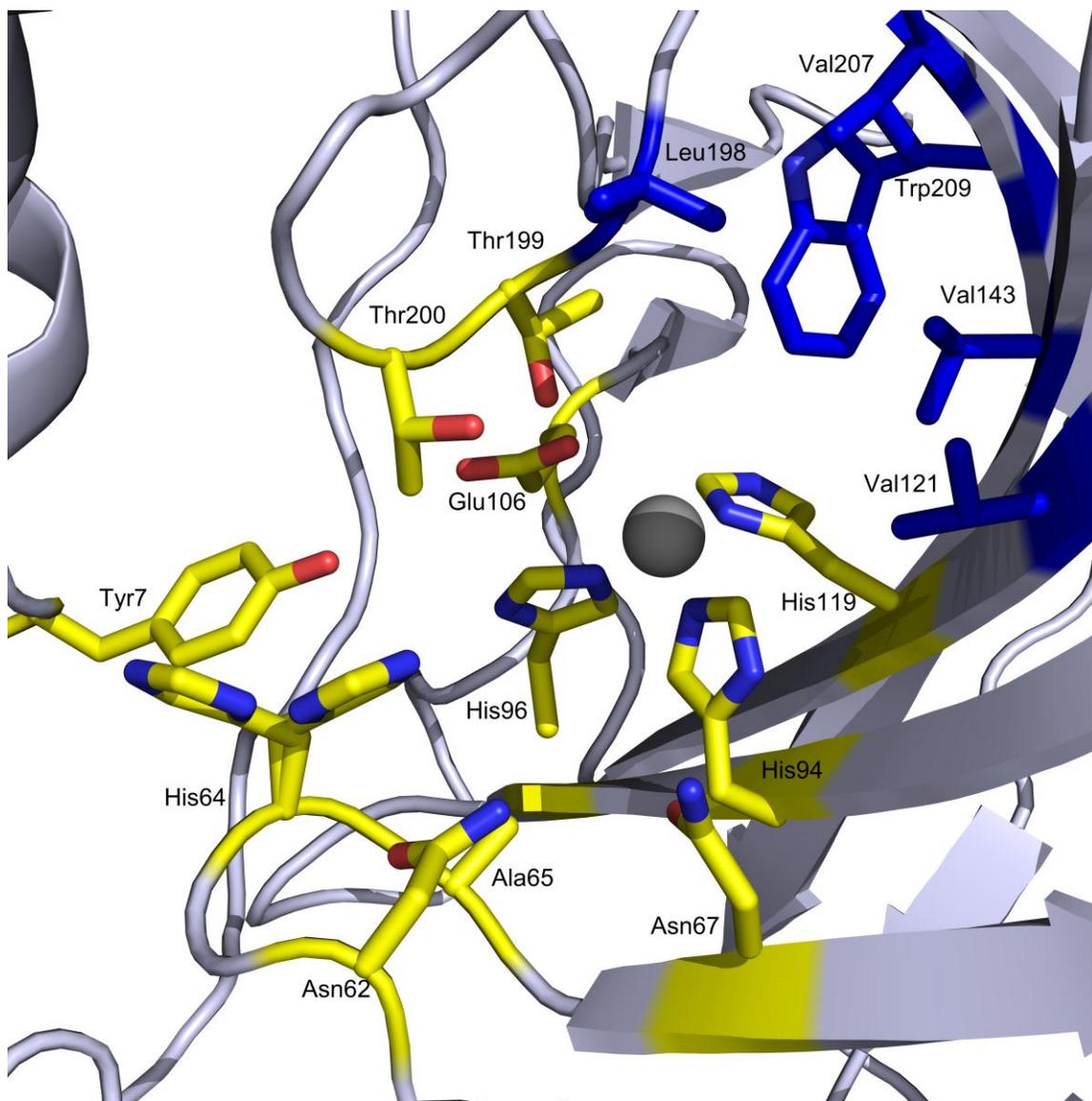


Figure 1-2. Cartoon representation of hCA II active site residues (PDB ID: 3KS3; 54). The active site residues are represented as sticks. To highlight the dual nature of the active site, the hydrophilic active site residues are shown in yellow, and the hydrophobic residues are shown in blue. The zinc ion is represented as a gray sphere. Oxygen atoms are shown in red, and nitrogen atoms in blue.

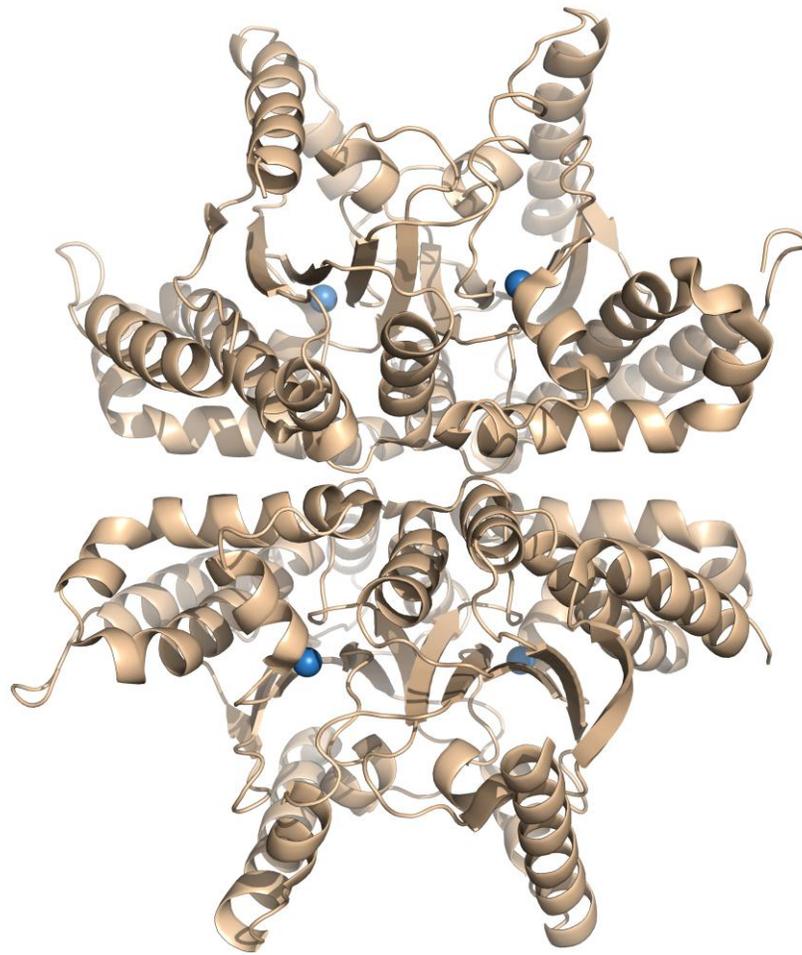


Figure 1-3. Cartoon representation of the β -CA expressed in *Haemophilus influenzae* bacterium (PDB ID: 2A8C; 19). The biological assembly of the protein is a tetramer, and each monomer has its own active site. The zinc ions are shown as blue spheres.



Figure 1-4. Cartoon representation of Cam, the γ -CA expressed in the thermophile *Methanosarcina thermophila* (PDB ID: 1QRG; 24). The biological assembly of Cam is a trimer, with each chain presented in a different color (A – red, B – green, and C – cyan). The zinc ion in chain A, shown as a gray sphere, is coordinated by two histidines (His81a, His122a) in chain A, and one histidine in chain C (His117c). His81b, His122b, His117a, and His81c, His122c, His 117b, coordinate the zinc ion in chains B and C, respectively. Zinc-coordinating histidines are represented as sticks, with the oxygen atoms shown in red, and nitrogen atoms shown in blue.

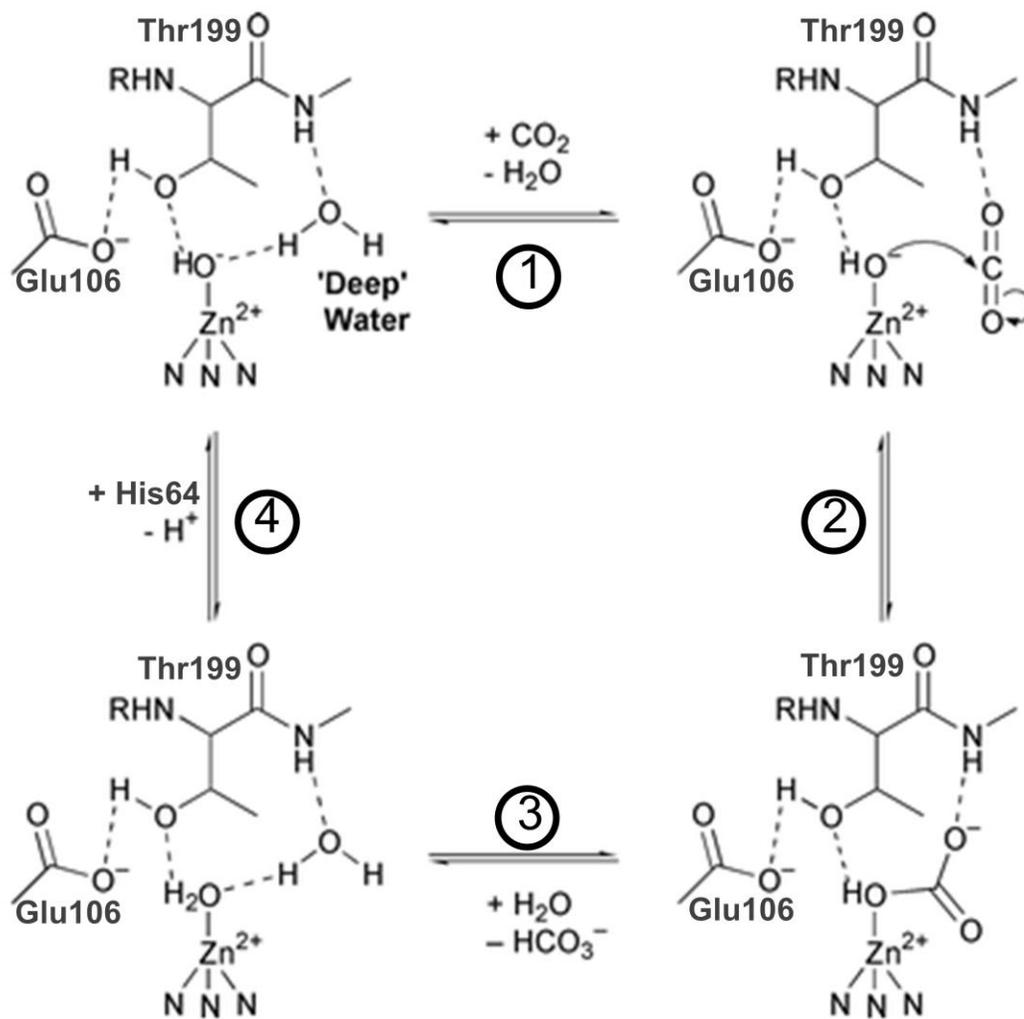


Figure 1-5. Hydrogen bonding network involved in the zinc-hydroxide mechanism of hCA II. 1) Glu106 hydrogen bonds to the hydroxyl group of Thr199, which acts as a hydrogen-bond acceptor to the zinc-bound hydroxide, orientating the electron pair for nucleophilic attack. Also, the backbone amide group of Thr199 polarizes the CO_2 by hydrogen bonding. 2) After nucleophilic attack on CO_2 by the zinc-bound hydroxide, the bicarbonate ion is formed and, 3) is removed from the active site by displacement of a water molecule. 4) His64 serves as a proton shuttle to transfer the proton from the zinc-bound water to the bulk solvent, regenerating the enzyme for catalysis. Modified from Sprigings and Hall, 2001 (55).

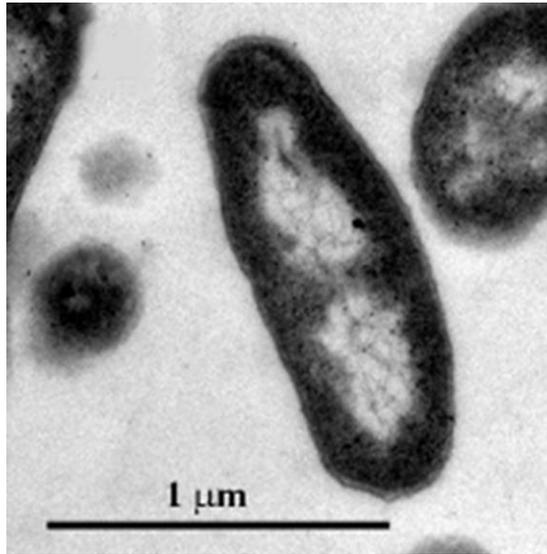


Figure 1-6. Transmission electron micrograph of *Thiomicrospira crunogena* XCL-2 cells. The *T. cru* cells are colorless, spiral-shaped, gram-negative bacteria ($0.4 - 0.5 \times 1.5 - 3.0 \mu\text{m}^2$) that contain a single flagellum (not shown in figure). Provided by Dr. Kathleen Scott.

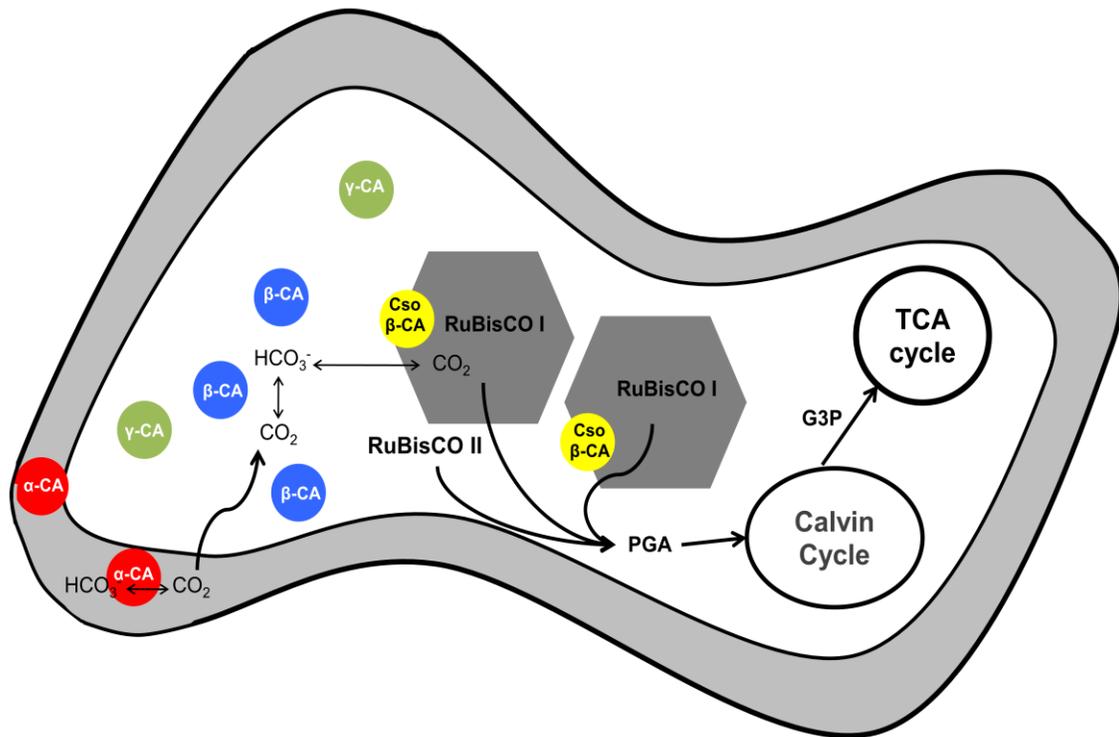


Figure 1-7. Putative model for the interaction of four carbonic anhydrases in the *T. cru* gammaproteobacterium. The periplasmic α -CA (red) could sequester CO_2 or bicarbonate to prevent diffusion out of the cell. As CO_2 diffuses into the cytoplasm/ bicarbonate is transported back into the cell, the cytoplasmic β -CA (blue) could maintain an equilibrium of the CO_2 /bicarbonate pool, until CO_2 is used by form II ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO II) or enters the carboxysome. The carboxysomal β -CA ($\text{Cso } \beta\text{-CA}$; yellow) is involved in carbon fixation, probably by sequestering CO_2 in the carboxysome in order to be fixated into phosphoglyceric acid (PGA) by RuBisCO I , which has a low affinity for CO_2 and low turnover rates (51). As part of the carbon fixation, PGA continues on the Calvin-Benson-Bassham cycle and is converted to glyceraldehyde 3-phosphate (G3P), which then enters a reductive tricarboxylic acid cycle (TCA) to form starch, pyruvate, and other carbon skeletons required for normal cell function (45).

CHAPTER 2 MATERIALS AND METHODS

Expression and Purification of *T. crunogena* α -CA

A pET SUMO plasmid containing the *T. cru* α -CA gene was transformed into BL21(DE3) *E. coli* cells through standard protocols for high yield protein expression, and the transformed cells were incubated at 37°C in a modified Luria Broth with 2x yeast extract and tryptone (2XYT broth) containing 50 μ g/mL kanamycin at 200 rpm to an OD₆₀₀ of 0.6 - 1.0 AU (46, 56). *T. cru* α -CA expression was induced for 18 h at 18°C by addition of isopropyl β -D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. The cells were harvested at 4°C by centrifugation at 4000 rpm for 10 min and were kept overnight at -20°C. The cell pellets were resuspended in lysis buffer (0.1 M Tris-HCl pH 9.0, 0.2 M Na₂SO₄) with 1 mg/mL lysozyme and 50 μ g/mL DNase I, and were lysed at 4°C stirring overnight. The samples were then centrifuged for 70 min at 15000 rpm.

T. cru α -CA was purified through affinity chromatography using a p-(aminomethyl)benzenesulfonamide (p-AMBS) resin as has been previously described for hCA II purification (57), with several adjustments. Briefly, the cell lysate was loaded onto the affinity column equilibrated with 10 column volumes (CV) of the lysis buffer. The column was washed with at least 20 CV of the lysis buffer and 20 CV of the lysis buffer at pH 7.0, until the absorbance of the flow-through measured at 280 nm was below the detection limit for proteins, 0.1 AU, after each wash. The protein was eluted with a sodium azide buffer (50 mM Tris-HCl pH 7.8, 0.4 M sodium azide), after which the eluted samples were buffer-exchanged and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit to remove the sodium azide. Further purification of the protein

required a second step of p-AMBS affinity chromatography. The protein samples were desalted and concentrated using the Amicon centrifugal filter. The final sample concentration was determined by UV-Vis spectroscopy at 280 nm, using an extinction coefficient of $42985 \text{ M}^{-1} \text{ cm}^{-1}$, calculated from the amino acid sequence (58), followed by SDS-PAGE analysis to assess purity (59).

Gel Filtration Chromatography

A 30 μM sample of *T. cru* α -CA was analyzed on a GE Healthcare ÄKTA Fast Protein Liquid Chromatography (FPLC) protein purification system, equipped with a prepacked HiPrep 16/60 Sephacryl S-200 High Resolution gel filtration column (Exclusion range: 5 to 250 kDa; GE Healthcare Biosciences AB, Sweden) previously equilibrated with 50 mM Tris-HCl pH 7.0, 100 mM NaCl buffer. Data acquisition and processing were performed using the FPLC UNICORN software. The protein was eluted with the aforementioned buffer at a 0.1 mL/min flow rate, collecting 1 mL fractions. Protein fractions were detected by absorbance at 280 nm, and stored at 4°C. The Sephacryl S-200 column was calibrated by using Gel Filtration Standard molecular weight markers from Bio-Rad (Cat. no. 151-1901), following the provided instructions.

Crystallization and X-ray Data Collection of *T. cru* α -CA

Initial crystallization screening was performed in INTELLI-PLATE 96-well sitting drop vapor diffusion crystallization plates (Art Robbins Instruments, Sunnyvale, California, USA) using a Crystal Gryphon Protein Crystallization System (Art Robbins Instruments, Sunnyvale, California, USA) with four commercial screens: Crystal Screen, Crystal Screen 2, PEG/Ion Screen, and PEG/Ion2 Screen (Hampton Research, Aliso Viejo, California, USA). Drops consisting of ~0.3 mM protein sample (in 50 mM Tris-HCl pH 7.8, 100 mM NaCl) and precipitant solution at two different ratios (1:1 and 2:3

protein/precipitant solution) were equilibrated at 17°C against a reservoir containing 100 μ L of the precipitant solution. Small protein crystals were obtained in 2% v/v Tacsimate pH 4.0, 0.1 M Sodium acetate trihydrate pH 4.6, 16% w/v PEG 3350 (Solution 31 from PEG/Ion2 Hampton Screen). The crystals were cryoprotected by rapidly immersing it into the precipitant solution with 20% glycerol prior to freezing in liquid N₂.

X-ray diffraction data was collected using an ADSC Quantum 270 CCD detector at the Cornell High Energy Synchrotron Source (CHESS) beamline F1ZC ($\lambda = 0.9177$ Å). The data sets were collected at a crystal-to-detector distance of 300 mm, 1° oscillation angle with an exposure time of 60s per image for 180 frames. The data were indexed, integrated and scaled to a 2.6 Å resolution with HKL2000 software (60).

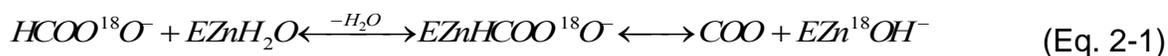
Structure Determination of *T. cru* α -CA

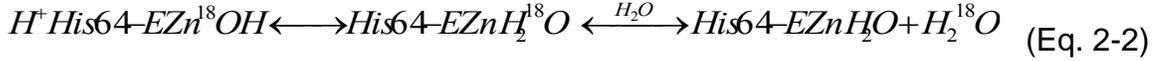
The crystal structure of hCA II (PDB ID: 3KS3; 54) was used to calculate the initial phases of the *T. cru* α -CA using the *PHENIX* AutoMR software (61, 62). The zinc and all solvent molecules were removed to decrease model bias. *PHENIX* AutoMR yielded a coordinate file with 4 molecules in the asymmetric unit, and two Fourier electron density map files, a $2F_o-F_c$ map which shows density observed from the new model and a F_o-F_c map which shows the difference between the observed density in the model and experimental data. The model obtained from AutoMR was truncated, where amino acid residues that had no density were removed, and additional residues were built into observed density. The truncated model was used for a second round of molecular replacement with *PHENIX* AutoMR. The coordinate file and electron density maps obtained from the second round of molecular replacement were used to build in the residues that were missing in the truncated model to complete the structure of *T. cru* α -CA. Structure refinement was completed by alternating the use of the *PHENIX* software

package with density modification using AutoBuild (63) to enhance the Fourier electron density maps and manual refitting of the model in *Coot* (64). 5% of the unique reflections were selected randomly before initiating refinement and were excluded from refinement for purposes of R_{free} calculation (65). The validity of the model was assessed by *PROCHECK* (66). Secondary structure moieties were assigned using the STRIDE program (67). Figures were made using PyMOL.

Oxygen-18 Exchange Kinetic Analysis

The ^{18}O -exchange experiments presented in this thesis were conducted by Dr. Chingkuang Tu in the Silverman laboratory. The method is based on the measurement of the depletion of ^{18}O from species of CO_2 at chemical equilibrium by membrane inlet mass spectrometry (68, 69). A continuous measure of various isotopic species of CO_2 is provided by CO_2 diffusing across a membrane permeable to dissolved gases, which is submerged in the reaction solution and connected by glass tubing to a mass spectrometer (Extrel EXM-200) (70). The catalyzed and uncatalyzed exchange of ^{18}O between CO_2 and water at chemical equilibrium were measured in the absence of added buffer at a total substrate concentration of 25 mM. The reaction is kept without added buffers to simplify data analysis, since buffers can interfere with proton transfer. The reaction solution was maintained at 25°C , and the ionic strength of the solution was kept at 0.2 M by addition of Na_2SO_4 . During the first independent stage of catalysis, there is a probability that dehydration of the labeled HCO_3^- will label the active site with ^{18}O (Eq. 2-1). During the second catalytic stage, the zinc-bound ^{18}O -labeled hydroxide is protonated, forming H_2^{18}O , which then diffuses into the solvent (Eq. 2-2).





The ^{18}O -exchange kinetic approach to measure the catalysis of CA yields two rates. The first, denoted R_1 , describes the rate of interconversion of carbon dioxide and bicarbonate at chemical equilibrium (Eq. 2-3). $k_{\text{cat}}^{\text{ex}}$ is a rate constant for maximal interconversion of substrate to product, in this case CO_2 and HCO_3^- , whereas $K_{\text{eff}}^{\text{CO}_2}$ is an apparent affinity constant for substrate to the enzyme, either CO_2 or HCO_3^- to carbonic anhydrase. The ratio $k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$ is equivalent to the catalytic efficiency (k_{cat}/K_M) of hydration obtained by steady-state methods (71).

$$\frac{R_1}{[E]} = \frac{k_{\text{cat}}^{\text{ex}} [\text{CO}_2]}{K_{\text{eff}}^{\text{CO}_2} + [\text{CO}_2]} \quad (\text{Eq. 2-3})$$

The second rate obtained by the ^{18}O -exchange method, $R_{\text{H}_2\text{O}}$, describes the rate at which ^{18}O -labeled water is released from the active site. $R_{\text{H}_2\text{O}}$ is dependent on donation of protons to the ^{18}O -labeled zinc-bound hydroxide by His64 as a second independent step of the catalysis by hCA II, or its equivalent in other isozymes, as shown in Eq. 2-2 (68, 72). In Eq. 2-4, k_B is the rate constant for proton transfer to the zinc-bound hydroxide, and $(K_a)_{\text{His64}}$ and $(K_a)_{\text{ZnH}_2\text{O}}$ are the ionization constants of the proton donor His64 and the zinc-bound water molecule, respectively.

$$\frac{R_{\text{H}_2\text{O}}}{[E]} = \frac{k_B}{\left[1 + \frac{(K_a)_{\text{His64}}}{[H^+]}\right] \left[1 + \frac{[H^+]}{(K_a)_{\text{ZnH}_2\text{O}}}\right]} \quad (\text{Eq. 2-4})$$

Equation 2-3 and 2-4 were fitted to the data by using non-linear least squares methods in Enzfitter (Biosoft).

Differential Scanning Calorimetry (DSC)

Differential scanning calorimetry (DSC) experiments were performed to assess the thermostability of *T. cru* α -CA using a VP-DSC microcalorimeter (Microcal, Inc., North Hampton, MA) with a cell volume of ~ 0.6 mL. *T. cru* α -CA samples were buffer-exchanged into solutions with pH values ranging from 4 to 9, at 1 pH unit intervals. A phosphate/citrate buffer was used for samples at pH values 4-6, and a 50 mM Tris HCl, 100 mM NaCl buffer was used for samples at pH values 7-9, all of which had a protein concentration of 30 μ M. The protein and buffer samples were degassed, while stirring, at 16°C for 20 minutes prior to data collection. DSC scans were collected from 30°C to 90°C with a scan rate of 60°C/hr. DSC scans for all the samples were performed in triplicate. A reference scan for each pH value, where buffer was placed in both the reference and sample cells, was also performed in order to subtract the contributing heat capacity of the buffer solution. After subtracting the reference and adjusting the baseline for each scan, the calorimetric enthalpies of unfolding were calculated by integrating the area under the peaks in the thermograms. The thermograms were fit to a non two-state reversible unfolding model to obtain van't Hoff enthalpies of unfolding (ΔH_v) (73). The melting temperature (T_M) values of the *T. cru* α -CA samples at the different pH values and of hCA II were obtained from the midpoints of the thermograms.

CHAPTER 3 RESULTS

The α -CA expressed in *Thiomicrospira crunogena* was successfully overexpressed and isolated with a usual yield of 10 mg per liter of bacterial culture. Small, rectangular crystals with approximate dimensions $0.5 \times 0.5 \times 0.01 \text{ mm}^3$ formed after 14 days in a variety of crystallization conditions containing PEG 3350 over a pH range of 4.0 – 7.4 (Figure 3-1). Crystals grown by hanging drop vapour diffusion method in 2% v/v tacsimate pH 4.0, 0.1 M sodium acetate trihydrate pH 4.6, 16% w/v PEG3350 at 17°C were cryoprotected in 20% glycerol and used to collect diffraction data both “in-house” and at CHESS synchrotron. Diffraction data at 2.6 Å resolution was collected at CHESS synchrotron. The crystals were shown to belong to the C2 space group, with unit cell parameters $a = 127.1$, $b = 102.2$, $c = 105.0 \text{ Å}$, $\beta = 127.3^\circ$ with an R_{sym} of 10.0%. Initial data processing with HKL2000 software (60) also provided the possibility of an F222 space group, with unit cell parameters $a = 102.2$, $b = 127.0$, $c = 167.2 \text{ Å}$, but with an R_{sym} of ~ 50%, which led to the assignment of C2 as the space group. Data-collection statistics and processing parameters are summarized in Table 3-1.

Considering the C2 space group, the unit cell parameters, and assuming the protein molecular weight was 33217 Da with 4 molecules in the crystallographic asymmetric unit, a Matthews coefficient (V_M) of $2.04 \text{ Å}^3 \text{ Da}^{-1}$ was calculated, with an estimated solvent content of 39.75% (74, 75).

The diffraction data were phased using the molecular replacement method with the structure of hCA II (PDB ID: 3KS3; 54) using *PHENIX* AutoMR. This yielded a unique solution, comprising of 4 molecules in the asymmetric unit with a final translation function Z-score (TFZ) of 17.9. A TFZ score above 8 usually indicates that the MR was

successful. The structure was then refined using standard protocols to 2.6 Å resolution with a R_{work} of 20.8% and R_{free} of 24.4%. The average B factor for the zinc molecules was 21.6 Å². The overall B factor for the main-chain and side-chain atoms were 30.3 Å² and 33.3 Å², respectively. A total of 104 solvent molecules were added, with an average B factor of 28.5 Å². Root-mean-square deviation (RMSD) values from ideal bond lengths and angles were 0.003 Å and 0.9°, respectively (Table 3-1).

The structure of *T. cru* α-CA contains 4 molecules in the asymmetric unit, which appear to be interacting as two separate dimers (Figure 3-2). Previous assays used to analyze the protein samples were performed under denaturing conditions and provided no information regarding the structure's quaternary structure. Therefore, the crystal structure was the first data to suggest the possibility of a dimer. Size-exclusion chromatography using a Sephacryl S-200 FPLC column, along with DSC studies performed later, strongly supported that the protein is dimeric in solution. The crystal structure of *T. cru* α-CA demonstrates the possibility of an interface between the two dimers, but there is no supportive evidence to suggest that the protein acts as a dimer of dimers. Hence, it is believed that the basic unit of the enzyme is a dimer, and the appearance of a dimer of dimers in the MR solution is a result of crystal packing.

Residues 75 – 304 are observed for all chains in the structure, except for Chain B, which contains three additional residues (Arg305, Asn306, and Ala307, which is supposed to be arginine, but it had no side chain density) at the C-terminus (Figure 3-2). No residues before Pro75 were observed for any of the chains, even after multiple rounds of refinement, which in addition with other computational data suggests that the N-terminus might be disordered (details in discussion).

T. cru α -CA was superimposed onto hCA II (PDB ID: 3KS3) using secondary structure matching in *Coot* with a RMSD of 1.455 Å. Comparison of the *T. cru* α -CA monomer with hCA II shows minor differences in the overall structure, mostly in the length of surface loops, β -strands, and turns (Figure 3-3), but there is a definite structural similarity between both enzymes. Each subunit counts with a zinc molecule tetrahedrally coordinated by His165, His167, His 184 and a water molecule, as is observed for hCAII and most α -CAs. The active site of the *T. cru* α -CA, shown larger in Figure 3-4, superimposes well over the hCA II active site and most of the active site residues are conserved, except for Ala65 and Asn67, which are Thr141 and Gln143 in *T. cru* α -CA, respectively.

The catalytic activity of the *T. cru* α -CA was also assayed and compared to that of hCA II. As mentioned in the Methods section, measurement of the catalytic activity by ^{18}O -exchange yields two rates. The first rate obtained is R_1 (Eq. 2-3), from which we determined $k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$, equivalent to the catalytic efficiency, for the forward direction as it is expressed in Eq. 2-1. The pH profile for this rate constant was fitted to a single ionization event, giving a maximal k_{cat}/K_M of $11 \pm 1 \mu\text{M}^{-1} \text{s}^{-1}$ and an estimated pK_a for the zinc-bound water of 6.6 ± 0.1 . The values obtained for hCA II and *T. cru* α -CA kinetic assays are listed in Table 3-2. The pH profile of *T. cru* α -CA for R_1 appeared qualitatively similar to the pH profile of hCA II for R_1 , as both had a bell-shaped curve with the maximal R_1 near neutral pH, though it can be appreciated that the pH profile of the *T. cru* α -CA is shifted towards the left, obtaining a maximal R_1 at a lower pH than hCA II (Figure 3-5). The pH profile of the catalytic efficiency (k_{cat}/K_M) for both enzymes was also remarkably similar, with a maximal catalytic efficiency at high pH,

corresponding to the reactivity of the zinc-bound hydroxide in the hydration reaction (Figure 3-6). However, there were significant quantitative differences between the kinetic parameters for the *T. cru* α -CA and hCA II. The catalytic efficiency of *T. cru* CA, as measured by $k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$, was reduced 10-fold as compared to hCA II, and the kinetic pK_a , which is an estimated value of the pK_a of the zinc-bound water, determined from the pH profile of $k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$ was lower by approximately one unit (Table 3-2).

The second rate obtained from the ^{18}O -exchange kinetic assays is $R_{\text{H}_2\text{O}}$ (Eq. 2-4), the rate of release of H_2^{18}O from the active site, which is dependent on proton transfer to the labeled zinc-bound hydroxide. Intramolecular proton transfer determines the values for $R_{\text{H}_2\text{O}}/[\text{E}]$, as has been demonstrated by pH profiles, kinetic isotopic effects, and chemical rescue experiments (69, 76). The pH profiles for $R_{\text{H}_2\text{O}}/[\text{E}]$ have a characteristic bell-shaped curve for most of the pH range covered in the studies (Figure 3-7), which is attributed to the transfer of a proton from the PSR His64 to the zinc-bound hydroxide (Eq. 2-2). The solid line in Figure 3-7 represents the fit of Eq. 2-4 to the data. A second, dashed line that appears to fit the data more accurately is the result of a double ionization model used to calculate the kinetic parameters for $R_{\text{H}_2\text{O}}/[\text{E}]$. In order to fit Eq. 2-4 to the data accurately, the pK_a values for the donor and acceptor have to be assigned, which meant using the pK_a value for the zinc-bound water as determined by ^{18}O -exchange for $k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$ in the hydration direction. The fit of Eq. 2-4 to the data yielded a rate constant for intramolecular proton transfer (k_B) of $0.30 \pm 0.05 \mu\text{s}^{-1}$. The rate constant k_B was more similar between *T. cru* CA and hCA II than k_{cat}/K_M (Table 3-2), suggesting that *T. cru* CA has a PSR similar to the side chain of His64 in hCA II for proton transfer. As the pH profile for R_1 , the pH profile of *T. cru* α -CA for $R_{\text{H}_2\text{O}}/[\text{E}]$ was

also shifted toward a lower pH (Figure 3-7). The pK_a of the zinc-bound water and the proton shuttle residue in *T. cru* α -CA are both 6.4 ± 0.1 , which could be a reason why proton transfer is slower than in hCA II.

T. cru α -CA activity was also assayed in the presence of various anions that can be found in the hydrothermal vents to determine whether the anions in the environment modulate its catalytic activity. Inhibition constants of iodide, chloride, and bromide were obtained for *T. cru* α -CA using the ^{18}O -exchange method (Figure 3-8) and compared to the inhibition constants for hCA II reported previously (Table 3-3) (77), demonstrating similar inhibition values for both enzymes. Inhibition by hydrogen sulfide (HS^-), one of the major sulfur species available at hydrothermal vents, was also obtained for *T. cru* α -CA and hCA II (Table 3-3). Interestingly, results show that HS^- actively inhibits both *T. cru* α -CA and hCA II with micromolar affinity, which somewhat conflicts with the sulfur-oxidizing nature of the *T. cru* gammaproteobacterium.

The thermal stability of *T. cru* α -CA was determined at different pH values ranging from 4 to 9 by DSC, and compared to the thermal stability of hCA II at pH 8. A preliminary temperature profile of *T. cru* α -CA was determined using ^{18}O -exchange kinetics by increasing the temperature during the assay, which showed that the thermal inactivation temperature was between 55 and 60°C (data not shown). The melting temperature (T_M) of hCA II, also the thermal inactivation temperature, is 59.5 ± 0.5 °C, and is observed as a single endothermic peak at the midpoint of the DSC curves representative of the main unfolding transition (78). Unlike the scans collected for hCA II, the thermograms collected for the *T. cru* α -CA samples presented two independent transitions, one at approximately 59°C and the other at 72°C, for all pH values at which

data were collected (Table 3-4). We attempted to collect data for *T. cru* α -CA at pH 4 in several occasions, keeping sample concentrations roughly the same, but we were unsuccessful, which led us to believe that the enzyme is unstable at pH 4. It was also determined by completing a reverse scan that only the first transition was reversible, suggesting that the first transition represents the dissociation of the dimer while the second transition represents the unfolding of the *T. cru* α -CA. The thermograms were initially fit to a two-state reversible unfolding model, the simplest model, to obtain van't Hoff enthalpies of unfolding (ΔH_v), but the model did not fit the DSC data accurately (73). This led us to exchange the two-state reversible model for a non two-state reversible unfolding model, indicative of an unfolding intermediate. After fitting the data to a non 2-state transition, the calorimetric enthalpies (ΔH), van't Hoff enthalpies (ΔH_v), and the melting temperatures were calculated and are listed in Table 3-4. The melting temperatures of the first and second transition were plotted as a function of pH (Figures 3-9 and 3-10) to determine if the changes in environmental pH affected the thermal stability of the *T. cru* α -CA. The T_M for the lower temperature transition increased roughly 1 – 2°C per increase in pH unit, while the T_M for the second transition increased approximately 3° throughout the pH profile. A linear relationship could be established between pH and temperature of the first transition, with a correlation coefficient (r) of 0.94. A linear fit was also applied to the second transition, but the r -value is much lower (0.80), which could indicate that there isn't a linear relationship between pH and temperature for the second transition. However, the pH profile for the melting temperatures for both transitions, more so for the first transition, demonstrated that

there is a direct relationship between pH and temperature for *T. cru* α -CA, as the thermal stability of the enzyme was enhanced as pH was increased.

Table 3-1. Data collection and refinement statistics for the crystallographic study of *T. cru* α -CA.

Data collection statistics	
Space group	C2
Unit-cell parameters (\AA , $^\circ$)	$a = 127.1, b = 102.2, c = 105.0$ $\beta = 127.3$
Resolution range (\AA)	20.00 – 2.60 (2.69 – 2.60)*
R_{sym} (%) ^a	10.0 (40.5)
$I/\sigma(I)$	9.2 (2.5)
Redundancy	3.7 (3.5)
Total number of measured reflections	120808
Total number of unique reflections	32859
R_{work} (%) ^b	20.8
R_{free} (%) ^c	24.5
V_M ($\text{\AA}^3 \text{Da}^{-1}$)	2.04
Residue Nos.	Chain A: 75 – 304 Chain B: 75 – 307 Chain C: 75 – 304 Chain D: 75 – 304
No. of Atoms	
Protein	7537
Zn	4
H ₂ O molecules	104
B factors (\AA^2), average	
Main-chain, Side-chain, Zn	Chain A: 28.5, 31.2, 19.6 Chain B: 28.1, 31.2, 18.0 Chain C: 32.2, 35.2, 22.7 Chain D: 32.2, 35.5, 26.3
Solvent	28.5
Ramachandran statistics (%)	
Most favored, Additionally allowed, Generously allowed	Chain A: 91.8, 6.2, 2.1 Chain B: 93.4, 5.6, 1.0 Chain C: 93.3, 6.2, 0.5 Chain D: 93.8, 5.7, 0.5
R.M.S.D. for bond lengths and angles (\AA , $^\circ$)	0.003, 0.9

* Values in parentheses refer to the highest resolution shell.

^a $R_{\text{sym}} = (\sum |I_{\text{hkl}} - \langle I_{\text{hkl}} \rangle| / \sum \langle I_{\text{hkl}} \rangle) \times 100$, where I_{hkl} is the intensity of an individual reflection and $\langle I_{\text{hkl}} \rangle$ is the average intensity for this reflection.

^b $R_{\text{work}} = [(\sum |F_{\text{obs}}| - |F_{\text{calc}}|) / \sum |F_{\text{obs}}|] \times 100$

^c R_{free} is calculated the same as R_{cryst} , except it uses 5% of the reflection data omitted from refinement.

Table 3-2. Comparison of maximal (pH independent) catalytic parameters for *T. cru* α -CA and hCA II.

Parameter	<i>T. cru</i> α -CA	hCA II ^a
$k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$ ($\mu\text{M}^{-1}\text{s}^{-1}$)	11.0 ± 0.1	120
k_{B} (μs^{-1})	0.30 ± 0.05	0.8
$\text{p}K_{\text{a ZnH}_2\text{O}}^1$	6.6 ± 0.1	6.9
$\text{p}K_{\text{a ZnH}_2\text{O}}^2$	6.4 ± 0.2	6.8
$\text{p}K_{\text{a PSR}}^2$	6.4 ± 0.2	7.2

^a(39). Standard errors are no larger than 20%.

¹Determined by ¹⁸O-exchange from calculation of $k_{\text{cat}}^{\text{ex}}/K_{\text{eff}}^{\text{CO}_2}$ in the hydration of CO₂.

²Determined from calculation of $R_{\text{H}_2\text{O}}/[\text{E}]$.

Table 3-3. Comparison of inhibition constants for *T. cru* α -CA and hCA II.

Inhibition constant	<i>T. cru</i> α -CA	hCA II ^a
K _i (Cl ⁻)	361 ± 27	200
K _i (I ⁻)	53 ± 3	26
K _i (Br ⁻)	242 ± 55	-
K _i (HS ⁻)	0.0011 ± 0.0001	0.0028 ± 0.0002

^a(77)

Table 3-4. Thermodynamic parameters of unfolding of *T. cru* α -CA at different pH values.

pH	First transition			Second transition		
	T_M ($^{\circ}\text{C}$)	ΔH (kcal/mol)	ΔH_v (kcal/mol)	T_M ($^{\circ}\text{C}$)	ΔH (kcal/mol)	ΔH_v (kcal/mol)
5	57.7 ± 0.7	130 ± 33	91 ± 17	68.1 ± 0.8	193 ± 34	75 ± 12
6	58.3 ± 0.4	85 ± 15	123 ± 20	69.5 ± 0.4	234 ± 18	69 ± 7
7	59.8 ± 0.5	211 ± 30	104 ± 15	69.9 ± 0.6	189 ± 30	102 ± 23
8	60.1 ± 0.3	129 ± 12	109 ± 11	71.9 ± 0.4	128 ± 13	103 ± 13
9	62.1 ± 0.4	242 ± 25	98 ± 11	71.0 ± 0.5	119 ± 24	132 ± 28
Control (hCA II, pH 8.0)	-	-	-	55.95 ± 0.05	170 ± 4	181 ± 5

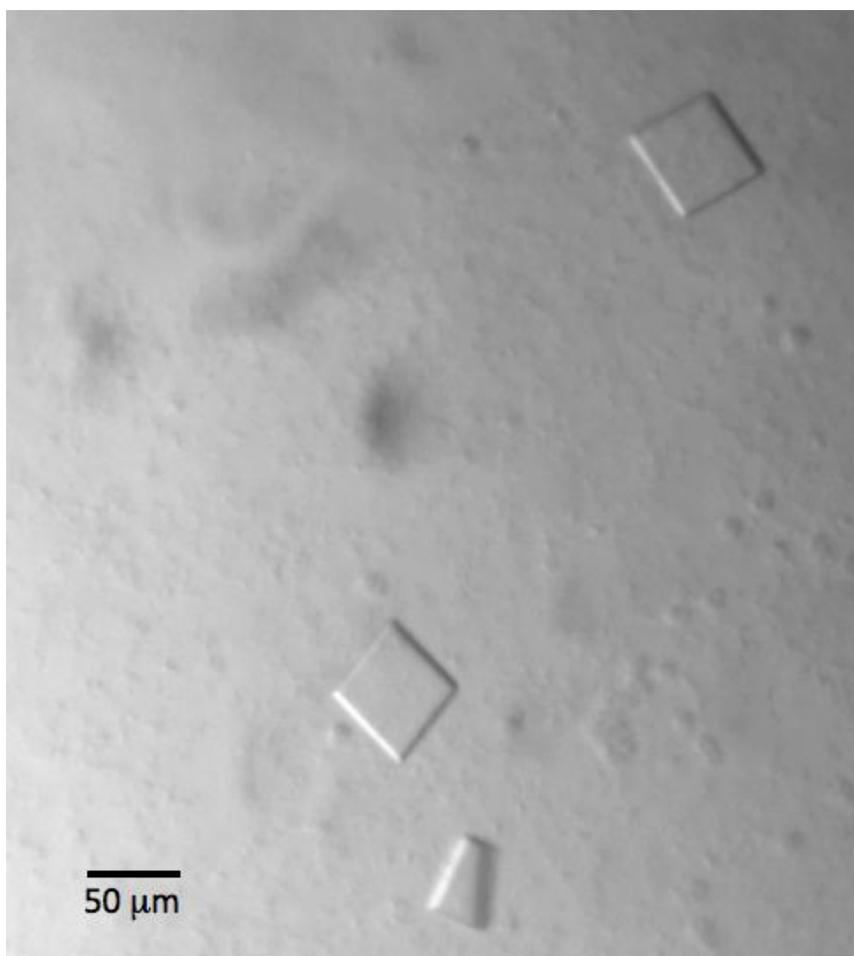


Figure 3-1. Crystals of *T. cru* α -CA grown in 2% v/v tacsimate pH 4.0, 0.1 M sodium acetate trihydrate pH 4.6, 16% w/v PEG 3350 at 17°C using the hanging-drop vapour-diffusion method. The crystal dimensions are approximately 0.05 x 0.05 x 0.01 mm³.

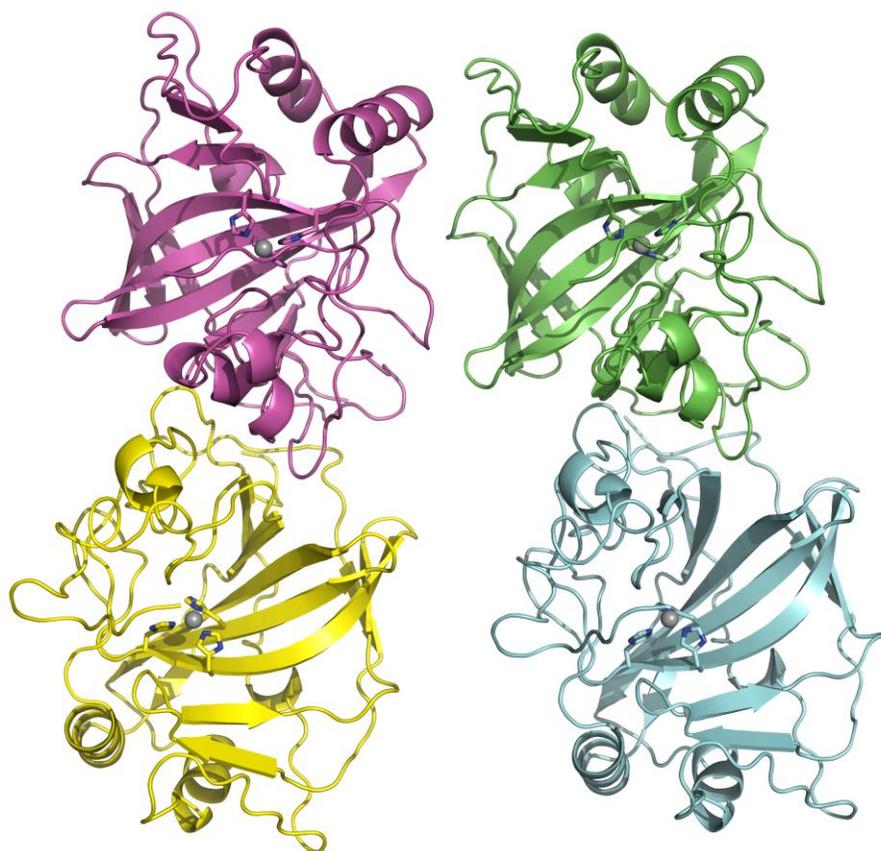


Figure 3-2. Crystal structure of *T. cru* α -CA. The crystal contained four molecules in the asymmetric unit (au). Cartoon representation of *T. cru* α -CA, with each chain within the au presented in a different color (A - purple, B - yellow, C - green, and D - blue). Chains A and B interact to form a dimer, as well as chains C and D. The zinc ions are presented as gray spheres.

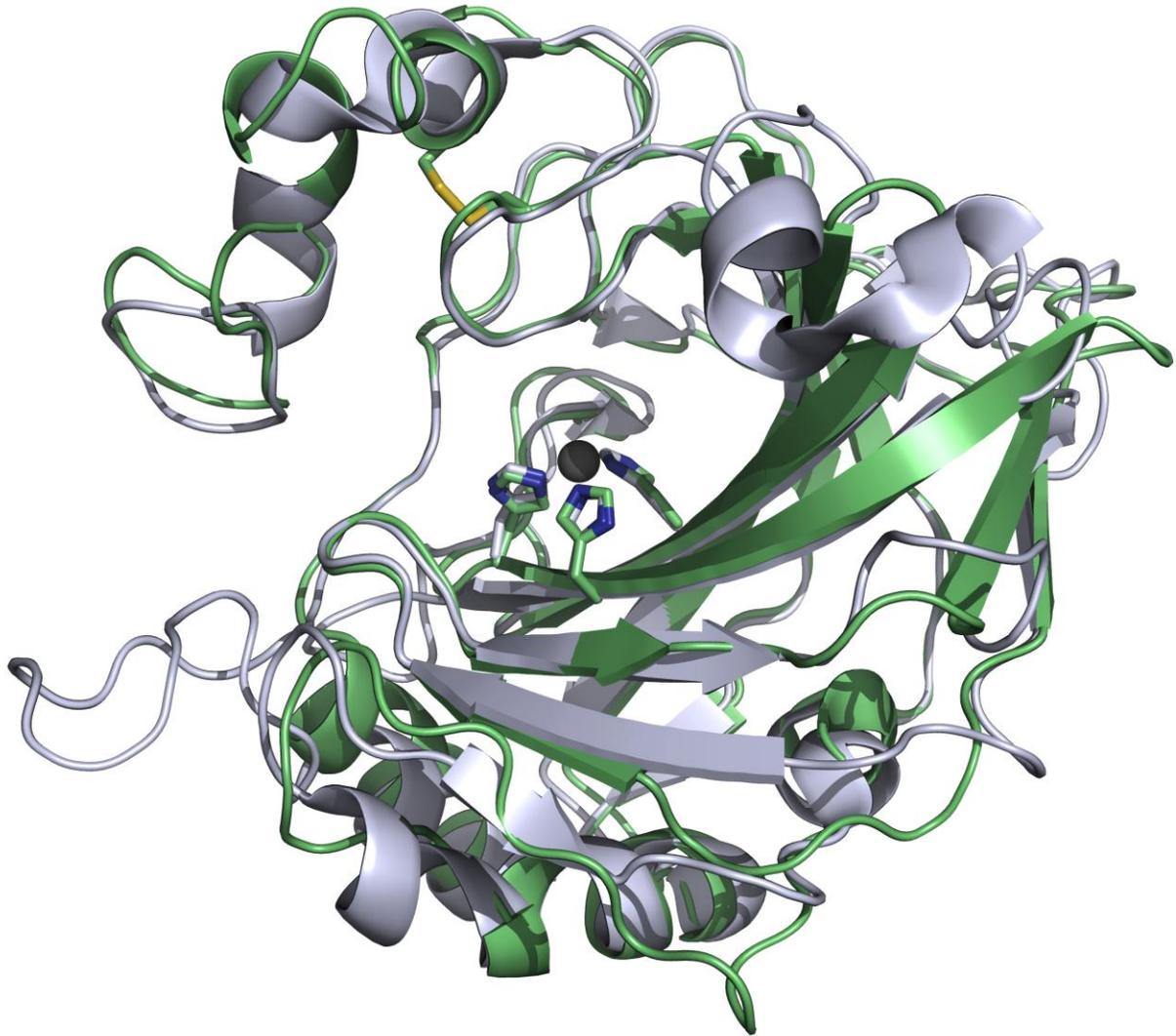


Figure 3-3. Superposition of the *T. cru* α -CA (green) monomer and the hCA II (gray) crystal structure (PDB ID: 3KS3; 54). The cartoon representation overlaps to show the similarities and differences between both structures, with an RMSD of 1.455 Å. The coordinating histidines and the disulfide bond are presented as sticks. The nitrogen is shown in navy blue, sulfur in yellow, and the zinc ion is shown as a gray sphere.

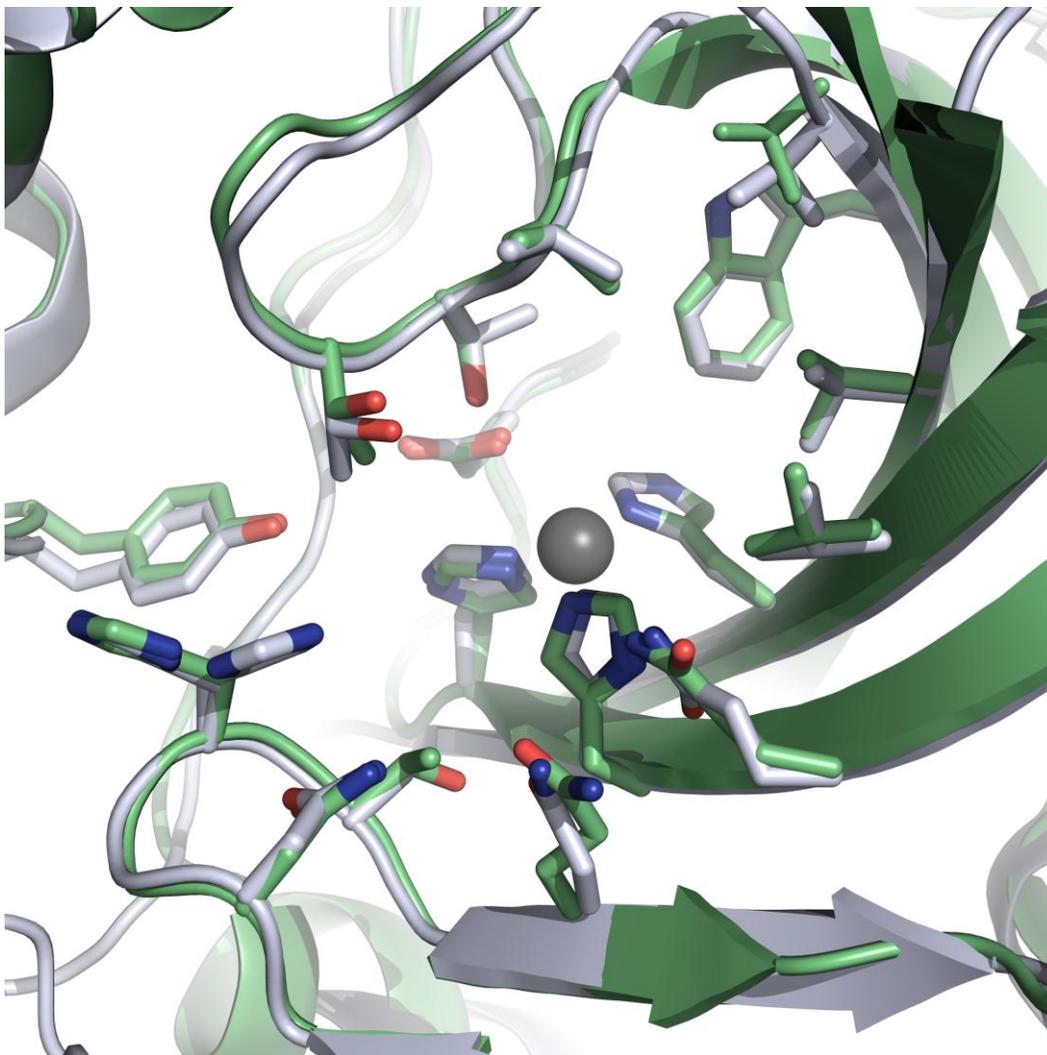


Figure 3-4. Close-up stick representation of superimposed active site residues in *T. cru* α -CA (green) and hCA II (gray) (PDB ID: 3KS3; 54). Oxygen atoms are shown in red, nitrogen atoms in navy blue, and the zinc ion is represented as a gray sphere.

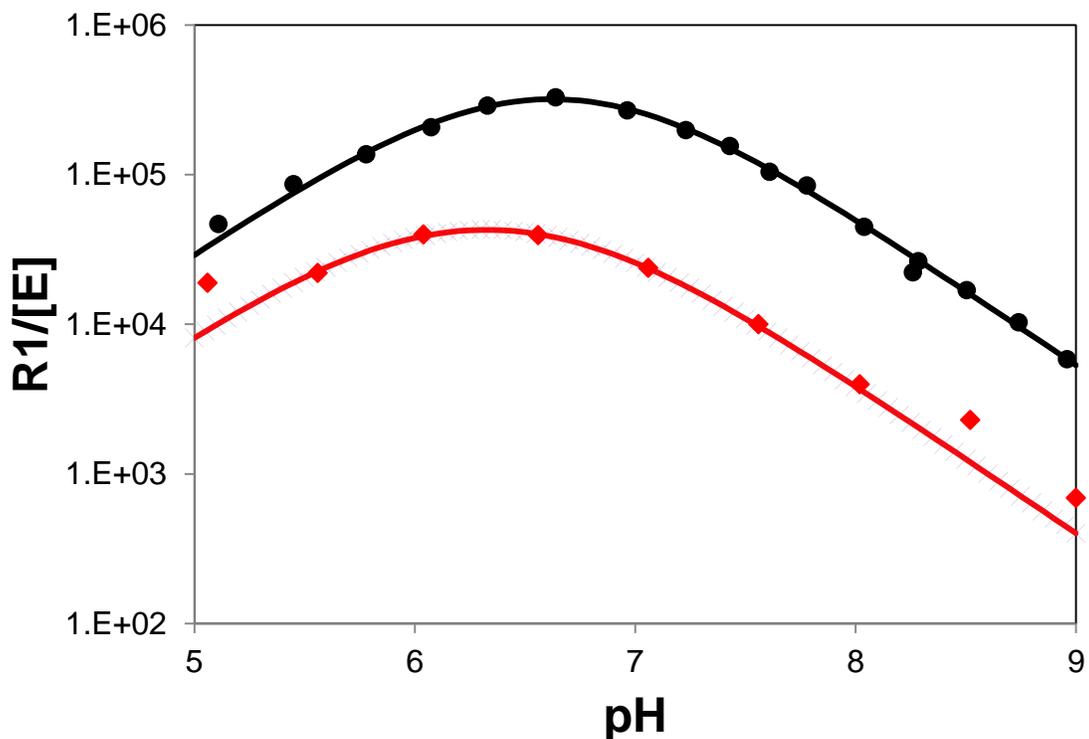


Figure 3-5. The pH profiles for R_1 for the hydration of CO_2 catalyzed by *T. cru* α -CA (red, \blacklozenge) and hCA II (black, \bullet). Data were collected by ^{18}O -exchange by CO_2 and water measured at 25 °C. Total concentration of all species of CO_2 was 25 mM and sodium sulfate was added to maintain ionic strength at 0.2 M. The solid lines are a fit of Eq. 2-3 to the data.

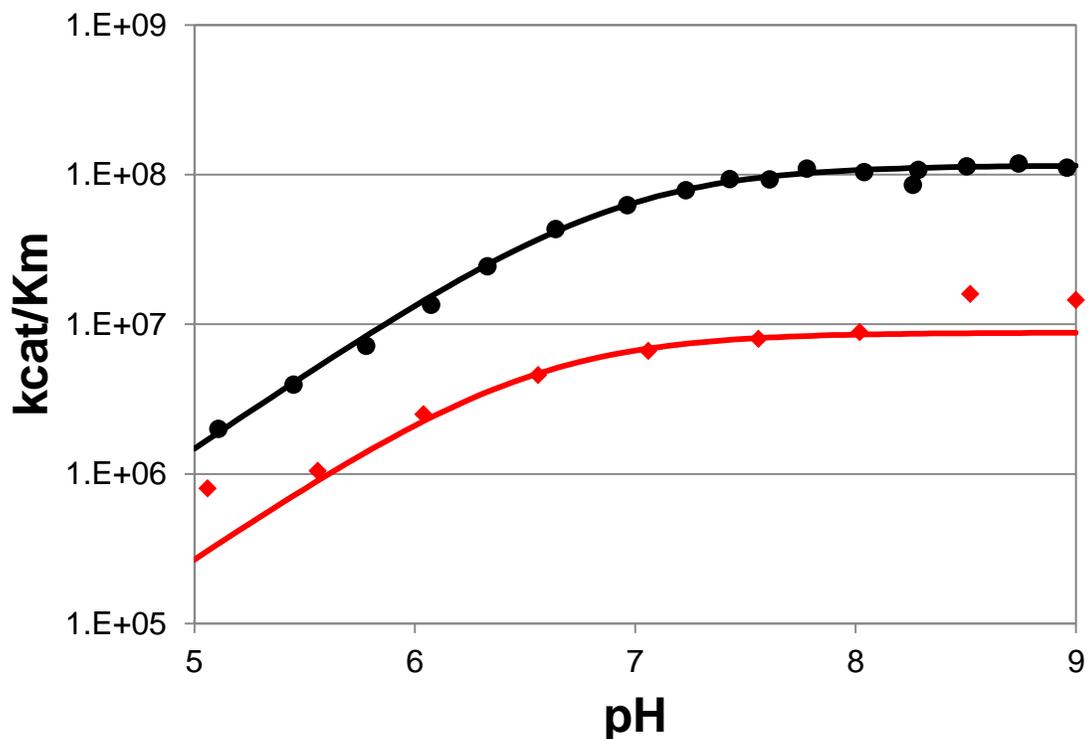


Figure 3-6. The pH profiles for $k_{cat}^{exch}/K_{eff}^{CO_2}$ ($M^{-1}s^{-1}$) for the hydration of CO_2 catalyzed by *T. cru* α -CA (red, ◆) and hCA II (black, ●). The same experimental conditions were used as in Figure 3-5. The solid lines are a fit of a single ionization model to the data.

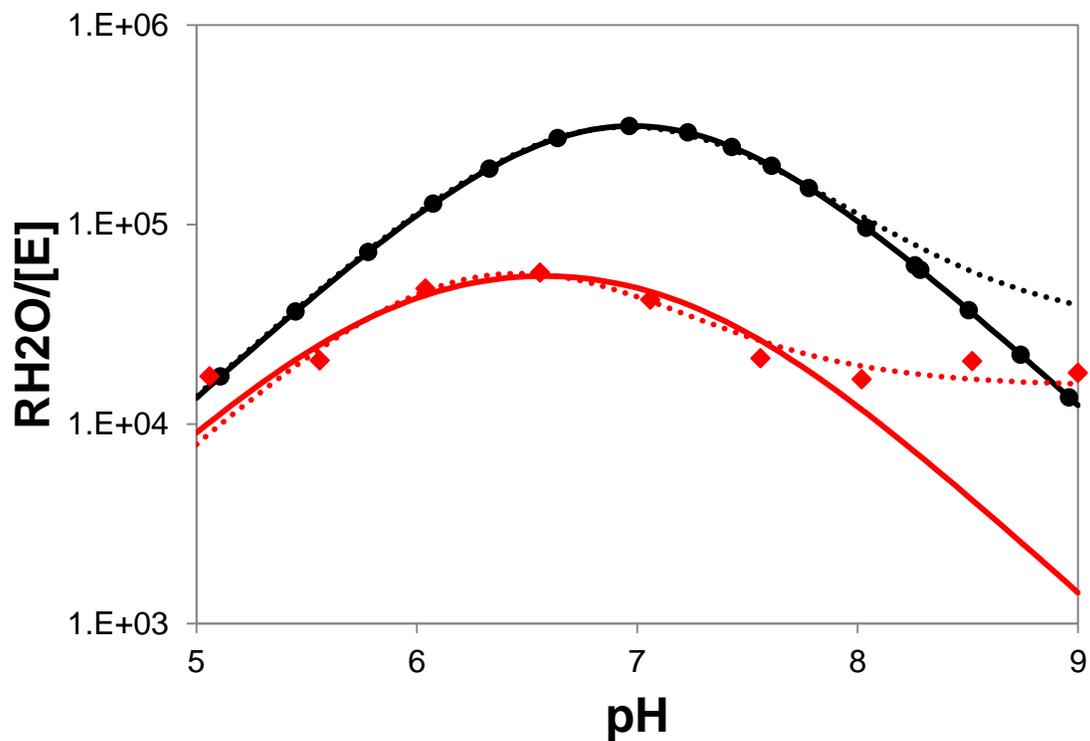


Figure 3-7. The pH profile for $R_{H_2O}/[E]$ (μs^{-1}) for proton transfer in the dehydration direction catalyzed by *T. cru* α -CA (red, \blacklozenge) and hCA II (black, \bullet). The same experimental conditions as in Figure 3-5 were used. The solid lines are a fit of Eq. 2-4 to the data, and the dashed lines are a fit of a double ionization model to the data.

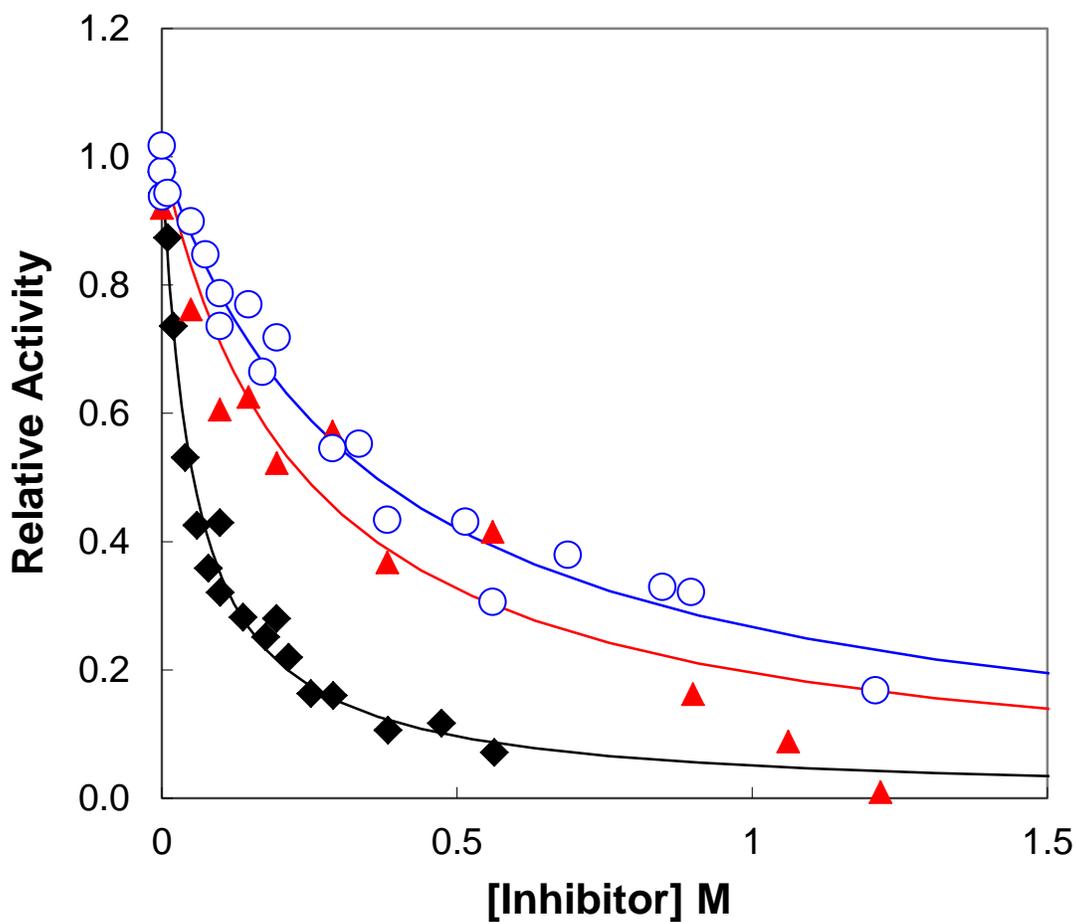


Figure 3-8. The inhibition of *T. cru* α -CA by iodide (black, \blacklozenge), chloride (blue, \circ), and bromide (red, \blacktriangle). Total concentration of all species of CO_2 was 25 mM in a solution containing 100 mM HEPES at pH 7.6 and 25 °C. Na_2SO_4 was added to maintain ionic strength at a minimum of 0.2 M.

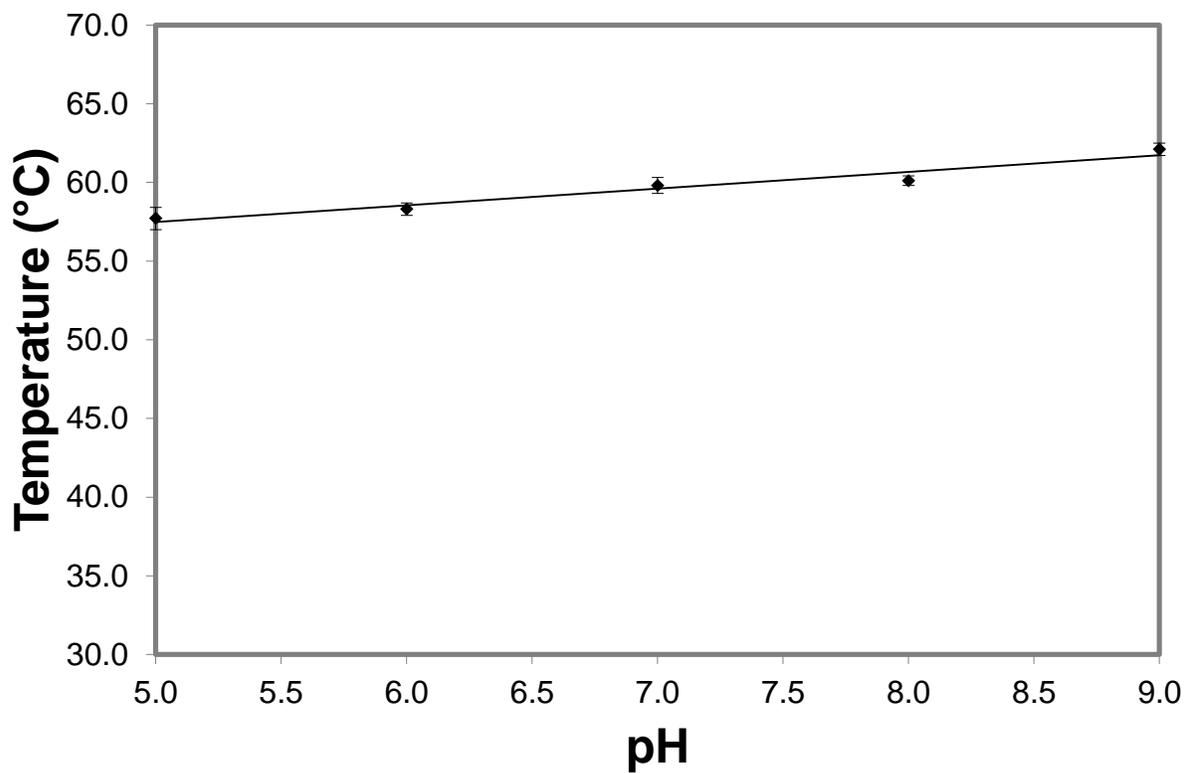


Figure 3-9. The pH profile of the melting temperature (TM) of the dissociation of dimeric *T. cru* α -CA. DSC thermograms were collected at pH range 5 – 9 for *T. cru* α -CA in triplicate. After data processing, the TM were averaged and plotted as a function of pH. The solid line is a linear fit to the data with a correlation coefficient (r) of 0.94. Error bars are shown.

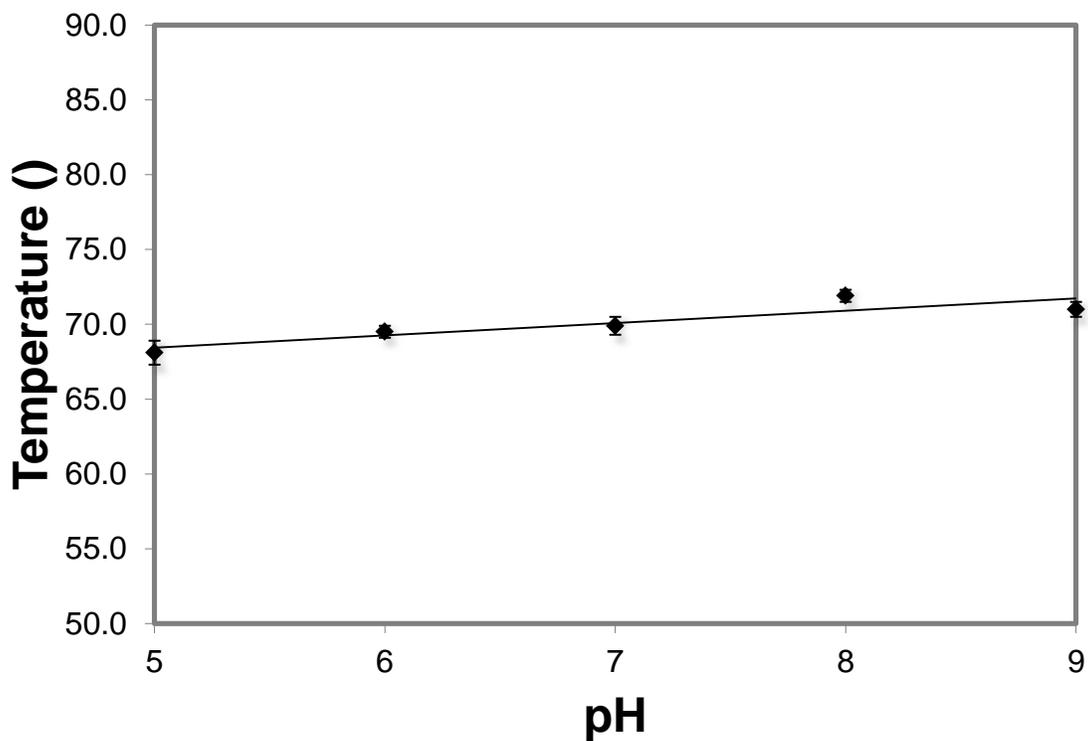


Figure 3-10: The pH profile of the melting temperature (TM) for the unfolding of dimeric *T. cru* α -CA. DSC thermograms were collected at pH range 5 – 9 for *T. cru* α -CA in triplicate. After data processing, the TM were averaged and plotted as a function of pH. The solid line is a linear fit to the data with a correlation coefficient (r) of 0.80.

CHAPTER 4 DISCUSSION

The goal of this research project was to characterize the α -class CA expressed in the *T. cru* gammaproteobacterium by structural, kinetic, and biophysical means to understand its physiological role. Measurement of ^{18}O -exchange at chemical equilibrium by membrane inlet mass spectrometry was used to obtain the kinetic rates of catalytic activity for *T. cru* α -CA, which demonstrated that it is a fast enzyme, and differential scanning calorimetry was used to determine the increased thermal stability of the *T. cru* α -CA as compared to hCA II. X-ray crystallography allowed us to determine the molecular structure for the *T. cru* α -CA, which provided structural evidence supporting its fast activity and thermal stability, and revealed the possible dimeric nature of the protein. The interpretation and combination of the results obtained with these techniques, along with the comparison to various well-studied α -CA isoforms, particularly hCA II, has led us to establish a structure-function relationship of the *T. cru* α -CA that can provide further insight to its physiological role in *T. cru*.

Crystal Structure of *T. cru* α -CA

The crystal structure obtained for the *T. cru* α -CA contained four molecules in the asymmetric unit, with 16 molecules total per unit cell (Figure 4-1). The different chains observed were labeled A, B, C, and D. Each chain contains 230 amino acids, from residue 75 to 304 of the full-length sequence, except for chain B, which contains 3 additional residues at the C-terminus. Superimposition of chains B, C, and D to chain A, which we selected as the reference chain, showed an RMSD for the superimposition of the corresponding C α atoms (residues 75 to 304) of 0.229, 0.228, and 0.243 Å, respectively. Given that the molecules show no significant structural differences, any

discussion regarding the monomeric unit of *T. cru* α -CA can be applied to all chains in the asymmetric unit.

The electron density of some polar residues at the surface of the molecules is weak or unobserved, but the density for the main chain is clearly observed in the final $2F_o - F_c$ map and the density-modified map created using AutoBuild. There were significant peaks in the Fourier difference map at the active site, indicating the presence of the zinc ion and the zinc-bound water molecule. The model contains 104 water molecules, with an average B-factor of 28.4 \AA^2 . Due to the resolution of the data, only a few water molecules could be added with absolute certainty, resulting in a reduced number of solvent molecules. This affects our ability to make observations regarding the water network in *T. cru* α -CA, which potentially be used to explain differences in the catalytic activity in comparison to hCA II.

The geometry of the final model was verified using *PROCHECK* (66). The RMSD values from the ideal bond lengths and angles were within the acceptable limits. The ramachandran plot generated by *PROCHECK* (Figure 4-2) shows that over 90% of the dihedral angles were within the most favored region, while the rest were in the allowed region, except for 1.0% which were in the generously allowed region. Furthermore, the average B-factors of the main-chain and side-chain atoms are 30.3 \AA^2 and 33.3 \AA^2 , respectively, with an average B-factor for the zinc atom at 21.6 \AA^2 , suggesting that the structure is correctly built.

The *T. cru* α -CA monomer has a roughly ellipsoidal shape of approximately $44 \times 40 \times 41 \text{ \AA}^3$ in size. As can be observed in Figure 4-3, the core of the molecule is formed by a ten-stranded antiparallel β -sheet, a fold characteristic of α -CAs. The active site of

the protein is located in a large conical cavity extending from the surface of the molecule to the center, with the zinc ion tetrahedrally coordinated by His165, His167 and His184, along with a water molecule, found at the base of the cavity (Figure 4-4). The active site is partitioned into two different environments, a hydrophobic patch formed by residues Val186, Val196, Leu251, Val260, and Trp262, corresponding to the hydrophobic patch in hCA II, and an area lined by hydrophilic residues Tyr80, Asn138, His140, Thr141, Gln143, Thr252, and Thr253. Two residues that are part of the hydrophilic patch in hCA II, Ala65 and Asn67, are not conserved in the active site of *T. cru* α -CA; instead, they are replaced by polar residues threonine (Thr141) and glutamine (Gln143) respectively, conserving the hydrophilic nature. Therefore, the same dual nature that characterizes the α -class CAs active site is maintained in the *T. cru* α -CA. Furthermore, the zinc-bound water and the water molecule termed “deep-water”, which hydrogen bonds to the amide nitrogen of Thr199 in hCA II (79), are observed in the active site of *T. cru* α -CA.

Superimposition of *T. cru* α -CA with hCA II highlights the structural differences between both enzymes (Figure 3-3), which include changes in secondary structure elements, differences in surface loop lengths, and displacement of certain loops and residues. Residues pertaining to hCA II are noted in parenthesis. One of the first differences noted is the difference in secondary structural elements. As mentioned previously, the α -CA fold formed by the ten-stranded antiparallel β -sheet is observed in *T. cru* α -CA and superimposes well over the hCA II core, along with a few other short β -strands present in the structure. However, there are some variations regarding the lengths and positions of some of the β -strands, and other strands in hCA II are not

identified as such by the secondary structure assignment program STRIDE (67). For example, the β -strand formed by residues Lys(39) and Tyr(40) in hCA II, which is part of the 10 β -strands that form the core, is present as a loop in *T. cru* α -CA. However, a β -strand formed by residues Phe177 – Asn178 found in *T. cru* α -CA is classified as a loop in hCA II. Also, a short β -strand formed by residues Ser(173) – Asp(175) in hCA II is absent in *T. cru* α -CA, but a larger β -strand not observed in hCA II containing residues Ile226 – Lys231 appears next to where the missing β -strand is in hCA II. Furthermore, two β -strands that are conserved between both proteins appear to be displaced in the crystal structures. The β -strand formed by residues Phe124 – Tyr127 is displaced 2.4 Å from the corresponding β -strand in hCA II formed by residues Leu(47) – Ser(50). Another β -strand containing residues Tyr151 – Ile154 is found 2.0 Å from the similar β -strand in hCA II, with residues Val(78) – Leu(81). However, none of the displacements or changes in β -strands disrupt the orientation or position of the active site residues.

The helical content of *T. cru* α -CA is not conserved as well as the β -sheet when compared to hCA II. The primary differences in the helical regions are the missing helices in *T. cru* α -CA. A helix containing residues Thr(125) – Asp(139) in hCA II is not observed in *T. cru* α -CA, and instead is replaced by a shorter, hairpin loop containing residues Asp189 – Gly192. Some residues in this helix have been implied in stabilizing the strain observed in the loop containing the cis-Pro(202) and catalytic residue Thr(199) in hCA II (80). Loss of this particular helix could result in a decrease of conformational stability. However, *T. cru* α -CA contains a disulfide bridge that stabilizes the Thr(199) loop. The disulfide bond will be discussed later in further detail. A shorter helix, composed of residues Pro(181) – Leu (184), is also missing from the *T. cru* α -CA

structure. Besides, the missing helices, other differences included displacement of α -helices and variations in length, similar to the β -strands. Most notably, the helix formed by residues Val273 – Leu283 is displaced 2.3 Å from the corresponding helix consisting of residues Ser(220) – Arg(227) in hCA II.

The largest variability observed between the structures of *T. cru* α -CA and hCA II is due to the difference of surface loops, particularly loops that are shorter or absent in *T. cru* α -CA. In *T. cru* α -CA, there is only one small, hairpin loop with three residues (Ala93, Pro94, Glu95) added to the N-terminal segment in *T. cru* α -CA, joining two short helices, but missing in hCA II, which instead has the two helices connected by residues Phe20 and Pro21. However, most of the differences in surface loops are observed as deletions in *T. cru* α -CA. The major deletions in *T. cru* α -CA are the surface loops formed by residues Gly(98) – Gln(103) and Lys(228) – Met(241), which are extended from behind the active site out to the surface and near the opening of the active site cavity, respectively. The loop Gly(98) – Gln(103) is close to the helix formed by Ser(220) – Arg(227). Deletion of the loop could account for the displacement of the helix Val273 – Leu283, which corresponds to the helix Ser(220) – Arg(227) in hCA II, as the helix shifted 2.3 Å toward the space occupied by the loop present only in hCA II. Deletion of the surface loop containing residues Lys(228) – Met(241), which is located near the opening of the active site, could also account for the displacement of helix Val273 – Leu283. Also, all the surface loop deletions, along with the deletion of helices Thr(125) – Asp(139) and Pro(181) – Leu(184), result in the appearance of a more compact α -CA with a more solvent-accessible active site (Figure 4-5).

The N-terminus of *T. cru* α -CA

The gene for *T. cru* α -CA encodes a full-length sequence of 315 amino acids. However, out of the 315 residues encoded, only 292 amino acid residues should be observed. *T. cru* α -CA was identified as a periplasmic protein, as its sequence contains the signal peptide NVAAP at the N-terminus, identified by Signal P software, and its periplasmic location was previously confirmed (46, 52, 53). However, even after cleavage of the signal peptide, 64 residues are still unaccounted for in the crystal structure, 52 of which are at the N-terminus and the rest located at the C-terminus.

Before we were able to crystallize the protein and obtain the 3-D structure, we used sequence alignments and homology modeling with hCA II to have a model to work with and establish our hypotheses regarding catalysis and thermal stability (Figure 4-6). In the homology model we obtained, the 52 amino acids at the N-terminus were not taken into consideration to build the model; the N-terminus was missing. We used the NCBI Protein Blast to search the databases for possible similar regions between all available protein sequences, including those without reported 3-D structures, and the 52 missing residues at the N-terminus of *T. cru* α -CA. We found that the sequence (MAAPLIDLGAELAKKQAQKSAATQSAVPEKESATKVAEKQKEPEEKAKPEPKK) is unique among known protein sequences to date. The N-terminal region appears to be of hydrophilic nature, as it contains 20 charged residues, among them 11 lysines, 8 glutamates, and one aspartic acid, and 9 polar residues, which are four glutamines, three serines, and two threonines. The 11 lysines could confer a more basic, rather than acidic nature to the N-terminus, which could potentially be the reason why we observe a lower pK_a for the zinc-bound water in *T. cru* α -CA. The lower catalytic efficiency of *T. cru*

α -CA that we observed as the result of the kinetic analysis may therefore be a result of decreased nucleophilicity of the zinc-bound hydroxide due to the more acidic pK_a .

After obtaining the crystal structure, and noticing that the 52 residues at the N-terminus were missing, we hypothesized that the missing segment could be intrinsically disordered. Intrinsically disordered regions in proteins have been associated with biological functions regarding signaling, molecular recognition, and regulation (81, 82). We used the meta-predictor for intrinsically disordered residues PONDR-FIT to predict the disordered regions of *T. cru* α -CA (81). Figure 4-7 demonstrates that some residues of the missing N-terminal region (Ser46 – Val48, and Ala54 – Glu63), and the residues 305-315 at the C-terminus missing from the crystal structure, cross the dashed lines at 0.5 of the Y-axis, and are therefore predicted as disordered by PONDR-FIT. Therefore, although the region is not completely disordered, the various stretches that are disordered within the N-terminus could increase the flexibility of that region, leading to the absence in the crystal structure.

The Disulfide Bond

The monomers in the *T. cru* α -CA structure show an intramolecular disulfide bond between Cys99 and Cys256, the only two cysteines in each monomer (Figure 4-3). The disulfide bond was formed spontaneously, most likely during the oxidative folding process in the periplasm of *E. coli* (83). There is little to no information available on the oxidative folding of *T. cru*, but it is expected that the gammaproteobacterium have the machinery and conditions necessary to form disulfide bonds as part of protein folding. This disulfide bond connects a helix at the N-terminal segment with a loop containing Thr252, which is part of the hydrogen bond network essential for CO₂ hydration (84). Superimposition of the active site of *T. cru* α -CA with hCA II (Figure 3-4) reveals that

Thr252, corresponding to Thr199 in hCA II, along with the other residues in the loop are aligned almost perfectly. Protein folding studies of hCA II have demonstrated that the conformation of the 198-206 loop is strained at Pro202, due to the adoption of a cis-conformation that could potentially be stabilized by interactions with the residues at the 123-139 helix (80, 85). However, in *T. cru* α -CA, the segment corresponding to the 123-139 helix in hCA II is replaced with a short loop. Therefore, the presence of the disulfide bond could potentially stabilize the loop with residues Thr252 and Pro254.

The same intramolecular disulfide bond corresponding to the linkage by Cys99 and Cys256 in *T. cru* α -CA has been observed in other α -CA monomers, such as hCA IV, hCA VI, hCA IX, hCA XII, NGCA, and AoCA (12, 86-89, 100). The presence of the disulfide bond in hCA IV was proved to confer the decreased susceptibility of the protein to denaturation by 5% SDS (90, 91). Guanidine-HCl (GdnHCl) induced denaturation experiments performed with NGCA also demonstrated that the inactivation of NGCA occurred at lower concentrations of GdnHCl upon addition of the reducing agent tris(2-carboxyethyl)phosphine (TCEP), decreasing from 2.1 M GdnHCl to 1.2 M GdnHCl, which is comparable to the inactivation of hCA II by GdnHCl (at 0.9 M GdnHCl) (92, 93). The reduction of the disulfide bond in NGCA led to decreased conformational stability, making it more sensitive to denaturation/inactivation by GdnHCl. It was inferred in subsequent papers regarding other α -CAs that the presence of the disulfide bonds accounted for increased conformational stability. Disulfide bonds in other globular proteins, such as Ribonuclease I, have also been linked to increased stability when compared to the reduced proteins (94-96). Furthermore, a study using an hCA II variant with an engineered disulfide bond at residues Cys23 and Cys203, equivalent to the

disulfide bond observed in the aforementioned α -CA monomers and in *T. cru* α -CA, dramatically increased the conformational stability of hCA II by 3.7 kcal/mol, as calculated from an increase in the unfolding midpoint from 0.9 to 1.7 M GdnHCl (97).

DSC studies were performed as part of this study to determine the thermal stability of *T. cru* α -CA. From the scans we collected at pH 5-9, we observed that the melting temperature (T_M) of the α -CA went from 68.1 to 71.9°C as the pH was increased, establishing a linear relationship between pH and temperature. The melting temperature of hCA II, also equal to its thermal inactivation temperature, is 59.5 ± 0.5 °C at pH 7.8 (78). However, we decided to set up a control of hCA II in the same buffer as *T. cru* α -CA, in order to allow direct comparison at pH 8.0. Our results show that at pH 8.0, *T. cru* α -CA has a T_M of 71.9 ± 0.4 °C, whereas hCA II has a T_M of 55.96 ± 0.05 °C. The 15.94°C difference between both T_M 's demonstrates that *T. cru* α -CA has greater thermal stability than hCA II. Furthermore, the linear relationship between pH and temperature is not observed for hCA II (data not shown). Thermostability has long been associated with the structural and conformational stability of proteins, along with amino acid composition, hydrogen bonding and solvent accessibility (98, 99). By obtaining a larger T_M , *T. cru* α -CA is more thermodynamically stable, and therefore has greater conformational stability than hCA II.

Comparison of hCA II and *T. cru* α -CA structures revealed that there are minor structural differences, particularly shorter/missing surface loops and small changes in secondary structure. Having shorter surface loops provides *T. cru* α -CA with a slightly more compact structure, which could certainly enhance the folding stability. There could be other factors influencing the enhanced stability observed in *T. cru* α -CA, such as bulk

solvent interactions and hydrogen bonding networks, as well. However, we believe that the disulfide bond present between Cys99 and Cys256 is the major contributor of the increased conformational stability of *T. cru* α -CA.

The Dimeric Interface

The crystal structure for *T. cru* α -CA contained 4 molecules in the asymmetric unit, which appeared to be interacting in pairs as a dimer or possibly a dimer of dimers. In order to determine the multimeric nature of the protein, we proceeded to use size-exclusion chromatography to estimate the molecular weight of *T. cru* α -CA. The protein eluted with an approximate molecular weight of 70 kDa, two times the molecular weight of the monomer, suggesting that the protein is a dimer in solution. Further evidence pointing to a dimeric interface was obtained from the DSC studies, which revealed two transition events, at 58°C and 71°C, representing the dissociation of the dimer and the unfolding of the monomers, respectively.

T. cru α -CA forms a homodimer in solution, (Figure 4-8), deviating from most α -class CAs, which are mainly monomeric. Crystal structures have been reported for several other α -CA isoforms that are also homodimeric in nature, which are the human CA VI, CA IX, and CA XII, the α -CA in the unicellular green alga *Chlamydomonas reinhardtii* (Cr- α CA1), and the fungal α -CA expressed in *Aspergillus oryzae* (AoCA) (87-89, 100, 101). For the purposes of this discussion, we will compare the dimer interface of *T. cru* α -CA to the dimer interface of each of the previously listed α -CAs. We used the Protein Interfaces, Surfaces and Assemblies (PISA) server to analyze and compare all possible interfaces of *T. cru* α -CA, in order to discriminate between real interfacing structures and interfaces resulting from crystal packing (102). We also used the PISA

server to identify the interacting residues at the dimer interfaces of *T. cru* α -CA and all other dimeric α -CAs presented in this discussion.

Table 4-1 presents all possible interfaces present in the crystal structure of *T. cru* α -CA (Figure 4-9). The PISA server provides a list of parameters that ascertain the likelihood of a real interface interaction. For example, PISA provides the symmetry operators of each interface, which describe the symmetry operation applied to one of the monomers in order to form the designated interface. In other words, the symmetry operators state how the monomer is moved in order to form the interface. From the results stated in the table regarding the symmetry operators, it is clear that the interfaces between chains B – D, A – D, and B – C are not real interacting interfaces, as the interfaces created by these monomers are formed after applying a symmetry operator different to x, y, z.

Other parameters used to discriminate between real and artificial interfaces are the gain in solvation free energy (ΔG_{solv}) upon formation of the interface, the P-value assigned to the ΔG_{solv} , and the complexation significance score (CSS). The P-value assigned to ΔG_{solv} can basically be regarded as a measurement of interface specificity; the lower the P-value is, the greater the probability that the interface is interaction-specific. The CSS value, on the other hand, indicates how significant the interface interactions are for the association of the dimer on a scale of 0 – 1, where 0 implies that the interface is a result of crystal packing and 1 implies that the interface plays an essential role in complex formation. The results obtained from PISA demonstrate that the largest positive protein affinity is at the interfaces formed by chains A – B and C – D, which results in a more negative ΔG_{solv} when the interface is formed. Also, the lowest P-

values are observed for interfaces A – B and C – D. Furthermore, the CSS values for the interfaces A – B and C – D are near 1, indicating that the interactions at these interface are significant for the dimer association. On the other hand, the interface between A – C, while obtaining modest values for ΔG_{solv} and the P-score, has a CSS value of 0.000, indicating that the interface is only a result of crystal packing, which is supported by the data obtained by gel filtration and DSC that there is only a dimer present in solution, and not a dimer of dimers.

We also wished to analyze the residue-specific interactions at the dimer interface of *T. cru* α -CA, and compare these to the residue-specific interactions observed at the dimer interface of the other α -CAs in order to determine if the same interface is observed in any other isoform. Table 4-2 lists the hydrogen bond interactions observed at the *T. cru* α -CA dimer interface. Note that the table only refers to chains A and B in the crystal structure, but the same residue-specific interactions are also observed at the interface of dimer C – D. Out of the 18 hydrogen bond interactions observed at the dimer interface, 11 of them involve ionizable residues, which are Asp304, Lys246, and Glu258. Increasing the pH of the protein sample could ionize all three amino acids and strengthen the interaction formed by these residues. Although the pK_a of lysine is around 10 when exposed to free solvent, there is reason to believe that the actual pK_a of Lys246 can be lower, due to the microenvironment formed by surrounding residues (Figure 4-10). Microenvironments formed in active sites, protein pockets, or at interfaces can dramatically affect a residue's pK_a (103). In *T. cru* α -CA, the ionizable residues are found buried at the dimer interface, where Lys246 is almost completely buried (90%), and Asp 304 and Glu258 are 50% and 70% buried, respectively. Interactions involving

surrounding residues, some of which are hydrophobic, could decrease the pK_a of Lys148, effectively deprotonating the Lys148 upon increase of the overall pH. Therefore, involvement of ionizable groups at the dimeric interface can explain the increase in the T_M of the dissociation of the dimer, as the hydrogen bond interactions at the interface are strengthened when pH is increased from 5 to 9.

Hydrophobic residues were also found to interact at the dimer interface. The major hydrophobic residues involved in residue-specific interactions at the interface are Val116, Leu121, Gly259, Val260, and Leu299. Interestingly, Val260 is also part of the hydrophobic patch at the active site of *T. cru* α -CA, corresponding to Val(207) in hCAII, and is involved in CO₂ binding (2, 104). The side-chain of Val260 is displaced 1.2 Å away from the active site toward the CO₂ binding site, possibly due to its involvement at the dimer interface (Figure 3-4). The displacement of Val260 could potentially affect the CO₂ binding orientation, which could decrease the effectiveness of the nucleophilic attack by the zinc-bound hydroxide. Therefore, the interaction of Val260 at the dimer interface could explain the reduced catalytic efficiency of *T. cru* α -CA.

We compared the dimer interface of *T. cru* α -CA to the interface of the other dimeric α -CAs by superimposition of the dimeric structures and PISA analysis (Table 4-3). The results obtained from PISA show that unlike the other α -CAs, the interactions at the dimer interface of *T. cru* α -CA dictate the formation of the dimer. However, the lower CSS scores presented by the other α -CAs do not imply that the association of the dimers is a result of crystal packing, as they have been proved to be dimeric in nature, but that the interface of the dimers is of a more hydrophilic nature (102). Except for hCA IX, all dimeric α -CAs compared in this study exceed the 860 Å² cutoff value of surface

area buried per monomer used to discriminate between monomers and homodimers (105). However, hCA IX, along with Cr- α CA1, has an intermolecular disulfide bond at Cys41 bridging the two monomers together. Therefore, a reason why these have CSS values closer to 0, while still being classified as dimers, is because the assembly of the complex depends greatly on the formation of the intermolecular disulfide bond.

We compared the interfaces of all dimeric α -CAs by superimposition of the structures on *T. cru* α -CA using *Coot* (Table 4-4). *Coot* was not able to superimpose both monomers for any of the dimeric α -CAs onto *T. cru* α -CA, superimposing instead only one of the monomeric units for each structure. Interestingly, we found that *T. cru* α -CA forms a unique dimer interface, which is why *Coot* was not able to accurately superpose the dimeric unit without distorting the secondary structures. The *T. cru* α -CA dimer interface is formed mostly by interactions of residues at the N- and C-terminus, with each monomer related by a 2-fold symmetry axis at the interface. Figures 4-11 to 4-15 illustrate the dimer interface of hCA VI, hCA IX, hCA XII, Cr- α CA1, and AoCA in relation to the interface of *T. cru* α -CA. Additionally, all residues forming part of the interface of the dimeric α -CAs are listed in Table 4-5.

As can be observed in the figures, all other dimeric α -CAs demonstrate different residue-specific interactions involved in the association of the dimer. Furthermore, superimposition of all dimers onto *T. cru* α -CA revealed that none of the dimeric α -CAs share the same interface (Figure 4-16). The dimeric α -CA interface that most closely resembles the orientation of the dimer association of *T. cru* α -CA is hCA IX, where even some residues appear to overlap in the interface. However, most of the residues that

are involved in the dimer association of hCA IX do not form part of the interface of *T. cru*
 α -CA.

Table 4-1. Possible interfaces observed in the crystal structure of *T. cru* α -CA as determined by PISA (102).

Interfaces	Symmetry operator ^a	Interface Area ^b , Å ²	No. Interfacing Residues ^c	$\Delta G_{\text{solv.}}$ ^d , kcal/mol	$\Delta G_{\text{solv.}}$ P-Value ^e	CSS ^f
A – B	x, y, z	897.9	63 (31-32)	-9.7	0.123	0.963
C – D	x, y, z	882.6	66 (34-32)	-8.6	0.176	0.963
A – C	x, y, z	529.5	33 (17-16)	-5.7	0.197	0.000
B – D	$x - \frac{1}{2}, y + \frac{1}{2}, z$	506.8	35 (18-17)	-5.6	0.197	0.000
A – D	$-x + \frac{3}{2}, y - \frac{1}{2}, -z + 2$	155.3	14 (6-8)	-0.9	0.480	0.000
B – C	$-x + 2, y, -z + 3$	261.5	20 (5-15)	+1.1	0.684	0.000

^aSymmetry operation applied to the 2nd interfacing structure to obtain the interface, specified in fractional space relative to the initial structure position.

^bCalculated as difference in total accessible surface areas of isolated and interfacing structures divided by two.

^cValues in parenthesis correspond to number of residues from each monomer.

^dSolvation free energy gain upon formation of the interface. $\Delta G_{\text{solv.}} < 0$ corresponds to positive protein affinity by hydrophobic interfaces (doesn't include effect of salt bridges or H-bonds across the interface).

^eMeasure of interface specificity.

^fIndicator of interface significance for assembly formation.

Table 4-2. Hydrogen bond interactions at the *T. cru* α -CA dimer interface.

Monomer A	Monomer B	Distance (Å)
Asn104 – O ^{δ1}	Ser248 – O ^Y	3.6
Ala115 – O	Ser120 – O ^Y	3.9
Gly117 – O	Thr119 – N	3.1
Thr119 – O ^{Y1}	Asp304 – N	3.0
Ser120 – O ^Y	Gly117 – N	3.3
Glu258 – O ^{ε1}	Asn104 – N ^{δ2}	3.6
Ala295 – O	Lys246 – N	2.5
Asp304 – O ^{δ1}	Ser120 – O ^Y	2.7
Asp304 – O ^{δ1}	Ser120 – N	3.8
Asp304 – O ^{δ1}	Thr119 – O ^{Y1}	2.3
Ser248 – O ^Y	Asn104 – O ^{δ1}	3.2
Thr119 – N	Gly117 – O	3.3
Asp304 – N	Thr119 – O ^{Y1}	3.0
Gly117 – N	Ser120 – O ^Y	3.4
Asn104 – N ^{δ2}	Glu258 – O ^{ε1}	3.4
Lys246 – N	Ala295 – O	2.4
Ser120 – O ^Y	Asp304 – O ^{δ2}	3.2
Thr119 – O ^{Y1}	Asp304 – O ^{δ2}	3.1

Table 4-3. Comparison of the interfaces of *T. cru* α -CA and other dimeric α -CAs.

Dimeric α -CA	Interface Area ^a , Å ²	No. Interfacing Residues ^b	$\Delta G_{\text{solv.}}$ ^c kcal/mol	$\Delta G_{\text{solv.}}$ P-value ^d	CSS ^e	No. Hydrogen Bonds
<i>T. cru</i> α -CA	897.9	63 (31-32)	- 9.7	0.123	0.963	18
hCA VI	869.3	54 (28-26)	- 9.8	0.552	0.000	11
hCA IX	801.6	54 (27-27)	- 9.6	0.170	0.000	2
hCA XII	1130.2	72 (35-37)	+ 1.1	0.783	0.000	17
<i>Cr</i> - α CA1	1493.3	84 (42-42)	- 17.7	0.068	0.221	21
<i>Ao</i> CA	1812.2	93 (46-47)	- 18.4	0.189	0.265	23

^aCalculated as difference in total accessible surface areas of isolated and interfacing structures divided by two.

^bValues in parenthesis correspond number of residues from each monomer.

^cSolvation free energy gain upon formation of the interface. $\Delta G_{\text{solv.}} < 0$ corresponds to positive protein affinity by hydrophobic interfaces (doesn't include effect of salt bridges or H-bonds across the interface).

^dMeasure of interface specificity.

^eIndicator of interface significance for assembly formation.

Table 4-4. Comparison of dimer interfaces by secondary structure matching (SSM) superposition of *T. cru* α -CA using *Coot* (64).

	Sequence identity (%)	No. aligned residues	Bond-length RMSD (Å)
hCA VI (3FE4)	32.52	206	1.740
hCA IX (3IAI)	35.50	214	1.374
hCA XII (1JCZ)	33.18	217	1.426
<i>Cr</i> - α CA (3B1B)	33.48	221	1.611
<i>Ao</i> CA (3Q31)	29.47	207	1.606

Table 4-5. List of interfacing residues of dimeric α -CAs.

Residue type	<i>T. cru</i> α -CA	hCA VI	hCA IX	hCA XII	<i>Cr</i> - α CA	AoCA
Hydrophilic	Asn104	His38	Ser41	Tyr7	Cys21	Thr40
	Ala115	Tyr39	Pro42	Glu13	Lys24	Asn47
	Gly117	Glu77	Pro84	Asn14	Asn45	Tyr49
	Thr119	Pro79	Arg86	Ser15	Asn71	Glu53
	Ser120	Ser90	Tyr88	Ser17	Thr167	Thr126
	Ser248	Asp150	Ser124	Lys18	Lys175	Pro127
	Lys246	Gln153	Glu133	Cys23	Tyr177	His131
	Glu258	Asp154	Arg137	His34	Pro178	Glu134
	Ala295	Tyr189	Pro138	Asp36	Gln207	Glu135
	Asp304	Pro190	Glu195	Asn99	Asn210	His136
		Gln192	Arg254	Asp102	Asn285	Pro138
		Arg193	Glu257	His103	Arg288	Gln160
		Thr194		Ser110	Gln358	Glu163
		Thr195		Gln112	Asn362	Tyr239
		Pro223		Asn243	Pro363	Asn240
				Arg246	Tyr365	Lys243
				Gln274		Lys247
				Gln249		Tyr248
				Lys150		Tyr252
				Asp252		Thr253
					Gln254	
					Asn255	
					Glu264	
					Lys269	
Hydrophobic	Val116	Ile151	Phe27	Phe8	Ile22	Gly41
	Leu121	Gly191	Ala39	Gly9	Gly43	Leu42
	Gly259		Ala43	Gly24	Leu44	Leu46
	Val260		Gly85	Ile37	Leu171	Val68
	Leu299		Ala127	Gly111	Ala173	Phe99
			Phe128	Ile242	Gly174	Leu161
			Gly136	Phe245	Ile176	Val236
			Gly139	Lys248	Leu208	Leu262
			Gly253		Phe281	Leu263
			Val255		Trp284	Val265
					Leu293	Ala266
					Val354	Ala267
					Ala355	Leu270
					Phe356	
					Phe360	
Cystines	-	-	Cys41	-	Cys21	-

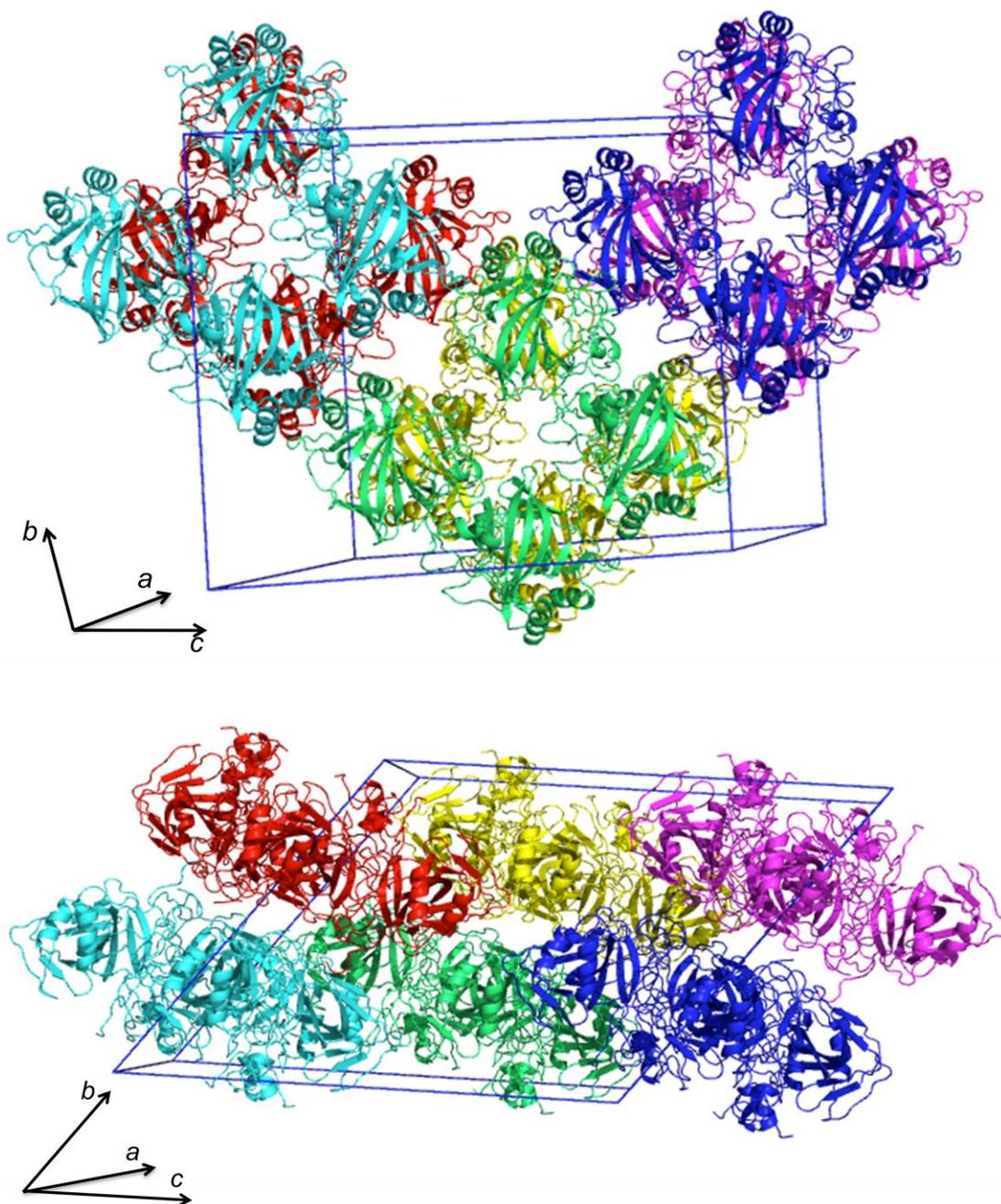


Figure 4-1. Representation of the crystal packing of *T. cru* α -CA. A unit cell (denoted by the blue box) for *T. cru* α -CA contains 16 chains, 4 molecules per asymmetric unit (au). Each molecule in the au is depicted in a different color.

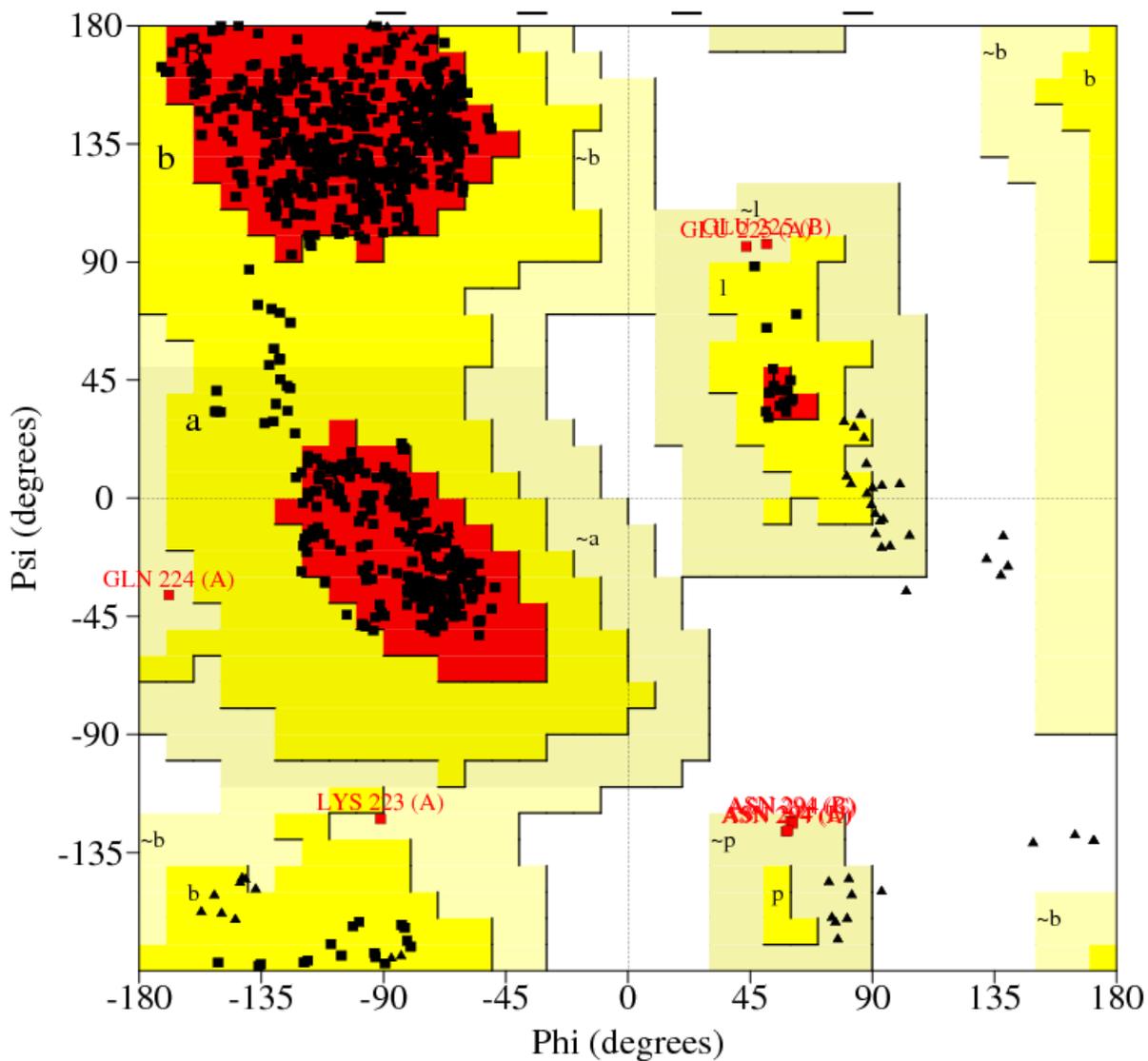


Figure 4-2. Ramachandran plot of *T. cru* α -CA residues obtained from *PROCHECK* (66). The Ramachandran plot statistics show that over 90% of the dihedral angles of *T. cru* α -CA are within the most favored regions, while the rest are in the additionally allowed regions, except for 1.0% of the angles that fall within the generously allowed regions.

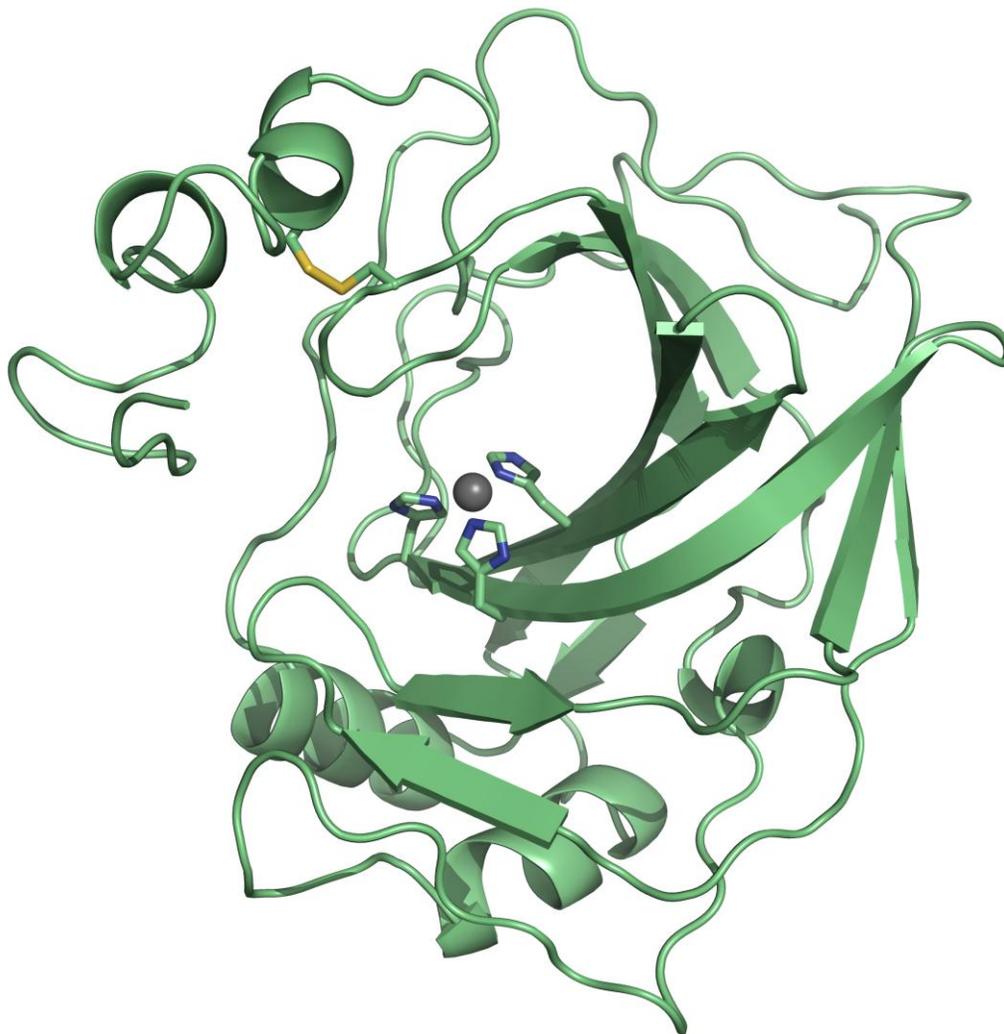


Figure 4-3. Cartoon representation of *T. cru* α -CA. The structure is oriented in such a manner as to look down into the active site cavity, so the β -sheet fold characteristic of α -CAs can be seen better. Zinc-coordinating histidines and the disulfide bond represented by sticks. Nitrogen atoms are shown in navy blue, and sulfur atoms in yellow. Zinc ion is shown as a gray sphere.

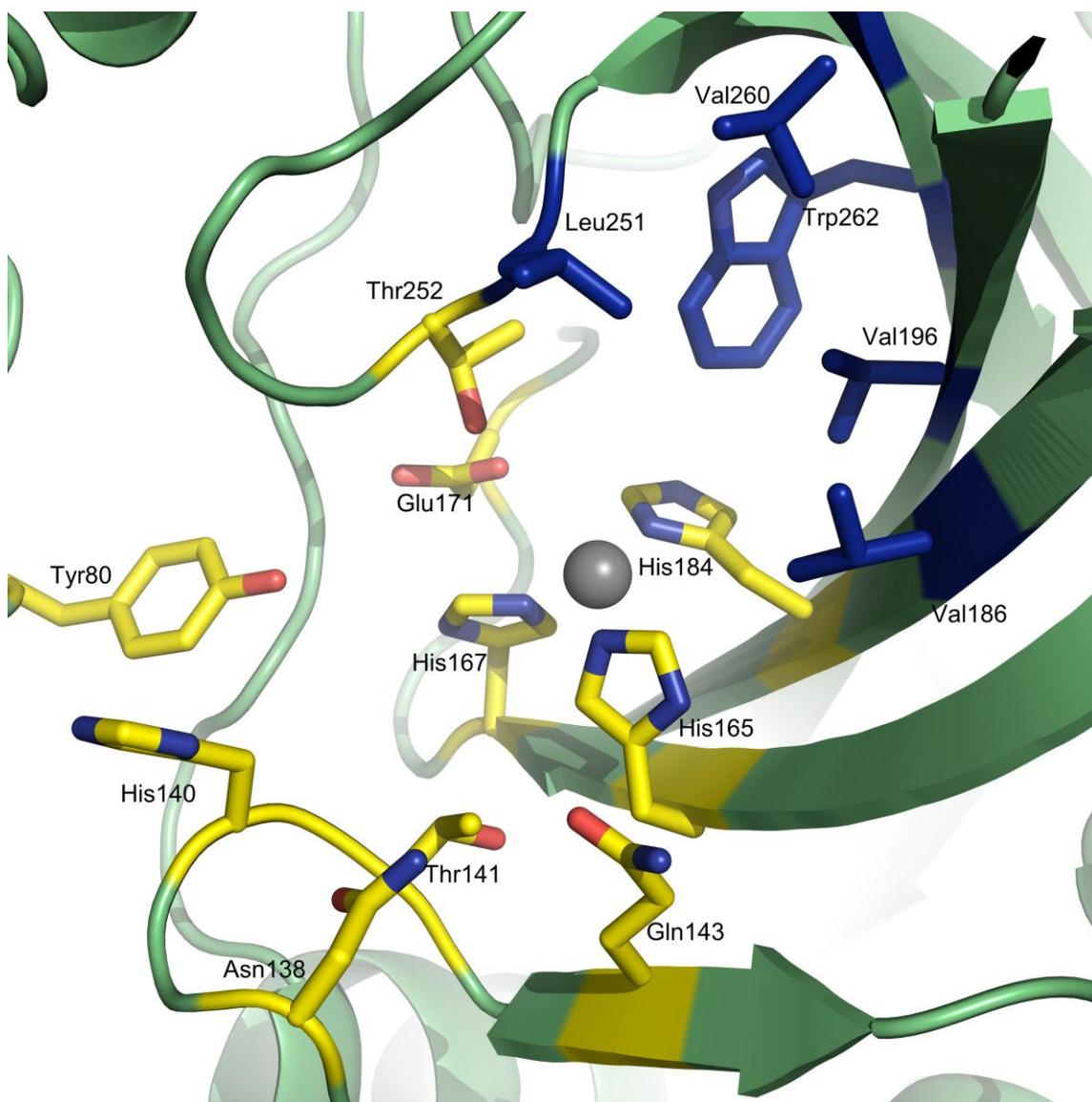


Figure 4-4. Close-up stick representation of the *T. cru* α -CA active site. Residues Tyr80, Asn138, His140, Thr141, Gln143, Glu171, and Thr252 form the hydrophilic side of the active site (yellow) which aids in proton transfer and bicarbonate removal, while residues Val186, Val196, Leu251, Val260, and Trp262 line the hydrophobic side (blue), also the CO₂ binding site. Nitrogen atoms are shown in navy blue; oxygen in red, and the zinc ion as a gray sphere.

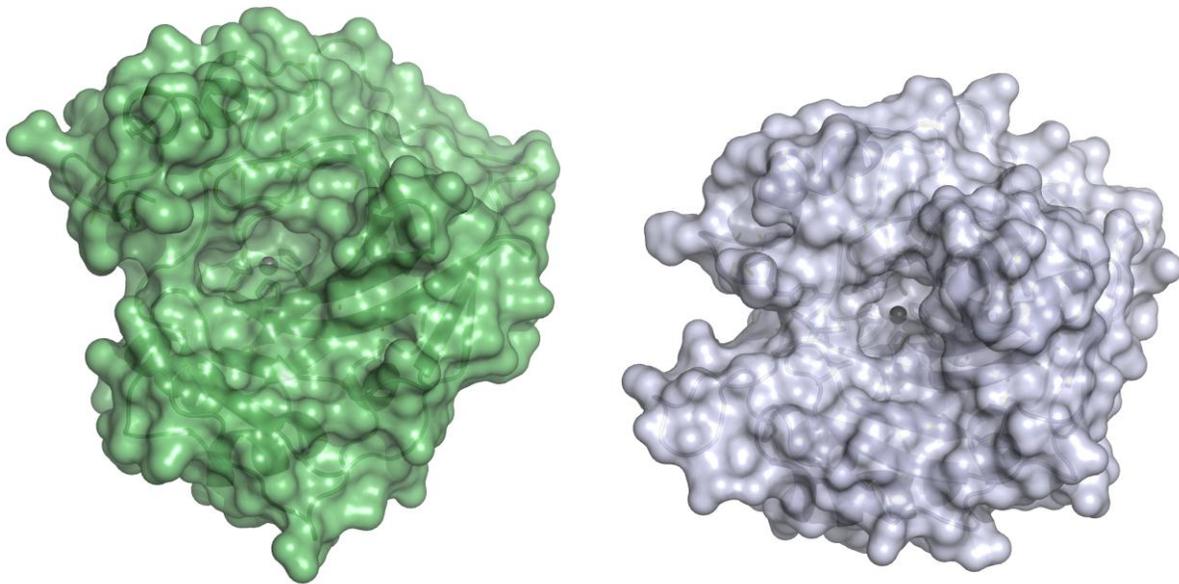


Figure 4-5. Comparison of *T. cru* α -CA (green) and hCA II (gray) (PDB ID: 3KS3; 54). The deletion of surface loops in *T. cru* α -CA lead to a structure with a wider gap at the active site cavity, increasing solvent accessibility. Zinc ion is shown as a gray sphere.

```

hCA II -----
T. cru α-CA MKKRFSFIFIFLVALPLYSANNVAAPLIDLGAEAKKQAQKSAATQSAVPE 50

hCA II -----MSHHWGYGKHNGPEHWHKDFPIAK--- 24
T. cru α-CA KESATKVAEKQKEPEEKAKPEPKPPHWGYFGEEGPQYWGELAPEFSTCK 100
                . **** .:***::* : * .

hCA II -GERQSPVDIDHTHTAKYDPSLKPLSVSYDQATSLRILNNGHAFNVEFDDS 73
T. cru α-CA TGKNQSPINLKPQTAVGTTSLPGFDVYY-RETALKLINNGHTLQVNIPLG 149
                *:***:::..:*** ** :.* * :*:***::*****::*::: .

hCA II QDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKCKYAAELHLVHW 123
T. cru α-CA SYIKIN-----GHRYELLQYHFHTP-----SEHQRDGFNYPMEMHLVHK 188
                . : . *.*:*** ** * :*. *:***

hCA II NTKYGDGPKAVQPPDGLAVLGIPLKVGSAKPLQKVVDVLDSIKTKGKSA 173
T. cru α-CA -----DGDGNLAVIAILFQEGEENETLAKLMSFLPQTLKKQEIH 227
                : ..***:.*::: *. : * *::...* . .* :

hCA II DFTNFDPRGLLPESLDYWTYPGSLTTPPLLECVTWIVLKEPISVSSEQVL 223
T. cru α-CA ESVKIHPAKFFPADKKFYKYSGLTTPPCSEGVYWMVFKQPIQASVTQLE 277
                : :...* ::* . :...*.***** * * *:***:***..* *:

hCA II KFRKLNPNGEPEELMVDNWRPAQPLKNRQIKASF- 260
T. cru α-CA KMHEYLGSNARPVQRQNARTLLKSWPDRNRANTVYEFY 315
                *::: .. :. . . : * :** ..

```

Figure 4-6. Sequence alignment of hCA II and *T. cru* α-CA determined by ClustalW (106, 107). The sequence alignment demonstrates the additional N-terminal region formed by residues 23 – 74 in *T. cru* α-CA, as well as the predicted gaps in the *T. cru* α-CA sequence compared to hCA II.

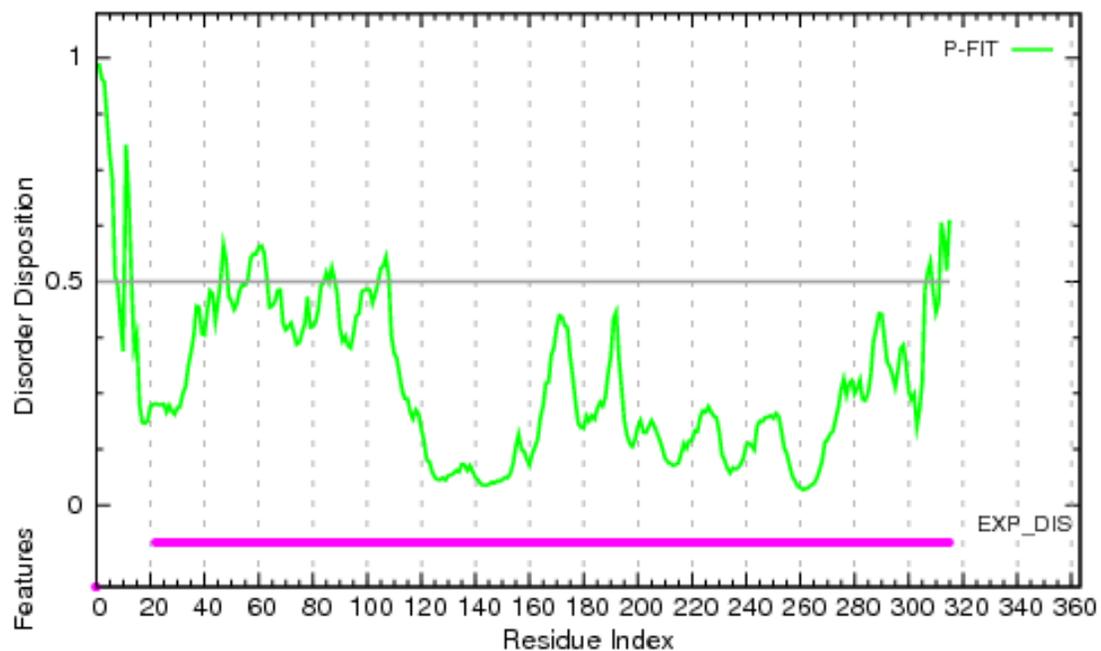


Figure 4-7. Disorder probability plot obtained from PONDR-FIT (81). The plot shows that some regions, particularly in the N-terminus and at the C-terminus, are probably disordered. Disorder in some regions could cause the absence of these and adjacent regions in crystal structures.

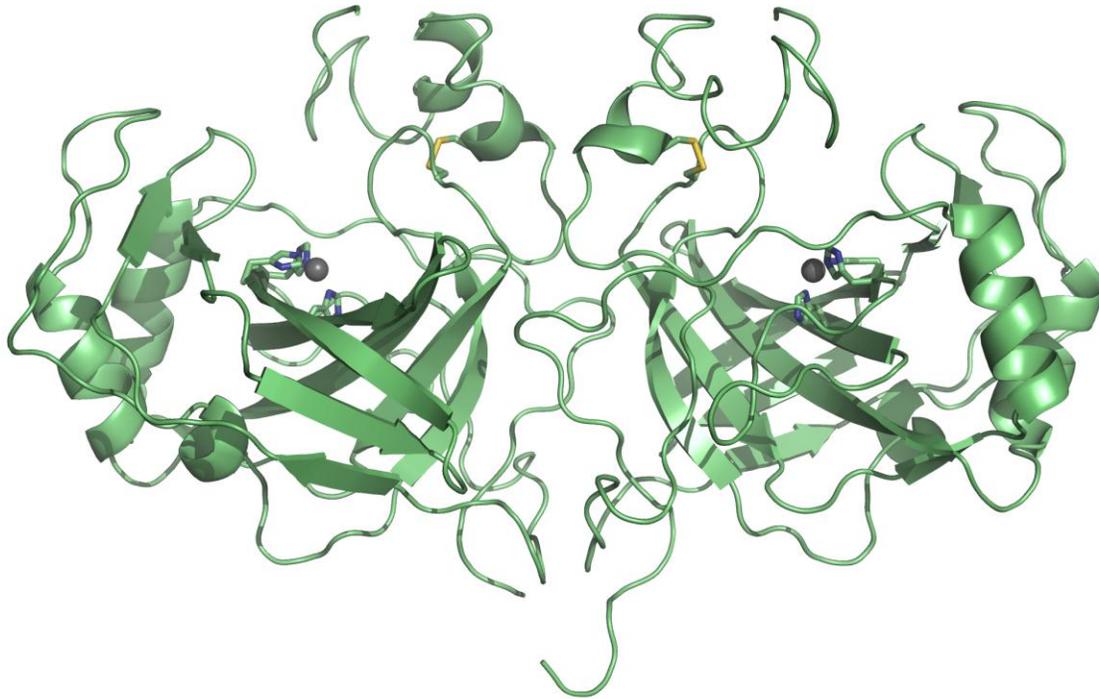


Figure 4-8. Cartoon representation of the dimer formed by chains A (left) and B (right) of *T. cru* α -CA. The monomers in the homodimer are related by a 2-fold axis along the N- and C-terminus. The zinc-coordinating histidines and the disulfide bond are shown as sticks. Nitrogen atoms are shown in navy blue, sulfur atoms in yellow, and the zinc ion as a gray sphere.

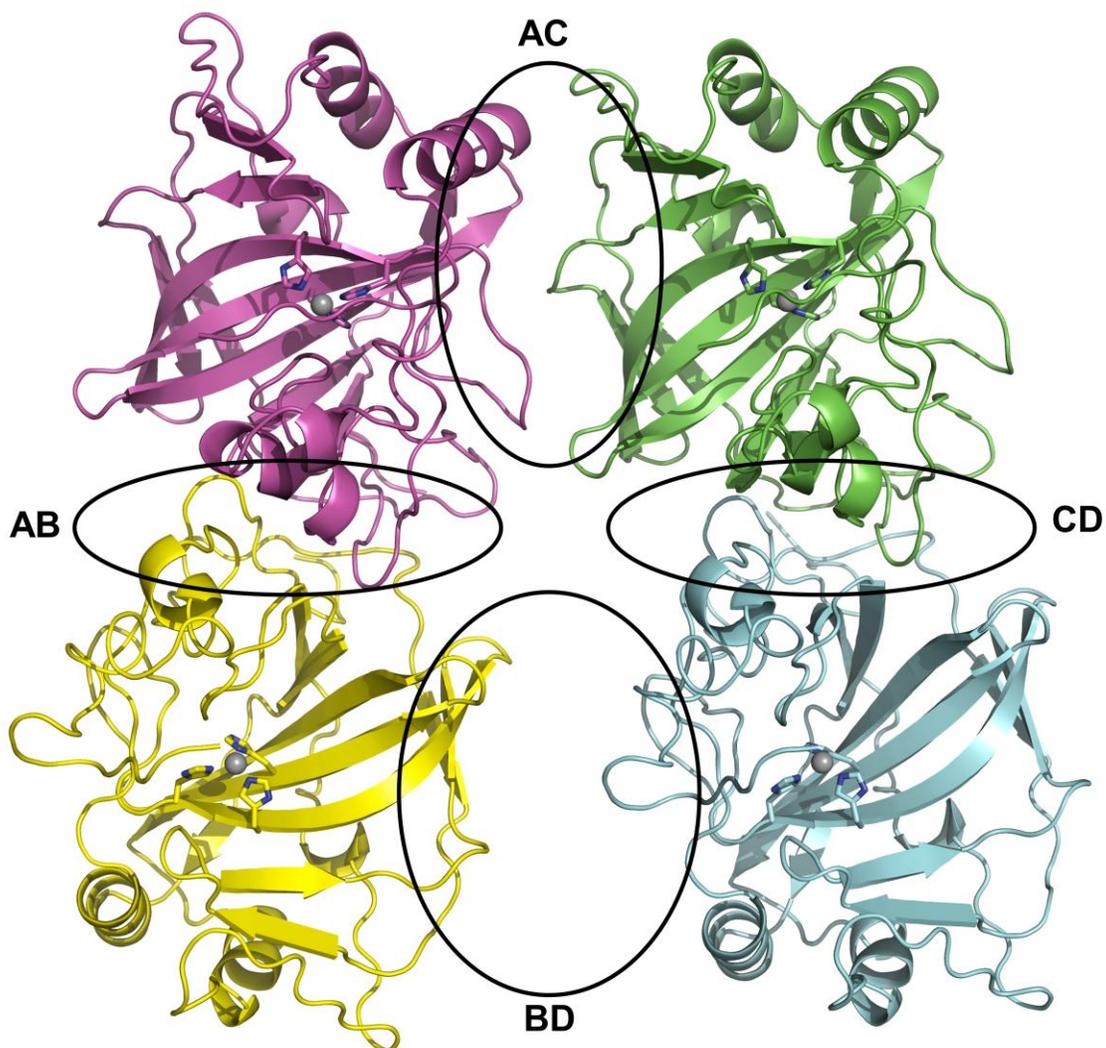


Figure 4-9. Possible interfaces present at the crystal structure of *T. cru* α -CA. The PISA server (102) analyzed the interfaces that could potentially form between the chains within the au of *T. cru* α -CA. The black circles drawn above represent some of the possible interfaces. AB, AC, CD, and BD represent the interface formed between chains A and B, chains A and C, chains C and D, and chains B and D, respectively.

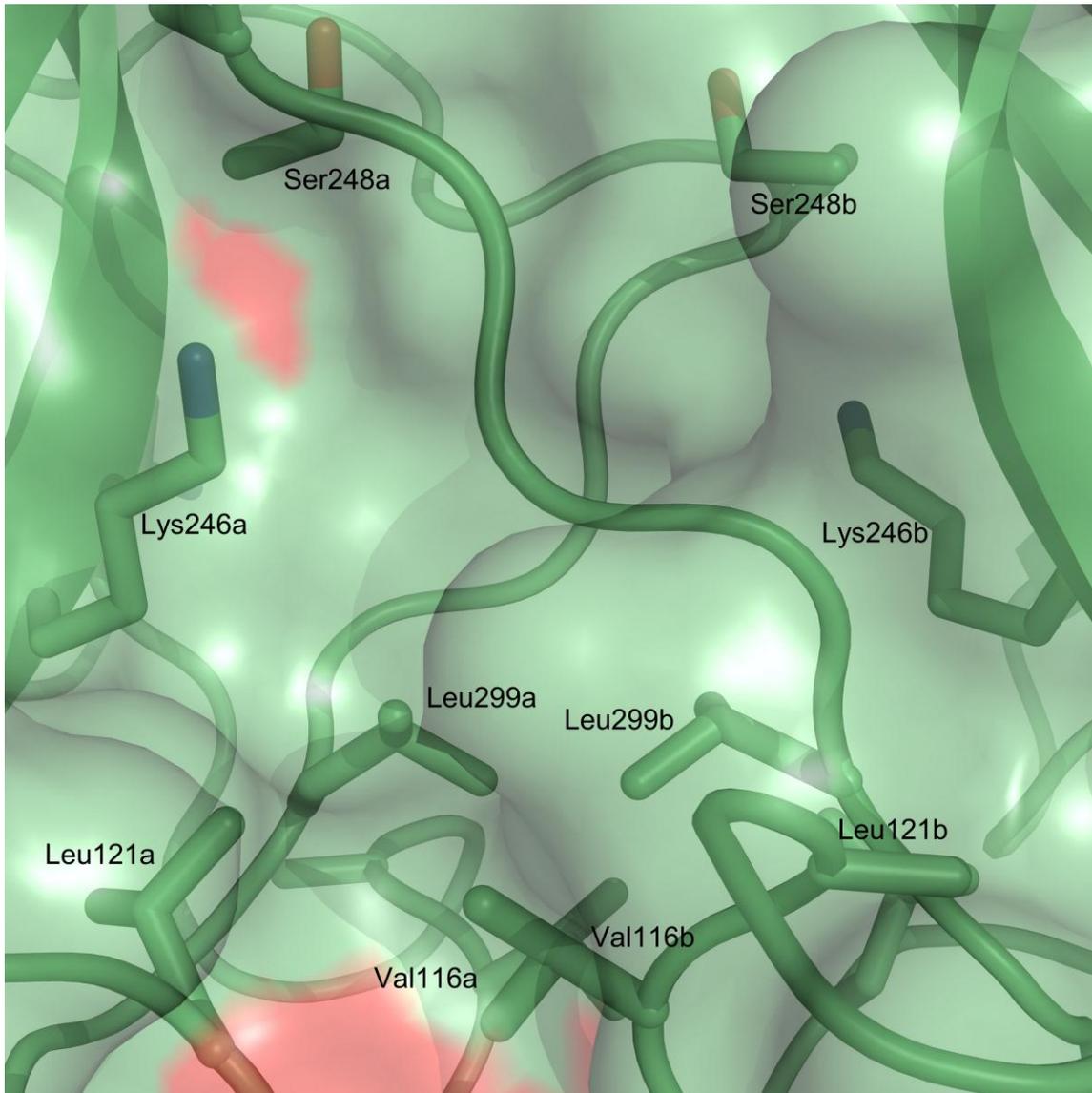


Figure 4-10. Buried lysines at the *T. cru* α -CA dimer interface. The microenvironment formed by the surrounding hydrophobic residues Val 116, Leu121, and Leu299 can affect the pK_a of the buried lysines. Nitrogen is shown in blue, and oxygen in red.

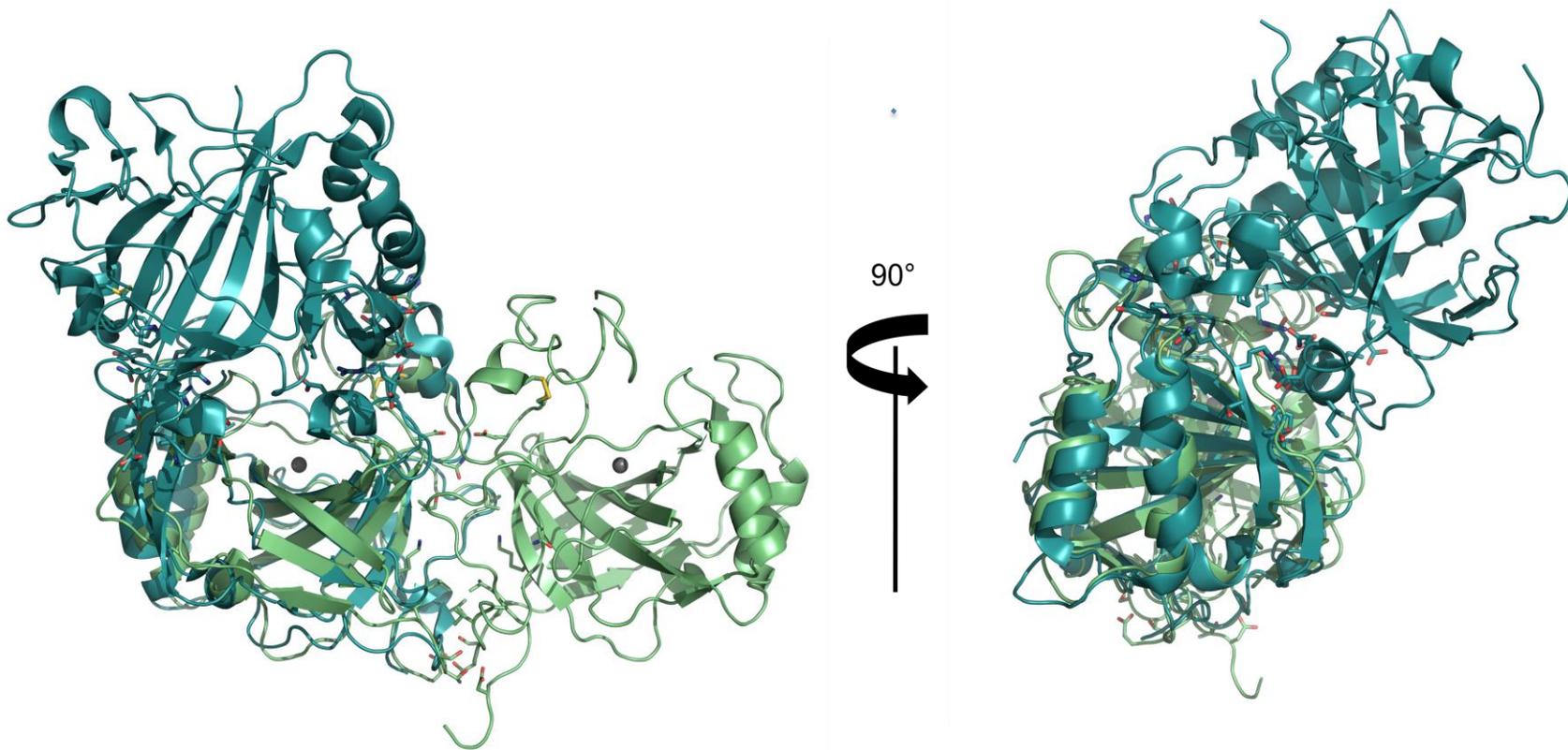


Figure 4-11. Superposition of the dimeric α -CAs *T. cru* α -CA (green) and hCA VI (teal) (PDB ID: 3FE4; 100). The dimer interface of hCA VI is very different from that of *T. cru* α -CA, as it appears to be right on top of the active site cavity. Zinc ion is shown as gray sphere. Interface residues are shown as sticks. Nitrogen atoms are shown in blue, sulfur in yellow, and oxygen in red.

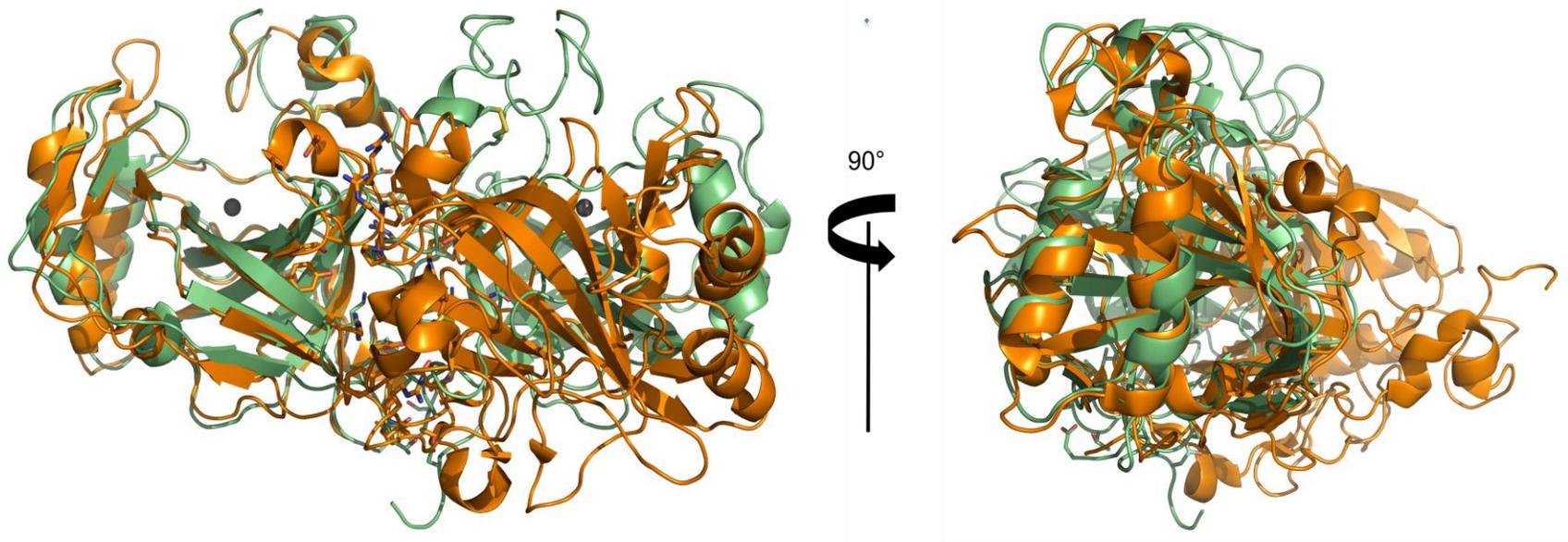


Figure 4-12. Superimposition of *T. cru* α-CA (green) and hCA IX (orange) (PDB ID: 3IAI; 87) dimeric structures. Some residues involved in the dimer association of hCA IX appear in the interface of *T. cru* α-CA as well. Nitrogen atoms are shown in navy blue, sulfur atoms in yellow, oxygen atoms in red, and the zinc ions are represented by gray spheres.

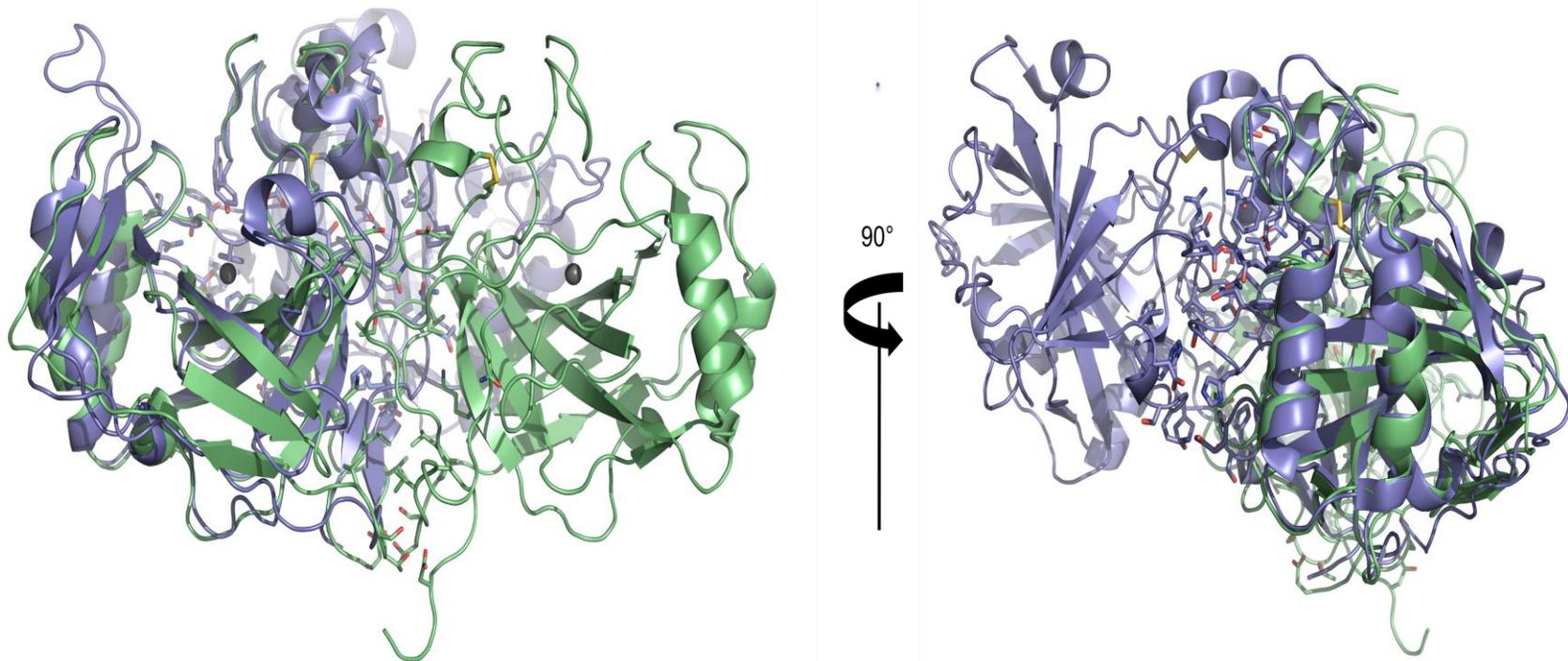


Figure 4-13. Superimposition of *T. cru* α -CA (green) and hCA XII (blue) (PDB ID: 1JCZ; 88). The dimer interface of *T. cru* α -CA and hCA XII are also very different. Nitrogen atoms are shown in navy blue, sulfur atoms in yellow, oxygen atoms in red, and the zinc ions are represented by gray spheres.

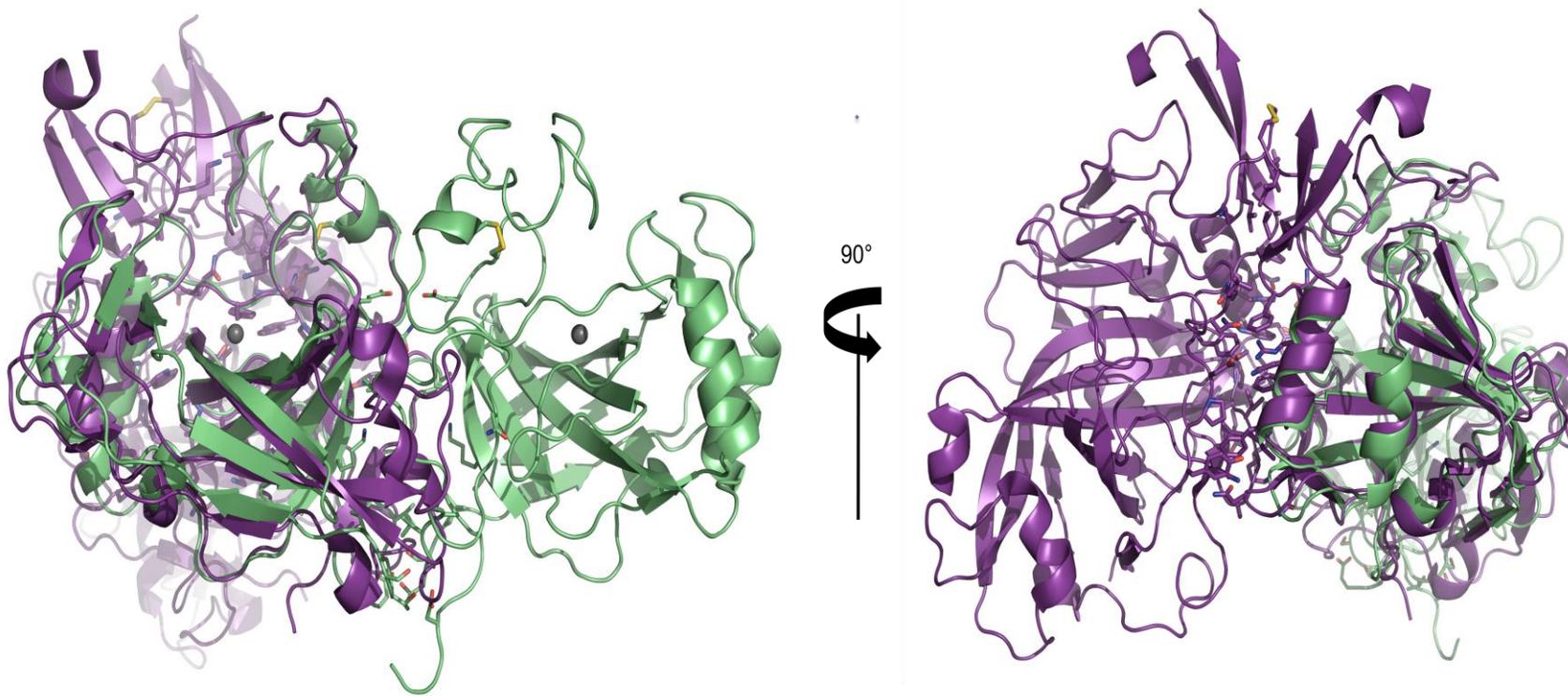


Figure 4-14. Superposition of *T. cru* α -CA (green) and *Cr*- α CA1 (purple) dimers (PDB ID: 3B1B; 101). The dimer interfaces of *T. cru* α -CA and *Cr*- α CA1 also differ in orientation. Nitrogen atoms are shown in navy blue, sulfur atoms in yellow, oxygen atoms in red, and the zinc ions are represented by gray spheres.

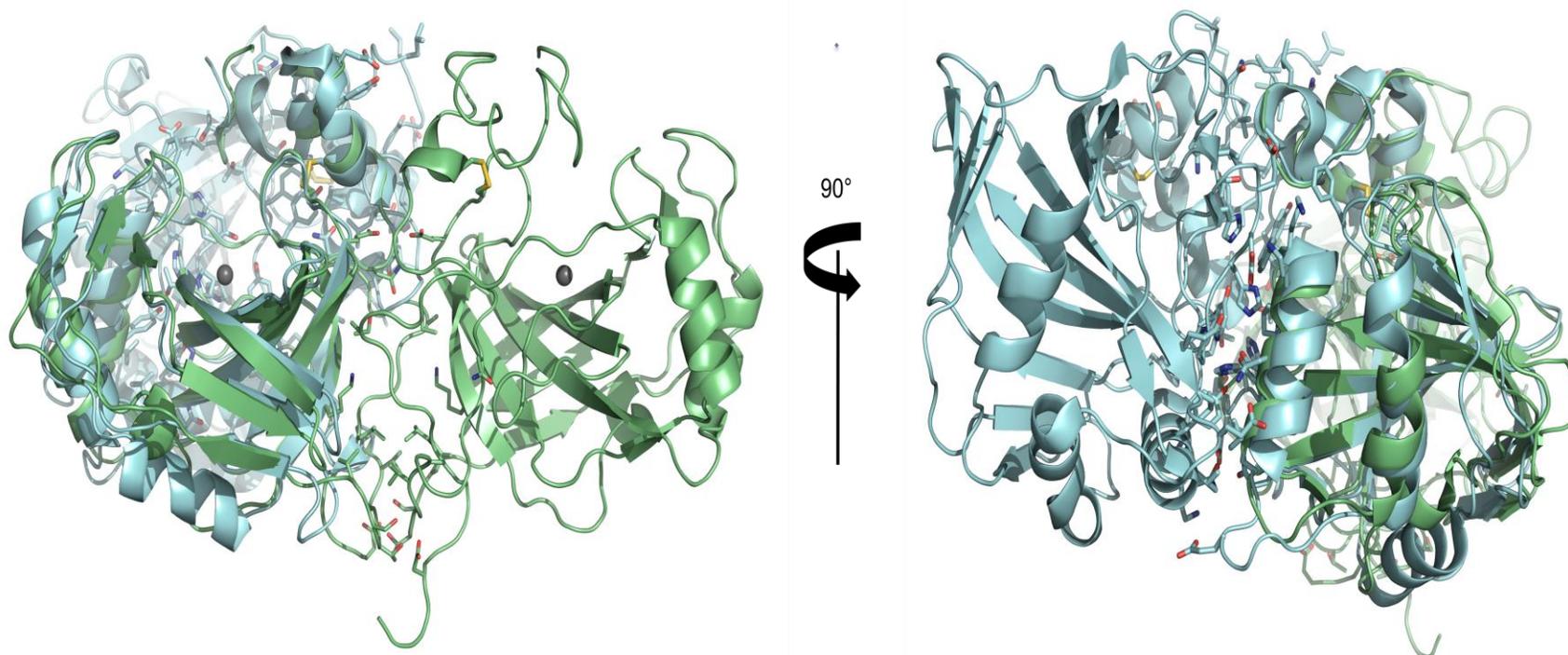


Figure 4-15. Superimposition of *T. cru* α -CA (green) and AoCA (cyan) dimers (PDB ID: 3Q31; 89). The residues at the dimer interface of *T. cru* α -CA and AoCA differ, leading to a different orientation of dimer association. Nitrogen atoms are shown in navy blue, sulfur atoms in yellow, oxygen atoms in red, and the zinc ions are represented by gray spheres.

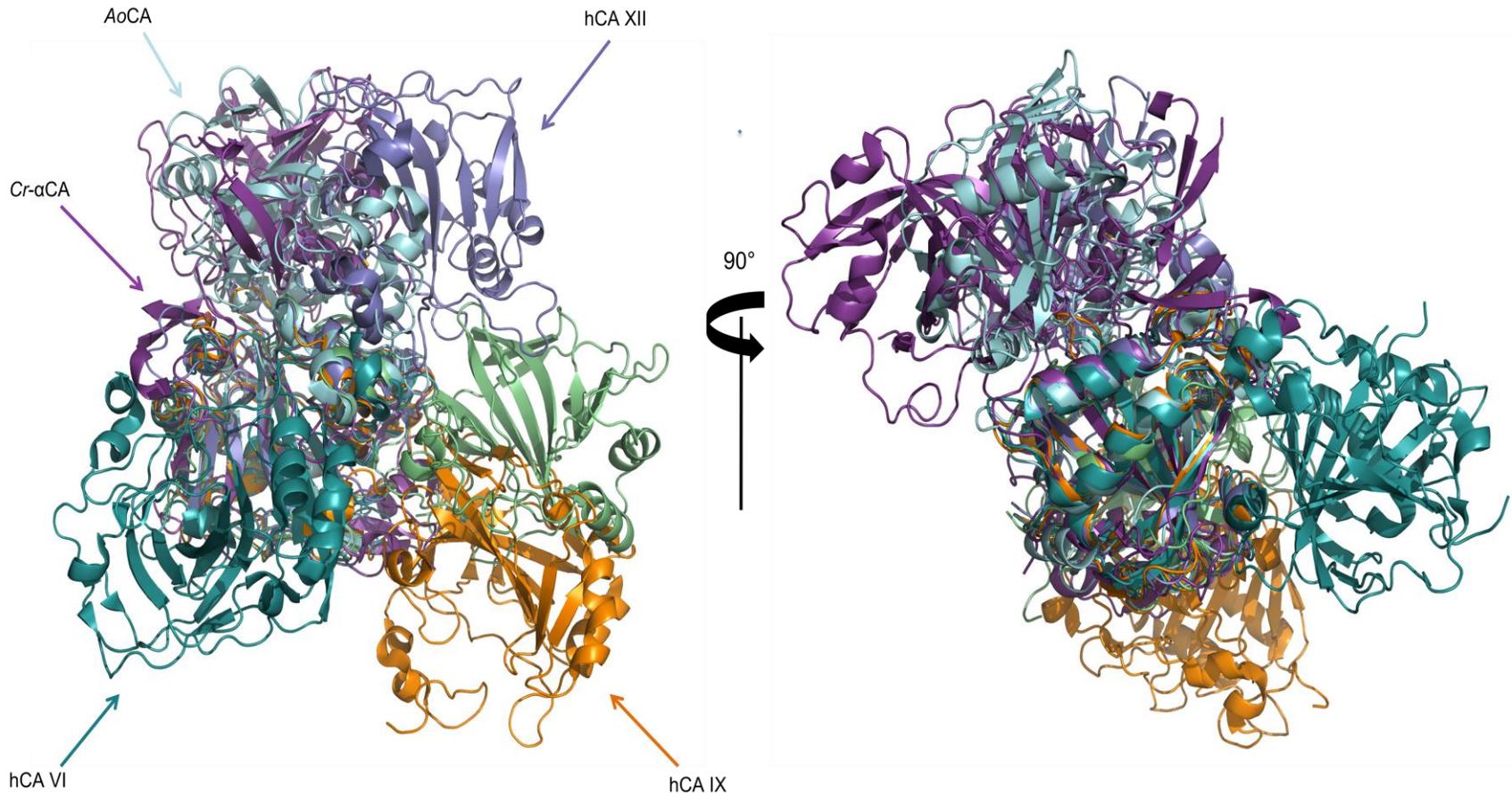


Figure 4-16. Superimposition of all dimeric α -CAs on *T. cru* α -CA (green). On the figure to the right, it can be observed that all of the α -CAs superimposed well on one of the monomers. However, none of the dimers superimpose on each other, so all of the dimers form a unique interface. The view has been altered for clarity. Zinc ions are shown as gray spheres.

CHAPTER 5 CONCLUSIONS

Studying the structural, kinetic, and biophysical characteristics of the α -CA expressed in the *T. cru* gammaproteobacterium provided an insight to increase thermal stability of α -CAs that could be used for industrial applications, such as CO₂ scrubbers or artificial lungs. *T. cru* α -CA is an active isoform of the α -class CAs that shows increased thermal and pH-independent stability as compared to other α -CAs. Results from this study, along with studies performed with thermostable CAs NGCA and hCA IV (90-93), suggest that the disulfide bond present between residues Cys99 and Cys256 is the major contributor of the increased thermal stability of *T. cru* α -CA. The disulfide bond, along with the pH-independent stability observed between pH values 5 – 9, are adaptations that allow the *T. cru* α -CA to remain fully functional in the erratic environment that dominates the hydrothermal vents, and could potentially be used to engineer CAs with increased thermal and pH stability. Further studies need to be completed to determine the structural or physical features that stabilize the protein at different pH values.

Unlike most α -class CAs, *T. cru* α -CA forms a homodimer in solution. Comparable to the melting temperature of the unfolding of *T. cru* α -CA, the temperature corresponding to the dissociation of the dimer was directly proportional to the pH. An increased dimer association, which would result in a greater T_M , suggested the presence of ionizable groups at the interface. A closer look at the interfacing residues revealed a buried lysine (Lys246) could have a decreased pK_a that would result in deprotonation of the ϵ -amine group at lower pH, enhancing the hydrogen bonds present at the interface. Comparison of the dimeric interface of *T. cru* α -CA with other dimeric α -

CAs demonstrated that *T. cru* α -CA forms a unique interface. Furthermore, we observed that none of the dimeric α -CAs used for comparison in this study share the same interface.

The crystal structure of *T. cru* α -CA lacks 52 amino acid residues at the N-terminus, a region that appears to be unique among known proteins to date. The region is of highly hydrophilic content with a basic nature that could potentially affect the nucleophilicity of the zinc-bound water by decreasing its pK_a, resulting in a decreased catalytic efficiency as compared to hCA II. Since the 52 N-terminus residues were missing in both the crystal structure and previous homology models, we assessed the possibility of an intrinsically disordered region within *T. cru* α -CA and observed that the N-terminus was predicted to be partially disordered. Additional studies could be performed in order to further understand the disordered nature of the missing N-terminal region and determine if this region might be involved in other biological functions.

Overall, the *T. cru* α -CA is comparable to hCA II in structure and catalytic activity, with major differences in the thermal and pH stability. The similarities in structure and activity to hCA II suggest that the biological roles of *T. cru* α -CA are fairly similar to those of hCA II. Therefore, *T. cru* α -CA could very well be involved in CO₂ transport within the periplasm to the cytoplasm and pH homeostasis. *T. cru* α -CA could also be involved in sequestering the CO₂ and bicarbonate in the periplasm to prevent CO₂ from diffusing out of the cell. The differences in thermal stability between *T. cru* α -CA and hCA II suggest that the enhanced thermal and pH stability of the *T. cru* α -CA is vital for the growth and survival of *T. cru*.

In conclusion, this project describes the structure, catalysis, and biophysical nature of the *T. cru* α -CA with the intent to identify the physiological roles of the α -CA within the *T. cru* gammaproteobacterium. Several roles have been suggested, and will require whole-cell studies to confirm the physiological relevance of these roles. Additional studies are also required to structurally, kinetically, and biophysically characterize the other CAs expressed in *T. cru*. Characterization of the other CAs, along with whole-cell studies, can provide the information needed to assess the biological interactions of these CAs and determine their physiological functions, as well as the relevance of the biological interactions to the growth and survival of *T. cru*.

LIST OF REFERENCES

1. Duda, D. M. and McKenna, R. (2006) Carbonic Anhydrases (α -Class). Handbook of Metalloproteins. New York: John Wiley & sons, Ltd.
2. Domsic, J. F., Avvaru, B. S., Kim, C. U., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N., and McKenna, R. (2008) Entrapment of carbon dioxide in the active site of carbonic anhydrase II, *J. Biol. Chem.* 283, 30766-30771.
3. Chegwidde, W. R., Carter, N. D., and Edwards, Y.H. (2000) *The carbonic anhydrases, New Horizons*, Switzerland.
4. Meldrum, N. U., and Roughton, F. J. W. (1933) Carbonic anhydrase. Its preparation and properties, *J. Physiol. (Lond)* 80, 113-142.
5. Henry, R. P. and Swenson, E. R. (2000). The distribution and physiological significance of carbonic anhydrase in vertebrate gas exchange organs. *Respir. Physiol.* 121, 1-12.
6. Maren, T. H. (1984), The general physiology of reactions catalyzed by carbonic anhydrase and their Inhibition by sulfonamides. *Annals of the New York Academy of Sciences* 429, 568–579.
7. Sly, W. S. and Hu, P. Y. (1995) Human carbonic anhydrases and carbonic anhydrase deficiencies, *Annu. Rev. Biochem.* 64, 375-401.
8. Shively, J. M., Keulen, G., and Meijer, W. G. (1998). Something from almost nothing: Carbon dioxide fixation in chemoautotrophs. *Annu. Rev. Microbiol.* 52, 191-230.
9. Badger, M.R., Price, G.D. (1994). The role of carbonic anhydrase in photosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Bio.* 45, 369-392.
10. Tripp, B.C., Smith, K., and Ferry, J.G. (2001) Carbonic anhydrase: New insights for an ancient enzyme, *J. Biol. Chem.* 276, 48615-48618.
11. Hewett-Emmett, D., and Tashian, R. E. (1996) Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families, *Mol. Phylogenet. Evol.* 5, 50-77.
12. Huang, S., Xue, Y., Sauer-Eriksson, E., Chirica, L., Lindskog, S., and Jonsson, B. H. (1998). Crystal structure of carbonic anhydrase from *Neisseria gonorrhoeae* and its complex with the inhibitor acetazolamide. *Journal of Molecular Biology*, 283, 301-310.

13. Kalloniati, C., Tsikou, D., Lampiri, V., Fotelli, M. N., Rennenberg, H., Chatzipavlidis, I., Fasseas, C., Katinakis, P., and Flemetakis, E. (2009) Characterization of a *Mesorhizobium loti* α -type carbonic anhydrase and its role in symbiotic nitrogen fixation. *J. Bacteriol.* *191*, 2593–2600.
14. Puskás, L. G., Inui, M., Zahn, K., and Yukawa, H. (2000) A periplasmic, α -type carbonic anhydrase from *Rhodospseudomonas palustris* is essential for bicarbonate uptake. *Microbiology*, *146*, 2957-2966
15. Cronk, J. D., Endrizzi, J. A., Cronk, M. R., O'Neill, J. W. and Zhang, K. Y.J. (2001). Crystal structure of *E. coli* β -carbonic anhydrase, an enzyme with an unusual pH-dependent activity. *Protein Science* *10*, 911–922.
16. Strop, P., Smith, K. S., Iverson, M., Ferry, J. G., and Rees, D. C. (2001) Roles of the conserved aspartate and arginine in the catalytic mechanism of an archaeal β -class carbonic anhydrase. *J. Biol. Chem.* *276*, 10299–10305.
17. Mitsuhashi, S., Mizushima, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Moriyama, H., Ueki, T., Miyachi, S., and Tsukihara, T. (2000) X-ray structure of beta-carbonic anhydrase from the red alga, *Porphyridium purpureum*, reveal a novel catalytic site for CO₂ hydration. *J. Biol. Chem.* *275*, 5521–5526.
18. Kimber, M. S., and Pai, E. F. (2000) The active site architecture of *Pisum sativum* β -carbonic anhydrase is a mirror image of that of α -carbonic anhydrases. *EMBO J.* *19*, 1407–1418
19. Cronk, J. D., Rowlett, R. S., Zhang, K. Y. J., Tu, C. K., Endrizzi, J. A., Lee, J., Gareiss, P. C., and Preiss, J. R. (2006) Identification of a novel noncatalytic bicarbonate binding site in eubacterial beta-carbonic anhydrase. *Biochemistry* *45*, 4351-4361.
20. Smith, K. S., Jakubzick, C., Whittam, T. S., and Ferry, J. G. (1999) Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 15184 –15189.
21. Smith, K. S., Coper, N. J., Stalhandske, C., Scott, R. A., and Ferry, J. G. (2000) Structural and kinetic characterization of an archaeal β -class carbonic anhydrase. *J. Bacteriol.* *182*, 6605–6613
22. Alber, B. E., and Ferry, J. G. (1994) A carbonic-anhydrase from the archaeon *Methanosarcina thermophila*. *Proc. Natl. Acad. Sci. U.S.A.* *91*, 6909–6913.
23. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. C. (1996) A left-handed beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*. *EMBO J.* *15*, 2323–2330.

24. Iverson, T. M., Alber, B. E., Kisker, C., Ferry, J. G., and Rees, D. C. (2000) A closer look at the active site of γ -class carbonic anhydrases: High-resolution crystallographic studies of the carbonic anhydrase from *Methanosarcina thermophila*. *Biochemistry* 39, 9222-9231.
25. Parisi, G., Fornasari, M., and Echave, J. (2000) Evolutionary analysis of γ -carbonic anhydrase and structurally related proteins. *Mol Phylogenet Evol.* 14, 323–334.
26. Ferry, J. G. (2010) The gamma class of carbonic anhydrases. *Biochim Biophys Acta.* 1804, 374-381.
27. Alber, B. E., Colangelo, C. M., Dong, J., Stalhandske, C. M. V., Baird, T. T., Tu, C., Fierke, C. A., Silverman, D. N., Scott, R. A., Ferry, J. G. (1999) Kinetic and spectroscopic characterization of the gamma carbonic anhydrase from the methanoarchaeon *Methanosarcina thermophila*. *Biochemistry.* 38, 13119–13128.
28. MacAuley, S. R., Zimmerman, S. A., Apolinario, E. E., Evilia, C., Hou, Y., Ferry, J. G., Sowers, K. R. (2009) The archetype γ -class carbonic anhydrase (Cam) contains iron when synthesized *in vivo*. *Biochemistry* 48, 817–819.
29. Tripp, B. C., Bell, C. B., Cruz, F., Krebs, C., Ferry, J. G. (2004) A role for iron in an ancient carbonic anhydrase. *J. Biol. Chem.* 279, 6683–6687.
30. Lindskog, S. (1997) Structure and mechanism of carbonic anhydrase. *Pharmacol. Ther.* 74, 1-20.
31. Silverman, D. N., and Lindskog, S. (1988) The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc. Chem. Res.* 21, 30-36.
32. Zimmerman, S. A., and Ferry, J. G. (2006) Proposal for a hydrogen bond network in the active site of the prototypic γ -class carbonic anhydrase. *Biochemistry* 45, 5149-5157.
33. Merz, K. M. (1990) Insights into the function of the zinc hydroxide-Thr199-Glu106 hydrogen bonding network in carbonic anhydrases. *J. Mol. Biol.* 214, 4: 799–802.
34. Krebs, J. F., Ippolito, J. A., Christianson, D. W., and Fierke, C. A. (1993) Structural and functional importance of a conserved hydrogen bond network in human carbonic anhydrase II. *J. Biol. Chem.* 268, 27458-27466.
35. Liang, Z., Xue, Y., Behravan, G., Jonsson, B. H., and Lindskog, S. (1993) Importance of the conserved active-site residues Try7, Glu106 and Thr199 for the catalytic function of human carbonic anhydrase II. *European Journal of Biochemistry*, 211, 821–827.

36. Krabs, J. F. and Fierke, C. A. (1993) Determinants of catalytic activity and stability of carbonic anhydrase II as revealed by random mutagenesis. *J. Biol. Chem.* 268, 948-954.
37. Rowlett, R. S., Chance, M. R., Wirt, M. D., Sidelinger, D. E., Royal, J. R., Woodroffe, M., Wang, Y. F. A., Saha, R. P., and Lam, M. G. (1994) Kinetic and structural characterization of spinach carbonic anhydrase. *Biochemistry* 33, 13967-13976.
38. Rowlett, R. S., Tu, C., Murray, P. S., and Chamberlin, J. E. (2004) Examination of the role of Gln-158 in the mechanism of CO₂ hydration catalyzed by β -carbonic anhydrase from *Arabidopsis thaliana*. *Arch. Biochem. Biophys.* 425, 25-32.
39. Fisher, S. Z., Tu, C., Bhatt, D., Govindasamy, L., Agbandje-McKenna, M., McKenna, R., and Silverman, D. N. (2007) Speeding up proton transfer in a fast enzyme: Kinetic and crystallographic studies on the effect of hydrophobic amino acid substitutions in the active site of human carbonic anhydrase II. *Biochemistry* 46, 3803-3813.
40. Tu, C. K., Silverman, D. N., Forsman, C., Jonsson, B. H., and Lindskog, S. (1989) Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* 28, 7913-7918.
41. Fisher, Z., Hernandez Prada, J. A., Tu, C., Duda, D., Yoshioka, C., An, H., Govindasamy, L., Silverman, D. N., and McKenna, R. (2005) Structural and kinetic characterization of active-site histidine as a proton shuttle in catalysis by human carbonic anhydrase II. *Biochemistry* 44, 1097-1105.
42. Silverman, D. N., and McKenna, R. (2007) Solvent-mediated proton transfer in catalysis by carbonic anhydrase. *Acc Chem Res* 40, 669-675.
43. Earnhardt, J. N., Wright, S. K., Qian, M., Tu, C., Laipis, P. J., Viola, R. E., and Silverman, D. N. (1999) Introduction of histidine analogs leads to enhanced proton transfer in carbonic anhydrase V. *Arch. Biochem. Biophys.* 361, 264-270.
44. Smith, K. S., Ingram-Smith, C., Ferry, J. G. (2002) Roles of the Conserved Aspartate and Arginine in the Catalytic Mechanism of an Archaeal β -Class Carbonic Anhydrase. *J. Bacteriol.* 184, 4240-4245.
45. Scott, K. M., Sievert, S. M., Abril, F. N., Ball, L. A., Barrett, C. J., et al. (2006) The genome of deep-sea vent chemolithoautotroph *Thiomicrospira crunogena* XCL-2. *PLoS Biol* 4, e383.
46. Dobrinsky, K. P., Boller, A. J., and Scott, K. M. (2010) Expression and function of four carbonic anhydrase homologs in the deep-sea chemolithoautotroph *Thiomicrospira crunogena*. *Appl. Environ. Microbiol.* 76, 3561-3567.

47. Von Damm, K. L. (2000) Chemistry of hydrothermal vent fluids from 9°–10°N, East Pacific Rise: “Time zero,” the immediate post-eruptive period. *J. Geophys. Res.*, *105*, 11203–11222.
48. Wirsen, C. O., Brinkhoff, T., Kuever, J., Muyzer, G., Molyneaux, S., and Jannasch, H. W. (1998) Comparison of a new *Thiomicrospira* strain from the Mid-Atlantic Ridge with known hydrothermal vent isolates. *Appl. Environ. Microbiol.* *64*, 4057-4059.
49. Jannasch, H. W., Wirsen, C. O., Nelson, D. C., Robertson, L. A. (1985) *Thiomicrospira crunogena* sp. nov., a colorless sulfur-oxidizing bacterium from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* *35*, 422–424.
50. Bury-Mone, S., Mendz, G., Ball, G., Thibonnier, M., Stingl, K., Ecobichon, C., Ave, P., Huerre, M., Labitne, A., Thiberge, J., and Reuse, H. D. (2008) Roles of α and β carbonic anhydrases of *Helicobacter pylori* in the urease-dependent response to acidity and in colonization of the murine gastric mucosa. *Infect. Immun.* *76*, 497–509.
51. Badger, M. R., and Bek, E. J. (2008) Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO₂ acquisition by the CBB cycle. *J. Exp. Bot.* *59*, 1525-1541.
52. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* *340*, 783–795.
53. Petersen, T. N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* *8*, 785-786.
54. Avvaru, B.S., Kim, C.U., Sippel, K.H., Gruner, S.M., Agbandje-McKenna, M., Silverman, D.N., and McKenna, R. (2010). A short, strong hydrogen bond in the active site of human carbonic anhydrase II. *Biochemistry* *49*, 249-251.
55. Sprigings, T. G., and Hall, D. C. (2001) A simple carbonic anhydrase model which achieves catalytic hydrolysis by the formation of an ‘enzyme–substrate’-like complex. *J. Chem. Soc., Perkin Trans. 2*, 2063-2067.
56. Forsman, C., Behravan, G., Osterman, A., and Jonsson, B. H. (1988) Production of active human carbonic anhydrase II in *E. coli*. *Acta Chem. Scand. B* *42*, 314-318.
57. Khalifah, R. G., Strader, D. J., Bryant, S. H., Gibson, S. M. (1977) Carbon-13 nuclear magnetic resonance probe of active-site ionizations in human carbonic anhydrase B. *Biochemistry.* *17*, 2241–2247.

58. Gill, S. C, and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* 182, 319-326.
59. Díaz-Torres, N., González, G., Biswas, S., Scott, K., and McKenna, R. (2012) Preliminary X-ray crystallographic analysis of [alpha]-carbonic anhydrase from *Thiomicrospira crunogena* XCL-2. Manuscript submitted for publication.
60. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymology* 276, 307-326.
61. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D66*, 213-231.
62. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* 40, 658-674.
63. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. (2008) Iterative model building, structure refinement, and density modification with the PHENIX AutoBuild Wizard. *Acta Cryst D64*, 61-69.
64. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010) Features and development of Coot. *Acta Cryst. D66*, 486-501.
65. Brunger, A.T. (1992) "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures". *Nature* 355, 472-475.
66. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.
67. Frishman, D., and Argos, P. (1995) Knowledge-Based Protein Secondary Structure Assignment. *Proteins* 23, 566-579.
68. Silverman, D. N. (1982) Carbonic anhydrase: oxygen-18 exchange catalyzed by an enzyme with rate-contributing proton-transfer steps, *Methods Enzymol* 87, 732-752.
69. Silverman, D. N., Tu, C., Chen, X., Tanhauser, S. M., Kresge, A. J., and Laipis, P. J. (1993) Rate-equilibria relationships in intramolecular proton transfer in human carbonic anhydrase III. *Biochemistry* 32, 10757-10762.

70. Tu, C. K., Swenson, E. R., and Silverman, D. N. (2007) Membrane inlet for mass spectrometric measurement of nitric oxide. *Free Radical Biology and Medicine* 43, 1453-1457.
71. Simonsson, I., Jonsson, B. H., and Lindskog, S. (1979) C-13 NMR study of carbon dioxide-bicarbonate exchange catalyzed by human carbonic anhydrase-C at chemical-equilibrium. *European Journal of Biochemistry* 93, 409-417.
72. Tu, C.K., Silverman, D.N., Forsman, C., Jonsson, B.H., and Lindskog, S. (1989) Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* 28, 7913-7918.
73. Bruylants, G., Wouters, J., and Michaux, C. (2005) Differential scanning calorimetry in life science: Thermodynamics, stability, molecular recognition and application in drug design. *Current Medicinal Chemistry* 12, 2011-2020.
74. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* 33, 491-497.
75. Kantardjieff, K. A., and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA and protein-nucleic acid complex crystals. *Protein Science* 12, 1865-1871.
76. An, H., Tu, C., Duda, D., Montanez-Clemente, I., Math, K., Laipis, P. J., McKenna, R., and Silverman, D. N. (2002) Chemical rescue in catalysis by human carbonic anhydrases II and III. *Biochemistry* 41, 3235-3242.
77. Maren, T. H., and Sanyal, G. (1983) The Activity of Sulfonamides and Anions against the Carbonic-Anhydrases of Animals, Plants, and Bacteria. *Annual Review of Pharmacology and Toxicology* 23, 439-459.
78. Mikulski, R., Domsic, J. F., Ling, G., Tu, C., Robbins, A. H., Silverman, D. N. and McKenna, R. (2011) Structure and catalysis by carbonic anhydrase II: Role of active-site tryptophan 5. *Arch. Biochem. Biophys.* 516, 97-102.
79. Eriksson, A. E., Jones, T. A. and Liljas, A. (1988) Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins* 4, 274–282.
80. Fransson, C., Freskgård, P. O., Herbertsson, H., Johansson, A., Jonasson, P., Mårtensson, L. G., Svensson, M., Jonsson, B. H., and Carlsson, U. (1992) Cis-trans isomerization is rate-determining in the reactivation of denatured human carbonic anhydrase II as evidenced by proline isomerase. *FEBS Lett.* 296, 90–94.
81. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K. and Uversky, V.N. (2010) PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* 1804, 996–1010.

82. Dunker, A. K., Brown, C. J., Lawson, J., Iakoucheva L. M., and Obradović, Z. (2002) "Intrinsic disorder and protein function," *Biochemistry*, 41, 6573–6582.
83. Kadokura, H., Katzen, F., and Beckwith, J. (2003) Protein disulfide bond formation in prokaryotes. *Annual Review of Biochemistry* 42, 111–135.
84. Xue, Y. F., Liljas, A., Jonsson, B. H., and Lindskog, S. (1993) Structural-Analysis of the Zinc Hydroxide-Thr-199-Glu-106 Hydrogen-Bond Network in Human Carbonic Anhydrase-II, *Proteins-Structure Function and Genetics* 17, 93-106.
85. Jonasson, P., Aronsson, G., Carlsson, U., and Jonsson, B. H. (1997) Tertiary structure formation at specific tryptophan side chains in the refolding of human carbonic anhydrase II, *Biochemistry* 36, 5142-5148.
86. Stams, T., Nair, S. K., Okuyama, T., Waheed, A., Sly, W.S., and Christianson, D. W. (1996) Crystal structure of the secretory form of membrane-associated human carbonic anhydrase IV at 2.8 Å resolution. *Proc. Natl. Acad. Sci. USA* 93, 13589–13594.
87. Alterio, V., Hilvo, M., Di Fiore, A., Supuran, C. T., Pan, P., Parkkila, S., Scaloni, A., Pastorek, J., Pastorekova, S., Pedone, C., Scozzafava, A., Monti, S. M., De Simone, G. (2009) Crystal structure of the extracellular catalytic domain of the tumor-associated human carbonic anhydrase IX *Proc. Natl. Acad. Sci. U.S.A.* 106, 16233–16238.
88. Whittington, D. A., Waheed, A., Ulmasov, B., Shah, G. N., Grubb, J. H., Sly, W. S., and Christianson, D. W. (2001) Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells, *Proc Natl Acad Sci U S A* 98, 9545-9550.
89. Cuesta-Seijo, J. A., Borchert, M. S., Navarro-Poulsen, J. C., Schnorr, K. M., Mortensen, S. B., Lo Leggio, L. (2011) Structure of a dimeric fungal α -type carbonic anhydrase. *FEBS Letters*, 585, 1042-1048.
90. Waheed, A., Okuyama, T., Heyduk, T., Sly, W.S. (1996) Carbonic anhydrase IV: Purification of a secretory form of the recombinant human enzyme and identification of the positions and importance of its disulfide bonds. *Archives of Biochemistry and Biophysics* 333, 432–438.
91. Whitney, P. L., and Briggler, T. V. (1982) Membrane-associated carbonic anhydrase purified from bovine lung. *J. Biol. Chem.* 257, 12056-12059.
92. Elleby, B., Chirica, L.C., Tu, C., Zeppezauer, M., and Lindskog, S. (2001) Characterization of carbonic anhydrase from *Neisseria gonorrhoeae*. *Eur. J. Biochem.* 268, 1613-1619.

93. Mårtensson, L. G., Jonsson, B. H., Andersson, M., Kihlgren, A., Bergenhem, N. and Carlsson, U. (1992) Role of an evolutionarily invariant serine for the stability of human carbonic anhydrase II. *Biochim. Biophys. Acta* 1118, 179-186.
94. Pace, C. N. (1990) Conformational stability of globular proteins. *Trends Biochem. Sci.* 15, 14–17.
95. Pace, C. N., Grimsley, G. R., Thomson, J. A., and Barnett, B. J. (1988) Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* 263, 11820–11825.
96. Trivedi, M. V., Laurence, J. S., Siahaan, T. J. (2009) The role of thiols and disulfides on protein stability. *Curr. Protein Pept. Sci.* 10, 614–625.
97. Mårtensson, L. G., Karlsson, M., and Carlsson, U. (2002) Dramatic stabilization of the native state of Human carbonic anhydrase II by an engineered disulfide bond. *Biochemistry* 41, 15867-15875.
98. Pechkova, E., Sivozhelezov, V., and Nicolini, C. (2007) Protein thermal stability: The role of protein structure and aqueous environment. *Archives of Biochemistry and Biophysics*, 466, 40-48.
99. Vogt, G., Woell, S., and Argos, P. (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *J Mol Biol.* 269, 631–643.
100. Kannan, K. K., Liljas, A., Waara, I., Bergstén, P. C., Lövgren, S., Strandberg, B., Bengtsson, U., Carlbom, U., Fridborg, K., Järup, L., and Petef, M. (1972) Crystal structure of human erythrocyte carbonic anhydrase C. VI. The three-dimensional structure at high resolution in relation to other mammalian carbonic anhydrases. *Cold Spring Harb. Symp. Quant. Biol.* 36, 221-31.
101. Suzuki, K., Yang, S. Y., Shimizu, S., Morishita, E. C., Jiang, J., Zhang, F., Hoque, M. M., Sato, Y., Tsunoda, M., Sekiguchi, T., and Takénaka, A. (2011) The unique structure of carbonic anhydrase α -CA1 from *Chlamydomonas reinhardtii*. *Acta Cryst. D67*, 894-901.
102. Krissinel, E, and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774-797.
103. Mehler, E. L., Fuxreiter, M., Simon, I., and Garcia-Moreno, E. B. (2002) The role of hydrophobic microenvironments in modulating pK_a shifts in proteins. *Proteins* 48, 283–292.
104. Aldred, P., Fu, P., Barrett, G., Penschow, J. D., Wright, R. D., Coghlan, J. P., and Fernley, R. T. (1991) Human secreted carbonic anhydrase: cDNA cloning, nucleotide sequence, and hybridization histochemistry. *Biochemistry* 30, 569-575.

105. Ponstingl, H., Henrick, K. and Thornton, J. M. (2000) Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins Struct. Funct. Genet.* 41, 47–57.
106. Larkin, M. A, Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins D. G. (2007) ClustalW and ClustalX version 2.0. *Bioinformatics* 23(21): 2947-2948.
107. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research* 38 Suppl, W695-699.

BIOGRAPHICAL SKETCH

Natalia has been involved in research since she was a sophomore at the University of Puerto Rico Rio Piedras' Campus, where she received the Minority Access to Research Careers Undergraduate Student Training Award. She used to work with snake venom phospholipases A_2 under the guidance of Dr. Reginald Morales. During her junior year, she completed a summer internship at Brown University as part of the Leadership Alliance Summer Research Early Identification Program, and worked with Dr. Kimberly Mowry, studying Vg1 mRNA localization in *Xenopus laevis*' oocytes. At the conclusion of her internship, she presented her projects at two national meetings. Natalia graduated from the University of Puerto Rico with a Bachelor of Science in Chemistry, and was awarded *Magna cum Laude* honors.

Natalia joined the McKenna laboratory in the spring of 2011. Her thesis project was based on the characterization of an α -carbonic anhydrase expressed in the deep-sea vent *Thiomicrospira crunogena* gammaproteobacterium. During her leisure time, Natalia volunteered at a local hospital a few hours a week, and played catcher for the IDP softball team "Human Centipedes". After obtaining her master's degree, Natalia aspires to venture into a career in medicine.