

UNDERSTANDING SUBSTRATE BINDING AND MOVEMENT IN
PROLINE CATABOLIC ENZYMES

A Dissertation presented to the Faculty of the Graduate School at the University of
Missouri-Columbia

In Partial Fulfillment of the Requirements for the Degree
Doctor of Philosophy

By TRAVIS PEMBERTON

Dr. John J. Tanner
Dissertation Supervisor

JULY 2014

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation
entitled

UNDERSTANDING LIGAND BINDING AND MOVEMENT IN PROLINE CATABOLIC ENZYMES

presented by Travis Pemberton,

a candidate for the degree of doctor of philosophy of Chemistry,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor John J. Tanner

Professor Lesa Beamer

Professor Jason Cooley

Professor John Smith

Professor Kent Gates

ACKNOWLEDGEMENTS

This dissertation represents the work I have accomplished during my time at the University of Missouri, Columbia. Here I would like to take a moment and acknowledge those who helped me along the way.

First and foremost, I would like to thank my advisor Dr. Jack Tanner. The year I joined many people wanted to join his lab and I was one of the people he allowed in his group. I joined this group knowing very little about crystallography, but Dr. Tanner teaches in a way that makes even this complex subject understandable. When hitting road blocks in my research he always offered advice and guidance. His line of questioning and rationale forced me to approach problems in a more scientific way. The constant attention to detail, double checking, and redouble checking results, while frustrating, has instilled in me the importance of being absolutely sure of something before proclaiming it. His dedication to work was apparent from his late nights in lab to working through the weekends. He teaches by example, experience, and thoughtful guidance. I have greatly enjoyed working for and learning under Dr. Tanner, and with his guidance I believe I have become a better scientist.

I would also like to thank Dr. Dhiraj Srivastava. He taught me the basics in lab and was always around to help with questions on lab protocol, problems in my research, and troubleshooting projects. On that same token, I would also like to thank Dr. Harkewal Singh who was always willing to lend his expertise and offer a new perspective on issues I ran in to. Both, Dr. Srivastava and Dr. Singh, were extremely helpful and made lab an enjoyable place to work. I would like to thank Min Luo, who started at the

same time as me, for always being around to discuss new ideas about research projects and being supportive when things were not going well. Finally, I would like to thank my committee members: Dr. Lesa Beamer, Dr. Jason Cooley, and Dr. Kent Gates. I tried not to bother them with too much, but their thoughts, suggestions, and guidance are greatly appreciated.

TABLE OF CONTENTS

Acknowledgements.....	ii
List of Figures.....	xiii
List of Tables.....	xvi
Chapter 1: Introduction.....	1
Summary of Research.....	6
References.....	7
Chapter 2: Proline: Mother Nature’s Cryoprotectant Applied to Macromolecular Crystallography.....	11
Abstract.....	12
2.1 Introduction.....	13
2.2 Materials and Methods.....	14
2.2.1 Loop Diffraction Tests.....	14
2.2.2 Crystallization and Cryoprotection.....	15
2.2.2.1 General Cryoprotection Procedure.....	15
2.2.2.2 Lysozyme.....	16
2.2.2.3 Xylose Isomerase.....	16
2.2.2.4 Histidine Acid Phosphatase.....	18
2.2.2.5 Δ^1 -Pyrroline-5-Carboxylate Dehydrogenase.....	18
2.2.3 Data Collection and Refinement.....	19
2.3 Results.....	20
2.3.1 Loop Diffraction Tests.....	20

2.3.2 Diffraction Data.....	21
2.3.3 Ordered Proline Molecules.....	24
2.4 Discussion	25
Acknowledgements.....	27
References	27
Figure Legends.....	36
Chapter 3: Structural Basis of Substrate Selectivity of Δ1-Pyrroline-5-Carboxylate Dehydrogenase (ALDH4A1): Semialdehyde Chain Length.....	43
Abstract.....	44
3.1 Introduction.....	45
3.2 Experimental Procedures.....	47
3.2.1 Crystallization and Soaking.....	47
3.2.2 X-ray Data Collection and Refinement	48
3.2.3 Steady-state Inhibition Kinetics	48
3.3 Results	49
3.3.1 5- and 4-Carbon Carboxylate Ligands: Glutarate and Succinate	49
3.3.2 3- and 2-Carbon Ligands: Malonate, Glyoxylate, and Acetate	51
3.3.3 Inhibition of MmP5CDH.....	52
3.4 Discussion	53
Acknowledgements.....	57
References	57

Figure Legends.....	64
Chapter 4: Crystal Structure and Unexpected Oligomeric State of Put2p, the Mitochondrial Δ^1-Pyrroline-5-Carboxylate Dehydrogenase from <i>Saccharomyces cerevisiae</i>.....	75
Abstract.....	76
4.1 Introduction.....	77
4.2 Experimental Procedures.....	79
4.2.1 Subcloning and Mutagenesis.....	79
4.2.2 Expression and Purification of Put2p.....	79
4.2.3 Crystallization of Put2p.....	80
4.2.4 Expression, Purification, and Crystallization of Tag-free HsP5CDH.....	81
4.2.5 X-ray Diffraction Data Collection, Phasing, and Refinement.....	82
4.2.6 Small-Angle X-ray Scattering of Put2p.....	83
4.2.7 Analytical Ultracentrifugation.....	84
4.2.8 Steady-State Kinetics.....	85
4.3 Results.....	85
4.3.1 Steady-State Kinetics Measurements.....	85
4.3.2 Protomer Structure of Put2p.....	86
4.3.3 Structure of Put2p-NAD ⁺ Complex.....	87
4.3.4 Structure of Monoclinic HsP5CDH.....	88
4.3.5 Oligomeric State and Quaternary Structure of Put2p.....	88
4.3.6 Hexamerization Hot Spot of Put2p.....	90
4.4 Discussion.....	91

Acknowledgements	93
References	94
Figure Legends.....	100
Chapter 5: Kinetic and Structural Characterization of Tunnel-Perturbing Mutants in Bradyrhizobium japonicum Proline Utilization A (PutA)	110
Abstract	111
5.1 Introduction.....	112
5.2 Experimental Procedures.....	114
5.2.1 Chemicals	114
5.2.2 Site-Directed Mutagenesis.....	114
5.2.3 Expression and Purification of BjPutA Proteins	115
5.2.4 Steady-State Kinetic Assays.....	116
5.2.5 NAD ⁺ Binding to BjPutA.....	118
5.2.6 Single-Turnover Kinetic Experiments.....	120
5.2.7 Crystal Structure Determination and Structure Analysis	120
5.3 Results	122
5.3.1 Rationale for Channel-Blocking Mutagenesis and Purification of BjPutA Mutant Enzymes	122
5.3.2 Channeling Assays of BjPutA Mutants	123
5.3.3 Steady-State Kinetic Properties of BjPutA Wild-type Mutants	124
5.3.4 Single Turnover Rapid-Reaction Kinetics.....	125
5.3.5 Alternative P5CDH Substrates	126

5.3.6 Crystal Structures of D778Y, D779Y, and D779W	126
5.4 Discussion	128
Acknowledgements	132
References	133
Figure Legends.....	142
Chapter 6: Understanding Access to an Internal Cavity Discovered in Bradyrhizobium japonicum Proline Utilization A (PutA)	153
Abstract	154
6.1 Introduction.....	155
6.2 Materials and Methods.....	157
6.2.1 Cloning and PCR	157
6.2.2 Protein Expression and Purification	157
6.2.3 Protein Crystallization	159
6.2.4 Data Collection and Refinement	160
6.2.5 Steady-State Kinetics.....	160
6.3 Results	161
6.3.1 Proline Reduced BjPutA Structure.....	161
6.3.2 E460A and E197A.....	164
6.3.3 R456M.....	166
6.4 Discussion	167
6.4.1 Proline Reduced Wild-type BjPutA	167
6.4.2 E197A.....	168

6.4.3 E460A.....	170
6.4.4 R456M.....	171
6.5 Conclusion.....	172
References	172
Figure Legends.....	180
Chapter 7: Structure and Role of Mg²⁺ Binding of the Apoptosis-Linked Protein (ALG-2)	192
Abstract	193
7.1 Introduction.....	194
7.2 Materials and Methods.....	196
7.2.1 Proline Preparation for Crystal Trays.....	196
7.2.2 Crystallization Trials	196
7.2.3 Cryoprotection.....	198
7.2.4 Data Collection and Processing.....	199
7.3 Results	200
7.3.1 Enzyme Precipitation.....	200
7.3.2 Diffraction Quality and Cryoprotection	200
7.3.3 Metal Binding Sites	202
7.4 Discussion	203
7.5 Future Work.....	207
References	208
Figure Legends.....	212

APPENDIX

A1. Expression, Purification, Crystallization, and Small Angle X-ray Scattering (SAXS) of the Δ1-986 <i>Bradyrhizobium japonicum</i> Proline Utilization A (BjFlap)	220
A1.1 Rationale.....	221
A1.2 Materials and Methods.....	221
A1.2.1 Protein Expression and Purification	221
A1.2.2 Crystallization.....	223
A1.2.3 Data Collection and Refinement.....	224
A1.2.4 SAXS	224
A1.2.5 PISA.....	225
A1.3 Results.....	226
A1.3.1 Crystallography	226
A1.3.2 SAXS	226
A1.4 Discussion	227
A1.4.1 Structure Analysis of the Flap Mutant.....	227
A1.4.2 C-Terminal Deletion Disruption of Tetrameric Oligomerization.....	228
A1.4.3 Possible Mutations for Disruption of the Tetramer	230
Figure Legends.....	233
A2. Protein Expression, Purification, Crystallization and SAXS Analysis of <i>Legionella pneumophila</i> PutA (LpPutA)	236
A2.1 Rationale.....	237

A2.2 Materials and Methods.....	238
A2.2.1 Protein Expression and Purification	238
A2.2.2 Crystallography	238
A2.2.3 SAXS	239
A2.3 Results.....	239
A2.3.1 Expression and Crystallography	239
A2.3.2 Molecular Replacement and PISA	240
A2.3.3 SAXS	241
A2.4 Discussion	242
Figure Legends.....	245
A3. Expression, Purification, Crystallization and SAXS	
Analysis of <i>Desulfovibrio vulgaris</i> PutA (DvPutA)	249
A3.1 Rationale.....	250
A3.2 Materials and Methods.....	251
A3.2.1 Protein Expression and Purification	251
A3.2.2 Crystallography	252
A3.2.3 SAXS	253
A3.3 Results.....	253
A3.3.1 Expression and Crystallography	253
A3.3.2 Molecular Replacement and PISA	254
A3.3.3 SAXS	254
References	255

Figure Legends.....	260
Vita	264

LIST OF FIGURES

Figure 2.1: X-ray Diffraction Images of Proline Cryoprotectants	38
Figure 2.2: Proline in Lysozyme Structure	39
Figure 2.3: Proline in Xylose Isomerase Structure	40
Figure 2.4: Proline in Histidine Acid Phosphatase Structure	41
Figure 2.5: Proline in P5CDH Structure	42
Figure 3.1: Proline and Hydroxy-Proline Catabolism Pathways	67
Figure 3.2: Overall MmP5CDH Structure Complexed with Glutarate	68
Figure 3.3: Ligands Used for MmP5CDH Crystal Soaking	69
Figure 3.4: Binding of 5- and 4-Carbon Ligands to MmP5CDH	70
Figure 3.5: Binding of 3- and 2-Carbon Ligands to MmP5CDH	71
Figure 3.6: Essential Elements of Semialdehyde recognition by MmP5CDH	72
Figure 3.7: Comparison of L-Proline and Glyoxylate to MmP5CDH.....	73
Figure 3.8: Steady-state Kinetics of MmP5CDH	74
Figure 4.1: Proline Catabolism Pathway	102
Figure 4.2: Structure of Put2p.....	103
Figure 4.3: Electron Density Evidence of Binding Loop Disorder	104
Figure 4.4: Electron Density and Interactions for NAD ⁺ in Put2p.....	105
Figure 4.5: SAXS Analyses of Put2p	106
Figure 4.6: Hexamer Hot Spot of Put2p	107
Figure 4.7: Sedimentation Velocity Analysis of Put2p W193A Mutant	108
Figure 4.8: SAXS Analyses of Put2p W193A Mutant	109
Scheme 5.1: Proline Catabolism Pathway.....	138

Figure 5.1: Tunnel System and Cavities Calculated for BjPutA	145
Figure 5.2: Channeling Assays of BjPutA Wild-Type and Mutants	146
Figure 5.3: Channeling Assays with Increasing Concentrations of Mutants.....	147
Figure 5.4: NAD ⁺ Binding to BjPutA.....	148
Figure 5.5: Electron Density Maps and Local Conformational Changes	149
Figure 5.6: Constriction of the Channeling Tunnel by D778Y	150
Figure 5.7: Constriction of the Channeling Tunnel in D779Y	151
Figure 5.8: Constriction of the Channeling Tunnel in D779YW	152
Scheme 6.1: Proline Catabolism Pathway.....	177
Figure 6.1: Reduced FAD from the Proline Soaked BjPutA Structure	183
Figure 6.2: Channeling Assays of BjPutA Wild-Type and Mutants	184
Figure 6.3: Movement of the Ion-Pair Gate in BjPutA.....	185
Figure 6.4: Reduced and Oxidized BjPutA Cavities	186
Figure 6.5: Ion-Pair Comparison of the BjPutA R456M Mutants.....	187
Figure 6.6: Movement of Glycine 196 in Reduced BjPutA	188
Figure 6.7: Unstructuring of Helix 8 in the BjPutA E197A Mutant	189
Figure 6.8: Specific Movement Within Helix 8 in E197A Mutant	190
Figure 6.9: Steady-State Kinetics of E197A Mutant	191
Figure 7.1: Form 1 ALG-2 Crystals.....	213
Figure 7.2: Form 2 ALG-2 Crystals.....	214
Figure 7.3: Diffraction Pattern of Form 1 ALG-2	215
Figure 7.4: Diffraction Pattern of Form 2 ALG-2	216
Figure 7.5: Overall Fold and Structure of the ALG-2 Dimer	217

Figure 7.6: Magnesium Binding Pockets in ALG-2	218
Figure A1.1: SAXS Analyses of BjFlap	234
Figure A1.2: Missing C-terminal of BjFlap.....	235
Figure A2.1: SAXS Analyses of LpPutA	246
Figure A2.2: Oligomer found by PISA	247
Figure A2.3: Oligomer formed by crystallographic symmetry	248
Figure A3.1: Diffraction Pattern of DvPutA	261
Figure A3.2: SAXS Analyses of DvPutA.....	262
Figure A3.3: Oligomer Found by PISA	263

LIST OF TABLES

Table 2.1: Proline-Based Cryobuffers	31
Table 2.2: Data Collection and Refinement Statistics for HEWL.....	32
Table 2.3: Data Collection and Refinement Statistics for XI	33
Table 2.4: Data Collection and Refinement Statistics for FtHAP	34
Table 2.5: Data Collection and Refinement Statistics for MmP5CDH	35
Table 3.1: Data Collection and Refinement Statistics MmP5CDH Ligand Soaks.....	61
Table 3.2: Kinetic Parameters for the Inhibition of MmP5CDH.....	62
Table 3.3: Inferred Geometrical Parameters for Active and Inactive Semialdehydes	63
Table 4.1: Data Collection and Refinement Statistics for Put2p	99
Table 5.1: Primers Used for Site-Directed Mutagenesis for BjPutA Tunnel Mutants	139
Table 5.2: PRODH Kinetic Parameters	139
Table 5.3: P5CDH Kinetic and NAD ⁺ Binding Parameters	140
Table 5.4: Data Collection and Refinement Statistics for BjPutA Tunnel Mutants	141
Table 6.1: Primers for Site-Directed Mutagenesis for BjPutA Ion-Pair Mutants.....	178
Table 6.2: Data Collection and Refinement Statistics for BjPutA Ion-Pair Mutants	179
Table 7.1: Data Collection and Refinement Statistics for ALG-2.....	210
Table 7.2: ALG-2 Binding Pocket Coordination Comparison	211
Table A1.1: Data Collection and Refinement Statistics for BjFlap.....	232
Table A2.1: MolRep Space Group Determination Results	243
Table A2.2: Data Collection Statistics for LpPutA	244
Table A3.1: MoleRep Space Group Determination	258
Table A3.2: Data Collection Statistics for DvPutA.....	259

Chapter 1
Introduction

Proline is a unique amino acid. It has the highest solubility of any amino acid reaching a solubility limit of 6.5 M [H₂O 20 °C] (1). The side chain bonds with the amino group, creating a secondary amine rather than a primary amine. This unique side chain creates a pyrrolidine ring which is thought to contribute heavily to its diverse functions and roles.

Although proline has never been investigated as a possible cryo-additive for macromolecular cryocrystallography, there is extensive literature on its unique properties and unusual functions (1-7). Plants, bacteria, invertebrates, protozoa and algae all accumulate proline during water stress, including cold acclimation to increase freezing tolerance (7). Proline also can increase solubility of some compounds, including proteins and be added to guard against denaturation during dehydration (4). Further, cells accumulate proline while experiencing water stress, which fits with its known role as an effective osmolyte. When cells are under water stress, dehydration, high salt, or low temperature, they up regulate proline synthesis and down regulate proline degradation. Proline has also been known to increase viability of cells after freezing when added exogenously, usually better than other cryoprotectants such as glycerol (6).

The unique function of proline has led cells to evolve specialized enzymes to control proline metabolism. Proline catabolism to glutamate happens in a two-step, four electron oxidation catalyzed by two enzymes. First, proline is oxidized to Δ^1 -pyrroline-5-carboxylate (P5C) the flavin dependent proline dehydrogenase (PRODH). PRODH uses flavin adenine dinucleotide (FAD) as an electron acceptor. Reduction of FAD to FADH₂ disrupts the aromaticity of the isoalloxazine ring and reorganizes the ribityl chain. P5C then goes through a non-enzymatic, pH dependent hydrolysis to glutamate γ -

semialdehyde (GSA) (8). The final two electron oxidation is catalyzed by P5C dehydrogenase (P5CDH), which oxidizes GSA to glutamate using NAD^+ as the electron acceptor.

PRODH and P5CDH are conserved in bacteria and eukaryotes. They are arranged into three different domains, or branches, based on sequence analysis of PRODHs (9, 10). In branches 1 and 2 the two enzymes are combined into a single polypeptide; a bifunctional enzyme known as proline utilization A (PutA) (11). PutA appears to be limited to Gram-negative bacteria. Branch 1 houses PutAs from alpha-, beta-, and gamma-proteobacteria and ranges in chain length from 999 to about 1400 residues. The overall sequence identity within this branch is about 38%. Branch 1 further has a subset of trifunctional PutAs that have a DNA-binding domain located in the N-terminus of the enzyme. These enzymes tend to be longer, about 1270-1361 residues, and have the third function of repressing transcription of the *put* regulon. High levels of proline in the bacterium cause the enzyme to disengage from the regulon and allow transcription of PutA. Branch 2 contains only bifunctional enzymes from delta- and epsilon-proteobacteria as well as cyanobacteria with a chain length of 980-1300 residues. The sequence identity within this branch is as low as 23% suggesting a more diverse group, even without the presence of trifunctional enzymes (9, 12).

To date the structures of *Bradyrhizobium japonicum* PutA (BjPutA) (13, 14) from branch 1 and *Geobacter sulfurreducens* PutA (GsPutA) (15) from branch 2 have been determined. BjPutA was the first structure solved and was the only full length structure solved for some time. This structure revealed the presence of a large, internal, water-filled cavity that linked the PRODH domain with the P5CDH domain. This tunnel is

about 75 Å in length and implies the channeling of the intermediate P5C/GSA. Steady-state kinetics and a fluorescence assay also had results indicative of channeling the intermediate between active sites. The BjPutA PRODH active site was blocked from the main cavity by a universally conserved ion pair between R456 and E197. It was hypothesized that this ion pair was a gate that allowed access to the tunnel from the PRODH active site. Unfortunately, due to the high concentration of ammonium sulfate, 2 M, and the presence of sulfate in the active site, no structural information from ligand soaking with proline analogs and inhibitors could be performed (13).

GsPutA revealed a similar cavity with kinetics matching the results expected for channeling. GsPutA had the added advantage of being higher resolution (1.8, as compared to 2.1 Å) and having a precipitating solution that did not contain ammonium sulfate. This gave more in depth structural analysis to be done and saw the conformational changes revolved around the breaking of the ion pair and lending strength to the idea of this being a gate to the main cavity (15).

Branch 3 contains Gram-positive (3A) and eukaryotic (3B) enzymes. However, these enzymes are monofunctional and encoded on separate and distinct genes (9). There is no presence of a large internal cavity linking the two and to date there is no clear evidence suggesting that these two monofunctional enzymes interact and channel substrate. In eukaryotes, PRODH is found interacting with the inner mitochondrial membrane (16, 17), while P5CDH is found as a soluble enzyme in the mitochondrial matrix (18, 19). The product of proline catabolism, glutamate, can go through deamination and enter the TCA cycle as α -ketoglutarate. Ornithine gets converted to GSA/P5C by ornithine amino transferase, allowing cross talk between the urea cycle and

proline metabolism. Further, P5C produced from PRODH can be transported out of the mitochondria to the cytosol where it can be reduced back to proline by P5C reductase and transported back into the mitochondria (20). This redox shuttling transfers redox equivalents from the pentose phosphate pathway to the electron transport chain and could explain eukaryotes desire to uncouple the PRODH and P5CDH activities.

Further, monofunctional P5CDH is an aldehyde dehydrogenase and belongs to the aldehyde dehydrogenase (ALDH) superfamily. The ALDH family is huge with hundreds of distinct genes found in all forms of life (21). Specifically, P5CDH is ALDH4A1 and has the characteristic Rossmann fold found for binding NAD^+ (22). P5CDH has been known to have a broad substrate specificity, oxidizing several semialdehydes (23). The best example is its ability to oxidize both P5C and hydroxy-P5C (24), however unique and distinct PRODH enzymes are needed to oxidize proline and hydroxyproline. Several common mutations of the human P5CDH gene have been reported to cause a loss of function resulting in the metabolic disorder type II hyperprolinemia (25-27). PRODH and P5CDH genes are activated by p53, a known tumor suppressor which indicates proline catabolism in apoptosis, stress, and even cancer (28, 29).

Summary of Research.

My doctorate work has been separated into several parts. First, investigation into the channeling found in PutA enzymes. This study involved solving the structure of several mutant enzymes of BjPutA including D779Y, Δ 1-986, E460A, and E197A. Further, soaking of the wild type structure with a high concentration shed light on the movement of the proposed ion-pair gate. This soak also spurred on research that ultimately found that proline was usable as a cryoprotectant for macromolecular crystallography. This study led to crystallization and structure determination of BjPutA, several BjPutA mutants, lysozyme, glucose isomerase, histidine acid phosphatase and mouse P5CDH. Several studies were done on P5CDH, specifically the mouse P5CDH. This study involved ligand soaking with varying sizes and chain lengths of semialdehydes to better understand the enzymes substrate selectivity. The result was the characterization of a binding loop that played a critical role in anchoring the substrate in the correct orientation for catalytic activity. Studies on the yeast P5CDH (Put2) were also carried out, but the crystal structure showed an inherent flexibility of this binding loop characterized by the mouse studies. Lastly extensive small angle X-ray scattering (SAXS) was performed to determine how we can influence oligomeric states.

References

1. Troitzsch, R. Z., Martyna, G. J., McLain, S. E., Soper, A. K., and Crain, J. (2007) Structure of aqueous proline via parallel tempering molecular dynamics and neutron diffraction, *The journal of physical chemistry. B* *111*, 8210-8222.
2. Troitzsch, R. Z., Vass, H., Hossack, W. J., Martyna, G. J., and Crain, J. (2008) Molecular mechanisms of cryoprotection in aqueous proline: light scattering and molecular dynamics simulations, *The journal of physical chemistry. B* *112*, 4290-4297.
3. McLain, S. E., Soper, A. K., Terry, A. E., and Watts, A. (2007) Structure and hydration of L-proline in aqueous solutions, *The journal of physical chemistry. B* *111*, 4568-4580.
4. Rudolph, A. S., and Crowe, J. H. (1986) A calorimetric and infrared spectroscopic study of the stabilizing solute proline, *Biophysical journal* *50*, 423-430.
5. Troitzsch, R. Z., Tulip, P. R., Crain, J., and Martyna, G. J. (2008) A simplified model of local structure in aqueous proline amino acid revealed by first-principles molecular dynamics simulations, *Biophysical journal* *95*, 5014-5020.
6. Withers, L. A., and King, P. J. (1979) Proline: A Novel Cryoprotectant for the Freeze Preservation of Cultured Cells of *Zea mays* L, *Plant physiology* *64*, 675-678.
7. Yoshiba, Y., Kiyosue, T., Nakashima, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1997) Regulation of levels of proline as an osmolyte in plants under water stress, *Plant & cell physiology* *38*, 1095-1102.
8. Bearne, S. L., and Wolfenden, R. (1995) Glutamate gamma-semialdehyde as a natural transition state analogue inhibitor of *Escherichia coli* glucosamine-6-phosphate synthase, *Biochemistry* *34*, 11515-11520.
9. Tanner, J. J. (2008) Structural biology of proline catabolism, *Amino acids* *35*, 719-730.

10. White, T. A., Krishnan, N., Becker, D. F., and Tanner, J. J. (2007) Structure and kinetics of monofunctional proline dehydrogenase from *Thermus thermophilus*, *The Journal of biological chemistry* 282, 14316-14327.
11. Ratzkin, B., and Roth, J. (1978) Cluster of genes controlling proline degradation in *Salmonella typhimurium*, *Journal of bacteriology* 133, 744-754.
12. Singh, R. K., and Tanner, J. J. (2012) Unique structural features and sequence motifs of proline utilization A (PutA), *Front Biosci (Landmark Ed)* 17, 556-568.
13. Srivastava, D., Schuermann, J. P., White, T. A., Krishnan, N., Sanyal, N., Hura, G. L., Tan, A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2010) Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*, *Proceedings of the National Academy of Sciences of the United States of America* 107, 2878-2883.
14. Schuermann, J. P., White, T. A., Srivastava, D., Karr, D. B., and Tanner, J. J. (2008) Three crystal forms of the bifunctional enzyme proline utilization A (PutA) from *Bradyrhizobium japonicum*, *Acta crystallographica. Section F, Structural biology and crystallization communications* 64, 949-953.
15. Singh, H., Arentson, B. W., Becker, D. F., and Tanner, J. J. (2014) Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-channeling tunnel and the quinone-binding site, *Proceedings of the National Academy of Sciences of the United States of America* 111, 3389-3394.
16. Brunner, G., and Neupert, W. (1969) Localisation of proline oxidase and Delta-pyrroline-5-carboxylic acid dehydrogenase in rat liver, *FEBS letters* 3, 283-286.
17. Wang, S. S., and Brandriss, M. C. (1987) Proline utilization in *Saccharomyces cerevisiae*: sequence, regulation, and mitochondrial localization of the PUT1 gene product, *Molecular and cellular biology* 7, 4431-4440.
18. Brandriss, M. C., and Krzywicki, K. A. (1986) Amino-terminal fragments of delta 1-pyrroline-5-carboxylate dehydrogenase direct beta-galactosidase to the mitochondrial matrix in *Saccharomyces cerevisiae*, *Molecular and cellular biology* 6, 3502-3512.

19. Haslett, M. R., Pink, D., Walters, B., and Brosnan, M. E. (2004) Assay and subcellular localization of pyrroline-5-carboxylate dehydrogenase in rat liver, *Biochimica et biophysica acta* 1675, 81-86.
20. Hagedorn, C. H., and Phang, J. M. (1986) Catalytic transfer of hydride ions from NADPH to oxygen by the interconversions of proline and delta 1-pyrroline-5-carboxylate, *Archives of biochemistry and biophysics* 248, 166-174.
21. Sophos, N. A., and Vasiliou, V. (2003) Aldehyde dehydrogenase gene superfamily: the 2002 update, *Chemico-biological interactions* 143-144, 5-22.
22. Liu, Z. J., Sun, Y. J., Rose, J., Chung, Y. J., Hsiao, C. D., Chang, W. R., Kuo, I., Perozich, J., Lindahl, R., Hempel, J., and Wang, B. C. (1997) The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold, *Nature structural biology* 4, 317-326.
23. Forte-McRobbie, C. M., and Pietruszko, R. (1986) Purification and characterization of human liver "high Km" aldehyde dehydrogenase and its identification as glutamic gamma-semialdehyde dehydrogenase, *The Journal of biological chemistry* 261, 2154-2163.
24. Adams, E., and Frank, L. (1980) Metabolism of proline and the hydroxyprolines, *Annual review of biochemistry* 49, 1005-1061.
25. Efron, M. L. (1965) Familial Hyperprolinemia. Report of a Second Case, Associated with Congenital Renal Malformations, Hereditary Hematuria and Mild Mental Retardation, with Demonstration of an Enzyme Defect, *The New England journal of medicine* 272, 1243-1254.
26. Geraghty, M. T., Vaughn, D., Nicholson, A. J., Lin, W. W., Jimenez-Sanchez, G., Obie, C., Flynn, M. P., Valle, D., and Hu, C. A. (1998) Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia, *Human molecular genetics* 7, 1411-1415.
27. Valle, D., Goodman, S. I., Applegarth, D. A., Shih, V. E., and Phang, J. M. (1976) Type II hyperprolinemia. Delta1-pyrroline-5-carboxylic acid dehydrogenase deficiency in cultured skin fibroblasts and circulating lymphocytes, *The Journal of clinical investigation* 58, 598-603.

28. Donald, S. P., Sun, X. Y., Hu, C. A., Yu, J., Mei, J. M., Valle, D., and Phang, J. M. (2001) Proline oxidase, encoded by p53-induced gene-6, catalyzes the generation of proline-dependent reactive oxygen species, *Cancer research* 61, 1810-1815.
29. Liu, Y., Borchert, G. L., Donald, S. P., Diwan, B. A., Anver, M., and Phang, J. M. (2009) Proline oxidase functions as a mitochondrial tumor suppressor in human cancers, *Cancer research* 69, 6414-6422.

Chapter 2:
Proline: Mother Nature's Cryoprotectant
Applied to Macromolecular Crystallography

Abstract

L-Proline is one of Mother Nature's cryoprotectants. Plants and yeast accumulate proline under freeze-induced stress and the use of proline in the cryopreservation of biological samples is well established. Here, it is shown that L-proline is also a useful cryoprotectant for protein crystallography. Proline was used to prepare crystals of lysozyme, xylose isomerase, histidine acid phosphatase and 1-pyrroline-5-carboxylate dehydrogenase for low-temperature data collection. The crystallization solutions in these test cases included the commonly used precipitants ammonium sulfate, sodium chloride and polyethylene glycol and spanned the pH range 4.6–8.5. Thus, proline is compatible with typical protein-crystallization formulations. The proline concentration needed for cryoprotection of these crystals is in the range 2.0–3.0 *M*. Complete data sets were collected from the proline-protected crystals. Proline performed as well as traditional cryoprotectants based on the diffraction resolution and data-quality statistics. The structures were refined to assess the binding of proline to these proteins. As observed with traditional cryoprotectants such as glycerol and ethylene glycol, the electron-density maps clearly showed the presence of proline molecules bound to the protein. In two cases, histidine acid phosphatase and 1-pyrroline-5-carboxylate dehydrogenase, proline binds in the active site. It is concluded that L-proline is an effective cryoprotectant for protein crystallography.

2.1. Introduction

X-ray diffraction data collection at cryogenic temperatures (~ 100 K) is nearly universal in macromolecular crystallography today. As water occupies a substantial fraction (~ 0.5) of the volume of macromolecular crystals, crystalline ice typically forms upon exposure of the untreated protein crystal to cryogenic temperatures. Crystalline ice can degrade diffraction quality, and the intense powder diffraction rings from microcrystalline ice compromise the adjacent protein reflections. This problem is typically prevented by adding a high concentration of a solute, known as the cryoprotectant, to the mother liquor or stabilizing solution (1). Commonly used cryoprotectants include glycerol, ethylene glycol, 2-methyl-2,4-pentanediol, low-molecular-weight polyethylene glycols (PEGs) and sucrose (2-4). Lithium salts and sodium malonate can also be used as cryoprotectants in some cases (5, 6). More recently, the osmolytes trimethylamine *N*-oxide, sarcosine and betaine have been used successfully as cryoprotective agents for protein crystals (7, 8).

A motivation for the current study is nature's use of the amino acid L-proline as a cryoprotectant. Plants accumulate proline in response to environmental stresses, including freezing temperatures (9-11). For example, early studies on coastal bermudagrass shoots showed that drought stress caused an increase of proline from less than 0.1 mg per gram of dry weight to over 15 mg per gram (12). Also, a 500-fold increase in free proline to levels as high as 60 mM has been observed in water-stressed tomato-plant cells (13). Proline also protects yeast against freeze stress (14-16). Gene-knockout studies have shown that disruption of proline catabolism in yeast improves freeze tolerance and that the mutant yeast strains accumulate up to 9% of the cell's dry

weight in proline (16). In addition, the freeze tolerance of certain fly larva is a consequence of elevated levels of proline (17, 18). In some cold-acclimated larvae, for example, the proline concentration reaches 147 mM (17).

The role of proline in freeze tolerance *in vivo* has prompted the use of the amino acid in the cryopreservation of biological samples *in vitro*. For example, cultured cells of maize have been freeze-preserved in 10%(w/v) proline (19). Also, proline at 27 mM has been used for the preservation of ram sperm (20). Additionally, low levels of proline [1%(w/v)] have been used in conjunction with other solutes in the cryopreservation of human stem cells (21). To our knowledge, proline has not been used as a cryoprotectant for protein crystals.

Here, we demonstrate the use of proline in the cryoprotection of crystals of hen egg-white lysozyme, xylose isomerase, histidine acid phosphatase and 1-pyrroline-5--carboxylate dehydrogenase. Proline was found to perform as well as traditional cryoprotectants in these cases.

2.2. Materials and Methods

2.2.1 Loop diffraction tests

Loop diffraction tests were performed to determine the proline concentration range needed to prevent the formation of crystalline ice in the presence of the two common precipitants ammonium sulfate and PEG 3350. For these tests, stock solutions of 6.5 M proline, 4 M ammonium sulfate and 50%(w/v) PEG 3350 in water were prepared. We note that the pH of 6.5 M aqueous proline is 6.5. Several solutions containing various concentrations of proline and either ammonium sulfate or PEG 3350 were then made by mixing appropriate volumes of the stock solutions and water. Hampton Research 20 μm

diameter nylon loops (0.5 mm loop size) were dipped into these solutions, flash-cooled in a cryogenic gaseous N₂ stream (Rigaku X-stream 2000 set at 123 K) and exposed to X-rays from a Rigaku RUH3R rotating-anode source coupled to an R-Axis IV⁺⁺ detector. The cryostream was blocked with a thin plastic card while the loop was mounted on the goniometer and the card was quickly removed after the sample was in place. The loops were not blotted prior to transfer to the goniometer. The crystal-to-detector distance was 150 mm, which corresponds to 2.00 Å resolution at the detector edge. The exposure time was 1 min. Each solution was tested three times to ensure reproducibility.

2.2.2 Crystallization and cryoprotection

2.2.2.1. General cryoprotection procedure

The crystals were cryoprotected using the *in situ* serial transfer method described previously (2). In our implementation of this method, cryoprotection is initiated by adding 20–40 µl of a buffer containing a low concentration of the cryoprotectant to the sitting drop in which the crystal was grown (Cryschem plate). The solution bathing the crystal is then mixed by drawing 20 µl liquid into the pipette and expelling it back into the drop 3–5 times. Next, 20 µl buffer is removed, 20 µl fresh buffer is added and the solution is mixed. The cycle of removal, replacement and mixing is repeated until schlieren lines are no longer evident upon the addition of fresh buffer (typically 3–5 cycles). The entire procedure is repeated using a series of buffers with increasing amounts of the cryoprotectant until the final desired level of cryoprotectant is achieved. The time for cryoprotection was 5–10 min per crystal. The cryoprotected crystals were picked up with Hampton Research 20 µm nylon loops and vitrified by plunging them into liquid nitrogen. The crystals were not blotted prior to plunging into liquid nitrogen, nor

was the cold gas layer above the liquid-nitrogen surface removed (22). The thickness of the cold gas layer was estimated to be 1–2 cm.

2.2.2.2. Lysozyme

Hen egg-white lysozyme (HEWL) was purchased from Sigma–Aldrich (catalogue No. L7651) and a stock solution of 12.5 mg ml⁻¹ HEWL in 0.1 M sodium acetate buffer pH 4.6 was prepared. Tetragonal crystals were grown in 6 µl sitting drops (Crychem plates) at room temperature using a reservoir solution consisting of 0.5–0.9 M NaCl, 0.1 M sodium acetate pH 4.6. The protein:reservoir ratio in the drop was varied, with the largest crystals obtained at ratios of 4:2 and 3:3. HEWL crystals were cryoprotected in 2.8 M proline, 1.2 M NaCl, 0.1 M sodium acetate pH 4.6 (Table 1) *via* the *in situ* serial transfer method described above. The starting buffer was 0.5 M proline, 1.2 M NaCl, 0.1 M sodium acetate pH 4.6. The proline concentration in the buffer was increased stepwise to 1.0, 1.5, 2.0, 2.5 and finally 2.8 M.

As a control, HEWL crystals were also cryoprotected in ethylene glycol, a cryoprotectant that has previously been used with tetragonal HEWL crystals (23, 24). The starting buffer for *in situ* serial transfer cryoprotection was 10% (v/v) ethylene glycol, 0.9 M NaCl, 0.1 M sodium acetate pH 4.6. The ethylene glycol concentration was increased stepwise to 15, 20 and finally 22.5%.

2.2.2.3. Xylose isomerase

Xylose isomerase (XI) from *Streptomyces rubiginosus* was purchased from Hampton Research (catalogue No. HR7-102) as a crystalline suspension at 33 mg ml⁻¹. The protein was dialyzed into 10 mM HEPES, 1 mM MgCl₂ pH 7.0 and was concentrated to 25 mg ml⁻¹ using a centrifugal device. Crystals were grown at room temperature in

Cryschem sitting-drop plates using drops formed by mixing 1.5 μl each of the protein and reservoir solutions.

The *I222* crystal form of XI was grown using a reservoir solution consisting of 1.0–2.2 *M* ammonium sulfate, 0.1 *M* Tris–HCl pH 7.0–8.5. The crystals were cryoprotected using the *in situ* stepwise approach described in §2.2.1. The starting buffer consisted of 1.0 *M* proline, 2.0 *M* ammonium sulfate, 0.1 *M* Tris–HCl pH 8.5. The proline concentration in the buffer was increased by 1.0 *M* in one step to achieve a final cryobuffer of 2.0 *M* proline, 2.0 *M* ammonium sulfate, 0.1 *M* Tris–HCl pH 8.5 (Table 1; XI-1).

For comparison purposes, XI crystals were also cryoprotected in glycerol. We note that glycerol has been used previously for *I222* crystals of XI grown in ammonium sulfate (PDB entry 2GLK). The starting buffer for *in situ* serial transfer cryoprotection was 5% (v/v) glycerol, 2.0 *M* ammonium sulfate, 0.2 *M* MgCl₂, 0.1 *M* Tris pH 8.0. The concentration of glycerol in the buffer was increased to 20% in steps of 5%.

The *I222* crystal form of XI was also obtained using low ionic strength conditions consisting of 4–9% (w/v) PEG 4000, 0.2 *M* MgCl₂, 0.1 *M* Tris–HCl pH 7.0–8.5. The crystals were cryoprotected *in situ* by first replacing the mother liquor with 20 μl 1.0 *M* proline, 5% (w/v) PEG 4000, 0.2 *M* MgCl₂, 0.1 *M* Tris pH 8.5. The proline concentration was increased in 0.5 *M* steps to a final cryobuffer of 3.0 *M* proline, 5% PEG 4000, 0.2 *M* MgCl₂, 0.1 *M* Tris pH 8.5 (Table 1; XI-2). As a control experiment, this crystal form of XI was also cryoprotected in PEG 200. The starting buffer for cryoprotection was 5% (v/v) PEG 200, 5% (w/v) PEG 4000, 0.2 *M* MgCl₂, 0.1 *M* Tris pH 8.0. The PEG 200 concentration in the buffer was increased to 20% in steps of 5%.

2.2.2.4. Histidine acid phosphatase

The H17N/D261A double mutant of the histidine acid phosphatase from *Francisella tularensis* (FtHAP) was expressed using auto-induction (25) and was purified and crystallized as described previously (26, 27). Crystals were grown in sitting drops using a reservoir solution consisting of 1.7–2.2 M ammonium sulfate, 0.1 M bis-Tris pH 5.5–6.5. The drop size was 2 μ l and equal volumes of the protein and reservoir solutions were mixed. The space group of the crystals was $P4_1$, with unit-cell parameters $a = b = 62.0$, $c = 210.4$ Å. The starting buffer for *in situ* cryoprotection was 1.0 M proline, 2.0 M ammonium sulfate, 0.1 M bis-Tris pH 6.5. The proline concentration was increased in 0.5 M steps to a final cryobuffer of 2.0 M proline, 2.0 M ammonium sulfate, 0.1 M bis-Tris pH 6.5 (Table 1).

For comparison purposes, FtHAP crystals were also cryoprotected in glycerol as described previously (27). The starting buffer for cryoprotection was 5% (v/v) glycerol, 2.0 M ammonium sulfate, 0.1 M bis-Tris pH 6.25. The glycerol concentration was increased using *in situ* serial transfer to 10, 15 and finally 20%.

2.2.2.5. 1-Pyrroline-5-carboxylate dehydrogenase

1-Pyrroline-5-carboxylate dehydrogenase from *Mus musculus* (MmP5CDH) was expressed, purified and crystallized as recently described (28). MmP5CDH was crystallized using a reservoir consisting of 15–25% PEG 3350, 0.1 M bis-Tris pH 5.5–6.5. The space group of the crystals was $P2_12_12_1$, with unit-cell parameters $a = 84.9$, $b = 94.0$, $c = 132.4$ Å. The reservoir solution served as the initial buffer for *in situ* stepwise cryoprotection. Cryoprotection was achieved in a single step with a final buffer of 2.4 M proline, 25% PEG 3350, 0.1 M bis-Tris pH 6.25 (Table 1).

2.2.3 Data collection and refinement

X-ray diffraction data were collected from the HEWL and XI crystals using a rotating-anode X-ray source and an R-AXIS IV⁺⁺ detector. For each data set, the crystal-to-detector distance was 85 mm, the exposure time was 2 min per frame and the oscillation width was 0.5°. The resolution of the inscribed circle of the detector at 85 mm is 1.52 Å.

Data from FtHAP crystals were collected on beamline 4.2.2 of the Advanced Light Source using a NOIR-1 detector. Each data set consisted of 546 images with a crystal-to-detector distance of 160 mm, a detector offset of 8°, an oscillation width of 0.33° and an exposure time of 2 s per frame. The resolution limits at the top and bottom edges of the detector were 1.72 and 3.21 Å, respectively, while the resolution at the side of the detector was 2.25 Å.

Data from an MmP5CDH crystal were collected on beamline 24-ID-E of the Advanced Photon Source using an ADSC Q315 detector. The data set was collected using a continuous vector scan and consisted of 120 images obtained using a crystal-to-detector distance of 125 mm, an oscillation width of 1.0° and an exposure time of 1 s per frame at 100% transmittance. The resolution of the inscribed circle of the detector was 1.10 Å.

All data sets were integrated with *XDS* (29) and scaled with *SCALA* (30) via the *CCP4i* interface (31). Data-collection statistics are listed in Tables 2, 3, 4 and 5. The mosaicity values in Tables 2–5 are from the CORRECT step of *XDS*.

The data sets were used in structure refinement in order to identify proline molecules bound to the protein. Refinement was performed with *PHENIX* (32) starting from coordinates derived from the following structures: HEWL, PDB entry 2LYZ (33); XI, PDB entry 1XIF (Carrell *et al.*, 1994); FtHAP, PDB entry 3IT2 (27); Mmp5CDH, PDB entry 3V9J (28). In each case solvent molecules were removed prior to refinement. The refinement protocol consisted of rigid-body refinement followed by simulated annealing. The *B*-factor model consisted of an isotropic *B* factor for each non-H atom plus one TLS group per protein chain for HEWL, XI and FtHAP. Anisotropic *B* factors were used for the refinement of Mmp5CDH. After the first round of refinement, *Coot* (34) was used to adjust the protein model and to add water molecules. The model was then input to *PHENIX* for a second round of refinement. The resulting electron-density maps were inspected, proline molecules were added and *PHENIX* refinement was performed. This procedure was repeated as necessary. Refinement statistics are listed in Tables 2–5. Coordinates and structure factors for structures containing ordered proline molecules have been deposited in the PDB under the accession codes listed in Tables 2–5.

2.3. Results

2.3.1 Loop diffraction tests

Loop diffraction tests were performed to determine the proline concentration range needed for cryoprotection. A solution of 4 *M* proline prevents crystalline ice rings; lower proline levels afford cryoprotection when other solutes are present. In particular, Fig. 1 shows diffraction images from loops containing 1.0–3.0 *M* proline together with one of two common precipitants: ammonium sulfate (Fig. 1*a*) or PEG 3350 (Fig. 1*b*).

The data for mixtures of proline and ammonium sulfate suggest that the sum of the two solute concentrations should be at least 4 *M* to prevent crystalline ice formation (Fig. 1*a*). For example, a combination of 2.0 *M* ammonium sulfate and 2.0 *M* proline prevents crystalline ice formation. The diffraction pattern from a solution of 1.0 *M* ammonium sulfate and 3.0 *M* proline is also free of crystalline ice rings, but rings are evident when the proline concentration is lowered to 2.5 *M*. Likewise, a solution of 1.5 *M* ammonium sulfate and 2.5 *M* proline is cryoprotective, but a combination of 1.5 *M* ammonium sulfate and 2.0 *M* proline does not prevent crystalline ice formation.

Solutions containing proline and PEG 3350 were also tested (Fig. 1*b*). Crystalline ice formation is suppressed by 15% (*w/v*) PEG 3350 and 2.5 *M* proline. If the PEG concentration is increased to 20% the proline concentration may be decreased to 2.0 *M*. Increasing the PEG concentration further to 25% allows the proline concentration to be lowered to 1.5 *M*. Thus, a rule of thumb is that 0.5 *M* proline is roughly equivalent to 5% PEG 3350 in terms of cryoprotective capacity.

2.3.2 Diffraction data

HEWL is a standard test case for protein-crystallography methods development; therefore, we tested proline cryoprotection of HEWL crystals. A diffraction data set was collected from a HEWL crystal that had been cryoprotected in 2.8 *M* proline. To facilitate the assessment of proline as a cryoprotectant, a comparison data set was collected from another crystal that had been cryoprotected in 22.5% ethylene glycol. Both crystals diffracted to beyond 1.5 Å resolution using a rotating-anode source and the data sets were truncated at 1.5 Å resolution owing to the limitation of the minimum detector distance on our system (Table 2). The mosaicity of the proline-soaked crystal was about 50% higher

than that of the ethylene glycol-soaked crystal. Otherwise, the two data sets were comparable in terms of global indicators of quality such as R_{merge} , R_{meas} , $R_{\text{p.i.m.}}$ and $\langle I/\sigma(I) \rangle$. It is difficult to know whether such small differences in data-quality statistics are a consequence of the cryoprotectant or simply reflect crystal-to-crystal variation. Nevertheless, these data certainly show that proline does not substantially degrade the diffraction characteristics of HEWL crystals, which suggests that proline is a potentially promising cryoprotectant. Thus, proline passes the lysozyme test.

XI is another test case that is routinely used for methods development. The *I222* form of XI was grown under two different conditions corresponding to high (XI-1) and low (XI-2) ionic strength. XI-1 crystals were cryoprotected in 2.0 M proline and a data set was collected using a rotating-anode system (Table 3; XI-1-pro). A comparison data set was collected from an XI-1 crystal that had been cryoprotected in 20% glycerol (Table 3; XI-1-gol). Similarly, data sets from XI-2 crystals cryoprotected in proline (Table 3; XI-2-pro) and PEG 200 (Table 3; XI-2-peg) were also obtained. All four crystals diffracted to beyond 1.50 Å resolution. The mosaicities of the XI-1 crystals are about 0.2°, whereas those of the XI-2 crystals are 0.1°. The data-quality statistics are very similar for the XI-1-pro, XI-2-pro and XI-2-peg data sets. Interestingly, those for the XI-1-gol data set are noticeably worse. In particular, the *R* factors of this data set are about two times higher than those of the other three data sets. Likewise, the $\langle I/\sigma(I) \rangle$ for XI-1-gol is substantially lower (Table 3). These differences are most evident in the high-resolution bin. As with the HEWL data, it is difficult to know whether these differences reflect crystal variation or the cryoprotectant. However, the data appear to suggest that proline

performs as well as PEG 200 for XI-2 crystals and possibly better than glycerol for XI-1 crystals.

In addition to the two above-mentioned test cases, we investigated two other proteins available in our laboratory: FtHAP and MmP5CDH. FtHAP was crystallized at high ionic strength with ammonium sulfate, whereas MmP5CDH crystals grew at lower ionic strength with PEG 3350 as the precipitant. Thus, these two cases represent quite different regions of crystallization space and provide an opportunity to assess the generality of proline as a cryoprotectant.

Data from a proline-soaked FtHAP crystal were collected on Advanced Light Source beamline 4.2.2 (Table 4). Another data set using identical data-collection parameters was collected from a crystal that was cryoprotected in glycerol (Table 4). Both data sets were processed to 1.75 Å resolution. Although it is difficult to account for the effects of crystal-to-crystal variation, the data seem to suggest that proline is superior to glycerol for cryoprotecting FtHAP crystals. For example, the mosaicity of the proline-soaked crystal (0.18°) is slightly lower than that of the glycerol-soaked crystal (0.24°). Furthermore, the *R* factors for the proline data set are about half those for the glycerol data set. Also, the $\langle I/\sigma(I) \rangle$ of the proline data is about twice that of the glycerol data. These results suggest that proline is a good cryoprotectant for FtHAP crystals.

Finally, a 1.24 Å resolution data set was collected on Advanced Photon Source beamline 24-ID-E from a crystal of MmP5CDH that was cryoprotected in 2.4 *M* proline (Table 5). For reference, we have recently reported structures of MmP5CDH complexed with sulfate ion, the product glutamate and the cofactor NAD⁺ which were determined from crystals that were cryoprotected in 25% glycerol (28). The high-resolution limits of

those structures ranged from 1.30 Å for the sulfate complex to 1.50 Å for the glutamate and NAD⁺ complexes. The mosaicity of the proline-soaked crystal is 0.17°, whereas the mosaicities of the glycerol-soaked crystals were 0.12–0.15°. Thus, soaking with proline did not degrade the resolution or substantially increase the mosaicity of Mmp5CDH crystals.

2.3.3 Ordered Proline Molecules

Electron-density maps were inspected to identify ordered proline molecules bound to the protein and this analysis clearly indicated that proline molecules were bound to all four enzymes (Tables 2–5). One proline molecule binds to HEWL in a crystal contact region formed by residues 19, 22 and 24 of one protein and Arg114 of a symmetry-related protein (Fig. 8). One proline site was also identified for XI. Electron density for proline was observed at this location in both crystal forms, but the density was much stronger in the ammonium sulfate form (XI-1-pro). Proline binds in a water-filled trough on the surface of XI and interacts with Ser281 (Fig. 9).

The active site of FtHAP contains one proline molecule, which is wedged between Phe23 and Tyr135 and forms no direct hydrogen bonds to the enzyme (Fig. 2). Interestingly, Phe23 and Tyr135 form an aromatic clamp that binds the adenine base of the substrate 3'-AMP (27). As shown in Fig. 2(b), the proline molecule occupies the substrate adenine site. A sulfate ion is also bound in the active site and occupies the substrate phosphoryl binding pocket. Thus, proline and sulfate together appear to mimic the substrate 3'-AMP.

Seven proline molecules are bound to Mmp5CDH (Fig. 3). Three prolines are bound to each of the two proteins in the asymmetric unit (labeled 1, 2 and 3 in Fig. 3) and an

additional proline binds in a crystal contact (labeled 4 in Fig. 3). Proline 1 binds in the active site and forms hydrogen bonds to Gly512 and Ser513 (Fig. 3, inset 1). The electron density suggests that the bound proline possibly exhibits conformational disorder or has less than full occupancy. Interestingly, this location corresponds to the binding site for the aldehyde substrate (and the glutamate product) (28). In fact, the backbone of the bound proline superimposes almost perfectly with that of the product glutamate (Fig. 3, inset 1). That proline binds in this site is consistent with the fact that proline is a competitive inhibitor of P5CDH (35). Proline 2 binds in the crevice between the NAD⁺-binding and catalytic lobes and forms electrostatic interactions with Arg399 and Asp393 (Fig. 3, inset 2). Also, the nonpolar part of the pyrrolidine ring of proline 2 contacts Phe210. Proline 3 contacts both protomers of the dimer (Fig. 3). Proline 3 forms electrostatic interactions with Lys292 of one chain and the 303–308 loop of the other chain (Fig. 3, inset 3). Also, the pyrrolidine ring of proline 3 makes nonpolar contacts with Leu302 and Phe308 of one chain of the dimer and Ala525 of the other chain. Finally, proline 4 binds in a crystal contact region. This proline interacts with the backbone of Gln184 of the protein in the asymmetric unit and is linked *via* two water molecules to Glu46 of a symmetry-related molecule (Fig. 3, inset 4).

2.4. Discussion

The results reported here suggest that proline is a suitable cryoprotectant for protein crystals. Our test cases included two ‘model’ proteins, HEWL and XI, as well as a phosphatase and an aldehyde dehydrogenase. None of these crystals exhibited any signs of deterioration, such as cracking or melting, while soaking with proline. Furthermore, the diffraction images obtained from these crystals were free of crystalline ice rings.

Moreover, the diffraction quality is similar to that obtained from crystals of these enzymes cryoprotected with conventional reagents such as ethylene glycol, glycerol and PEG 200.

The crystallization recipes of our test cases are representative of those used in protein crystallography. The precipitants include ammonium sulfate, NaCl and PEG 3350, which are commonly used in protein crystallization. The pH values of the crystallization solutions span a wide range: 4.6–8.5. Thus, proline is compatible with the kinds of solutions that are typically used in protein crystallization, suggesting that proline is widely applicable as a cryoprotectant for protein crystals.

Electron-density maps clearly indicated ordered proline molecules bound to the protein in four of the five crystal structures. This result is expected because penetrating cryoprotectants frequently bind to proteins. In fact, as of 29 April 2012 the PDB contained 6620 entries with glycerol as a ligand, 2720 entries with ethylene glycol as a ligand and 702 entries with 2-methyl-2,4-pentanediol as a ligand. Typically, these compounds are present at 10–30% (v/v) in the cryobuffer, which corresponds to 1–5 *M*. This concentration range is similar to that of proline used here (2.0–3.0 *M*). Furthermore, like glycerol, ethylene glycol and 2-methyl-2,4-pentanediol, proline has both hydrogen-bond donors and acceptors. The similarities of proline to traditional cryoprotection agents suggest that one should expect to find proline bound to the protein.

Finally, we note that we have not attempted to perform a systematic head-to-head comparison between proline and other cryoprotectants. Such a study is challenging because it is difficult to account for crystal-to-crystal variation of the diffraction quality.

Rather, our goal was to demonstrate that proline is an effective and generally applicable cryoprotectant for protein crystals. Our data support this assertion.

Acknowledgements

This research was supported by the NIH grant GM065546. We thank Dr. Jay Nix and Dr. Jonathan Schuermann for help with data collection and processing. Part of this research was performed at the Advanced Light Source. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

References

1. Hope, H. (2001) *International Tables for Crystallography*, Vol. F, Dordrecht: Kluwer Academic Publishers.
2. Garman, E., and Owen, R. L. (2007) Cryocrystallography of macromolecules: practice and optimization, *Methods Mol Biol* 364, 1-18.
3. David W, R. (1997) [13] Practical cryocrystallography, In *Methods in enzymology* (Charles W. Carter, Jr., Ed.), pp 183-203, Academic Press.

4. Pflugrath, J. W. (2004) Macromolecular cryocrystallography--methods for cooling and mounting protein crystals at cryogenic temperatures, *Methods* 34, 415-423.
5. Rubinson, K. A., Ladner, J. E., Tordova, M., and Gilliland, G. L. (2000) Cryosalts: suppression of ice formation in macromolecular crystallography, *Acta crystallographica. Section D, Biological crystallography* 56, 996-1001.
6. Holyoak, T., Fenn, T. D., Wilson, M. A., Moulin, A. G., Ringe, D., and Petsko, G. A. (2003) Malonate: a versatile cryoprotectant and stabilizing solution for salt-grown macromolecular crystals, *Acta crystallographica. Section D, Biological crystallography* 59, 2356-2358.
7. Mueller-Dieckmann, C., Kauffmann, B., and Weiss, M. S. (2011) Trimethylamine N-oxide as a versatile cryoprotective agent in macromolecular crystallography, *Journal of Applied Crystallography* 44, 433-436.
8. Marshall, H., Venkat, M., Seng, N. S., Cahn, J., and Juers, D. H. (2012) The use of trimethylamine N-oxide as a primary precipitating agent and related methylamine osmolytes as cryoprotective agents for macromolecular crystallography, *Acta crystallographica. Section D, Biological crystallography* 68, 69-81.
9. Hare, P. D., Cress, W. A., and van Staden, J. (1999) Proline synthesis and degradation: a model system for elucidating stress-related signal transduction, *Journal of Experimental Botany* 50, 413-434.
10. Yoshiba, Y., Kiyosue, T., Nakashima, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1997) Regulation of levels of proline as an osmolyte in plants under water stress, *Plant & cell physiology* 38, 1095-1102.
11. Szabados, L., and Savoure, A. (2010) Proline: a multifunctional amino acid, *Trends in plant science* 15, 89-97.
12. Barnett, N. M., and Naylor, A. W. (1966) Amino Acid and protein metabolism in bermuda grass during water stress, *Plant physiology* 41, 1222-1230.
13. Handa, S., Bressan, R. A., Handa, A. K., Carpita, N. C., and Hasegawa, P. M. (1983) Solutes contributing to osmotic adjustment in cultured plant cells adapted to water stress, *Plant physiology* 73, 834-843.
14. Takagi, H. (2008) Proline as a stress protectant in yeast: physiological functions, metabolic regulations, and biotechnological applications, *Applied microbiology and biotechnology* 81, 211-223.
15. Morita, Y., Nakamori, S., and Takagi, H. (2002) Effect of proline and arginine metabolism on freezing stress of *Saccharomyces cerevisiae*, *Journal of bioscience and bioengineering* 94, 390-394.

16. Takagi, H., Sakai, K., Morida, K., and Nakamori, S. (2000) Proline accumulation by mutation or disruption of the proline oxidase gene improves resistance to freezing and desiccation stresses in *Saccharomyces cerevisiae*, *FEMS microbiology letters* 184, 103-108.
17. Kostal, V., Zahradnickova, H., and Simek, P. (2011) Hyperprolinemic larvae of the drosophilid fly, *Chymomyza costata*, survive cryopreservation in liquid nitrogen, *Proceedings of the National Academy of Sciences of the United States of America* 108, 13041-13046.
18. Kostal, V., Simek, P., Zahradnickova, H., Cimlova, J., and Stetina, T. (2012) Conversion of the chill susceptible fruit fly larva (*Drosophila melanogaster*) to a freeze tolerant organism, *Proceedings of the National Academy of Sciences of the United States of America* 109, 3270-3274.
19. Withers, L. A., and King, P. J. (1979) Proline: A Novel Cryoprotectant for the Freeze Preservation of Cultured Cells of *Zea mays* L, *Plant physiology* 64, 675-678.
20. Sanchez-Partida, L. G., Setchell, B. P., and Maxwell, W. M. (1998) Effect of compatible solutes and diluent composition on the post-thaw motility of ram sperm, *Reproduction, fertility, and development* 10, 347-357.
21. Freimark, D., Sehl, C., Weber, C., Hudel, K., Czermak, P., Hofmann, N., Spindler, R., and Glasmacher, B. (2011) Systematic parameter optimization of a Me(2)SO- and serum-free cryopreservation protocol for human mesenchymal stem cells, *Cryobiology* 63, 67-75.
22. Warkentin, M., Berejnov, V., Hussein, N. S., and Thorne, R. E. (2006) Hyperquenching for protein cryocrystallography, *J Appl Crystallogr* 39, 805-811.
23. Evans, G., and Bricogne, G. (2002) Triiodide derivatization and combinatorial counter-ion replacement: two methods for enhancing phasing signal using laboratory Cu K[alpha] X-ray equipment, *Acta Crystallographica Section D* 58, 976-991.
24. Retailleau, P., and Prange, T. (2003) Phasing power at the K absorption edge of organic arsenic, *Acta crystallographica. Section D, Biological crystallography* 59, 887-896.
25. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein expression and purification* 41, 207-234.
26. Felts, R. L., Reilly, T. J., Calcutt, M. J., and Tanner, J. J. (2006) Crystallization of a newly discovered histidine acid phosphatase from *Francisella tularensis*, *Acta crystallographica. Section F, Structural biology and crystallization communications* 62, 32-35.

27. Singh, H., Felts, R. L., Schuermann, J. P., Reilly, T. J., and Tanner, J. J. (2009) Crystal Structures of the histidine acid phosphatase from *Francisella tularensis* provide insight into substrate recognition, *Journal of molecular biology* 394, 893-904.
28. Srivastava, D., Singh, R. K., Moxley, M. A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2012) The three-dimensional structural basis of type II hyperprolinemia, *Journal of molecular biology* 420, 176-189.
29. Kabsch, W. (2010) Xds, *Acta crystallographica. Section D, Biological crystallography* 66, 125-132.
30. Evans, P. (2006) Scaling and assessment of data quality, *Acta crystallographica. Section D, Biological crystallography* 62, 72-82.
31. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite, *Acta crystallographica. Section D, Biological crystallography* 59, 1131-1137.
32. Adams, P. D., Afonine, P. V., Bunkoczi, 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 crystallographica. Section D, Biological crystallography* 66, 213-221.
33. Diamond, R. (1974) Real-space refinement of the structure of hen egg-white lysozyme, *Journal of molecular biology* 82, 371-391.
34. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallographica Section D* 60, 2126-2132.
35. Forte-McRobbie, C., and Pietruszko, R. (1989) Human glutamic-gamma-semialdehyde dehydrogenase. Kinetic mechanism, *The Biochemical journal* 261, 935-943.

Table 2.1. Proline-based cryobuffers

	Lysozyme	GI-1	GI-2	FtHAP	MmP5CDH
Buffer	0.1 M Na acetate	0.1 M Tris-HCl	0.1 M Tris- HCl	0.1 M Bis-Tris	0.1 M Bis-Tris
pH	4.6	8.5	8.5	6.5	6.25
Precipitant	1.2 M NaCl	2.0 M (NH ₄) ₂ SO ₄	5 % PEG 4000	2.0 M (NH ₄) ₂ SO ₄	25 % PEG 3350
Salt	N/A	N/A	N/A	N/A	0.2 M Li ₂ SO ₄
Proline	2.8 M	2.0 M	3.0 M	2.0 M	2.4 M

Table 2.2. Data collection and refinement statistics for HEWL.

Values in parentheses are for the outer resolution shell.

Cryoprotectant	2.8 M proline	22.5% ethylene glycol
Space group	$P4_32_12$	$P4_32_12$
Unit-cell parameters (Å)	$a = b = 78.0, c = 37.7$	$a = b = 78.9, c = 36.9$
Wavelength (Å)	1.54	1.54
Resolution (Å)	19.50-1.50 (1.58-1.50)	19.50-1.50 (1.58-1.50)
No. of observations	205088	228617
Unique reflections	19076	19082
$R_{\text{merge}}(I)$	0.028 (0.241)	0.028 (0.148)
$R_{\text{meas}}(I)$	0.029 (0.257)	0.029 (0.157)
$R_{\text{p.i.m.}}(I)$	0.008 (0.085)	0.008 (0.051)
Mean $I/\sigma(I)$	46.1 (7.6)	58.6 (12.9)
Completeness (%)	99.6 (97.6)	99.9 (99.5)
Multiplicity	10.8 (8.2)	12.0 (8.8)
Mosaicity (°)	0.21	0.13
No. of protein atoms	976	
No. of water molecules	99	
No. of proline molecules	1	
R_{cryst}	0.185	
$R_{\text{free}}^{\#}$	0.216	
<i>R.m.s. deviations</i> [†]		
Bond lengths (Å)	0.005	
Bond angles (°)	1.012	
<i>Ramachandran plot</i> [‡] , residues in		
Favored regions	126	
Allowed regions	1	
Outliers	0	
<i>Average B factor</i> (Å ²)		
Protein	16	
Water	25	
Proline	22	
PDB code	4E3U	

[#]5 % test set.

[†]Compared to the parameters of Engh and Huber (Engh & Huber, 1991).

[‡]The Ramachandran plot was generated with RAMPAGE (Lovell *et al.*, 2003).

Table 2.3. Data collection and refinement statistics for XI.

Values in parentheses are for the outer resolution shell.

	XI-1-pro	XI-1-gol	XI-2-pro	XI-2-peg
Cryoprotectant	2.0 M proline	20% glycerol	3.0 M proline	20% PEG 200
Space group	<i>I</i> 222	<i>I</i> 222	<i>I</i> 222	<i>I</i> 222
Unit-cell parameters (Å)	<i>a</i> = 92.5, <i>b</i> = 98.0, <i>c</i> = 102.4	<i>a</i> = 92.5, <i>b</i> = 98.2, <i>c</i> = 101.8	<i>a</i> = 92.8, <i>b</i> = 98.1, <i>c</i> = 102.7	<i>a</i> = 93.0, <i>b</i> = 97.8, <i>c</i> = 102.9
Wavelength (Å)	1.54	1.54	1.54	1.54
Resolution (Å)	19.5-1.50 (1.58-1.50)	19.5-1.50 (1.58-1.50)	19.7-1.50 (1.58-1.50)	19.7-1.50 (1.58-1.50)
No. of observations	470488	490567	463654	487593
Unique reflections	72416	72675	73511	73858
$R_{\text{merge}}(I)$	0.037 (0.183)	0.071 (0.552)	0.037 (0.174)	0.046 (0.160)
$R_{\text{meas}}(I)$	0.040 (0.203)	0.077 (0.608)	0.040 (0.194)	0.500 (0.178)
$R_{\text{p.i.m.}}(I)$	0.015 (0.085)	0.029 (0.249)	0.016 (0.082)	0.019 (0.076)
Mean $I/\sigma(I)$	30.1 (8.5)	20.6 (3.1)	31.3 (8.7)	27.5 (9.2)
Completeness (%)	97.8 (89.5)	98.5 (93.1)	98.8 (92.2)	99.0 (94.2)
Multiplicity	6.5 (5.2)	6.8 (5.4)	6.3 (4.9)	6.6 (4.8)
Mosaicity (°)	0.19	0.25	0.13	0.12
No. of protein atoms	3052		3086	
No. of water molecules	324		292	
No. of proline molecules	1		0	
R_{cryst}	0.164		0.172	
$R_{\text{free}}^{\#}$	0.179		0.187	
<i>R.m.s. deviations</i> [†]				
Bond lengths (Å)	0.006		0.006	
Bond angles (°)	1.107		1.066	
<i>Ramachandran plot</i> [‡] , residues in				
Favored regions	372		373	
Allowed regions	10		12	
Outliers	1		1	
<i>Average B factor</i> (Å ²)				
Protein	11		10	
Water	21		17	
Proline	20		-	
PDB code	4E3V			

[#]5 % test set.[†]Compared to the parameters of Engh and Huber (Engh & Huber, 1991).[‡]The Ramachandran plot was generated with RAMPAGE (Lovell *et al.*, 2003).

Table 2.4. Data collection and refinement statistics for FtHAP.

Values in parentheses are for the outer resolution shell.

Cryoprotectant	2.0 M proline	20% glycerol
Space group	$P4_1$	$P4_1$
Unit-cell parameters (Å)	$a = b = 62.0, c = 210.4$	$a = b = 62.0, c = 211.0$
Wavelength (Å)	1.00	1.00
Resolution (Å)	48.6-1.75 (1.84-1.75)	46.5-1.75 (1.84-1.75)
No. of observations	341312	308938
Unique reflections	78960	76942
$R_{\text{merge}}(I)$	0.032 (0.130)	0.065 (0.153)
$R_{\text{meas}}(I)$	0.036 (0.160)	0.072 (0.199)
$R_{\text{p.i.m.}}(I)$	0.015 (0.092)	0.030 (0.126)
Mean $I/\sigma(I)$	25.8 (5.7)	13.1 (3.1)
Completeness (%)	99.1 (95.3)	96.6 (81.1)
Multiplicity	4.3 (2.8)	4.0 (1.9)
Mosaicity (°)	0.18	0.24
No. of protein atoms	5035	
No. of water molecules	365	
No. of proline molecules	2	
R_{cryst}	0.182	
$R_{\text{free}}^{\#}$	0.203	
<i>R.m.s. deviations</i> [†]		
Bond lengths (Å)	0.006	
Bond angles (°)	1.007	
<i>Ramachandran plot</i> [‡] , residues in		
Favored regions	648	
Allowed regions	6	
Outliers	0	
<i>Average B factor</i> (Å ²)		
Protein	19	
Water	23	
Proline	24	
PDB code	4E3W	

[#]5 % test set.

[†]Compared to the parameters of Engh and Huber (Engh & Huber, 1991).

[‡]The Ramachandran plot was generated with RAMPAGE (Lovell *et al.*, 2003).

Table 2.5. Data collection and refinement statistics for MmP5CDH

Values in parentheses are for the outer resolution shell.

Cryoprotectant	2.4 M proline
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 85.0, b = 93.9, c = 132.4$
Wavelength (Å)	0.979
Resolution (Å)	45.6-1.24 (1.31-1.24)
No. of observations	1144056
Unique reflections	279378
$R_{\text{merge}}(I)$	0.051 (0.550)
$R_{\text{meas}}(I)$	0.058 (0.644)
$R_{\text{p.i.m.}}(I)$	0.027 (0.326)
Mean $I/\sigma(I)$	15.0 (2.3)
Completeness (%)	94.1 (87.3)
Multiplicity	4.1 (3.3)
Mosaicity (°)	0.17
No. of protein atoms	8304
No. of water molecules	882
No. of proline molecules	7
R_{cryst}	0.155
$R_{\text{free}}^{\#}$	0.180
<i>R.m.s. deviations</i> [†]	
Bond lengths (Å)	0.005
Bond angles (°)	1.078
<i>Ramachandran plot</i> [‡] , residues in	
Favored regions	1066
Allowed regions	19
Outliers	0
<i>Average B factor</i> (Å ²)	
Protein	12
Water	22
Proline	18
PDB code	4E3X

[#]5 % test set.

[†]Compared to the parameters of Engh and Huber (Engh & Huber, 1991).

[‡]The Ramachandran plot was generated with RAMPAGE (Lovell *et al.*, 2003).

Figure legends

Figure 2.1

X-ray diffraction images from loops containing proline and either (a) ammonium sulfate or (b) PEG 3350.

Figure 2.2

Electron density for proline bound to HEWL. The cage represents a simulated annealing σ_A -weighted $F_o - F_c$ omit map contoured at 3.0σ . The protein in the asymmetric unit is colored gray. The protein related by the crystallographic symmetry operation ($y-1/2, -x+1/2, z+1/4$) is colored cyan.

Figure 2.3

Electron density for proline bound to GI-1. The cage represents a simulated annealing σ_A -weighted $F_o - F_c$ omit map contoured at 3.0σ .

Figure 2.4

Proline bound to the active site of FtHAP. (a) Ribbon drawing of the dimer with the bound proline molecules drawn as spheres. (b) Electron density for proline bound to the active site of HAP. The cage represents a simulated annealing σ_A -weighted $F_o - F_c$ omit map contoured at 3.0σ . The 3'-AMP bound to FtHAP mutant D261A (PDB code 3IT3) is shown in thin cyan sticks.

Figure 2.5

Proline molecules bound to Mm5CDH. The two protomers of the dimer are colored red and blue. The insets show close-up views of the bound proline ligands. The cage represents a simulated annealing σ_A -weighted $F_o - F_c$ omit map contoured at 3.0σ . In inset 4, the protein in the asymmetric unit is colored gray, and the protein related by $(x - 1/2, -y + 1/2, -z)$ is colored cyan.

Fig 2.1.

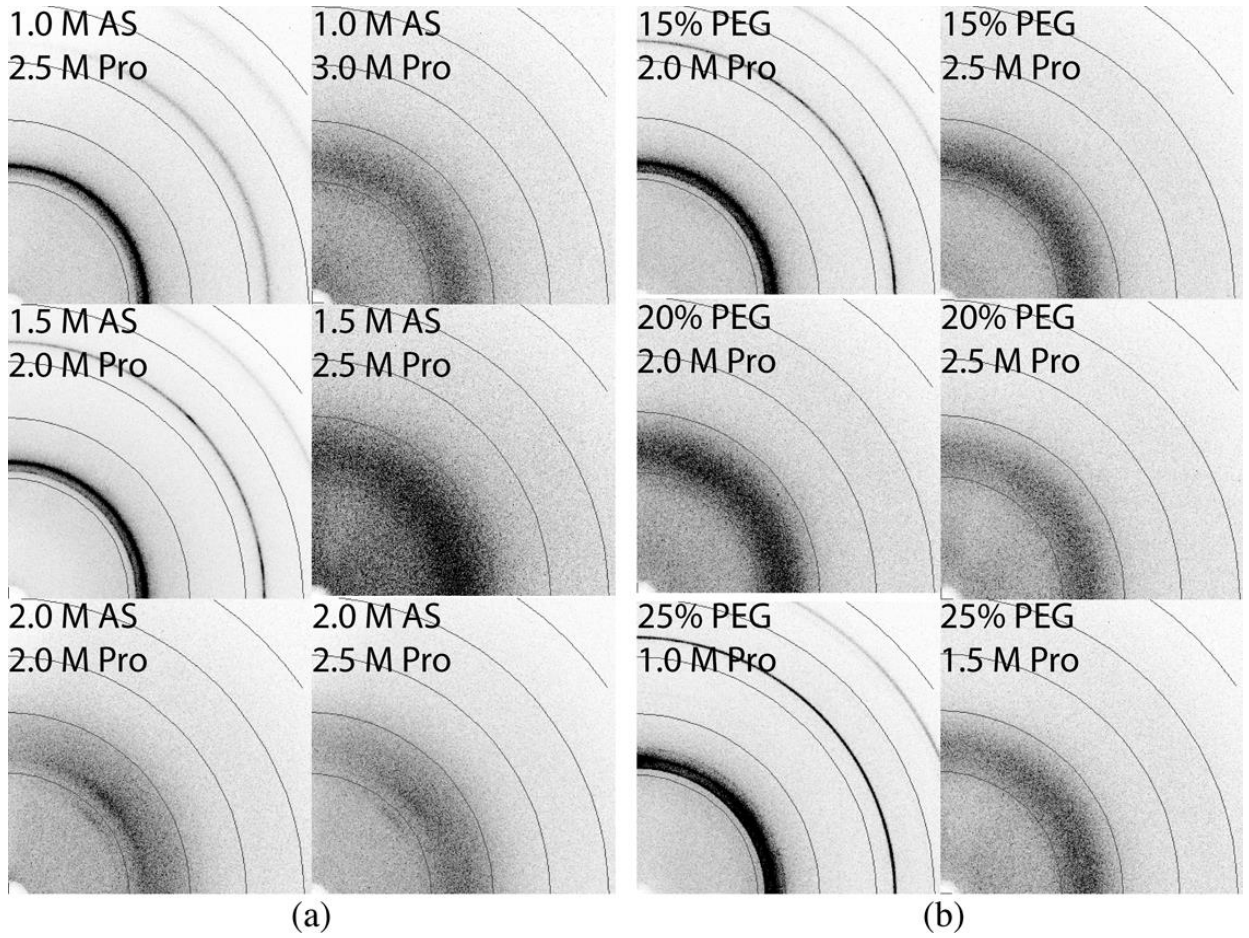


Fig 2.2.

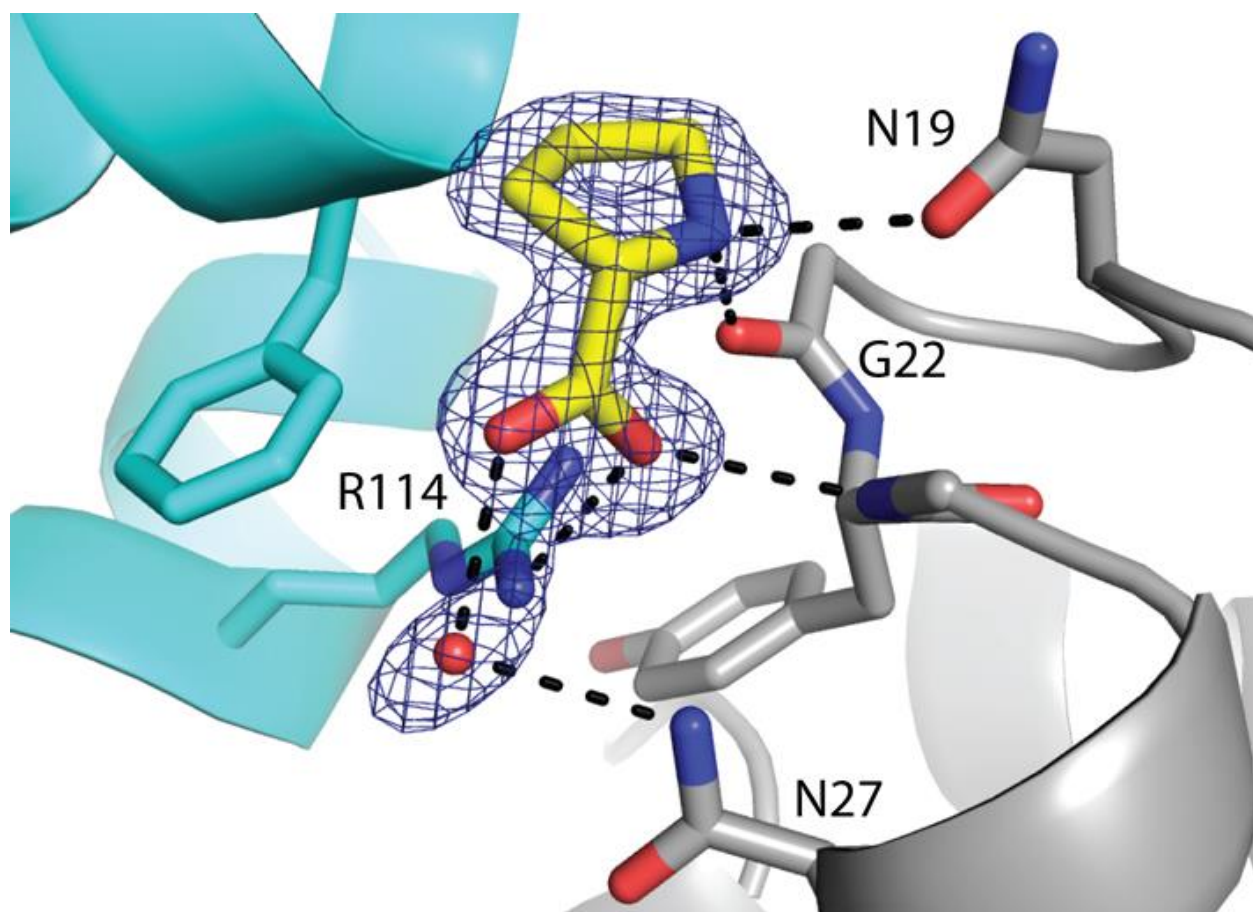


Fig 2.3.

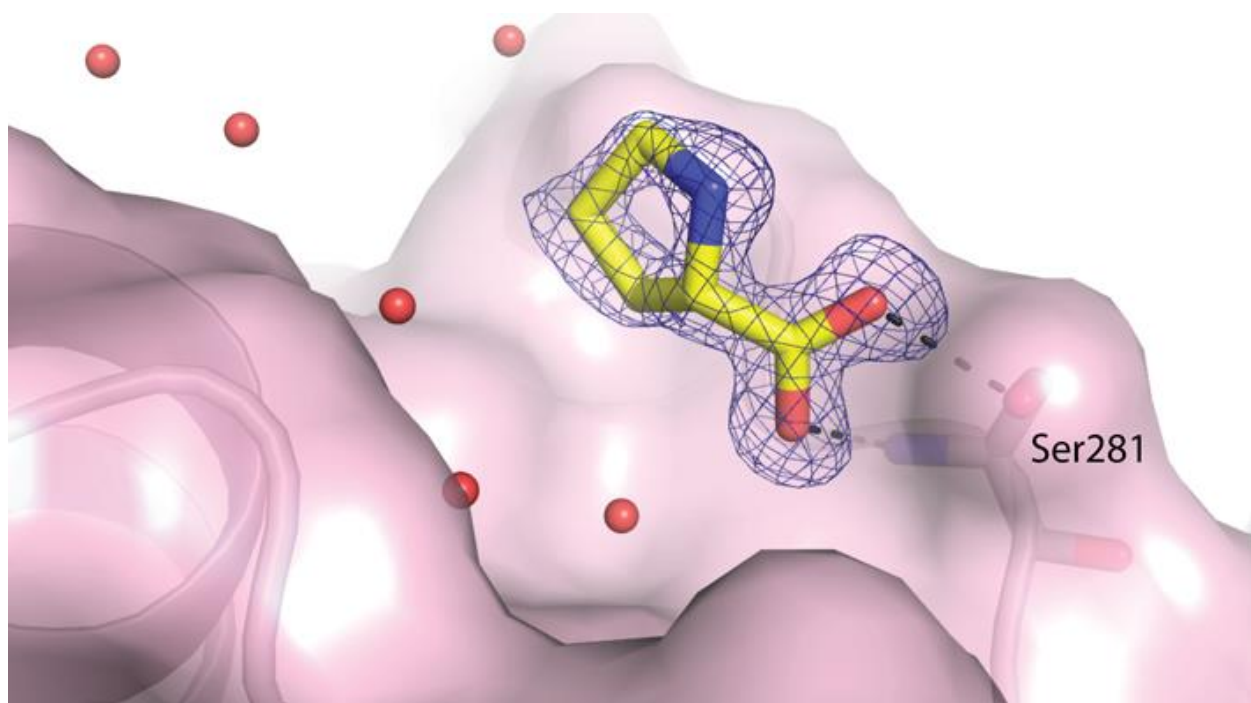


Fig 2.4.

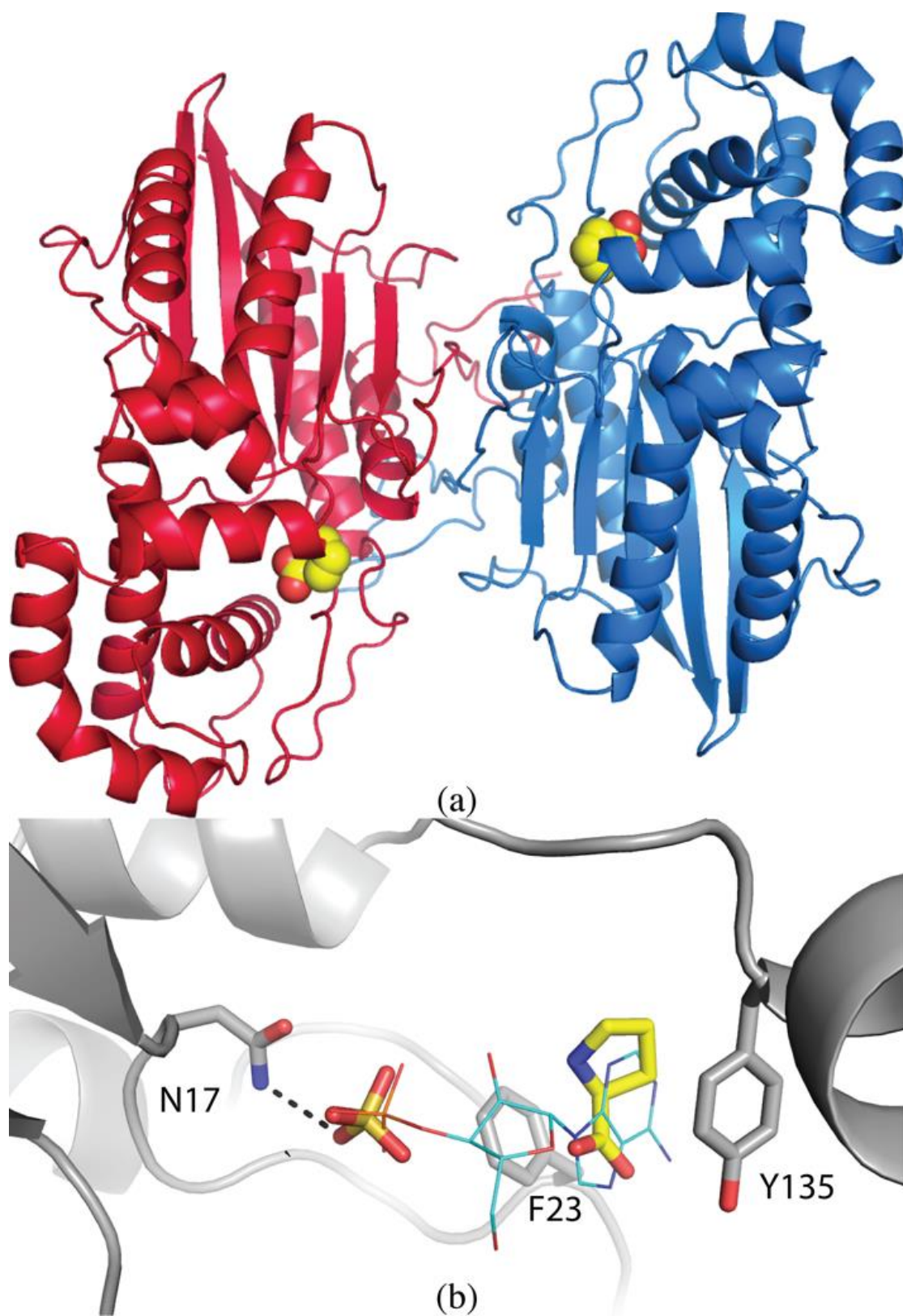
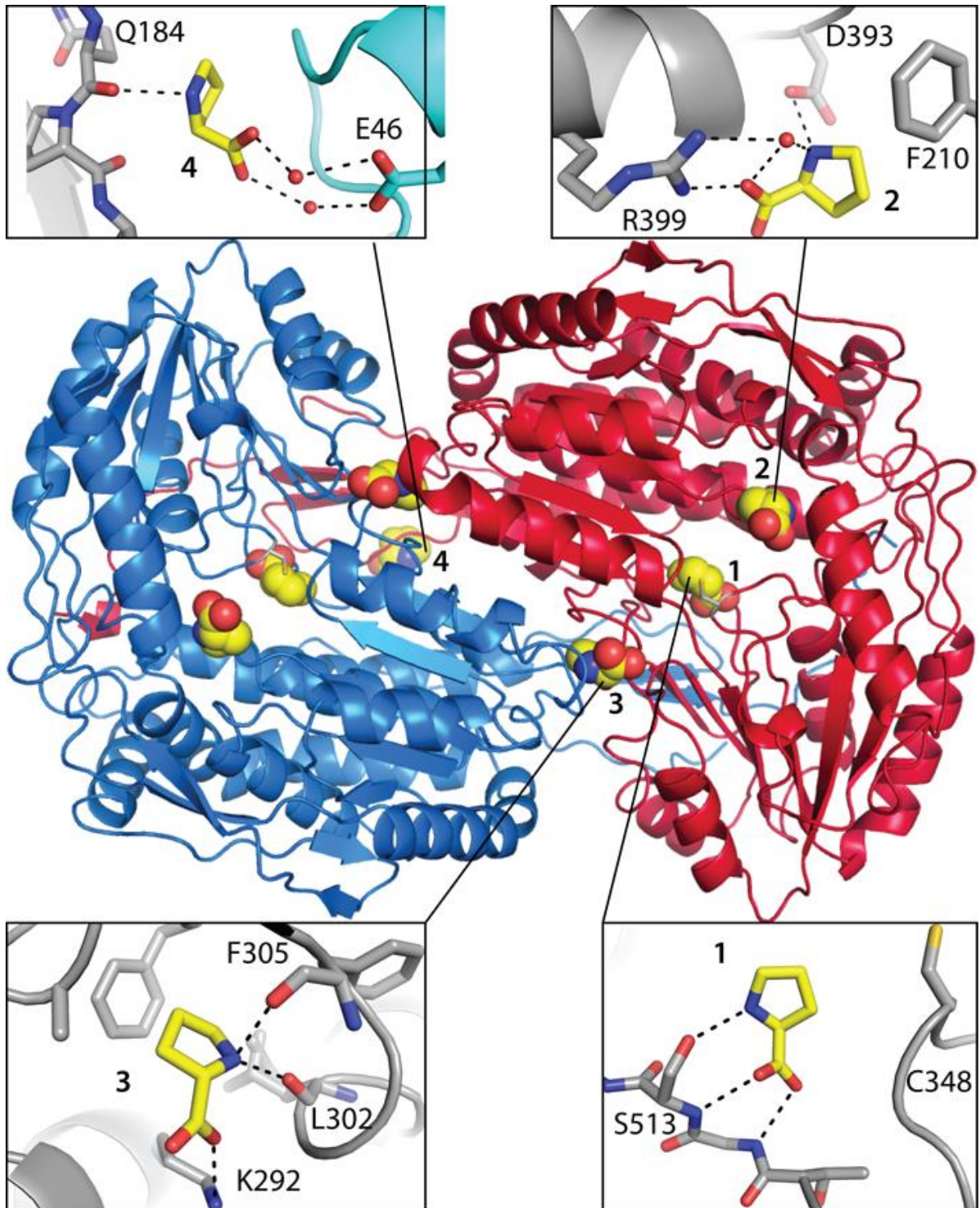


Fig 2.5.



Chapter 3:

Structural Basis of Substrate Selectivity of Δ^1 -Pyrroline-5-Carboxylate Dehydrogenase (ALDH4A1): Semialdehyde Chain Length

Abstract

The enzyme Δ^1 -pyrroline-5-carboxylate (P5C) dehydrogenase (aka P5CDH and ALDH4A1) is an aldehyde dehydrogenase that catalyzes the oxidation of γ -glutamate semialdehyde to L-glutamate. The crystal structures of mouse P5CDH complexed with glutarate, succinate, malonate, glyoxylate, and acetate are reported. The structures are used to build a structure-activity relationship that describes the semialdehyde carbon chain length and the position of the aldehyde group in relation to the cysteine nucleophile and oxyanion hole. Efficient 4- and 5-carbon substrates share the common feature of being long enough to span the distance between the anchor loop at the bottom of the active site and the oxyanion hole at the top of the active site. The inactive 2- and 3-carbon semialdehydes bind the anchor loop but are too short to reach the oxyanion hole. Inhibition of P5CDH by glyoxylate, malonate, succinate, glutarate, and L-glutamate is also examined. The K_i values are 0.27 mM for glyoxylate, 58 mM for succinate, 30 mM for glutarate, and 12 mM for L-glutamate. Curiously, malonate is not an inhibitor. The trends in K_i likely reflect a trade-off between the penalty for desolvating the carboxylates of the free inhibitor and the number of compensating hydrogen bonds formed in the enzyme-inhibitor complex.

3.1. Introduction

Δ^1 -pyrroline-5-carboxylate (P5C) dehydrogenase (P5CDH) is an NAD⁺-dependent aldehyde dehydrogenase that catalyzes the last step of proline catabolism, the oxidation of γ -glutamate semialdehyde (GSA) to L-glutamate (Fig. 1, upper reactions) (1). The aldehyde substrate for P5CDH in this pathway is generated by the hydrolysis of P5C, which is produced from the oxidation of proline by the flavoenzyme proline dehydrogenase (PRODH), known as proline oxidase in humans. Mammalian P5CDHs also function in hydroxyproline catabolism (2). This pathway begins with the oxidation of hydroxyproline by hydroxyproline oxidase and ends with the P5CDH-catalyzed oxidation of 4-hydroxyglutamate semialdehyde to 4-erythro-hydroxy-L-glutamate (Fig. 1, lower reactions) (3). In eukaryotes, P5CDH is a nuclear-encoded mitochondrial matrix enzyme, whereas in some bacteria P5CDH is combined with PRODH into a single bifunctional enzyme known as proline utilization A (1).

P5CDHs are of considerable biomedical importance. Mutations in the P5CDH gene (*ALDH4A1*) that abrogate enzyme activity cause the metabolic disorder type II hyperprolinemia (4-6). Both proline catabolic genes are activated by the tumor suppressor p53 (7, 8), indicating a role for the pathway in apoptosis, stress, and cancer (9-11). In this context, the coordination of P5CDH, proline oxidase, and the proline biosynthetic enzyme P5C reductase is required for maintaining the proper balance of mitochondrial reactive oxygen species (12). In the fungal pathogen *Cryptococcus neoformans*, P5CDH is required for optimal production of the major virulence factors (13). And bacterial P5CDHs are being considered as components of vaccines against *Staphylococcus aureus* (14).

P5CDH belongs to the aldehyde dehydrogenase (ALDH) superfamily and is known as ALDH4A1. The ALDH superfamily comprises hundreds of distinct genes from all three domains of life, including 19 human ALDHs (15). ALDHs share a common protein fold (Fig. 2) and catalytic mechanism for the oxidation of aldehydes to carboxylates. The mechanism (16-18) begins with the binding of the substrate aldehyde group in the oxyanion hole, which positions the C atom of the aldehyde for nucleophilic attack by the essential Cys residue. Nucleophilic attack results in formation of a hemithioacetal intermediate. Hydride transfer to NAD(P)^+ generates NAD(P)H and the thioacylenzyme intermediate. Finally, hydrolysis of the thioacylenzyme yields the carboxylate product and regenerates the Cys nucleophile.

P5CDH was first characterized in the late 1980s (19-22). In this early work performed on placental and liver P5CDHs, several potential substrates were tested in steady-state kinetic assays in order to help establish P5C/GSA as the physiological substrate (19, 21, 22). The human enzyme exhibits good activity with glutarate semialdehyde and adipate semialdehyde. The catalytic efficiencies (V/K_m) for these substrates at pH 7.0 are about 50 % of that of the physiological substrate, P5C/GSA. The shorter semialdehyde, succinate semialdehyde, was also found to be a substrate, but the efficiency was somewhat lower at 16 % of the value for P5C/GSA. Activity was undetectable with malonate semialdehyde, glyoxylate, and aspartate semialdehyde. These studies clearly suggested that semialdehyde chain length is an important factor in substrate selectivity, however; the structural basis for this aspect of substrate recognition has not been explored.

To better understand the basis of substrate selectivity by P5CDH, we have determined

the high resolution crystal structures of *Mus musculus* P5CDH (MmP5CDH, 92% identical to human P5CDH) complexed with carboxylate ligands having chain lengths ranging from 5 carbons to 2 carbons: glutarate (**1** in Fig. 3), succinate (**2**), malonate (**3**), glyoxylate (**4**), and acetate (**5**). The structures explain the relationship between semialdehyde chain length and enzyme activity.

3.2. Experimental procedures

3.2.1 Crystallization and soaking

MmP5CDH was expressed, purified, and crystallized as described previously (23). Briefly, the stock enzyme solution used for crystallization contained 10 mg/mL MmP5CDH in a buffer of 50mM Tris, 50mM NaCl, 0.5mM EDTA, 0.5mM THP, and 5% glycerol at pH 7.5. This His-tag was not removed. Crystallization experiments were performed in sitting drops at room temperature with drops formed by mixing 1 μ L of the enzyme solution and 1 μ L of reservoir solution. The latter consisted of 15 - 25 % (w/v) PEG 3350, 0.2 M Li_2SO_4 , and 0.1 M Bis-Tris at pH 5.0 - 6.5. The space group is $P2_12_12_1$ with unit cell lengths of $a = 85 \text{ \AA}$, $b = 94 \text{ \AA}$, and $c = 132 \text{ \AA}$. The asymmetric unit contains two protein molecules, which form a dimer.

Crystals of Mm5CDH complexed with the anions in Fig. 3 were obtained by soaking the aforementioned crystals in a solution containing cryobuffer (24% PEG 3350, 0.1 M Bis-Tris at pH 6.5, 15 - 20 % PEG 200) supplemented with high concentration of the ligand (440 mM acetate, 220 mM glyoxylate, 220 mM malonate, 270 mM succinate, or 270 mM glutarate). We note that the high concentration was necessary to displace the adventitious sulfate ion that binds to the active site. The crystals were soaked for about 15 minutes, harvested with Hampton loops, and flash-cooled in liquid nitrogen.

3.2.2 X-ray data collection and refinement

X-ray diffraction data were collected using a Rigaku rotating anode source with an R-AXIS IV++ detector. Each data set consisted of 240 – 260 images collected with an exposure time of 5 minutes per frame, oscillation width of 0.5°, and detector distance of 110 mm or 130 mm. The data were integrated in XDS (24) and scaled in SCALA (25). Data processing statistics are listed in Table 1.

Refinement with PHENIX (26) was initiated from a 1.3 Å resolution structure of MmP5CDH (PDB code 4V9J). A common test set of reflections was used for cross validation, which was based on the one used previously for refinements of MmP5CDH (23, 27). COOT was used for model building (28). The PHENIX elbow utility (29) was used to create ligand restraint files from ideal coordinates downloaded from PDB Ligand Expo (30). Refinement statistics are listed in Table 1.

3.2.3 Steady-state inhibition kinetics

P5CDH activity was measured at 20 °C by monitoring the production of NADH at 340 nm as described previously for HsP5CDH (23). The assay buffer contained 0.1 M sodium phosphate and 1 mM EDTA at pH 7.0. The enzyme concentration was 9.6 µg/ml (0.15 µM), except for the measurements performed in the presence of malonate, which were done with 3.2 µg/ml enzyme (0.05 µM). Inhibition of MmP5CDH by glyoxylate, malonate, succinate, glutarate, and L-glutamate was studied using succinate semialdehyde as the variable substrate at fixed NAD⁺ concentration of 1.0 mM. (We note that P5C/GSA is not commercially available.) The substrate range was 10 – 400 µM when glyoxylate, succinate or glutarate was the inhibitor, 10 – 300 µM with malonate as

the inhibitor, and 10 - 350 μ M with L-glutamate as the inhibitor. In each case, 10 – 12 substrate concentrations were used (Supplemental Figure S1). The inhibitors were present at the following concentrations: glyoxylate, 1 mM and 4 mM; malonate, 50 mM and 100 mM; succinate, 100 mM and 200 mM; glutarate, 100 mM and 200 mM; L-glutamate, 50 mM and 100 mM.

The kinetic data were analyzed by global fitting to the competitive (eq. 1), noncompetitive (eq. 2), and mixed (eq. 3) inhibition models using Origin 9 software. The fits to the competitive model had R^2 in the range 0.981 – 0.991 (Supplemental Figure S1), whereas the fits to the noncompetitive model were consistently worse ($R^2 < 0.956$). The fits to the mixed model ($R^2 = 0.980 - 0.992$) were indistinguishable from those of the competitive model, and since the competitive model has one fewer parameter than the mixed model, the competitive model was selected for the final determination of kinetic parameters.

$$V = V_{\max}[S]/\{ K_m(1 + [I]/K_i) + [S] \} \quad (1)$$

$$V = V_{\max}[S]/\{ (K_m + [S])(1 + [I]/K_i) \} \quad (2)$$

$$V = V_{\max}[S]/\{ K_m(1 + [I]/K_i) + [S](1 + [I]/\alpha K_i) \} \quad (3)$$

3.3. Results

3.3.1 5- and 4-carbon carboxylate ligands: glutarate and succinate

The location of the aldehyde substrate-binding site has been deduced from crystal structures of P5CDHs complexed with the product L-glutamate (23, 31). L-glutamate binds in the cleft between the catalytic and NAD⁺-binding domains, as demonstrated by

glutarate bound to MmP5CDH (Fig. 2). The γ -carboxylate of L-glutamate represents the aldehyde group of GSA and binds in the oxyanion hole near the catalytic Cys residue, while the backbone of L-glutamate is anchored to a loop that connects the final strand of the catalytic domain to the NAD⁺-binding domain. This loop is the second of two crossover peptides that connect the major domains of the enzyme and will be referred to as the anchor loop.

The canonical aldehyde-binding mode is exhibited by glutarate, the longest ligand used in the present study (Fig. 4A). One of the carboxylate groups of glutarate binds to the anchor loop by forming hydrogen bonds with Gly512 and Ser513. This carboxylate also forms a hydrogen bond with Ser349 of the catalytic loop. These interactions are identical to those formed by the α -carboxylate of L-glutamate (Fig. 4C), and by inference, GSA. The three methylene carbons of glutarate are sandwiched between Phe512 and Phe520 (Fig. 2, inset), which is also reminiscent of the L-glutamate complex. These residues are part of a more general aromatic box that is common to ALDHs (32). The upper carboxylate of glutarate binds near catalytic Cys348 and thus represents the aldehyde group of glutarate semialdehyde (Fig. 4A). One of the O atoms of the upper carboxylate occupies the oxyanion hole and accepts hydrogen bonds from the side chain of Asn211 and the backbone of Cys348. The other O atom of the upper carboxylate extends toward a solvent cavity and represents the atom derived from nucleophilic attack of water on the thioacyl intermediate. The interactions of the upper carboxylate of glutarate are nearly identical to those of the γ -carboxylate of L-glutamate (Fig. 4C).

The position of the upper carboxylate is consistent with efficient nucleophilic attack by Cys348. In the glutarate complex, Cys348 has $\chi_1 = -65^\circ$, which is the most favored

rotamer for Cys but not the one that is populated during nucleophilic attack ($\chi_1 = 55^\circ$). We note that the $\chi_1 = -65^\circ$ rotamer is typically observed in P5CDH structures that lack NAD^+ , as is the case here. Rotation of Cys348 to $\chi_1 = 55^\circ$ brings the nucleophile into attack position and places the S atom 2.3 Å from the carboxylate C atom (Fig. 4C). For reference, this distance is also 2.3 Å in the MmP5CDH-L-glutamate complex. The structural similarity between the glutarate and L-glutamate poses is consistent with glutarate semialdehyde being a good substrate for human P5CDH (19, 21, 22).

The conformation of succinate is close to the canonical one. As observed for glutarate, the lower carboxylate group mimics the backbone of L-glutamate and is anchored to Gly512, Ser513, and Ser349 (Fig. 4B). The upper carboxylate, however, deviates from the canonical conformation. One of the O atoms of the upper carboxylate is in the oxyanion hole, as expected for a product, but the other one engages Ser349 rather than pointing to the water-filled cavity. As a result, when Cys348 is rotated into attack position ($\chi_1 = 55^\circ$), the distance between the S atom and the carboxylate C atom is 2.6 Å, which is 0.3 Å longer than observed for L-glutamate and glutarate. The suboptimal placement of the upper carboxylate of succinate is consistent with succinate semialdehyde being a less efficient substrate than GSA and glutarate semialdehyde (19, 21, 22).

3.3.2 3- and 2-carbon ligands: malonate, glyoxylate, and acetate

The pose of malonate (Fig. 5A) deviates substantially from the canonical pose. Although malonate binds to the anchor loop and Ser349, it is too short to occupy the oxyanion hole. Rather, a water molecule fills the oxyanion hole. As a result, the upper carboxylate of malonate is too far from Cys348 for nucleophilic attack. Specifically,

when Cys348 is rotated into the attack rotamer, the S-C distance is 3.2 Å. This result is consistent with the lack of enzymatic activity observed with malonate semialdehyde (21).

Electron density maps suggest that glyoxylate has two conformations that are related by a 180° rotation around the C-C bond axis (Fig. 5B). In both conformations, the carboxylate group is bound to the anchor residues. In conformation A (occupancy 0.6), the carbonyl is directed toward the substrate entrance channel and hydrogen bonds to Ser513. In conformation B (occupancy 0.4), the carbonyl points in the direction of the oxyanion hole. As observed with malonate, glyoxylate is too short to simultaneously occupy the anchor site and the oxyanion hole. This result is consistent with the lack of activity observed with glyoxylate (21).

The carboxylate group of acetate binds to the aldehyde backbone anchor residues (Fig. 5C). Obviously, if acetaldehyde binds similarly, the aldehyde group is out of reach of the Cys nucleophile. The fact that activity, albeit weak, is observed with acetaldehyde suggests that the acetate complex is not representative of enzyme-acetaldehyde complex.

3.3.3 Inhibition of Mm5CDH

The discovery that glyoxylate and malonate bind in the active site, yet neither glyoxylate nor malonate semialdehyde is a substrate for mammalian P5CDHs (21, 22), suggested the possibility that glyoxylate and malonate could be inhibitors. This idea was tested with steady-state kinetic measurements. As a control, we verified that enzymatic activity was undetectable with glyoxylate as the substrate, in agreement with previous studies (21, 22). The analogous check with malonate semialdehyde was not performed due to the lack of a commercial source for this compound.

Glyoxylate is a sub-millimolar competitive inhibitor of MmP5CDH (Table 2). Initial

rate data could be fit satisfactorily to the competitive inhibition model (Supplemental Figure S1A), yielding K_i of 0.27 mM.

Surprisingly, malonate does not appear to be an inhibitor of MmP5CDH. Inhibition was not apparent even when malonate was present at 50 - 100 mM (Supplemental Figure S1B). Inhibition studies using succinate, glutarate, and L-glutamate were also performed. All three compounds are millimolar competitive inhibitors. The estimated K_i values are 58 mM for succinate, 30 mM for glutarate, and 12 mM for L-glutamate. The K_i for L-glutamate agrees with value of 14 mM determined previously for human P5CDH (20).

3.4. Discussion

Pioneering studies of mammalian P5CDHs showed that the length of the semialdehyde chain is an important determinant of substrate selectivity (21). The catalytic efficiency is highest for the 5-carbon substrates GSA and glutarate semialdehyde, somewhat lower for the 4-carbon succinate semialdehyde, and essentially zero with 3- and 2-carbon semialdehydes malonate semialdehyde and glyoxylate.

The structures reported here, along with that of the L-glutamate complex reported previously, provide a satisfying explanation for the substrate selectivity of mammalian P5CDH. Analysis of these structures suggests a structure-activity relationship consisting of four distances that describe the E-S complex (Figure 6, Table 3). These parameters include the distance between the two terminal C atoms of the bound substrate (L_S), the nucleophilic attack distance (d_N), and the distances between the aldehyde carbonyl oxygen atom and the two hydrogen bond donors of the oxyanion hole (d_{OH1} , d_{OH2}). GSA is the most efficient substrate for P5CDH, and thus the distance parameters from the enzyme-L-glutamate complex represent the optimal case.

The distances describing the E-S complex for glutarate semialdehyde are nearly identical to those of GSA (Table 3), which is consistent with the observation that glutarate semialdehyde also has high catalytic efficiency. The somewhat higher efficiency of GSA is likely due to the fact that it forms an extra hydrogen bond that is not possible for glutarate semialdehyde (noted in Figure 4C).

The parameters describing the E-S complex of succinate semialdehyde differ from those of the 5-carbon substrates GSA and glutarate semialdehyde. In particular, L_S is 1.7 Å shorter, d_N is 0.3 Å longer, and the oxyanion hole hydrogen bonding distances are 0.1 - 0.2 Å longer. These values reflect the fact that the aldehyde of succinate semialdehyde is not optimally positioned for nucleophilic attack and are consistent with the lower efficiency of succinate semialdehyde.

The deduced parameters for the inactive semialdehydes, malonate semialdehyde and glyoxylate, are clearly suboptimal. The chain lengths are only 2.5 Å and 1.5 Å for malonate semialdehyde and glyoxylate, respectively, and since these ligands bind to the anchor loop, the aldehyde groups are poorly positioned for nucleophilic attack. In particular, the attack distances are 3.2 and 4.4 Å for malonate semialdehyde and glyoxylate, respectively. Furthermore, the oxyanion hole distances are 3.6 – 5.7 Å, which is beyond optimal hydrogen bonding range. The parameters describing malonate semialdehyde and glyoxylate are thus consistent with the lack of activity observed with these semialdehydes. In summary, active semialdehyde substrates share the common feature of being long enough to span the distance between the anchor loop and the oxyanion hole, which is required to position the aldehyde group for nucleophilic attack by the catalytic Cys.

Absent structural information, the lack of catalytic activity observed with malonate semialdehyde and glyoxylate might lead one to suspect that these compounds and their respective dicarboxylates do not bind the enzyme active site. Our studies show otherwise; glyoxylate and malonate mimic actual substrates by binding to the anchor loop (Figures 5A, 5B).

Glyoxylate was found to be a competitive inhibitor ($K_i = 0.27$ mM). In fact, it is the most potent inhibitor reported to date for P5CDH. The pose of glyoxylate is similar to that of another known P5CDH inhibitor, L-proline. L-proline is a competitive inhibitor (with GSA) of human P5CDH with K_i of 3 mM (20), and we previously determined the structure of MmP5CDH complexed with L-proline (27). Glyoxylate and L-proline bind to P5CDH similarly and form identical hydrogen bonds with the enzyme (Figure 7). Note that the amine of proline and the aldehyde oxygen atom of glyoxylate (conformation A) form a hydrogen bond with the hydroxyl of Ser513. An analogous interaction is present in the L-glutamate complex (Figure 4C) but absent in the glutarate, succinate, and malonate complexes.

The trends in K_i (i.e., affinity) of the ligands investigated here are perhaps unexpected but nonetheless consistent with the structures. Glyoxylate has the highest affinity of the ligands tested. The extra hydrogen bond with Ser513 may partly account for this observation. Also, glyoxylate is accommodated in the active site in two orientations, suggesting a more favorable (or less unfavorable) entropy of binding. The desolvation of carboxylate versus aldehyde groups provides another possible explanation. The free energy of hydration of the acetate ion is approximately -80 kcal/mol, whereas that of acetaldehyde is only about -4 kcal/mol (33-35), implying that a larger desolvation penalty

must be paid to bind a carboxylate group to the enzyme than an aldehyde group. Thus, binding a dicarboxylate ligand, such as malonate, presumably requires a larger desolvation penalty than glyoxylate, which contains only one carboxylate group. The higher affinity of glutarate compared to succinate likely reflects the better oxyanion hole hydrogen bonding of glutarate (Table 3). The higher affinity of L-glutamate compared to glutarate is probably due to the extra hydrogen bond between the amine of L-glutamate and Ser513. Comparing the K_i values for L-glutamate and glutarate suggests that this hydrogen bond contributes about -0.5 kcal/mol to the free energy of binding. Finally, it is puzzling that malonate does not bind to MmP5CDH in solution. This result perhaps reflects the fact that whereas two carboxylate groups must be desolvated, only the lower carboxylate forms compensating hydrogen bonds when bound to the enzyme (Figure 5A). The upper carboxylate fails to engage the hydrogen bond donors of the oxyanion. These ideas about the trade-off between the desolvation penalty of the inhibitor and the compensating hydrogen bonds formed in the enzyme-inhibitor complex could potentially aid the discovery of new ALDH inhibitors (32, 36).

Acknowledgements

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health via Grant GM065546. We thank Dr. Dhiraj Srivastava for advice on expression, purification, and crystallization of MmP5CDH.

References

1. Tanner, J. J. (2008) Structural biology of proline catabolism, *Amino Acids* 35, 719-730.
2. Adams, E., and Frank, L. (1980) Metabolism of proline and the hydroxyprolines, *Annu. Rev. Biochem.* 49, 1005-1061.
3. Valle, D., Goodman, S. I., Harris, S. C., and Phang, J. M. (1979) Genetic evidence for a common enzyme catalyzing the second step in the degradation of proline and hydroxyproline, *J. Clin. Invest.* 64, 1365-1370.
4. Geraghty, M. T., Vaughn, D., Nicholson, A. J., Lin, W. W., Jimenez-Sanchez, G., Obie, C., Flynn, M. P., Valle, D., and Hu, C. A. (1998) Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia, *Hum. Mol. Genet.* 7, 1411-1415.
5. Phang, J. M., Hu, C. A., and Valle, D. (2001) Disorders of proline and hydroxyproline metabolism, In *Metabolic and molecular basis of inherited disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.), pp 1821-1838, McGraw Hill, New York.
6. Valle, D., Goodman, S. I., Applegarth, D. A., Shih, V. E., and Phang, J. M. (1976) Type II hyperprolinemia. Delta1-pyrroline-5-carboxylic acid dehydrogenase deficiency in cultured skin fibroblasts and circulating lymphocytes, *J. Clin. Invest.* 58, 598-603.
7. Yoon, K. A., Nakamura, Y., and Arakawa, H. (2004) Identification of ALDH4 as a p53-inducible gene and its protective role in cellular stresses, *J. Hum. Genet.* 49, 134-140.
8. Polyak, K., Xia, Y., Zweier, J. L., Kinzler, K. W., and Vogelstein, B. (1997) A model for p53-induced apoptosis, *Nature* 389, 300-305.

9. Phang, J. M., Donald, S. P., Pandhare, J., and Liu, Y. (2008) The metabolism of proline, a stress substrate, modulates carcinogenic pathways, *Amino Acids* 35, 681-690.
10. Phang, J. M., Pandhare, J., and Liu, Y. (2008) The metabolism of proline as microenvironmental stress substrate, *J Nutr* 138, 2008S-2015S.
11. Phang, J. M., Liu, W., Hancock, C., and Christian, K. J. (2012) The proline regulatory axis and cancer, *Front. Oncol.* 2, 60.
12. Liang, X., Zhang, L., Natarajan, S. K., and Becker, D. F. (2013) Proline Mechanisms of Stress Survival, *Antioxid. Redox. Signal.*
13. Lee, I. R., Lui, E. Y., Chow, E. W., Arras, S. D., Morrow, C. A., and Fraser, J. A. (2013) Reactive Oxygen Species Homeostasis and Virulence of the Fungal Pathogen *Cryptococcus neoformans* Requires an Intact Proline Catabolism Pathway, *Genetics*, doi: 10.1534/genetics.1113.150326.
14. Lijek, R. S., Luque, S. L., Liu, Q., Parker, D., Bae, T., and Weiser, J. N. (2012) Protection from the acquisition of *Staphylococcus aureus* nasal carriage by cross-reactive antibody to a pneumococcal dehydrogenase, *Proc. Natl. Acad. Sci. USA* 109, 13823-13828.
15. Sophos, N. A., and Vasiliou, V. (2003) Aldehyde dehydrogenase gene superfamily: the 2002 update, *Chem. Biol. Interact.* 143-144, 5-22.
16. Perez-Miller, S. J., and Hurley, T. D. (2003) Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase, *Biochemistry* 42, 7100-7109.
17. Steinmetz, C. G., Xie, P., Weiner, H., and Hurley, T. D. (1997) Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion, *Structure* 5, 701-711.
18. Farres, J., Wang, T. T., Cunningham, S. J., and Weiner, H. (1995) Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis, *Biochemistry* 34, 2592-2598.
19. Farres, J., Julia, P., and Pares, X. (1988) Aldehyde oxidation in human placenta. Purification and properties of 1-pyrroline-5-carboxylate dehydrogenase, *Biochem. J.* 256, 461-467.
20. Forte-McRobbie, C., and Pietruszko, R. (1989) Human glutamic-gamma-semialdehyde dehydrogenase. Kinetic mechanism, *Biochem. J.* 261, 935-943.
21. Forte-McRobbie, C. M., and Pietruszko, R. (1986) Purification and characterization of human liver "high Km" aldehyde dehydrogenase and its identification as glutamic gamma-semialdehyde dehydrogenase, *J Biol Chem* 261, 2154-2163.

22. Small, W. C., and Jones, M. E. (1990) Pyrroline 5-carboxylate dehydrogenase of the mitochondrial matrix of rat liver. Purification, physical and kinetic characteristics, *J Biol Chem* 265, 18668-18672.
23. Srivastava, D., Singh, R. K., Moxley, M. A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2012) The Three-Dimensional Structural Basis of Type II Hyperprolinemia *J. Mol. Biol.* 420, 176-189.
24. Kabsch, W. (2010) XDS, *Acta Crystallogr. D Biol. Crystallogr.* 66, 125-132.
25. Evans, P. (2006) Scaling and assessment of data quality, *Acta Cryst. D62*, 72-82.
26. Adams, P. D., Afonine, P. V., Bunkoczi, 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 Crystallogr. D Biol. Crystallogr.* 66, 213-221.
27. Pemberton, T. A., Still, B. R., Christensen, E. M., Singh, H., Srivastava, D., and Tanner, J. J. (2012) Proline: Mother Nature's cryoprotectant applied to protein crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 68, 1010-1018.
28. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Cryst. D60*, 2126-2132.
29. Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation, *Acta Crystallogr. D Biol. Crystallogr.* 65, 1074-1080.
30. Feng, Z., Chen, L., Maddula, H., Akcan, O., Oughtred, R., Berman, H. M., and Westbrook, J. (2004) Ligand Depot: a data warehouse for ligands bound to macromolecules, *Bioinformatics* 20, 2153-2155.
31. Inagaki, E., Ohshima, N., Takahashi, H., Kuroishi, C., Yokoyama, S., and Tahirov, T. H. (2006) Crystal structure of *Thermus thermophilus* Delta1-pyrroline-5-carboxylate dehydrogenase, *J. Mol. Biol.* 362, 490-501.
32. Riveros-Rosas, H., Gonzalez-Segura, L., Julian-Sanchez, A., Diaz-Sanchez, A. G., and Munoz-Clares, R. A. (2013) Structural determinants of substrate specificity in aldehyde dehydrogenases, *Chem. Biol. Interact.* 202, 51-61.
33. Pearson, R. G. (1986) Ionization potentials and electron affinities in aqueous solution, *J. Am. Chem. Soc.* 108, 6109-6114.
34. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1997) Parametrized Model for Aqueous Free Energies of Solvation Using Geometry-Dependent Atomic Surface Tensions with Implicit Electrostatics, *J. Phys. Chem. B* 101, 7147-7157.

35. Wang, J., Wang, W., Huo, S., Lee, M., and Kollman, P. A. (2001) Solvation Model Based on Weighted Solvent Accessible Surface Area, *J. Phys. Chem. B* *105*, 5055-5067.
36. Koppaka, V., Thompson, D. C., Chen, Y., Ellermann, M., Nicolaou, K. C., Juvonen, R. O., Petersen, D., Deitrich, R. A., Hurley, T. D., and Vasiliou, V. (2012) Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application, *Pharmacol. Rev.* *64*, 520-539.
37. Weiss, M. (2001) Global indicators of X-ray data quality, *J. Appl. Cryst.* *34*, 130-135.
38. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation, *Proteins* *50*, 437-450.
39. DeLano, W. L. (2002) *The PyMOL User's Manual*, DeLano Scientific, Palo Alto, CA, USA.

Table 3.1. Data Collection and Refinement Statistics^a

	Acetate	Glyoxylate	Malonate	Succinate	Glutarate
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 84.9 b = 93.9 c = 132.1	a = 84.9 b = 94.2 c = 132.1	a = 84.9 b = 94.0 c = 132.1	a = 84.9 b = 94.1 c = 132.3	a = 84.9 b = 94.0 c = 132.2
Wavelength	1.541	1.541	1.541	1.541	1.541
Resolution (Å)	19.64 – 1.67 (1.76 – 1.67)	19.67 – 1.67 (1.76 – 1.67)	19.65 – 1.67 (1.76 – 1.67)	19.68 – 1.67 (1.76 – 1.67)	19.66 – 1.81 (1.91 – 1.81)
No. of observations	529175	541047	524924	551588	459704
Unique reflections	119816	120763	120594	118798	94657
$R_{\text{merge}}(\text{I})^{\text{b}}$	0.067 (0.297)	0.068 (0.259)	0.076 (0.436)	0.062 (0.363)	0.092 (0.676)
$R_{\text{meas}}(\text{I})^{\text{b}}$	0.076 (0.350)	0.077 (0.305)	0.086 (0.520)	0.070 (0.422)	0.116 (0.789)
$R_{\text{pim}}(\text{I})^{\text{b}}$	0.035 (0.181)	0.035 (0.157)	0.040 (0.278)	0.031 (0.209)	0.052 (0.397)
Mean $I/\sigma(I)$	14.1 (3.9)	14.8 (4.7)	13.3 (2.8)	16.0 (3.6)	11.3 (2.2)
Completeness (%)	98.2 (89.0)	98.7 (91.9)	98.5 (91.2)	97.0 (87.9)	98.3 (90.9)
Multiplicity	4.4 (3.3)	4.5 (3.5)	4.4 (3.2)	4.6 (3.5)	4.9 (3.7)
No. of protein atoms	8189	8185	8204	8208	8203
No. of waters	660	742	748	799	711
No. of PEG fragments	2	2	2	2	2
No. of ligands	4	3	2	2	2
R_{cryst}	0.168	0.159	0.170	0.164	0.166
$R_{\text{free}}^{\text{c}}$	0.194	0.183	0.198	0.191	0.203
R.m.s. deviations [†]					
bond lengths (Å)	0.006	0.006	0.006	0.006	0.006
bond angles (°)	1.041	1.054	1.032	1.037	1.029
Ramachandran plot ^d (%)					
favored	98.1	98.1	98.3	98.0	98.2
allowed	1.9	1.9	1.7	2.0	1.8
outliers	0	0	0	0	0
Average B (Å ²)					
protein	14.4	12.5	14.6	15.2	16.7
ligand	17.8	18.6	18.7	18.1	24.3
water	20.6	19.5	21.2	22.8	22.3
PEG	24.1	21.5	24.2	27.5	29.0
PDB Code	4LGZ	4LH0	4LH1	4LH2	4LH3

^aValues for the outer resolution shell of data are given in parenthesis.

^bDefinitions of R_{merge} , R_{meas} , and R_{pim} can be found in Weiss (37).

^cCommon 5% test set.

^dThe Ramachandran plot was generated with RAMPAGE (38).

Table 3.2. Kinetic parameters for the inhibition of MmP5CDH^a

	k_{cat} (s ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (s ⁻¹ M ⁻¹)	K_{i} (mM)
glyoxylate	0.160 ± 0.002	26 ± 2	6200 ± 500	0.27 ± 0.02
succinate	0.150 ± 0.002	31 ± 2	4800 ± 300	58 ± 4
glutarate	0.150 ± 0.004	40 ± 4	3800 ± 400	30 ± 3
L-glutamate	0.150 ± 0.004	28 ± 3	5400 ± 600	12 ± 1

^aEnzyme activity was measured at 20 °C with succinate semialdehyde as the variable substrate and the NAD⁺ concentration fixed at 1 mM.

Table 3.3. Inferred geometrical parameters for active and inactive semialdehydes

	L_S^a (Å)	d_N^b (Å)	d_{OH1}^c (Å)	d_{OH2}^c (Å)	Relative k_{cat} / K_m^d
GSA	5.0	2.3	2.7	3.1	1.0
Glutarate semialdehyde	5.0	2.3	2.6	3.1	0.5
Succinate semialdehyde	3.3	2.6	2.8	3.3	0.2
Malonate semialdehyde	2.5	3.2	3.6	5.4	0
Glyoxylate	1.5	4.4	5.0	5.7	0

^aSemialdehyde chain length, defined as the distance between the two terminal C atoms of the substrate in the E-S complex (Figure 6).

^bNucleophilic attack distance, defined as the distance between the C atom of the substrate aldehyde group and the S atom of Cys348 (Figure 6).

^cOxyanion hole hydrogen bond distances. d_{OH1} is the distance between the O atom of the substrate aldehyde group and the N atom of Cys348. d_{OH2} is distance between the O atom of the substrate aldehyde group and the side chain N atom of Asn211 (Figure 6).

^dData from reference (21) expressed relative to P5C/GSA.

Figure Legends

Figure 3.1. The reactions of proline (upper) and hydroxyproline (lower) catabolism in mammals.

Figure 3.2. Structure of MmP5CDH complexed with glutarate (cyan). The NAD⁺-binding, catalytic, and oligomerization domains are colored red, blue, and green, respectively. The side chains of selected active site residues are shown. The inset shows the essential elements of substrate recognition. This figure and others were created with PyMOL (39).

Figure 3.3. Ligands used in crystal structure determinations: (1) glutarate, (2) succinate, (3) malonate, (4) glyoxylate, and (5) acetate.

Figure 3.4. The binding of products to MmP5CDH. Electron density and interactions for (A) glutarate and (B) succinate. The cages represent simulated annealing σ_A -weighted $F_o - F_c$ omit maps contoured at 3.0σ . (C) Stereographic view (relaxed) of a superposition of the active sites of MmP5CDH complexed with L-glutamate (salmon carbons, black dashes, PDB 3V9K), glutarate (green carbons, green dashes), and succinate (gold carbons, gold dashes). The protein for the L-glutamate complex is shown

in white. The red dashes represent the inferred nucleophilic attack distances (2.3 Å for GSA and glutarate semialdehyde; 2.6 Å for succinate semialdehyde).

Figure 3.5. The binding of 3- and 2-carbon ligands to MmP5CDH. Electron density and interactions for (A) malonate, (B) glyoxylate, and (C) acetate. The cages represent simulated annealing σ_A -weighted $F_o - F_c$ omit maps contoured at 3.0 σ . In panel B, conformations A and B of glyoxylate are colored gold and cyan, respectively.

Figure 3.6. Cartoon representation of the essential elements of semialdehyde recognition by MmP5CDH.

Figure 3.7. Comparison of the binding of L-proline and glyoxylate to MmP5CDH. The L-proline complex is colored gray with black hydrogen bonds. The glyoxylate complex is colored gold with gold hydrogen bonds. Conformations A and B of glyoxylate are colored gold and cyan, respectively.

Figure 3.8. Steady-state inhibition kinetics of MmP5CDH. The following potential inhibitors were tested: (A) glyoxylate ($K_i = 0.27 \pm 0.02$ mM), (B) malonate (no inhibition observed), (C) succinate ($K_i = 58 \pm 4$ mM), (D) glutarate ($K_i = 30 \pm 3$ mM), and (E) L-glutamate ($K_i = 12 \pm 1$ mM). The curves represent global fits to the competitive inhibition model. For malonate (B), the global fit collapses to the Michaelis-Menten equation, consistent with no observed inhibition.

Fig 3.1.

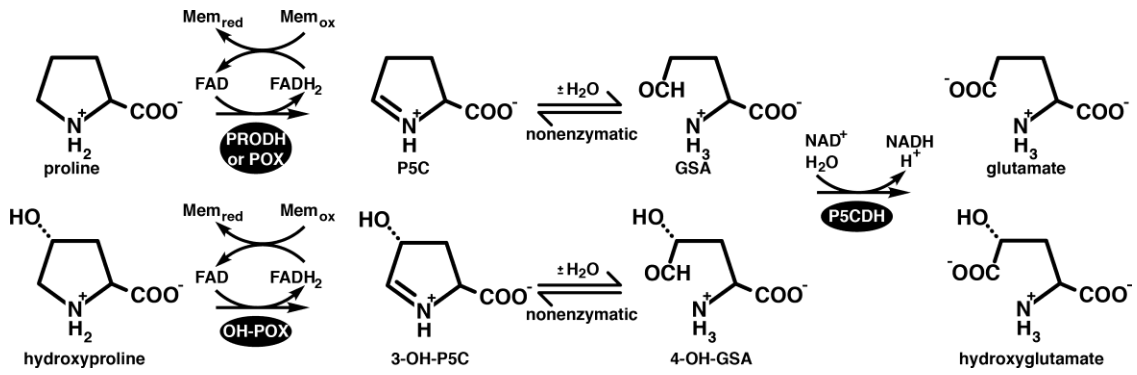


Fig 3.2.

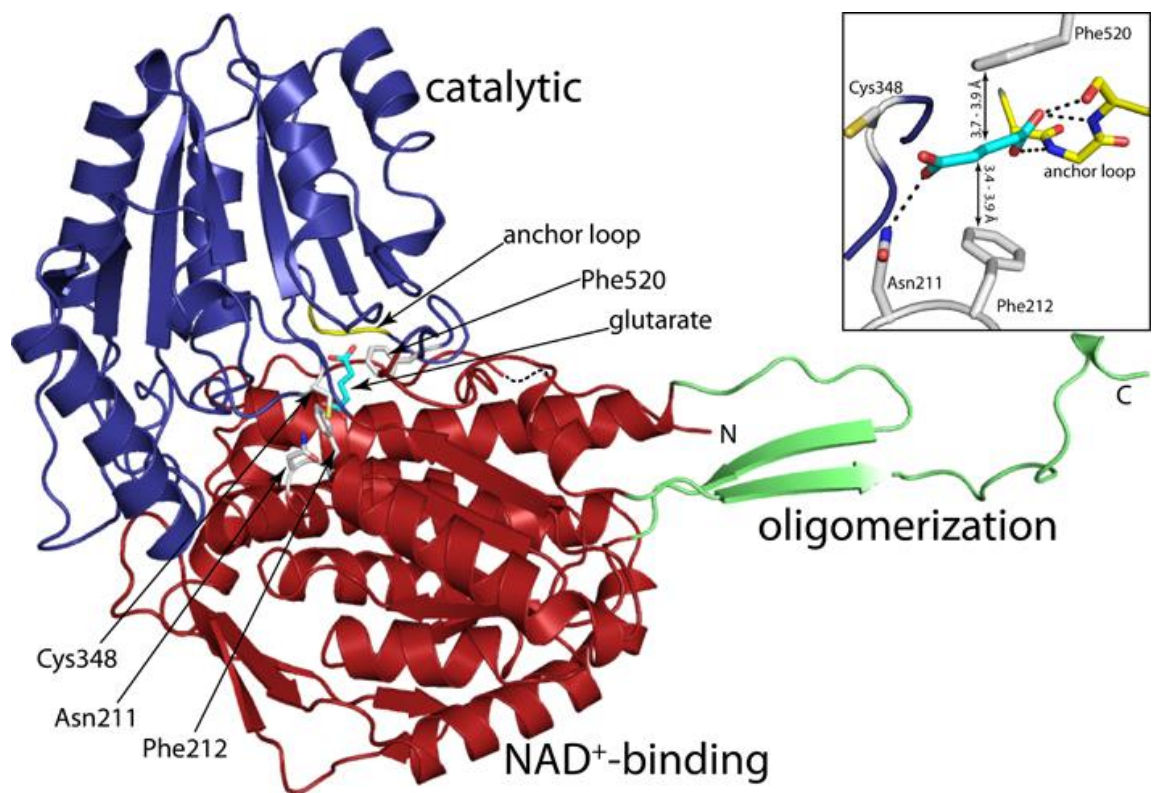


Fig 3.3.

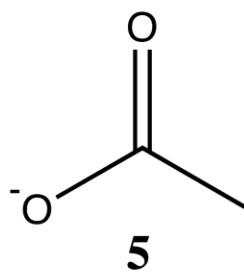
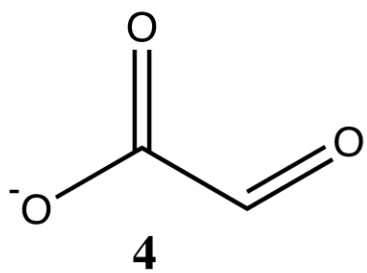
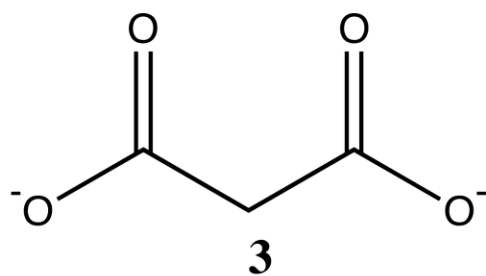
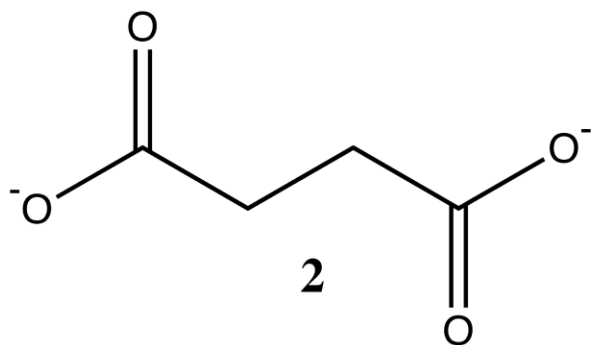
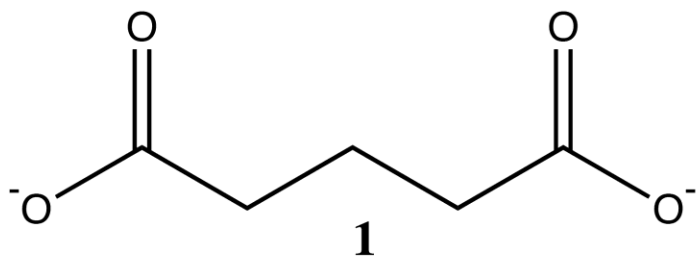


Fig 3.4.

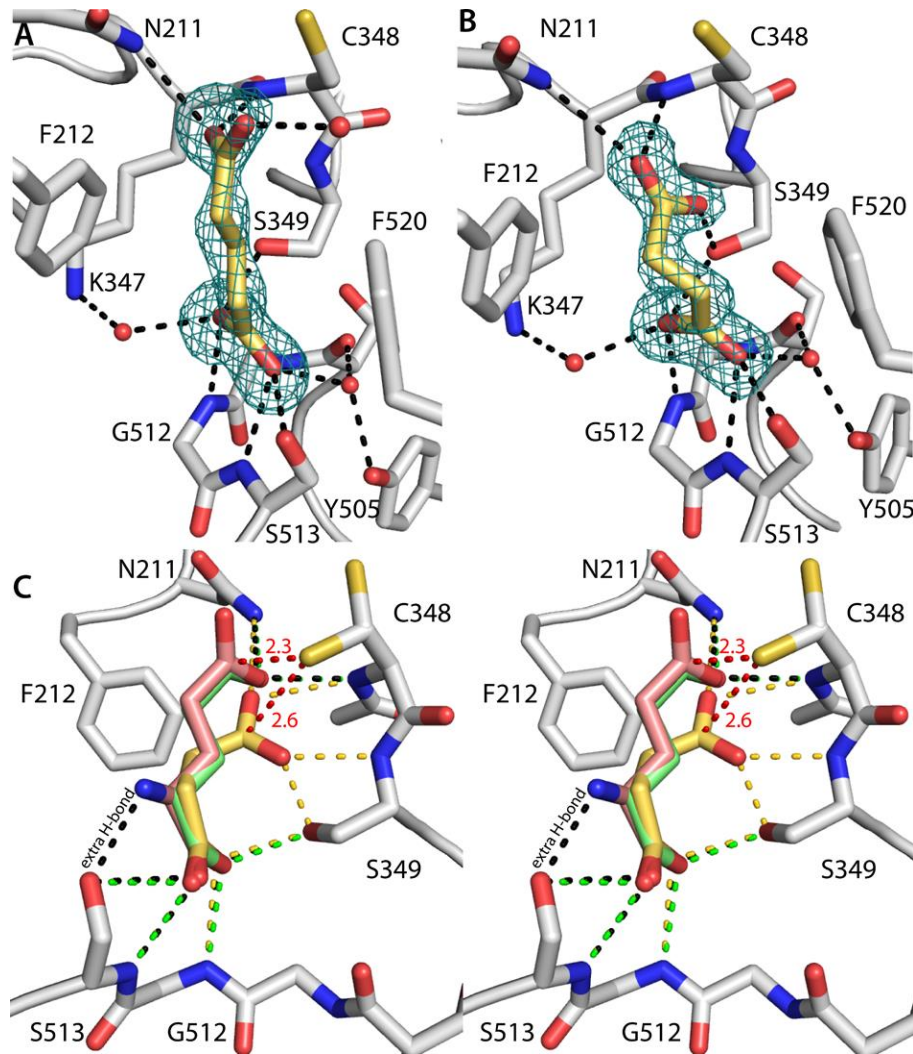


Fig 3.5.

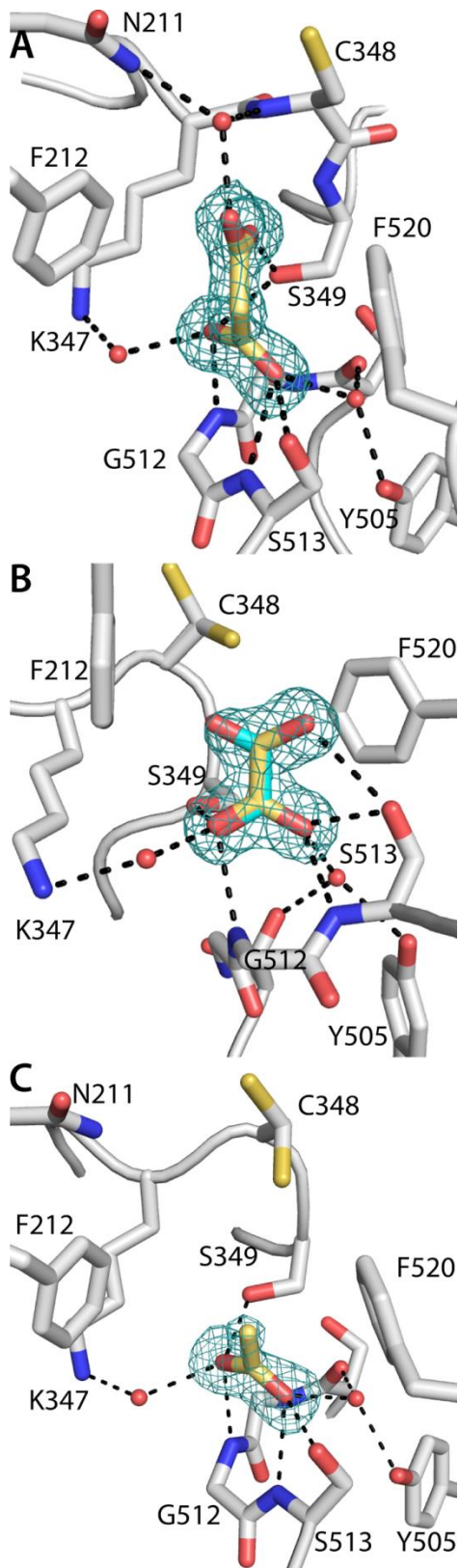


Fig 3.7.

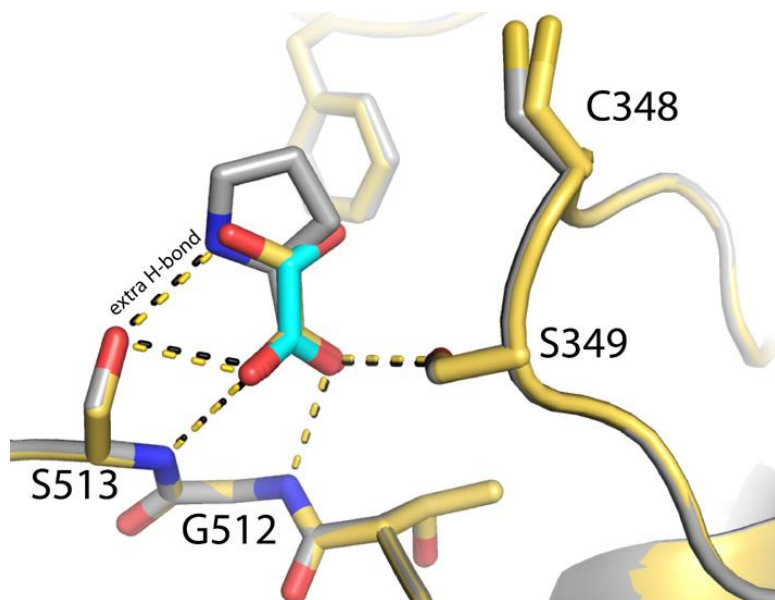
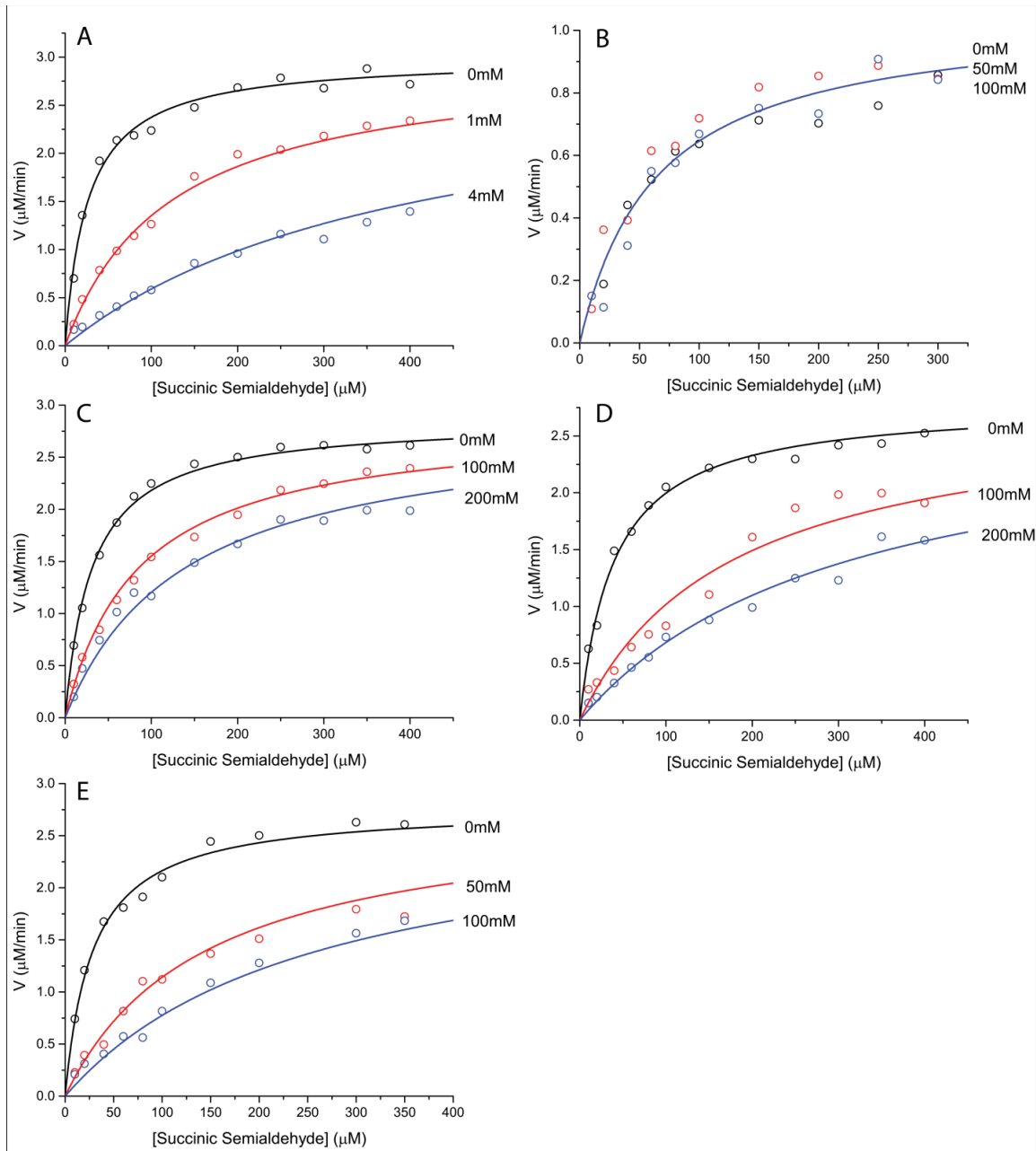


Fig 3.8.



Chapter 4:

Crystal Structure and Unexpected Oligomeric State of Put2p, the Mitochondrial Δ^1 -Pyrroline-5-Carboxylate Dehydrogenase from *Saccharomyces cerevisiae*

Abstract

The proline catabolic enzyme $\Delta(1)$ -pyrroline-5-carboxylate dehydrogenase (ALDH4A1) catalyzes the NAD(+)-dependent oxidation of γ -glutamate semialdehyde to l-glutamate. In *Saccharomyces cerevisiae*, ALDH4A1 is encoded by the PUT2 gene and known as Put2p. Here we report the steady-state kinetic parameters of the purified recombinant enzyme, two crystal structures of Put2p, and the determination of the oligomeric state and quaternary structure from small-angle X-ray scattering and sedimentation velocity. Using $\Delta(1)$ -pyrroline-5-carboxylate as the substrate, catalytic parameters k_{cat} and K_m were determined to be 1.5 s⁻¹ and 104 μ M, respectively, with a catalytic efficiency of 14000 M⁻¹ s⁻¹. Although Put2p exhibits the expected aldehyde dehydrogenase superfamily fold, a large portion of the active site is disordered in the crystal structure. Electron density for the 23-residue aldehyde substrate-binding loop is absent, implying substantial conformational flexibility in solution. We furthermore report a new crystal form of human ALDH4A1 (42% identical to Put2p) that also shows disorder in this loop. The crystal structures provide evidence of multiple active site conformations in the substrate-free form of the enzyme, which is consistent with a conformational selection mechanism of substrate binding. We also show that Put2p forms a trimer-of-dimers hexamer in solution. This result is unexpected because human ALDH4A1 is dimeric, whereas some bacterial ALDH4A1s are hexameric. Thus, global sequence identity and domain of life are poor predictors of the oligomeric states of ALDH4A1. Mutation of a single Trp residue that forms knob-in-hole interactions across the dimer-dimer interface abrogates hexamer formation, suggesting that this residue is the center of a protein-protein association hot spot.

4.1. Introduction

The mitochondrial enzymes Put1p and Put2p catalyze proline catabolism in *Saccharomyces cerevisiae* (Figure 1). Put1p is a flavin - dependent dehydrogenase that catalyzes the oxidation of L-proline to Δ^1 -pyrroline-5-carboxylate (P5C) and directly couples proline oxidation to reduction of ubiquinone (1). The hydrolysis of P5C yields the substrate for Put2p, γ -glutamate semialdehyde (GSA). Put2p catalyzes the NAD⁺-dependent oxidation of γ -glutamate semialdehyde to L-glutamate. Put2p is a P5C dehydrogenase (P5CDH, EC 1.5.1.12) and is also known as ALDH4A1 in reference to its membership in family 4 of the aldehyde dehydrogenase (ALDH) superfamily. In total, the proline catabolic pathway affects a 4-electron oxidation of proline, and these reactions enable yeast to use proline as the sole source of nitrogen.

Several studies have shown that proline is important in various organisms for stress protection (2). Proline accumulation in yeast leads to improved freeze tolerance (3). In contrast to proline, accumulation of P5C is associated with toxicity effects and induced cell death in yeast (4, 5). Coordinated expression of *PUT1* and *PUT2* is important for avoiding buildup of P5C. The *PUT1* and *PUT2* genes are upregulated by proline in nitrogen limiting conditions and together with glutamate dehydrogenase, enable yeast to utilize proline as nitrogen source (6). The mechanism of P5C toxicity involves inhibition of mitochondrial respiration and increased production of reactive oxygen species (4). Interestingly, a N-acetyltransferase enzyme known as Mpr1 acetylates P5C thereby diminishing its accumulation and harmful effects in yeast (7).

Mammalian P5CDHs have been characterized biochemically. Early studies on the human (HsP5CDH) and rat enzymes established the order of substrate binding and

product release, as well as substrate preferences (8-10). More recent work on HsP5CDH and mouse P5CDH (MmP5CDH) provided additional details of the kinetic mechanism and the molecular basis of type II hyperprolinemia,(11) a metabolic disorder caused by a defect in HsP5CDH function (12-14).

The three-dimensional structures of P5CDHs have been studied intensely. Tahirov's group reported the first P5CDH structure using the enzyme from *Thermus thermophilus* (TtP5CDH) (15, 16). We subsequently determined the structures of HsP5CDH and MmP5CDH (11, 17), and the New York Structural Genomics Research Consortium has deposited structures of two *Bacillus* P5CDHs in the Protein Data Bank (PDB codes 3QAN and 3RJL). Collectively, the structures confirmed P5CDH as a member of the ALDH superfamily, contributed additional information about the catalytic mechanism, and provided insight into the structural basis of cofactor and semialdehyde selectivity. We also determined the solution oligomeric states and quaternary structures of several P5CDHs using a variety of biophysical techniques (18). The human, mouse, and *Bacillus* enzymes are dimeric in solution, whereas TtP5CDH and *Deinococcus radiodurans* P5CDH (DrP5CDH) form trimer-of-dimers hexamers in solution. Thus, the protein fold, but not the oligomeric state, is conserved within the ALDH4A1 family.

Although the yeast PUT genes were identified over 30 years ago, the Put1p and Put2p enzymes have remained relatively understudied. Wanduragala et al. reported the purification and characterization of Put1p in 2010 (1), but analogous studies of Put2p have not appeared in the literature. We therefore have created a recombinant bacterial expression system for Put2p, measured the steady-state kinetic parameters of the purified

enzyme, determined the crystal structure, and deduced the oligomeric state and quaternary structure in solution from small-angle X-ray scattering (SAXS).

4.2. Experimental Procedures

4.2.1 Subcloning and Mutagenesis

The *PUT2* gene encoding the 575-residue preprotein was obtained from Dr. Marjorie Brandriss of the Rutgers New Jersey Medical School and subcloned into vectors pET14b using NdeI and XhoI and pKA8H between NdeI and BamHI restriction sites resulting in pET14b-PUT2 and pKA8H-PUT2 constructs. DNA encoding residues 20-575 (According to the sequence I got from W193A and the one I have of the wt gene, it should start at 20), which lacks the N-terminal mitochondrial targeting peptide (19), was amplified from pKA8H using AAACATATGAAGCCCCCAAGCACATAAG as the forward primer and the T7 terminator as the reverse primer. The resulting fragment was subcloned into pKA8H using NdeI and BamHI to give the construct pKA8H-PUT2 Δ 22. The expressed protein has an N-terminal His₈ tag and Tobacco Etch Virus protease (TEVP) cleavage site.

The W193A mutant of Put2p was created from pKA8H-PUT2 Δ 22 with the Quick-Change II site-directed mutagenesis kit (Agilent) using the forward primer (5'-GGTACAATCTACCGTTTGCTGACAGATCTGC-3') and reverse primer (5'-GCAGATCTGTCAGCAAACGGTAGATTGTACC-3'). The mutation was confirmed using DNA sequencing.

4.2.2 Expression and Purification of Put2p

Put2p and Put2p mutant W193A were expressed in BL21(DE3)pLysS cells. Cells were grown at 37 °C and 250 rpm until an OD₆₀₀ of 0.6 and induced with 0.5 mM IPTG

for 8 hours at 22 °C at 200 rpm. The cells were harvested by centrifugation at 3500 rpm for 30 minutes and resuspended in 20 mM HEPES, 60 mM NaCl, and 5 % glycerol at pH 8.0. Cells were quick frozen in liquid nitrogen and stored at -80 °C until further purification.

Thawed cells were broken by sonication and centrifuged at 16500 rpm for one hour. The supernatant was applied to a HisTrap HP column that had been charged with Ni²⁺ and equilibrated with 20 mM HEPES, 300 mM NaCl, and 5 % glycerol at pH 8.0. The column was washed with equilibration buffer supplemented with 30 mM imidazole, and the enzyme was eluted with 300 mM imidazole. Fractions containing Put2p were pooled and dialyzed into 50 mM Tris, 0.5 mM EDTA, 0.5 mM THP, and 5 % glycerol at pH 8.0 and loaded onto a HiTrap Q anion exchange column. Put2p was eluted with a linear gradient of NaCl from 0 to 1 M. Fractions were analyzed with SDS-PAGE, pooled, and dialyzed into 50 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, and 5 % glycerol at pH 8.0.

4.2.3 Crystallization of Put2p

Crystallization trials were performed in sitting drops at 20 °C using drops formed by mixing 1 µL each of the enzyme and reservoir solutions. Crystal screening using commercially available kits (Hampton Research) was used to identify promising crystallization conditions.

Put2p expressed from pKA8H-PUT2Δ22, including the N-terminal His tag, was used for crystallization. Optimized crystals were grown using the enzyme at 8 mg/mL and a reservoir containing 8 - 14 % (w/v) polyethylene glycol (PEG) 3350, 0.1 M (NH₄)₂SO₄, and 0.1 M Bis-Tris at pH 5.5 - 6.5. The crystals were cryoprotected with 15 % PEG 3350,

0.1 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Bis-Tris at pH 6.0, and 20 % glycerol. The cryoprotected crystals were harvested with nylon loops (Hampton Research) and plunged into liquid nitrogen. The space group is $P6_3$ with $a = 109 \text{ \AA}$ and $c = 181 \text{ \AA}$. The asymmetric unit contains one dimer and 51 % solvent (20).

Crystals of Put2p complexed with NAD^+ were obtained by adding 10mM NAD^+ to a 7 mg/mL protein stock. Optimized crystals were grown from combining this stock and a reservoir containing 8-14% PEG 3350, 0.1 M $(\text{NH}_4)_2\text{SO}_4$ and 0.1 M Bis-Tris at pH 5.5 - 6.5. These crystals were cryoprotected with 12% PEG 3350, 0.1 M $(\text{NH}_4)_2\text{SO}_4$, 10mM NAD^+ , 0.1 M Bis-Tris at pH 5.5 and 15% PEG 200. The cryoprotected crystals were harvested with nylon loops and plunged into liquid nitrogen. The space group is $P6_3$ with $a = 108 \text{ \AA}$ and $c = 181 \text{ \AA}$. The asymmetric unit was the same as before.

4.2.4 Expression, Purification, and Crystallization of Tag-free HsP5CDH

HsP5CDH residues 18-563 with an N-terminal His tag was expressed and purified as described previously (11). The His tag was removed with thrombin as follows. The purified protein was dialyzed into thrombin cleavage buffer (20 mM Tris, 150 mM NaCl, 2.5 mM CaCl_2 , and 5 % glycerol at pH 8.0) and incubated with thrombin (1 unit of thrombin per 4 mg of HsP5CDH) for one day at 4 °C. Thrombin was removed using a HiTrap Benzamidine FF affinity column. The cleaved His tag and residual unprocessed HsP5CDH were removed by passing the sample through a Ni^{2+} -charged HiTrap HP column. The tag-free enzyme was collected in the flow-through, dialyzed into 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, and 5 % glycerol at pH 8.0, and concentrated to 6 mg/mL for crystallization trials.

Monoclinic crystals of HsP5CDH were grown using a reservoir of 20 - 25 % (w/v) PEG3350, 0.2 M MgCl₂, and 0.1 M HEPES at pH 7.0 - 8.0. The crystals were cryoprotected with 25 % (w/v) PEG 3350, 0.2 M MgCl₂, 0.1 M HEPES at pH 7.5, and 20 % (v/v) PEG 200. The cryoprotected crystals were harvested with nylon loops and plunged into liquid nitrogen. The space group is $P2_1$ with $a = 92.0 \text{ \AA}$, $b = 121.3 \text{ \AA}$, $c = 93.4 \text{ \AA}$, and $\beta = 104.2^\circ$. The asymmetric unit contains two dimers and 39 % solvent (20).

4.2.5 X-ray Diffraction Data Collection, Phasing, and Refinement

Crystals were analyzed at NE-CAT beam lines 24-ID-C and 24-ID-E at the Advanced Photon Source using a Quantum 315 detector (Table 1). The Put2p data set was obtained at 24-ID-C and consisted of 105 frames collected with an oscillation width of 1° , detector distance of 225 mm, and exposure time of 1.0 s/frame at 7 % transmission. The data were processed to 1.95 \AA resolution using HKL2000 (21). Data from Put2p-NAD⁺ crystals were collected 24-ID-E and processed with XDS (22). The 2.2 \AA resolution data set consisted of 70 frames collected with an oscillation width of 1° , detector distance of 200 mm, and exposure time of 1.0 s/frame at 55 % transmission. Data from HsP5CDH crystals were collected at beam line 24-ID-C. The data set used for refinement consisted of 180 frames collected with an oscillation width of 1° , detector distance of 250 mm, and exposure time of 1.0 s/frame at 4 % transmission. The data were processed to 1.95 \AA resolution using HKL2000.

The phase problem was solved using molecular replacement as implemented in MOLREP (23). The search model for Put2p was derived from a 1.3 \AA resolution structure of mouse P5CDH (PDB code 3V9J (11), 43 % sequence identity over 543 residues). The search model for HsP5CDH was obtained from a 2.5 \AA resolution

structure of HsP5CDH (PDB code 3V9G(11)). The models from molecular replacement were used as the starting points for several rounds of model building with COOT (24) and refinement with PHENIX (25). The refined Put2p structure was used as the starting point for refinement of the Put2p-NAD⁺ structure.

4.2.6 Small-Angle X-ray Scattering of Put2p

In preparation for SAXS studies, wild-type Put2p was purified in a similar manner as that used for crystallization. However, after the first Ni column, TEVP was added to the pooled fractions and incubated at 28 °C for 2 hours. The protein was then placed in a dialysis buffer, overnight, consisting of 20mM HEPES, 100mM NaCl and 5% glycerol at pH 8.2. The protein was then run over a Ni column again and the protein eluted at 30mM imidazole. The fractions were collected, pooled, and placed in a dialysis buffer of 50mM Tris, pH 7.5, 0.5mM EDTA, 0.5mM THP, and 5% glycerol overnight. The protein was then run over an anion exchange column and dialyzed by the same procedure for the crystallization. For the last step of purification this protein was run on a gel filtration column. The aggregate was discarded and homogeneous protein was collected. The fractions were pooled and concentration was estimated using the BCA method to be 3 mg/mL. The protein was concentrated in a Millipore centrifugal filter unit to 7 mg/mL. This protein was used for SAXS and the buffer from the flow through of the centrifugal unit was used as a reference.

The Put2p mutant W193A was also analyzed with SAXS. The sample was prepared for SAXS analysis as follows. W193A followed the same purification process as the wild type with two exceptions. First, the final buffer used was 50mM Tris, 300mM NaCl, 0.5mM EDTA, 0.5mM THP, and 5% glycerol at pH 7.5. Second, This protein was run

over a size exclusion chromatography column and no aggregation peak was observed. Fractions were collected and pooled and concentrated to a final concentration of 3mg/mL.

SAXS experiments were performed at SIBYLS beamline 12.3.1 of the Advanced Light Source through the mail-in program (26, 27). For each sample, scattering intensities were measured at three nominal protein concentrations. Data were collected for each protein concentration at exposure times of 0.5 s, 1.0 s, 3.0 s, and 6.0 s. The scattering curves collected from the protein samples were corrected for background scattering using intensity data collected from the reference buffer.

The SAXS data were analyzed as follows. Composite scattering curves were generated with PRIMUS (28) by scaling and merging the background-corrected low q region data from the 0.5 s exposure (wild-type) or 0.1 s exposure (W193A) with the high q region data from the 3.0 s exposure (wild-type) or 6.0 s exposure (W193A). PRIMUS was also used to perform Guinier analysis. FoXS was used to calculate theoretical scattering profiles from atomic models.(29) GNOM was used to calculate pair distribution functions (30). The SASTBX server (31) was used for shape reconstruction calculations.

4.2.7 Analytical Ultracentrifugation

W193A was analyzed by sedimentation velocity using a Beckman XL-I analytical ultracentrifuge. The protein was purified as described above. The His-tag was removed by adding 1mg of TEVP to the pooled fractions. This sample was incubated at 28 °C for 2 hours and then moved into dialysis at 4 °C for about 14 hours. Following dialysis to equilibrium against the buffer (50 mM Tris, 300 mM NaCl, 0.5 mM EDTA, 0.5 mM

THP, and 5% glycerol at pH 7.5), the protein solution was diluted with dialysis buffer to yield an absorbance (at 280 nm) of 0.95 in a 1.0 cm path length cuvette. A 400 μL aliquot was loaded into the sample compartment of a double-sector cell, assembled with a 1.2 cm charcoal-epon centerpiece and quartz windows. A slightly larger volume (430 μL) of dialysis buffer was placed in the reference compartment. The sample cell was placed in the rotor and allowed to equilibrate to 20 °C in the rotor chamber for two hours under vacuum, prior to beginning the experiment. The sample was then sedimented at 30000 rpm, acquiring data at 4 min intervals, until a total of 170 scans had been collected. The resulting data set was analyzed by SEDFIT v9.4 (32), using the continuous $c(s)$ and continuous $c(M)$ distribution models, allowing the frictional ratio to float.

4.2.8 Steady-State Kinetics

Put2p activity was measured by monitoring the formation of NADH at 20 °C as described previously for human P5CDH (11). The assay buffer contained 50 mM potassium phosphate (pH 7.5) and 25 mM NaCl. Put2p enzyme concentration was 0.4 μM (25.6 $\mu\text{g/ml}$) and NAD^+ concentration was held constant at 0.2 mM. Initial velocity data were collected using L-P5C concentrations from 1 to 300 μM . Initial velocity data were fitted to the Michaelis-Menten equation to estimate K_m and k_{cat} .

4.3. Results

4.3.1 Steady-State Kinetics Measurements

Brandriss and coworkers previously measured Put2p enzymatic activity in cellular and mitochondrial extracts to verify the identity of the enzyme and determine the subcellular location (33-35). To our knowledge, the catalytic properties of the purified enzyme have not been studied. We therefore performed steady-state kinetic

measurements of recombinant Put2p. Kinetic parameters for Put2p were $K_m = 104 \pm 4$ μM L-P5C and $k_{cat} = 1.5 \pm 0.3 \text{ sec}^{-1}$ with $k_{cat}/K_m = 14423 \text{ M}^{-1} \text{ s}^{-1}$. The measured activity of Put2p is lower than human P5CDH which has a K_m of 32 μM L-P5C and k_{cat} of 10.0 s^{-1} ($k_{cat}/K_m = 312500 \text{ M}^{-1} \text{ s}^{-1}$) (11).

4.3.2 Protomer Structure of Put2p

The 1.95 Å resolution crystal structure of Put2p was determined (Table 1). Put2p exhibits the expected ALDH fold, which consists of three domains (Figure 2A). The N-terminal half of the polypeptide chain contains a Rossmann fold domain and binds NAD^+ (residues 43-181, 200-319, 536-553). The catalytic domain (residues 320-535) interrupts the NAD^+ -binding domain and exhibits an open α/β fold featuring a twisted 7-stranded β -sheet with all but one strand in parallel. This domain furnishes the essential cysteine nucleophile (Cys351), which is situated at the interface of the NAD^+ -binding and catalytic domains. The third domain is a bipartite β substructure consisting of residues 182-199 and 554-565 that protrudes from the NAD^+ -binding domain. This domain is involved in oligomerization. As with other ALDHs, Put2p forms a dimer in which the oligomerization domain of one protomer engages the catalytic domain of the other protomer (Figure 2B). The asymmetric unit contains one dimer. As described below, three Put2p dimers assemble into a hexamer in solution.

The most conspicuous feature of the Put2p structure is that which is unseen. ALDHs have a ~25 residue peptide that connects the last strand of the catalytic domain to the second piece of the oligomerization domain. This section corresponds to residues 523 - 547 of Put2p. Electron density for these residues is very weak, implying substantial conformational disorder (Figure 3A). The lack of density for this part of the protein

gives the impression of a large void running lengthwise through the middle of the dimer (Figure 3A).

The omission of residues 524 – 546 from the Put2p model is notable because this region of ALDHs anchors the aldehyde substrate in the active site (Figure 3B). In other P5CDHs (and ALDHs in general) the 25-residue aldehyde-binding loop folds into a 2-stranded antiparallel β -sheet that forms the floor of the aldehyde substrate entrance tunnel. In P5CDHs, the beginning of the peptide provides critical hydrogen bonding groups that anchor the backbone of GSA, as well as a conserved Phe that packs against the aliphatic chain of GSA (Figure 3B) (11, 15, 17). The observation of disorder in the aldehyde anchor peptide suggests the hypothesis that in solution the anchor peptide samples conformations other than the one that binds the substrate.

4.3.3 Structure of Put2p-NAD⁺ Complex

The structure of Put2p complexed with NAD⁺ was determined (Table 1). Electron density for the ADP group of NAD⁺ is strong in both protomers (Figure 4). Density for the nicotinamide ribose is strong in one protomer and somewhat weaker in the other one. Density for the nicotinamide is diffuse in both protomers, implying disorder, and thus the nicotinamide was omitted from the final model deposited in the PDB. Disorder of the nicotinamide has been reported for ALDHs (11, 15, 36).

NAD⁺ binds in the expected location, with the adenine ring wedged between valine and proline residues belonging to helices of the Rossmann fold domain (Figure 4). The adenine ribose forms hydrogen bonds with Lys234, and the pyrophosphate interacts with Ser288. The nicotinamide ribose hydrogen bonds with Glu458. These interactions are also observed in the structure of MmP5CDH-NAD⁺ (11).

Binding of the cofactor is accompanied by conformational changes outside of the anchor peptide (Figure 4). Phe460 rotates around χ_1 by 96° to avoid steric clash with the nicotinamide ribose. Glu458 rotates by 120° , enabling a hydrogen bond with the nicotinamide ribose. These conformational changes have not been reported for other P5CDHs.

4.3.4 Structure of Monoclinic HsP5CDH

To further explore the idea that the aldehyde anchor peptide is flexible in solution, the structure of a second eukaryotic P5CDH without a ligand bound in the aldehyde site was determined. The structure of ligand-free HsP5CDH was determined from a new monoclinic crystal form that was obtained only after cleaving the N-terminal His tag.

As in the Put2p structure, electron density for the aldehyde anchor peptide of monoclinic HsP5CDH (residues 514-535) is very weak (Figure 3C). Electron density for these residues is essentially absent in chains A and D, and the entire loop was omitted. In the other two chains in the asymmetric unit, the binding of Mg^{2+} from the crystallization buffer stabilizes residues 522 - 535 in a non-native conformation, but the other residues of the loop remain disordered. Thus, the disordered active site is not an idiosyncrasy of Put2p.

4.3.5 Oligomeric State and Quaternary Structure of Put2p

The oligomeric state of Put2p in solution was investigated using SAXS (Figure 5). Guinier plots of data from three samples having different protein concentrations yield a radius of gyration (R_g) in the range 47 – 48 Å. Calculations of the pair distribution function suggest a maximum particle dimension of 140 – 150 Å and R_g of 45 Å. The R_g of the Put2p dimer in asymmetric unit (Figure 2B) is only 31.5 Å, which suggests that a

higher order oligomer is formed in solution. Furthermore, the theoretical SAXS curve calculated from the dimer deviates substantially from the experimental one (Figure 5A).

The Put2p crystal lattice was inspected to identify a higher order assembly that is consistent with the SAXS data. Application of the crystallographic three-fold rotation to the dimer in the asymmetric unit generates a point group 32 hexamer with R_g of 43.9 Å (Figure 5B), which is close to the experimental R_g of 48 Å. The scattering profile calculated from the Put2p hexamer shows good agreement with the experimental profile (Figure 5A). Slightly better agreement is obtained with the hexamer of TtP5CDH (Figure 5A), which includes the sections that are disordered in Put2p. Consideration of a mixture of hexamers and dimers using minimum ensemble search (37) did not improve the fit.

The hexamer oligomeric state was confirmed by estimating the molecular mass from SAXS data using the volume of correlation (V_c) invariant (38). The V_c of Put2p is 1459 Å², which corresponds to a molecular mass of 363 kDa. This value is within 6 % of the expected molecular mass of 387 kDa for a hexamer of the His-tagged Put2p protein that was used for SAXS. It is concluded that the predominant form of Put2p in solution is the trimer-of-dimers hexamer observed in the crystal lattice.

Finally, shape reconstruction calculations are also consistent with the hexamer. The shape reconstruction was performed with the SASTBX server (31), which makes no explicit assumptions about the molecular mass, number of residues, or oligomeric state of the particle under consideration. The envelope is consistent with the crystallographic hexamer (Figure 5B).

4.3.6 Hexamerization Hot Spot of Put2p

We previously identified a hexamerization hot spot in bacterial P5CDHs. The hot spot is located in the interface between two dimers and is formed by $\alpha 3$ of one molecule and the oligomerization domain of another molecule. An Arg residue of $\alpha 3$ (e.g., Arg100 of TtP5CDH) was shown to be essential for hexamer formation, as mutation of this residue to Ala abrogates hexamerization of TtP5CDH and DrP5CDH. Interestingly, Ala126 of Put2p replaces this important arginine, while Trp193 of the Put2p oligomerization domain occupies the space corresponding to the Arg side chain of TtP5CDH. The indole of Trp193 fits into a hole formed by several residues of $\alpha 3$ (Phe121, Tyr122, Ser125, Ala126, Leu129) as well as Val174 of $\alpha 5$ (Figure 6). We therefore hypothesized that Trp193 is the center of the hexamerization hot spot of Put2p, analogous to the essential Arg of Tt5CDH and DrP5CDH. This idea was tested by creating the Put2p variant W193A.

The oligomeric state of W193A was determined using sedimentation velocity. Several of the velocity profiles are displayed in Figure 7A, together with the corresponding least-squares fits. Analysis of the entire data set with SEDFIT yielded a symmetric sedimentation coefficient distribution centered at 4.65 S (Figure 7B). The corresponding molecular mass distribution is centered at 121 kDa (Figure 7C), which is within 2 % of the theoretical M of the W193A dimer (124 kDa). Thus, mutation of Trp193 to Ala appears to have disrupted the hexamer, as intended.

W193A was also analyzed using SAXS (Figure 8). Guinier plots of data from three samples having different protein concentrations yield an R_g in the range of 34.5 – 34.7 Å. Calculations of the pair distribution function suggest a maximum particle dimension of

110-115 Å and an R_g of 34.7 Å. As stated above the R_g of the Put2p dimer in the asymmetric unit is 31.5 Å, which is close to the experimental R_g . The theoretical SAXS curve calculated from the Put2p dimer is in good agreement with the experimental one, while the hexamer curve shows significant deviation (Figure 8A). Also, shape reconstruction calculations, performed with the SASTBX server, are consistent with the crystallographic dimer (Figure 8B). Thus, the SAXS data are also consistent with W193A being primarily dimeric in solution.

4.4. Discussion

It is tempting to speculate that the Put2p and HsP5CDH structures reported here represent the enzyme conformation prior to aldehyde binding, implying that aldehyde binding stabilizes the active site. This interpretation is supported by the fact that all other structures of eukaryotic P5CDHs (human and mouse) have ligands bound in the aldehyde site, including sulfate ion from the crystallization buffer (PDB entries 3V9J, 3V9H) (11), the product glutamate (3V9K) (11), and proline (4E3X) (39), which is a competitive inhibitor with respect to GSA. Although a ligand was not modeled in two other structures (3V9L, 3V9G), a strong electron density feature consistent with a sulfate ion is nevertheless evident. Thus, none of the previously determined structures of eukaryotic P5CDHs are truly representative of the aldehyde-free conformation. Bacterial P5CDH structures, however, provide a counterargument; a few bacterial P5CDHs have an ordered active site in the absence of a ligand in the aldehyde site (e.g., PDB entries 2EHQ, 2BHP, 3RJL). Nevertheless, the observation that the aldehyde anchor peptide is disordered in two different P5CDHs that were crystallized in unrelated lattices suggests that the disorder may be functionally relevant, at least for eukaryotic P5CDHs. At the very least,

the structures demonstrate that the aldehyde-binding loop is inherently flexible, possibly sampling active and inactive conformations in solution. In this view, substrate binding to Put2p would be an example of conformational selection.

Disorder in enzyme active sites on the scale described here has precedent, and retinal dehydrogenase II is the most germane example. Also a member of the ALDH superfamily, retinal dehydrogenase II catalyzes the oxidation of retinal to retinoic acid, and is thus involved in retinoic acid signaling. As in Put2p, the aldehyde binding loop is disordered in the crystal structure of retinal dehydrogenase II (PDB code 1BI9). It has been proposed that the disordered loop plays a role in discriminating retinal from smaller potential substrates (40, 41). It is likewise possible that the flexibility of ALDH4A1 plays a role in enforcing substrate selectivity.

The discovery that Put2p is hexameric is unexpected. Put2p and HsP5CDH are eukaryotic P5CDHs that share 42 % amino acid sequence identity, yet Put2p is hexameric and HsP5CDH is dimeric (11). The trimer-of-dimers hexamer described here for Put2p is also formed by TtP5CDH and DrP5CDH (18), although Put2p is only 30 % identical in sequence to these bacterial P5CDHs. Curiously other bacterial P5CDHs, such as those from *Bacillus* species are dimeric (18). Thus, global sequence identity and domain of life are poor predictors of the oligomeric states of P5CDHs. Instead, we suggest that the hot spot theory of protein-protein association provides a better explanation for the oligomeric states of P5CDHs (42, 43).

We previously proposed the existence of a hot spot that is essential for hexamer formation of P5CDH, based on biophysical and mutagenesis studies of the two hexameric P5CDHs TtP5CDH and DrP5CDH (18). The hot spot involves residues in the

oligomerization domain and $\alpha 3$. Bacterial hexameric P5CDHs have an essential Arg that forms intermolecular electrostatic interactions in the hot spot. Interestingly, this key Arg residue is Ala in Put2p (Figure 6), and so the discovery that Put2p is hexameric in solution was quite unexpected. The Put2p structure revealed Trp193 that forms nonpolar, knob-in-hole interactions in the hot spot region (Figures 6B, 6C). Mutation of Trp193 to Ala abrogates hexamer formation, suggesting that these nonpolar interactions substitute for the electrostatic interactions that stabilize the bacterial P5CDH hexamers. Thus, the hot spot appears to be a unique region of P5CDH three-dimensional structural space that must be occupied by residues capable of forming attractive intermolecular forces in order for hexamerization to occur. Interestingly, the precise nature of these forces appears to be less important than their location.

ASSOCIATED CONTENT

Accession Codes

Atomic coordinates and structure factors have been deposited in the Protein Data Bank as entries 4OE6 (Put2p), 4OE4 (Put2p-NAD⁺), and 4OE5 (HsP5CDH).

ACKNOWLEDGEMENTS

We thank Dr. Marjorie Brandriss of the Rutgers New Jersey Medical School for providing the PUT2 clone, Dr. Jonathan Schuermann of the Northeastern Collaborative Access Team at the Advanced Photon Source for help with diffraction data collection and

processing, and Kevin Dyer of the SIBYLS beamline of the Advanced Light Source for collecting the SAXS data. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by a grant from the National Institute of General Medical Sciences (P41 GM103403) from the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357. Part of this work was conducted at the the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the National Institute of Health project MINOS (R01GM105404).

References

1. Wanduragala, S., Sanyal, N., Liang, X., and Becker, D. F. (2010) Purification and characterization of Put1p from *Saccharomyces cerevisiae*, *Arch Biochem Biophys* 498, 136-142.
2. Liang, X., Zhang, L., Natarajan, S. K., and Becker, D. F. (2013) Proline Mechanisms of Stress Survival, *Antioxid. Redox. Signal.*
3. Morita, Y., Nakamori, S., and Takagi, H. (2002) Effect of proline and arginine metabolism on freezing stress of *Saccharomyces cerevisiae*, *J Biosci Bioeng* 94, 390-394.
4. Nishimura, A., Nasuno, R., and Takagi, H. (2012) The proline metabolism intermediate Delta1-pyrroline-5-carboxylate directly inhibits the mitochondrial respiration in budding yeast, *FEBS Lett.* 586, 2411-2416.
5. Nomura, M., and Takagi, H. (2004) Role of the yeast acetyltransferase Mpr1 in oxidative stress: regulation of oxygen reactive species caused by a toxic proline catabolism intermediate, *Proc. Natl. Acad. Sci. USA* 101, 12616-12621.

6. Magasanik, B., and Kaiser, C. A. (2002) Nitrogen regulation in *Saccharomyces cerevisiae*, *Gene* 290, 1-18.
7. Nasuno, R., Hirano, Y., Itoh, T., Hakoshima, T., Hibi, T., and Takagi, H. (2013) Structural and functional analysis of the yeast N-acetyltransferase Mpr1 involved in oxidative stress tolerance via proline metabolism, *Proc Natl Acad Sci U S A* 110, 11821-11826.
8. Forte-McRobbie, C., and Pietruszko, R. (1989) Human glutamic-gamma-semialdehyde dehydrogenase. Kinetic mechanism, *Biochem. J.* 261, 935-943.
9. Forte-McRobbie, C. M., and Pietruszko, R. (1986) Purification and characterization of human liver "high Km" aldehyde dehydrogenase and its identification as glutamic gamma-semialdehyde dehydrogenase, *J Biol Chem* 261, 2154-2163.
10. Small, W. C., and Jones, M. E. (1990) Pyrroline 5-carboxylate dehydrogenase of the mitochondrial matrix of rat liver. Purification, physical and kinetic characteristics, *J Biol Chem* 265, 18668-18672.
11. Srivastava, D., Singh, R. K., Moxley, M. A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2012) The Three-Dimensional Structural Basis of Type II Hyperprolinemia *J. Mol. Biol.* 420, 176-189.
12. Valle, D., Goodman, S. I., Applegarth, D. A., Shih, V. E., and Phang, J. M. (1976) Type II hyperprolinemia. Delta1-pyrroline-5-carboxylic acid dehydrogenase deficiency in cultured skin fibroblasts and circulating lymphocytes, *J. Clin. Invest.* 58, 598-603.
13. Geraghty, M. T., Vaughn, D., Nicholson, A. J., Lin, W. W., Jimenez-Sanchez, G., Obie, C., Flynn, M. P., Valle, D., and Hu, C. A. (1998) Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia, *Hum. Mol. Genet.* 7, 1411-1415.
14. Phang, J. M., Hu, C. A., and Valle, D. (2001) Disorders of proline and hydroxyproline metabolism, In *Metabolic and molecular basis of inherited disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.), pp 1821-1838, McGraw Hill, New York.
15. Inagaki, E., Ohshima, N., Takahashi, H., Kuroishi, C., Yokoyama, S., and Tahirov, T. H. (2006) Crystal structure of *Thermus thermophilus* Delta1-pyrroline-5-carboxylate dehydrogenase, *J. Mol. Biol.* 362, 490-501.
16. Inagaki, E., Ohshima, N., Sakamoto, K., Babayeva, N. D., Kato, H., Yokoyama, S., and Tahirov, T. H. (2007) New insights into the binding mode of coenzymes: structure of *Thermus thermophilus* [Delta]1-pyrroline-5-carboxylate dehydrogenase complexed with NADP⁺, *Acta Cryst. F* 63, 462-465.

17. Pemberton, T. A., and Tanner, J. J. (2013) Structural Basis of Substrate Selectivity of 1-Pyrroline-5-Carboxylate Dehydrogenase (ALDH4A1): Semialdehyde Chain Length, *Arch Biochem Biophys*, under review.
18. Luo, M., Singh, R. K., and Tanner, J. J. (2013) Structural determinants of oligomerization of delta(1)-pyrroline-5-carboxylate dehydrogenase: identification of a hexamerization hot spot, *J. Mol. Biol.* 425, 3106-3120.
19. Martin, O., Brandriss, M. C., Schneider, G., and Bakalinsky, A. T. (2003) Improved anaerobic use of arginine by *Saccharomyces cerevisiae*, *Appl. Environ. Microbiol.* 69, 1623-1628.
20. Matthews, B. W. (1968) Solvent content of protein crystals, *J. Mol. Biol.* 33, 491-497.
21. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol* 276, 307-326.
22. Kabsch, W. (2010) XDS, *Acta Crystallogr. D Biol. Crystallogr.* 66, 125-132.
23. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *J. Appl. Crystallogr.* 30, 1022-1025.
24. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Cryst. D60*, 2126-2132.
25. Adams, P. D., Afonine, P. V., Bunkoczi, 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 Crystallogr. D Biol. Crystallogr.* 66, 213-221.
26. Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., 2nd, Tsutakawa, S. E., Jenney, F. E., Jr., Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S. J., Scott, J. W., Dillard, B. D., Adams, M. W., and Tainer, J. A. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS), *Nat. Methods* 6, 606-612.
27. Classen, S., Hura, G. L., Holton, J. M., Rambo, R. P., Rodic, I., McGuire, P. J., Dyer, K., Hammel, M., Meigs, G., Frankel, K. A., and Tainer, J. A. (2013) Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source, *J. Appl. Crystallogr.* 46, 1-13.
28. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *J. Appl. Crystallogr.* 36, 1277-1282.

29. Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles, *Nucleic Acids Res.* 38, W540-544.
30. Svergun, D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr.* 25, 495-503.
31. Liu, H., Hexemer, A., and Zwart, P. H. (2012) The Small Angle Scattering ToolBox (SASTBX): an open-source software for biomolecular small-angle scattering, *J. Appl. Crystallogr.* 45, 587-593.
32. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling, *Biophys. J.* 78, 1606-1619.
33. Brandriss, M. C., and Krzywicki, K. A. (1986) Amino-terminal fragments of delta 1-pyrroline-5-carboxylate dehydrogenase direct beta-galactosidase to the mitochondrial matrix in *Saccharomyces cerevisiae*, *Mol. Cell. Biol.* 6, 3502-3512.
34. Brandriss, M. C., and Magasanik, B. (1979) Genetics and physiology of proline utilization in *Saccharomyces cerevisiae*: enzyme induction by proline, *J. Bacteriol.* 140, 498-503.
35. Brandriss, M. C. (1983) Proline utilization in *Saccharomyces cerevisiae*: analysis of the cloned PUT2 gene, *Mol. Cell. Biol.* 3, 1846-1856.
36. Perez-Miller, S. J., and Hurley, T. D. (2003) Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase, *Biochemistry* 42, 7100-7109.
37. Pelikan, M., Hura, G. L., and Hammel, M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering, *Gen. Physiol. Biophys.* 28, 174-189.
38. Rambo, R. P., and Tainer, J. A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering, *Nature* 496, 477-481.
39. Pemberton, T. A., Still, B. R., Christensen, E. M., Singh, H., Srivastava, D., and Tanner, J. J. (2012) Proline: Mother Nature's cryoprotectant applied to protein crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 68, 1010-1018.
40. Lamb, A. L., and Newcomer, M. E. (1999) The structure of retinal dehydrogenase type II at 2.7 Å resolution: implications for retinal specificity, *Biochemistry* 38, 6003-6011.
41. Bordelon, T., Montegudo, S. K., Pakhomova, S., Oldham, M. L., and Newcomer, M. E. (2004) A disorder to order transition accompanies catalysis in retinaldehyde dehydrogenase type II, *J Biol Chem* 279, 43085-43091.
42. Bogan, A. A., and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces, *J. Mol. Biol.* 280, 1-9.

43. Moreira, I. S., Fernandes, P. A., and Ramos, M. J. (2007) Hot spots--a review of the protein-protein interface determinant amino-acid residues, *Proteins* 68, 803-812.
44. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation, *Proteins* 50, 437-450.

Table 4.1. X-ray Diffraction Data Collection and Refinement^a

	Put2p	Put2p-NAD ⁺	HsP5CDH
Space group	$P6_3$	$P6_3$	$P2_1$
Unit cell parameters (Å, °)	$a = 109.0, c = 181.2$	$a = 108.0, c = 181.0$	$a = 92.0, b = 121.3, c = 93.4, \beta = 104.2$
Wavelength	0.979	0.979	0.979
Resolution (Å)	50.0 – 1.95 (2.02 – 1.95)	93.6 – 2.17 (2.29 – 2.17)	50.0 – 1.95 (2.02 – 1.95)
Observations	579452	277102	489835
Unique reflections	88348	63148	139296
$R_{merge}(I)$	0.061 (0.525)	0.057 (0.583)	0.049 (0.453)
Mean I/σ	32.1 (3.4)	16.5 (2.3)	24.2 (2.2)
Completeness (%)	99.9 (100.0)	99.8 (98.09)	96.5 (93.0)
Multiplicity	6.6 (6.5)	4.4 (4.4)	3.5 (3.3)
No. of protein residues	960	960	2054
No. of atoms	7611	7485	15436
No. of NAD ⁺ atoms	0	62	0
No. of water molecules	199	177	241
R_{cryst}	0.181 (0.208)	0.180 (0.285)	0.194 (0.229)
R_{free}^b	0.207 (0.236)	0.220 (0.319)	0.237 (0.302)
rmsd bond lengths (Å)	0.007	0.007	0.007
rmsd bond angles (°)	1.04	1.09	1.00
Ramachandran plot ^c			
Favored (%)	98.62	98.09	98.23
Outliers (residues)	0	0	0
Average B (Å ²)			
Protein	30.5	40.0	33.3
NAD ⁺	-	42.6	-
Water	29.1	35.7	27.2
Coordinate error (Å) ^d	0.21	0.27	0.24
PDB code	4OE6	4OE4	4OE5

^aValues for the outer resolution shell of data are given in parenthesis.

^b5 % test set. A common set was used for refinement of the Put2p structures.

^cThe Ramachandran plot was generated with RAMPAGE.⁽⁴⁴⁾

^dMaximum likelihood-based coordinate error estimate reported by PHENIX.

Figure Legends

Figure 4.1. The reactions catalyzed by the yeast proline catabolic enzymes Put1p and Put2p.

Figure 4.2. Structure of Put2p. (A) The protomer is shown with the catalytic domain in green, NAD⁺-binding domain in blue, and oligomerization domain in gold. The red section represents the aldehyde-binding loop of MmP5CDH, which is disordered in Put2p. (B) Structure of the Put2p dimer in the asymmetric unit. The two protomers are colored blue and gold. Red sections represent the aldehyde-binding loop of MmP5CDH, which is disordered in Put2p

Figure 4.3. Electron density evidence for disorder in the aldehyde-binding loops of Put2p and HsP5CDH. (A) The Put2p dimer with $2F_o - F_c$ electron density (1.0σ). The red sections represent the aldehyde-binding loops of MmP5CDH. The green spheres represent the product glutamate bound to MmP5CDH. (B) Close-up view of the expected conformation of the aldehyde-binding loop in Put2p. Put2p is colored gray. The aldehyde-binding loop (pink), bound glutamate ligand (green), and NAD⁺ of MmP5CDH have been fitted onto Put2p. (C) A dimer of monoclinic HsP5CDH with $2F_o - F_c$ electron density (1.0σ). The red sections represent the aldehyde-binding loops of MmP5CDH, which are disordered or form a nonnative conformation in HsP5CDH. The green spheres represent the product glutamate bound to MmP5CDH.

Figure 4.4. Electron density and interactions for NAD⁺ bound to Put2p. The cage represents a simulated annealing σ_A -weighted $F_o - F_c$ omit map contoured at 3.0σ . The conformations of Glu458 and Phe460 prior to NAD⁺ binding are shown in cyan.

Figure 4.5. SAXS analyses of Put2p. (A) Experimental and calculated SAXS curves. The inset shows a Guinier plot spanning the range of $qR_g = 0.508 - 1.29$. The linear fit of the Guinier plot has R^2 of 0.9995. (B) Superposition of the SAXS shape reconstruction and the Put2p hexamer generated from crystallographic symmetry. Two orthogonal views are shown.

Figure 4.6. Hexamer hot spot of Put2p. (A) A view of the hexamer showing the location of Trp193 (shown in spheres). The six chains have different colors. (B and C) Close-up views of the environment around Trp193. The chains are colored as in panel A.

Figure 4.7. Sedimentation velocity analysis of W193A. (A) Experimental velocity profiles (circles) and least squares fits. (B) Sedimentation coefficient distribution. (C) Molecular mass distribution.

Figure 4.8. SAXS analysis of W193A. (A) Experimental and calculated SAXS curves. The inset shows a Guinier plot spanning the range of $qR_g = 0.346 - 1.34$. The linear fit of the Guinier plot has R^2 of 0.997. (B) Superposition of the SAXS shape reconstruction and the Put2p dimer from the asymmetric unit. Two orthogonal views are shown.

Fig 4.1.

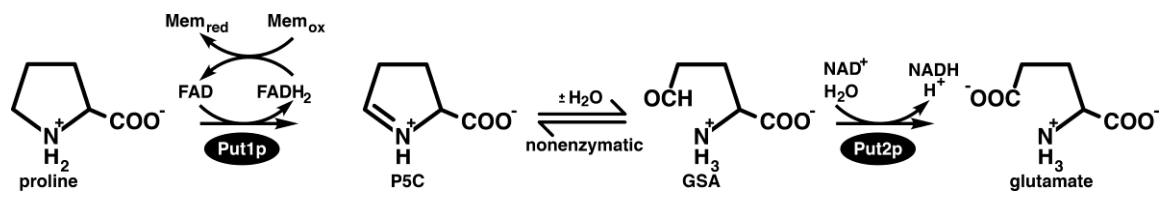


Fig 4.2.

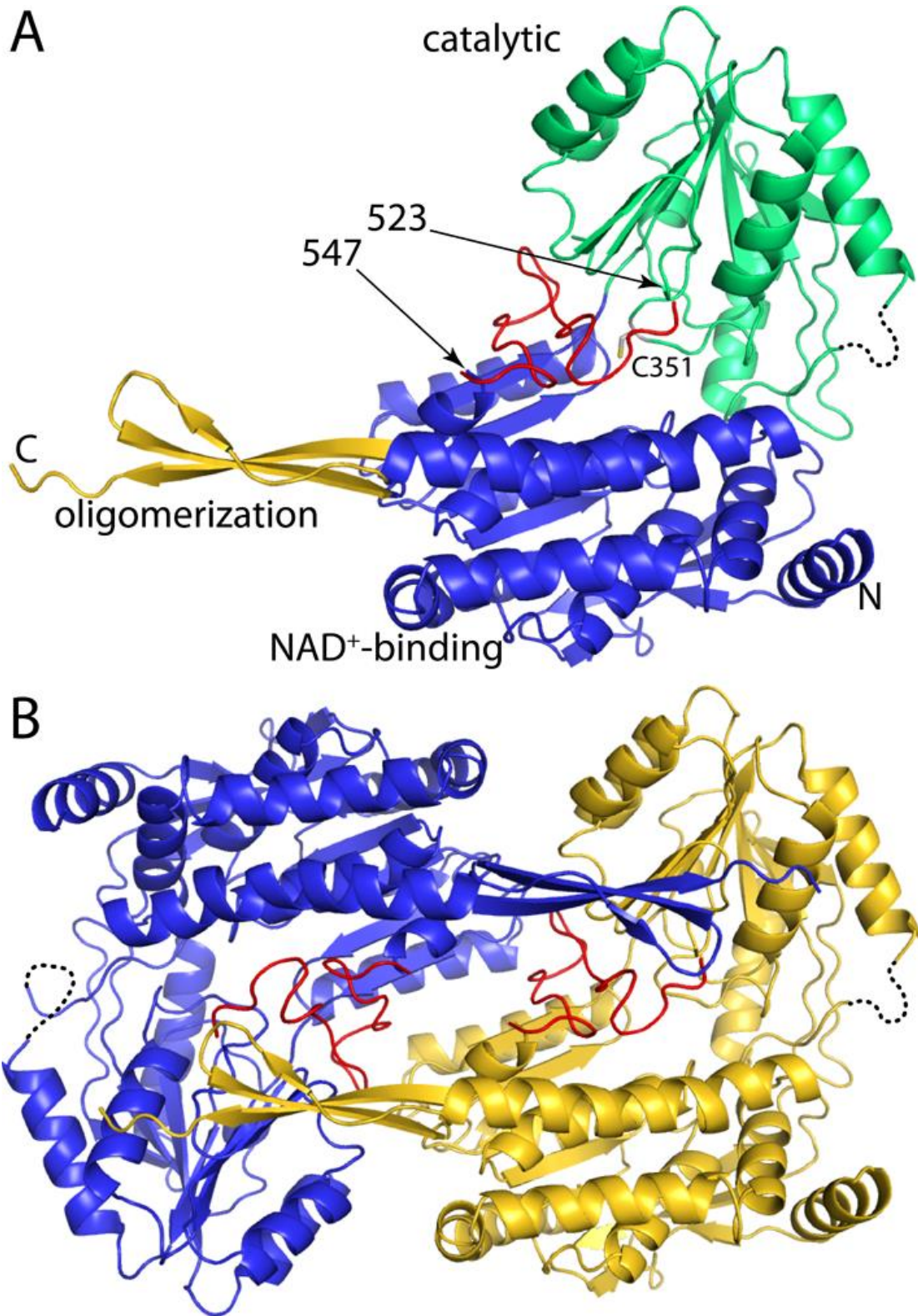
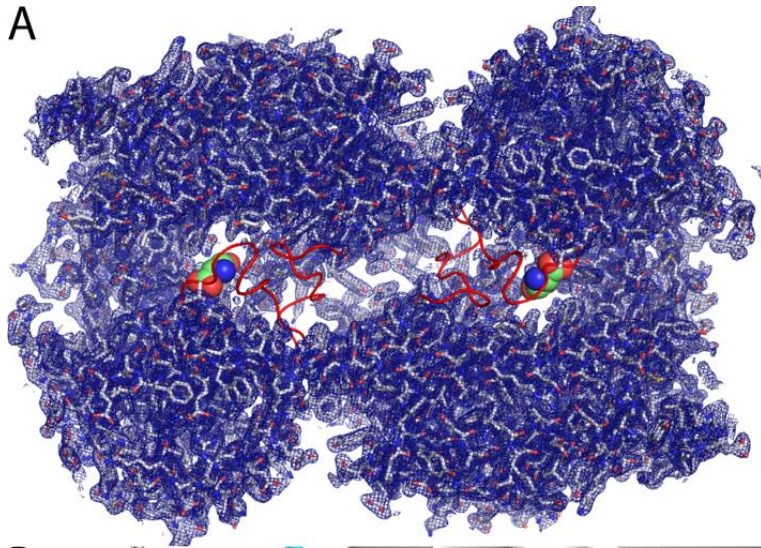
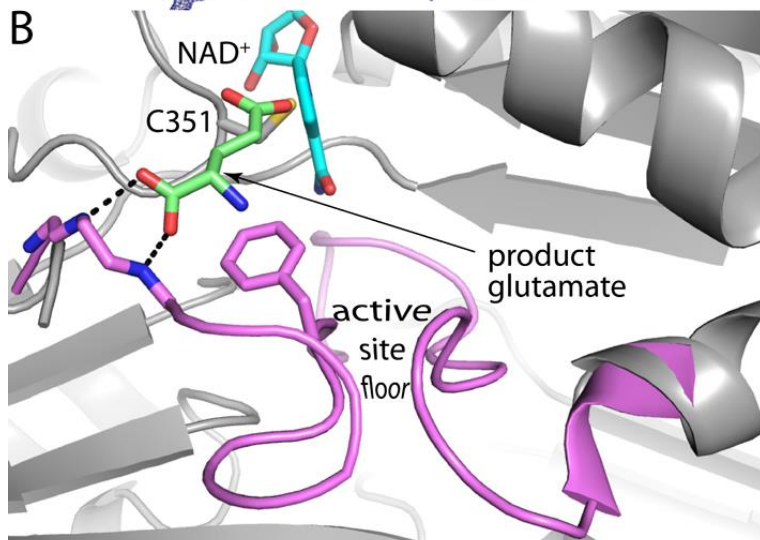


Fig 4.3.

A



B



C

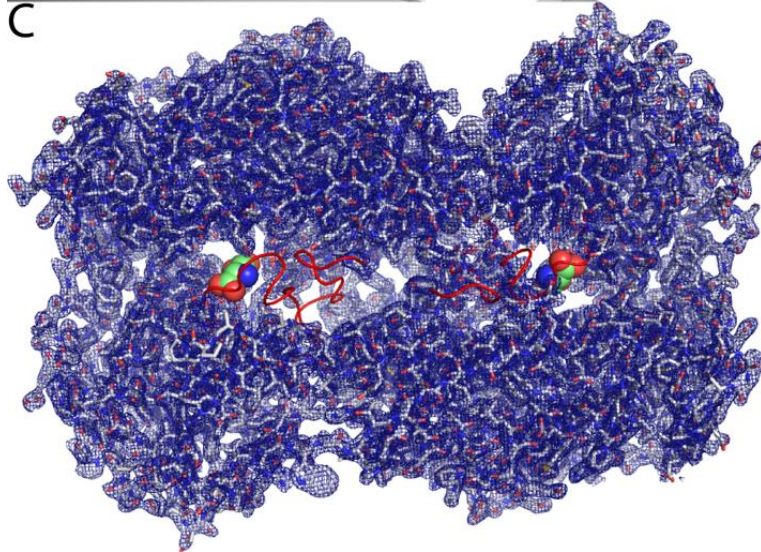


Fig 4.4.

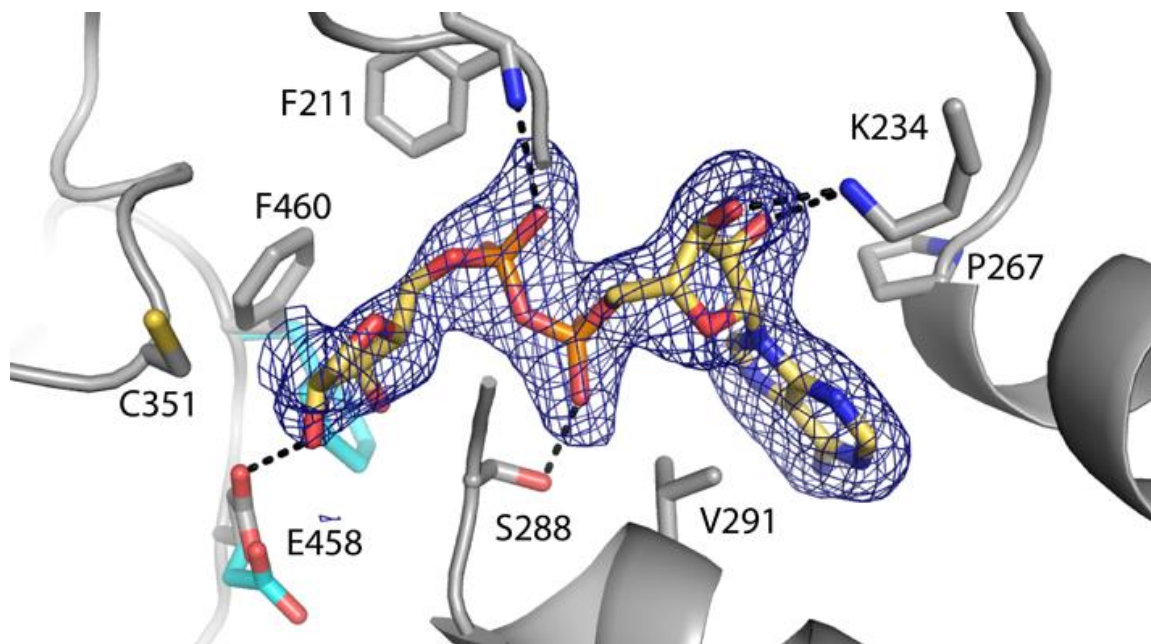


Fig 4.5.

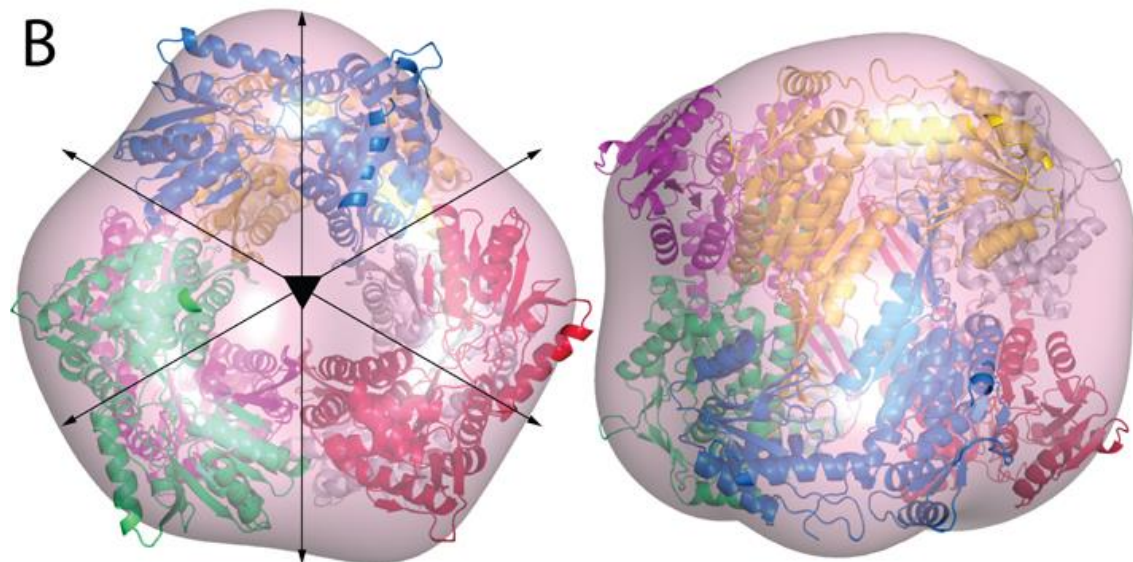
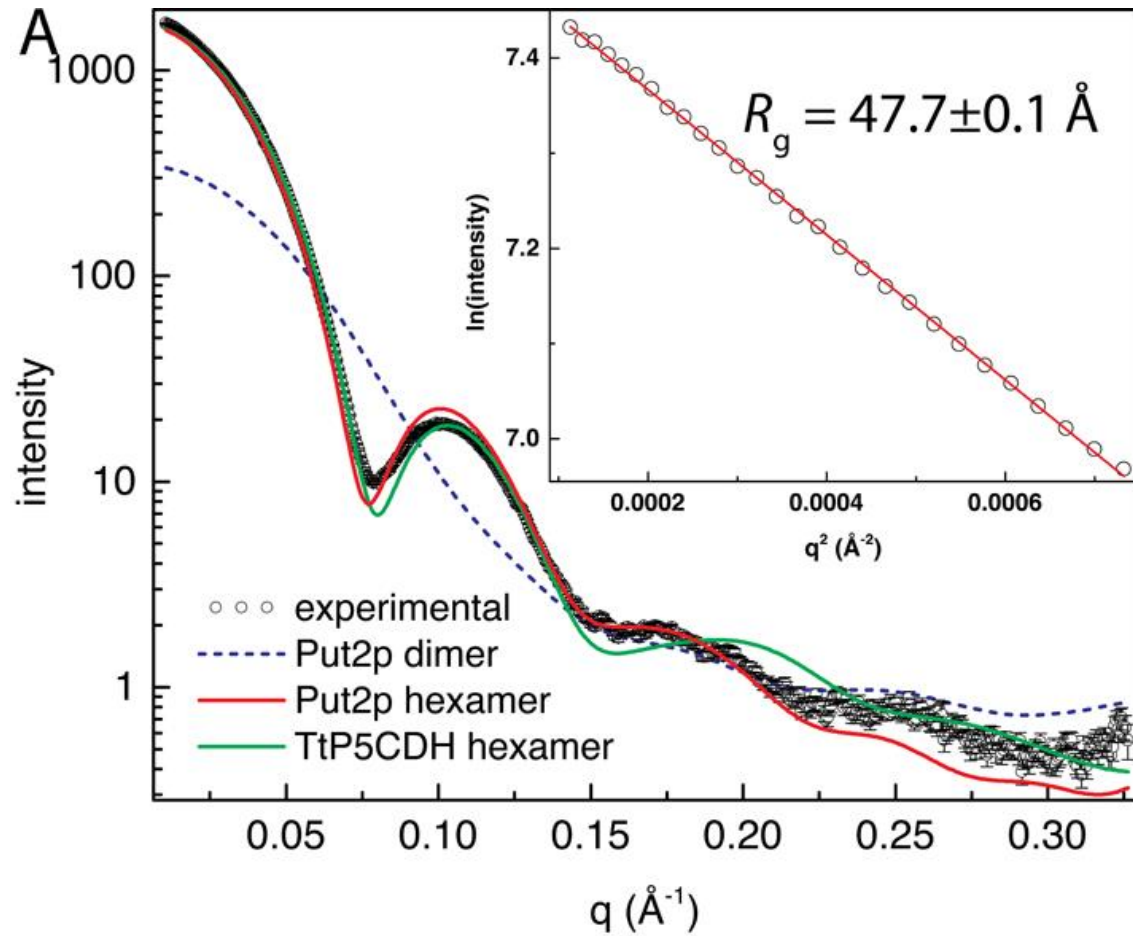


Fig 4.6.

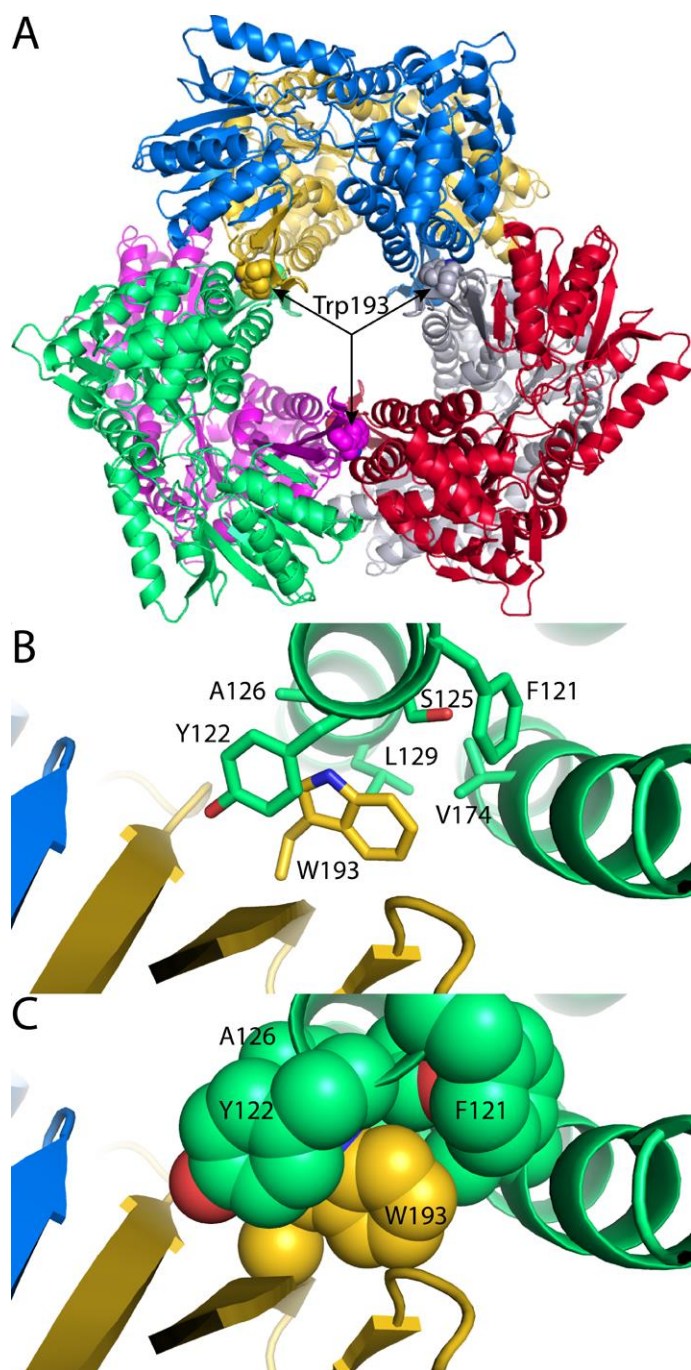


Fig 4.7.

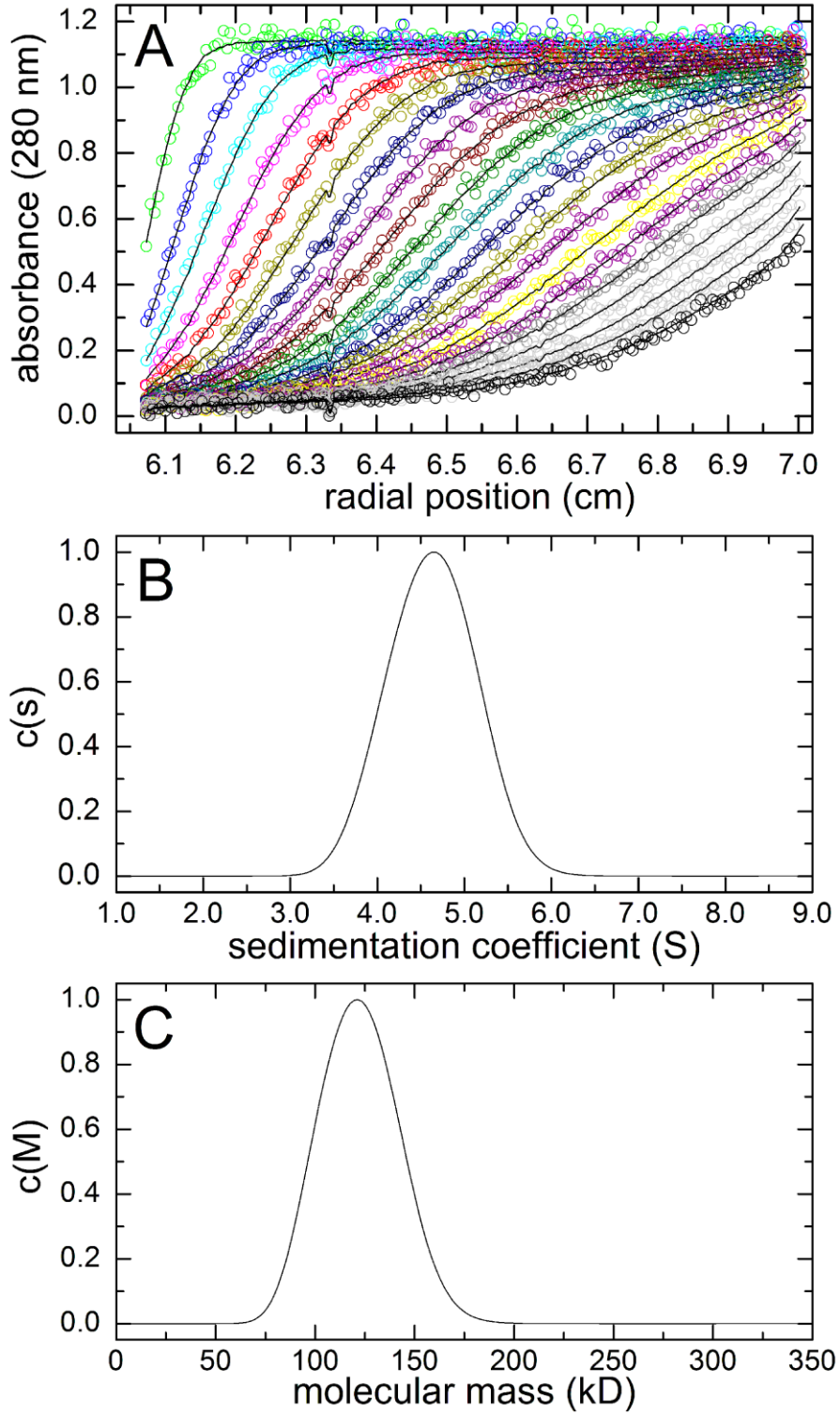
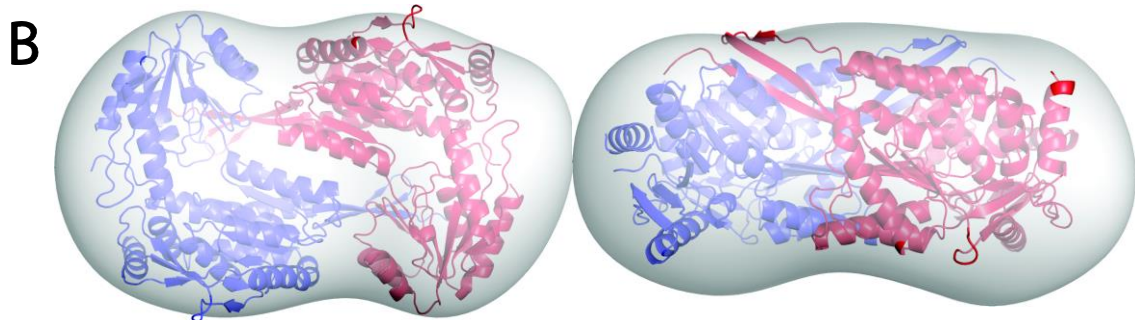
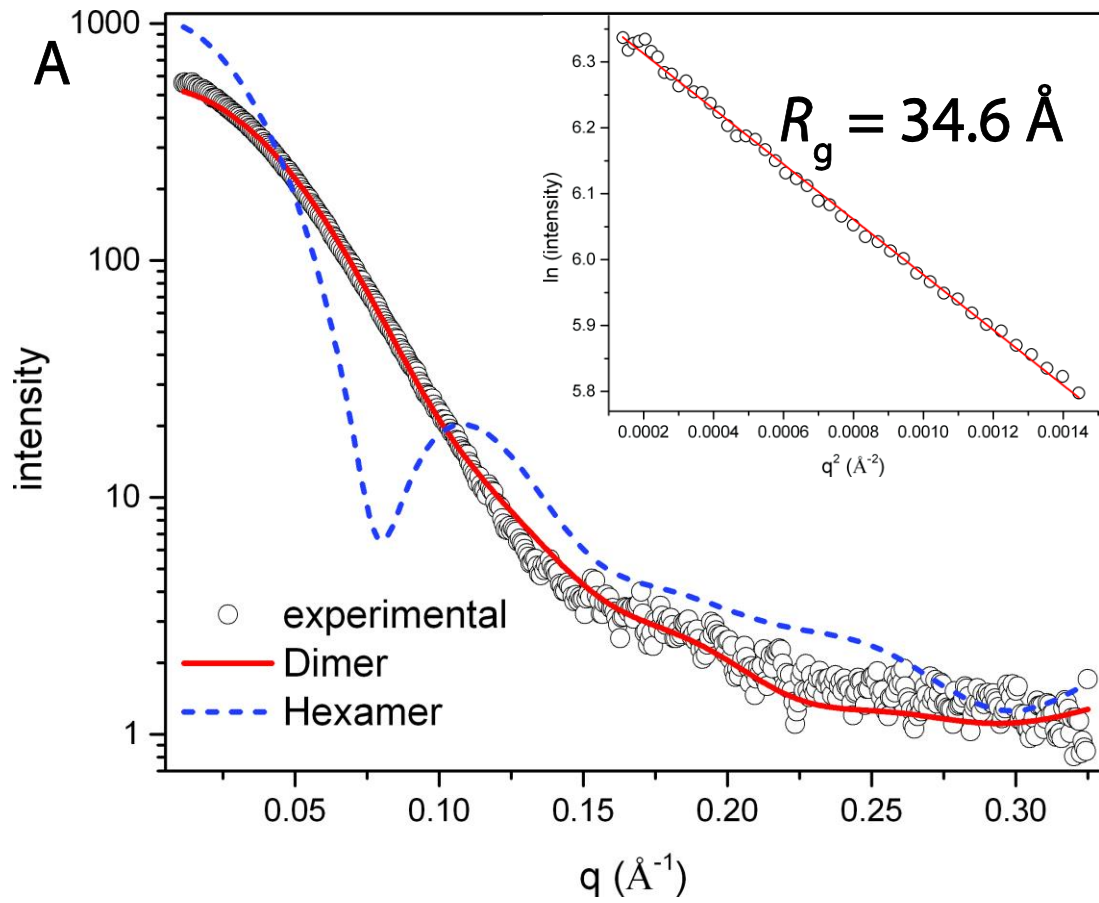


Fig 4.8.



Chapter 5:

Kinetic and Structural Characterization of Tunnel-Perturbing Mutants in *Bradyrhizobium japonicum* Proline Utilization A (PutA)

Abstract

Proline utilization A (PutA) from *Bradyrhizobium japonicum* (BjPutA) is a bifunctional flavoenzyme that catalyzes the oxidation of proline to glutamate using fused proline dehydrogenase (PRODH) and Δ^1 -pyrroline-5-carboxylate dehydrogenase (P5CDH) domains. Recent crystal structures and kinetic data suggest an intramolecular channel connects the two active sites, promoting substrate channeling of the intermediate Δ^1 -pyrroline-5-carboxylate/glutamate- γ -semialdehyde (P5C/GSA). In this work, the structure of the channel was explored by inserting large side chain residues at four positions along the channel in BjPutA. Kinetic analysis of the different mutants revealed replacement of D779 with Tyr (D779Y) or Trp (D779W) significantly decreased the overall rate of the PRODH-P5CDH channeling reaction. X-ray crystal structures of D779Y and D779W revealed that the large side chains caused a constriction in the central section of the tunnel thus likely impeding travel of P5C/GSA in the channel. The D779Y and D779W mutants have similar PRODH activity relative to wild-type BjPutA but exhibit significantly lower P5CDH activity, suggesting that exogenous P5C/GSA enters the channel upstream of Asp779. Replacement of nearby Asp778 with Tyr (D778Y) did not impact BjPutA channeling activity. Consistent with the kinetic results, the X-ray crystal structure of D778Y shows that the main channel pathway is not impacted, however, an off-cavity pathway is closed off from the channel. These findings provide evidence that the off-cavity pathway is not essential for substrate channeling in BjPutA. Structural analysis also shows that the tunnel constrains GSA in a linear conformation with the aldehyde directed toward the P5CDH active site.

5.1. Introduction

The proline catabolic pathway catalyzes the oxidation of proline to glutamate (Scheme 1). In the first step, proline dehydrogenase (PRODH) uses an FAD cofactor to remove two electrons (as H^+) from proline, resulting in Δ^1 -pyrroline-5-carboxylate (P5C). P5C then undergoes a non-enzymatic hydrolysis, which opens the pyrroline ring to create glutamate- γ -semialdehyde (GSA). Finally, GSA is oxidized to glutamate by the NAD^+ -dependent P5C dehydrogenase (P5CDH) to complete the overall 4-electron oxidation process. Proline and proline metabolism are important for the pathogenicity of *Helicobacter pylori* and *H. hepaticus* (1, 2), energy production in procyclic trypanosomes (3, 4), and regulation of metabolites linked to pathogenesis in *Photothabdus* and *Xenorhabdus* (5). In humans, inborn errors in proline catabolism lead to hyperprolinemia disorders, and defects in PRODH are linked to schizophrenia (6, 7). Also, PRODH is regulated by p53 and has been shown to function in tumor suppression (8).

PRODH and P5CDH are combined into a single polypeptide chain known as proline utilization A (PutA) in Gram-negative bacteria and *Corynebacterium* (9). The covalent linking of enzymes catalyzing consecutive reactions in a metabolic pathway affords the possibility of substrate channeling, i.e., the intermediate is transferred between the enzymes without equilibrating with the bulk medium. Several physiological benefits of substrate channeling versus free diffusion have been identified. For example, channeling improves kinetic efficiency by decreasing the transit time between active sites and preventing the loss of intermediates (10, 11). Thus, channeling enzymes can operate at maximum rates when cellular substrate concentrations are below saturating levels (12). Also, labile intermediates can be concealed from the bulk environment, preventing decay

or interaction with other molecules (13, 14). Finally, channeling can influence metabolic flux by segregating intermediates from competing pathways (15).

Substrate channeling of P5C/GSA in proline catabolism may be necessary to retain proper metabolic flux and avoid metabolic futile cycling (14). In addition, free P5C/GSA is an inhibitor of three different enzymes in *Escherichia coli* including glucosamine-6-phosphate synthase, cytidine 5'-triphosphate synthase, and the amidotransferase domain of carbamoyl phosphate synthetase (16-18). P5C has also been shown to form adducts with other metabolites such as oxaloacetic acid, pyruvic acid, and acetoacetic acid (19). Consistent with the physiological importance of controlling P5C/GSA release, kinetic studies have firmly established substrate channeling in PutAs. Early studies of *Salmonella typhimurium* PutA using ¹⁴C-labeled proline are consistent with a channeling mechanism (20). More recent steady-state and rapid reaction transient time measurements of PutAs from *Bradyrhizobium japonicum* (BjPutA) and *Geobacter sulfurreducens* (GsPutA) also indicate substrate channeling (21, 22). Furthermore, a comprehensive analysis of the complete kinetic mechanism of *E. coli* PutA showed that substrate channeling is rate limiting, and the rate constant for the channeling step is slowest during the first enzyme turnover and increases with subsequent turnovers, establishing PutA as a new example of a hysteretic enzyme (23).

With the kinetic data firmly demonstrating substrate channeling in PutA, the goal of this study is to gain insight into the structural basis of channeling. The crystal structures of BjPutA and GsPutA revealed that the two active sites are separated by a linear distance of 41 - 45 Å, implying that substrate channeling involves substantial movement of the

intermediate P5C/GSA (21, 22). Analysis of potential channeling pathways predicts a curved, 75-Å long tunnel that connects the two active sites (Figure 1).

Here we use site-directed mutagenesis, kinetics, and X-ray crystallography to gain further insight into the structural features that facilitate substrate channeling in BjPutA. Several residues between the two active sites have been mutated in an effort to obstruct molecular traffic. Kinetic and structural analysis of the mutant enzymes shows that channeling is hindered in some of the variants but not others, which provides information about the pathway traversed by the intermediate. Furthermore, steric considerations suggest that GSA is threaded through the tunnel in a linear conformation, with the aldehyde group facing the P5CDH end of the tunnel. This aspect of substrate channeling in PutA might be considered an example of shape selective catalysis.

5.2. Experimental Procedures

5.2.1 Chemicals

All chemicals were purchased from Sigma-Aldrich or Fisher Scientific unless otherwise noted. (DL)-P5C (50/50 mixture) was synthesized according to the method of Williams and Frank and stored in 1 M HCl at 4 °C. The concentration of (DL)-P5C was determined as previously reported (24, 25). *E. coli* strain BL21 (DE3) pLysS was purchased from Novagen, and DH5α strain was purchased from Invitrogen. All experiments used Nanopure water.

5.2.2 Site-Directed Mutagenesis

Mutagenic primers (Table 1) were purchased from Integrated DNA Technologies or Eurofins MWG Operon. GeneTailor Mutagenesis Kit (Invitrogen) was used to generate all mutants except T348Y and D779Y (QuikChange II kit, Agilent Technologies).

Mutant plasmids were transformed into DH5 α cells, and the resulting plasmids were sequenced by Eurofins MWG Operon to confirm the mutations.

5.2.3 Expression and Purification of BjPutA Proteins

BjPutA wild-type and mutant proteins were expressed as reported previously, except that induction with isopropyl β -D-1-thiogalactopyranoside was performed at 20 °C for 16 h (26). Cells were harvested by centrifugation and frozen at -80°C. Frozen cells were resuspended in 50 ml binding buffer (20 mM Tris-base, 0.5 M NaCl, 5 mM imidazole, 10% glycerol, pH 7.9) and 100 μ M flavin at 4°C. Protease inhibitors ϵ -amino-N-caproic acid (3 mM), phenylmethylsulfonyl fluoride (0.3 mM), leupeptin (1.2 μ M), tosyl phenylalanyl chloromethyl ketone (48 μ M), and tosyllysine chloromethyl ketone hydrochloride (78 μ M) were added, and cells were disrupted via sonication. The cell lysate was centrifuged for 1 h at 19,000 rpm in a JA-20 rotor (Beckman) and filtered through a 0.2 μ m filter (VWR). Cell-free lysate was loaded onto a Ni-NTA Superflow resin (Qiagen) equilibrated with binding buffer. Wash buffer (60 mM imidazole) followed by elution buffer (500 mM imidazole) were applied to the column. Elution fractions containing PutA protein were pooled and dialyzed into buffer containing 50 mM Tris (pH 7.5), 10 mM NaCl, 0.5 mM EDTA, and 10% glycerol and loaded onto an anion exchange column (HiTrap Q HP column, GE Life Sciences) equilibrated with dialysis buffer. BjPutA proteins were eluted using a linear 0-1 M NaCl gradient (1 L) in dialysis buffer. Purified enzyme was then dialyzed into a final buffer of 50 mM Tris (pH 7.5), 50 mM NaCl, 0.5 mM EDTA, 0.5 mM Tris(3-hydroxypropyl)phosphine, and 10 % glycerol. The His-tag was retained in the subsequent kinetic experiments. The amount of flavin bound in the purified proteins was quantified as described previously ($\epsilon_{451\text{nm}} = 13.62 \text{ mM}^{-1}$

$^1 \text{ cm}^{-1}$ for bound flavin) (26). Protein concentration was determined from the amount of bound flavin to normalize for differences in flavin content, and the protein was flash-frozen in liquid nitrogen and stored at $-80 \text{ }^\circ\text{C}$.

5.2.4 Steady-State Kinetic Assays

Steady-state kinetic assays were performed at $23 \text{ }^\circ\text{C}$. Kinetic parameters for the PRODH domain were determined for proline and ubiquinone-1 (CoQ₁) by following reduction of CoQ₁ at 278 nm ($\epsilon_{278\text{nm}} = 14.5 \text{ mM}^{-1} \text{ cm}^{-1}$) (Table 2) (27). All assays were performed in 50 mM potassium phosphate buffer (pH 7.5) with 0.5 μM PutA enzyme. The K_m and k_{cat} values for proline were determined by varying proline (1-200 mM) while holding CoQ₁ constant (250 μM) and CoQ₁ kinetic parameters were determined by varying CoQ₁ (10-350 μM) while holding proline fixed at 150 mM. Data were collected on a Hi-Tech Scientific SF-61DX2 stopped-flow instrument using a 0.15 cm pathlength. Initial velocities were fit to the Michaelis-Menten equation using SigmaPlot 12.0.

Kinetic parameters of P5CDH activity were determined for P5C/GSA (Table 3) using exogenous (DL)-P5C and 0.25 μM PutA enzyme. (DL)-P5C was neutralized with 10 M NaOH immediately prior to assays. The concentration of L-P5C is considered to be half the total (DL)-P5C concentration. To maintain a constant ionic strength, all assays with exogenous (DL)-P5C were performed in 600 mM NaCl as described previously (23). K_m and k_{cat} for P5C/GSA were determined by varying L-P5C (0.01 - 6 mM) while holding NAD^+ constant at 0.2 mM in 50 mM potassium phosphate (pH 7.5, 600 mM NaCl). The effective concentration of GSA was estimated from the pH dependence of the P5C-GSA equilibrium reported previously (16). Initial velocities were determined by following

NAD⁺ reduction at 340 nm ($\epsilon_{340\text{nm}} = 6200 \text{ M}^{-1} \text{ cm}^{-1}$). All assays were performed using a Powerwave XS 96 well plate reader (Biotek).

P5CDH activity of BjPutA wild-type and mutants D779A/Y/W using smaller aldehyde substrates relative to GSA was tested. All assays were performed in 50 mM potassium phosphate (pH 7.5, 25 mM NaCl) containing 0.2 mM NAD⁺ and variable concentrations of succinate semialdehyde (0.05-20 mM) and propionaldehyde (5-500 mM). For assays with succinate semialdehyde, the concentrations of BjPutA wild-type and mutant D779A were 0.25 μM while mutants D779W and D779Y were at 1 μM . For propionaldehyde, the concentrations of BjPutA wild-type and mutant D779A were 0.25 μM , D779W was 1 μM , and D779Y was 2 μM . Initial velocities were determined by following NAD⁺ reduction at 340 nm. All assays were performed using a Powerwave XS 96 well plate reader (Biotek).

The coupled PRODH-P5CDH activity of BjPutA wild-type and mutants was monitored by following NADH formation at 340 nm or by fluorescence excitation at 340 nm and monitoring fluorescence emission at 460 nm using a Cary Eclipse fluorescence spectrophotometer. Assays were performed at 23°C in 50 mM potassium phosphate buffer (pH 7.5, 25 mM NaCl, 10 mM MgCl₂) containing 0.1 mM CoQ₁, 0.2 mM NAD⁺, 40 mM proline, and BjPutA enzyme as previously described (28). To determine the kinetic parameters of the overall PRODH-P5CDH reaction for BjPutA wild-type and the D778Y mutant, assays were performed by varying proline (0.1-550 mM) and holding CoQ₁ (0.1 mM) and NAD⁺ (0.2 mM) fixed in 50 mM potassium phosphate buffer (pH 7.5, 25 mM NaCl). NADH formation was followed at 340 nm. Data were fitted to a

substrate inhibition equation (eq 1) using SigmaPlot 12.0, where [S] is the substrate concentration and K_i is the substrate inhibition constant.

$$v = \frac{V_{max}[S]}{K_m + [S] + \left(\frac{[S]^2}{K_i}\right)} \quad (1)$$

5.2.5 NAD⁺ binding to BjPutA

The binding of NAD⁺ to BjPutA was determined by intrinsic tryptophan fluorescence quenching of BjPutA as described (23). Tryptophan fluorescence was excited at 295 nm and fluorescence emission was recorded at 330 nm. Increasing concentrations of NAD⁺ (0-20 μ M) were added to BjPutA (0.25 μ M) in 50 mM potassium phosphate (pH 7.5). The inner filter effect caused by the absorption of incident light by NAD⁺ at 295 nm was corrected using eq 2. (27) F_{corr} and F_{obs} are the corrected and observed fluorescence, respectively, and A_{ex} and A_{em} are the absorbance values of NAD⁺ at the excitation and emission wavelengths, respectively.

$$F_{corr} = F_{obs} 10^{\left(\frac{A_{ex} + A_{em}}{2}\right)} \quad (2)$$

A dissociation constant (K_d) for the BjPutA-NAD⁺ complex was determined by plotting the fraction of BjPutA bound by NAD⁺ (θ) versus free [NAD⁺] using eq 3, where n is number of binding sites.

$$\theta = \frac{n[\text{NAD}^+]_{free}}{K_d + [\text{NAD}^+]_{free}} \quad (3)$$

The concentration of free NAD⁺ was determined using eq 4.

$$[\text{NAD}^+]_{free} = [\text{NAD}^+]_{total} - \theta[\text{BjPutA}]_{total} \quad (4)$$

The value of θ is obtained from the fluorescence measurements $\frac{(F_0 - F)}{(F_0 - F_{max})}$, where F_0 is fluorescence intensity without NAD⁺, F is fluorescence intensity in the presence of

NAD^+ , and F_{max} is the maximum fluorescence intensity under saturating concentrations of NAD^+ .

NAD binding to wild-type BjPutA was also estimated by isothermal titration calorimetry (ITC). Titrations were performed at 4°C using a MicroCal VP-ITC microcalorimeter. Wild-type BjPutA was dialyzed into a buffer composed of 50 mM Tris (pH 7.5), 50 mM NaCl, 0.5 mM EDTA, and 10% glycerol. A NAD^+ stock solution of 0.5 mM was made in dialysis buffer. For each titration, $23.4\ \mu\text{M}$ BjPutA was titrated with 2 μl injections (40 total) of 0.5 mM NAD^+ at 160 sec intervals, while stirring at 310 revolutions per min (rpm). Data were analyzed using a one-site binding model with Origin ITC Analysis software provided with the instrument.

Prior to performing the above assays, the amount of NAD^+ bound to purified BjPutA was estimated by HPLC. BjPutA was denatured with 5% vol/vol trichloroacetic acid and centrifuged at 13,000 rpm for 5 min to release bound FAD and NAD^+ cofactors. Samples were then filtered with a 0.45 micron filter before being loaded onto the column. FAD and NAD^+ were separated on a C18 column using 50 mM potassium phosphate (pH 5.3) and 100% methanol. The cofactors were eluted using a flow rate of 1 ml/min with 5 min of isocratic phosphate buffer, followed by a 25 min linear gradient to 50% methanol, and finally a 5 min linear gradient to 75% methanol. Both cofactors were detected at 280 nm. NAD^+ and FAD eluted from the column at 7.9 and 16.6 min, respectively. The concentration of NAD^+ was determined using standard solutions of NAD^+ (10, 25, 50, 100, and 200 μM). From this analysis, it was estimated that 74% of purified BjPutA contained bound NAD^+ . Thus, the NAD^+ binding experiments report on the remaining 26% of BjPutA that purified without NAD^+ bound.

5.2.6 Single-Turnover Kinetic Experiments

Single-turnover experiments were performed at 21 °C under anaerobic conditions as described previously (28). Briefly, equal volumes of BjPutA enzyme (21.3 μM wild-type and 17.9 μM D779Y) were pre-incubated with 0.1 mM NAD⁺ in 50 mM potassium phosphate (pH 7.5, 25 mM NaCl) and rapidly mixed with 40 mM proline in 50 mM potassium phosphate buffer (pH 7.5, 25 mM NaCl) (all concentrations reported as final concentrations after mixing) (29). Anaerobic conditions were achieved by degassing buffer, substrate, and enzyme solutions by performing repeated vacuum/nitrogen cycles followed by addition of protocatechuate dioxygenase (PCD) (0.05 U/mL) and protocatechuic acid (PCA) (100 μM), which scrub dissolved oxygen. All enzyme manipulations were performed in an anaerobic glovebox (Belle Technology) prior to the experiments. Rapid-reaction experiments were performed with a Hi-Tech Scientific SF-61DX2 stopped flow instrument equipped with a photodiode array detector. The stopped-flow mixing cell and tubing were thoroughly washed and incubated overnight with PCA/PCD buffer before loading stopped-flow syringes with anaerobic substrate and enzyme solutions. Multiwavelength data (300-700 nm) were recorded and single wavelength traces of FAD (451 nm) and NAD⁺ (340 nm) were extracted and fitted to a single exponential equation to estimate observed rate constants for FAD and NAD⁺ reduction as previously reported (28).

5.2.7 Crystal Structure Determination and Structure Analysis

BjPutA wild-type and mutants were expressed, purified, and crystallized as described previously for wild-type BjPutA (30). Briefly, crystals were grown in sitting drops at room temperature in the presence of ~ 2 M ammonium sulfate and cryoprotected with

glycerol. For some of the mutants, microseeding was used with a seed stock made initially by crushing crystals of the wild-type enzyme. Seed stocks made from crystals of the mutant enzymes were used in subsequent rounds of crystallization trials. The space group is *C2* with a BjPutA dimer in the asymmetric unit. X-ray diffraction data sets were collected at beamline 4.2.2 of the Advanced Light Source using a NOIR-1 detector. The data were integrated with MOSFLM(31) and scaled with SCALA (32). Refinements in PHENIX (33) were initiated from models derived from the structure of wild-type BjPutA (Protein Data Bank (PDB) entry 3HAZ). COOT (34) was used for model building. The structures were validated with MolProbity (35) and the PDB (36) validation server. Data collection and refinement statistics are listed in Table 5.

The substrate-channeling cavity/tunnel system was analyzed and visualized with VOIDOO (37), which characterizes cavities, and MOLE (38, 39), which finds tunnels that connect cavities to the bulk medium. Hydrogen atoms were added to the protein with the WHAT IF web services prior to these calculations (40). VOIDOO was run in probe-occupied mode (option O) with a probe radius of 2.9 Å, which approximates P5C/GSA. This radius was chosen based on molecular volume calculations performed with VOIDOO; P5C and GSA have volumes of 104 and 124 Å³, respectively, which correspond to spheres of radii 2.9 and 3.1 Å. Mole was run with default options and using Arg456 of the PRODH active site as the starting point.

Models of P5C and GSA were built into the cavity/tunnel system to understand the steric relationships and estimate the number of intermediates that the system accommodates. The starting models were downloaded from the NCBI PubChem database (compound identification numbers 193305 (GSA) and 11966181 (P5C)). A

model of P5C bound in the BjPutA PRODH active site was built using the structure of GsPutA complexed with the proline analog L-tetrahydro-2-furoic acid (PDB code 4NMA). A model of GSA bound in the BjPutA P5CDH active site was built using the structure of mouse P5CDH complexed with glutamate (PDB code 3V9K). Models of GSA were fitted manually into the tunnel between the two active sites and the off-pathway cavity.

5.3. Results

5.3.1 Rationale for Channel-Blocking Mutagenesis and Purification of BjPutA

Mutant Enzymes

The BjPutA dimer (PDB entry 3HAZ) was analyzed with the PyMOL plug-in CAVER (41, 42) and MOLE 2.0 to identify residues lining the cavity/tunnel system that, upon mutation to a larger side chain, might eliminate sections of the channeling apparatus. Using starting points in the PRODH site, the programs identified several channels leading to the bulk solvent, including some that connect the two active sites (Figure 1A). (Although the tunnel appears to be open to the bulk medium as shown for the protomer in Figure 1A, we note that it is buried by the dimerization flap of the corresponding protomer in the tetramer that forms in solution.) This tunnel features a prominent central section that runs between and parallel to two helices, the α 5a helix of the PRODH domain (residues 346-356) and the 770s helix of the P5CDH domain (residues 773-785). Side chains of these helices contribute to the walls of the tunnel. The central section is 25 Å in length and 4 – 8 Å in diameter, and can accommodate 2-3 molecules of GSA (Figure 1B).

Analysis with the program VOIDOO also identifies a cavity that is connected to the central section of the predicted tunnel (Figure 1C). This “off-pathway” cavity has a volume of $\sim 700 \text{ \AA}^3$, which is sufficient to accommodate another 2-3 molecules of GSA.

Four residues lining the central section of the tunnel were selected for mutagenesis: Thr348, Ser607, Asp778, and Asp779. Thr348 and Ser607 sit near the beginning and end of the central section, respectively, while Asp778 and Asp779 are closer to the middle of the central section, near the off-pathway cavity (Figure 1B). Each of the targeted residues were mutated to Tyr, which retains polarity while increasing steric bulk. Additionally, Asp779 was mutated to Trp and Ala. The Trp mutation further increases side chain bulk, whereas Ala decreases size and removes the functional property of the side chain carboxylate.

All six BjPutA mutant proteins, T348Y, S607Y, D778Y, and D779Y/W/A were purified and shown to have flavin spectra similar to wild-type BjPutA with flavin peak absorbances at 380 and 451 nm. From the flavin absorbance spectra, the percent bound flavin was estimated to be 74-99% per monomer for the mutants, which is similar to 79% bound flavin for wild-type BjPutA.

5.3.2 Channeling Assays of BjPutA Mutants

The impact of the mutations on channeling was evaluated by measuring coupled PRODH-P5CDH activity. The assay involves monitoring the progress curve of NADH production from proline and determining whether an initial lag phase is apparent in NADH formation (28). As shown in Figure 2, NADH production by wild-type BjPutA does not exhibit a perceptible lag time, which is consistent with channeling. The progress curves of NADH formation with BjPutA mutants T348Y, S607Y, D778Y, and D779A

likewise show no substantial lag phase indicating that substrate channeling is unperturbed in these mutants (Figure 2). The linear rate of NADH formation achieved with these mutants is similar to wild-type ($\sim 1.4 \mu\text{M}/\text{min}$) at the same enzyme concentration ($0.187 \mu\text{M}$). No significant NADH formation, however, was observed with the BjPutA mutants D779Y and D779W (Figure 2).

Mutants D779Y and D779W were then assayed using up to 10-fold higher concentration of enzyme ($1.87 \mu\text{M}$) and fluorescence spectroscopy to detect NADH formation (Figure 3). Increasing D779Y concentration to 10-fold higher than wild-type BjPutA ($0.187 \mu\text{M}$) resulted in a similar rate of NADH formation suggesting that the coupled PRODH-P5CDH activity of D779Y is about 10-fold lower than wild-type BjPutA (Figure 3A). At 10-fold higher D779W, NADH formation remained very slow indicating that the D779W mutant is severely impaired (Figure 3B).

5.3.3 Steady-state Kinetic Properties of BjPutA Wild-type and Mutants

The PRODH and P5CDH kinetic parameters were then determined for BjPutA wild-type and the mutants. The steady-state kinetic parameters of the PRODH domain were determined using proline and CoQ_1 as substrates (Table 2). Similar k_{cat}/K_m values (within \sim two-fold) were found for BjPutA wild-type and all the mutants except for D778Y. D778Y exhibited comparable K_m values for proline (91 mM) and CoQ_1 ($82 \mu\text{M}$), but the k_{cat} value was nearly 9-fold lower than wild-type BjPutA resulting in a significantly lower k_{cat}/K_m . This result was unexpected because D778Y exhibited similar activity to that of wild-type BjPutA in the channeling assays (Figure 2).

The P5CDH kinetic parameters were also determined for BjPutA wild-type and mutants (Table 3). The k_{cat}/K_m values for P5CDH activity in the mutants were similar to

wild-type BjPutA except for mutants D779Y and D779W. The k_{cat}/K_m values for D779Y and D779W were 81-fold and 941-fold lower than wild-type BjPutA. To determine whether perturbations in NAD^+ binding account for the severe loss in P5CDH activity, NAD^+ binding was measured for wild-type BjPutA and the mutants (Table 3). For wild-type BjPutA, dissociation constants (K_d) of 0.6 μM and 1.5 μM were determined by intrinsic tryptophan fluorescence quenching (Figure 4A) and ITC (Figure 4B), respectively. The K_d values of NAD^+ binding to the BjPutA mutants were shown by intrinsic tryptophan fluorescence quenching to be similar to wild-type BjPutA (Table 3). Thus, NAD^+ binding is unchanged in the mutants suggesting that the severe decrease in P5CDH activity of D779Y and D779W is not caused by alterations in the Rossmann fold domain.

5.3.4 Single Turnover Rapid-reaction Kinetics

To further corroborate impaired channeling activity in the D779Y mutant, single turnover experiments were performed anaerobically without an electron acceptor for the flavin cofactor. In this experiment, PutA enzyme and NAD^+ were rapidly mixed with proline and the absorbance spectrum was recorded (Figure 5). Observed rate constants for FAD reduction and NADH formation were estimated by single exponential fits of absorbance changes at 451 nm and 340 nm, respectively. The observed rate constant for FAD reduction was faster for the BjPutA mutant D779Y (0.46 s^{-1}) than wild-type BjPutA (0.18 s^{-1}). In contrast, the observed rate constant for NADH formation is 10-fold slower in D779Y (0.003 s^{-1}) relative to wild-type BjPutA (0.03 s^{-1}) consistent with severely impaired P5CDH activity.

5.3.5 Alternative P5CDH Substrates

The potential tunnel constriction in the mutants D779Y/W was explored by measuring P5CDH activity with smaller aldehyde substrates. Table 4 shows the kinetic parameters of BjPutA wild-type and mutants D779A/Y/W with exogenous GSA/P5C and smaller substrates succinate semialdehyde and propionaldehyde. Succinate semialdehyde contains one less carbon and no amino group, whereas propionaldehyde is a three-carbon aldehyde. The $k_{\text{cat}}/K_{\text{m}}$ values were significantly lower for each enzyme using the smaller substrates (Table 4). To assess whether succinate semialdehyde and propionaldehyde are more effective substrates in the mutants relative to P5C/GSA, the $k_{\text{cat}}/K_{\text{m}}$ ratio of wild-type BjPutA and each mutant ($k_{\text{cat}}/K_{\text{m}}(\text{WT})/k_{\text{cat}}/K_{\text{m}}(\text{mut})$) was determined for all the substrates. For D779A, the $k_{\text{cat}}/K_{\text{m}}(\text{WT})/k_{\text{cat}}/K_{\text{m}}(\text{mut})$ ratio remained ~ 1 with each substrate. For the mutants D779Y and D779W, the ratios of $k_{\text{cat}}/K_{\text{m}}(\text{WT})/k_{\text{cat}}/K_{\text{m}}(\text{mut})$ were 81 and 941, respectively, with P5C/GSA. The $k_{\text{cat}}/K_{\text{m}}(\text{WT})/k_{\text{cat}}/K_{\text{m}}(\text{mut})$ ratios decreased to 30 (D779Y) and 38 (D779W) with succinate semialdehyde, suggesting that relative to GSA/P5C this smaller substrate more readily accesses the P5CDH active site in the mutants D779Y and D779W. A further decrease in the $k_{\text{cat}}/K_{\text{m}}(\text{WT})/k_{\text{cat}}/K_{\text{m}}(\text{mut})$ ratio, however, was not observed with propionaldehyde.

5.3.6 Crystal structures of D778Y, D779Y, and D779W

The structures of D778Y, D779Y, and D779W were determined at 2.2 – 2.3 Å resolution (Table 5). The electron density features representing the mutated side chains are strong in all three mutant enzymes (Figure 6A, B, C). The mutations induce rotations of neighboring side chains but otherwise have minimal impact on the protein structure (Figure 6D). In the wild-type enzyme structure, Asp778 and Arg200 are within 2.8 Å of

each other and form an ion pair; the mutation of Asp778 to the larger Tyr would result in steric clash in the absence of conformational changes. Clash is avoided because Tyr778 has rotated by 100° around χ_1 relative to Asp778 of the wild-type enzyme. This movement is accompanied by rotation of Arg200 into the space occupied by the carboxylate of Asp778 in the wild-type enzyme. In contrast to D778Y, mutation of Asp779 to Tyr or Trp does not change χ_1 . Nevertheless, these mutations cause rotations of His919 and Gln775 to avoid steric clash with the new, bulkier side chain at residue 779 (Figure 6D). Aside from these local perturbations, no other significant structural changes are evident. In particular, the active site structures are essentially unchanged.

Mutation of Asp778 to Tyr substantially changes the off-pathway cavity located near the central section of the predicted channeling pathway. Asp778 borders this cavity in wild-type B_jPutA (Figure 1C). Because of the aforementioned 100° rotation around χ_1 , the phenol ring of Tyr778 invades the space corresponding to the off-pathway cavity of the wild-type enzyme (Figure 7). The presence of Tyr778 in this region reduces the volume of the cavity by 70 % to 200 Å³, so that just a residual cavity remains (Figure 7, blue surface). Furthermore, the close approach of Tyr778 to Arg356 severs the connection between the cavity and the predicted channeling tunnel (using a 2.9 Å probe). Thus, the structure suggests that P5C/GSA molecules that are moving through the tunnel of D778Y cannot enter the off-pathway cavity.

In contrast to D778Y, the mutation of Asp779 to Tyr constricts the predicted channeling tunnel without affecting the off-cavity pathway (Figure 8). The side chain of Tyr779 pokes into the space corresponding to the central section of the tunnel in the wild-type enzyme (Figure 8A). As a result, the predicted tunnel of D779Y has a 2.0-Å

invagination near the phenol hydroxyl (Figure 8B). This narrowing of the tunnel reflects a decrease in distance between the 770s and α 5a helices. In particular, the distance between the side chains of residue 779 and Lys351 decreases from 9.3 Å in the wild-type enzyme to only 6.8 Å in D779Y. Thus, the gap between these side chains decreases by about 2.5 Å, which accounts for the invagination of the tunnel near Tyr779.

The mutation of Asp779 to Trp similarly reshapes the predicted channeling tunnel (Figure 9). As in D779Y, the bulky side chain of Trp779 penetrates the space corresponding to the tunnel in the wild-type enzyme (Figure 9A). Also, Gln775, which has rotated relative to the wild-type enzyme, protrudes into the tunnel just upstream from Trp779. The invasion of these residues into the tunnel reshapes the predicted channeling pathway, essentially shaving a 2 Å slice off one side of the tunnel (Figure 9B).

5.4. Discussion

Introducing residues with bulkier side chains in a predicted channeling path is a useful approach for validating substrate channeling and exploring the structural architecture of an interconnecting path between active sites. In tryptophan synthase, substitution of β Cys170 with Trp in the tunnel pathway significantly hindered passage of the indole intermediate between active sites and also impacted communication between subunits (43). In the bifunctional enzyme dethiobiotin synthetase (DTBS)-diaminopelargonic acid aminotransferase (DAPAT-AT) from *Arabidopsis*, two mutations were made in a crevice on the surface connecting the two active sites (44). The surface crevice was proposed to be a channel pathway for movement of the intermediate from DAPA-AT to DTBS. Mutation of two crevice residues, Ser360 to Tyr and Ile793 to Trp resulted in long lag times (~ 10-12 min) for product formation, whereas no lag phase was

observed with the wild-type enzyme. These results were consistent with the predicted function of the crevice as a channeling path.

Here, we substituted four residues at different points along the predicted channeling path in BjPutA with bulkier side chains. Although Thr348 and Ser607 are located at apparent bottleneck regions and Asp778 points toward the middle of the channel, Tyr substitutions of these residues did not impact PRODH-P5CDH channeling activity in BjPutA. Only replacement of Asp779 with Tyr or Trp disrupted coupled PRODH-P5CDH activity. Substitution of Asp779 with Ala did not diminish channeling, indicating that the carboxylate group of Asp779 is not critical for channel function. The decrease in substrate channeling activity of the D779Y/W mutants correlates with a significant drop in P5CDH activity whereas the PRODH activity of the mutants is similar to wild-type BjPutA. The X-ray crystal structures of the D779Y/W mutants show that the PRODH and P5CDH domains are essentially unchanged from wild-type BjPutA. The only structural perturbations are in the side chain conformations of residues near Asp779. Thus, the severely impaired substrate channeling and P5CDH activities of the D779Y/W mutants are likely caused by local effects of substituting a larger side chain in the channel. Replacing Asp779 with Tyr decreased the internal width of the predicted channeling path between the 770s and $\alpha 5a$ helices by 2.5 Å or ~ 25%. In D779W, the Trp residue curves into the channel by 2.0 Å. These changes result in a narrowing of the tunnel that is sufficient to disrupt substrate channeling and illustrates that the channel structure is finely tuned for transporting P5C/GSA. The results with D779Y/W also validate the tunnel in BjPutA identified by X-ray crystallography as the path for channeling the P5C/GSA intermediate.

An outstanding question in PutA enzymes is how P5C/GSA accesses the P5CDH active site. Because the X-ray crystal structures of D779Y/W show no changes in the P5CDH active site relative to wild-type BjPutA, the significantly lower P5CDH activity of the D779Y/W mutants indicates exogenous P5C enters the tunnel upstream of Asp779 possibly via the PRODH active site. If P5C/GSA were able to enter the P5CDH active site from a point downstream of Asp779, the P5CDH activity of the D779Y/W mutants would be expected to be similar to the wild-type enzyme. These results indicate that exogenous P5C/GSA must access the P5CDH domain through the channel. The kinetic results using smaller aldehydes as exogenous substrates are consistent with this interpretation. Although the activity of D779W with succinate semialdehyde is still lower than wild-type BjPutA, the difference in k_{cat}/K_m between wild-type BjPutA and D779W is reduced by 25-fold relative to GSA.

Even though it neighbors Asp779, replacing Asp778 with Tyr did not diminish the substrate channeling and P5CDH activities of BjPutA. Similar to the D779Y/W mutants, the X-ray crystal structure of D778Y shows no changes in the PRODH and P5CDH domains as only perturbations in local residues of the channel were observed. Introducing a bulkier side chain at Asp778 appears to close the off-pathway cavity from the main channeling path. The coupled PRODH-P5CDH activity of the D778Y mutant is similar to wild-type BjPutA demonstrating that the off-pathway cavity is not required for substrate channeling. The function of the off-cavity pathway in substrate channeling thus remains unknown. An interesting finding with the D778Y mutant was its significantly lower PRODH activity. This result may provide additional evidence of a communication link between the PRODH domain and the channel. Recently we have shown in PutA from *E.*

coli that a substrate channeling step becomes activated during enzyme turnover thereby increasing overall PRODH-P5CDH activity by nearly 40-fold (23). PutA also undergoes a conformational change upon flavin reduction, with a conserved ion pair (Arg456-Glu197) proposed to act as a gate between the PRODH domain and the main channeling pathway (28, 45). Residues that are critical for communication between the PRODH domain and the channel are unknown, but the findings with D778Y suggest that the 770s helix may be involved. Despite having 9-fold lower PRODH activity, D778Y exhibited similar substrate channeling activity to that of wild-type BjPutA consistent with the rate of the coupled PRODH-P5CDH reaction being limited by a channeling step as found previously for *E. coli* PutA (23).

Structural analysis of the channeling path in BjPutA provides new insight into how P5C/GSA is shuttled between the PRODH and P5CDH active sites. Our results suggest that the off-pathway cavity is dispensable for channeling, which implies that the intermediate is constrained to travel through the cylindrical middle section of the tunnel that runs parallel to the $\alpha 5a$ and 770s helices (Figure 1B). The dimensions of this section are consistent with a maximum of 2-3 intermediates simultaneously occupying the middle section. Furthermore, since the tunnel diameter is similar to the length scales of P5C and GSA, rotational and torsional motions of the intermediates are constrained. In particular, it is unlikely that P5C or GSA can flip orientation while in the tunnel, and torsional motion of GSA is probably restricted. Thus, if the hydrolysis reaction occurs upstream of the P5CDH active site, GSA likely travels through the tunnel with the aldehyde group directed toward the P5CDH active site, as shown in Figure 1B.

Potentially the amino and carboxylic groups of GSA may have a critical role in properly directing its movement and orientation in the tunnel.

ASSOCIATED CONTENT

Accession Codes

Atomic coordinates and structure factors have been deposited in the Protein Data Bank as entries 4Q71 (D779W), 4Q72 (D779Y), and 4Q73 (D778Y).

AUTHOR INFORMATION

Corresponding Author

* Tel: (402) 472-9652. Fax: (402) 472-7842. E-mail: dbecker3@unl.edu.

Funding

† Research reported in this publication was supported by NIH Grants GM065546 and P30GM103335 and is a contribution of the University of Nebraska Agricultural Research Division, supported in part by funds provided by the Hatch Act.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

We thank Dr. Jay Nix of beamline 4.2.2 for help with data collection and processing. Part of this work was conducted at the Advanced Light Source, which is supported by the

Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

ABBREVIATIONS

CoQ₁, ubiquinone-1; D778Y, site-directed mutant of BjPutA in which Asp778 is replaced with Tyr; D779A, D779Y, and D779W, site-directed mutants of BjPutA in which Asp779 is replaced with Ala, Tyr, and Trp, respectively; S607Y, site-directed mutant of BjPutA in which Ser607 is replaced with Tyr; T348Y, site-directed mutant of BjPutA in which Thr348 is replaced with Tyr; BjPutA, proline utilization A from *Bradyrhizobium japonicum*; FAD, flavin adenine dinucleotide; GSA, glutamate- γ -semialdehyde; PRODH, proline dehydrogenase; PCD, protocatechuate dioxygenase; PCA, protocatechuic acid; P5C, Δ^1 -pyrroline-5-carboxylate; P5CDH, Δ^1 -pyrroline-5-carboxylate dehydrogenase; PutA, proline utilization A; ITC, isothermal titration calorimetry.

References

1. Nakajima, K., Inatsu, S., Mizote, T., Nagata, Y., Aoyama, K., Fukuda, Y., and Nagata, K. (2008) Possible involvement of *putA* gene in *Helicobacter pylori* colonization in the stomach and motility, *Biomed. Res.* 29, 9-18.
2. Krishnan, N., Doster, A. R., Duhamel, G. E., and Becker, D. F. (2008) Characterization of a *Helicobacter hepaticus putA* mutant strain in host colonization and oxidative stress, *Infect. Immun.* 76, 3037-3044.
3. van Weelden, S. W., Fast, B., Vogt, A., van der Meer, P., Saas, J., van Hellemond, J. J., Tielens, A. G., and Boshart, M. (2003) Procylic *Trypanosoma brucei* do not use Krebs cycle activity for energy generation, *J. Biol. Chem.* 278, 12854-12863.
4. Bringaud, F., Riviere, L., and Coustou, V. (2006) Energy metabolism of trypanosomatids: adaptation to available carbon sources, *Mol. Biochem. Parasitol.* 149, 1-9.

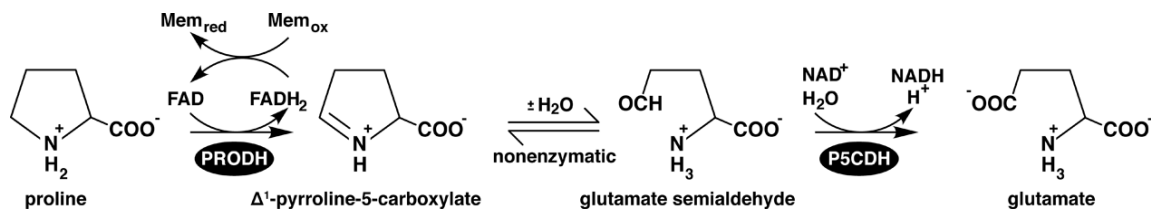
5. Crawford, J. M., Kontnik, R., and Clardy, J. (2010) Regulating alternative lifestyles in entomopathogenic bacteria, *Curr. Biol.* *20*, 69-74.
6. Willis, A., Bender, H. U., Steel, G., and Valle, D. (2008) PRODH variants and risk for schizophrenia, *Amino Acids* *35*, 673-679.
7. Chakravarti, A. (2002) A compelling genetic hypothesis for a complex disease: PRODH2/DGCR6 variation leads to schizophrenia susceptibility, *Proc. Natl. Acad. Sci. U. S. A.* *99*, 4755-4756.
8. Phang, J. M., Donald, S. P., Pandhare, J., and Liu, Y. (2008) The metabolism of proline, a stress substrate, modulates carcinogenic pathways, *Amino Acids* *35*, 681-690.
9. Tanner, J. J., and Becker, D. F. (2013) PutA and proline metabolism, In *Handbook of Flavoproteins* (Hille, R., Miller, S. M., and Palfey, B., Eds.), pp 31-56, Walter de Gruyter, Boston.
10. Ovadi, J. (1991) Physiological significance of metabolic channelling, *J. Theor. Biol.* *152*, 1-22.
11. Easterby, J. S. (1981) A generalized theory of the transition time for sequential enzyme reactions, *Biochem. J.* *199*, 155-161.
12. Rudolph, J., and Stubbe, J. (1995) Investigation of the mechanism of phosphoribosylamine transfer from glutamine phosphoribosylpyrophosphate amidotransferase to glycylamide ribonucleotide synthetase, *Biochemistry* *34*, 2241-2250.
13. Huang, X., Holden, H. M., and Raushel, F. M. (2001) Channeling of substrates and intermediates in enzyme-catalyzed reactions, *Annu. Rev. Biochem.* *70*, 149-180.
14. Arentson, B. W., Sanyal, N., and Becker, D. F. (2012) Substrate channeling in proline metabolism, *Front Biosci.* *17*, 375-388.
15. Anderson, K. S. (1999) Fundamental mechanisms of substrate channeling, *Methods Enzymol.* *308*, 111-145.
16. Bearne, S. L., and Wolfenden, R. (1995) Glutamate gamma-semialdehyde as a natural transition state analogue inhibitor of *Escherichia coli* glucosamine-6-phosphate synthase, *Biochemistry* *34*, 11515-11520.
17. Bearne, S. L., Hekmat, O., and Macdonnell, J. E. (2001) Inhibition of *Escherichia coli* CTP synthase by glutamate gamma-semialdehyde and the role of the allosteric effector GTP in glutamine hydrolysis, *Biochem. J.* *356*, 223-232.
18. Thoden, J. B., Huang, X., Raushel, F. M., and Holden, H. M. (1999) The small subunit of carbamoyl phosphate synthetase: snapshots along the reaction pathway, *Biochemistry* *38*, 16158-16166.

19. Farrant, R. D., Walker, V., Mills, G. A., Mellor, J. M., and Langley, G. J. (2001) Pyridoxal phosphate de-activation by pyrroline-5-carboxylic acid. Increased risk of vitamin B6 deficiency and seizures in hyperprolinemia type II, *J. Biol. Chem.* 276, 15107-15116.
20. Surber, M. W., and Maloy, S. (1998) The PutA protein of *Salmonella typhimurium* catalyzes the two steps of proline degradation via a leaky channel, *Arch. Biochem. Biophys.* 354, 281-287.
21. Srivastava, D., Schuermann, J. P., White, T. A., Krishnan, N., Sanyal, N., Hura, G. L., Tan, A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2010) Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*, *Proc. Natl. Acad. Sci. USA* 107, 2878-2883.
22. Singh, H., Arentson, B. W., Becker, D. F., and Tanner, J. J. (2014) Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-channeling tunnel and the quinone binding site *Proc. Nat. Acad. Sci. U. S. A.*, in press.
23. Moxley, M. A., Sanyal, N., Krishnan, N., Tanner, J. J., and Becker, D. F. (2014) Evidence for Hysteretic Substrate Channeling in the Proline Dehydrogenase and Delta1-Pyrroline-5-carboxylate Dehydrogenase Coupled Reaction of Proline Utilization A (PutA), *J. Biol. Chem.* 289, 3639-3651.
24. Williams, I., and Frank, L. (1975) Improved chemical synthesis and enzymatic assay of delta-1-pyrroline-5-carboxylic acid, *Anal. Biochem.* 64, 85-97.
25. Moxley, M. A., and Becker, D. F. (2012) Rapid reaction kinetics of proline dehydrogenase in the multifunctional proline utilization A protein, *Biochemistry* 51, 511-520.
26. Krishnan, N., and Becker, D. F. (2005) Characterization of a bifunctional PutA homologue from *Bradyrhizobium japonicum* and identification of an active site residue that modulates proline reduction of the flavin adenine dinucleotide cofactor, *Biochemistry* 44, 9130-9139.
27. Moxley, M. A., Tanner, J. J., and Becker, D. F. (2011) Steady-state kinetic mechanism of the proline:ubiquinone oxidoreductase activity of proline utilization A (PutA) from *Escherichia coli*, *Arch. Biochem. Biophys.* 516, 113-120.
28. Srivastava, D., Schuermann, J. P., White, T. A., Krishnan, N., Sanyal, N., Hura, G. L., Tan, A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2010) Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*, *Proc. Natl. Acad. Sci. U. S. A.* 107, 2878-2883.
29. Patil, P. V., and Ballou, D. P. (2000) The use of protocatechuate dioxygenase for maintaining anaerobic conditions in biochemical experiments, *Anal. Biochem.* 286, 187-192.

30. Schuermann, J. P., White, T. A., Srivastava, D., Karr, D. B., and Tanner, J. J. (2008) Three crystal forms of the bifunctional enzyme proline utilization A (PutA) from *Bradyrhizobium japonicum*, *Acta Cryst. F64*, 949-953.
31. Leslie, A. G. (2006) The integration of macromolecular diffraction data, *Acta Cryst. D62*, 48-57.
32. Evans, P. (2006) Scaling and assessment of data quality, *Acta Cryst. D62*, 72-82.
33. Adams, P. D., Afonine, P. V., Bunkoczi, 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 Crystallogr. D Biol. Crystallogr.* 66, 213-221.
34. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Cryst. D Biol. Crystallogr.* 66, 486-501.
35. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* D66, 12-21.
36. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res.* 28, 235-242.
37. Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta Cryst. D50*, 178-185.
38. Berka, K., Hanak, O., Sehnal, D., Banas, P., Navratilova, V., Jaiswal, D., Ionescu, C. M., Svobodova Varekova, R., Koca, J., and Otyepka, M. (2012) MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels, *Nucleic Acids Res.* 40, W222-227.
39. Petrek, M., Kosinova, P., Koca, J., and Otyepka, M. (2007) MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels, *Structure* 15, 1357-1363.
40. Hekkelman, M. L., Te Beek, T. A., Pettifer, S. R., Thorne, D., Attwood, T. K., and Vriend, G. (2010) WIWS: a protein structure bioinformatics Web service collection, *Nucleic Acids Res.* 38, W719-723.
41. Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., and Damborsky, J. (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures, *PLoS Comput. Biol.* 8, e1002708.

42. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
43. Anderson, K. S., Kim, A. Y., Quillen, J. M., Sayers, E., Yang, X. J., and Miles, E. W. (1995) Kinetic characterization of channel impaired mutants of tryptophan synthase, *J. Biol. Chem.* 270, 29936-29944.
44. Cobessi, D., Dumas, R., Pautre, V., Meinguet, C., Ferrer, J. L., and Alban, C. (2012) Biochemical and structural characterization of the Arabidopsis bifunctional enzyme dethiobiotin synthetase-diaminopelargonic acid aminotransferase: evidence for substrate channeling in biotin synthesis, *The Plant cell* 24, 1608-1625.
45. Brown, E. D., and Wood, J. M. (1993) Conformational change and membrane association of the PutA protein are coincident with reduction of its FAD cofactor by proline, *J. Biol. Chem.* 268, 8972-8979.
46. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation, *Proteins* 50, 437-450.

Schemes



Scheme 5.1. Overall reaction catalyzed by proline utilization A (PutA). Flavin-dependent proline dehydrogenase (PRODH) catalyzes the oxidation of proline to Δ^1 -pyrroline-5-carboxylate (P5C). P5C undergoes a non-enzymatic hydrolysis, resulting in glutamate- γ -semialdehyde (GSA). GSA is oxidized to glutamate by P5C dehydrogenase (P5CDH) using an NAD^+ cofactor.

Table 5.1. Primers used for Site-Directed Mutagenesis

Mutant	Primers
T348Y	Fwd 5'-GCGCCTATTGGGACT <u>AC</u> GAGATCAAGCGCGCG-3' Rev 5'-CGCGCGCTTGATCTCGTAGTCCCAATAGGCGC-3'
S607Y	Fwd 5'-AGACGCTCGACGATGCGCTCT <u>AT</u> GAGCTGCGCG-3' Rev 5'-GAGCGCATCGTCGAGCGTCTTGCCGCCCTCG-3'
D778Y	Fwd 5'-GCTGCCGGAGCAGGTCGCCT <u>AC</u> GACGTTGTCACC-3' Rev 5'-GGCGACCTGCTCCGGCAGCGCGGTGGC ATCG-3'
D779A	Fwd 5'-TGCCGGAGCAGGTCGCCGACG <u>CC</u> GTTGTCACCTCC 3' Rev 5'-GTCGGCGACCTGCTCCGGCAGCGCGGTGGC-3'
D779W	Fwd 5'-TGCCGGAGCAGGTCGCCGACT <u>TGG</u> GTTGTCACCTCC-3' Rev 5'-GTCGGCGACCTGCTCCGGCAGCGCGGTGGC-3'
D779Y	Fwd 5'-CCGGAGCAGGTCGCCGACT <u>AC</u> GTTGTCACCTCCGC-3' Rev 5'-GCGGAGGTGACAACGTAGTCGGCGACCTGCTCCGG-3'

Table 5.2. PRODH Kinetic Parameters

BjPutA	Proline ^a			CoQ ₁ ^b		
	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (M ⁻¹ s ⁻¹)	K_m (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (M ⁻¹ s ⁻¹)
wild-type	43 ± 5	3.1 ± 0.1	72 ± 8.6	105 ± 6	2.9 ± 0.1	27619 ± 1713
T348Y	30 ± 2	1.8 ± 0.1	60 ± 4.0	59 ± 2	1.9 ± 0.1	32203 ± 1204
S607Y	46 ± 6	1.6 ± 0.1	35 ± 4.8	131 ± 16	2.0 ± 0.1	15267 ± 1987
D778Y	91 ± 38	0.36 ± 0.07	4.0 ± 1.8	82 ± 15	0.33 ± 0.02	4024 ± 775
D779A	56 ± 7	1.8 ± 0.1	32 ± 4.2	188 ± 22	2.5 ± 0.1	13297 ± 1725
D779Y	43 ± 2	2.7 ± 0.1	63 ± 3.1	56 ± 2	3.1 ± 0.1	55357 ± 2102
D779W	30 ± 4	1.9 ± 0.1	63 ± 8.6	109 ± 12	2.3 ± 0.1	21100 ± 2593

^a1-200 mM proline, 250 μM CoQ₁, 0.5 μM enzyme, 50 mM potassium phosphate (pH 7.5)

^b150 mM proline, 10-350 μM CoQ₁, 0.5 μM enzyme, 50 mM potassium phosphate (pH 7.5)

Table 5.3. P5CDH Kinetic and NAD⁺ Binding Parameters

BjPutA	k_{cat} (s ⁻¹) ^a	K_{m} (mM) ^a	$k_{\text{cat}}/K_{\text{m}}$ (M ⁻¹ s ⁻¹)	K_{d} (μM, NAD ⁺) ^b
wild-type	3.4 ± 0.1	0.42 ± 0.04	8095 ± 822	0.60 ± 0.04
T348Y	4.2 ± 0.2	0.42 ± 0.04	10000 ± 1017	0.75 ± 0.06
S607Y	4.5 ± 0.2	0.48 ± 0.03	9375 ± 664	1.00 ± 0.04
D778Y	3.8 ± 0.1	0.38 ± 0.02	10000 ± 567	0.67 ± 0.04
D779A	5.0 ± 0.1	0.38 ± 0.03	13157 ± 1102	0.64 ± 0.05
D779Y	0.02 ± 0.01	0.20 ± 0.03	100 ± 16	0.65 ± 0.04
D779W	0.003 ± 0.001	0.35 ± 0.15	8.6 ± 4	0.78 ± 0.05

^a0.01-6 mM L-P5C, 0.2 mM NAD⁺, 0.25 μM enzyme, 50 mM potassium phosphate (pH 7.5, 600 mM NaCl)

^bfrom fluorescence quenching with 0.1-25 μM NAD⁺, 0.25 μM enzyme, 50 mM potassium phosphate (pH 7.5)

Table 5.4. X-ray Diffraction Data Collection and Refinement^a

	D779W	D779Y	D778Y
space group	C2	C2	C2
unit cell parameters (Å, °)	<i>a</i> = 166.9 <i>b</i> = 195.3 <i>c</i> = 108.8 β = 121.6	<i>a</i> = 167.1 <i>b</i> = 196.0 <i>c</i> = 108.7 β = 121.4	<i>a</i> = 166.1 <i>b</i> = 195.1 <i>c</i> = 108.4 β = 121.5
wavelength (Å)	1.000	1.000	1.000
diffraction resolution (Å)	32.0 - 2.20 (2.32 - 2.20)	32.0 - 2.30 (2.42 - 2.30)	46.9 - 2.30 (2.42 - 2.30)
no. of observations	549668	490658	485882
no. of unique reflections	149604	130815	130019
$R_{merge}(I)$	0.106 (0.464)	0.103 (0.515)	0.095 (0.524)
$R_{meas}(I)$	0.124 (0.556)	0.120 (0.602)	0.112 (0.612)
$R_{pim}(I)$	0.063 (0.302)	0.061 (0.310)	0.058 (0.314)
mean I/σ	6.8 (2.1)	8.1 (2.2)	10.0 (2.5)
completeness (%)	99.9 (99.3)	99.3 (98.8)	99.9 (100)
multiplicity	3.7 (3.3)	3.8 (3.6)	3.7 (3.8)
no. of protein chains	2	2	2
no. of protein residues	1943	1943	1941
no. of protein atoms	14390	14386	14490
no. of FAD atoms	106	106	106
no. of water molecules	531	296	419
no. of sulfate ions	6	6	8
no. of glycerol molecules	4	3	4
R_{cryst}	0.208	0.216	0.195
R_{free}^b	0.241	0.251	0.235
rmsd bond lengths (Å)	0.008	0.008	0.009
rmsd bond angles (deg.)	1.102	1.107	1.106
Ramachandran plot ^c			
favored (%)	98.8	98.1	98.1
outliers (residues)	2	2	0
average B -factors (Å ²)			
protein	31.5	38.9	34.5
FAD	20.0	29.3	25.2
water	28.5	31.8	30.4
sulfate	61.4	67.6	74.3
glycerol	36.5	47.3	45.3
coordinate error (Å) ^d	0.27	0.31	0.28
PDB accession code	4Q71	4Q72	4Q73

^aValues for the outer resolution shell of data are given in parenthesis.

^b5 % random test set. A common set was used for refinement of all structures.

^cThe Ramachandran plot was generated with RAMPAGE.(46)

^dMaximum likelihood-based coordinate error estimate reported by PHENIX.

Figures and Legends

Figure 5.1. The tunnel/cavity system of BjPutA. (A) BjPutA protomer with PRODH in blue, P5CDH in pink, and oligomerization flap in green. The FAD and NAD⁺ are shown in yellow and green sticks, respectively. Catalytic Cys792 of the P5CDH active site is indicated. The gray surface represents the predicted channeling pathway calculated with MOLE. The $\alpha 5a$ and 770s helices are colored gold and cyan, respectively. We note that in a tetramer of BjPutA, the dimerization flap of one protomer covers the tunnel of the other protomer. (B) Details of the predicted channeling pathway. The predicted path from MOLE is shown in mesh. Models of P5C and GSA in the tunnel are shown for scale (green). (C) Another view of the tunnel/cavity system, with the predicted channeling tunnel calculated from MOLE in gray mesh and the off-pathway cavity calculated using VOIDOO in red mesh.

Figure 5.2. Channeling assays of BjPutA wild-type and mutants. Assays were performed in 50 mM potassium phosphate (pH 7.5, 25 mM NaCl, 10 mM MgCl₂) with 0.187 μ M BjPutA enzyme, 40 mM proline, 100 μ M CoQ₁, and 200 μ M NAD⁺.

Figure 5.3. Channeling assays with increasing concentrations of D779Y (A) and D779W (B). NADH formation was monitored using fluorescence by exciting at 340 nm and recording the emission at 460 nm. Assays were performed with BjPutA wild-type (0.187 μ M) and increasing concentration of mutants (0.187-1.87 μ M) in 50 mM potassium phosphate (pH 7.5, 25 mM NaCl, 10 mM MgCl₂) containing 40 mM proline, 100 μ M CoQ₁, and 200 μ M NAD⁺.

Figure 5.4. NAD⁺ binding to BjPutA. (A) Wild-type BjPutA (0.25 μM) was titrated with increasing concentrations of NAD⁺ (0-20 μM NAD⁺) in 50 mM potassium phosphate buffer (pH 7.5). Inset is a plot of the change in tryptophan fluorescence versus NAD⁺ fitted to a single site binding isotherm. A K_d value of 0.60 ± 0.04 μM was estimated for NAD⁺-BjPutA. (B) ITC analysis of NAD⁺ binding to wild-type BjPutA. Upper panel is the raw data of wild-type BjPutA (23.4 μM) titrated with increasing amounts of NAD⁺ in 50 mM Tris buffer (pH 7.5). Lower panel is the integration of the titration data. The binding of NAD⁺ to BjPutA is shown to be exothermic and a best fit of the data to a single site binding isotherm yielded a K_d value of 1.5 ± 0.2 μM.

Figure 5.5. Electron density maps and local conformational changes. (A) Electron density map for D778Y. (B) Electron density map for D779Y. (C) Electron density map for D779W. (D) Superposition of BjPutA (gray), D778Y (gold), D779Y (cyan), and D779W (magenta). The cages in panels A-C represent simulated annealing σ_A -weighted $F_o - F_c$ omit maps contoured at 2.5σ .

Figure 5.6. Constriction of the channeling tunnel by Tyr779 in D779Y. (A) The gray cylinder represents the channeling pathway calculated from the wild-type BjPutA structure (PDB code 3HAZ) using MOLE, and the view is from the P5CDH active site looking through the tunnel toward the PRODH site. (B) Comparison of the predicted channeling pathway of wild-type BjPutA (gray surface) and D779Y (red mesh).

Figure 5.7. Constriction of the channeling tunnel by Trp779 in D779W. (A) The gray cylinder represents the channeling pathway calculated from the wild-type BjPutA structure (PDB code 3HAZ) using MOLE, and the view is from the P5CDH active site looking through the tunnel toward the PRODH site. (B) Comparison of the predicted channeling pathway of wild-type BjPutA (gray surface) and D779W (red mesh).

Figure 5.8. Constriction of the channeling tunnel by Trp779 in D779W. (A) The gray cylinder represents the channeling pathway calculated from the wild-type BjPutA structure (PDB code 3HAZ) using MOLE, and the view is from the P5CDH active site looking through the tunnel toward the PRODH site. (B) Comparison of the predicted channeling pathway of wild-type BjPutA (gray surface) and D779W (red mesh).

Fig 5.1.

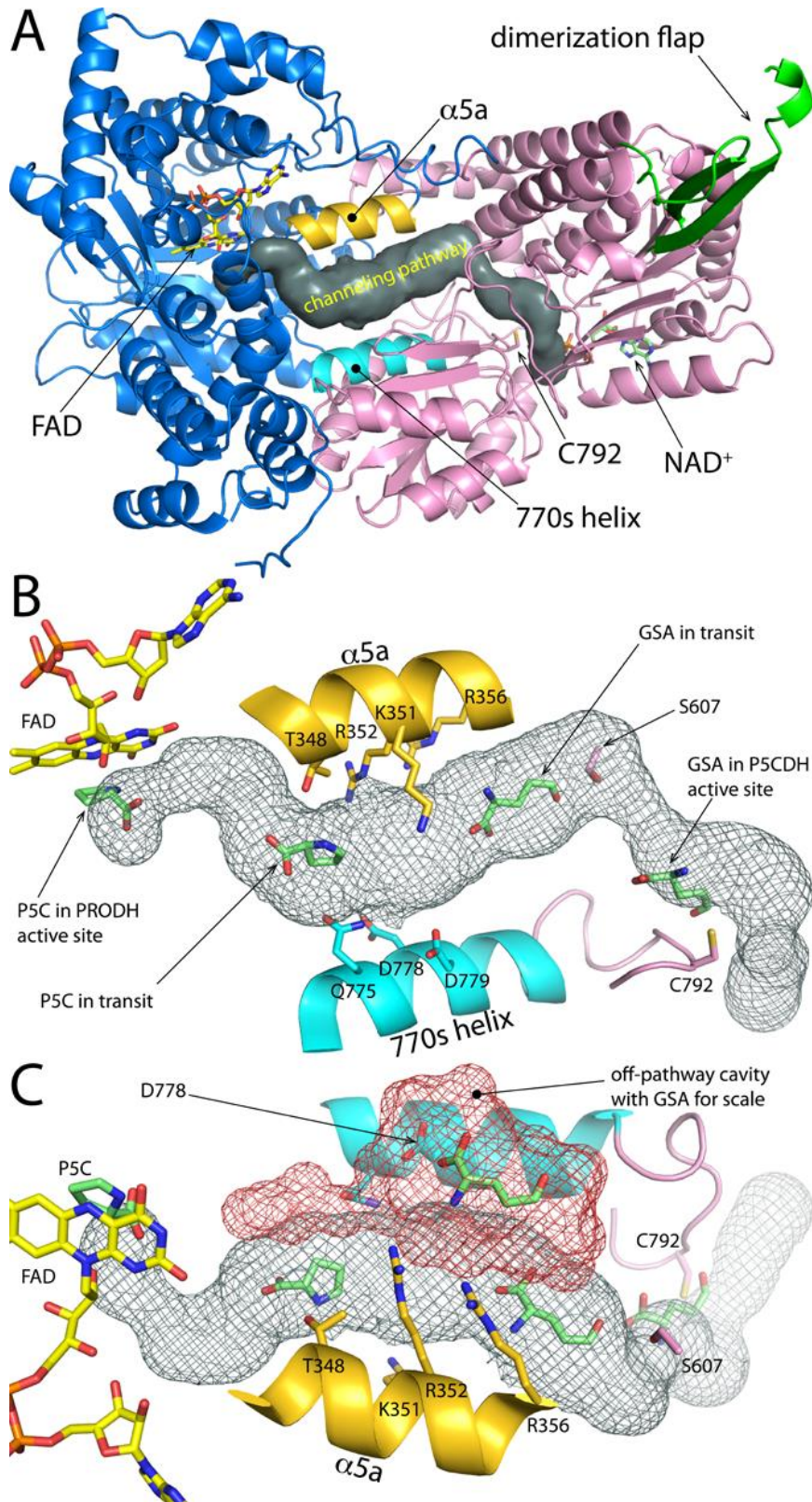


Fig 5.2.

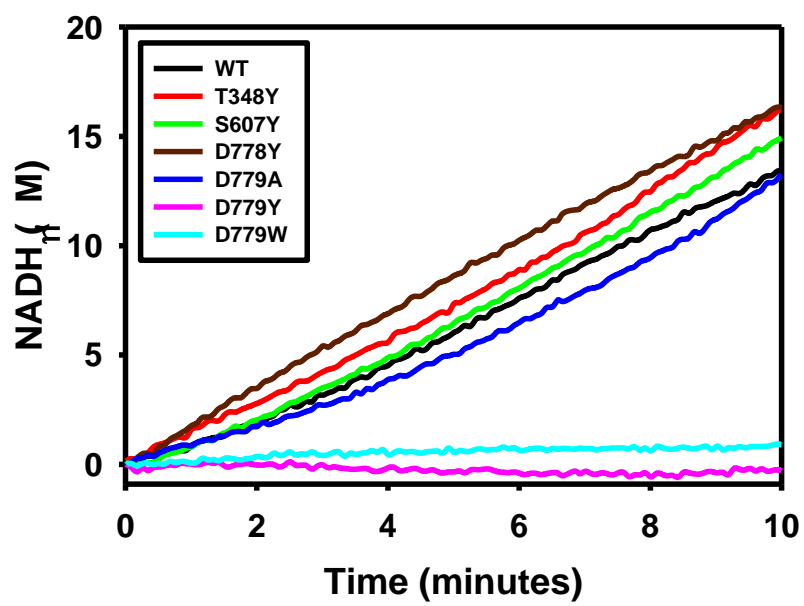


Fig 5.3.

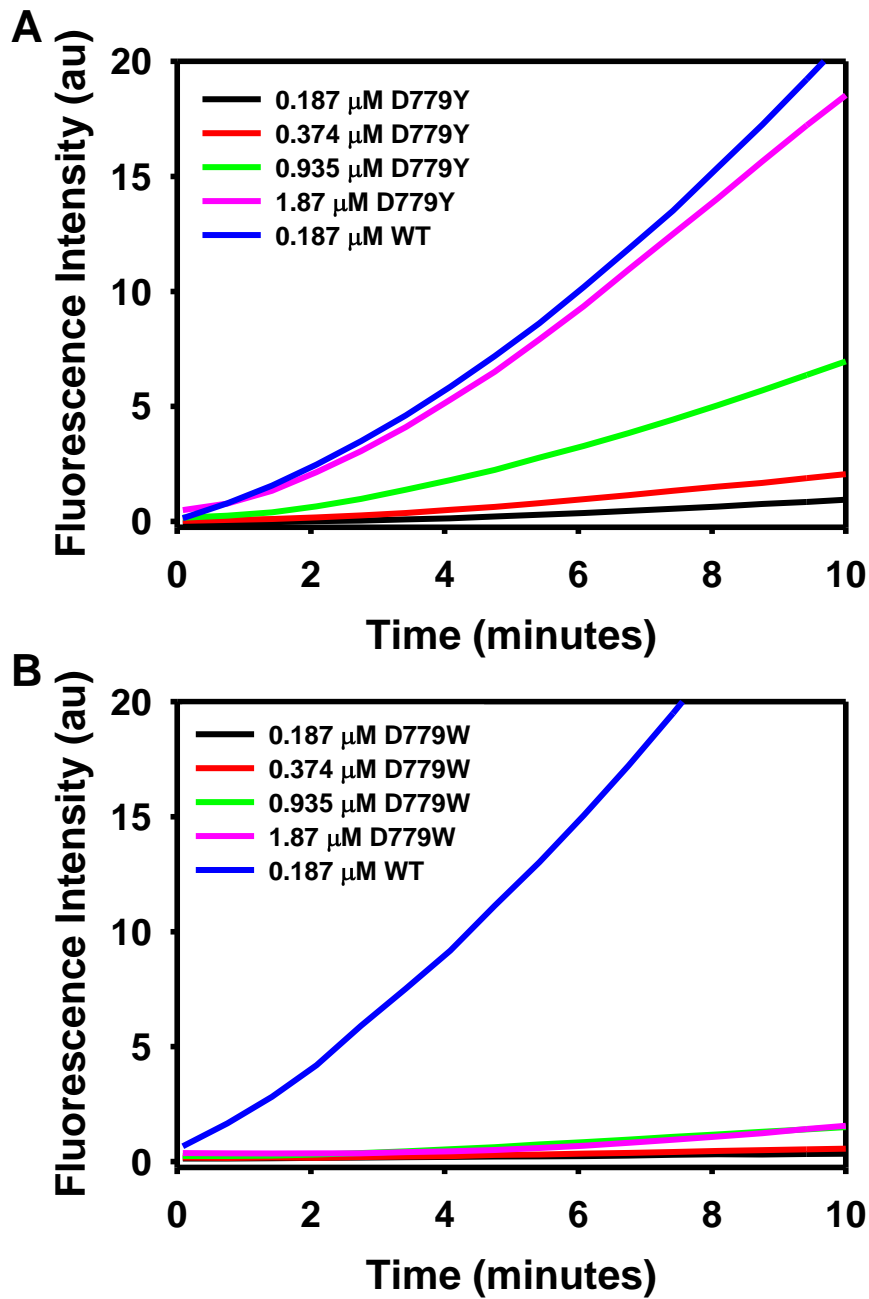


Fig 5.4.

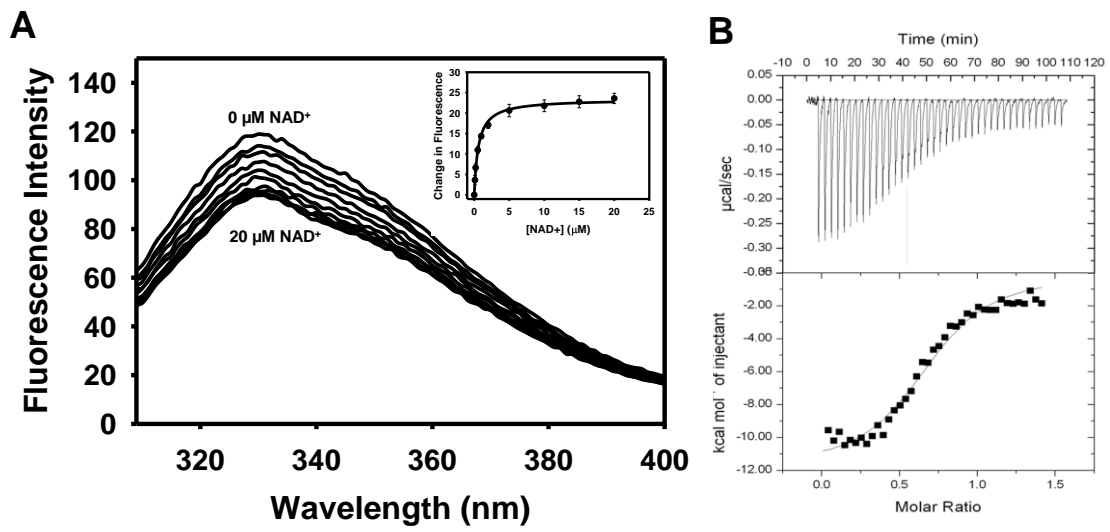


Fig 5.5.

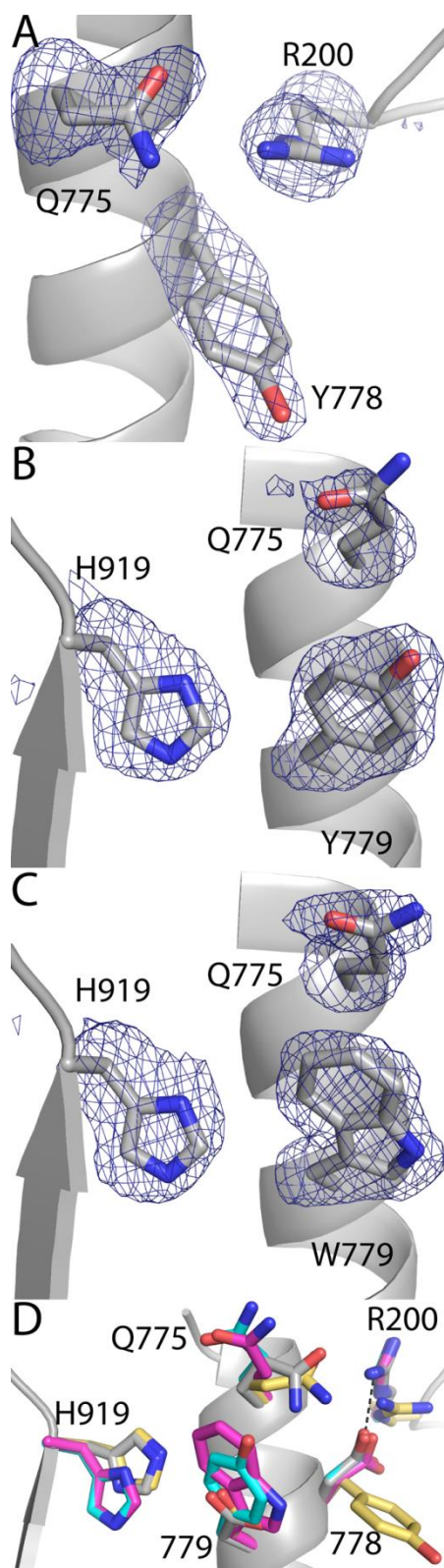


Fig 5.6.

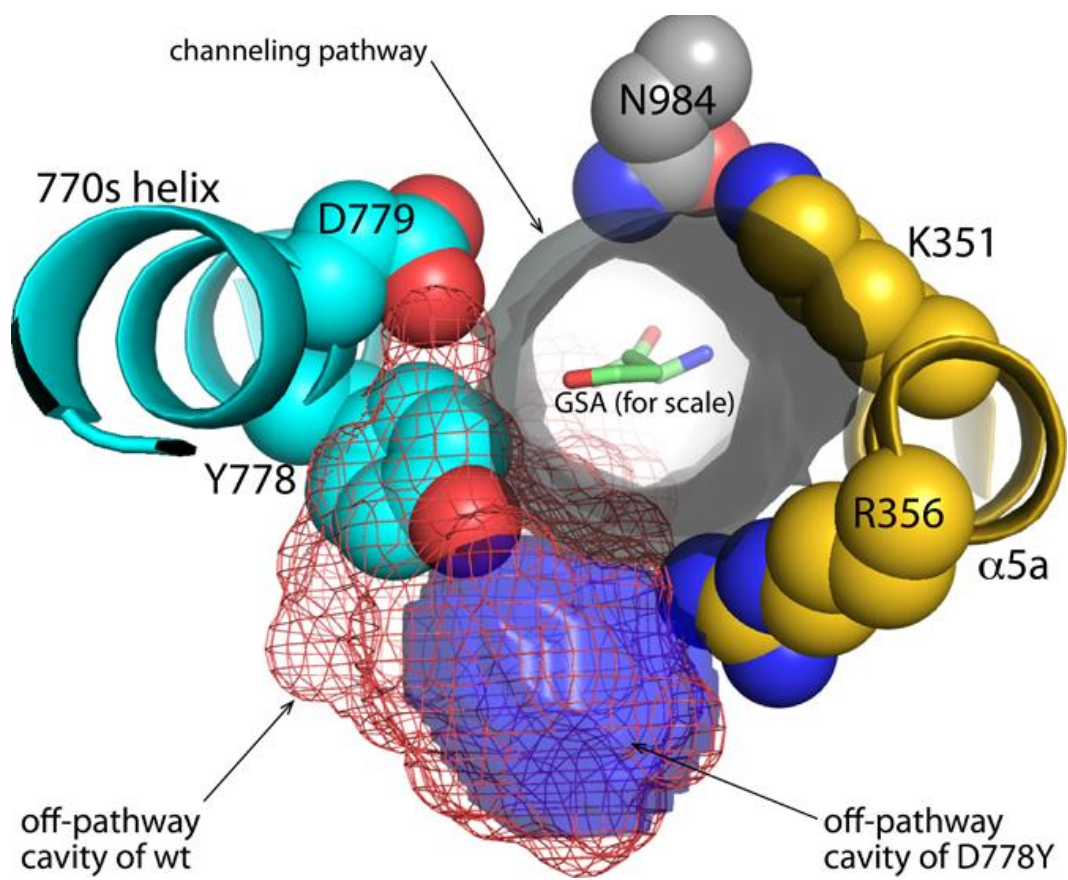


Fig 5.7.

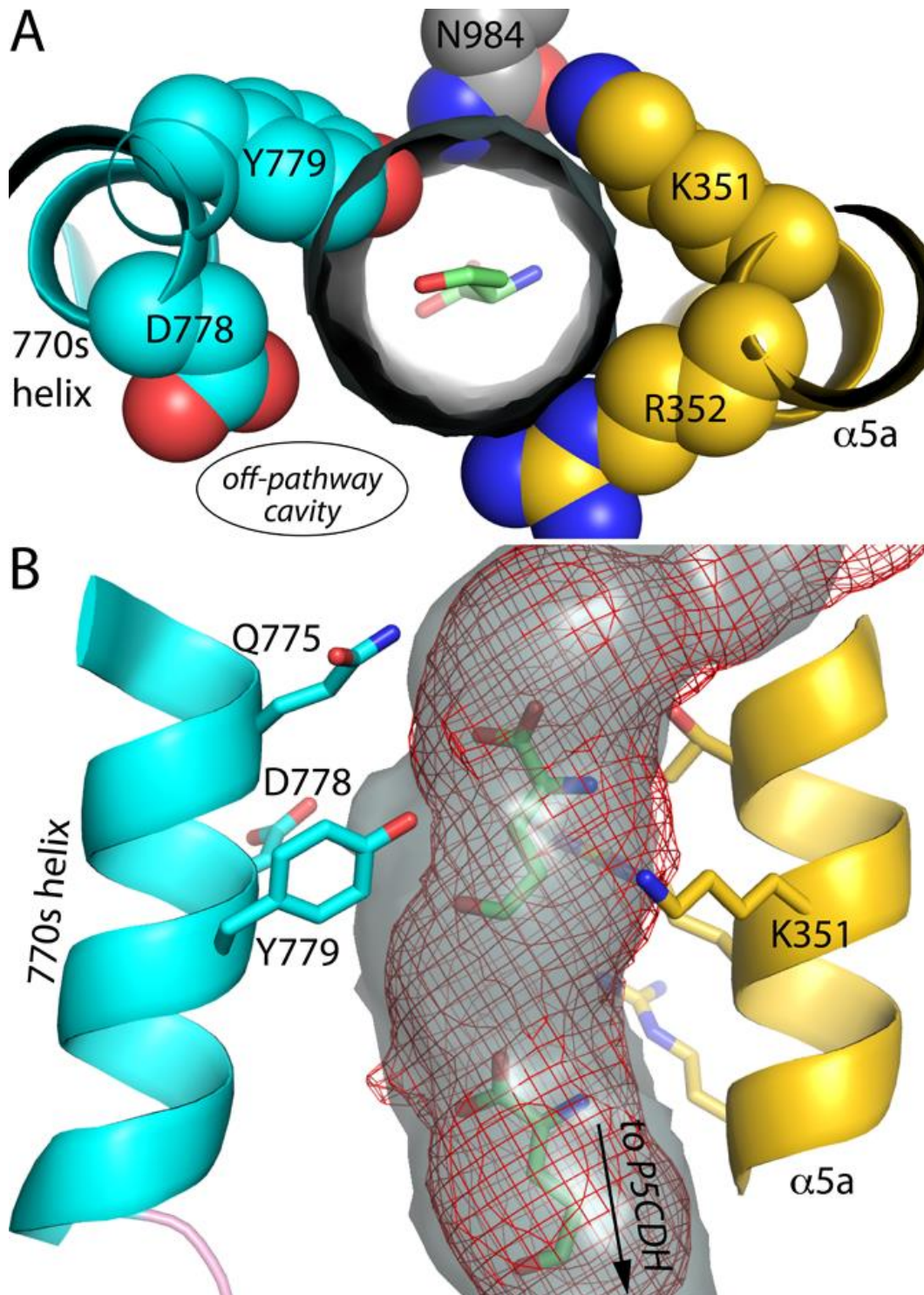
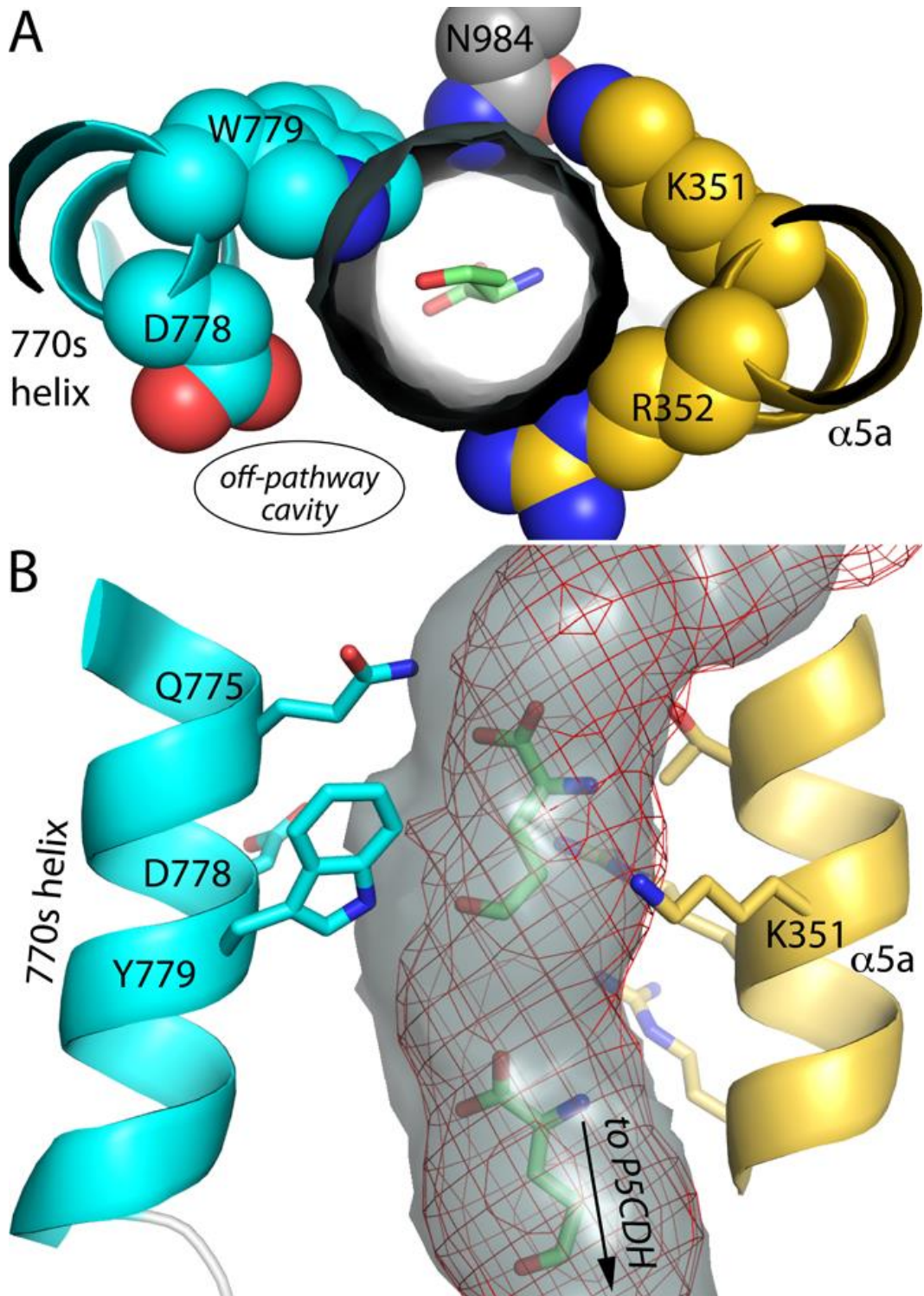


Fig 5.8.



Chapter 6:

Understanding Access to an Internal Cavity Discovered in
Bradyrhizobium japonicum Proline Utilization A (PutA)

Abstract

The two step oxidation of proline to glutamate is catalyzed through two enzymes, proline dehydrogenase (PRODH) which catalyzes the oxidation of proline to Δ^1 -pyrroline-5-carboxylate (P5C) and P5C dehydrogenase (P5CDH) which catalyzes the oxidation of glutamic γ -semialdehyde (GSA) to glutamate using NAD^+ as a cofactor. GSA and P5C exists in a pH dependent equilibrium with one another. In Gram-negative bacteria, these two enzymes are fused together into a single, bifunctional enzyme known as proline utilization A (PutA). The structure of PutA from *Bradyrhizobium japonicum* (BjPutA) has been solved previously and the structure illuminated the enzyme housing a large internal cavity. In conjunction with kinetic experiments, this cavity was proposed as a channel that moved the intermediate between the two active sites. Access to this channel for the intermediate from the PRODH domain was proposed to be modulated by an ion pair gate (R456-E197). Here we show that reduction of the flavins by proline breaks this ion pair and opens the gate, allowing access to the cavity. Further, disruption of the ion pairs via mutations severely hurts enzyme function (E460A and E197A) or completely kills it (R456M). Destruction of the ion pair also induces large conformational changes in the important helix 8, implicated in large movements between the closed and opened states.

6.1. Introduction

The proline catabolism pathway catalyzes the two-step oxidation of proline to glutamate (Scheme 1). Proline dehydrogenase (PRODH) catalyzes the first oxidation by transferring 2 electrons, in the form of H^+ , from proline to the FAD cofactor. The product, Δ^1 -pyrroline-5-carboxylate (P5C), goes through a non-enzymatic hydrolysis to glutamate- γ -semialdehyde (GSA). GSA is the substrate for the NAD^+ -dependent P5C dehydrogenase (P5CDH), which catalyzes the final oxidation to glutamate.

In Gram-negative bacteria PRODH and P5CDH are combined into a single, bifunctional enzyme known as proline utilization A (PutA) (1). The original full length structure of the enzyme showed the presence of an internal, water filled cavity linking the two active sites together (2). An internal cavity connecting the two active sites is indicative of substrate channeling which is the transfer of an intermediate between two active sites without equilibration to the bulk solvent (3). Channeling provides an increase in kinetic efficiency by providing a direct path between active sites and eliminating diffusion into the surrounding environment, resulting in a decrease in transit time (4). As seen in dimethylglycine oxidase (DMGO) reactive intermediates cannot be allowed to diffuse to bulk solvent without devastating results (5). Enzymes, such as DMGO, utilize channels to prevent intermediate escape (6-8).

P5C/GSA release needs to be tightly controlled as free P5C is a known inhibitor of three separate enzymes in *Escherichia coli*, cytidine 5'-triphosphate synthase, the amidotransferase domain of carbamoyl phosphate synthase, and glucosamine-5-phosphate synthase (9-11). It is also known to form adducts with important metabolites such as pyruvic acid (12). Such a volatile and interactive intermediate needs control and

indeed initial studies monitoring production with a ^{14}C -labeled proline of *Salmonella typhimurium* PutA hinted at the idea of substrate channeling (13). More recent studies using *Bradyrhizobium japonicum* (BjPutA) and *Geobacter sulfurreducens* (GsPutA) have further cemented the idea of substrate channeling (2, 14). Further, a complete kinetic mechanistic analysis was performed on *E. coli* PutA and brought to light PutA as a hysteretic enzyme, with the first turnover being slower than subsequent turnovers (15).

The kinetic analysis clearly establishes that transfer of the intermediate happens through a substrate channeling mechanism. Both structures of BjPutA and GsPutA indicate an internal cavity separated from the PRODH active site by a conserved ion pair between an arginine and glutamate (R456 and E197 in BjPutA) (2, 14). The initial hypothesis put forward was that this ion pair acts as a gate that closes upon proline binding and opens to release the product P5C into the internal cavity. The GsPutA oxidized and reduced structures confirmed this by providing snap shots of an open and closed form of the gate, which is a common mechanism seen in tryptophan synthase, glutamine amidotransferase, and glucosamine-6-phosphate synthase (16-18). A BjPutA-lactate structure (unpublished) has shown that a second ion pair between E460 and R457 also forms during substrate binding. Sequence alignments have also revealed a conserved glycine (G196) before the conserved glutamate (E197) involved in the gate acting ion-pair. The GE is seen in bifunctional and trifunctional PutAs but also the monofunctional PRODH (19).

The original hypothesis of a gating mechanism facilitating entrance into the internal cavity is established in GsPutA (14). Here we study the reduction of the oxidized BjPutA enzyme using the natural substrate, proline. Also, disruption of the two ion pairs

via mutations is analyzed to determine their influence on substrate binding and the gating mechanism. This study utilizes mutation, kinetic, and X-ray crystallography to gain further insight into how conformational changes in the PRODH domain allow access to the cavity. Proline soaking of the crystals is performed to monitor conformational changes that occur when the active enzyme interacts with the real substrate. The ion pairs have been disrupted with mutations in an effort to establish their role in substrate binding and movement into the cavity. Kinetic and structural analysis of the mutant enzymes shows large changes to the helix 8 that highlight their role in stability of the PRODH active site.

6.2. Materials and Methods

6.2.1 Cloning and PCR

Mutagenic primers were purchased from Integrated DNA Technologies. QuikChange II kit (Agilent Technologies) was used to generate all mutants. Mutant plasmids were transformed into DH5 α cells, and the resulting plasmids were sequenced by DNA Core facilities to confirm the mutations. Plasmid was purified from DH5 α cells using the QIAprep Spin miniprep Kit (Qiagen) and the plasmid was stored at -20 °C. The plasmid, pKA8H, contains an N-terminal His tag that is cleavable with Tobacco Etch Virus protease (TEVP).

6.2.2 Protein Expression and Purification

Plasmid was retransformed into BL21DE3(pLysS) cells, plated on LB Agar plates, and incubated at 37 °C for about 12 hours. Afterwards they were removed from incubation and stored at 4 °C. Single colonies were picked and used in about 10 mL of LB Broth for a starter culture and allowed to grow overnight. The starter culture was used

to inoculate 1 L of LB Broth and then put in a shaker at 37 °C and 250 rpm. When the cells reached an OD₆₀₀ of about 0.8, the cultures were induced with 50 μM isopropyl β-D-1-

Thiogalactopyranoside (IPTG) and the temperature reduced to 20 °C and rpm dropped to 200(20). The cells were then incubated for another 16 hours at the lower temperature. Cells were harvested by centrifugation, resuspended in 50 mM HEPES, 150 mM NaCl and 5% glycerol at a pH of 8.0, and flash frozen at -80 °C.

Frozen cells were thawed and 0.3 mM protease inhibitor phenylmethanesulfonylfluoride (PMSF) was added. The cells were disrupted via 4 rounds of sonication, with each round consisting of 30 seconds of use, and then 2 minutes to cool down. The cell lysate was centrifuged for 1 hour at 16,500 rpm in a SS-34 rotor.

The supernatant was decanted off and filtered through a 0.2 μm filter. This filtered cell lysate was loaded onto a Ni-NTA column equilibrated with 20 mM HEPES, 50 mM NaCl and 5% glycerol at pH 8.2. The column was washed with the same buffer with 30 mM imidazole to remove impurities and then 300 mM imidazole at which point the protein eluted off the column. Fractions were collected in 2 mL aliquots. The fractions were pooled and TEV protease was added to the pooled fractions at a final concentration of 1 mg/mL. The enzyme was allowed to incubate with TEVP for about an hour and a half at 28 °C. After, the enzyme was placed in a dialysis buffer containing 20 mM HEPES, 150 mM NaCl, and 5% glycerol at 4 °C overnight. The following day the enzyme was removed and loaded back onto a Ni-NTA column equilibrated with the same binding buffer as before. Again the column was washed with 30 mM imidazole, at which point the enzyme eluted and fractions were collected in 2 mL aliquots. The fractions were

tested for purity on SDS-PAGE. The fractions were determined to be pure, pooled, and placed in dialysis of the pre-crystallization buffer containing 50 mM TRIS, 0.5 mM EDTA, 150 mM NaCl, 0.5 mM THP, and 5% glycerol at pH 7.5 at 4 °C overnight.

After removing from dialysis, protein concentration was checked using the BCA method. Centrifugal filters (Millipore) with a 30 kDa pore size membrane were used to increase the concentration to 8 mg/mL which was verified again using the BCA method.

6.2.3 Protein Crystallization

Crystallization of mutant BjPutAs occurred in similar conditions as described for BjPutA previously (21). Crystals were grown in sitting drops at room temperature in the presence of 1.7-2.1 M ammonium sulfate and Tris buffer at pH 7.5-9.0. Microseeding was performed with a seed stock made from crushing wild-type enzyme crystals, and vortexing the samples from there. All mutants used a 100-fold dilution of the seed stock to achieve maximum crystal growth(22). Subsequent optimizations utilized the crystals of the respective mutants grown at room temperature with a seed stock prepared in the same manner. Crystals were cryoprotected in 25-30% glycerol and then flash cooled by dunking in liquid nitrogen.

Wild-type BjPutA crystals were grown in the same conditions as described for the 3HAZ structure (21). These crystals were cryoprotected with 2 M proline which was achieved through 0.5 M incremental steps at a starting concentration of 0.5 M proline(23). The cryoprotection in proline took approximately 10-15 minutes to reach the desired concentration of 2 M. These crystals were then looped and plunged and stored in liquid nitrogen. If many crystals were being pulled from the same drop, the cryobuffer being used was replaced with fresh solution after every 2-3 crystals.

6.2.4 Data Collection and Refinement

The space group is $C2$ with a BjPutA dimer in the asymmetric unit. X-ray diffraction data sets were collected at beamline 4.2.2 of the Advanced Light Source (ALS) using a NOIR-1 detector. The data were integrated with MOSFLM (24) and scaled with SCALA (25). The E460A and E197A mutants were collected using the ALS 4.2.2 beamline on their new CMOS-based Taurus-1 detector. XDS (26) was used for integration and SCALA was used for scaling. Refinements in PHENIX (27) were initiated from models derived from the structure of wild-type BjPutA (Protein Data Bank (PDB) entry 3HAZ). COOT (28) was used for model building. The structures were validated with MolProbity (29) and the PDB validation server. Data collection and refinement statistics are listed in Table 2.

The changes in the internal cavity were monitored using VOIDOO (30) which characterizes cavities. Specifically, the probe-occupied (option O) calculation was used with a probe radius of 1.4 Å for water, which rolls a probe along the internal walls of the cavity and maps out all space that was occupied by the rolling of this probe. It was run with a grid spacing of 0.5 Å and a starting point in the middle of the internal cavity, near residue 779. MOLE(31), which finds tunnels that connect to the bulk solvent, was used to identify and characterize the tunnel system found inside this enzyme. The Mole 2.0 server(32) was used with a starting location of the sulfate ion in the PRODH active site.

6.2.5 Steady-State Kinetics

Steady-state kinetic assays were performed at room temperature. Kinetic parameters for the PRODH domain were determined for proline. The assay used is based on reduction of dichlorophenolindophenol (DCPIP) as described previously for PutA

proteins. One unit of activity is the quantity of enzyme that transfers electrons from 1 μmol of substrate to DCPIP per minute at room temp (33). The enzyme concentration for E460A and E197A were 1.45 nM and 1.64 nM, respectively. Proline concentration was varied from as low as 1 mM all the way up to 1.2 M. All data were plotted and fitted using Origin 9.0.

6.3. Results

6.3.1 Proline Reduced BjPutA Structure

The original BjPutA structure was crystallized in the presence of 2 M ammonium sulfate and the structure revealed a sulfate bound in the active sites of both PRODH and P5CDH. The 3HAZ structure illuminated arginine 456 (R456) interacting with both a sulfate in the active site and also glutamate 197 (E197). Any attempts to soak in inhibitors were met with failure as the high concentration of sulfate out competed with the much smaller concentrations of inhibitors used. However, proline is a known cryoprotectant and can be used in concentrations equal to and higher than that of the sulfate present. As a result crystals were soaked with 2 M proline for cryoprotection. BjPutA crystals have a distinct yellow color resulting from the FAD cofactor. As proline concentrations increased, the crystals visibly turned colorless from the edges inward, finally becoming completely colorless between 1 and 1.5 M proline.

. This color change points to reduction of the FAD which causes a butterfly bend in the isoalloxazine ring leading to the loss of color. This butterfly bend is widely associated with the reduced state of the cofactor. After hydride transfer to the N5 the aromaticity is disrupted and the isoalloxazine ring adopts a lower energy butterfly bend (Figure 1). This characteristic bend is seen in the proline reduced structure, as compared

to the planar ring seen in 3HAZ. The two ends point towards the re-face, away from the active site. The ribityl chain adopts the reduced conformation described previously for the N-propargylglycine (PPG) inactivated PRODH domain of EcPutA. In fact, the only major difference between the flavins in these two structures is the isoalloxazine ring has a more severe bend in the PPG inactivated structure. The proline reduced BjPutA flavins mimics very closely what the GsPutA PPG-inactivated and dithionite reduced structures show. The ribityl chain is oriented in the same way, but the bending is more severe in the GsPutA structures, at about 27° (14).

The reduced form of the enzyme has a proline occupying the P5CDH active site (Figure 2). The carboxyl group occupies a similar area as the sulfate that was seen in the 3HAZ structure. One oxygen from the carboxyl group makes hydrogen bonds with the guanidinium group of arginine 791 (R791) and the back bone amino group of glycine 946 (G946). The other oxygen of the carboxyl group interacts with the same amino group on G946 and the backbone amino group of alanine 947. This interaction is with amino acids on the binding loop as described previously (34). The proline ring points towards the catalytic cysteine and sits between two phenylalanines, F659 and F954. The aliphatic chain of glutamate is also known to sit here.

In this proline reduced structure, there is no density for the sulfate in the PRODH active site. The guanidinium group of R456 now points away from the active site and makes a 3.1 \AA hydrogen bond with the backbone carbonyl of alanine 344. The carboxyl group of E197 now points away from the active site, into the internal cavity and makes a 3.2 \AA hydrogen bond with arginine 200 as well as a water mediated hydrogen bond with

arginine 284 (Figure 3A). The original 2.8 Å hydrogen bond seen for this ion pair is broken and the side chains now sit about 6.8 Å away from each other.

To analyze the possibility of this being a snapshot of an open conformation of the gate, Mole calculations were run on both the wild-type and proline reduced structures (Figure 3). Both could find their way into the cavity and back out, but how they entered the cavity was markedly different. The wild-type mole calculations found a narrow passage that passed by R456 and a hole at the C-terminal of Chain B, which is known as the flap. This route seems to briefly leave the channel of the enzyme, which would leave the intermediate solvent exposed and is inconsistent with the mechanism of substrate channeling. It finds many, probably water channels out of the PRODH domain, but when entering the internal cavity from the PRODH domain, it always enters through this “alternate” pathway.

The mole calculations performed on the proline reduced structure never found this pathway (Figure 4). Instead, the PRODH domain connects to the main cavity through the now open gate. In fact, every pathway into the main channel from the PRODH active site enters through the open gate, the alternate pathway is never seen in these mole calculations. The 6.8 Å hole is just big enough for P5C to pass through, which is estimated to have a sphere diameter of approximately 5.8 Å.

The VOIDOO calculations on both the wild-type and proline reduced structures show a very similar result as those seen from the Mole calculations (Figure 5). The wild-type structure stops at the closed gate and does not enter the PRODH domain when the point of origin is within the internal cavity. Also, when the origin is moved to the PRODH domain, VOIDOO cannot map an escape route into the channel. This is not true

for the proline reduced structure, however. Instead the introduction of a 6.8 Å gap where the gate used to be allows VOIDOO to connect the PRODH active site and internal cavity.

The backbone amine group of glycine 196 (G196) flips about 180 degrees and forces a conformational change of the back bone carbonyl group of leucine 195 (Figure 5). The amine group of G196 now faces in towards the active site, while the carbonyl of 195 has flipped 180 degrees and faces away from the active site. Helix 8 looks very similar to that of the wild type structure, except the side chain of glutamate 460 (E460) is completely disordered. Arginine 457 (R457), an arginine implicated in substrate binding as well adopts the same conformation as that seen in the 3HAZ structure.

6.3.2 E460A and E197A

E197A mutant was made to disrupt the ion pair shown to be acting as a gate to the cavity. E197 typically makes a 2.8 Å bond with R456 to create this gate, however severing this side chain makes such interactions impossible. The R456 adopts a conformation that varies from the wild-type. The R456 guanidinium group in E197A does not point at the sulfate in the active site. Instead of one of the terminal nitrogens of the guanidinium group interacting with the sulfate, as seen in the wild-type, the nitrogen connecting the side chain (NE) makes a 2.2 Å hydrogen bond with the sulfate (Figure 6). The NH1 actually faces towards the now absent E197 side chain.

Starting at leucine 459, the backbone conformation completely changes and disrupts alpha helix formation. The E460 side chain actually moves about 5.4 Å and the back bone has a shift of about 3.6 Å (Figure 7). This glutamate appears to swing into the active site with one of the oxygens being only 2.7 Å away from the sulfate. Asparagine

461 (N461) moves into the space that E460 occupies in the wild-type structure. This constitutes about a 5.6 Å shift that leads to the side chain moving out of the proximity of glutamate 59 (E59). The side chain of E59 orders as opposed to the wild-type where it is completely disordered. Further, E59 seems to make a new pair with one of the dual conformations of R457 as described below (Figure 7). Glycine 462 shifts about 6.5 Å away from where the wild-type structure is resulting in the largest relocation in helix 8. Alanine 463 only has a shift of about 1.6 Å but the side chain in the mutant structure adopts a new conformation than that of the wild-type, which strangely sits outside of the Ramachandran allowed regions. It should be noted, though, that the electron density at the C-terminal of helix 8 is weak, and it is possible that A463 simply cannot be positioned correctly due to lack of data. Starting at asparagine 464, the mutant and wild-type overlap again and the helix structure returns.

Arginine 457 in this case adopts a dual conformation. Conformation A makes a 2.7 Å hydrogen bond with tyrosine 453. It is similar to the conformation seen in the wild-type; however it is shifted about 1.5 Å towards the active site. Conformation B faces away from the active site and points directly at the newly ordered E59. The two terminal nitrogens on the guanidinium ion make a 2.6 and 3.2 Å hydrogen bonds with the two oxygens from the carboxyl group of the glutamate. The two conformations are just about 50/50 occupancy based off of PHENIX occupancy refinement.

The E460A mutant, however, the differences are harder to interpret. The active site only has one obvious change: the R456 guanidinium group is disordered. E197 looks the same as the wild-type but with the guanidinium of R456 disordered, no salt bridge is present. There is a sulfate in the active site, like in the wild-type and FAD is in the

oxidized form. With this in mind it is probably that R456 adopts a very similar orientation as that of the wild type. Helix 8, surprisingly, is hard to analyze. Residues 459-464, the C-terminal of the helix, are disordered with very little density. E460 has poor density in most BjPutA structures, including that of 3HAZ. This resulted in the exact position of this side chain to not be as concrete as even residues surrounding it. Also, there are no obvious interactions that would hold E460 in place. The presence of the side chain of E59 makes it seem as though the conformation that it is adopting is that seen in E197A. However, it is possible that the mutation has simply made this loop more mobile which, in turn, has allowed E59 to occupy space, previously taken by the helix.

Both of these glutamate mutants appear to severely reduce enzyme activity. Originally, the ubiquinone assay was attempted, as it is a more direct measurement of the enzyme activity. However, activity was so low using this assay, that it was not usable. Better success was seen using the DCPIP assay, but even in this case the activity was low. E460A appears to be the most severe as kinetic parameters were unable to be determined even when trying an upwards concentration of 1.2M proline. E197A had a K_m of 798 mM, k_{cat} of 0.080, with a k_{cat}/K_m of 0.0001.

6.3.3 R456M

R456M kills PRODH activity but leaves P5CDH active. R456 is important for substrate binding and in its absence the substrate can't be bound in the PRODH domain and activity is lost. Data for the structure for the non-proline soaked form was collected by Dhiraj and later refined by myself. I also collected data for the proline soaked structure. The methionine disrupts the ion pair seen in the wild-type and also substrate binding. E197 adopts the conformation seen in the proline reduced structure, hydrogen

bonding with R200. This is seen in both the proline and non-proline soaked R456M structures. Also the glycine 196 flip described for the proline reduced structure is seen in, again, both structures. The largest difference seen between the proline and non-proline soaked structures is the orientation of the M456. In the non-proline soaked structure the last carbon of the side chain points towards the active site. However, in the proline soaked structure, the side chain actually bends away from the active site, adopting a conformation similar to that in the proline soaked wild-type structure described above. It is worth noting that the FAD in both R456M structures is oxidized and that soaking of these crystals with proline did not yield colorless crystals.

6.4. Discussion

6.4.1 Proline Reduced wild-type BjPutA

In the original structure paper for BjPutA, it was hypothesized that the R456-E197 ion pair was a gate that moved and allowed the product of the PRODH domain access to the internal cavity. Mole calculations on the wild-type structure show that there is a path into the cavity that takes a narrow, circuitous route that moves towards the flap and briefly becomes solvent exposed. This route seems to defeat the purpose of a channel by increasing risk of losing the volatile product of the PRODH domain to bulk solvent. That coupled with a narrowing of the path as it passes by helix 8 to move around the R456, makes this path of travel possible but unlikely.

In the proline reduced structure, the R456-E197 pair breaks as the two side chains make large conformational changes to move away. Mole calculations only ever find its way into the internal cavity through the now open gate. The alternate pathway seen in the wild-type calculations never emerges during the mole calculations on the reduced

enzyme. The gap entrance to the cavity is a snug, but large enough opening of 6.8 Å. R456, that binds the substrate, actually swings inward into the cavity in what could be a way to further direct the substrate into the cavity. It is conceivable, that the substrate is not released by R456 until entering the internal cavity. It is known that the hydrolysis of P5C to GSA happens within the cavity; however the channel becomes quite narrow for such a reaction to occur. Perhaps being anchored in place upon entrance to a water filled cavity allows for the hydrolysis to happen, and allows release from the R456 to channel to the P5CDH domain. Further, if the released product travels through the ion pair instead of moving around, there is less risk of solvent exposure as the movement of the R456 is towards that of the flexible C-terminal flap.

6.4.2 E197A

Mutation of R456 is known to kill enzyme activity. However, it was unclear what the outcome would be if removing the glutamate from the ion pair. E197 does not directly interact with substrate; however, as part of the proposed gate, it had the potential to disrupt activity completely. E197A was made to test how the loss of the ion pair, but not the residue responsible for substrate binding would affect the overall activity. According to the kinetics, enzyme function is not lost but severely hampered by the removal of the glutamate side chain. Guess and inferences from the known structures of BjPutA can explain these kinetics as removing the interaction that stabilizes the needed arginine for substrate binding. It is possible that with the loss of the gate the substrate could more freely enter the cavity or leave the active site, making diffusion a larger obstacle to overcome during substrate binding. Thus, R456 becomes more flexible and the binding

affinity for proline becomes worse as substrate has to enter and interact at the right time for the arginine to be in position.

Interestingly, the structure of this mutant had large and unpredicted changes in it. R456 is actually interacting with a sulfate bound in the active site. This side chain occupies a similar space as the wild-type enzyme and would be well within hydrogen bonding distance of the glutamate 197 side chain if it were present. More intriguing is the effect that this mutation had on helix 8. Helix 8 is home to both of the important arginine residues, 456 and 457, which bind the substrate. It is also where glutamate 460 is located, which forms an ion pair with R457 when a ligand is bound. Arginine 457 in particular now adopts two different conformations, one that models the wild-type and another that creates a new ion pair with the previously disordered glutamate 59.

However, starting at leucine 458, the structure of the alpha helix is disrupted. E460 swings into the active site and makes a close contact with the sulfate bound, at about 2.8 Å. It is unclear why two negative charges of oxygen would want this close interaction. N461 moves into the space previously occupied by glutamate 460. It is the movement of this residue in particular that allows for E59 to order and form a salt bridge with R457. Glycine 461 makes the largest shift of 6.5 Å, and A463 while only shifting 1.5 Å. The unstructuring of the loop with such large conformational changes was unexpected. Helix 8 is already known to be mobile and flexible, seen in GsPutA to make large shifts between the relaxed and substrate bound states. With the loss of E197, perhaps this flexibility allowed the enzyme to adopt a less favorable conformation to allow enzyme activity to continue. The E460 swinging into the active site might be a reflection of the need for the enzyme to adapt.

6.4.3 E460A

After seeing the importance in this residue in pairing with R457 during substrate binding, it seemed worthwhile to compare the disruption of both ion pairs through the glutamate to alanine mutations. In this case, mutation of E460A severely hinders the activity of the enzyme. Kinetic parameters could not be established as activity was very low and hardly above noise levels below 1 M proline. Removal of this residue just about kills the enzyme.

The PRODH active site looks very similar to that of the wild-type structure with sulfate bound. There is density for the sulfate, R457 occupies the same orientation as the wild-type and E197 is in a similar orientation as the wild-type, well within range to ion pair with R456. The guanidinium group of R456, however, is disordered. It was unexpected to see this arginine disordered and a sulfate in the active site. This could be significant as the kinetics are clear in poor enzyme activity. The absence of only the guanidinium ion, however, most likely means that this is an artifact.

Helix 8 was a different story. After seeing the unstructuring effects that the E197A mutant had at the C-terminal of the helix, a similar effect was anticipated for the E460A mutant. However, instead of interrupting the helix formation the same residues, 459-464 are completely disordered. Disordering of these residues indicates an increase in flexibility. The side chain of E460 was already known to be mobile as density is generally poor. The disordering of surrounding residues, however, shows that E460 not only plays an important role in substrate binding via a salt bridge, but also anchors helix 8 into the right location.

As compared with the E197A that has an unstructured but ordered C-terminal to the helix and lower but still measurable activity, it would seem to indicate that the helix 8 being ordered is needed for enzyme activity. The glutamate half of the salt bridge, while important, is less important than an intact helix 8. It is unclear why the truncation of the side chain of E460 would cause such a severe disordering of the surrounding helix. Due to alanine's propensity for forming alpha helices(35), the complete disruption with the introduction of this mutation was surprising. With the wild-type E460 making no clear and important hydrogen bonds there seems to be a lack of explanation why the mutation caused such a severe change.

6.4.4 R456M

The two structures of R456M, glycerol cryoprotected and proline soaked yielded very similar results. The mutation of the arginine yields a dead enzyme. As a result the gate is already oriented in an open position. With the PRODH domain, the proline soak does not have the same effect as it does on the catalytically active wild-type. During the proline soak the FAD, in crystal, never reduces as visible by the lack of the crystal changing color and the lack of FADH density in the structure.

The VOIDOO calculations on these structures yield very similar results as the proline soaked wild-type structure. VOIDOO finds its way through the proposed gate from the internal cavity, which is completely absent in the oxidized structure. The Mole calculations mimic those of the proline soaked wild-type structure as well. Mole finds its way through the internal cavity via the open gate. Put simply, R456 is an important residue for substrate binding and controlling flow of product into the channel. Without an

arginine present, not only is there no activity, but there is now a large hole where the gate used to segregate the PRODH active site and internal cavity.

6.5. Conclusion

In the PNAS paper describing the first full length structure of PutA, access to the internal cavity was hypothesized to happen via breaking the ion pair of R456 and E197A. Through mutation and ligand soaking the mechanism of this gate was probed. Indeed the ion pair is disrupted and the gate swings inward towards the cavity creating a gap of about 6.8 Å and connecting the internal cavity and the PRODH active site. Further the E197A and E460A mutations disrupting helix 8 points to the helix having an important role in catalytic activity and is likely sensitive to or involve in the gate opening. The movement of this gate allows for substrate to inter the channel without being exposed to bulk solvent, which is a key function for efficient channeling.

References

1. Tanner, J. J. (2008) Structural biology of proline catabolism, *Amino acids* 35, 719-730.
2. Srivastava, D., Schuermann, J. P., White, T. A., Krishnan, N., Sanyal, N., Hura, G. L., Tan, A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2010) Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*, *Proceedings of the National Academy of Sciences of the United States of America* 107, 2878-2883.
3. Ovadi, J. (1991) Physiological significance of metabolic channelling, *Journal of theoretical biology* 152, 1-22.
4. Easterby, J. S. (1981) A generalized theory of the transition time for sequential enzyme reactions, *The Biochemical journal* 199, 155-161.

5. Tralau, T., Lafite, P., Levy, C., Combe, J. P., Scrutton, N. S., and Leys, D. (2009) An internal reaction chamber in dimethylglycine oxidase provides efficient protection from exposure to toxic formaldehyde, *The Journal of biological chemistry* 284, 17826-17834.
6. Leys, D., Basran, J., and Scrutton, N. S. (2003) Channelling and formation of 'active' formaldehyde in dimethylglycine oxidase, *The EMBO journal* 22, 4038-4048.
7. Huang, X., Holden, H. M., and Raushel, F. M. (2001) Channeling of substrates and intermediates in enzyme-catalyzed reactions, *Annual review of biochemistry* 70, 149-180.
8. Arentson, B. W., Sanyal, N., and Becker, D. F. (2012) Substrate channeling in proline metabolism, *Front Biosci (Landmark Ed)* 17, 375-388.
9. Bearne, S. L., and Wolfenden, R. (1995) Glutamate gamma-semialdehyde as a natural transition state analogue inhibitor of Escherichia coli glucosamine-6-phosphate synthase, *Biochemistry* 34, 11515-11520.
10. Bearne, S. L., Hekmat, O., and Macdonnell, J. E. (2001) Inhibition of Escherichia coli CTP synthase by glutamate gamma-semialdehyde and the role of the allosteric effector GTP in glutamine hydrolysis, *The Biochemical journal* 356, 223-232.
11. Thoden, J. B., Huang, X., Raushel, F. M., and Holden, H. M. (1999) The small subunit of carbamoyl phosphate synthetase: snapshots along the reaction pathway, *Biochemistry* 38, 16158-16166.
12. Farrant, R. D., Walker, V., Mills, G. A., Mellor, J. M., and Langley, G. J. (2001) Pyridoxal phosphate de-activation by pyrroline-5-carboxylic acid. Increased risk of vitamin B6 deficiency and seizures in hyperprolinemia type II, *The Journal of biological chemistry* 276, 15107-15116.
13. Surber, M. W., and Maloy, S. (1998) The PutA protein of Salmonella typhimurium catalyzes the two steps of proline degradation via a leaky channel, *Archives of biochemistry and biophysics* 354, 281-287.
14. Singh, H., Arentson, B. W., Becker, D. F., and Tanner, J. J. (2014) Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-

channeling tunnel and the quinone-binding site, *Proceedings of the National Academy of Sciences of the United States of America* 111, 3389-3394.

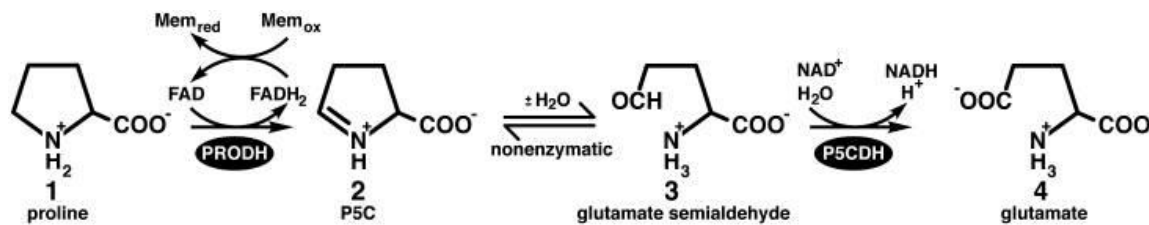
15. Moxley, M. A., Sanyal, N., Krishnan, N., Tanner, J. J., and Becker, D. F. (2014) Evidence for hysteretic substrate channeling in the proline dehydrogenase and Delta1-pyrroline-5-carboxylate dehydrogenase coupled reaction of proline utilization A (PutA), *The Journal of biological chemistry* 289, 3639-3651.
16. Dunn, M. F., Niks, D., Ngo, H., Barends, T. R., and Schlichting, I. (2008) Tryptophan synthase: the workings of a channeling nanomachine, *Trends in biochemical sciences* 33, 254-264.
17. Mouilleron, S., Badet-Denisot, M. A., and Golinelli-Pimpaneau, B. (2008) Ordering of C-terminal loop and glutaminase domains of glucosamine-6-phosphate synthase promotes sugar ring opening and formation of the ammonia channel, *Journal of molecular biology* 377, 1174-1185.
18. Mouilleron, S., Badet-Denisot, M. A., Badet, B., and Golinelli-Pimpaneau, B. (2011) Dynamics of glucosamine-6-phosphate synthase catalysis, *Archives of biochemistry and biophysics* 505, 1-12.
19. Luo, M., Arentson, B. W., Srivastava, D., Becker, D. F., and Tanner, J. J. (2012) Crystal structures and kinetics of monofunctional proline dehydrogenase provide insight into substrate recognition and conformational changes associated with flavin reduction and product release, *Biochemistry* 51, 10099-10108.
20. Krishnan, N., and Becker, D. F. (2005) Characterization of a bifunctional PutA homologue from *Bradyrhizobium japonicum* and identification of an active site residue that modulates proline reduction of the flavin adenine dinucleotide cofactor, *Biochemistry* 44, 9130-9139.
21. Schuermann, J. P., White, T. A., Srivastava, D., Karr, D. B., and Tanner, J. J. (2008) Three crystal forms of the bifunctional enzyme proline utilization A (PutA) from *Bradyrhizobium japonicum*, *Acta crystallographica. Section F, Structural biology and crystallization communications* 64, 949-953.
22. Bergfors, T. (2003) Seeds to crystals, *Journal of structural biology* 142, 66-76.
23. Pemberton, T. A., Still, B. R., Christensen, E. M., Singh, H., Srivastava, D., and Tanner, J. J. (2012) Proline: Mother Nature's cryoprotectant applied to protein

- crystallography, *Acta crystallographica. Section D, Biological crystallography* 68, 1010-1018.
24. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, *Acta crystallographica. Section D, Biological crystallography* 67, 271-281.
 25. Evans, P. (2006) Scaling and assessment of data quality, *Acta crystallographica. Section D, Biological crystallography* 62, 72-82.
 26. Kabsch, W. (2010) Xds, *Acta crystallographica. Section D, Biological crystallography* 66, 125-132.
 27. Adams, P. D., Afonine, P. V., Bunkoczi, 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 crystallographica. Section D, Biological crystallography* 66, 213-221.
 28. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallographica Section D* 60, 2126-2132.
 29. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta crystallographica. Section D, Biological crystallography* 66, 12-21.
 30. Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta crystallographica. Section D, Biological crystallography* 50, 178-185.
 31. Petrek, M., Kosinova, P., Koca, J., and Otyepka, M. (2007) MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels, *Structure* 15, 1357-1363.
 32. Berka, K., Hanak, O., Sehnal, D., Banas, P., Navratilova, V., Jaiswal, D., Ionescu, C. M., Svobodova Varekova, R., Koca, J., and Otyepka, M. (2012) MOLEonline

2.0: interactive web-based analysis of biomacromolecular channels, *Nucleic acids research* 40, W222-227.

33. Becker, D. F., and Thomas, E. A. (2001) Redox properties of the PutA protein from *Escherichia coli* and the influence of the flavin redox state on PutA-DNA interactions, *Biochemistry* 40, 4714-4721.
34. Pemberton, T. A., and Tanner, J. J. Structural basis of substrate selectivity of Δ 1-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): Semialdehyde chain length, *Archives of biochemistry and biophysics*.
35. Chakrabarty, A., Kortemme, T., and Baldwin, R. L. (1994) Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions, *Protein science : a publication of the Protein Society* 3, 843-852.

SCHEMES



Scheme 6.1. The overall 4 step oxidation of proline catabolism to glutamate. Proline dehydrogenase catalyzes the oxidation of proline to Δ^1 -pyrroline-5-carboxylate (P5C) utilizing FAD as an electron acceptor. P5C goes through a nonenzymatic hydrolysis to glutamate γ -semialdehyde (GSA). The NAD-dependent P5C dehydrogenase (P5CDH) catalyzes the final oxidation of GSA to glutamate and NADH.

Table 6.1. Primers for Site-Directed Mutagenesis.

Mutant	Primers (5'-3')
E460A	Fwd: CGGCGACTGCTGGCGAACGGTGCCAAC Rev: GTTGGCACCGTTCGCCAGCAGTCGCCG
E197A	Fwd: CATGCTCGGCGCAGGCGCGCGCAC Rev: GTGCGCGCGCCTGCGCCGAGCATG

Table 6.2. X-ray Diffraction Data Collection and Refinement.

	BjPutA-Proline	BjPutA R456M	BjPutA R456M-proline	BjPutA E197A	BjPutA E460A
Beamline	ALS	ALS	ALS	ALS	ALS
Space group	C2	C2	C2	C2	C2
Unit cell parameters (Å, °)	$a = 167.8$ $b = 196.1$ $c = 109.0$ $\beta = 121.7$	$a = 167.0$ $b = 195.8$ $c = 108.7$ $\beta = 121.5$	$a = 167.16$ $b = 195.20$ $c = 109.19$ $\beta = 121.46$	$a = 167.3$ $b = 196.0$ $c = 109.0$ $\beta = 121.4$	$a = 166.5$ $b = 195.8$ $c = 108.7$ $\beta = 121.4$
Wavelength	1.0000	1.0000	1.0000	1.0000	1.0000
Resolution (Å)	37.84 – 2.20 (2.32 – 2.20)	29.89 – 2.30 (2.42 – 2.30)	29.96 – 2.03 (2.14 – 2.03)	62.15 – 2.10 (2.14 – 2.10)	61.97 – 2.30 (2.34 – 2.30)
Observations	461836	500901	708040	642493	486982
Unique reflections	138636	131581	190232	171782	130904
$R_{merge}(I)$	0.091 (0.480)	0.132 (0.495)	0.124 (0.589)	0.086 (0.611)	0.072 (0.509)
$R_{meas}(I)$	0.109 (0.648)	0.154 (0.577)	0.146 (0.723)	0.119 (0.839)	0.099 (0.698)
$R_{pim}(I)$	0.058 (0.430)	0.079 (0.296)	0.076 (0.410)	0.082 (0.572)	0.068 (0.476)
Mean I/σ	8.4 (2.0)	11.4 (2.4)	6.5 (2.5)	9.4 (1.8)	11.6 (2.1)
Completeness (%)	91.7 (63.0)	100.0 (100.0)	99.1 (93.9)	98.8 (98.3)	99.7 (99.8)
Multiplicity	3.3 (2.1)	3.8 (3.8)	3.7 (3.2)	3.7 (3.5)	3.7 (3.5)
No. of protein chains	2	2	2	2	2
No. of protein residues	1938		1941		
No. of protein atoms	14631	14477	14681	14372	14285
No. of FAD atoms	106	106	106	106	106
No. of proline atoms	80	0	96	0	0
No. of water molecules	706	541	980	692	
R_{cryst}	0.2300 (0.3559)	0.2066	0.2217 (0.3856)	0.1926	0.2238
R_{free}^b	0.2681 (0.3693)	0.2459	0.2634 (0.4165)	0.2239	0.2618
Rmsd bond length, Å ^c	0.010	0.007	0.007	0.009	0.009
Rmsd bond angles, °	1.392	1.082	1.046	1.080	1.145
Ramachandran plot ^d					
Favored (no. residues)	1883		1891		
Allowed (no. residues)	49		38		
Outliers (no. residues)	6		12		
Average B-factor (Å ²)					
Protein	28.14	31.21	26.72	35.15	43.315
FAD	23.92	21.90	18.65	26.47	37.017
Proline	42.09	-	45.71	-	-
water	37.75	29.32	28.86	36.77	
Coordinate error (Å) ^e	0.35		0.33		
PDB code					

Figure Legends

Figure 6.1. Reduced FAD electron density from proline soaked BjPutA. A) In yellow is the reduced FAD seen in the proline reduced BjPutA structure. The mesh represents an F_o-F_c simulated annealing omit map. B) View of the butterfly bend of the isoalloxazine ring seen in the reduced structure.

Figure 6.2. View of proline (pink) found in the P5CDH active site. NAD (cyan) sits above the dual conformation cysteine 792. Proline sits on the other side of the active site, with the pyrroline ring positioned between phenylalanine 659 and 954. The carboxyl group of proline occupies a similar location as sulfate found in 3HAZ, interacting with the backbone amino group of A947 and G946 and the guanidinium ion of R791.

Figure 6.3. Movement of the R456-E197 ion pair gate. A) The blue surface represents the MOLE calculation starting at the sulfate for the proline reduced BjPutA structure. B) The pink surface represents the MOLE calculation starting at the sulfate for the BjPutA wild-type (3HAZ) structured. C) The superposition of the wild-type (gray) and proline-reduced (green) ion pair, with the proline reduced MOLE tunnel (blue) present.

Figure 6.4. VOIDOO calculation comparisons of the oxidized and reduced form of the wild-type enzyme. A) Shows the PRODH domain and the surface represents the VOIDOO calculation for the original 3HAZ structure. The side chains of the ion pair for the wild-type are colored cyan and FAD is colored yellow. B) The red mesh represents the extra space found during the VOIDOO calculation overlaid with that of the 3HAZ

VOIDOO calculation (surface). The gray side chains now represent the proline-reduced ion pair. C) View from the internal cavity into the PRODH domain, clearly showing the gate preventing VOIDOO from entering the active site.

Figure 6.5. A) The oxidized gate (green) superimposed with the glycerol cryoprotected R456M structure (gray) and the proline soaked (pink) R456M structure. The surface represents the VOIDOO calculated for the proline-soaked wild-type enzyme. B) The proline soaked wild-type structure (cyan) superimposed with the glycerol cryoprotected R456M structure (gray) and the proline soaked R456M structure (pink). The surface represents the VOIDOO calculated for the proline soaked wild-type structure.

Figure 6.6. View of the movement of glycine 196 and glutamate 197. The 3HAZ structure (green) superimposed with the proline soaked structure (gray). The dark blue mesh represents simulated annealing σ_A -weighted $F_o - F_c$ omit maps contoured at 3.0σ .

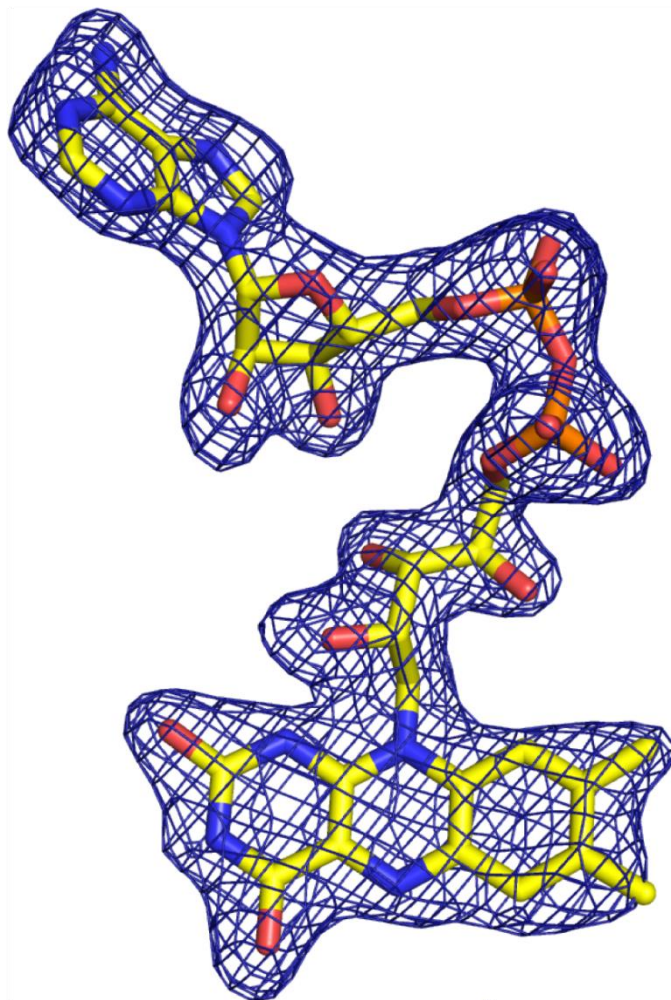
Figure 6.7. A) Superposition of 3HAZ (gray) and the E197A (cyan) structures highlighting the unstructuring effect that the E197A mutant has on helix 8. B) The dark blue mesh represents simulated annealing σ_A -weighted $F_o - F_c$ omit maps contoured at 3.0σ .

Figure 6.8. Superposition of 3HAZ (gray) and the E197A mutant (cyan) structures highlighting residues involved in the unstructured of helix 8. A) All residues experiencing a significant movement in helix 8 in the E197A mutation. B) R457 dual conformation and ordering of the side chain of E59 in the mutant. Shift of the E460 residue (C) and the N461 (D).

Figure 6.9. The Michaelis-Menten curve for the E197A structure. The Y-axis represents the V_{max} while the X-axis plots the proline concentration used.

Fig 6.1.

A



B

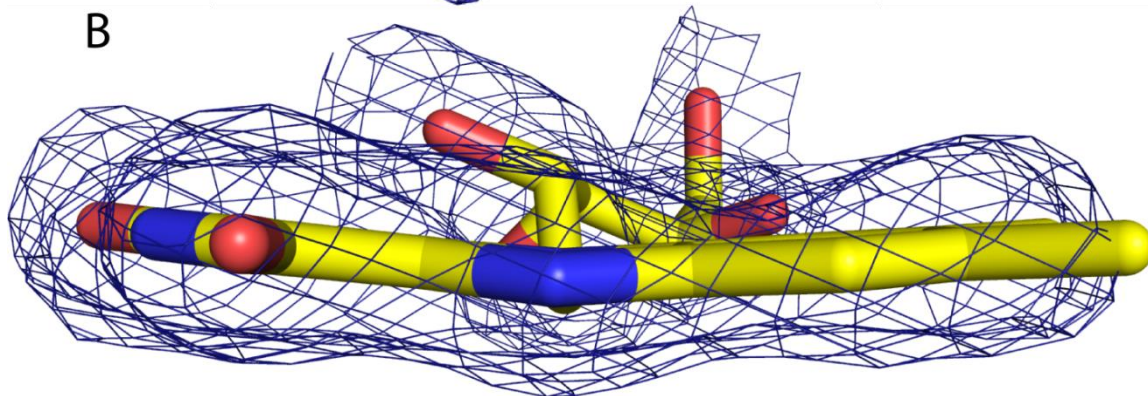


Fig 6.2.

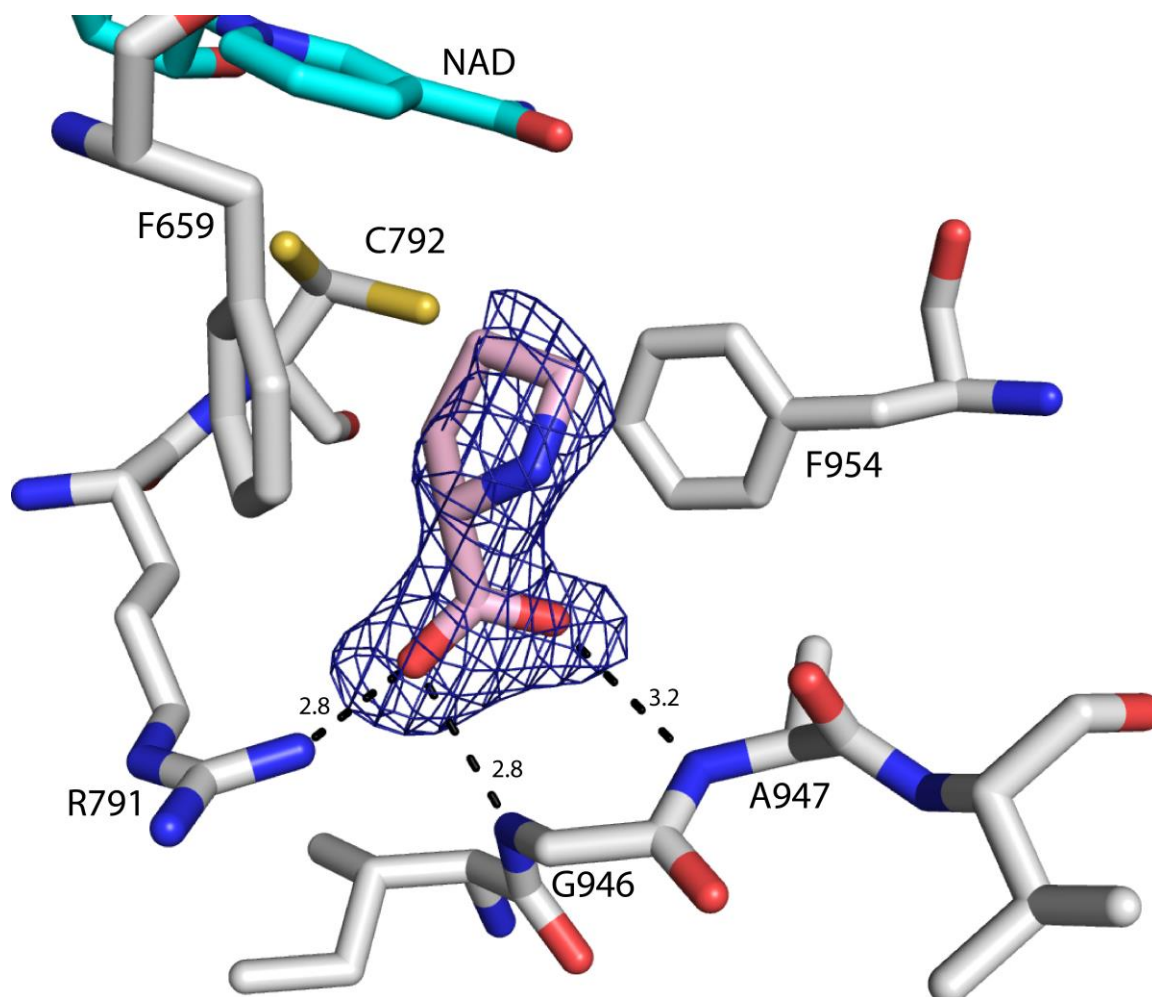


Fig 6.3.

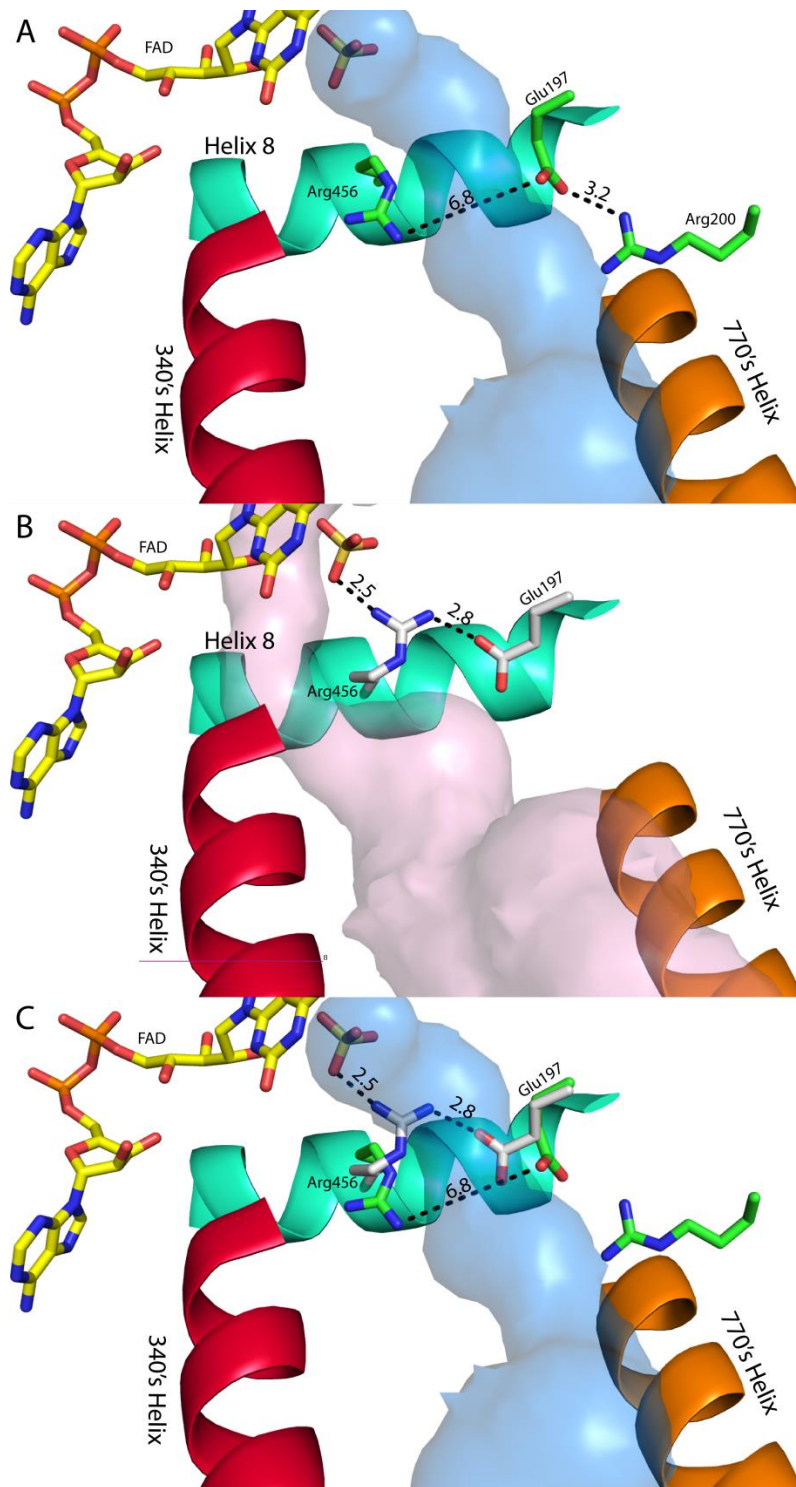


Fig 6.4.

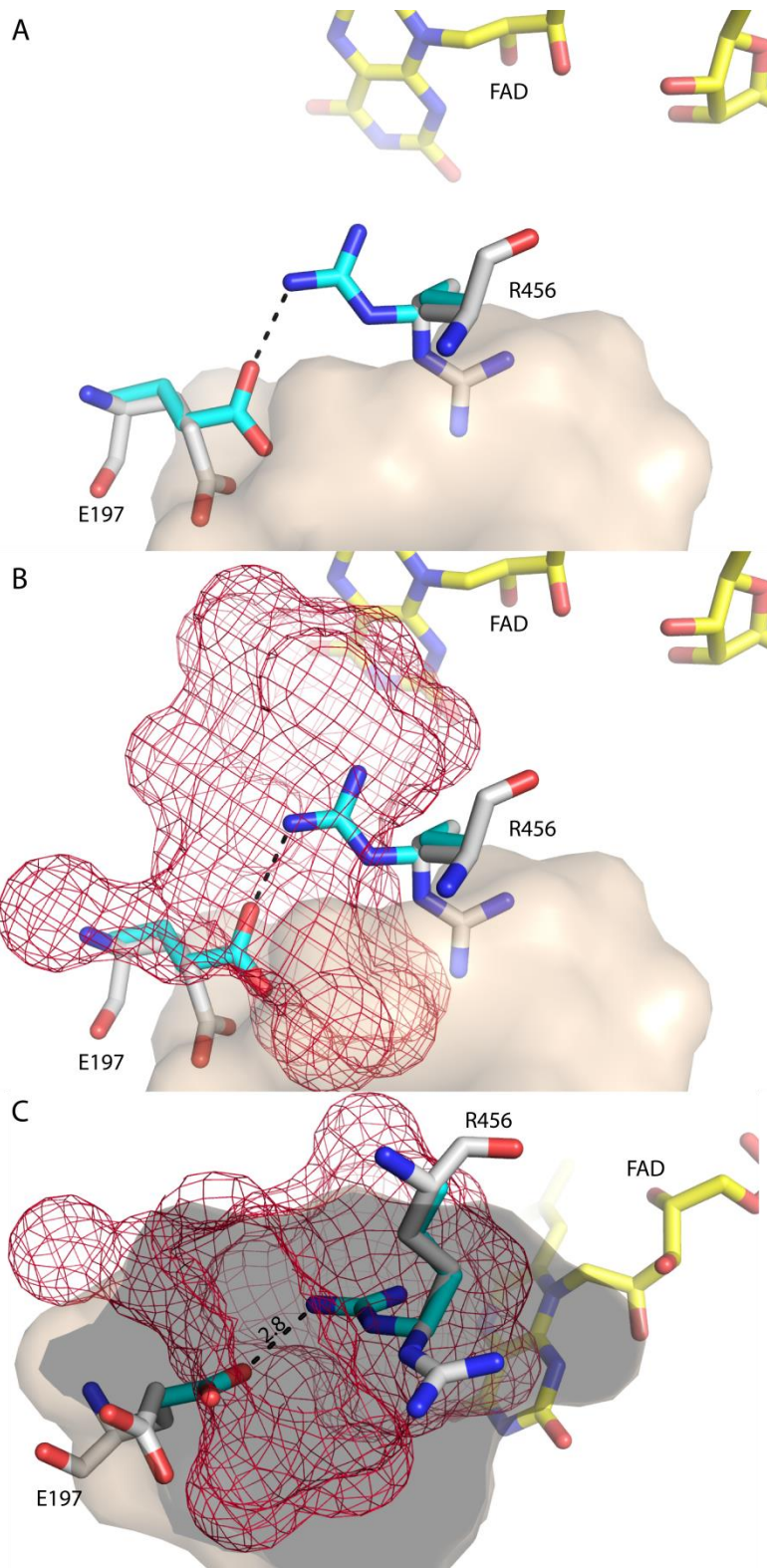


Fig 6.5.

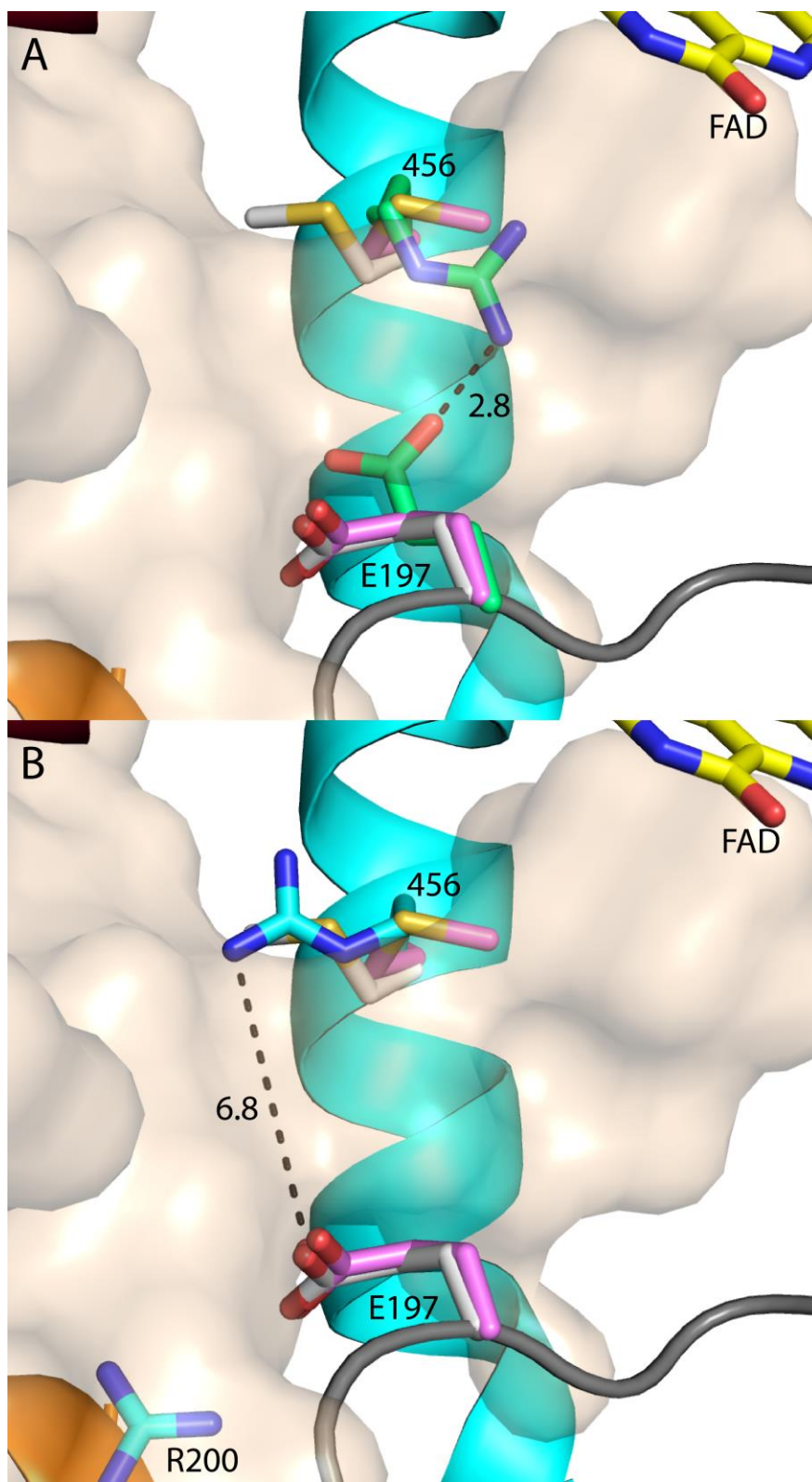


Fig 6.6.

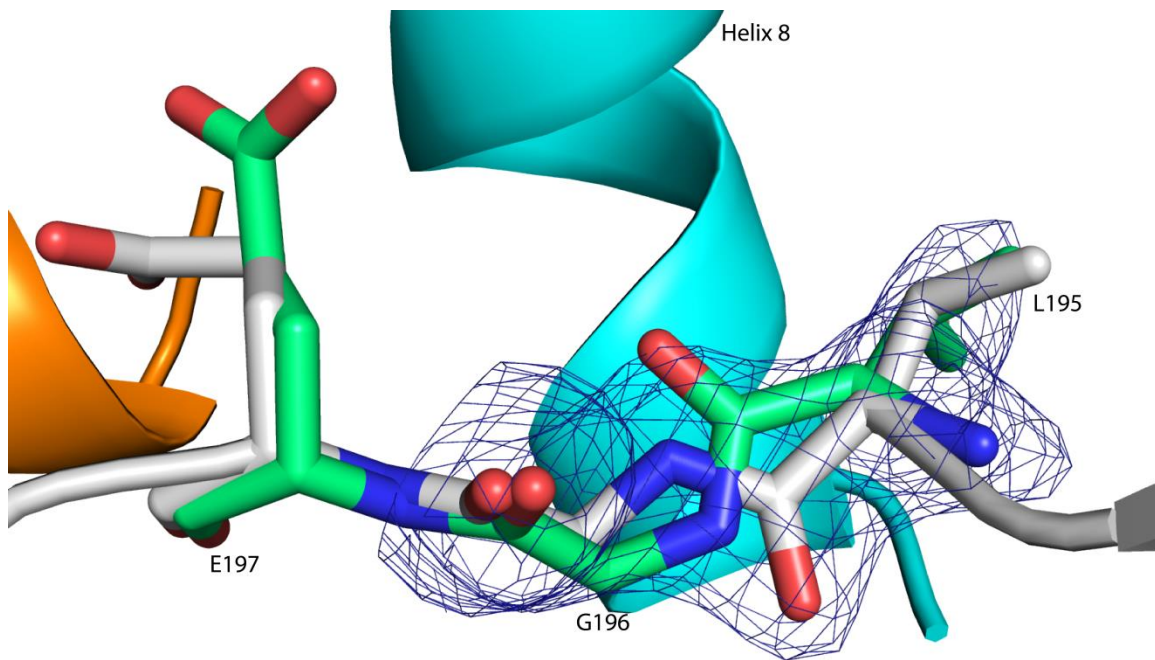


Fig 6.7.

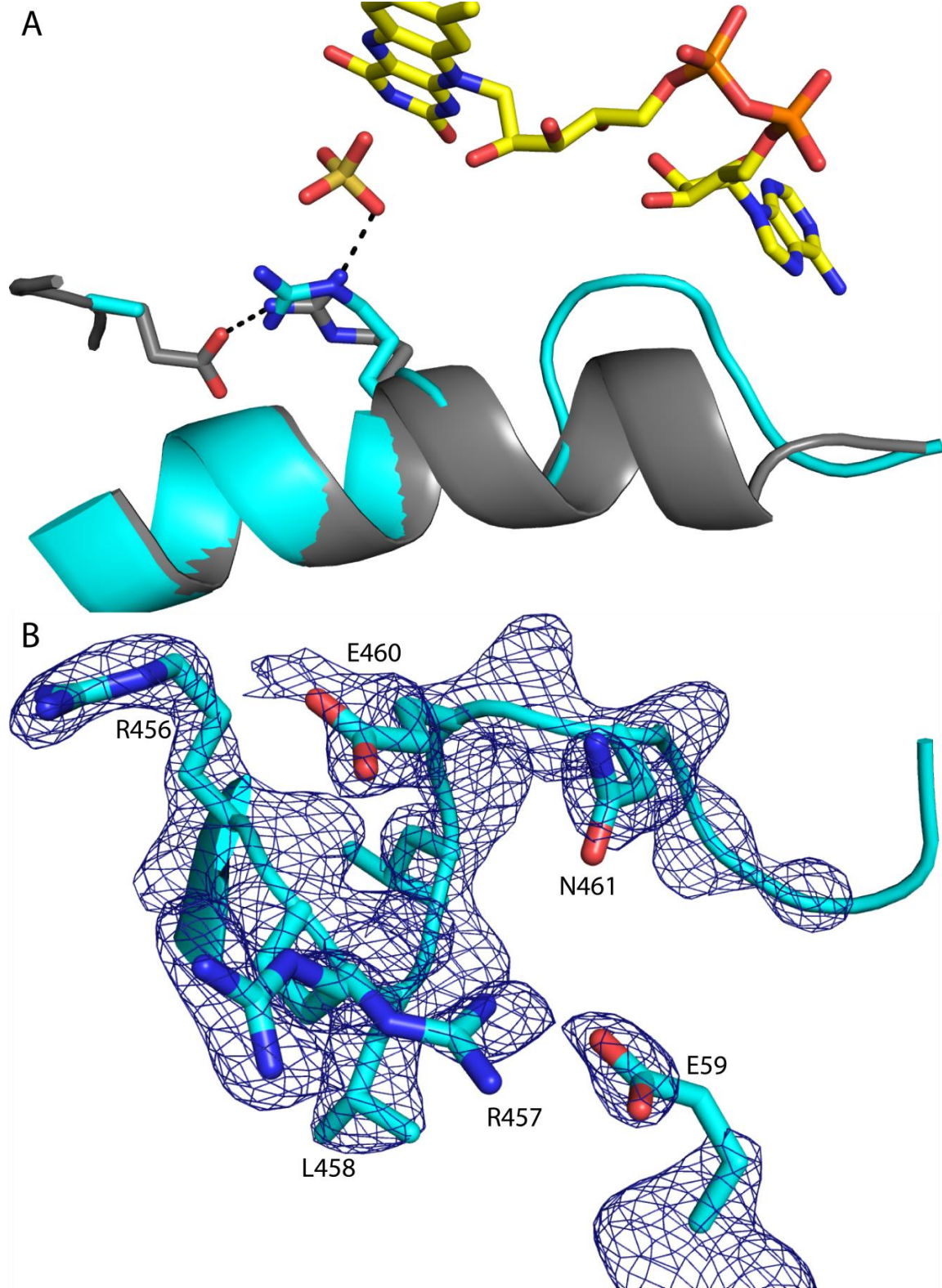


Fig 6.8.

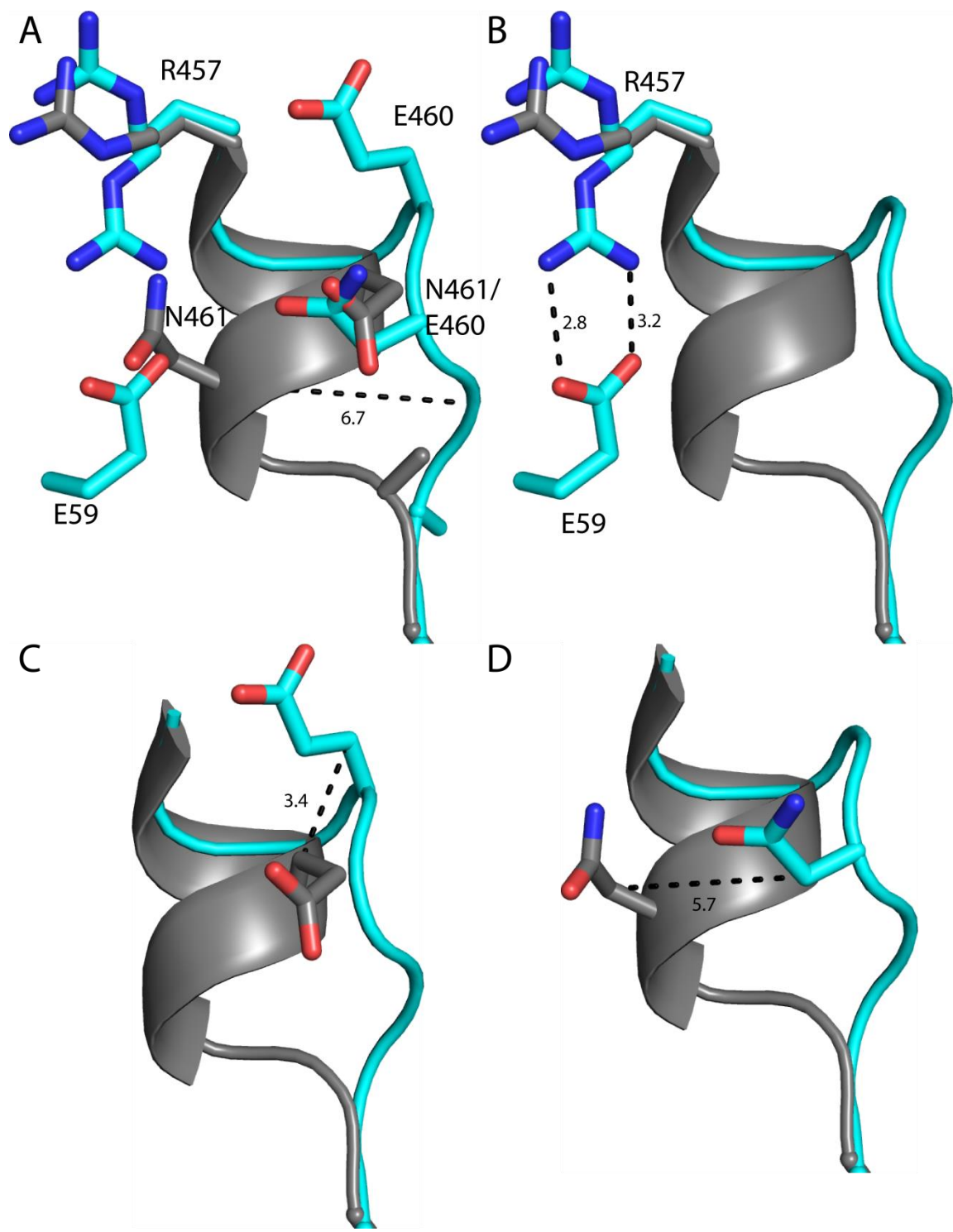
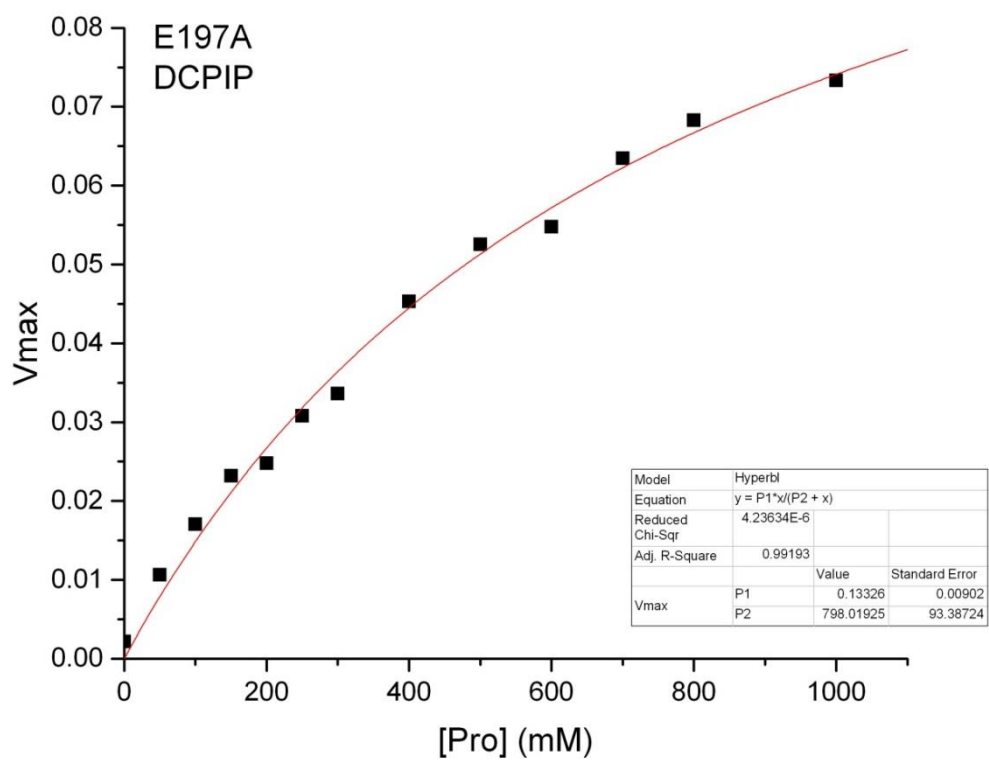


Fig 6.9.



Chapter 7:
Structure and Role of Mg²⁺ Binding of the
Apoptosis-Linked Protein (ALG-2)

Abstract

Apoptosis-linked gene protein (ALG-2) is a penta-EF-hand calcium binding protein that induces apoptosis during high calcium concentrations in cells. Structures of this enzyme in the calcium bound, zinc bound, calcium free, and Alix bound ALG-2 have been solved previously. It was previously thought that these binding sites do not bind magnesium; however, recent studies show that this enzyme binds magnesium, albeit at a much lower affinity than calcium. This study aims to elucidate the structure of the magnesium bound ALG-2. However, it also highlights some of the problems not always seen in crystallography. Specifically, how do you distinguish between the electron density of different possible ions in the binding site? This question came about when close inspection of the “calcium-free” structure revealed that sodium ions were bound in all of the binding sites. This chapter outlines the methods used to try to answer this question and lays out further work that could be done to answer this question.

7.1. Introduction

Apoptosis-linked gene 2 (ALG-2) is a 22-kDa penta-EF-hand (PEF) protein, characterized by the presence of five EF-hands which are helix-loop-helix Ca^{2+} binding motifs(1). In ALG-2 only three of these sites are thought to be functional; EF1, EF3, and EF5. EF1 and 3 are considered high affinity sites, binding Ca^{2+} at micromolar levels(2, 3). The fifth site, EF5, is located in the C-terminal and is the oligomerization domain, interacting with another EF5 to form a homodimer, and is considered a low affinity site. Dimer formation, however, happens in a Ca^{2+} -independent manner(4). Canonical EF-hands have a calcium ion with pentagonal bipyramidal coordination with five ligands on the yz plane and two ligands on the x axis. Five of the coordination spots are occupied with the oxygens of side chains (x , y , z and a bidentate $-z$). A sixth is from a peptide carbonyl at position $-y$ and the seventh is from a water molecule located at $-x$ (5). These enzymes interact Ca^{2+} -dependently with a multitude of proteins with proline-rich areas.

In this family, ALG-2 and peflin belong to group 1 PEF proteins. These vary from group 2 PEF proteins, calpain and grancalcin, with variations within the EF1 and EF5 binding sites. Group 2 enzymes have a one-residue deletion within the EF1 site and have a substitution of the x site with from asparagine to alanine, while group 1 does not have this. Group 2 also has an additional two residues present in the EF5 binding site. Additionally, these enzymes have an N-terminal proline/glycine rich region, with ALG-2 having the shortest of these, about 22 residues. Removal of this N-terminal region has led to an increased crystallization chance, with many of the deposited structures having this region removed(5).

Upon Ca^{2+} binding in EF3, conformation changes expose a hydrophobic region of ALG-2 allowing for the enzyme to interact with proline-rich regions of proteins. One protein in particular, Alix or AIP1, is a 95 kDa peptide that has three proline-rich regions at the N-terminal, middle, and C-terminal. The C-terminal region interacts with ALG-2 at a region containing four PxY repeating motifs. Mutation of either the proline or tyrosine to alanine results in a loss of binding of Alix to ALG-2(6-8).

Crystal structures of mouse and human ALG-2 have been solved(7, 9). The human structures, in particular, have a Ca^{2+} -bound, Ca^{2+} -unbound, Zn^{2+} -bound, and an alix-ALG-2 complex structure. Comparisons between the bound and unbound structures provide important insights into the understanding of the binding. In short, binding of calcium to EF3 induces conformational changes in R125, exposing a hydrophobic pocket and allowing for the proline-rich, C-terminal area of the Alix peptide to bind(7). These comparisons between the calcium bound and the non-calcium bound structures, however, are not entirely accurate. Upon inspection of the Ca^{2+} -free form of the enzyme, there is a Na^+ bound in EF1, EF3, and EF5. In the Zn^{2+} -bound structure, EF5 is occupied not by Zn^{2+} but Na^+ as well. The issue of the presence of sodium in all of the binding pockets is not addressed, but could be an indication of why dimer formation is Ca^{2+} -independent.

ALG-2 is only thought of as Ca^{2+} -activated, which has low concentrations in vivo compared with another divalent cation, Mg^{2+} . An increase in Ca^{2+} concentration in the cell inducing apoptosis regulated by ALG-2 works as long as Mg^{2+} doesn't also bind and activate ALG-2. Binding studies done by Dr. Henzl have shown that ALG-2 does have an affinity for Mg^{2+} , and while not as high as it is for Ca^{2+} , it is not insignificant. This study aims to understand the differences in the structures of Ca^{2+} -bound and Mg^{2+} -bound ALG-

2 and why only one of these cations leads to programmed cell death. This study focuses heavily on the crystallization of enzyme devoid of Ca^{2+} , with only the presence of Mg^{2+} .

7.2. Materials and Methods

7.2.1 Protein Preparation for Crystal Trays.

Protein was given to me in a buffer containing 150 mM NaCl, 25 mM HEPES at pH 7.8, 5 mM EGTA, and 1 mM MgCl_2 . The ALG-2 was already frozen and at a concentration of 10 mg/mL. The protein was thawed using two different methods. The first was holding the epitube under water at room temperature to unthaw it quickly. The second method, the one more commonly used, was to place the epitube containing the protein on ice, and let it thaw slowly overtime.

The thawed ALG-2 at room temperature was clear, regardless of how it was unthawed. However, when storing the enzyme at 4°C , a white precipitate formed in the solution. When the enzyme was left at room temperature, after this this precipitate appeared, the precipitate disappeared, and reappeared again when stored at 4°C .

7.2.2 Crystallization Trials

Based off of previous crystallization of the Ca^{2+} and the Ca^{2+} -free form, two conditions were tested to attempt to reproduce crystals. First, a condition containing 14-24% PEG 8,000, 0.2 M magnesium acetate (MgOAc), and 0.1 M cacodylate pH 5.5-7.0 was attempted. The wells were set up in a 1:1 ratio with 10 and 5 mg/mL protein solutions being used. Trays were set up later after increasing the PEG 8,000 concentration to 26-36%. The second condition tried was 6-16% isopropanol, 0.2 M MgOAc , and 0.1 M cacodylate at pH 5.5-7.0. The wells were again set up with a 1:1 μL

ratio using 10 and 5 mg/mL enzyme. A second set of trays was set up, increasing the isopropanol concentration range to 18-28%. None of these optimizations yielded crystals.

Crystal screening using Index and Crystal Screen (CS) I and II was done in a 96 well plate. Reservoir size was 100 μ L with a drop size of 1 μ L of enzyme to 1 μ L reservoir. Screens were set up at both 10 and 5 mg/mL at room temperature. Initial possible hits were seen in CS II 43, CS II 10, CS II 19, CS I 1, and CS I 14. These were promising as 3 of the conditions contained MPD, a precipitant reported previously. However, none of these conditions were reproducible. Crystal screens were then set up using Index and CSI and II under Al's Oil. These were set up by adding 1 μ L of enzyme to the well followed by 8 μ L of Al's Oil, and finally 1 μ L of the precipitant solution. These trays were wrapped with parafilm and incubated at room temperature. None of these screens yielded crystals. Screening took place again in 96 well plates only this time set up at 4°C. These trays were set up the exact same as the 96 well trays at room temperature except the trays were placed in the cold room ahead of time and allowed to cool down for about 10 minutes before setting up the trays. This time there was a hit in 30% isopropanol and 0.1 M Tris pH 7.5.

Optimization of this hit was set up with a range of 22-32% isopropanol and 0.1 M Tris at pH 7.0-8.5. These trays yielded very large, very thin clear crystals that tended to cluster together, referred to as Form 1 (Figure 1). A second batch of these trays was set up using 22-32% isopropanol, 0.1 M Tris at pH 7.0-8.5, and 0.2 M $MgCl_2$. These crystals, Form 2, were short, thick rod shaped crystals (Figure 2). Both crystal forms took about 2-3 days to grow to maximum size.

7.2.3 Cryoprotection

Cryoprotecting these crystals proved difficult. The first crystals shot were cryoprotected in about 5% MPD, 30% isopropanol, 0.1 M Tris pH 7.5 and 0.2 M MgCl_2 (salt only for Form 2). Any concentration above 5% cracked and destroyed these crystals and ice ring diffraction was still present.

Several problems presented themselves upon review of the cryoconditions. First, isopropanol has a tendency to evaporate quickly while cryoprotecting. Evaporation of the precipitant could lead to problems cracking of the crystals. Second, these crystals were very temperature sensitive, with crystals breaking down into just precipitate after being at room temperature for a few minutes. Finally, it could very well be the problem of the chosen cryoprotectant not being suitable for this condition.

The first of these problems tested was that of simply changing the cryoprotectant. First, instead of adding a new molecule as a cryoprotectant, increasing the concentration of isopropanol might work. Loop tests were performed by dipping a loop in the prepared cryoprotectant and flash frozen in the X-stream of gaseous nitrogen. At 40% isopropanol, the appearance of ice seemed to be removed, leaving instead a waxy-looking drop on the pin. However, upon exposing the 40% isopropanol mixture to X-rays, ice ring diffraction was still present. Second, 2 M proline was tested with 30% isopropanol. This was indeed cryoprotective with no ice ring diffraction present. At low concentrations of proline, both forms of the crystals appeared to be stable. However, upon increasing the concentration of proline to 2 M, the crystals completely dissolved. Sadly, at 1 M proline, where the crystals were stable, the solution was not cryoprotective. Finally, the cryoprotectant that

worked was 25% glycerol with 30% isopropanol, 0.1 M Tris at pH 7.5, and either 1mM MgCl₂ or 200mM MgCl₂.

This cryoprotectant was chilled on ice before use. Chilling was done to try to remove temperature-dependent degradation of the crystals. Then, 20 μ L of the full 25% cryoprotectant was directly used in the well. The soak time was long enough to allow the Schlerian lines to settle and then crystals were looped, flash frozen, and stored in liquid nitrogen.

7.2.4 Data Collection and Processing

The first data set (Table 1) was collected at ALS beamline 4.2.2. 180 degrees was collected with an exposure time of 0.6 seconds, an oscillation width of 0.1 degrees, and a distance of 280 mm. The second data set was collected on the R-Axis IV++ home source and Rigaku detector. 170 degrees of data was collected with 20 minutes exposure times, 0.5 degree oscillation width, and a detector distance of 100 mm. Finally, the last two data sets were also collected at ALS beamline 4.2.2. 180 degrees of data was collected with an exposure time of 0.6 seconds, 0.1 degrees of oscillation, and a detector distance of 290 mm.

Data integration was performed using XDS(10). Aimless in the CCP4 package was used for space group determination and scaling(11). The spacegroup was determined to be $P2_122_1$ with a unit cell of 75, 54, and 48 Å. However, convention is to have the unique axis in the terminal position. Using Reindex program, in the CCP4 suite, a matrix transformation was performed and moved the space group to $P2_12_12$ and with the unit cell being 48, 75, and 54 Å. Molecular replacement was not needed because the Ca²⁺-free

model, previously published (2ZND), was in the same spacegroup and used for the R_{free} flags.

2ZND was used a model for refinement in PHENIX(12). Initial refinement was done with rigid body, TLS, NCS, and simulated annealing. The structures and electron density was viewed and adjusted in COOT(13). Significant density for something in all three of the Ca^{2+} binding sites (EF1, EF3, and EF5) was seen. Water was modelled into this density. A second refinement was performed without rigid body or simulated annealing. The newly refined structure was then plugged into the STAN server. According to WASP(14), on the STAN server, the three waters in the Ca^{2+} were determined to be either lithium or magnesium. In COOT, these waters were removed and replaced with Mg^{2+} and the occupancy was set to 0.5 and PHENIX refined the occupancy.

7.3. Results

7.3.1 Enzyme Precipitation

To begin, an interesting issue arises with the pure enzyme provided. After thawing the protein solution is clear at room temperature, but at 4°C a white precipitate forms. Since the crystals are grown at 4°C some trays were set up with the precipitate present and some where the precipitate had been spun off with a short spin in the centrifuge. Interestingly, the best crystals were grown from the enzyme that had the precipitate still present.

7.3.2 Diffraction Quality and Cryoprotection

Only Form 1 gave diffraction out to a reasonable resolution (Figure 3). Form 2 diffraction (Figure 4) was very poor out to only about 4 Å. However, upon initial diffraction it was unclear why Form 2 was diffracting poorly. With substantial ice rings

and the knowledge that the cryoprotectant caused the crystals to crack at concentrations higher than 5% MPD, the poor diffraction could have been a direct result of these subpar conditions for the crystal.

Form 1, while it had substantial ice rings, diffracted beyond 2.1 Å. Initially XDS could not process the data. However, after excluding the ice rings XDS processed the data and find the correct space group with no issue. The result was $P2_122_1$ space group with unit cell dimensions 48, 54, and 75 Å. However, convention is to put the unique axis at the terminal end. Reindex in the CCP4 suite can perform the transformation of (b,c,a) – (a,b,c), which puts the space group in the proper convention and the unit cell dimensions are now 75, 48, 54 Å. This is the same space group as was seen from the Ca^{2+} -free form (2ZND), with the exception of the a dimension becoming 4 Å longer.

Data collection and refinement statistics are listed in Table 1. The low completeness would be troubling if ice rings had not been excluded. However, because of the substantial ice ring presence, a large piece of the data needed to be excluded and a reduction in the completeness was seen. One perplexing result was that of a low amount of unique reflections, especially in the inner shell. Usually, there are quite a few of unique reflections in this bin, but it was as low as 771 reflections in this group of data sets.

Perhaps equally surprising is the unusually high I/σ . There are many checks to deciphering good data quality, and one common method for evaluating if the correct resolution range has been picked is seen in the mean I/σ . It is a general rule of thumb that I/σ is cut off at a value of 2.0, though the view of this has changed to believe it is too conservative of an estimate(15). However, when attempting to push to higher resolution

to get to the cutoff, the diffraction statistics become poor with high R_{merge} values and low completeness. It does not appear to be a result of twinning, as the tests for twinning came back fine and the structures looked great. It is worth noting that there was very intense diffraction, perhaps contributing to the observed high signal to noise ratio.

7.3.3 Metal Binding Sites

The binding sites, EF1, EF3, and EF5, all originally had positive $F_o - F_c$ map density. When comparing these binding sites with previously published structures, it was clear that any number of ions could be bound. The biggest concern was the presence of Na^+ in the binding sites of the Ca^{2+} - free form(7). With EGTA in the buffer, Ca^{2+} should not be a concern(16). However, Na^+ and Mg^{2+} are both present and with Na^+ binding evident in what structure it will be difficult to tell if it is truly Mg^{2+} bound, or which of the three binding sites it is bound in.

To help answer this question the STAN server was used. The STAN server specifically runs three different calculations, but the one of interest is called WASP. WASP will look at all the water molecules within a structure and give a list of waters that could possibly be either Li^+ , Na^+ , Mg^{2+} , or Ca^{2+} . Interestingly, the program does not seem to be able to distinguish between Li^+ and Mg^{2+} or Na^+ and Ca^{2+} . However, when replacing the Ca^{2+} and the Na^+ with water, WASP will say that those particular waters should be either calcium or sodium. As a result these binding sites were modelled with water and then run through refinement. These refined structures were then used in the STAN server, and WASP gave an output that said the waters in the binding pockets should actually be either Li^+ or Mg^{2+} . With no lithium present, magnesium was modelled in.

Refinement in PHENIX left the Mg²⁺ modelled into all three active sites left occupancy of about 60% for all three sites. Another interesting question arose from this; is it possible that the sites have a dual occupancy of Na⁺ and Mg²⁺? Or perhaps even a triple occupancy with water as the third molecule? With no definitive suggestion that a triple occupancy would be occurring, Occam's razor insists to throw this idea out. When modelling a dual occupancy, the answer does not become clear. Both ions go to an almost equal 50/50 split, with Mg²⁺ having a slightly higher occupancy, usually around 54-58%. However, with the B-factors being low for both the single occupancy and the dual occupancy structures, it is impossible to determine which the correct answer is. It is worth noting that when EF-3 was modelled as a dual occupancy of magnesium and sodium, the magnesium moves slightly out of the binding pocket making a very long (3.2 Å) interaction. However, when using a single occupancy of just magnesium, the ion moves back into the binding pocket and makes interactions with expected distances.

A summary of the coordination of the magnesium in comparison with that of sodium and calcium from previously solved structures is displayed in Table 2. Interestingly, the largest difference seen is that of glutamate 114 in EF3. In the case of the calcium bound structure, the glutamate coordinates with the ion at 2.3 and 2.5 Å. However, in the magnesium structure this glutamate is now much further away with a coordination of 4.3 and 4.8 Å, and this is seen in the sodium structure as well with the sodium being 3.7 and 4.3 Å from the two oxygens on the glutamate.

7.4. Discussion

Calcium in EF1 has pentagonal bipyramidal coordination, with long bond lengths to E47 and S40. Actually moves completely away from S40 and makes a closer

interaction with OE1 from E47. Magnesium seems to make all the same coordinations as calcium. For calcium, a coordination number of 7 (needed for pentagonal bipyramidal) is fairly common, about 30%, in this resolution range. Magnesium, on the other hand, only has a 7 coordination number less than 5% of the time, much preferring coordination of 5 and 6(17).

Calcium has a preferred coordination length of about 2.3-2.5 Å. Distances shorter than 2.2 Å are much less common and almost non-existent below 2.0 Å. Coordination bond lengths greater than 2.75 Å are uncommon. For EF1 this puts all of the coordination distances reported for the Ca²⁺ bound structure well within the expected ranges with only one unusual distance of 2.8 Å. Magnesium, however, has a preferred coordination bond length of 1.9-2.1 Å. The frequency of coordination lengths shorter than 1.8 Å is almost non-existent, and lengths greater than 2.5 Å become very rare. This gives the Mg²⁺ bound in the EF1 pocket 4 coordination bonds with lengths that fit perfectly in the expected range. However, two lengths 2.8 (OE2E47) and 3.0 Å (OS40) are rarely seen with magnesium coordination. The other interaction with E47 is at 2.3 Å, and while much more reasonable but still without the most common distances.

Further, bidentate coordination with either glutamate or aspartate for calcium with a coordination number of 7 is a common occurrence. In fact, bidentate coordination becomes increasingly more common with calcium as the metal binding with increasing coordination number. However, bidentate coordination is not as common with magnesium. With a coordination number of 7, which as stated earlier is rare for magnesium, bidentate coordination is less common than coordination with single amino

acid oxygens. However, with a smaller coordination number than 7, bidentate coordination becomes almost non-existent.

All of this points to Mg^{2+} not coordinating the active site in the same way as Ca^{2+} . Instead, it is possible that only aspartate 36, aspartate 38, valine 42, and water make the inner-sphere coordination, bonding directly with the metal. The two oxygens from glutamate 47 and the oxygen from serine 40 are only loosely interacting with the metal ion and are in the Mg^{2+} outer-sphere. All 4 inner-sphere coordinations are within the same plane making the actual coordination square planar and giving the Mg^{2+} a coordination of 4, which is seen about 15% of the time in this resolution range.

The EF3 binding site geometry has an obvious and drastic change between Mg^{2+} and Ca^{2+} . Calcium has a coordination number of 6, with bidentate coordination with E114 and no interaction with a water present when magnesium is bound. Unlike in EF1, all of the distances (2.3-2.5 Å) are within the most commonly seen range for calcium. Magnesium, like with calcium, has a coordination number of 6 with bond lengths from 1.9-2.1 Å, right in the most common range. However, with the short bond distances with residues 103, 105, 107, and 109, magnesium is too far away from glutamate 114 to interact directly with it. Instead, magnesium coordinates with a water molecule that hydrogen bonds with glutamate 114. Also, magnesium coordinates with an extra water molecule, not present in the calcium bound structure. Calcium not interacting with water is not hugely surprising in that coordination with water is a very small fraction of seen calcium coordination with a coordination number of both 6 and 7. However, at a coordination number of 6, magnesium has the lowest fraction seen coordinating with water. Interactions with water are clearly not unheard of, it is surprising that calcium does

not interact with that extra water to increase its coordination number from 6 to 7, which occupy almost equal observed percentages in the PDB.

In fact, magnesium models closely what is seen with the Na⁺-bound structure. Surprisingly, Na⁺ has similar bond lengths as Ca²⁺, however sodium has moved away from the bidentate coordinating E114. Also, the sodium does not appear to have a water mediated coordination with this glutamate as magnesium does. Based on sodium and magnesium occupying a similar location that is different from calcium, but sodium having equivalent bond lengths as calcium, it is not just magnesium having smaller coordination bond lengths and moving further into the pocket. Also, with calcium and sodium ions having a similar size, it seems unlikely that it is a conformational change in EF3 to accommodate the smaller magnesium ion. It is unclear why the movement away from glutamate happens.

Nothing overtly strange happens within binding pocket EF5. Both calcium and magnesium have the same coordination number, 6, interact with all of the same residues, and all have their respective expected bond lengths. EF5 is considered the low affinity binding site and is located within the oligomerization domain. This could explain why sodium is also seen in this binding site in not only the Ca²⁺-free form but also in the Zn²⁺-bound enzyme. It could also explain why metal binding in this binding site is almost identical in all forms of the enzyme.

However, with sodium being able to interact in all of the same binding sites as those for the metals of interest, how at 2.0 Å resolution can magnesium and sodium be distinguished from one another? The STAN server uses an algorithm to determine if modelled water molecules are actually metal ions. This server gave the first indication as

to what was bound in the active site. For the Ca^{2+} -free form the server found that there was either Na^+ or Ca^{2+} in these binding pockets. However, when using the Mg^{2+} structure with water modelled into all of the active sites, the server finds either Mg^{2+} or Li^+ . With the absence of lithium in the crystallization solutions used, then it would be safe to assume that magnesium ions were those in the binding sites. With the divalent and monovalent ions having similar ionic radii, then it only follows that this algorithm would not be able to distinguish between the two. However, it does seem to be able to distinguish between sodium and magnesium.

7.5. Future Work

To further prove that that it is magnesium bound over sodium a few experiments have been considered. The first and most direct method is to remove all traces of sodium from the protein buffer and crystallization condition. This, unfortunately, means no buffers used which changes the crystallization conditions and has the potential to not yield similar crystals to those crystallized with sodium.

Another experiment is to use lutetium, a lanthanide with a similar radius as magnesium, to replace magnesium in the binding pocket. Lanthanides will have a, roughly, 1000x higher binding affinity than magnesium. A simple soak should work, but adding lutetium to the crystallization condition may also be successful. Lutetium will be obvious within the anomalous scattering of X-rays. With lutetium being confirmed, then a comparison between modelling sodium and lutetium can be made with the current magnesium structure.

The concern is lanthanides are notoriously insoluble, especially in the pH range that crystallization of these crystals takes place. A pH between 4-6 is where lutetium is

soluble, but not the pH range that these crystals exist. Further, with a 3+ ionic state, lutetium, if it binds, may cause unrelated conformational changes. With a distorted binding site, then a comparison between the two binding sites would be impossible.

If lutetium fails, then manganese could be tried as well. Manganese is known to bind all the same binding pockets that magnesium does. Manganese would be obvious also based off of anomalous scattering within the crystal, and with a +2 charge, manganese should not cause the potential conformational changes that a +3 lutetium would.

References

1. Vito, P., Lacana, E., and D'Adamio, L. (1996) Interfering with apoptosis: Ca(2+)-binding protein ALG-2 and Alzheimer's disease gene ALG-3, *Science* 271, 521-525.
2. Maki, M., Yamaguchi, K., Kitaura, Y., Satoh, H., and Hitomi, K. (1998) Calcium-induced exposure of a hydrophobic surface of mouse ALG-2, which is a member of the penta-EF-hand protein family, *Journal of biochemistry* 124, 1170-1177.
3. Tarabykina, S., Moller, A. L., Durussel, I., Cox, J., and Berchtold, M. W. (2000) Two forms of the apoptosis-linked protein ALG-2 with different Ca(2+) affinities and target recognition, *The Journal of biological chemistry* 275, 10514-10518.
4. Maki, M., Narayana, S. V., and Hitomi, K. (1997) A growing family of the Ca²⁺-binding proteins with five EF-hand motifs, *The Biochemical journal* 328 (Pt 2), 718-720.
5. Maki, M., Suzuki, H., and Shibata, H. (2011) Structure and function of ALG-2, a penta-EF-hand calcium-dependent adaptor protein, *Science China. Life sciences* 54, 770-779.
6. Shibata, H., Yamada, K., Mizuno, T., Yorikawa, C., Takahashi, H., Satoh, H., Kitaura, Y., and Maki, M. (2004) The penta-EF-hand protein ALG-2 interacts with a region containing PxY repeats in Alix/AIP1, which is required for the subcellular punctate distribution of the amino-terminal truncation form of Alix/AIP1, *Journal of biochemistry* 135, 117-128.
7. Suzuki, H., Kawasaki, M., Inuzuka, T., Okumura, M., Kakiuchi, T., Shibata, H., Wakatsuki, S., and Maki, M. (2008) Structural basis for Ca²⁺-dependent

formation of ALG-2/Alix peptide complex: Ca²⁺/EF3-driven arginine switch mechanism, *Structure* 16, 1562-1573.

8. Shibata, H., Suzuki, H., Kakiuchi, T., Inuzuka, T., Yoshida, H., Mizuno, T., and Maki, M. (2008) Identification of Alix-type and Non-Alix-type ALG-2-binding sites in human phospholipid scramblase 3: differential binding to an alternatively spliced isoform and amino acid-substituted mutants, *The Journal of biological chemistry* 283, 9623-9632.
9. Jia, J., Tarabykina, S., Hansen, C., Berchtold, M., and Cygler, M. (2001) Structure of apoptosis-linked protein ALG-2: insights into Ca²⁺-induced changes in penta-EF-hand proteins, *Structure* 9, 267-275.
10. Kabsch, W. (2010) Xds, *Acta crystallographica. Section D, Biological crystallography* 66, 125-132.
11. Evans, P. (2006) Scaling and assessment of data quality, *Acta crystallographica. Section D, Biological crystallography* 62, 72-82.
12. Adams, P. D., Afonine, P. V., Bunkoczi, 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 crystallographica. Section D, Biological crystallography* 66, 213-221.
13. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallographica Section D* 60, 2126-2132.
14. Nayal, M., and Di Cera, E. (1996) Valence screening of water in protein crystals reveals potential Na⁺ binding sites, *Journal of molecular biology* 256, 228-234.
15. Luo, Z., Rajashankar, K., and Dauter, Z. (2014) Weak data do not make a free lunch, only a cheap meal, *Acta crystallographica. Section D, Biological crystallography* 70, 253-260.
16. Matsuda, S., and Yagi, K. (1980) The apparent binding constants of Ca²⁺ to EGTA and heavy meromyosin, *Journal of biochemistry* 88, 1515-1520.
17. Zheng, H., Chruszcz, M., Lasota, P., Lebioda, L., and Minor, W. (2008) Data mining of metal ion environments present in protein structures, *Journal of Inorganic Biochemistry* 102, 1765-1776.

Table 7.1. Data collection and Refinement Statistics.

	5% MPD Mg ²⁺	Home Source	1mM Mg ²⁺
<i>Spacegroup</i>	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2
<i>Unit-cell parameters</i> (Å)	<i>a</i> = 75.8 <i>b</i> = 48.6 <i>c</i> = 53.9	<i>a</i> = 75.7 <i>b</i> = 48.5 <i>c</i> = 54.0	<i>a</i> = 76.1 <i>b</i> = 48.6 <i>c</i> = 54.3
<i>Wavelength</i>	1.0000	1.541	1.0000
<i>Resolution</i>	48.6 – 2.1 (2.2 – 2.1)	19.1 – 2.3 (2.4 – 2.3)	54.3 – 1.9 (2.0 -1.9)
<i>No. of observations</i>	68495 (5798)	59829 (6010)	110235 (5557)
<i>Unique reflections</i>	10352 (1081)	9286 (1020)	16134 (986)
<i>R_{merge} (I)</i>	0.028 (0.084)	0.053 (0.301)	0.025 (0.168)
<i>R_{meas} (I)</i>	0.033 (0.102)	0.063 (0.365)	0.029 (0.202)
<i>R_{pim} (I)</i>	0.017 (0.058)	0.034 (0.204)	0.015 (0.110)
<i>Mean I/σ(I)</i>	41.6 (14.2)	23.0 (5.6)	45.9 (9.5)
<i>Completeness (%)</i>	85.1 (99.7)	99.8 (99.9)	98.3 (86.1)
<i>Multiplicity</i>	6.6 (5.4)	6.4 (5.9)	6.8 (5.6)

Table 7.2. Comparison of the different coordination lengths between the Mg²⁺-bound structure, the Na⁺-bound structure (2ZND) and the Ca²⁺-bound structure (2ZN9).

EF1	Distance			EF3	Distance			EF5	Distance		
	Mg ²⁺ ₊	Na ⁺	Ca ²⁺		Mg ²⁺	Na ⁺	Ca ²⁺		Mg ²⁺ ₊	Na ⁺	Ca ²⁺
OD1 ^{D36}	2.0	2.3	2.3	OD1 ^{D103}	2.1	2.4	2.3	OD1 ^{D169}	2.0	2.4	2.5
OD1 ^{D38}	1.9	2.3	2.2	OD2 ^{D105}	2.1	2.4	2.5	OD2 ^{D171}	1.9	2.4	2.3
OG ^{S40}	3.0	-	2.8	OG ^{S107}	2.1	2.5	2.5	OD1 ^{D173}	1.9	2.3	2.2
O ^{V42}	2.2	2.2	2.4	O ^{M109}	2.0	2.4	2.4	O ^{W175}	2.2	2.3	2.5
H ₂ O	1.9	2.6	2.3	H ₂ O	2.0	2.3	ND	H ₂ O	1.9	2.5	2.3
OE1 ^{E47}	2.3	2.3	2.8	OE1 ^{E114}	-	3.7	2.3	H ₂ O	2.2	2.6	2.4
OE2 ^{E47}	2.8	2.8	2.6	OE2 ^{E114}	-	4.3	2.5				

Figure Legends

Figure 7.1. Form 1 ALG-2 des23 crystals. Grown in sitting drop, vapor diffusion plates using 22-32% isopropanol and 0.1 M Tris pH 7.0-8.5.

Figure 7.2. Form 2 ALG-2 des23 crystals. Grown in sitting drop, vapor diffusion plates using 22-32% isopropanol, 0.1 M Tris pH 7.0-8.5, and 0.2 M MgCl₂.

Figure 7.3. Diffraction pattern from Form 1 crystals of ALG2.

Figure 7.4. Diffraction pattern from Form 2 crystals of ALG2.

Figure 7.5. Overall structure of the symmetry-made dimeric ALG-2 enzyme. In red is the chain in the asymmetric unit and blue is the chain made by symmetry. The orange spheres represent the magnesium modelled into EF1, EF3, and EF5.

Figure 7.6. A) EF1 binding pocket with magnesium in orange. This pocket in particular appears to make a pentagonal bipyramidal coordination, which is unusual for Mg²⁺. B) EF3 binding pocket with magnesium (orange) making the expected octahedral coordination. C) EF5 binding pocket with magnesium (orange) making the expected octahedral coordination.

Fig 7.1.



Fig 7.2.



Fig 7.3.

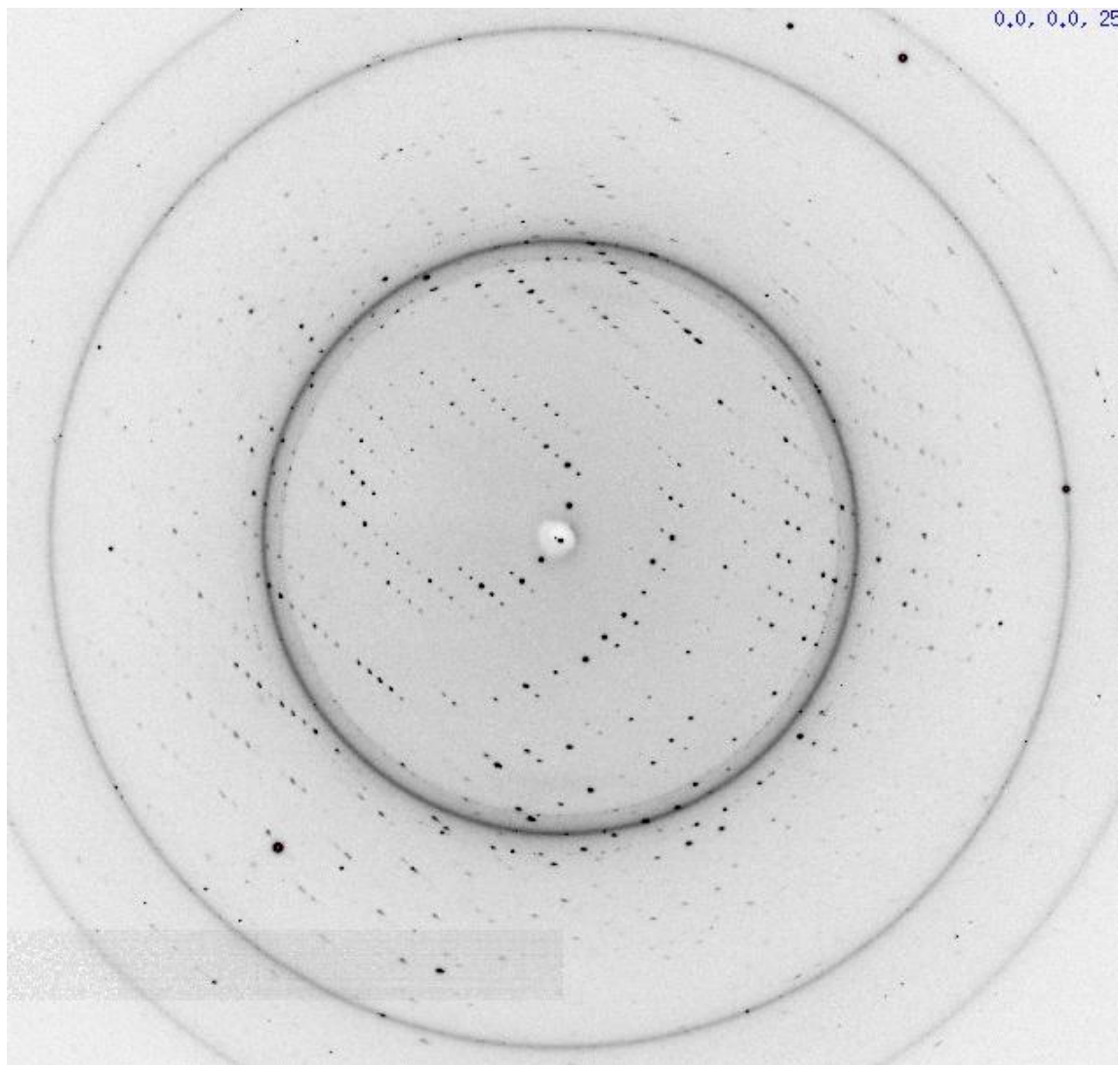


Fig 7.4.

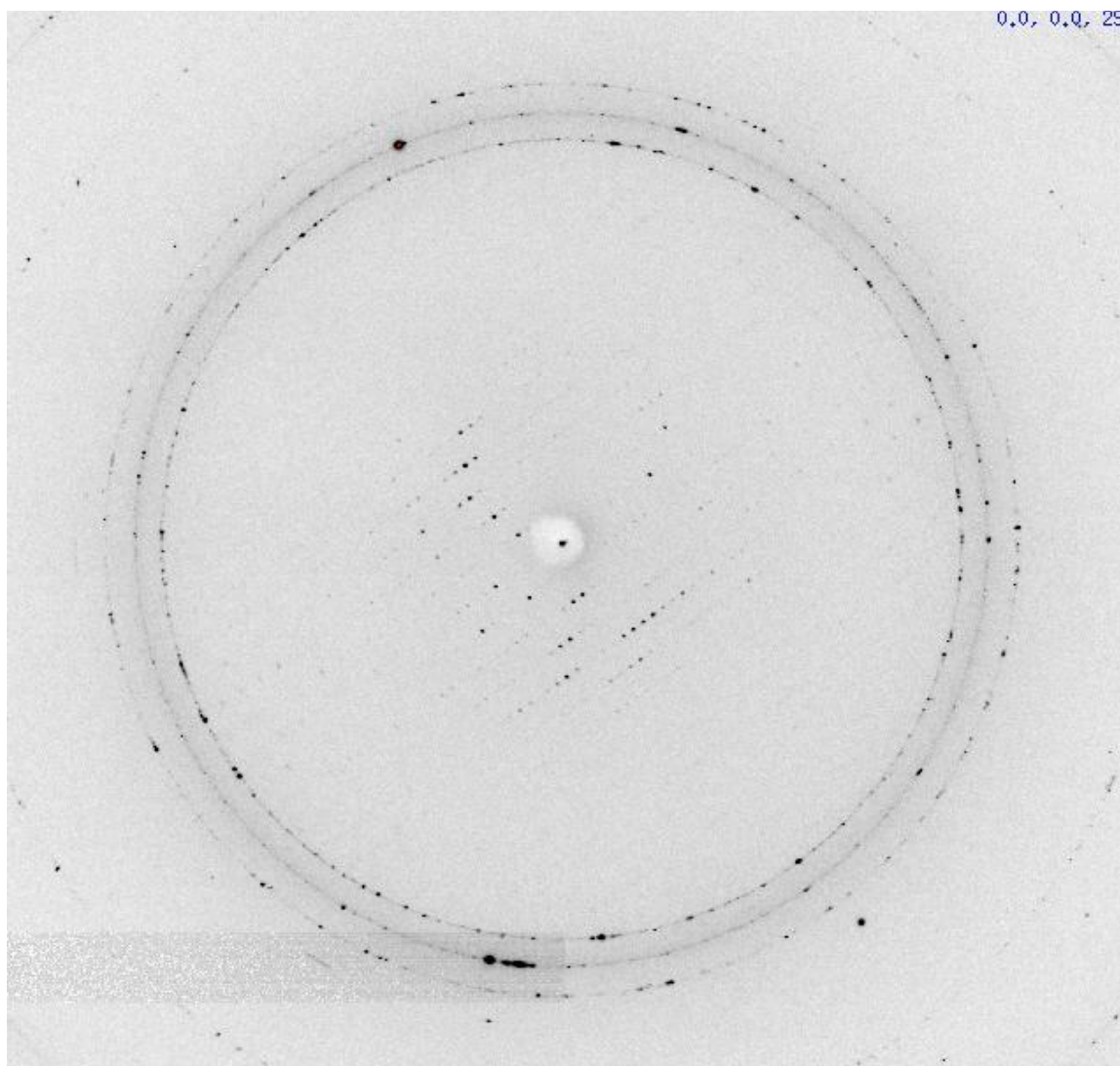


Fig 7.5.

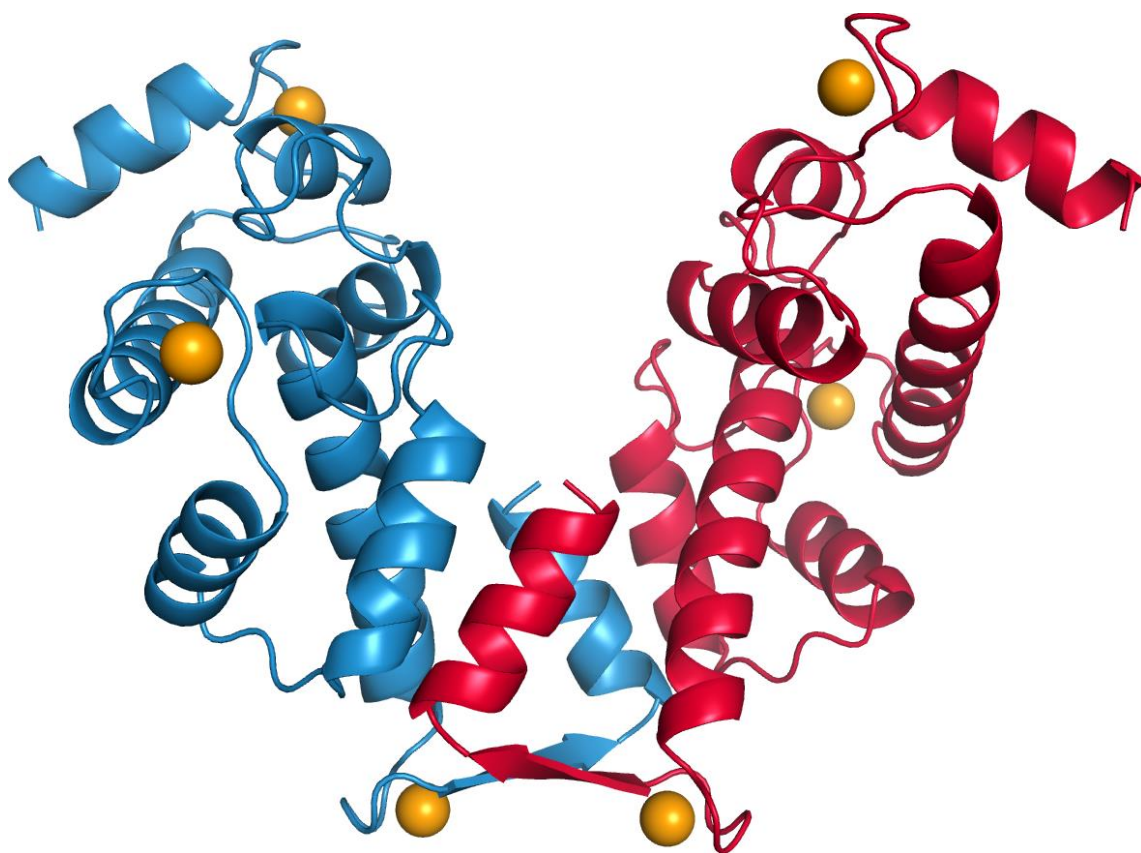
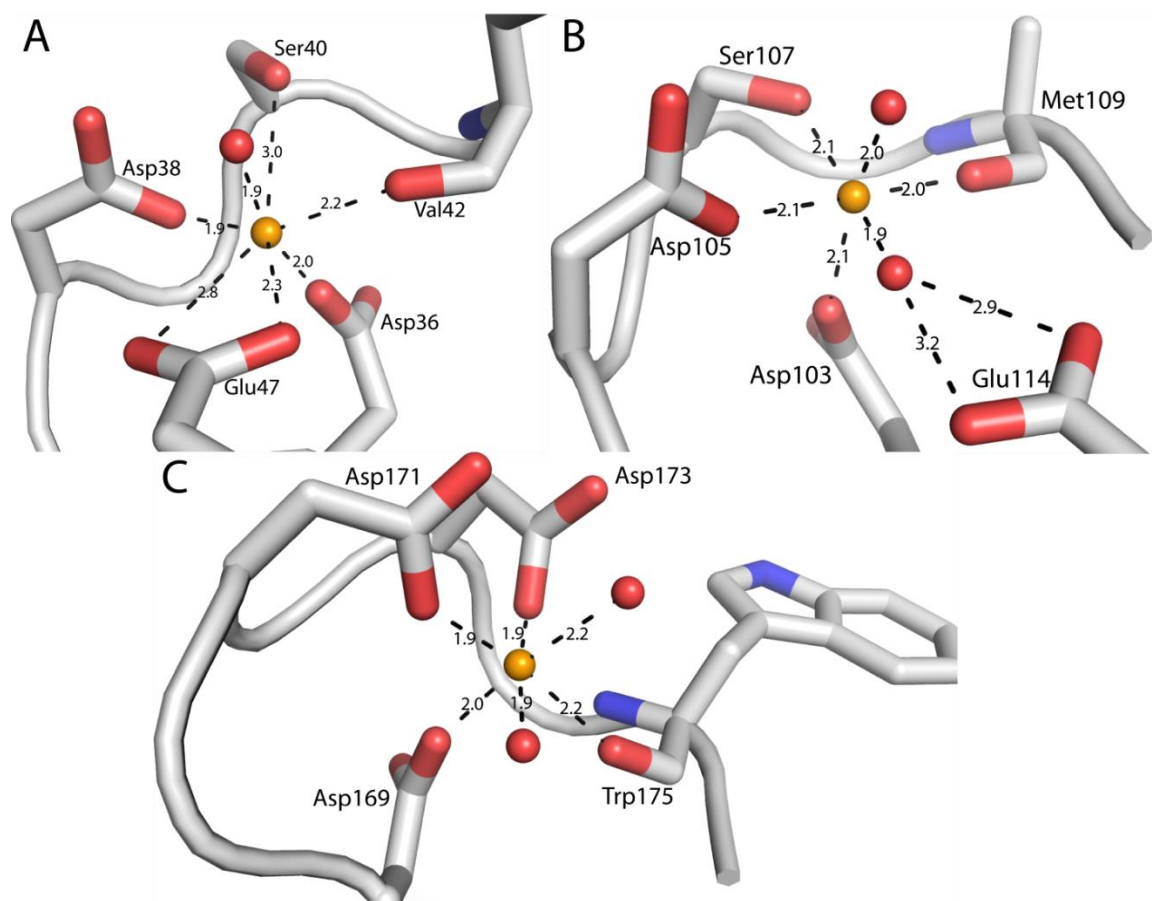


Fig 7.6.



APPENDIX

A1.

EXPRESSION, PURIFICATION, CRYSTALLIZATION, AND
SMALL ANGLE X-RAY SCATTERING (SAXS) OF THE Δ 1-
986 *BRADYRHIZOBIUM JAPONICUM* PROLINE
UTILIZATION A (BJFLAP)

A1.1. RATIONALE

The original BjPutA structure was published in PNAS in 2010. The structure revealed an internal, water filled cavity that upon further investigation allowed for channeling of the product from the PRODH domain to the active site of the P5CDH domain without contact with bulk solvent. Further, it was discovered that the protomer of this enzyme had an elongated C-terminal domain that became disordered after residue 989, 10 total disordered residues. Alone, the protomer had a large hole to the cavity, which would have made not only channeling difficult but contact with bulk solvent inevitable(1).

Upon dimerization the C-terminal of one protomer covers the whole of another protomer. This means that for channeling to be possible dimerization has to occur, otherwise a large, gaping hole allows for leakage out of the channel. To test this idea, perhaps a large hole can be manufactured and then channeling can be tested to see if efficiency is reduced. It was decided that the very C-terminal of this enzyme would be removed. A truncation was made removing the last 13 residues. The truncation was 1-986, removing residues 987-999, and it's commonly referred to in the lab as BjFlap. Positive results in this hypothesis would confirm that dimerization is needed and would be further evidence that channeling is indeed occurring.

A1.2. MATERIALS AND METHODS

A1.2.1 Protein Expression and Purification

This protein was expressed and purified as described previously for the wild-type(2). The plasmid containing BjFlap was retransformed into BL21 DE3 pLysS cells. Colonies from these plates were placed in 10mL and 100 µg/mL ampicillin of LB Broth

and allowed to grow overnight. These starting cultures were used to inoculate 1L of LB Broth and 100 µg/mL ampicillin. Inoculated media was grown at 37 °C and 250 rpm until reaching an OD₆₀₀ of 0.8. At this time, the media was induced using 500 µM IPTG and grown overnight at 25 °C and 200 rpm.

The next day the cell cultures spun at 3500 rpm in an SLC-6000 rotor in a Sovall fixed-angle centrifuge. The pelleted cells were resuspended in 20 mM HEPES pH 8.2, 150 mM NaCl, and 5% glycerol. The resuspended cells were flash frozen and stored at -80 °C for later use.

Frozen cells were thawed and broken using sonication with cycles of 30 seconds on and 2 minutes off to allow time to chill. Four of these were performed and the cell lysate was centrifuged in an SS-34 rotor at 16,500 rpm for 45 minutes. The cells were removed, the supernatant decanted into a fresh centrifugal tube and spun again for 45 minutes at 16,500 rpm. This supernatant was decanted and then filtered through a 0.45 µm syringe-driven filter.

The filtered supernatant was then loaded onto a Ni-NTA column equilibrated with 20 mM HEPES at pH 8.2, 150 mM NaCl, and 5% glycerol using an AKTA-prime FPLC. The loaded column was washed using the 30 mM imidazole to remove loosely binding contaminating enzymes. The enzyme was then eluted off the column using 300 mM imidazole. Fractions were collected in 2 mL aliquots. The fractions were pooled and about 1 mg/mL TEV protease was added. This mixture was allowed to incubate at 25 °C for 1 hour and then placed in a dialysis buffer containing 20 mM HEPES at pH 8.2, 150 mM NaCl, and 5% glycerol and allowed to dialyze overnight.

Following dialysis the enzyme was loaded, again, on a Ni-NTA column using the

afore mentioned binding buffer. However, this time during elution, only 30 mM imidazole was needed to remove the enzyme from the column. 2 mL fractions were collected and SDS-PAGE was run to check for purity. The enzyme was placed in a dialysis of 50 mM Tris pH 7.5, 0.5 mM EDTA, 0.5 mM THP, and 5% glycerol in 4 °C overnight. The following day the enzyme was loaded onto a Hi-Trap anion exchange column with a binding buffer that was the same as the dialysis buffer. A gradient of 0-1 M NaCl was run over 100 mL. The enzyme eluted around 300 mM NaCl and fractions were collected of about 500 µL. The fractions were tested for purity using SDS-PAGE and pure fractions were pooled. Pooled fractions were then dialyzed against the precrystallization buffer consisting of 50 mM Tris at pH 7.5, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, and 5% glycerol.

A1.2.2 Crystallization

Initial crystallography trials were attempted as previously reported for the wild-type enzyme. Sitting drop trays containing 1.6-2.2 M ammonium sulfate and a pH range of 7.5-9.0 was attempted for crystallization. High concentrations of 8-10 mg/mL of BjFlap were tried, however, none of these worked. Another set of crystal trays were set up increasing the ammonium sulfate concentration with an upper limit of 2.8 M. However, these trays only had precipitate in them.

Burning the end of a 0.1-10 µL pipet tip, and allowing it to cool created a nice pestle to crush wild-type BjPutA crystals. 10 µL of the crushed enzyme were placed into 90 µL of mother liquor, to create a 10x dilution stock. This stock was vortexed for about 30 seconds to break up any large crystal chunks that might be present. From the stock a 100x and 1,000x dilution were made to further control nucleation.

Crystal trays in the original range were set up again, this time with varied concentrations of 4-8 mg/mL BjFlap. This time after mixing the enzyme and mother liquor in 1:1 ratio, a horse hair that had been dipped in a seed stock was swiped through each well, with a brief dip in ethanol at the end. Each stock solution of seed 10x, 100x, and 1000x was tried(3). Crystal trays were incubated at room temperature and after a couple days these trays had yellow, plate crystals present. They took about a week to grow to maximum dimensions. Subsequent optimizations utilized the crystals of the respective mutants grown at room temperature with a seed stock prepared in the same manner. Crystals were cryoprotected in 25-30% glycerol and then flash cooled by dunking in liquid nitrogen.

A1.2.3 Data Collection and Refinement

The space group is $C2$ with a BjPutA dimer in the asymmetric unit. X-ray diffraction data sets were collected at beamline 4.2.2 of the Advanced Light Source using a NOIR-1 detector. The data were integrated with MOSFLM(4) and scaled with SCALA(5). Refinements in PHENIX(6) were initiated from models derived from the structure of wild-type BjPutA (Protein Data Bank (PDB) entry 3HAZ). COOT(7) was used for model building. The structures were validated with MolProbity(8) and the PDB validation server. Data collection and refinement statistics are listed in Table 1.

A1.2.4 SAXS

Enzyme for SAXS was purified in the same way as that for purification with one extra step added; running the enzyme was run through size exclusion chromatography. Any aggregate was discarded and homogenous enzyme was collected. Collected fractions were pooled and concentration was estimated using the BCA method(9). Enzyme was

concentrated up to about 8 mg/mL and again verified using BCA. This concentrated sample was used for SAXS and the flow through from concentration was used as a reference buffer.

SAXS experiments were performed at SIBYLS beamline 12.3.1 of the Advanced Light Source through the mail-in program(10, 11). For each sample, scattering intensities were measured at three nominal protein concentrations. Data were collected for each protein concentration at exposure times of 0.5 s, 1.0 s, 3.0 s, and 6.0 s. The scattering curves collected from the protein samples were corrected for background scattering using intensity data collected from the reference buffer.

The SAXS data were analyzed as follows. Composite scattering curves were generated with PRIMUS(12) by scaling and merging the background-corrected low q region data from the 0.5 s exposure with the high q region data from the 3.0 s exposure. PRIMUS was also used to perform Guinier analysis. FoXS(13) was used to calculate theoretical scattering profiles from atomic models. GNOM was used to calculate pair distribution functions. The SASTBX “Shape Up!” server was used for shape reconstruction calculations.

A1.2.5 PISA

To analyze the interfaces for oligomerization, the PISA server was used. The 3HAZ structure was plugged directly into the server from a search of the PDB. The different interfaces were examined for potential mutation sites using COOT. The same thing was done, again, by loading the BjFlap structure and analyzing the proposed interfaces.

A1.3. RESULTS

A1.3.1 Crystallography

The structure looks very similar to that of the wild-type enzyme. The PRODH domain has a sulfate bound and the ion-pair gate is in the closed state. The flavin is in the oxidized form as well. The P5CDH domain also, looks as if it was the wild-type enzyme as well. The most distinct difference is that residues 987 and 988 and 989 are missing. The structure ends 986, which was expected. Further when running mole calculations the escape routes run out near the C-terminal residues, 986. VOIDOO(14) calculations also show a widening of this hole near the C-terminal just on the large cavity side of the ion pair gate.

A1.3.2 SAXS

The oligomeric state of BjFlap was investigated using SAXS. Guinier plots of data from four samples having different protein concentrations yield a radius of gyration (R_g) in the range 44.5-47.5 angstroms. The least noisy data was consistently 46-47 Å. The R_g of the BjPutA dimer in the asymmetric unit is about 45 Å, which suggests that the BjFlap actually looks most like the dimer. For comparison, the tetramer of the wild-type BjPutA has an R_g value of 51.1, which is much larger than the experimental R_g values found.

More to the point, when investigating the scattering profile, BjFlap loses two important characteristics of the wild-type tetramer. First, the characteristic dip at about .075 q is completely gone. Second, and possibly more importantly, the low q range approaching zero no longer aligns well with the tetramer scattering profile.

SASTBX shape reconstructions no longer find any trace of the donut shaped

tetramer from the wild-type. Shape reconstructions fit well with the U-shaped dimer seen for many PutA dimers. The FoXS server picks the dimer as the best fit with a χ value of 3.00. The protomer (unexpected) had a χ value of 8.30 and the tetramer had a χ value of 20.36. However, when using the servers minimal ensemble search (MES) which can take multiple structures and find a combination of them to give the best result, reduces the χ value to 1.74 when combining 89% dimer and 11% tetramer.

A1.4. DISCUSSION

A1.4.1 Structure Analysis of the Flap Mutant

The structure of this mutant does not offer much information. The mutation was removed the last 13 residues of the enzyme, making BjFlap a truncation mutant with identical sequence as wild-type from residues 1-986. The problem is that the wild-type structure becomes disordered after 989. The last 10 residues of the wild-type structure have no density present. The only observable difference in the structure comparison between these two is missing the absence of 987-989 in the BjFlap. All other residues look similar.

VOIDOO and CAVER(15) calculations were done to determine if there is any difference in the internal cavity. VOIDOO performed on the wild-type structure frequently finds an exit from the cavity to the bulk solvent near the C-terminal of one protomer and the PRODH of the complementing protomer. As expected, VOIDOO calculations for the BjFlap find this same exit, however, the exit it finds appears to be a substantially larger opening. This leads credit to the hypothesis of the removal of this C-terminal makes a substantial hole, exposing the internal cavity to bulk solvent,

CAVER starts at designated coordinates within the structure and then finds the

fastest way out of the enzyme for a water molecule. Using this calculation for the wild-type structure there is an escape route out to the bulk solvent, but only finds this route once. The same calculations performed on BjFlap show upwards of 3 different escape routes near the C-terminal of the enzyme. This clearly indicated an enlarging of this hole to the cavity.

What is unclear is how much larger this hole is as compared to the wild-type. With the final 10 residues missing in the wild-type structure it is hard to tell if the access to the bulk solvent found by VOIDOO and CAVER is simply an artifact of the structure. A different space group with an ordered C-terminal might stop VOIDOO and CAVER from finding this hole. Unfortunately, another space group for this enzyme never gave usable data for structure determination. It is likely this hole is real based on the wild-type structure having approximately 70% efficiency. Substrate leakage is highly likely.

What is clear, though, is that the hole and CAVER escape routes found in the BjFlap structure are real. This gives some indication that the C-terminal is important to covering the cavity. Without the C-terminal the cavity is still covered but a large hole decreases efficiency of the channel and increases the likelihood of the substrate escaping to bulk solvent.

A1.4.2 C-Terminal Deletion Disruption of Tetrameric Oligomerization

Perhaps the most surprising of results of the oligomeric state. The wild-type enzyme is a tetramer in solution. When plugged into the FoXS server using the multiple ensemble search (MES) yields a mixture of 85% tetramer and 15% dimer. The dominant species, by far, is the tetramer, shown in the scattering profile with a characteristic dip and bump. The dimer, however, lacks that specific characteristic. Further when doing

shape reconstruction a donut shape appears which fits well with the dimer of dimer ensemble found through crystallographic symmetry.

However, the scattering profile of the BjFlap mutant is devoid of the dip and bump expected for a tetramer. Rather, when using the FoXS server the experimental scattering profile fits well with the dimer found in the asymmetric unit. When using MES the scattering profile fits best with a mixture of 11% tetramer and 89% dimer. Compared to the wild-type at the same concentration the ratios have been virtually reversed, with the mutant now favoring dimer. Further when using the SASTBX server for shape reconstruction, the envelope resembles that of the dimer found in the asymmetric unit, fitting with a 0.94 correlation coefficient. Tetramer formation is almost non-existent and does not play a dominant role in the overall shape of the enzyme.

As mentioned previously, the C-terminal domain is completely disordered in the crystal structure and there is no direct evidence to what it is doing. However, recently there have been two more structures of PutAs solved that of GsPutA(16) and that of SmPutA (unpublished), that both have a complete C-terminal. In the wild-type BjPutA structure the last visible residue, 989, is about 18-20 angstroms away from the tetramerization oligomerization domain. It is conceivable that the full length of the C-terminal could reach there. To get a better picture of what is happening the two new structures of homologous PutAs were superimposed on the BjPutA. However, the C-terminal of both enzymes pointed away from the oligomerization domain, and instead coiled up over the hole over the PRODH domain found in the BjPutA structure.

These may be a poor substitute for modeling the BjPutA C-terminal, however, because GsPutA is a dimer in solution and SmPutA is a monomer. In fact, to date the

only PutA that has been discovered to be a tetramer is BjPutA. It is entirely possible, considering it is a unique case, that the C-terminal plays a special role in this enzyme to stabilize tetramer formation.

A1.4.3. Possible Mutations for Disruption of the Tetramer

PISA indicates that a tetramer assembly is stable in solution. The tetramer is a dimer of dimers. Each dimer is a domain swap where the C-terminal of one protomer covers the internal cavity of the other protomer. The tetrameric assembly, however, appears to be largely loose stacking interactions near the PRODH domain. The two dimers interact at three separate domains all near the same area. Residues 46-66, 144-154, and 474-479 are the major oligomeric domains. However, even in these domains, there are virtually no direct hydrogen bonding interactions between dimers. Rather, the domains appear to just come into fairly close contact and sit near each other. Without PISA indicating this as a stable interaction and the SAXS data backing up the claim, this could have easily been dismissed as a crystal packing artifact.

Mutations to disrupt oligomer formation in the yeast and thermus thermophiles P5CDHs were relatively straight forward as there were clear interactions of side chains reaching out and interacting with areas on other protomers. Removing these interacting side chains disrupted oligomerization almost completely for these enzymes. However, the approach here is to take a small side chain and make it a large, obtrusive one to disrupt the tetramerization domain on these enzymes. Specifically, Ala66, Gly148, Pro150, and Val478 to tryptophan. The tryptophan introduces a very large and space filling side chain that may impede the somewhat loose interactions seen between these two dimers.

However, it might also be worthwhile investigating mutations similar to those

used before. Such as taking a larger, polar side chain and replacing it with a large, nonpolar side chain. There are a few interactions that, while not making a direct interaction between dimers, may be important for the interface to come together correctly. Glu46, Arg49, and Glu64 to tryptophan mutations may be able to interrupt tetramer formation as well by disrupting the interface.

Table A1.1. Data Collection and Refinement Statistics

	BjFlap
Spacegroup	C2
Unit-cell Parameters (Å)	$a = 167.1$ $b = 196.1$ $c = 108.8$ $\beta = 121.4$
Wavelength	1.0000
Resolution	38.69 – 2.75 (2.90- 2.75)
No. of Observations	284897 (38961)
Unique Reflections	77333 (11219)
R _{merge} (I)	0.146 (0.469)
R _{meas} (I)	0.171 (0.553)
R _{pim} (I)	0.088 (0.291)
Mean I/σ(I)	6.3 (2.1)
Completeness (%)	99.9 (99.7)
Multiplicity	3.7 (3.5)
No. of Proteins Atoms	14470
R _{cryst}	0.2136
R _{free}	0.2667
<i>R.m.s. Deviations</i>	
Bond Lengths (Å)	0.008
Bond Angles (°)	1.128
<i>Ramachandran Plot (%)</i>	
Favored	94.6
Allowed	4.6
Outliers	0.8
<i>Average B (Å²)</i>	
Protein	39.640
Water	24.14

Figure Legends

Figure A1.1. Summary of BjPutA C-terminal truncation SAXS data. A) The experimental scattering curve of BjFlap overlaid with the calculated scattering profile of the BjPutA dimer (Solid Blue), BjPutA tetramer (Solid Red), and an MES combination of 89% dimer and 11% tetramer (Dashed Yellow). The inset represents the Guinier Analysis plotted for all three concentrations as indicated. B) The shape reconstructions from the SASTBX Shape Up web server of BjFlap superimposed with the structure of the dimer of the structure of BjFlap.

Figure A1.2. The structure of Chain A of the BjPutA structure is shown. FAD is in yellow, NAD is in cyan, and the cavity calculated for the wild type is represented as the pink surface. A) The C-terminal flap motif of Chain B that covers the cavity hole of Chain A is shown for the wild-type in red and blue. B) The C-terminal flap motif of the truncated Flap mutant highlighting the enlarging of the hole seen in the cavity.

Fig A1.1.

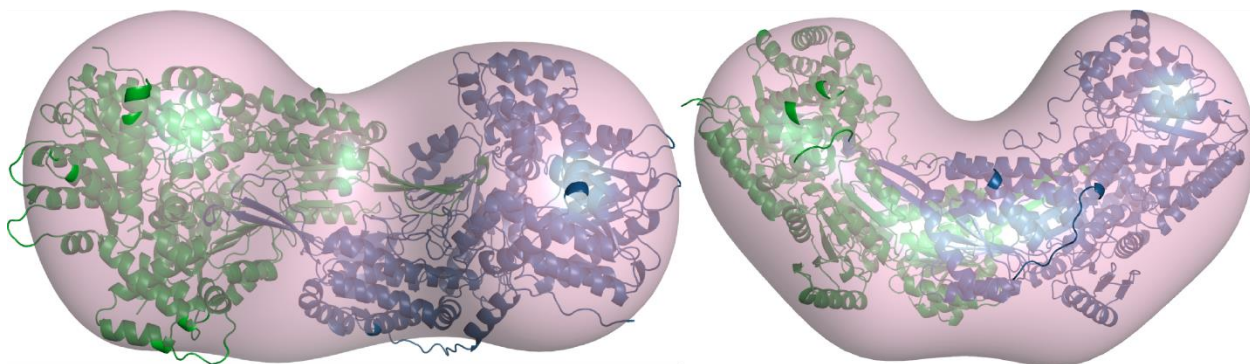
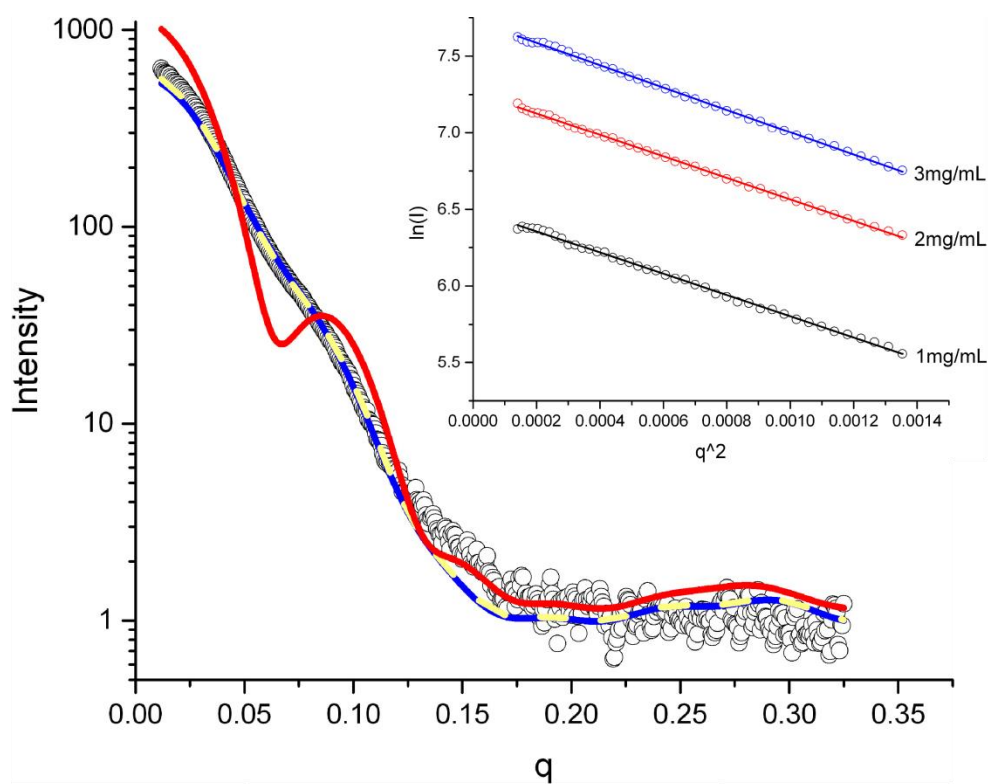
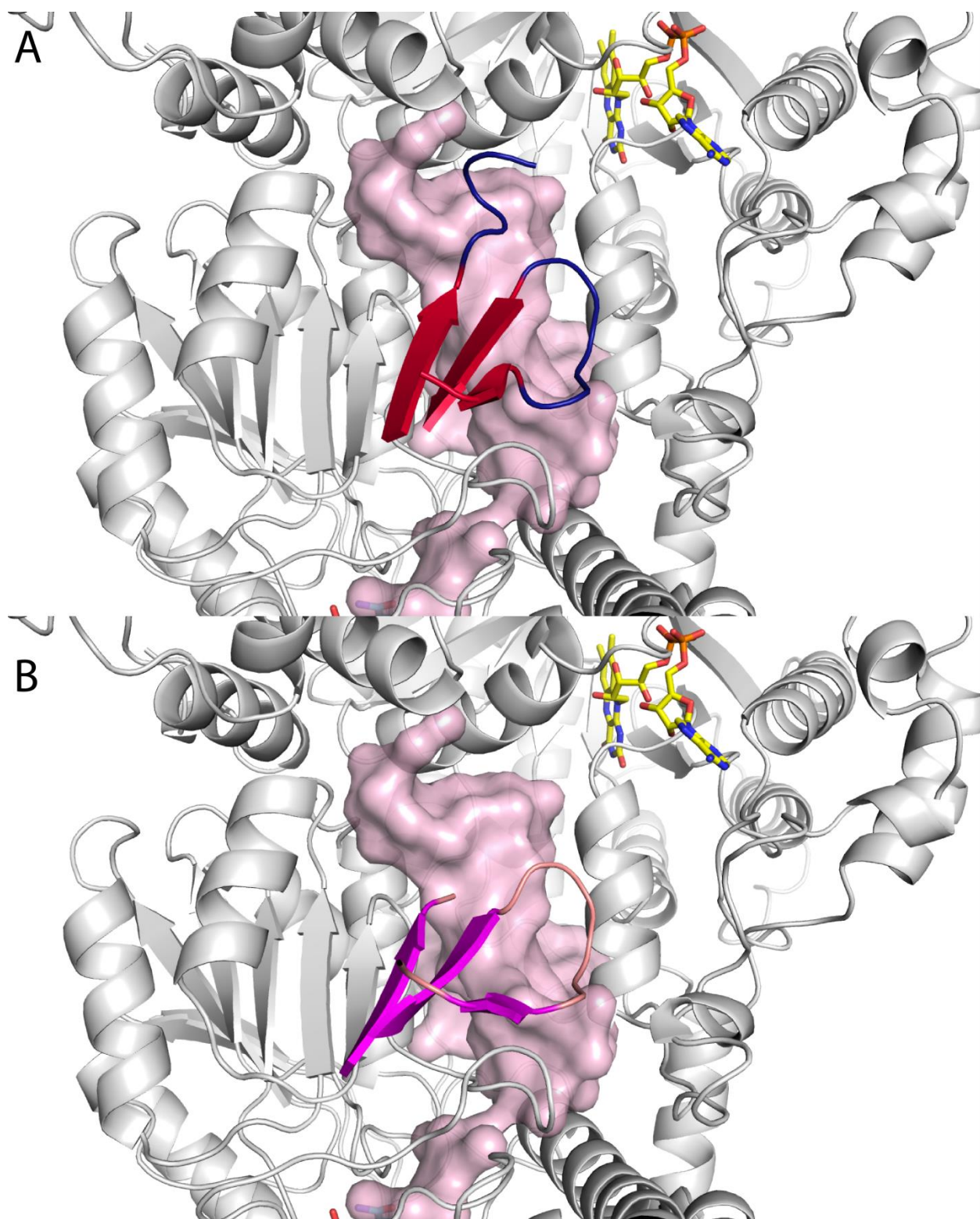


Fig A1.2.



A2.

PROTEIN EXPRESSION, PURIFICATION,
CRYSTALLIZATION AND SAXS ANALYSIS OF
LEGIONELLA PNEUMOPHILA PUTA (LPPUTA)

A2.1. RATIONALE

BjPutA was the first full length PutA that was crystallized. Homologous screening shows that many organisms have similar PutA enzymes. *Legionella pneumophila* PutA (LpPutA) was a minimalist PutA much like BjPutA, but was only about 30% identical in sequence. This gives the possibility of further characterizing questions that remained about BjPutA, such as characterization of the channel, how the C-terminal of the enzyme is ordered (10 residues were completely disordered), and how important the high order oligomeric state was.

Usable crystals of BjPutA were only ever seen and obtained in 2 M ammonium sulfate. Unfortunately, density for sulfate was seen in both the active site of PRODH and P5CDH. The high concentration of sulfate made it hard to displace with inhibitors and ligands that would help characterize the active sites and explain enzyme function. Homologs open up the possibility of crystallizing an enzyme in either the absence of salts that bind the active site, or at low enough concentrations that ligand soaking is possible.

The quality of diffraction data of BjPutA was another concern for homologous screening. BjPutA had severe anisotropic diffraction, which was evident by the 'smearing' of diffraction spots. This led to many crystals needed to obtain one usable data set, as data reduction of many data sets was impossible. This made obtaining structures of mutants a more arduous task. However, if a homolog could provide more reliable diffraction void of these issues, than characterization would be much easier.

A2.2. MATERIALS AND METHODS

A2.2.1 Protein Expression and Purification

LpPutA was expressed as described previously for BjPutA with a few exceptions(2). First, instead of BL21 DE3 (pLysS) cells for expression, BL21 DE3 Star cells were used. After induction the temperature was reduced to 25C and rpm reduced to 220. After pelleting and resuspending cells in the binding buffer for Ni-NTA, the cells were flash frozen and stored at -80 °C. Frozen cells were thawed in the presence of 5 protease inhibitors: 0.1 mM TPCK, 0.5 mM AEBSF, 0.005 mM E-64, 0.001 mM pepstain, and 0.01 mM leupeptin. Cells were disrupted in a French pressure cell at 110 MPa. The cell debris and insoluble material were removed by centrifugation in an SS-34 rotor at 16,500 rpm for 60 minutes and the supernatant decanted off. Purification of the enzyme is the same as that described for BjPutA previously(2).

A2.2.2 Crystallography

Screening attempts using Crystal Screen I and II, Index, and Wizard I-III were set up both in sitting drop, vapor diffusion trays and under Al's oil. Vapor diffusion trays were set up at 10 and 5 mg/mL in a ratio of 1 μ L protein to 1 μ L mother liquor. Under oil was set up in the same way with 8 μ L of Al's oil added to conceal the 2 μ L mixture. All trays were incubated at room temperature.

Wizard II - 42 with a solution of 30% PEG 400, 0.1 M HEPES (pH 7.5), and 0.2 M NaCl had a hit that was optimized. Optimizations of this hit were grown with a range of PEG 400 of 21%-36%, 0.1 M HEPES (pH 6.5 – 8.0) and either with or without NaCl. Crystals took about 2 weeks to appear and grow to a maximum dimension. Crystals were cryoprotected in 36% PEG 400 and 0.1 M HEPES (pH 7.5), mounted on a copper base

pin, flash cooled by plunging into liquid nitrogen, and stored in liquid nitrogen.

A2.2.3 SAXS

SAXS data were collected on Advanced Light Source beamline 12.3.1 through the mail-in program(10, 11). Before SAXS analysis, LpPutA was subjected to size-exclusion chromatography using a Superdex-200 column equilibrated with 0.05 M Tris (pH 8.0), 0.5 M NaCl, and 1 mM THP. The enzyme was then concentrated after SEC using a centrifugal filter. SAXS analysis was performed at three different concentrations (8 mg/mL, 5.3 mg/mL, and 2.7 mg/mL). Scattering intensities were measured with exposure times of 0.5, 1.0, and 3.0 s for each concentration. Scattering curves collected from the protein samples were corrected for background scattering using intensity from effluent of the centrifugal filter and concentration procedure. PRIMUS(12) was used to create a composite scattering curve for each protein concentration. Scaling and merging of the background corrected data with the high q region of the 3.0 s exposure and the low q region of the 1.0 s exposure. PRIMUS was also used to perform Guinier analysis. FoXS(13) was used to calculate theoretical scattering profiles from atomic models.

A2.3. RESULTS

A2.3.1 Expression and Crystallography

Many hits were seen for this enzyme, almost all of them were microcrystals and micro-needles. One crystal obtained was a needle, very long and thin, but was large enough to mount and shoot. This crystal grown in 36% PEG 400 and 0.1 M HEPES (pH 8.0) diffracted to 3.8 Å. Data collection statistics are listed in Table 1. Molecular replacement of the crystal was successful using the dimeric search model of the BjPutA 3HAZ structure. A dimer in the asymmetric unit was seen much like that of BjPutA.

However, instead of a dimer of dimers related by symmetry, a tetramer of dimers was seen related by crystallographic symmetry.

Unfortunately, pursuing higher resolution was cut short. After obtaining this crystal structure the protein became insoluble on subsequent expression attempts. Originally expression was high and very little was insoluble. The plasmid was retransformed into BL21 (DE3), pLysS, Rosetta, and Star cells. Star gave the best expression, but yields were still poor. Further auto induction was tried as plasmid was retransformed into AI cells. However, autoinduction was also unsuccessful.

The enzyme was housed in pET151 vector, which uses a topoisomerase to insert the gene into the plasmid. Attempts were made to remove the gene from this plasmid by placing BAMH1 and Ned1 cleavage sites on each end. However, these attempts were unsuccessful. Instead, protein obtained from expression in BL21 (DE3) Star cells was used to recreate the crystals seen previously. However, the crystal hits were not seen even after rescreening with both His-tag present and removed. Further attempts to pursue this crystal form were dropped.

A2.3.2 Molecular Replacement and PISA

To solve the phasing issue molecular replacement was done using MolRep. Pointless had already determined the Laue symmetry to be in I4, however the question remained if I4₁ was the real space group. However, at 3.8 Å resolution finding systematic absences will be difficult. Instead MolRep was used to determine which of these symmetries gave a better score (Table 1). A BjPutA monomer using chainsaw was used as the search model. The score and contrast from MolRep indicates that I4 is most likely the correct solution. The Matthews coefficient suggests that in I4 that the most likely two

or three monomers in the asymmetric unit. Two was used based off of knowledge of dimerization seen broadly in PutAs.

The results of the MolRep model show an octomer, or a tetramer of dimers, related by 4 fold symmetry. However, PISA analysis suggests that a dimer is what is stable in solution, implying that the presence of the octomer is a result of crystal packing. In contrast, the BjPutA tetramer was found to be stable in solution. Indeed, when the octomer scattering profile is calculated using the FoXS server, it does not match the experimental profile found for LpPutA.

A2.3.3 SAXS

The oligomeric state of LpPutA was investigated using SAXS. Guinier plots of three different protein concentrations gave R_g values of 46-48 Å. $P(r)$ distributions gave an even narrower range for R_g at 46-47 Å and gave a maximum particle dimension of 150-160. When compared to the R_g values of BjPutA (30% identical), LpPutA more closely resembles the U-shaped dimer, as opposed to the tetramer. The BjPutA scattering curve has a characteristic dip and peak around a q range of 0.05 which was absent from the LpPutA scattering profile.

Without a structure of LpPutA the only full length structure up to that point, BjPutA, was used as a model for comparison. The FoXS server indicates that the best fit is the BjPutA dimer with a chi value of 6.31. The protomer had a chi value of 24.67 and the tetramer had a chi value of 66.45. The MES only improves this chi to 5.16 when splitting between 13% protomer and 87% dimer. Shape reconstructions run with DAMMIN(17) and the SASTBX server both give good correlation with the dimer found in the asymmetric unit of BjPutA.

A2.4. DISCUSSION

BjPutA had been a tetramer, more specifically a dimer of dimers, in solution after the discovery of the existence of the tetramer by crystallographic symmetry in the structure. The low resolution structure obtained from LpPutA, assuming molecular replacement was correct, indicated that this enzyme was an octomer, or tetramer of dimers, related by crystallographic symmetry. This was a surprising find as such a high order oligomer is unusual. SAXS analysis was done and the results indicated a dimer in solution, meaning that the octomer seen was most likely a crystal artifact or an indication that molecular replacement did not work.

However, being a minimalist PutA the expectation was that LpPutA would mimic what was seen in BjPutA previously. However, as indicated by the SAXS data the presence of a tetramer was not seen. Even in the MES calculation there was no presence of a tetramer, rather a presence of dimer and protomer. The protomer seems unlikely as dimerization is needed to cover the cavity. However, the absence of a tetramer in solution led to question which of these enzymes the outlier was. As it turns out, BjPutA ended up being a unique case as tetramerization was not seen in any other PutAs studied, but rather most of the minimalist PutAs formed dimers in solution much like LpPutA.

Further, the tetramerization domain is almost completely disrupted by removing the C-terminal domain of BjPutA. The cause of this cannot be easily investigated as the C-terminal domain is disordered in the BjPutA structure. A comparison to a minimalist homolog such as LpPutA would have been perfect for modelling the C-terminal and perhaps shedding light on this question. It is unfortunate that the crystallization of LpPutA was not reproducible and had to be stopped.

Table A2.1. MolRep Space Group Determination Results.

	SG #	Space Group	Score	Contrast
1	79	<i>I</i> 4	0.252	9.676
2	80	<i>I</i> 4 ₁	0.195	3.899

Table A2.2. Data Collection Statistics

	LpPutA
Spacegroup	<i>I4</i>
Unit-cell Parameters (Å)	<i>a, b</i> = 199.80 <i>c</i> = 134.23
Wavelength	1.0000
Resolution	74.38 – 3.80 (4.01-3.08)
No. of Observations	103636 (15345)
Unique Reflections	26021 (3785)
R _{merge} (I)	0.189 (0.465)
R _{meas} (I)	0.255 (0.628)
R _{pim} (I)	0.170 (0.420)
Mean I/σ(I)	5.9 (3.0)
Completeness (%)	99.7 (99.9)
Multiplicity	4.0 (4.1)

Figure Legends

Figure A2.1. A) The scattering profile collected from SAXS represented by the black circles. The curves represent the calculated scattering curve for the dimer (blue) and the tetramer (green) superimposed by the FoXS server. The inset represents the Guinier plots for all three concentrations used. Their calculated R_g values are shown above the Guinier plots. B) The DAMMIN reconstructions superimposed with the B_jPutA model. The two figures are rotated 90° along the x axis from each other.

Figure A2.2 The stable oligomeric dimer as determined by PISA.

Figure A2.3 The octomer formed by four-fold symmetry in the crystal lattice.

Fig A2.1.

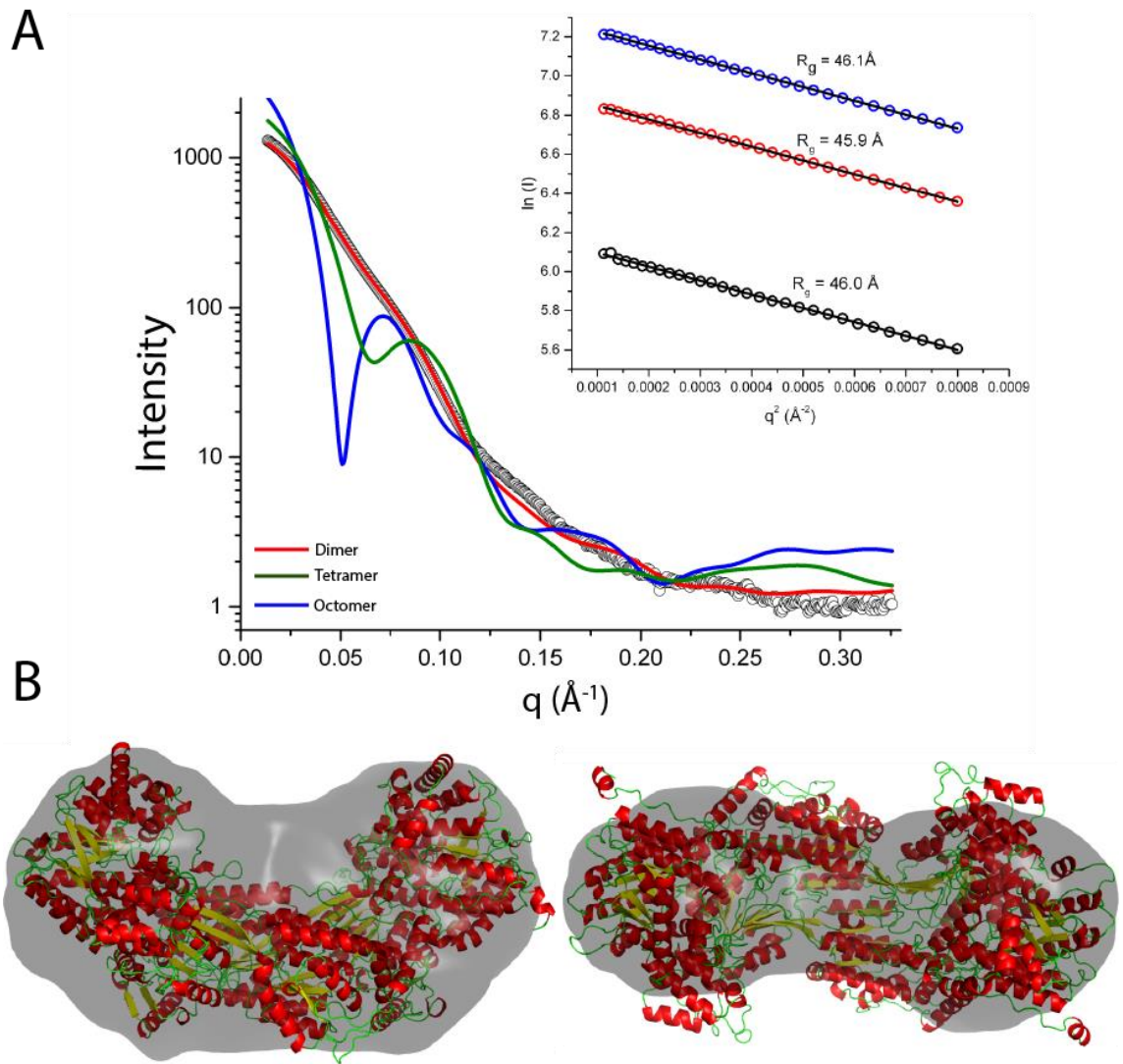


Fig. A2.2

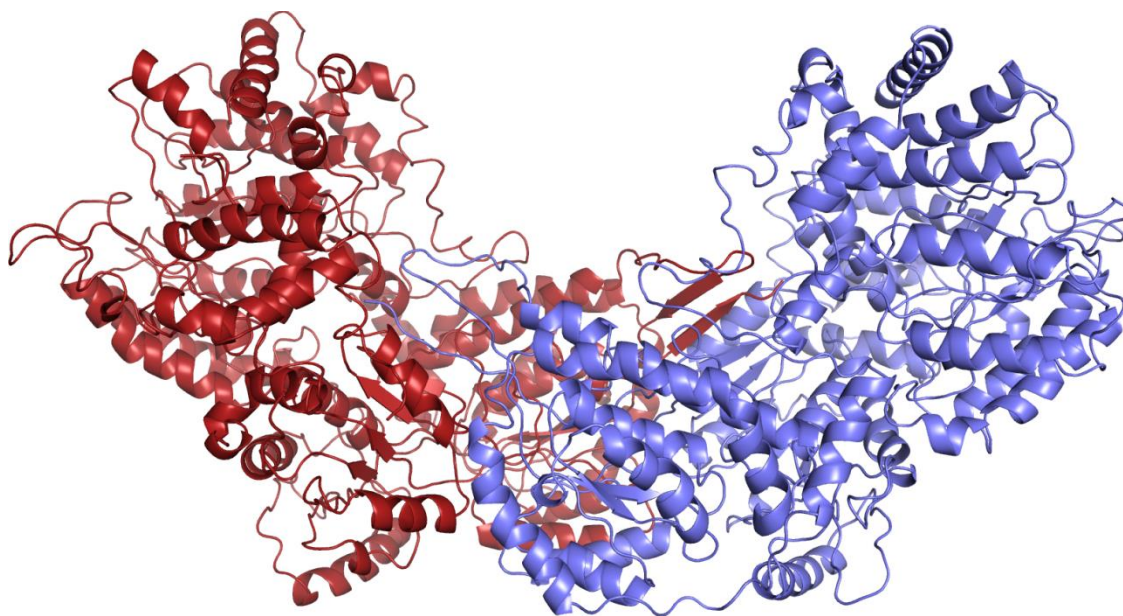
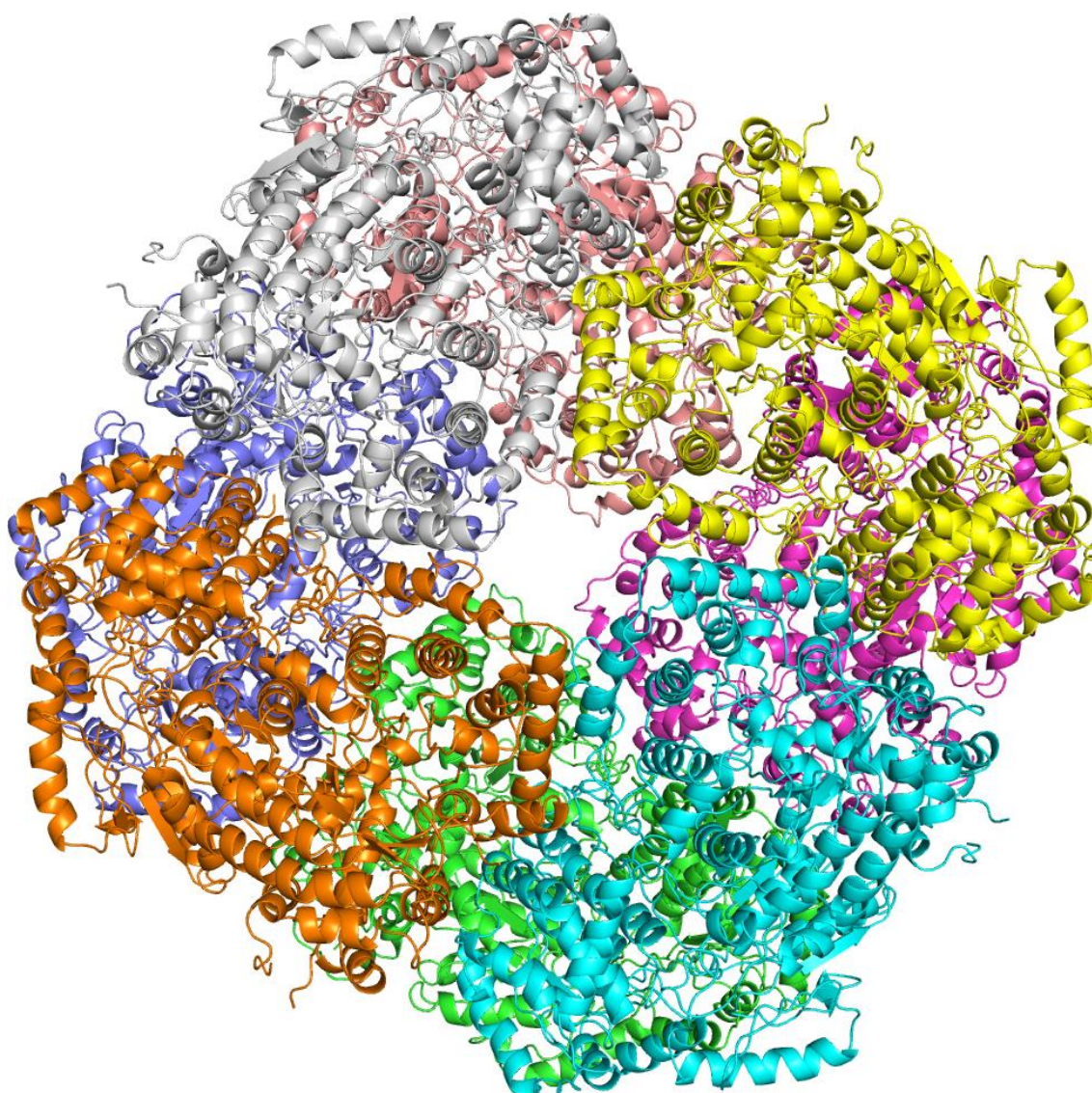


Fig. A2.3



A3.

EXPRESSION, PURIFICATION, CRYSTALLIZATION AND
SAXS ANALYSIS OF *DESULFOVIBRIO VULGARIS* PUTA
(DVPUTA)

A3.1. RATIONALE

Proline catabolic enzymes have been classified into three separate branches. Branch 3 houses the monofunctional enzymes, PRODH and P5CDH. The fused enzymes are divided into two separate branches, branches 1 and 2. Branch 1 is primarily PutAs from alpha-, beta-, and gamma- proteobacteria. This branch has PutAs ranging in size from 999-1400 residues with sequence identities greater than 38%. Branch 2 consists mostly of epsilon-proteobacteria and cyanobacteria PutAs. These range in size from 980-1300 residues and have a sequence identity greater than 23%. The sequence identities between the two branches are generally less than 30%. However, the active sites for PRODH and P5CDH are highly conserved, but with the large gap in identity it is clear these two branches have variations. Further branch 1 contains not only bifunctional enzymes, but have a subset of trifunctional PutAs as well, such as that for EcPutA. Trifunctional PutA has the additional function of repressing its own gene. Branch 2 does not have these larger trifunctional enzymes.

Up until recently, the only full length PutA structure solved was that of BjPutA, which is a minimalist, branch 1 PutA. EcPutA has some individual domain structures, but a full length trifunctional PutA has yet to be fully realized. Additionally, there are no branch 2 PutAs that have been fully sequenced. Considering that branch 2 appears to be a much more diverse grouping, structural data on this group of enzymes would be beneficial. Branch 2 houses the *Helicobacter pylori* and *hepaticus* PutAs(18), which have been shown to be more reactive with oxygen than other PutAs. This activity with oxygen is structurally unclear because BjPutA does not have such reactivity. *Desulfovibrio*

vulgaris PutA (DvPutA) is a branch 2 PutA and crystallization of this enzyme could lead to a better understanding of this diverse family of enzymes(19, 20).

Further, as with the BjPutA homolog, LpPutA, a new crystal form of a PutA that lacks a high concentration of sulfate would be beneficial. Ligand soaking and structural mutant studies have been challenging with the BjPutA because of its anisotropic diffraction and presence of sulfate in both active sites. A new crystal form could solve the poor diffraction problem and a smaller concentration of salt has the potential to yield better data and a deeper understanding of the mechanistic functions of this enzyme.

A3.2. MATERIALS AND METHODS

A3.2.1 Protein Expression and Purification

Protein was grown in LB Broth with 100 µg/mL kanamycin. The cultures were allowed to grow at 37 °C and an rpm of 250 to an OD600 of 0.5 and then induced with 50 µM IPTG. After induction the temperature was dropped to 18 °C and rpm reduced to 200. Cells were left overnight and pelleted the following morning in an SLC-6000 rotor at 3500 rpm. The pelleted cells were resuspended with the binding buffer for Ni-NDA containing 20 mM HEPES pH 8.2, 50 mM NaCl, and 5% glycerol. The cells were flash frozen in liquid nitrogen and stored at -80 °C.

Frozen cells were thawed and broken with 4 rounds of sonication. Each round involved 30 seconds on and 2 minutes of cooldown time for the sample. Debris and insoluble material was removed through centrifugation using an SS-34 rotor at 16,500 rpm. The supernatant was decanted and filtered through a 0.45 µm, syringe-driven filter.

This filtered supernatant was loaded onto a Ni-NTA column with the binding buffer described above. Much of the impurities were removed from sample while loading

as nothing bound. The column was then washed with 30 mM imidazole to remove extra impurities. After removing all impurities at 30 mM, the concentration was increased to 300 mM imidazole at which point DvPutA eluted off the column. Eluted enzyme was collected in 2 mL aliquots and then pooled. The His tag was not cleaved and the pooled fractions were put into a dialysis buffer containing 50 mM Tris, 0.5 mM EDTA, 0.5 mM THP and 5% glycerol at pH 7.5. The protein was allowed to dialyze for over 12 hours.

The dialyzed enzyme was loaded onto a Hi-trap anion exchange column with a binding buffer that is the same as the dialysis buffer. After the enzyme was loaded a gradient of 0-1 M NaCl was run over 100 mL. The protein eluted at approximately 300 mM NaCl and the eluted protein was collected in 500 μ L samples. These samples were checked for purity on SDS-PAGE. Pure samples were pooled and put into a pre-crystallization dialysis buffer consisting of 50 mM Tris (pH 7.5), 50 mM NaCl, 0.5 mM EDTA, and 0.5 mM THP. The enzyme was allowed to dialyze overnight and after removal enzyme concentration was checked using the BCA method. The enzyme was then concentrated up to about 13 mg/mL using a centrifugal concentrator and verified through another round of the BCA method(9).

A3.2.2 Crystallography

This enzyme was screen using Crystal Screen I and II, Index, and Wizard I-III in sitting drop, vapor diffusion trays. Wells consisted of 1 μ L of 10 mg/mL DvPutA mixed with 1 μ L of mother liquor. Crystal trays were incubated at room temperature and another set were incubated at 4°C. Microbatch trays under Al's oil were set up with the same 1:1 ratio of protein and mother liquor and incubated at room temperature.

A3.2.3 SAXS

SAXS data were collected on Advanced Light Source beamline 12.3.1 through the mail-in program. Before SAXS analysis, LpPutA was subjected to size-exclusion chromatography using a Superdex-200 column equilibrated with 0.05 M Tris (pH 8.0), 0.5 M NaCl, and 1 mM THP. The enzyme was then concentrated after SEC using a centrifugal filter. SAXS analysis was performed at three different concentrations (8 mg/mL, 5.3 mg/mL, and 2.7 mg/mL). Scattering intensities were measured with exposure times of 0.5, 1.0, and 3.0 s for each concentration. Scattering curves collected from the protein samples were corrected for background scattering using intensity from effluent of the centrifugal filter and concentration procedure. PRIMUS(12) was used to create a composite scattering curve for each protein concentration. Scaling and merging of the background corrected data with the high q region of the 3.0 s exposure and the low q region of the 1.0 s exposure. PRIMUS was also used to perform Guinier analysis. FoXS(13) was used to calculate theoretical scattering profiles from atomic models.

A3.3. RESULTS

A3.3.1 Crystallography

Several hits were obtained as a result of the many crystal screens set up and many of them were investigated. However, the most widely reproducible and largest crystals came from 6-16% PEG 8,000, 8% ethylene glycol, and 0.1 M HEPES pH 7.0 – 8.5. These were very large, hollow, and hexagonal-looking crystals. Hollow crystal types had been reported before, in fact hollow crystals were the only Put2 crystals that gave usable data(21). There were originally determined to be of the space group $P3$ and have unit cell parameters of $a, b = 163.6$, $c = 214.5$, and $\gamma = 120$. Unfortunately, these crystals had

pseudotranslational symmetry, a serious pathology in crystallography that makes the diffraction data unusable.

The other crystal forms never diffracted to a reasonable resolution and could never be grown much larger. This, in conjunction with GsPutA, another branch 2 PutA with 67% sequence identity with DvPutA, already having a high resolution structure made DvPutA undesirable and ultimately was dropped.

A3.3.2 Molecular Replacement and PISA

The 3.5 Å structure of DvPutA was determined using molecular replacement as implemented in MOLREP using the GsPutA dimer as a search model. MOLREP was first used to search all Laue group members, using a GsPutA monomer as a search model (Table 1). The scores indicate that the correct space group is likely $P6_122$. The matthew's coefficient suggests a dimer or monomer in the asymmetric unit. MOLREP was run again using the monomer of GsPutA as the search model. MOLREP determined that only a monomer exists in the asymmetric unit. Analysis using PISA indicated that a dimer was stable in solution, however it was not the expected dimer found in the SAXS data and seen in other PutAs. Instead of a domain swap of the C-terminals there is a loose association near the PRODH domains.

A3.3.3 SAXS

The oligomeric state was investigated using SAXS. The scattering profiles of DvPutA resembled those of GsPutA. Guinier analysis of the profile yields a radius of gyration (R_g) value of 43.6-43.9 Å. Without a structure of DvPutA to compare to, the closest homolog was used as a comparison, which is GsPutA with a sequence identity of 67%. The calculated R_g value for the dimer in the asymmetric unit of GsPutA is 44.6 Å.

The overlapping of the scattering curves reveals that the calculated GsPutA dimer scattering is a close fit with that of DvPutA, however the calculated BjPutA tetramer is a poor lap.

It is not that surprising to find DvPutA present as a dimer in solution. The initial discovery of GsPutA being dimeric in solution, was surprising since previously the full length PutA was seen as a tetramer. However, as more data has been collected on the oligomeric state of this family of enzymes, it is becoming apparent that BjPutA was the exception. With DvPutA being more identical to GsPutA and being a branch 2 PutA as opposed to the branch 1 BjPutA, the expected outcome was DvPutA being a dimer.

References

1. Srivastava, D., Schuermann, J. P., White, T. A., Krishnan, N., Sanyal, N., Hura, G. L., Tan, A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2010) Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*, *Proceedings of the National Academy of Sciences of the United States of America* 107, 2878-2883.
2. Schuermann, J. P., White, T. A., Srivastava, D., Karr, D. B., and Tanner, J. J. (2008) Three crystal forms of the bifunctional enzyme proline utilization A (PutA) from *Bradyrhizobium japonicum*, *Acta crystallographica. Section F, Structural biology and crystallization communications* 64, 949-953.
3. Bergfors, T. (2003) Seeds to crystals, *Journal of structural biology* 142, 66-76.
4. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, *Acta crystallographica. Section D, Biological crystallography* 67, 271-281.

5. Evans, P. (2006) Scaling and assessment of data quality, *Acta crystallographica. Section D, Biological crystallography* 62, 72-82.
6. Adams, P. D., Afonine, P. V., Bunkoczi, 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 crystallographica. Section D, Biological crystallography* 66, 213-221.
7. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta crystallographica. Section D, Biological crystallography* 66, 486-501.
8. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta crystallographica. Section D, Biological crystallography* 66, 12-21.
9. Walker, J. M. (1994) The bicinchoninic acid (BCA) assay for protein quantitation, *Methods Mol Biol* 32, 5-8.
10. Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., 2nd, Tsutakawa, S. E., Jenney, F. E., Jr., Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S. J., Scott, J. W., Dillard, B. D., Adams, M. W., and Tainer, J. A. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS), *Nature methods* 6, 606-612.
11. Classen, S., Hura, G. L., Holton, J. M., Rambo, R. P., Rodic, I., McGuire, P. J., Dyer, K., Hammel, M., Meigs, G., Frankel, K. A., and Tainer, J. A. (2013) Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source, *J Appl Crystallogr* 46, 1-13.
12. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *Journal of Applied Crystallography* 36, 1277-1282.
13. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments, *Biophysical journal* 105, 962-974.

14. Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta crystallographica. Section D, Biological crystallography* 50, 178-185.
15. Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., and Damborsky, J. (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures, *PLoS computational biology* 8, e1002708.
16. Singh, H., Arentson, B. W., Becker, D. F., and Tanner, J. J. (2014) Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-channeling tunnel and the quinone-binding site, *Proceedings of the National Academy of Sciences of the United States of America* 111, 3389-3394.
17. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, *Biophysical journal* 76, 2879-2886.
18. Krishnan, N., and Becker, D. F. (2006) Oxygen reactivity of PutA from *Helicobacter* species and proline-linked oxidative stress, *Journal of bacteriology* 188, 1227-1235.
19. Tanner, J. J. (2008) Structural biology of proline catabolism, *Amino acids* 35, 719-730.
20. Singh, R. K., and Tanner, J. J. (2012) Unique structural features and sequence motifs of proline utilization A (PutA), *Front Biosci (Landmark Ed)* 17, 556-568.
21. Pemberton, T. A., Srivastava, D., Sanyal, N., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2014) Structural studies of yeast Delta(1)-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): active site flexibility and oligomeric state, *Biochemistry* 53, 1350-1359.

Table A3.1 MolRep Space Group Determination

	SG #	Space Group	Score	Contrast
1	178	<i>P6₁22</i>	0.247	14.421
2	177	<i>P622</i>	0.079	2.548
3	179	<i>P6₅22</i>	0.153	4.209
4	180	<i>P6₂22</i>	0.076	1.661
5	181	<i>P6₄22</i>	0.080	1.719
6	182	<i>P6₃22</i>	0.150	4.726

Table A3.2. Data Collection Statistics.

	DvPutA
Spacegroup	<i>P</i> 6 ₁ 22
Unit-cell Parameters (Å)	a,b = 163.6 c = 214.9 γ = 120.0
Wavelength	1.541
Resolution	19.71 – 3.50 (3.69-3.50)
No. of Observations	181740 (26918)
Unique Reflections	20902 (3060)
R _{merge} (I)	0.406 (0.589)
R _{meas} (I)	0.430 (0.625)
R _{pim} (I)	0.140 (0.205)
Mean I/ σ (I)	7.1 (3.8)
Completeness (%)	95.9 (3.8)
Multiplicity	8.7 (8.8)

Figure Legends

Figure A3.1. Diffraction pattern of DvPutA collected on the Rigaku IV++ home source detector.

Figure A3.2. A) The scattering profile collected from SAXS represented by the black circles. The curves represent the calculated scattering curve for the dimer (blue) and the tetramer (green) superimposed by the FoXS server. The inset represents the Guinier plots for all three concentrations used. Their calculated R_g values are shown above the Guinier plots. B) The SASTBX reconstructions superimposed with the BjPutA model. The two figures are rotated 90° along the x axis from each other.

Figure A3.3. The oligomeric ensemble predicted by PISA to be stable in solution.

Fig A3.1.

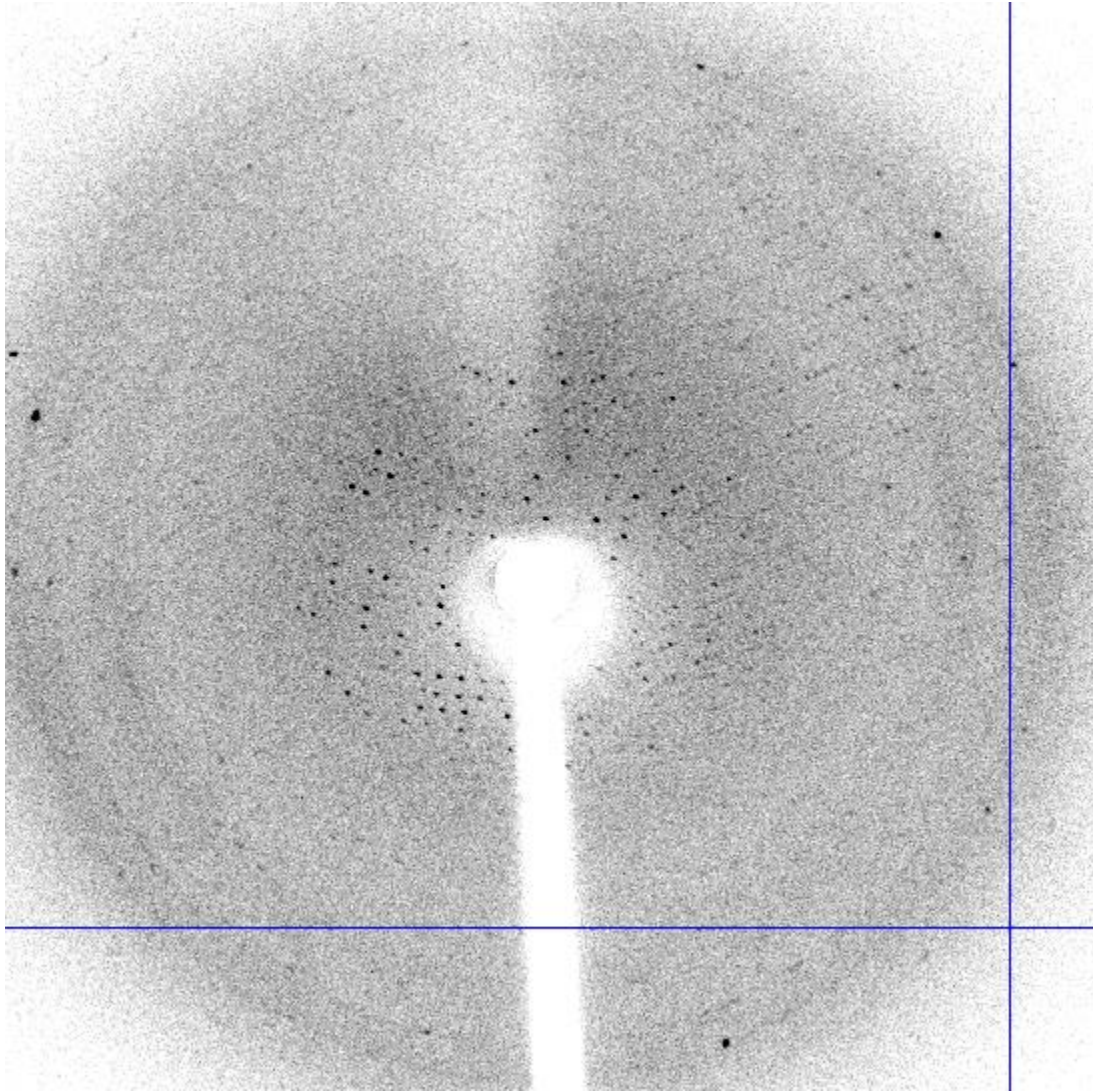
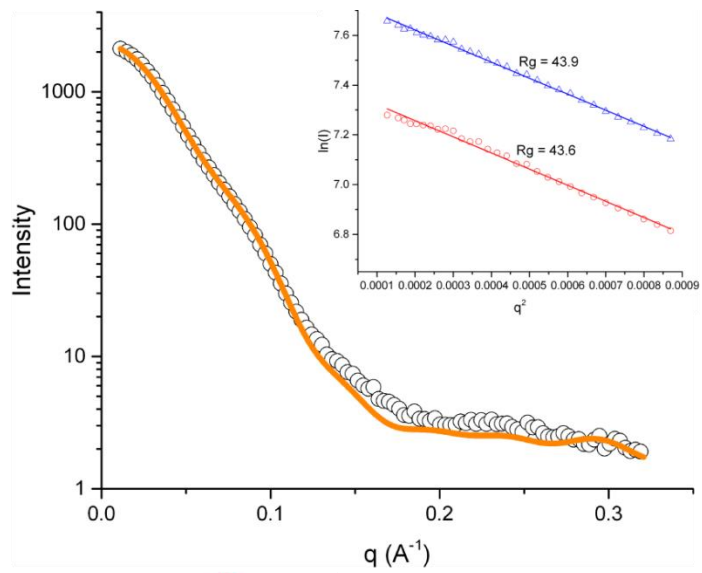


Fig A3.2.

A



B

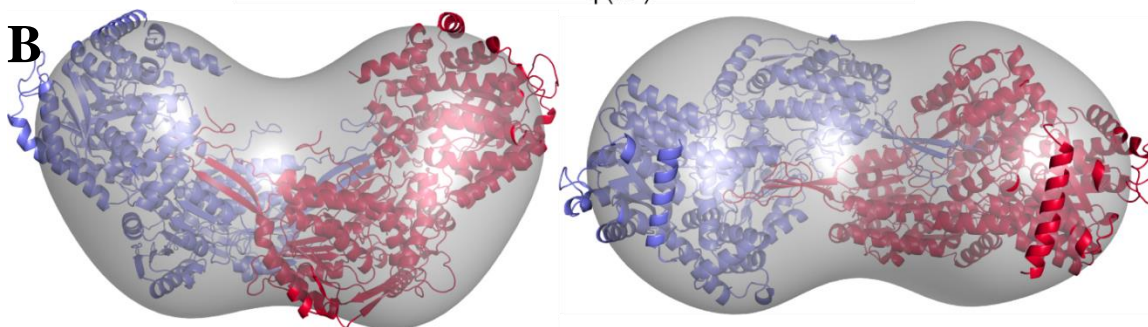
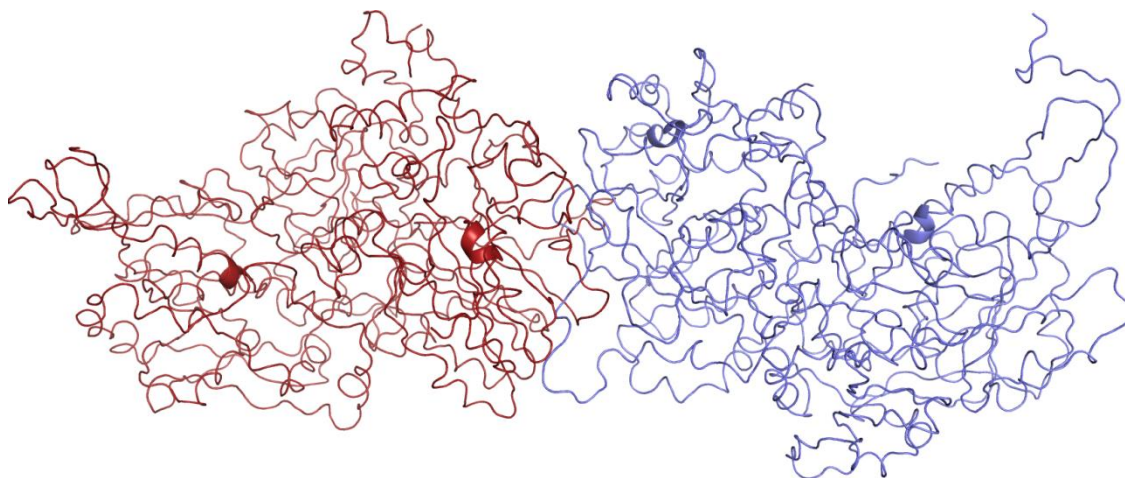


Fig. A3.3.



VITA

Travis Pemberton was born in 1987 in Forthworth, Texas. He grew up in Arizona where he attended high school at Mountain View High School. He then attended college at the University of Missouri where he majored in Biological Sciences. His original plan was to enter Veterinary School, however his interests drifted towards physical sciences and Chemistry. Travis played in Marching Mizzou, where he made many friends that he still has to this day. He graduated in 2005 with a Bachelors of Science in Biology. Upon entering the graduate program in Chemisty at the University of Missouri, he heard Dr. John Tanner talk about his work on structural biology. This was the research Travis knew he wanted to be a part of and joined Dr. Tanner's lab. Throughout his five year stint in the Tanner lab, Travis learned many new techniques, including the art of karaoke. In 2012, he published his first paper which was followed in 2013 by two more papers. In 2014, Travis successfully defended his dissertation and graduated with a Ph.D. Upon graduation, Travis accepted a postdoctoral position at the University of Pennsylvania in Philadelphia. He leaves Missouri after 9 years of school, and the five in Dr. Tanner's lab being some of the most memorable.