Structural and kinetic studies of bifunctional and monofunctional 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 Dhiraj Srivastava

Dr. John J. Tanner
Dissertation Supervisor

July 2011
The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

Structural and kinetic studies of bifunctional and monofunctional proline catabolic enzymes

presented by Dhiraj Srivastava,

a candidate for the degree of doctor of philosophy, and hereby certify that, in their opinion, it is worthy of acceptance.

Professor John J. Tanner

Professor Kent Gates

Professor Lesa Beamer

Professor Renee Jiji
Acknowledgement

There are several people who have shaped my career as a scientist but foremost, I would like to thank my advisor, Prof. John J. Tanner for his continuous guidance and support during my Ph.D. This work would have never been possible without his support. He not only helped me but also encouraged me to try new experiments and to use new techniques. I also want to thank him for several great suggestions about my postdoctoral career.

I should not forget my M.S. advisor, Prof. Rajiv Bhat for introducing me to protein science, my friend Neeraj Mishra for introducing me to several laboratory techniques, and teachers during my Master degree. After Jack, they had the biggest influence on my scientific thinking and career. I would also like to thank my father, Girijesh Srivastava for his trust and for his continuous support throughout my life. He always stands by me in my every decision and gives me freedom to choose my own path.

I thank former and current lab members of the Tanner group for creating such a nice work environment. Especially, I would like to thank Harkewal and John Larson for nice time, Dale for suggestion about gene cloning, and Travis, Shatadru and Ranjan for reading part of my thesis. At the end, I would like to thank Prof. Mike Henzl for helping my doing ITC and analytical ultracentrifugation and my committee members; Prof. Lesa Beamer, Prof. Renee Jiji, and Prof. Kent Gates, for suggestions about experiments and my thesis.

Dhiraj Srivastava
List of Figures

Figure 1.1: Schametic of proline catabolism ................................................................. 8
Figure 1.2: Different domain organization of proline dehydrogenase and P5C dehydrogenase .................................................................................................................. 8
Figure 1.3: Schematic of the proline metabolic pathway ............................................... 8
Figure 1.4: Induction of apoptosis by reactive oxygen species ..................................... 9
Figure 2.1: Spectroscopic analysis of the inactivation of full-length E. coli PutA by PPG. ........................................................................................................................................ 43
Figure 2.2: Kinetics of PutA inactivation by PPG at 25 °C using a dichlorophenolindophenol-based activity assay ................................................................. 44
Figure 2.3: The two views of PPG-inactivated PutA86-630. ...................................... 45
Figure 2.4: Electron density maps showing the (A) overall FAD conformation and covalent attachment to Lys329 and (B) curvature of the isoalloxazine ring. ........ 46
Figure 2.5: Conformational changes during PutA inactivation by PPG detected by limited proteolysis ........................................................................................................ 47
Figure 2.6: Physical binding of PPG-inactivated PutA to E. coli polar lipids ............. 48
Figure 2.7: The FAD conformations observed in three crystal structures of the E. coli PutA PRODH domain ......................................................................................... 49
Figure 2.8: Close-up view of structural differences between PPG-inactivated PutA86-630 and PutA86-669 complexed with THFA ......................................................... 50
Figure 2.9: Stereographic view of a superposition of the active sites of PPG-inactivated
PutA86-630 and PutA86-669 complexed with THFA.............................................. 51

Figure 2.10: Stereographic view of a comparison of PPG-inactivated PutA86-630,
PutA86-669 complexed with THFA, and TtPRODH............................................. 52

Figure 3.1: The reactions catalyzed by PRODH and P5CDH...................................... 79

Figure 3.2: Structure of BjPutA.................................................................................. 80

Figure 3.3: SAXS analysis and tetrameric structure of BjPutA.................................... 81

Figure 3.4: Two views of the cavity, differing in a rotation of ~180° around the vertical
axis. ......................................................................................................................... 82

Figure 3.5: Kinetic evidence for substrate channeling in BjPutA. .............................. 83

Figure S1: Stereo views of the two active sites of BjPutA........................................ 105

Figure S2: Stereo views of the two active sites of BjPutA........................................ 106

Figure S3: Analytical ultracentrifugation data for BjPutA in non-deuterated
and deuterated solvent............................................................................................. 107

Figure S4: Orthogonal views of the BjPutA tetramer............................................... 108

Figure S5: Stereographic view of the cavity............................................................... 109

Figure S6: Rapid-reaction kinetic data for native BjPutA and the mixed variants
non-channeling control.......................................................................................... 110

Figure S7: Residual analysis of stopped-flow kinetic data......................................... 111

Figure 4.1: Protomer structure of human P5C dehydrogenase................................. 157

Figure 4.2: Structure of human P5C dehydrogenase dimer...................................... 157

Figure 4.3: Structural alignment of human and mouse P5C dehydrogenase.............. 158

Figure 4.4: Structural alignment of mouse and Thermus thermophilus P5C
dehydrogenase....................................................................................................... 158
Figure 4.5: Protomer structure of human P5C dehydrogenase with dimer interface residues shown in red .............................................................................................. 159

Figure 4.6: Structure of Mouse P5C dehydrogenase active site showing its direct interactions with sulphate ion ........................................................................................................ 159

Figure 4.7: Structure of Mouse P5C dehydrogenase active site showing its interactions with bound product glutamate ...................................................................................... 160

Figure 4.8: Structural alignment of human P5C dehydrogenase in monoclinic and hexagonal crystal form ........................................................................................................ 160

Figure 4.9: Structure of mouse P5C dehydrogenase complexed with NAD$^+$ .............. 161

Figure 4.10: Structure of mouse P5C dehydrogenase complexed with NADH. .......... 161

Figure 4.11: Structure of human P5C dehydrogenase dimer showing single nucleotide polymorphism ............................................................................................................... 162

Figure 4.12: Structure of Mouse P5C dehydrogenase showing hydrogen-bonding network connecting Ser352 and Cys348 ................................................................. 163

Figure 4.13: Structural alignment of human P5C dehydrogenase (hexagonal form) with S352L mutant ........................................................................................................ 164

Figure 4.14: Stereo representation showing interaction of Glu447 with nicotinic ribose group and possible steric clash in S352L mutant .............................................. 164

Figure 4.15: Michaelis menten kinetics of native human P5CDH at different concentration of NAD$^+$ ........................................................................................................ 165

Figure 4.16: Michaelis menten kinetics of human P5CDH S352A mutant at different concentration of NAD$^+$ ...................................................................................... 165

Figure 4.17: Michaelis menten kinetics of human P5CDH V470I mutant at different concentration of NAD$^+$ ...................................................................................... 166
Figure 4.18: Raw ITC data of human P5C dehydrogenase S352L mutant……………..166

Figure 4.19: Raw ITC data of human P5C dehydrogenase. ........................................... 167

Figure 4.20: Binding isotherm for wild type human P5CDH generated by integrating each peak with respect to time after correcting for the dilution...................................... 167

Figure 4.21: Representative raw ITC data of human P5C dehydrogenase S352A mutant. .................................................................................................................................................................................. 168

Figure 4.22: Binding isotherm for S352A mutant generated by integrating each peak with respect to time after correcting for the dilution. ................................................................. 168

Figure 4.23: Schematics of coupled assay designed to study the oxidation of 3-OH-P5C by P5C dehydrogenase.................................................................169

Figure 4.24: NADH formation by human P5C dehydrogenase at different concentration of proline and hydroxy-proline. ................................................................. 169

Figure 4-25: NADH standard curve..................................................................................170

Figure 4-26: Equilibrium analytical ultracentrifugation data for human P5C dehydrogenase ............................................................................................................. 170

Figure 4-27: Equilibrium analytical ultracentrifugation data for mouse P5C dehydrogenase ............................................................................................................. 171

Figure 4-28: Equilibrium analytical ultracentrifugation data for Put2. ......................... 171

Figure 4-29: Structural alignment of different aldehyde dehydrogenase .................... 172
List of Tables

Table 2.1: Data collection and refinement statistics for PutA86-630 inactivated with NPPG .......................................................................................................................... 36

Table 3.1: Data Collection and Refinement Statistics for BjPutA................................. 84

Table S1: Data Collection and Phasing Statistics for Se-Met Derivative......................... 111

Table S2: Kinetic constants for native BjPutA and monofunctional BjPutA variants................................................................. 112

Table 4.1: Human aldehyde dehydrogenase genes with their respective enzyme, substrate, oligomeric state, and PDB code................................. 148

Table 4.2: Data collection and refinement statistics for HsP5CDH ............................... 149

Table 4.3: Data collection and refinement statistics for MmP5CDHb and Put2 .............. 150

Table 4.4: Kinetic parameters of wild type and mutant HsP5CDH................................. 151
Chapter 4 – Structure and kinetics of human P5C dehydrogenase and its mutants

Materials and Methods

A. Cloning, expression, and purification

1. HsP5CDH —
   a. Constructs
      (1) Parent constructs
      (2) Short construct
   b. Expression and purification
      (1) Parent construct
      (2) Short construct
      (3) Se-Met derivative of human P5C dehydrogenase

2. MmP5CDH
   a. Cloning
   b. Expression and purification
3. Put2 ................................................................. 121
   a. Cloning ................................................................. 121
   b. Expression and purification ........................................ 122
B. Crystallization .......................................................... 123
   1. HsP5CDH .......................................................... 123
      a. Crystal screening ............................................... 123
         (1) Hexagonal crystal form – ................................. 123
         (2) Monoclinic crystal form – ................................ 123
         (3) Crystallization of Se-Met derivative ..................... 123
         (4) Crystallization of HsP5CDH S352L ...................... 124
   2. MmP5CDH ......................................................... 124
   3. Put2 ................................................................. 124
C. X-ray data collection .................................................. 125
   1. Se-Met hexagonal .................................................. 125
D. Phasing and refinement ................................................ 126
   1. HsP5CDH .......................................................... 126
E. Kinetics ................................................................. 128
   1. Standard bi-substrate kinetics .................................. 128
   2. P5C vs. OH-P5C kinetics ........................................... 129
Results – ............................................................................... 131
Discussion: .......................................................................... 140
References – ........................................................................... 145
Chapter 1 - Introduction

Proline is unique among the twenty amino acids because of its cyclic nature. So nature has evolved unique enzymes for metabolism of proline. Proline is catabolised to glutamate by a set of two enzymes, proline dehydrogenase and pyrroline-5-carboxylate (P5C) dehydrogenase. Proline dehydrogenase catalyzed the oxidation of proline to P5C. In the process, FAD gets reduced, transfers its electron to the electron transport chain and subsequently reoxidized for the next round of catalysis. Pyrroline-5-carboxylate is hydrolyzed to glytamyl-\(\gamma\)-semialdehyde (GSA) by non-enzymatic hydrolysis, and is the actual substrate for P5C dehydrogenase. P5C dehydrogenase oxidizes glutamyl-\(\gamma\)-semialdehyde to glutamate with concomitant reduction of NAD\(^+\) to NADH (Fig. 1.1).

Structure and organization of proline catabolic enzymes –

Proline dehydrogenase and P5C dehydrogenase are present in all organisms from archaeabacteria to human. There are three different types of domain organization for these enzymes, which are present in different organisms\(^1\)(Fig. 1.2). In some bacteria, proline dehydrogenase and P5C dehydrogenase activities are on the same polypeptide chain. These bifunctional enzymes are called Proline utilization A (PutA). The examples are PutAs from *Bradyrhizobium japonicum*, *Rhodopseudomonas palustris*, and *Nitrobacter hamburgensis*. Besides proline dehydrogenase and P5C dehydrogenase activity, some PutA, known as trifunctional PutA, also have a DNA binding domain. In the absence of proline, DNA binding domain binds to the promoter of its own gene causing its repression\(^2\). Reduction of proline dehydrogenase domain by proline causes conformational changes in the flavin isalloxazine ring and ribityl side chain, inducing reorganization of the hydrogen bonds around flavin. These structural changes, in and
around FAD, decreases the affinity of PutA to DNA and increases its membrane association$^{3,4}$. Thus it has an autorepressor function. These trifunctional PutAs are present in bacteria such as *Escherichia coli*, *Salmonella typhimurium*, and *Bordetella*. Some bacteria and eukaryotes have these two catalytic activities on separate polypeptide chains for example Thermus thermophilus, yeast, and human. Monofunctional P5C dehydrogenases are 515-575 residues long. However, proline dehydrogenases are 300-550 residue long with the mammalian proline dehydrogenase being longest (551 residue) and bacterial proline dehydrogenases being shortest (300 residue).

The Structures of the DNA binding$^5$ and proline dehydrogenase domain of trifunctional PutA$^6$ from Escherichia coli have been solved by the Tanner group. The structure of the monofunctional proline dehydrogenase from *Thermus thermophilus*, has also been solved$^1$. The 300 residues of *Thermus thermophilus* proline dehydrogenase is the minimum requirement to form a catalytically active proline dehydrogenase. Besides this catalytic core, human proline dehydrogenase also has an about 250 residue long N-terminus of unknown function. Proline dehydrogenases adopt a distorted $\beta\alpha$8 barrel or TIM barrel. It has 8 parallel $\beta$ strands making the center of the barrel with 8 $\alpha$ helix arrayed outside the barrel next to each $\beta$ strand. Rather than sitting alongside the $\beta$8 strand, the $\alpha$8 helix sits on top of the barrel and provides sidechains that interacts with the substrate, proline$^1,6-7$. The FAD cofactor binds at the C-terminal end of the strands of the barrel with its re face tightly packed against the strands and se face available for the hydride transfer from the substrate proline.

**Proline catabolism in eukaryotes** –

In eukaryotes, proline dehydrogenase and P5C dehydrogenase are localized in mitochondria. While proline dehydrogenase is found attached to inner mitochondrial
membrane\textsuperscript{8,8b}, P5C dehydrogenase is a soluble enzyme present in the mitochondrial matrix\textsuperscript{9,10}. Proline metabolism connects three different metabolic pathways and involves two other amino acids\textsuperscript{11} (Fig. 1.3). Proline gets oxidize to glutamate, which can enter the TCA cycle as $\alpha$-ketoglutarate after deamination. Ornithine amino transferase converts ornithine to $\gamma$-glutamyl semialdehyde or pyrroline-5-carboxylate thus connecting proline metabolism to the urea cycle. Oxidation of proline to P5C in mitochondria, transport of P5C to cytosol where it is reduced to proline by P5C reductase and transport of proline back to mitochondria forms a redox shuttle which transfers redox equivalent from the pentose phosphate pathway to the electron transport chain\textsuperscript{12}.

**Role of proline dehydrogenase/oxidase and P5C dehydrogenase in diseases** –

Expression of proline dehydrogenase is induced by P\textsubscript{53}\textsuperscript{14} and peroxisome proliferator-activated receptor $\gamma$ (PPAR$\gamma$)\textsuperscript{15}. Further, a decrease in ATP:ADP ratio and an increase in AMP causes activation of AMP dependent protein kinase (AMPK). AMPK activation elicits the induction of proline dehydrogenase through the down-regulation of mTOR\textsuperscript{16}. Proline dehydrogenase, during the oxidation of proline to pyrroline-5-carboxylate, generates reactive oxygen species in a proline dependent manner\textsuperscript{17}. Reactive oxygen species generated, induces apoptosis through extrinsic and intrinsic pathway\textsuperscript{18} (Fig. 1.4).

The structure of P5C dehydrogenase from *Thermus thermophilus* has been solved by Inagaki *et. al.*\textsuperscript{19}. P5C dehydrogenase is an aldehyde dehydrogenase which catalyzes the oxidation of $\gamma$-glutamyl semialdehyde to glutamate. Several mutations in the human P5C dehydrogenase gene has been reported\textsuperscript{11}. Some of them are known to cause loss of P5C dehydrogenase activity and type II hyperprolinemia\textsuperscript{20}. 
Summary of research:

My Ph.D. research work is divided into three parts. In the first part, I have crystallized and solved the structure of *E. coli* PutA86-630 inactivated with N-Propargyl glycine. N-Propargly glycine irreversibly modifies FAD, mimicking the reduced state of FAD. This structure helped in understanding the conformational changes in the protein associated with FAD reduction. In the second part of my Ph.D work, I have crystallized and solved the structure of the bifunctional PutA from *Bradyrhizobium japonicum*. The structure shows a 40 Å long channel, which connects the proline dehydrogenase and P5C dehydrogenase active site and suggests the direct transfer of intermediate from one active site to the other. This phenomenon, known as substrate channeling, was supported by kinetics. Further, I used equilibrium analytical ultracentrifugation and small angle X-ray scattering to study the oligomeric state of *Bradyrhizobium japonicum* PutA in solution. In the last part, I have crystallized and solved the structure of P5C dehydrogenase from three different source; human, mouse, and yeast (*saccharomyces cerevisiae*). The structure of a human P5C dehydrogenase mutant was solved. Oligomeric states of these P5C dehydrogenases has been studied by equilibrium analytical ultracentrifugation. Effect of different mutations on the catalytic activity of enzyme was studied by kinetics and NAD⁺ binding affinity was measures by isothermal titration calorimetry.
References


Figure legends –

Figure 1.1 – Schematic of proline catabolism. Adopted from *Amino Acids* 2008, 35 (4), 719-30.

Figure 1.2 – Different domain organization of proline dehydrogenase and P5C dehydrogenase.

Figure 1.3– Schematic of the proline metabolic pathway connecting pentose phosphate pathway, urea cycle, and TCA cycle. Figure adopted from Phang et. al. 201013.

Figure 1.4 – Induction of apoptosis by reactive oxygen species, generated by proline oxidase. Figure taken from Phang et. al. 200818.
Figure 1.1: Schematic of proline catabolism

Figure 1.2: Different domain organization of proline dehydrogenase and P5C dehydrogenase.

- Trifunctional PutA (E. coli) 1320
- Bifunctional PutA (B. japonicum) 999
- Monofunctional Proline dehydrogenase (T. thermophilus) 307
- Monofunctional proline dehydrogenase (Human) 600
- Monofunctional P5C dehydrogenase (T. thermophilus, Human) 516

Figure 1.3: Schematic of the proline metabolic pathway
Figure 1.4: Induction of apoptosis by reactive oxygen species
Chapter 2 - Structure of the Proline Utilization A Proline Dehydrogenase Domain Inactivated by \textit{N}-propargylglycine Provides Insight into Conformational Changes Induced by Substrate Binding and Flavin Reduction$^†,‡$


Departments of Chemistry and Biochemistry, University of Missouri-Columbia, Columbia, MO, 65211, Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE 68588, and Division of Medicinal Chemistry, College of Pharmacy, The University of Texas, Austin TX, 78712

$^†$This research was supported by the National Institutes of Health Grants GM065546 (to JJT), GM061068 (to DFB), and GM041239 (to CPW). This research is a contribution of the University of Nebraska Agricultural Research Division, supported in part by funds provided through the Hatch Act. This publication was also made possible by NIH Grant Number P20 RR-017675-02 from the National Center for Research Resources. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

$^‡$Coordinates and structure factors have been deposited in the Protein Data Bank under Accession Number 3ITG.

$^§$Department of Chemistry, University of Missouri

$^‖$University of Nebraska
ABBREVIATIONS

PutA, proline utilization A; PRODH, proline dehydrogenase; P5C, Δ1-pyrrrole-5-carboxylate; P5CDH, Δ1-pyrrrole-5-carboxylate dehydrogenase; PPG, N-propargylglycine; TtPRODH, Thermus thermophilus proline dehydrogenase; PutA86-630, protein corresponding to residues 86 – 630 of E. coli PutA; PutA86-669, protein corresponding to residues 86 – 669 of E. coli PutA; MES, 2-(N-morpholino)ethanesulfonic acid; THFA, L-tetrahydoro-2-furoic acid; SDS, sodium dodecyl sulfate; HEPES-N, N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid.

ABSTRACT

Proline utilization A (PutA) from *Escherichia coli* is a flavoprotein that has mutually exclusive roles as a transcriptional repressor of the *put* regulon and a membrane-associated enzyme that catalyzes the oxidation of proline to glutamate. Previous studies have shown that the binding of proline in the proline dehydrogenase (PRODH) active site and subsequent reduction of the FAD trigger global conformational changes that enhance PutA-membrane affinity. These events cause PutA to switch from its repressor to enzymatic role, but the mechanism by which this signal is propagated from the active site to the distal membrane-binding domain is largely unknown. Here, it is shown that N-propargylglycine irreversibly inactivates PutA by covalently linking the flavin N(5) atom to the ε-amino of Lys329. Furthermore, inactivation locks PutA into a conformation that may mimic the proline reduced, membrane-associated form. The 2.15 Å resolution structure of the inactivated PRODH domain suggests that the initial events involved in broadcasting the reduced flavin state to the distal membrane binding domain include major reorganization of the flavin ribityl chain, severe (35 degree) butterfly bending of the isoalloxazine ring, and disruption of an electrostatic network involving the flavin N(5), Arg431, and Asp370. The structure also provides information about conformational changes associated with substrate binding. This analysis suggests that the active site is incompletely assembled in the absence of the substrate, and the binding of proline draws together conserved residues in helix 8 and the β1-α1 loop to complete the active site.
**Introduction:**

The oxidative conversion of proline to glutamate requires the coupled activities of proline dehydrogenase (PRODH) and Δ¹-pyrroline-5-carboxylate (P5C) dehydrogenase (P5CDH) (Scheme 1). In some bacteria, these activities are combined in a single polypeptide known as proline utilization A (PutA). The polypeptide chain lengths of PutAs exceed 1000 residues, with the PRODH and P5CDH domains appearing in the N-terminal and C-terminal halves of the protein, respectively (1, 2). Some PutAs, including *Escherichia coli* PutA, also serve as transcriptional repressors of the *put* regulon, which comprises the divergently transcribed genes encoding PutA and the high-affinity Na⁺/proline transporter, PutP, plus an intervening 419-bp control DNA region (1, 3). PutA represses transcription of *putA* and *putP* by recognizing, via an N-terminal ribbon-helix-helix domain, a consensus binding sequence that appears at five different locations in the *put* control DNA region. In the absence of proline, PutA occupies the five operator sites, thereby blocking transcription of the *put* genes (3). In the presence of proline, PutA associates with the membrane where it couples the oxidation of proline to P5C with reduction of the electron transport chain (Scheme 1). Following the nonenzymatic hydrolysis of P5C, the P5CDH domain catalyzes the NAD⁺-dependent oxidation of γ-glutamate semialdehyde to glutamate (Scheme 1) (4-6). Thus, the association of PutA with the inner membrane is not only essential for efficient catalytic turnover, but also increases expression of *putA* and *putP* by preventing PutA from engaging its operator sites in the *put* control DNA region (1).

The process by which PutA is transformed from a DNA-bound repressor to a membrane-associated enzyme, known as functional switching, involves conformational changes triggered by proline binding and reduction of the flavin cofactor (1, 7-9). Recently, conformational changes in the FAD were deduced from a structure of the
dithionite-reduced \textit{E. coli} PutA PRODH domain with hyposulfite bound in the proline binding site (7). Upon reduction, the FAD was observed to adopt a new conformation characterized by a significant “butterfly” bend (22°) of the isoalloxazine ring and rotation of the 2′-OH group of the ribityl chain resulting in formation of a new hydrogen bond between the 2′-OH and the FAD N(1) atom (7). The 2′-OH group of the FAD was subsequently demonstrated to act as a redox-sensitive toggle switch that helps control association of PutA with the membrane (7).

A second key interaction involves the FAD N(5) atom and Arg431. Arg431 is strategically placed within hydrogen bonding distance to the flavin N(5) and is thus potentially able to sense changes in electron density across the N(1)-N(5) enediamine system attendant to FAD reduction. Indeed, mutation of Arg431 to Met, or replacement of FAD with 5-deaza-FAD, abrogated the ability of proline to activate binding of PutA to the membrane (7). Thus, Arg431 was proposed to have an important role in transmitting redox signals out of the FAD active site (7), although the dithionite-reduced enzyme structure did not indicate any movement of this residue.

The conformational changes in the FAD deduced from the structure of the dithionite-reduced PRODH domain (bending of the isoalloxazine ring and rotation of the 2′-OH) presumably represent the initial step in transmission of the flavin redox status to a remote membrane-binding domain. Signal transmission likely involves Arg431, but other residues important for broadcasting the flavin redox state remain largely unknown.

Here, we have used mechanism-based inactivation of \textit{E. coli} PutA by \textit{N}-propargylglycine (PPG, 1 in Scheme 2) to elicit conformational changes similar to those involved in functional switching. This work was motivated by our recent study of the inactivation of PRODH from \textit{Thermus thermophilus} (TtPRODH) by PPG (10).
TtPRODH is a 307-residue monofunctional PRODH that has 27 % sequence identity to residues 235 - 569 of the 1320-residue *E. coli* PutA protein. The crystal structure of PPG-inactivated TtPRODH revealed a 3-carbon link between the ε–amine of Lys99 and the N(5) of FAD. In one scenario, the mechanism of inactivation is initiated by the oxidation of PPG to *N*-propargyliminoglycine (2 in Scheme 2) and concomitant reduction of the FAD. It is proposed that attack of 2 by the N(5) of the reduced flavin, rearrangement of a putative allenic species (3) to an imminium intermediate (4), and nucleophilic attack by an active site lysine residue (Lys99) on the imminium intermediate lead to the covalent FAD-Lys99 adduct (5). As a result, the enzyme is, effectively, locked into the reduced state. Sequence alignments show that Lys99 is highly conserved among monofunctional PRODHs and PutAs, with Lys329 being the corresponding residue in *E. coli* PutA. Since PPG locks TtPRODH in the reduced state, we investigated the possibility of using PPG to stabilize the reduced conformation of PutA in order to facilitate examination of this important conformation. Results of these studies are described herein.

**EXPERIMENTAL PROCEDURES**

**Materials.** *N*-propargylglycine (PPG) was synthesized as reported (10). *E. coli* polar lipid extracts, phosphatidylglycerol, and 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine were purchased from Avanti Polar Lipids Inc. and used without purification. Full-length *E. coli* PutA and an *E. coli* PutA PRODH domain construct containing residues 86-630 (PutA86-630) were expressed and purified as previously reported (8, 11, 12). All chemicals and buffers were purchased from Fisher Scientific and Sigma-Aldrich, unless stated otherwise. All experiments used Nanopure water.
**Crystallization.** Crystallization experiments were performed with PutA86-630, which was purified by Ni-iminodiacetic acid chromatography and HiTrapQ anion exchange chromatography as described previously (12). After ion exchange, the protein was dialyzed into 50 mM Tris-HCl buffer, 50 mM NaCl, 0.5 mM EDTA, and 5 % glycerol at pH 7.5. The sample was then concentrated to ~12 mg/mL as estimated by the bicinchoninic acid method (Pierce kit). The protein was incubated with PPG in a ratio of one mg of enzyme per one mg of PPG (~560 molar excess of PPG) for at least 30 min at 4 °C. Crystallization experiments were performed at room temperature using the method of vapor diffusion. Initial crystallization conditions were identified using sitting drops and commercially available crystal screen kits. These experiments revealed promising crystals in solutions containing PEG 5000 monomethyl ether, ammonium sulfate, and 2-(N-morpholino)ethanesulfonic acid (MES) buffer. Following a few rounds of optimization, further improvement was obtained using an additive screen approach in which crystal screen reagents served as additive solutions. The base condition consisted of 26 % (w/v) PEG 5000 monomethyl ether, 0.15 M ammonium sulfate, and 0.1 M MES buffer at pH 5.8. Reservoir solutions were created by mixing the base condition and crystal screen reagent in the ratio of three parts base condition to one part crystal screen solution. The best results were obtained in hanging drops using Index reagent 82 as the additive (0.2 M MgCl$_2$, 0.1 M Bis-Tris buffer, 25 % (w/v) PEG 3350, pH 5.5). In preparation for low-temperature data collection, the crystals were soaked in 22 % (w/v) PEG 5000 monomethyl ether, 112.5 mM ammonium sulfate, 50 mM MgCl$_2$, 9 % (w/v) PEG 3350, 100 mM MES buffer pH 5.8, and 20 % (v/v) PEG200, picked up with Hampton mounting loops and plunged into liquid nitrogen.

**X-ray Diffraction Data Collection, Processing, and Refinement.**

Diffraction data were collected at beamline 4.2.2 of the Advanced Light Source. The data set used for structure determination and refinement consisted of 200 frames collected with oscillation width of 0.5° per frame and detector distance of 160 mm. The data were
integrated with MOSFLM (13) through the iMosflm graphical interface, and scaled to 2.15 Å resolution with SCALA (14) using the CCP4i interface (15). The space group is $I2_12_12_1$ with unit cell dimensions of $a = 127.5$ Å, $b = 133.4$ Å, and $c = 133.6$ Å. There are two molecules in the asymmetric unit, which implies a solvent content of 47% and $V_m$ of 2.3 Å$^3$/Da (16). Data processing statistics are listed in Table 2.1. We note that this crystal form differs from the one used in previous crystallographic studies of the *E. coli* PutA PRODH domain, which has space group of $I222$ with unit cell dimensions of $a = 73$ Å, $b = 141$ Å, and $c = 146$ Å and one molecule in the asymmetric unit. (7, 12, 17-19).

The structure was solved using molecular replacement as implemented in MOLREP (20). The search model was derived from the coordinates of the *E. coli* PutA PRODH domain complexed with the inhibitor L-tetrahydro-2-furoic acid (THFA, PDB entry 1TIW, (19)). The structure was refined with PHENIX (21). Model building was performed with Coot (22).

The final model includes two protein chains, a covalently modified FAD cofactor bound to each protein chain, and 241 water molecules. Each protein chain includes residues 87-137, 258-555, and 561-610. As with all other structures of the *E. coli* PutA PRODH domain (7, 12, 18, 19), electron density for residues 138-257 is weak, and thus these residues were not modeled. As described in Results, inactivation by PPG induces substantial disorder in $\alpha 8$ (residues 545 - 562). Because of the weak electron density in this region of the structure, residues 549 - 555 were modeled as polyalanine, and residues 556 - 660 were not modeled. The two protein chains have identical conformations within experimental error (root mean square deviation of 0.12 Å for C$\alpha$ atoms). Likewise, the active site structures of the two chains are identical. Refinement statistics are listed in Table 2.1.
**Spectroscopic and Kinetic Characterizations of the Inactivation of Full-Length E. coli PutA.** Spectral changes associated with inactivation of full-length PutA by PPG were monitored by a Cary-100 spectrophotometer using the wavelength range 300-600 nm at 25 °C. The absorbance spectrum of oxidized PutA was measured in a quartz cuvette at 10 μM in 50 mM potassium phosphate buffer, pH 7.5 with 10% glycerol. A quartz cuvette with buffer served as the reference. A fresh 50 mM stock solution of PPG was added to the cuvette at a final concentration of 500 μM of PPG. The solution was then briefly mixed by inverting the cuvette. Spectra were collected immediately after returning the cuvette to the cuvette holder and at subsequent 20-s intervals for 0-3 min, 30-s intervals for 3-10 min, 60-s intervals for 10-60 min, and 120-s intervals for 60-90 min. After PPG treatment, the extent of inactivation exceeded 95 %. PPG-inactivated PutA was denatured by adding sodium dodecyl sulfate (SDS) to a final concentration of 0.2% (in 50 mM sodium phosphate buffer, 10% glycerol, pH 7.5) as described elsewhere (23), and the resulting flavin spectrum was recorded. Free FAD was then separated from denatured PutA by centrifugation using a 10-kDa cut-off membrane. The molar extinction coefficients for the flavin compounds with maximum absorbance at 378 and 397 nm were estimated using the flavin concentration determined for native PutA ($e_{451} = 12700 \text{ M}^{-1} \text{ cm}^{-1}$) prior to inactivation.

The kinetic parameters for PRODH activity of full-length PutA were determined as described previously for TtPRODH (10). Accordingly, a fresh 50 mM stock solution of PPG dissolved in 50 mM potassium phosphate buffer, pH = 7.5, 10% glycerol, was prepared. *E. coli* PutA (5 μM) was incubated at 25 °C with PPG at concentrations of 0.25, 0.5, 0.75, 1.0, and 2.5 mM. Aliquots were removed at 15-min intervals over a period of 120 min and assayed at 25 °C using a previously described PRODH assay.
based on the reduction of dichlorophenolindophenol with L-proline (80 mM) as the substrate. Inactivation parameters were estimated using the standard method of extracting half-lives from plots of the natural logarithm of activity versus time and making a Kitz and Wilson replot (24, 25).

**Conformational Changes of Full-Length PutA Accompanied by N-propargylglycine Inactivation.** Limited proteolysis of full-length PutA by chymotrypsin was performed as previously described with the following modifications (26). In all digests, PutA was pre-incubated with 5 mM proline or 5 mM THFA, for 5 min prior to adding chymotrypsin except for PPG, which was incubated with PutA for 30 min. Proline and THFA were dissolved in 50 mM potassium phosphate buffer (pH 7.5). PutA (1 mg/mL) was then digested with chymotrypsin (10 µg/mL) for 1 h at 25 °C in 50 mM potassium phosphate buffer (pH 7.5) containing 10% glycerol. The reactions were stopped by adding phenylmethylsulfonyl fluoride (2 mM final concentration) and hot SDS-PAGE sample buffer at a 1:2 (v/v) ratio of SDS buffer to sample. The samples (7.5 µg each) were analyzed by SDS-PAGE using 8% acrylamide separating gels. As a control, oxidized PutA (2.5 µg) without digestion was also run on the gel. Gels were visualized after Coomassie Blue G-250 staining and imaged using a Bio-Rad Gel Doc 2000 System. The molecular weight marker was the Precision Dual Color reagent (Bio-Rad).

**Assays of Full-Length PutA Binding to E. coli Polar Lipids.** The binding buffer used in these experiments consisted of 10 mM N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid containing 150 mM NaCl at pH 7.4 (HEPES-N buffer). E. coli polar lipid vesicles were prepared as previously described using a LiposoFast microextruder with a 100 nm polycarbonate filter at a final concentration of 5 mg/ml (8). After pre-incubation of full-length PutA (0.25 mg/ml) with
proline (5 mM, 5 min), THFA (5 mM, 5 min), or PPG (5 mM, 30 min) at 25 °C in 150 mM HEPES-N buffer, the freshly prepared E. coli polar lipid vesicles were added to the reaction mixtures to a final concentration of 0.8 mg/mL and incubated at room temperature for 1 h with slow rotation. The lipid-bound fraction was then separated from the solution phase using an air-driven ultracentrifuge (Airfuge, Beckman) at the maximum speed as previously described (27). The supernatant (soluble phase) was removed and denatured with SDS sample buffer and labeled as soluble fractions. The lipid pellet was resuspended in an equal volume of HEPES-N buffer, denatured with SDS sample buffer and labeled as lipid fraction. Non-specific sedimentation of the PutA proteins was not observed in the absence of the lipids. All the soluble and lipid fractions were analyzed with 10 % SDS-polyacrylamide gels, which were stained with Coomassie Blue G-250. Quantitative analysis of gels was performed with a Bio-Rad Gel Doc 2000 System.

**RESULTS**

**Spectral Analysis of the Inactivation of Full-Length E. coli PutA by PPG.** Similar to that observed with TtPRODH, PPG inactivates full-length *E. coli* PutA with a concomitant spectral change (Figure 2.1A). The absorption spectrum of PutA-bound FAD decreases at 451 nm after the addition of PPG. Coinciding with the decrease at 451 nm is a dramatic increase in absorbance at 378 nm (\(\varepsilon \sim 20800 \text{ M}^{-1} \text{ cm}^{-1}\)). The decrease at 451 nm indicates that the FAD is fully reduced, similar to reduction of the flavin by sodium dithionite or the substrate proline. The significant increase at 378 nm, which is not observed with proline or dithionite, suggests a new reduced flavin species is formed. The new species is most likely the Lys329-FAD covalent adduct shown in Scheme 2 (5). Formation of this adduct was confirmed by X-ray crystallography (\textit{vide infra}).
The covalently modified flavin could be separated from the inactivated enzyme by treating PPG-inactivated PutA with a 0.2% SDS solution and isolating the flavin component by membrane filtration. The isolated flavin exhibits a strong absorbance maximum at 397 nm (ε ~ 22750 M⁻¹ cm⁻¹) (Figure 2.1B, dotted curve), which is similar to the spectrum of PPG-inactivated PutA (Figure 2.1B, dashed curve) and distinctly different from the spectrum of oxidized PutA (Figure 2.1B, solid curve). This result suggests that the isolated flavin is covalently modified, albeit separated from the enzyme. Presumably, the Schiff base in (5) of Scheme 3 undergoes hydrolysis to produce (6) of Scheme 3 and the free enzyme as a result of treatment with the SDS buffer.

**Kinetic Analysis of the Inactivation of Full-Length E. coli PutA by PPG.** The inactivation of full-length PutA by PPG was monitored in a time and concentration-dependent manner, as shown in Figure 2.2. Inactivation profiles of PutA by five different PPG concentrations were used to estimate inactivation parameters. The standard method of extracting half-lives from plots of the natural logarithm of activity versus time was used to construct a Kitz and Wilson replot (Figure 2.2, inset). Note that the Kitz and Wilson replot intersects the ordinate above the origin, which suggests that inactivation proceeds with saturation, i.e., that a dissociable complex forms prior to inactivation (25). This complex of enzyme and PPG is analogous to the binding of enzyme and substrate in the E•S complex. The estimated values for $k_{\text{inact}}$ and $K_I$ are 0.13 min⁻¹ and 1.5 mM, respectively.

**Structural Characterization of the Enzyme-FAD Adduct.** The structure of an *E. coli* PutA PRODH domain construct (PutA86-630) inactivated by PPG was solved to confirm formation of a Lys-FAD covalent adduct, i.e., (5) in Scheme 2. The structure was determined at 2.15 Å resolution using molecular replacement (Table 2.1). As
expected, the enzyme displays the distorted ($\beta\alpha)_6$ barrel fold characteristic of the PRODH family (Figure 2.3). The FAD cofactor is bound at the C-termini of the strands of the barrel, with the re face of the isoalloxazine packed against strands 4-6 of the barrel.

Electron density maps clearly indicated that the flavin is covalently modified by PPG (Figure 2.4). Inactivation with PPG results in covalent attachment of Lys329 to the flavin N(5) via a three-carbon linkage (Figure 2.4A). Also, the flavin isoalloxazine ring is distinctly nonplanar, being bent by about 35° around the N(5)-N(10) axis (Figure 2.4B). Such deviations from planarity are referred to as "butterfly" conformations and are associated with reduced states of the flavin. We note that the isoalloxazine ring of the dithionite-reduced enzyme also exhibits a butterfly conformation (bend angle of 22°), and the isoalloxazine ring of the oxidized enzyme is planar. Both the covalent attachment of an active site lysine residue to the flavin N(5) and the butterfly conformation were observed in PPG-inactivated TtPRODH (10).

**Limited Proteolysis Studies of Full-Length E. coli PutA.** As reported previously, oxidized, reduced, and inhibited forms of full-length *E. coli* PutA have unique proteolytic fingerprints (26). Therefore, the PPG-inactivated enzyme was also subjected to limited proteolysis using chymotrypsin. As shown in Figure 2.5, oxidized PutA displays major proteolysis bands at 135 kDa and 111 kDa, whereas the proline-reduced enzyme displays bands at 119 kDa and 111 kDa. PutA inhibited by the non-reducing proline analog THFA exhibits a predominant band at ~ 90 kDa. These results are in excellent agreement with previous studies (26). PutA inactivated with PPG displays prominent bands at 119 kDa and 111 kDa, a pattern similar to that observed with proline-reduced PutA. These results suggest that the global structure of the PPG-inactivated
enzyme resembles that of the proline-reduced enzyme, and that PPG and proline trigger similar conformational changes in PutA.

**Lipid Associations of PPG-Inactivated Full-Length *E. coli* PutA.**

Physical lipid binding assays were performed in which full-length PutA was incubated with *E. coli* polar lipid vesicles, and the soluble and lipid bound protein fractions were separated by centrifugation. Results of this assay for the oxidized, proline-reduced, THFA-bound, and PPG-inactivated forms of PutA are shown in Figure 2.6. As expected, oxidized PutA is detected almost entirely in the soluble fraction, indicating that it does not associate with the lipid vesicles (Figure 2.6A). The corresponding soluble/lipid partitioning ratio is 98%/2% (Figure 2.6B). In contrast, proline-reduced PutA exhibits a soluble/lipid distribution of 35%/65%, indicating that reduction of the FAD by the substrate induces a dramatic increase in lipid binding. THFA also increases PutA-lipid associations relative to the inhibitor-free, oxidized enzyme, but to a lesser extent than proline. The soluble/lipid ratio for THFA-bound PutA is 57%/43%. These results are consistent with surface plasmon resonance studies (8) showing that oxidized PutA displays no binding to *E. coli* polar lipid vesicles, proline induces very tight binding ($K_D < 0.01$ nM), and THFA stimulates moderately high binding ($K_D = 34$ nM). PPG-inactivated PutA displays a soluble/lipid ratio of 34%/66%, which is comparable to that of the proline-reduced enzyme (Figure 2.6B). These results suggest that PPG causes PutA-lipid associations that are similar to those induced by proline.

**Comparison to Other PRODH Structures.** Since PPG mimics the effects of proline on PutA conformation and PutA lipid binding, the structure of PPG-inactivated PutA86-630 was compared to other PutA PRODH structures to gain insight into conformational changes involved in functional switching. Also, because this structure is
the only one of a PutA PRODH domain without a ligand occupying the proline binding site, such a comparison potentially provides clues about conformational changes associated with substrate binding. The relevant structures in the PDB are those of oxidized PutA86-669 (residues 86-669 of *E. coli* PutA) complexed with the proline analog THFA (PDB entry 1TIW) and dithionite-reduced PutA86-669 complexed with hyposulfite (PDB entry 2FZM).

The structure of PPG-inactivated PutA86-630 reveals a new conformation for the FAD ribityl chain of PutAs (Figure 2.7A). The 2'-OH is tucked below the pyrimidine ring and forms hydrogen bonds with the FAD N(1) and Gly435. The 3'-OH is rotated 66° from the 2'-OH and forms a hydrogen bond with the FAD ribose. The 4'-OH is oriented *trans* to the 3'-OH and forms a hydrogen bond with a conserved water molecule buried on the *re* side of the FAD. We note that this ribityl conformation is very similar to that of the FAD in oxidized and PPG-inactivated TtPRODH.

The ribityl hydroxyl groups are arrayed much differently in other structures of the PutA PRODH domain (Figures 2.7B, 2.7C). In the THFA complex, the 2'-OH and 3'-OH groups point into the proline binding site where they form hydrogen bonds with Arg556 and Glu559, respectively (Figure 2.7B). Both residues are conserved in PutAs and are located on α8. As described below, these critical residues are disordered in the PPG-inactivated enzyme. The 4'-OH of the THFA complex is *trans* to the 3'-OH and forms hydrogen bonds with Gly435 and the FAD ribose. The ribityl chain of the dithionite-reduced enzyme is identical to that of the THFA complex, except that the 2'-OH sits below the flavin N(1), as in the PPG-inactivated enzyme (Figure 2.7C).

Significant differences are also observed in the protein components of the active site. Since the THFA complex and the dithionite-reduced enzyme structures have identical
protein conformations, we will restrict the following analysis to a comparison with the THFA complex.

The PPG-inactivated enzyme differs from the other PutA PRODH structures in the conformations of two residues near the FAD N(5) atom, Arg431 and Asp370 (Figure 2.8). In the THFA complex, Arg431 donates a hydrogen bond to the FAD N(5) (3.1 Å) and forms an ion pair with Asp370 (also 3.1 Å). In the PPG-inactivated enzyme, the Arg431 - N(5) distance has increased to 4.2 Å, indicating disruption of this key interaction. Furthermore, electron density maps indicated that Asp370 occupies two conformations in the PPG-inactivated enzyme (Figure 2.8). One of these conformations corresponds to the one observed in the THFA complex (labeled B in Figure 2.8). In the other conformation (labeled A in Figure 2.8), Asp370 is rotated 180° away from Arg431 and toward a cluster of glutamate residues located on the surface of the protein (residues 372, 373, 375). When Asp370 is in conformation A, the ion pair with Arg431 and hydrogen bond with Asn368 are not formed.

Profound structural differences are also observed in two regions important for proline binding, α8 (residues 545-562) and the loop connecting β1 to α1 (residues 287-292). As shown in Figure 2.9, α8 is shifted by 3 Å away from the isoalloxazine ring and substantially disordered in the inactivated enzyme. Electron density in the region expected for α8 was quite weak, and although the density allowed the modeling of residues 545-555 in an α-helical conformation, residues 549-555 were modeled as polyalanine due to a complete absence of density for the side chains. Furthermore, density for residues 556-660 was not apparent, and these residues were omitted entirely from the final model (dashed curve in Figure 2.9). The conformation of the β1 - α1 loop is also quite different in the inactivated enzyme. In the PPG-inactivated enzyme, this loop
is shifted away from the active site by 4 Å, and the side chain of Glu289 is disordered (Figure 2.9). As a result, the Glu289-Arg555 ion pair, which is observed in all other PutA PRODH structures, is broken in the PPG-inactivated enzyme.

Interestingly, the conformations of α8 and the β1-α1 loop are reminiscent of those in the structure of oxidized TtPRODH (PDB code 2G37, (28)), the only other PRODH structure without a ligand bound in the proline binding site. A comparison of these two ligand-free enzymes with the PutA86-669/THFA complex is shown in Figure 2.10. As in PPG-inactivated PutA86-630, α8 of TtPRODH is shifted away from the isoalloxazine ring, the β1-α1 loop is withdrawn from the active site, and the conserved ion pair (Glu65-Arg288) is broken. Consequently, the active sites of the ligand-free enzymes are more open than in the PutA86-669/THFA complex.

**DISCUSSION**

**Mechanism-Based Inactivation of PRODH.** Mechanism-based inactivation of PRODH was first reported in 1993, but this area of research is nascent nonetheless. In the first study, Trisch et al. found that 4-methylene-L-proline decreased the PRODH activity of rat mitochondrial extracts (29). This area of research lay dormant until our report of the inactivation of TtPRODH by PPG in 2008 (10). In our current study, we have further shown that PPG is also an inactivator of the PRODH domain within PutA (Figure 2.2). The kinetic constants for PutA inactivation ($k_{\text{inact}} = 0.13 \text{ min}^{-1}, \ K_i = 1.5 \text{ mM}$) are similar to those of TtPRODH ($k_{\text{inact}} = 0.43 \text{ min}^{-1}, \ K_i = 0.8 \text{ mM}$). Furthermore, as with TtPRODH, a dissociable complex is likely formed between PutA and PPG before inactivation of the enzyme. The structures of the PPG-inactivated active sites of PutA and TtPRODH are also similar. In both cases, the flavin is highly bent and covalently attached to an active site lysine by a 3-carbon linkage. These similarities suggest that
PutA and TtPRODH share a common mechanism of inactivation by PPG. The fact that PPG inactivates both monofunctional PRODH and PutA suggests that it could be a broad-based inactivator of the PRODH family. A critical test of this idea will be whether PPG inactivates monofunctional eukaryotic PRODHs.

**Insight into Conformational Changes of PutA Induced by Substrate Binding and Flavin Reduction.** PPG-inactivated *E. coli* PutA exhibits spectral features, proteolysis patterns, and lipid binding properties similar to those of proline-reduced PutA, implying that PPG and proline induce analogous conformational changes in PutA. Therefore, we suggest that the structure of PPG-inactivated PutA86-630 is relevant for understanding the conformational changes involved in functional switching of PutA. In particular, the structure may provide insight into the changes occurring in the vicinity of the flavin that help to broadcast the reduced flavin redox state to the distal membrane-binding domain.

The PPG-inactivated enzyme structure may also be relevant for understanding conformational changes associated with substrate binding, or equivalently, product release, because it is the first one of a PutA PRODH domain without a ligand in the proline binding site. All structures of the oxidized PutA PRODH domain have a competitive inhibitor, such as THFA, L-lactate, or acetate bound in the active site. These structures are good models of the Michaelis complex between the oxidized enzyme and the substrate, and they show that proline is completely buried in the active site, implying that protein conformational changes must accompany substrate binding and product release. The structure of the dithionite-reduced enzyme did not provide insight into such conformational changes because a hyposulfite ion, an oxidation product of dithionite, occupies the proline site (7). Because the structure of the empty active site has proven to
be elusive, the nature of the conformational changes attendant to substrate binding and product release are currently unknown.

Our analysis of the available structures indicates four regions of interest with regard to conformational changes associated with flavin reduction and substrate binding: the FAD cofactor, residues near the FAD N(5) atom (Arg431, Asp370), α8, and the β1-α1 loop.

Involvement of the FAD ribityl chain in functional switching is consistent with previous biochemical and structural studies. Membrane-binding studies with mutant PutA proteins and PutA reconstituted with 2-deoxy-FAD showed that the 2'-OH group acts as a redox-sensitive switch that controls membrane association (7). The structures of the PPG-inactivated and dithionite-reduced enzymes indicate that reduction of the flavin induces a 90° rotation in this group. Additionally, the former structure reveals a more substantial reorganization of the ribityl chain involving all three hydroxyl groups and a larger distortion of the isoalloxazine ring (butterfly bend). Although it seems clear that the bending of the isoalloxazine ring and rotation of the 2'-OH group are associated with functional switching, whether the conformational changes of the 3'-OH and 4'-OH groups are consequences of the reduced flavin state or empty proline binding site remain to be determined.

It should be noted that the experimental details of crystal preparation might explain some of the differences between the flavin conformations observed in the structures of the PPG-inactivated and dithionite-reduced enzymes. The crystal used to solve the former structure was grown from inactivated enzyme, and thus conformational changes induced by inactivation were allowed to proceed in solution. In contrast, the latter structure was solved from an oxidized crystal that was soaked in dithionite prior to freeze-trapping in liquid nitrogen. In this case, the preformed crystal lattice likely
prohibited extensive conformational changes. Furthermore, the use of high dithionite concentration resulted in a hyposulfite ion binding in the active site.

The structure of PPG-inactivated PutA86-630 also implicates Arg431 in functional switching. In the THFA complex, Arg431 is poised above the hydride acceptor of the FAD and is thus positioned to sense the redox state of the flavin. The PPG-inactivated enzyme structure implies that reduction induces rupture of the N(5)-Arg431 hydrogen bond. This result is satisfying because it provides an explanation for previously reported results showing that replacement of Arg431 with Met or reconstitution of PutA with 5-deaza-FAD eliminated reductive activation of PutA-membrane binding (7). It was suggested at that time that Arg431 plays an important role in transmitting redox signals out of the PRODH active site that lead to activation of membrane binding and transcription of the *put* regulon, although the structure of the dithionite-reduced enzyme did not indicate any change in this residue. The structure reported here confirms the importance of Arg431 in functional switching and suggests that rupture of the N(5)-Arg431 hydrogen bond might be among the early events that occur in transmitting the membrane association signal from the flavin to the membrane-binding domain.

The PPG-inactivated enzyme structure suggests that Asp370 might also have a role in functional switching. The structure shows that flavin reduction weakens the interaction between Asp370 and Arg431, allowing the former residue to sample two side chain conformational states that are related by a rotation of 180° around $\chi_1$. This result suggests the new hypothesis that transmission of the functional switching signal out of the active site involves correlated movement of Asp370 and Arg431.

Finally, the structure of the PPG-inactivated enzyme is unique among PutA PRODH domain structures in that helix 8 is substantially unfolded, the conserved Arg555-Glu289
ion pair is broken, and the β1-αl loop is withdrawn from the active site. Consequently, the active site is partially disassembled, and the si face of the isoalloxazine ring, i.e., the proline binding site, is exposed to solvent. We suggest that the disassembled active site reflects the absence of a bound proline analog, rather than the reduced state of the flavin, because α8 and the β1-αl loop are critical for substrate binding. In particular, Tyr552, Arg555, and Arg556 of α8 directly contact the substrate, Glu559 of α8 stabilizes Arg556 through ion pairing, and Glu289 of the β1-αl loop ion pairs with Arg555 (Figure 2.9). Furthermore, the open active site of the PPG-inactivated enzyme bears a striking resemblance to that of oxidized TtPRODH, the only other PRODH structure with an empty proline binding site (Figure 2.10).

We thus suggest that the collective structural information on the PutA PRODH domain and TtPRODH provides clues about conformational changes attendant to substrate binding. The structures imply that the active site might be incompletely assembled in the absence of the substrate, and the binding of proline stabilizes residues of α8 and draws the β1-αl loop into the active site in order to connect the loop with α8 via the conserved ion pair. In this scenario, the ion pair might function as a gate that closes and opens in response to substrate binding and product release, respectively. Further experiments are needed to test these ideas. In this regard, structures of the ligand-free, oxidized PutA PRODH domain and a monofunctional PRODH complexed with THFA would be especially enlightening.
ACKNOWLEDGMENTS

We thank Dr. Jay Nix of Advanced Light Source beamline 4.2.2 for help with data collection. 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.
REFERENCES


Table 2.1: Data collection and refinement statistics for PutA86-630 inactivated with NPPGa

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>space group</td>
<td>$I2_{1}2_{1}2_{1}$</td>
</tr>
<tr>
<td>unit cell lengths (Å)</td>
<td>$a = 127.5, b = 133.4, c = 133.6$</td>
</tr>
<tr>
<td>wavelength (Å)</td>
<td>1.0000</td>
</tr>
<tr>
<td>diffraction resolution (Å)</td>
<td>31.5 – 2.15 (2.25 - 2.15)</td>
</tr>
<tr>
<td>no. of observations</td>
<td>243987</td>
</tr>
<tr>
<td>no. of unique reflections</td>
<td>60418</td>
</tr>
<tr>
<td>redundancy</td>
<td>4.0 (3.3)</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>97.9 (98.9)</td>
</tr>
<tr>
<td>$R_{\text{merge}}(I)$</td>
<td>0.094 (0.489)</td>
</tr>
<tr>
<td>$R_{\text{pim}}(I)$</td>
<td>0.053 (0.303)</td>
</tr>
<tr>
<td>average $I/\sigma$</td>
<td>10.3 (3.3)</td>
</tr>
<tr>
<td>Wilson B-factor (Å$^2$)</td>
<td>26</td>
</tr>
<tr>
<td>no. of protein chains</td>
<td>2</td>
</tr>
<tr>
<td>no. of atoms</td>
<td>6480</td>
</tr>
<tr>
<td>no. of protein residues</td>
<td>800</td>
</tr>
<tr>
<td>no. of FAD atoms</td>
<td>106</td>
</tr>
<tr>
<td>no. of water molecules</td>
<td>241</td>
</tr>
<tr>
<td>$R_{\text{cryst}}$</td>
<td>0.193 (0.250)</td>
</tr>
<tr>
<td>$R_{\text{free}}$</td>
<td>0.234 (0.292)</td>
</tr>
<tr>
<td>rmsd$^c$</td>
<td></td>
</tr>
<tr>
<td>bond lengths (Å)</td>
<td>0.006</td>
</tr>
<tr>
<td>bond angles (deg.)</td>
<td>0.98</td>
</tr>
<tr>
<td>Ramachandran plot$^d$</td>
<td></td>
</tr>
<tr>
<td>Favored (%)</td>
<td>98.1</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>1.9</td>
</tr>
<tr>
<td>Outliers (%)</td>
<td>0.0</td>
</tr>
<tr>
<td>average B-factors (Å$^2$)</td>
<td></td>
</tr>
<tr>
<td>protein</td>
<td>34</td>
</tr>
<tr>
<td>FAD</td>
<td>26</td>
</tr>
<tr>
<td>modified Lys329</td>
<td>27</td>
</tr>
<tr>
<td>water</td>
<td>31</td>
</tr>
<tr>
<td>coordinate error (Å)$^e$</td>
<td>0.52</td>
</tr>
<tr>
<td>PDB accession code</td>
<td>3ITG</td>
</tr>
</tbody>
</table>

$^a$Values for the outer resolution shell of data are given in parenthesis.
$^b$5% random test set.
$^c$Compared to the Engh and Huber parameters (30).
$^d$The Ramachandran plot was generated with RAMPAGE (31).
$^e$Maximum likelihood-based coordinate error reported by PHENIX.
Scheme 2.1

Scheme 2.2
Scheme 2.3

Scheme 2.3
Figure legends

Figure 2.1. Spectroscopic analysis of the inactivation of full-length *E. coli* PutA by PPG. (A) Flavin spectral changes of *E. coli* PutA induced by inactivation by PPG. The spectrum of oxidized PutA is represented as the thick curve and has maxima at $\lambda = 451$ nm and 378 nm. Spectra were acquired immediately after adding 500 $\mu$M PPG and were continually collected for up to 90 min as described in Experimental Procedures. For clarity, only 14 spectra are shown, with the dashed curve indicating the last spectrum. Note that the maximum at $\lambda = 451$ nm disappears (decreases) and a new maximum centered at $\lambda = 378$ nm appears (increases) as time advances. The inset shows the absorbance recorded at 451 nm (solid circles) and 378 nm (open circles) during the 90-min inactivation process. (B) Denaturation of PPG-inactivated PutA. Spectra are: PutA (7.35 $\mu$M) before inactivation (solid curve), PutA inactivated with 500 $\mu$M PPG for 90 min (dashed curved), and the flavin species isolated after treatment of PPG-inactivated PutA with 0.2 % SDS (dotted curve).

Figure 2.2. Kinetics of PutA inactivation by PPG at 25 °C using a dichlorophenolindophenol-based activity assay. The percent activity remaining after incubation with PPG is plotted as a function of time for four inactivator concentrations (0.25 - 2.5 mM). The inset shows the replot of the half-life of inactivation as a function of the reciprocal concentration of PPG. The inactivation parameters obtained from fitting are $k_{\text{inact}} = 0.13 \pm 0.05$ min$^{-1}$ and $K_1 = 1.5 \pm 0.2$ mM.
Figure 2.3. The two views of PPG-inactivated PutA86-630. Strands and helices are colored pink and blue, respectively. The dashed curves denote disordered segments of the polypeptide chain. The covalent FAD-Lys329 adduct is colored green. Red coloring denotes regions displaying large conformational differences from other PutA PRODH domain structures (i.e., α8 and the β1-α1 loop).

Figure 2.4. Electron density maps showing the (A) overall FAD conformation and covalent attachment to Lys329 and (B) curvature of the isoalloxazine ring. The cage represents a simulated annealing $\sigma_A$-weighted $F_o - F_c$ map contoured at 3σ. Prior to calculation of the map, the FAD and side chain of Lys329 were omitted, and simulated annealing refinement was performed with PHENIX.

Figure 2.5. Conformational changes during PutA inactivation by PPG detected by limited proteolysis. Purified PutA (1 mg/mL) was incubated for 5 min each with proline (5 mM), THFA (5 mM), or for 30 min with PPG (5 mM), in 50 mM potassium phosphate (pH 7.5) followed by digestion with chymotrypsin (10 µg/mL) for 1 h at 23 °C. The reactions were quenched with phenylmethylsulfonyl fluoride and hot SDS sample buffer after 1 h. After complete denaturation in SDS buffer, 7.5 µg of each digestion was loaded onto an 8 % polyacrylamide denaturing gel for analysis. As a control, undigested ligand-free PutA was also loaded (2.5 µg). The gel was stained with Coomassie Blue to visualize the major products. A molecular mass marker (M) was loaded in the first lane.

Figure 2.6. Physical binding of PPG-inactivated PutA to E. coli polar lipids. (A) PutA (0.25 mg/ml) was pre-incubated with HEPES-N buffer plus 5 mM proline, 5 mM THFA
or 5 mM PPG for 5 min, 5 min or 30 min respectively, and then incubated with freshly prepared *E. coli* polar lipids (0.8 mg/ml) for 1 hr at room temperature. The soluble and lipid fractions were then separated by Air-fuge ultracentrifugation, denatured with SDS buffer, and analyzed with SDS-PAGE. As a control, an equal amount of denatured PutA without treatment was also loaded. (B) Quantitative analysis of PutA bands for the soluble (gray bars) and lipid (grey hatched bars) fractions normalized as a percentage of total PutA using a Bio-Rad 2000 system.

**Figure 2.7.** The FAD conformations observed in three crystal structures of the *E. coli* PutA PRODH domain: (A) PPG-inactivated PutA86-630, (B) oxidized PutA86-669 complexed with THFA (cyan), and (C) dithionite-reduced PutA86-669. Helix α8 is shown in red for the latter two enzymes; this helix is disordered in PPG-inactivated PutA86-630. Note that the reduced flavins are colored green (panels A, C), whereas the oxidized flavin is colored yellow (panel A). The three flavins are oriented such that the AMP moieties are superimposed.

**Figure 2.8.** Close-up view of structural differences between PPG-inactivated PutA86-630 (green) and PutA86-669 complexed with THFA (white), highlighting the movements of Arg431 and Asp370. The occupancy values for the two conformations of Asp370 in the inactivated enzyme are given in parentheses. Orange dashed lines represent hydrogen bonds in the PPG-inactivated enzyme. Black dashed lines represent hydrogen bonds of the THFA complex. The cage represents a simulated annealing $\sigma_A$-weighted $F_o - F_c$ map contoured at 3σ. Prior to calculation of the map, the side chains of Asp370 and Arg431 were omitted, and simulated annealing refinement was performed with PHENIX.
Figure 2.9. Stereographic view of a superposition of the active sites of PPG-inactivated PutA86-630 (green) and PutA86-669 complexed with THFA (white). The THFA ligand is colored cyan. The thin dashed lines indicate hydrogen bonds of the THFA complex. The thick dashed curve denotes the disordered section of α8 of the PPG-inactivated enzyme. For clarity, conformation B of Asp370 of the inactivated enzyme has been omitted.

Figure 2.10. Stereographic view of a comparison of PPG-inactivated PutA86-630 (green), PutA86-669 complexed with THFA (white), and TtPRODH (orange). The THFA ligand is colored cyan. The thin dashed lines indicate the Arg555-Glu289 ion pair observed in the THFA complex. Note that Glu289 is disordered in the PPG-inactivated enzyme (modeled as Ala), and the corresponding residue of TtPRODH (Glu65) points out of the active site and into the solvent. The thick dashed curve denotes the disordered section of α8 of the PPG-inactivated enzyme.
Figure 2.1: Spectroscopic analysis of the inactivation of full-length *E. coli* PutA by PPG.
Figure 2.2: Kinetics of PutA inactivation by PPG at 25 °C using a dichlorophenolindophenol-based activity assay.
Figure 2.3: The two views of PPG-inactivated PutA86-630.
Figure 2.4: Electron density maps showing the (A) overall FAD conformation and covalent attachment to Lys329 and (B) curvature of the isoalloxazine ring.
Figure 2.4: Conformational changes during PutA inactivation by PPG detected by limited proteolysis.
Figure 2.5: Physical binding of PPG-inactivated PutA to *E. coli* polar lipids.

A

<table>
<thead>
<tr>
<th></th>
<th>Ctrl</th>
<th>OX</th>
<th>PRO</th>
<th>THFA</th>
<th>PPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>S</td>
<td>L</td>
<td>S</td>
<td>L</td>
<td>S</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th></th>
<th>Partitioning (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl</td>
<td>100</td>
</tr>
<tr>
<td>OX</td>
<td>L</td>
</tr>
<tr>
<td>PRO</td>
<td>S</td>
</tr>
<tr>
<td>THFA</td>
<td>L</td>
</tr>
<tr>
<td>PPG</td>
<td>S</td>
</tr>
</tbody>
</table>
Figure 2.6: The FAD conformations observed in three crystal structures of the *E. coli* PutA PRODH domain.
Figure 2.7: Close-up view of structural differences between PPG-inactivated PutA86-630 and PutA86-669 complexed with THFA
Figure 2.8: Stereographic view of a superposition of the active sites of PPG-inactivated PutA86-630 and PutA86-669 complexed with THFA.
Figure 2.9: Stereographic view of a comparison of PPG-inactivated PutA86-630, PutA86-669 complexed with THFA, and TiPRODH.
Structure of the Proline Utilization A Proline Dehydrogenase Domain Inactivated by N-propargylglycine Provides Insight into Conformational Changes Induced by Substrate Binding and Flavin Reduction

Dhiraj Srivastava, Weidong Zhu, William H. Johnson, Jr., Christian P. Whitman, Donald F. Becker, and John J. Tanner
Chapter 3 - Structural and Kinetic Evidence for Substrate Channeling in the Bifunctional Proline Utilization A Flavoenzyme

Dhiraj Srivastava, Jonathan P. Schuermann, Tommi A. White, Navasona Krishnan, Greg L. Hura, Anmin Tan, Michael T. Henzel, Donald F. Becker and John J. Tanner

aDepartment of Chemistry, University of Missouri-Columbia, Columbia, MO, 65211, USA
bDepartment of Biochemistry, University of Missouri-Columbia, Columbia, MO, 65211, USA
cDepartment of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
dAdvanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

1Corresponding author: John J. Tanner, Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA. Phone: 573-884-1280. Fax: 573-882-2754. E-mail: tannerjj@missouri.edu.

Manuscript information: 4.5 text pages (incl. references and figure legends), 5 figures, 1 table.

Abbreviations: PRODH, proline dehydrogenase; P5C, Δ^{1}-pyrroline-5-carboxylate; P5CDH, Δ^{1}-pyrroline-5-carboxylate dehydrogenase; PutA, proline utilization A; BjPutA, Bradyrhizobium japonicum proline utilization A; EcPutA, Escherichia coli proline utilization A; POX, human proline oxidase/dehydrogenase; PDB, Protein Data Bank.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3HAZ).
Abstract

The bifunctional proline catabolic flavoenzyme, proline utilization A (PutA), catalyzes the oxidation of proline to glutamate via the sequential activities of FAD-dependent proline dehydrogenase (PRODH) and NAD⁺-dependent Δ¹-pyrroline-5-carboxylate dehydrogenase (P5CDH) domains. Although structures for some of the domains of PutA are known, a structure for the full-length protein, heretofore, has not been solved. Here we report the 2.1 Å resolution crystal structure of PutA from *Bradyrhizobium japonicum*, along with data from small-angle X-ray scattering, analytical ultracentrifugation, and steady-state and rapid-reaction kinetics. PutA forms a ring-shaped tetramer in solution having a diameter of 150 Å. Within each protomer, the PRODH and P5CDH active sites face each other at a distance of 41 Å and are connected by a large, irregularly shaped cavity. Kinetics measurements show that glutamate production occurs without a lag phase, indicating that the intermediate, Δ¹-pyrroline-5-carboxylate, is preferably transferred to the P5CDH domain rather than released into the bulk medium. The structural and kinetic data imply that the cavity serves both as a microscopic vessel for the hydrolysis of Δ¹-pyrroline-5-carboxylate to glutamate semialdehyde and a protected conduit for the transport of glutamate semialdehyde to the P5CDH active site.

Introduction:

Oxidation of amino acids is a central part of energy metabolism. In eukaryotes and bacteria, proline is oxidized (or catabolized) to glutamate by the conserved enzymes, proline dehydrogenase (PRODH) and Δ¹-pyrroline-5-carboxylate (P5C) dehydrogenase (Fig. 3.1). PRODH catalyzes the FAD-dependent oxidation of proline (1 in Fig. 3.1) to P5C (2). The intermediate of proline catabolism, P5C, is hydrolyzed nonenzymatically to glutamate γ-semialdehyde (3), which is subsequently oxidized to glutamate (4) by NAD⁺-dependent P5C dehydrogenase (P5CDH). In humans, proline oxidation is important for cellular redox control, apoptosis, and suppression of cancer (1). Inborn errors in PRODH and P5CDH result in types I and II hyperprolinemia disorders (2), and deficiency in PRODH activity is thought to be a risk factor for schizophrenia (3). Moreover, human PRODH (known as POX) is a p53-induced enzyme that plays a crucial role in apoptosis by serving as a proline-dependent, superoxide generator (1). Other organisms have evolved the ability to harness energy almost exclusively from proline under certain circumstances. Examples include *Helicobacter pylori* in the human gut (4, 5), procyclic trypanosomes in the midgut of insects (6, 7), and the tsetse fly during flight (8, 9).

Intriguingly, whereas PRODH and P5CDH are separate enzymes in eukaryotes and some bacteria, they are fused into a single polypeptide chain in other bacteria. The fused enzymes, known as proline utilization A, or PutA, are membrane-associated proteins containing 1000-1300 amino acid residues, with the PRODH and P5CDH domains located in the N- and C-terminal halves of the polypeptide chain, respectively (10, 11). Some PutAs, such those from *Escherichia coli* (EcPutA) and *Salmonella typhimurium*, also have a regulatory function. These "trifunctional" PutAs repress transcription of the *put* regulon, which contains the genes for PutA and the proline transporter PutP. The
DNA-binding domain of trifunctional PutAs is a ribbon-helix-helix domain located in the N-terminal 50 residues (12-14).

PutAs have proven to be challenging structural biology targets. The first structures of PutA domains appeared over two decades after Ratzkin and Roth's discovery that PRODH and P5CDH were encoded by a single gene in *S. typhimurium* (15). These include crystal structures of the PRODH (16-19) and DNA-binding (13, 14) domains of EcPutA, and a solution structure of the DNA-binding domain of *Pseudomonas putida* PutA (20).

Whereas the domain structures have informed us about the individual functions of PutA, they have not provided insight into how PutAs coordinate multiple functions. Here we report the first crystal structure of a full-length bifunctional PutA. The structure and accompanying kinetic data indicate that P5C undergoes hydrolysis to glutamate semialdehyde *en route* to the P5CDH domain without leaving the confines of the protein.

**Results**

**Tertiary Structure.** Several PutAs were screened for crystallizability, which eventually led us to focus structure determination efforts on the 999-residue PutA from *Bradyrhizobium japonicum* (BjPutA). The structure of BjPutA was solved using anomalous dispersion diffraction data collected from hexagonal crystals of the Se-Met derivative (Table S1) and was subsequently refined to 2.1 Å resolution using native diffraction data collected from a centered monoclinic crystal (Table 3.1).

The protomer consists of the two catalytic domains plus three smaller, ancillary domains denoted N-terminal arm, α-helix bundle, and linker (Fig. 3.2A). The asymmetric unit of the C2 crystal form contains two protomers, which form a U-shaped
dimer having dimensions of 150 Å x 70 Å x 80 Å (Fig. 3.2B). The polypeptide chain begins with a 51-residue arm-like extension that traverses ~100 Å as it wraps around the PRODH domain (Fig. 3.2A, yellow). The arm is followed by the α-helix bundle domain, which comprises six tightly packed α-helices (Fig. 3.2A, green) and contacts both catalytic domains. The PRODH and P5CDH domains follow the α-helix bundle and are separated by an 83-residue linker.

The PRODH domain resembles that of EcPutA in both the overall fold and details of the active site. This domain has a distorted (βα)₈ barrel fold (Fig. 3.2, magenta strands, cyan helices), which is considered to be diagnostic of the PRODH family (21). The FAD cofactor is bound noncovalently at the C-termini of the strands (Figs. 3.2A, 3.2B). The re face packs tightly against strands 4 - 6 of the barrel, allowing proline to bind at the si face (Fig. S1). Residues near the isoalloxazine and proline-binding site are highly conserved in PutAs, and many of these residues have similar conformations in BjPutA and EcPutA (Fig. S2A). BjPutA and EcPutA also have identical FAD conformations (Fig. S2A). Finally, there appears to be a sulfate ion bound in the active site. Comparison to the structure of the EcPutA PRODH domain complexed with the proline analog L-tetrahydro-2-furoic acid indicates that the sulfate ion occupies the binding site for the substrate carboxylate group (Fig. S2A).

The P5CDH domain (Fig. 3.2A, red) is similar to other aldehyde dehydrogenases (22, 23). It has two lobes, one that binds NAD⁺ using a non-classical Rossmann fold and another that furnishes catalytic Cys792. In addition to the two lobes, there is a β-strand substructure protruding from the Rossmann fold sub-domain (β-flap, residues 628 - 646, 977 - 989, see Fig. 3.2A). The β-flap stabilizes the U-shaped dimer by forming an intermolecular β-sheet interaction with the P5CDH domain of the other protomer (Fig.
As with the PRODH active site, there is a sulfate ion bound in the P5CDH active site. Comparison with the structure of the monofunctional P5CDH from *Thermus thermophilus* shows that the sulfate ion occupies the binding site for the carboxylate group of the substrate glutamate semialdehyde (Fig. S2B).

**Quaternary Structure.** The oligomeric state of BjPutA in solution was examined by analytical ultracentrifugation. Data from eighteen parallel sedimentation equilibrium experiments were fit globally to a single-species model, enabling simultaneous estimation of the partial specific volume and molecular weight (Fig. S3). The best-fit values are 0.772 mL/g (0.767 - 0.777 mL/g) and 442 kDa (428 - 455 kDa). The partial specific volume is significantly larger than the value of 0.74 mL/g calculated from the sequence, which is consistent with the observation of a large solvent-filled cavity inside the protein (vide infra). The estimated molecular weight is within 3 % of the theoretical molecular weight of a tetramer of 430 kDa.

Small-angle X-ray scattering (SAXS) was used to determine the arrangement of protomers within the tetramer. A typical SAXS profile is shown in Fig. 3.3A (black solid curve). Guinier analysis suggested a radius of gyration of 56 Å. Calculations of the pair distribution function yielded estimates for the radius of gyration and the maximum particle dimension of 51 Å and 140 Å, respectively. Shape reconstructions calculated with GASBOR (24) indicated that BjPutA forms a ring-shaped particle in solution (Fig. 3.3A, inset).

The $C_2$ lattice was inspected to identify a tetramer of BjPutA having the size and shape of the SAXS reconstruction. Such an assembly could be built by applying the two-fold crystallographic symmetry transformation ($-x, y, -z, -1$) to the U-shaped dimer of the
asymmetric unit (Figs. 3.3B, S4). Further validation of this particular tetramer as the one formed in solution was obtained from calculations of the scattering curve from atomic models using CRY SOL (25). The curve calculated from the ring-shaped tetramer displays satisfactory agreement with the experimental one (Fig. 3.3A, compare red and black solid curves). In contrast, the curves calculated from the U-shaped dimer (Fig. 3.3A, dashes), and two other tetrameric assemblies that do not form rings (Fig. 3.3A, dots), show poor agreement with the experimental data. Finally, we note that the tetrameric ring is also present in lattices of the hexagonal and orthorhombic crystal forms of BjPutA.

The tetramer is a dimer of dimers having 222 symmetry, which can be described by three mutually perpendicular two-fold axes, denoted P, Q, and R, intersecting at the centroid of the particle (Figs. 3.3B, S4). We adopt the convention of labeling the four chains O, P, Q, and R, such that the P-, Q-, and R-axes transform the O chain into the P, Q, and R chains, respectively.

There are two types of interface in the tetramer, which bury a total of 8360 Å² of surface area. The larger one is generated by rotation around the P-axis (Fig. 3.2B). This interface features the swapped β-flap interaction shown in Fig. 3.2C and is a well-known feature of aldehyde dehydrogenases (26). The other interface of the tetramer is unique to BjPutA and contributes 1760 Å² of interfacial area. It is generated by rotation around the Q-axis and joins two U-shaped dimers to form a ring such that all four β-flaps face the hole of the ring (Fig. 3.3B). The last helix of the N-terminal arm, the α-helical bundle domain, and part of the linker domain form this interface.

**Internal Cavity Implies Substrate Channeling.** Within each protomer, the
PRODH and P5CDH active sites face each other at a distance of 41 Å. Between them is an irregularly shaped cavity that connects the si face of the isoalloxazine to the catalytic cysteine of the P5CDH domain (Figs. 3.2C, 3.4). The cavity volume was estimated with VOIDOO (27) to be 1400 Å³, which is equivalent to about 170 water molecules. All domains except the arm contribute to the cavity (Fig. 3.2C). The β-flap evidently serves as a lid that covers the cavity of the P-related protomer (Figs. 3.2C, 3.2D); analysis with the program MOLE (28) indicates potential openings (3 - 5 Å wide) to the bulk medium in this region (Fig. 3.2D). Because the cavity connects the active sites of a protomer, the structure implies the potential for substrate channeling within, rather than between, protomers.

The cavity has three sections, corresponding to the two active sites and a large, intervening chamber (Fig. 3.4). The middle chamber is 24 Å by 14 Å in its two largest dimensions, and 3-7 Å in the third dimension. It has a volume of about 1325 Å³, which represents over 90 % of the total volume of the entire cavity system. Thus, the middle chamber is large enough to accommodate P5C (molecular volume 102 Å³) and glutamate semialdehyde (120 Å³). The active sites connect to the middle chamber via constrictions that have widths approximately corresponding to the diameter of water molecule.

The connection between the PRODH active site and the middle chamber is particularly interesting because of its circuitous path (Fig. 3.4). It runs down the ribityl chain of the FAD and then passes over a potential opening to the bulk medium and under the side chain of Arg456 before connecting to the middle chamber. A more direct path into the middle chamber appears to be blocked by the ion-pairing residues Arg456 and Glu197. We note that both residues are identically conserved in PutAs, and Arg456 plays an essential role in substrate binding by forming a critical ionic interaction with the
carboxylate of the substrate.

Finally, the walls of the cavity are highly ionic, and thus the cavity itself is very hydrophilic. The cavity walls are lined with twelve arginine, three lysine, ten glutamate, and seven aspartate residues, which are represented by the blue (Arg, Lys) and red (Gly, Asp) patches in Fig. 3.4.

**Kinetic Evidence for Substrate Channeling.** Kinetic measurements were performed to explore the possibility of substrate channeling. The results of a P5C trapping assay are shown in Fig. 3.5A. The trapping experiment uses an excess of $o$-aminobenzaldehyde ($o$-AB), which forms a yellow complex with P5C (29). In the absence of NAD$^+$, formation of the $o$-AB-P5C complex is evident, indicating that P5C is released into bulk solvent. In contrast, with NAD$^+$ present (0.2 mM), formation of the $o$-AB-P5C complex is significantly diminished, clearly showing that P5C is not released to an appreciable extent from the functional enzyme.

The time-dependence of NADH production from proline was measured to see whether a lag phase was present. NADH is generated by wild type BjPutA at a linear rate without an apparent lag phase (Fig. 3.5B, solid curve), which is consistent with substrate channeling. This assay was repeated using an equimolar mixture of BjPutA monofunctional variants R456M and C792A as a non-channeling control. PRODH activity is disabled in R456M, but the P5CDH activity is identical to that of the native enzyme (Table S2). Conversely, the C792A mutation abolishes P5CDH activity but does not affect PRODH activity. If channeling occurs within the protomer, as suggested by the structure, an equimolar mixture of R456M and C792A should not exhibit the kinetic hallmarks of channeling. As shown in Fig. 3.5B (dashed curve), a lag phase of approximately 8 min was observed for the mixture of monofunctional variants. The lag in NADH production is
interpreted as the time required for a steady-state level of P5C to accumulate in solution (30). The substantial difference between the native enzyme and the non-channeling control in this assay is consistent with the hypothesis that substrate channeling occurs in BjPutA.

Rapid-reaction kinetic measurements were also performed. BjPutA, proline, and NAD$^+$ were rapidly mixed, and the absorbance spectra of the FAD cofactor and NADH product were recorded using a stopped-flow instrument. For the non-channeling control, reduction of FAD is observed, but no change in absorbance at 340 nm occurs, i.e., NADH is not produced (Fig. 3.5C). In contrast, for the native enzyme, NADH is generated following FAD reduction (Fig. 3.5D). The fact that NADH is produced in a time during which no NADH is generated by the non-channeling control suggests that channeling occurs.

**Discussion**

The BjPutA structure reveals, for the first time, the architecture of a PutA protein. The most significant result is that the two active sites are separated by 41 Å and oriented so that the exit path of the PRODH active site faces the substrate entrance tunnel of the P5CDH active site. The cavity connecting the two active sites strongly suggests the possibility of substrate channeling within the protomer, which is supported by kinetic measurements.

Although substrate channeling is not unique to PutA, some aspects of channeling in PutA are different from other bifunctional enzymes. In particular, the irregular shape of the cavity is rather different from the tunnels that channel indole and ammonia in tryptophan synthase and carbamoyl phosphate synthetase (31, 32). The PutA cavity better resembles that of dimethylglycine oxidase, which is also a flavoenzyme that channels an imine intermediate (33, 34).

The demonstration of substrate channeling in BjPutA is consistent with kinetic data
reported by Surber and Maloy a decade ago for the trifunctional PutA from *S. typhimurium* (35), the only other study of channeling in PutAs. As with BjPutA, initial production of NADH from proline occurred without a lag phase for *S. typhimurium* PutA. Surber and Maloy also used \(^{14}\text{C}\)-labeled proline to show that exogenous P5C does not efficiently compete for P5CDH activity in the presence of proline, clear evidence of channeling. The fact that channeling has been demonstrated in both bifunctional and trifunctional PutAs suggests that it may be a conserved aspect of PutAs.

Significantly, this study implies that the P5C produced by PutA is hydrolyzed to glutamate semialdehyde within the confines of the protein rather than in the bulk medium. It is tempting to suggest that the hydrolysis reaction occurs in the middle chamber, since it is large enough to accommodate glutamate semialdehyde. The middle chamber could also facilitate this reaction by shifting the equilibrium between P5C and glutamate semialdehyde to favor the latter species. The equilibrium has a strong pH-dependence, with the reverse reaction (toward P5C) favored above pH 6.5 (36, 37). This transition point corresponds to the \(pK_a\) of the pyrrolinium moiety of P5C, indicating that protonation of the pyrrolinium favors hydrolysis. The highly ionic character of the walls of the middle chamber could create an electrostatic field inside the cavity that effectively increases the \(pK_a\) of the pyrrolinium moiety, thus favoring the formation of glutamate semialdehyde. Future studies involving mutagenesis of charged residues lining the cavity will test this hypothesis.

The structure of the cavity suggests that protein dynamics might play a role in channeling. For example, the conserved Arg456-Glu197 ion pair blocks a direct route from the active site to the middle chamber (Fig. 3.4). If P5C formation and concomitant reduction of the FAD induce rupture of the ion pair, the PRODH cavity would coalesce
with the middle chamber, opening the direct route. Thus, the ion pair could potentially act as a gate that opens and closes in response to product release and substrate binding, respectively. This idea is supported by the observation that the conserved ion pair is formed in the structure of the EcPutA PRODH domain complexed with THFA (Fig. S2A) and broken in the structure of ligand-free monofunctional T. thermophilus PRODH (38). We also note that there are several examples of gates and constrictions blocking channeling pathways in other bifunctional enzymes (39-42).

The discovery that BjPutA forms a tetramer was unexpected. In the only other biophysical study of the oligomeric state of a PutA, Brown and Wood concluded that EcPutA is an elongated dimer in solution (43), which contrasts the donut shape of BjPutA. An important difference between the two proteins is that EcPutA functions as a transcriptional repressor of the put regulon, in addition to being a bifunctional enzyme. We previously showed that the repressor domain has a ribbon-helix-helix fold (12-14). It thus appears that bifunctional and trifunctional PutAs differ in quaternary structure, which was unanticipated. This observation raises the question of how the oligomeric states of PutAs are related to function. The relationship seems clear for trifunctional PutAs; ribbon-helix-helix domains must dimerize to bind DNA. In particular, an intermolecular β-sheet of the dimer inserts into the major groove of DNA so that side chains of the sheet make specific contacts with DNA bases. The significance the U-shaped dimer of BjPutA is also clear. The β-flap of one protomer covers the substrate channel of the other protomer of the dimer. On the other hand, the role of the tetramer for bifunctional PutAs is, as yet, unknown.

Finally, our work raises the question of the biological relevance of substrate channeling in proline catabolism. Channeling potentially affords the advantages of
decreased transit time between enzymes, protection of labile intermediates, and isolation of intermediates from competing enzymatic reactions, among others (30-32). Channeling in PutA may be particularly beneficial for bacteria that have adapted to use proline as a fuel source. One example is *H. pylori*, which utilizes proline as a preferred energy source in the gut environment, where proline levels in the gastric juice of infected patients are 10-fold higher compared to uninfected individuals (5, 44). In this case, channeling likely enables efficient use of an available fuel. Also, channeling segregates the P5C produced by proline catabolism from the proline biosynthetic pathway, which also uses P5C as an intermediate. In absence of channeling, a futile cycle of interconversion between proline and P5C could occur. Channeling might also be relevant in humans, because P5C is involved in a number of diverse biological functions, including apoptosis, oxidative stress, and RNA translation initiation (45-48). Although human PRODH (POX) and P5CDH are distinct (i.e., not fused) mitochondrial enzymes, the fact that the two enzymes are fused in PutA, along with the clear demonstration of channeling in PutA, suggests that the human enzymes may interact and engage in substrate channeling in order to protect P5C from competing biological pathways.

**Materials and Methods**

**Protein Preparation, Crystallization, and Structure Determination.**

Preparation of native BjPutA and the growth of three crystal forms have been described (49). See SI Text for detailed information regarding the preparation and crystallization of selenomethionyl BjPutA, phasing by anomalous dispersion using the hexagonal crystal form (Table S1), and molecular replacement calculations for the orthorhombic and $C_2$ crystal forms. The structure was refined to 2.1 Å resolution in space group $C_2$ (Table 3.1). The final model includes one BjPutA dimer consisting of 1957 of the expected
1998 residues, two FAD cofactors, and two NAD\(^+\) cofactors. Chain O includes residues 1-51, 54-122, and 127-989. Chain P includes residues 2-50, 55-122, 128-536, and 542-989. The missing residues appear in the chain termini, apparently flexible loops (50s loop, 120s loop), and the linker between the PRODH and P5CDH domains (chain P).

**Equilibrium Analytical Ultracentrifugation.** These experiments were performed using a Beckman XL-I Optima analytical ultracentrifuge equipped with an An50Ti rotor. Initial studies indicated that the partial specific volume was larger than the value of 0.74 mL/g predicted from the amino acid sequence. Therefore, the differential sedimentation equilibrium method of Edelstein and Schachman was used to simultaneously determine the partial specific volume and the molecular weight (50). Details of these experiments are provided in SI Text.

**Small-Angle X-ray scattering.** Small-angle X-ray scattering (SAXS) experiments were performed at beamline 12.3.1 of the Advanced Light Source. Details of these studies can be found in SI Text.

**Kinetic Measurements.** Steady-state and rapid-reaction methods were used to assess substrate channeling, as described in detail in SI Text.

**Acknowledgements**
We thank Dr. Jay Nix of ALS beamline 4.2.2 and Dr. Stephan L. Ginell of the APS SBC beamlines for assistance with data collection and processing. This research was supported by NIH grants GM065546, GM061068 and P20 RR-017675. 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. Results shown in this report are derived, in part, from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.
**Footnotes**

Author contributions: J.J.T., D.F.B., G.L.H., and M.T.H designed research; D.S., J.P.S., T.A.W., N.K., and A.T. performed research; all authors analyzed data; and J.J.T., D.F.B., and M.T.H wrote the paper.
References


Figure Legends

Fig. 3.1. The reactions catalyzed by PRODH and P5CDH.

Fig. 3.2. Structure of BjPutA. (A) Ribbon drawing of the protomer. The domains are color coded as follows: yellow, N-terminal arm; green α-helix bundle domain; cyan/magenta, PRODH domain; slate, linker; and red, P5CDH domain. FAD and NAD$^+$ are drawn as sticks in yellow and green, respectively. The β-flap is colored blue. The silver surface represents the internal cavity that connects the two active sites. (B) The dimer of the asymmetric unit, colored as in panel A. FAD and NAD$^+$ are drawn as spheres. (C) Close-up view of the dimer showing the region near the β-flap. The coloring scheme of panel A is used. (D) Surface representation of the β-flap region, shown with the same orientation and coloring scheme as in panel C.

Fig. 3.3. SAXS analysis and tetrameric structure of BjPutA. (A) Experimental and simulated scattering profiles. The solid black curve represents a typical experimental scattering curve for BjPutA. The solid red curve represents the profile calculated from the ring-shaped tetramer shown in panel B. The dashed magenta curve was calculated from the dimer of the asymmetric unit. The two dotted curves were calculated from other tetrameric assemblies that are not shaped like a ring. The inset shows an ab initio shape restoration calculated from the experimental SAXS data using GASBOR. (B) Superposition of the SAXS reconstruction and the crystallographic tetramer. The four chains are labeled O (red), P (blue), Q (green), and R (orange). FAD and NAD$^+$ are shown in yellow and gray spheres, respectively.
Fig. 3.4. Two views of the cavity, differing in a rotation of ~180° around the vertical axis. The cavity is represented as a semi-transparent surface and is colored to indicate the locations of positively (blue) and negatively (red) charged residues lining the cavity. The tubes represent pathways identified by the program MOLE. The green tube guides the eye from the PRODH active site to the P5CDH active site. The orange tubes represent possible pathways leading to the bulk medium.

Fig. 3.5. Kinetic evidence for substrate channeling in BjPutA. (A) P5C trapping experiments performed in the presence and absence of NAD⁺. (B) NADH formation by native BjPutA (solid curve) and an equimolar mixture of monofunctional variants R456M and C792A (dashed curve). The dotted curve shows the prediction of a free diffusion model for the consecutive PRODH and P5CDH activities in BjPutA. (C) Rapid reaction kinetics with the monofunctional BjPutA variants using proline as the substrate. The experiment was performed anaerobically in the absence of an external electron acceptor for the reduced flavin. An equimolar mixture of BjPutA variants C792A and R456M (15 µM each), NAD⁺ (0.1 mM), and proline (50 mM) were rapidly mixed and monitored by stopped-flow multiwavelength absorption. Curves 1-6: 0.014, 0.042, 0.07, 0.098, 0.18, and 3.3 s. Inset: The observed rate of FAD reduction at 450 nm (filled circles) was estimated to be 4.6 ± 0.3 s⁻¹ by best-fit of the experimental data. No NADH formation was observed at 340 nm (solid line). (D) Rapid reaction kinetics of native BjPutA using proline as the substrate. The experiment was performed anaerobically in the absence of an external electron acceptor for the reduced flavin. PutA (30 µM), NAD⁺ (0.1 mM), and proline (50 mM) were rapidly mixed and monitored by stopped-flow multiwavelength
absorption. Curves 1-8: 0.013, 0.026, 0.039, 0.052, 0.065, 0.078, 0.18, and 3.3 s. Inset: Plot of FAD reduction (circles) and NADH formation (triangles) monitored at 450 nm and 340 nm, respectively. The estimated rate constants for reduction of FAD and NAD$^+$ are $3.9 \pm 0.5$ s$^{-1}$ and $0.7 \pm 0.03$ s$^{-1}$, respectively.

**Table Legends**

**Table 3.1.** Data Collection and Refinement Statistics
Figure 3.1: The reactions catalyzed by PRODH and P5CDH.

1. proline
2. Δ1-pyrroline-5-carboxylate
3. glutamate semialdehyde
4. glutamate

PRODH catalyzes the reaction from proline to Δ1-pyrroline-5-carboxylate. P5CDH catalyzes the reaction from glutamate semialdehyde to glutamate.
Figure 3.2: Structure of BjPutA
Figure 3.3: SAXS analysis and tetrameric structure of BjPutA.
Figure 3.4: Two views of the cavity, differing in a rotation of ~180° around the vertical axis.
Figure 3.5: Kinetic evidence for substrate channeling in BjPutA.
Table 3.1: Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>C2</td>
</tr>
<tr>
<td>$a$, $b$, $c$, Å</td>
<td>166.8, 195.8, 108.7</td>
</tr>
<tr>
<td>$\beta$, °</td>
<td>121.5</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>43 – 2.1 (2.2 – 2.1)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>172549</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.8 (3.7)</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>99.9 (99.9)</td>
</tr>
<tr>
<td>$R_{\text{merge}}$</td>
<td>0.103 (0.412)</td>
</tr>
<tr>
<td>$I/\sigma$</td>
<td>12.1 (2.7)</td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>14647</td>
</tr>
<tr>
<td>No. of FAD and NAD$^+$ atoms</td>
<td>194</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>795</td>
</tr>
<tr>
<td>$R_{\text{cryst}}$</td>
<td>0.200 (0.225)</td>
</tr>
<tr>
<td>$R_{\text{free}}$†</td>
<td>0.233 (0.281)</td>
</tr>
<tr>
<td>rmsd bond lengths, Å</td>
<td>0.005</td>
</tr>
<tr>
<td>rmsd bond angles, °</td>
<td>0.86</td>
</tr>
<tr>
<td>Ramachandran – favored</td>
<td>1901</td>
</tr>
<tr>
<td>Ramachandran – allowed</td>
<td>39</td>
</tr>
<tr>
<td>Ramachandran – outliers</td>
<td>3</td>
</tr>
<tr>
<td>Average B for protein, Å$^2$</td>
<td>35</td>
</tr>
<tr>
<td>Average B for FAD, Å$^2$</td>
<td>24</td>
</tr>
<tr>
<td>Average B for NAD$^+$, Å$^2$</td>
<td>32</td>
</tr>
<tr>
<td>Average B for water, Å$^2$</td>
<td>34</td>
</tr>
</tbody>
</table>

32. Values for the outer resolution shell of data are given in parentheses.
†5% random test set.
Supporting Information

Crystal Structure of the Bifunctional Proline Utilization A Flavoenzyme from Bradyrhizobium japonicum

Dhiraj Srivastava\textsuperscript{a}, Jonathan P. Schuermann\textsuperscript{a}, Tommi A. White\textsuperscript{b}, Navasona Krishnan\textsuperscript{c}, Nikhilesh Sanyal\textsuperscript{c}, Greg L. Hura\textsuperscript{d}, Anmin Tan\textsuperscript{b}, Michael T. Henzl\textsuperscript{b}, Donald F. Becker\textsuperscript{c} and John J. Tanner\textsuperscript{a,b,1}

\textsuperscript{a}Department of Chemistry, University of Missouri-Columbia, Columbia, MO, 65211, USA
\textsuperscript{b}Department of Biochemistry, University of Missouri-Columbia, Columbia, MO, 65211, USA
\textsuperscript{c}Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
\textsuperscript{d}Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

\textsuperscript{1}Corresponding author: John J. Tanner, Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA. Phone: 573-884-1280. Fax: 573-882-2754. E-mail: tannerjj@missouri.edu.
SI Text

**Protein Expression, Purification, and Crystallization.** PutA from *Bradyrhizobium japonicum* (BjPutA) was expressed and purified as described previously (1). The selenomethionyl (Se-Met) derivative was expressed using the metabolic inhibition method (2). The derivative was purified like the native enzyme except that all buffers used after the immobilized metal ion affinity chromatography step contained 5 mM DTT.

Three crystal forms of BjPutA were used in this work: hexagonal, orthorhombic, and centered monoclinic (1). The hexagonal form was grown in PEG and has space group *P*622 with one molecule in the asymmetric unit. This crystal form was used for experimental phasing using the Se-Met derivative. The orthorhombic form was also grown in PEG. The space group is *P*212121 with four molecules in the asymmetric unit. This crystal form exhibits apparent tetragonal symmetry due to pseudo-merohedral twinning. The orthorhombic crystals diffracted weakly to 2.8 Å resolution (1). This crystal form was useful for verifying the quaternary structure of BjPutA. The C2 form was grown in 2 M ammonium sulfate, has one dimer in the asymmetric unit, and diffracted to 2.1 Å resolution. This crystal form was used for the final refinement.

**X-ray Diffraction Data Collection.** The hexagonal crystal form was used for experimental phasing. Data from Se-Met hexagonal crystals were collected at beamline 19ID of the Advanced Photon Source and processed with HKL2000 (3). Two crystals were used for phasing with the methods of single wavelength anomalous diffraction (SAD, crystal 1) and multiwavelength anomalous dispersion (MAD, crystal 2). For crystal 1, a data set was collected at the energy corresponding to the experimentally-
determined peak of $f''$ ($\lambda = 0.97945\text{Å}$). The data set consisted of 480 frames collected with a detector distance of 240 mm and an oscillation width of 0.5°. The SAD data set was processed to 2.75 Å resolution, was over 20-fold redundant and displayed high completeness (Table S1). For crystal 2, three data sets were collected at energies corresponding to the peak of $f''$, ascending inflection point, and high energy remote. Each consisted of 200 frames collected with a detector distance of 240 mm and an oscillation width of 0.5°. The data sets collected from crystal 2 have high-resolution limits of 2.9 – 3.0 Å, greater than 10-fold redundancy, and over 90 % completeness (Table S1).

The C2 form was used for high resolution refinement. A 2.1 Å resolution data set was collected at beamline 4.2.2 of the Advanced Light Source. The data set consisted of 720 frames collected with detector distance of 130 mm and oscillation width of 0.25°. The reflections were integrated with MOSFLM version 6.2.6 (4) and merged with SCALA version 3.2.25 (5). Data processing statistics are listed in Table 1 of the main article.

**Experimental Phasing.** Se sites were identified using SOLVE (6). Sixteen out of the eighteen expected sites were found using the SAD data set, and fifteen sites were found using the MAD data (Table S1). The resulting SAD and MAD phase sets had figures of merit of 0.40 and 0.47, respectively. Density modification in RESOLVE (6) improved both figures of merit to 0.7. RESOLVE was also used for automated model building. Using either the SAD or MAD phases after density modification, RESOLVE provided a nearly complete backbone trace for the C-terminal half of the protein (P5CDH domain). Density in the N-terminal half of the protein, however, was relatively weak,
and the RESOLVE-generated model for the N-terminal half consisted of only seven \(\alpha\)-helices. In total, RESOLVE correctly identified the amino acid sequence for 75 residues using SAD phases and 67 residues using MAD phases. A composite model was constructed using the RESOLVE models from SAD and MAD phasing, and this model served as the starting point for several rounds of model building in COOT (7) and refinement against the Se-Met peak data sets in CNS (8), PHENIX (9), and REFMAC5 (10). Eventually, a partial model containing 726 residues was obtained, which was used as the search model to solve the \(C_2\) and \(P2_12_12_1\) crystal forms with molecular replacement. This model included most of the linker and P5CDH domains, but only the last helix of the arm domain, none of the \(\alpha\)-helix bundle domain, and just over half of the PRODH domain.

**Molecular Replacement and Refinement.** These calculations were performed with MOLREP (11) using a 726-residue model from experimental phasing. For the \(C_2\) form, a clear solution having two molecules in the asymmetric unit was obtained, which had a correlation coefficient of 0.32 and \(R_{\text{cryst}}\) of 0.52. Rigid-body refinement lowered \(R_{\text{cryst}}\) to 0.45 for reflections to 2.1 Å resolution. Simulated annealing refinement in PHENIX further reduced \(R_{\text{cryst}}\) to 0.37 (\(R_{\text{free}} = 0.40\)) with rmsd bonds of 0.006. The resulting electron density map clearly showed the regions of the structure that were problematic in the hexagonal form, such as the FAD cofactor (Fig. S1).

Using the \(P2_12_12_1\) data, a clear molecular replacement solution having four molecules in the asymmetric unit was obtained, which had a correlation coefficient of 0.33 and \(R_{\text{cryst}}\) of 0.47 for reflections to 3.5 Å resolution. The four chains of the asymmetric unit are arranged in the ring-shaped assembly shown Fig. S4, providing additional validation of
this tetramer as the one formed in solution.

**Refinement and Final Model.** The C2 data set was used for refinement to 2.1 Å resolution. The model from molecular replacement was extended and improved with several rounds of model building in COOT and simulated annealing refinement in PHENIX. The final model includes one BjPutA dimer consisting of 1957 of the expected 1998 residues, two FAD cofactors, and two NAD\(^+\) cofactors (Table 1 of main article). Chain O includes residues 1-51, 54-122, and 127-989. Chain P has residues 2-50, 55-122, 128-536, and 542-989. The missing residues appear in the chain termini, apparently flexible loops (50s loop, 120s loop) and the linker between the PRODH and P5CDH domains (chain P). Electron density maps suggested partial occupancy for NAD\(^+\), which was estimated to be 0.55. The solvent model includes 795 water molecules, 13 sulfate ions, and six glycerol molecules. The sulfate ions presumably derive from the crystallization solution, which contained 2 M ammonium sulfate. Glycerol was present in the protein stock and cryoprotection buffers at 5 % and 25 %, respectively.

**Structure analysis.** Several programs and servers were used to analyze structures. Protein cavities were identified with VOIDOO (12). FLOOD (12) was used to estimate the number of water molecules accommodated by the cavities and to aid visualization of cavities. Tunnels connecting cavities with bulk solvent were identified with MOLE (13). Interfaces were analyzed with the PISA protein interface server (14). Other analysis programs used include CNS (8), COOT (7), and PyMOL (15).

**Analytical Ultracentrifugation.** The quaternary structure of BjPutA in solution
was analyzed using equilibrium analytical ultracentrifugation. Data were acquired at 20
°C using a Beckman XL-I Optima analytical ultracentrifuge equipped with an An50Ti
rotor. Initial studies indicated that the partial specific volume (\( \bar{\nu} \)) was larger than the
value of 0.742 mL/g predicted from the amino acid sequence. Therefore, the differential
sedimentation equilibrium method of Edelstein and Schachman was used to
simultaneously determine \( \bar{\nu} \) and the molecular weight, \( M \) (16). In this method, two
parallel experiments are performed using protein dialyzed into H\(_2\)O and D\(_2\)O solutions.
The data are fit to an appropriate model that takes into account the increased density of
the latter solution and the increase in protein mass due to exchange of protons for
deuterons. In the present case, satisfactory agreement was obtained using the following
single-species model:

\[
c(r) = c(r_o) \exp \left[ \frac{kM\omega^2(1 - \bar{\nu}\rho/k)}{2RT} \left( r^2 - r_o^2 \right) \right] + BL
\]  

(1)

In this equation, \( c(r) \) is the total protein concentration at radial position \( r \); \( c(r_o) \) is the
protein concentration at an arbitrary reference position, \( r_o \); \( M \) is the molecular weight; \( \omega \)
is the radial velocity (rad/s); \( \bar{\nu} \) is the partial specific volume of the protein; \( \rho \) is the
solvent density; \( R \) is the gas constant; \( T \) is the absolute temperature; \( BL \) is a baseline
offset; and \( k \) is the ratio of the molecular weight of the protein in the deuterated buffer to
that in the nondeuterated buffer. By definition, \( k \) is 1.0 for the nondeuterated sample.
For the deuterated sample (99.8 % D\(_2\)O), \( k \) was estimated to be 1.01616, based on
analysis of the amino acid sequence using CNS.

Prior to centrifugation, two identical protein samples having concentrations of 4.3
mg/mL were dialyzed into nondeuterated and deuterated buffers containing 50 mM Tris-HCl, 50 mM NaCl, 0.5 mM EDTA, 5 % glycerol, 0.1 mM FAD, pH 7.5. For each sample, data were collected at three protein concentrations and three rotor speeds (4000 rpm, 6000 rpm, 8000 rpm) using Rayleigh interference optics. Measured solvent densities were 1.020533 g/mL and 1.120807 g/mL for the two buffers. The eighteen sets of data were fit globally to equation 1 using Origin 8. Analysis of $\chi^2$ using the $F$-statistic, in which $M$ was determined at fixed values of $\nu$, was used to estimate 68 % confidence intervals for $M$ and $\nu$.

**Small-angle X-ray scattering.** Small angle X-ray scattering (SAXS) experiments were performed at beamline 12.3.1 of the Advanced Light Source. Prior to analysis, samples of BjPutA (6 mg/mL) were dialyzed against 50 mM Tris-HCl, 50 mM NaCl, 0.5 mM EDTA, 5 % glycerol, pH 7.5. Scattering intensities ($I$) were measured at three protein concentrations (6 mg/mL, 3 mg/mL, 1.5 mg/mL) as a function of the scattering vector $Q = 4\pi \sin \theta / \lambda$, where $2\theta$ is the scattering angle and $\lambda$ is the wavelength of the incident beam. Exposure times of 0.8 s and 8 s were used. The scattering curves collected from the protein samples were corrected for background scattering using intensity data collected from the dialysis buffer. Forward scattering intensities ($I(0)$) were calculated from Guinier plots (ln $I$ versus $Q^2$) (17) normalized for concentration (mass/volume). The radius of gyration ($R_g$) was determined by Guinier analysis from scattering curves extrapolated to zero concentration using the program PRIMUS (18). Composite scattering curves were generated with PRIMUS by scaling and merging background-corrected high $Q$ region data from higher concentration samples with low $Q$ region zero-extrapolated data. Scattering curves were subjected to indirect Fourier
transform using the GNOM program (19) to yield the pair distribution function ($P©$), from which $R_g$ and the maximum particle dimension ($D_{max}$) were estimated. GNOM output files were used to calculate ab initio shape reconstructions with DAMMIN (20) and GASBOR (21). At least ten independent reconstructions were performed for each data set. The two programs yielded similar results. Individual models were aligned by SUPCOMB (22) and averaged by DAMAVER (23) to yield the most probable particle shape. Theoretical scattering curves were calculated from atomic coordinates using CRYSOL (24).

**Kinetic Measurements.** Steady-state and rapid-reaction methods were used to assess substrate channeling. To facilitate interpretation of the kinetic data, the assays were also performed with a non-channeling control consisting of a solution containing equal concentrations of the two BjPutA monofunctional mutant enzymes C792A and R456M. Cys792 is the predicted catalytic Cys residue in the P5CDH domain. Arg456 of the PRODH active site is predicted to form a critical ion pair with the carboxylate group of the substrate proline (Fig. S2A). Thus, C792A and R456M are predicted to be monofunctional enzymes, with PRODH activity disabled in R456M and P5CDH activity disabled in C792A. These predictions were verified by measuring the PRODH and P5CDH activities of the individual mutant enzymes using steady-state assays as described for PutA proteins (25, 26). Kinetic parameters for the native and mutant enzymes are listed in Table S2. Measurements were performed in 50 mM potassium phosphate buffer (pH 7.5) at 23 °C. PRODH activity was determined by proline:ubiquinone oxidoreductase assays using CoQ₁ (0.1 mM), 0 – 200 mM proline and by following quinone reduction at 278 nm ($\varepsilon$=14.5 mM⁻¹ cm⁻¹). P5CDH activity was
monitored by following NAD$^+$ reduction at 340 nm as previously described (26). The effective concentrations of P5C and glutamate semialdehyde at pH 7.5 were estimated from the data of Bearne and Wolfenden on the pH dependence of the equilibrium between the ring and open chain forms of P5C (27). The C792A mutant has no P5CDH activity, whereas the PRODH activity is unaffected. Conversely, the R456M mutant has no detectable PRODH activity but exhibits P5CDH activity similar to that of native BjPutA. Thus, R456M and C792A are monofunctional variants, as expected.

The appearance of the intermediate P5C in the bulk medium was studied using a trapping assay based on o-aminobenzaldehyde (o-AB). P5C forms the yellow dihydroquinazolinium complex with o-AB, which is detected by absorbance at $\lambda = 443$ nm ($\varepsilon = 2900 \text{ M}^{-1}\text{cm}^{-1}$ (28)). For this assay, BjPutA (0.19 µM) was incubated with proline (40 mM), o-AB (4.0 mM), and inverted membrane vesicles from E. coli (putA$^-$ strain) in the absence and presence of 0.2 mM NAD$^+$. The assay buffer contained 50 mM phosphate buffer, 10 mM MgCl$_2$, and 25 mM NaCl at pH 7.5.

Assays designed to detect a lag phase in the progress curve for the production of NADH from proline were also performed. Observation of a lag phase in the progress curve would be inconsistent with channeling. The formation of NADH was monitored at 340 nm in assays performed at 23 °C containing BjPutA (20 µg), proline (40 mM), ubiquinone (0.1 mM), and NAD$^+$ (0.2 mM) in a buffer containing 50 mM potassium phosphate, 25 mM NaCl, and 10 mM MgCl$_2$ at pH 7.5. Identical assays were performed using 20 µg each of the monofunctional BjPutA variants C792A and R456M. Spectral interference caused by changes in quinone absorbance at 340 nm was corrected for by subtracting progress curves from assays performed under identical conditions in the absence of NAD$^+$. 

93
The progress curve for two non-interacting PRODH and P5CDH enzymes was simulated as described previously for a diffusion model using equation 2 (29).

\[ [\text{NADH}] = v_1 t + \left( \frac{v_1}{v_2} \right) K_m (e^{-v_2 t / K_m} - 1) \] (2)

In this equation, \( v_1 \) is the rate \((1.6 \ \mu\text{M min}^{-1})\) of PRODH activity under the reaction conditions. The parameters \( v_2 \) \((21 \ \mu\text{M min}^{-1})\) and \( K_m \) \((0.18 \text{ mM})\) are the steady-state kinetic parameters \( V_{\text{max}} \) and the Michaelis constant, respectively, for P5CDH activity of the R456M mutant.

The reaction of proline with BjPutA was also studied by rapid-reaction kinetic measurements under single turnover conditions. BjPutA \((28.52 \ \mu\text{M}, \text{in 50 mM K-phosphate, 50 mM NaCl, pH 7.5})\) was preincubated with 0.2 mM NAD\(^+\) and then rapidly mixed with an equal volume of 80 mM proline in 50 mM potassium phosphate buffer \((\text{pH 7.5})\) with 50 mM NaCl. For the non-channeling control reaction, an equimolar mixture of R456M \((59.6 \ \mu\text{M})\) and C792A \((59.6 \ \mu\text{M})\) \((\text{both mutants in 50 mM K-phosphate, 50 mM NaCl, pH 7.5})\) was preincubated with 0.2 mM NAD\(^+\) and then rapidly mixed with an equal volume of 80 mM proline in 50 mM potassium phosphate buffer \((\text{pH 7.5})\) with 50 mM NaCl. The final concentrations of the reagents after mixing were 40 mM proline, 50 mM K-phosphate \((\text{pH 7.5})\), 50 mM NaCl, and 0.1 mM NAD\(^+\). The enzyme concentrations after mixing were 14.26 \(\mu\text{M}\) (native PutA) and 29.8 \(\mu\text{M}\) each of the monofunctional mutants C792A and R456M. Thus, the concentration of functional active sites in the experiment using the native enzyme \((14.26 \mu\text{M PRODH and 14.26 \muM P5CDH})\) was similar to that of the non-channeling control experiment \((14.9 \mu\text{M PRODH and 14.9 \muM P5CDH actives sites from C792A and R456M, respectively})\).
The reduction reactions of FAD and NAD$^+$ were monitored by stopped-flow multiwavelength absorption in the range of 300-700 nm with a Hi-Tech Scientific SF-61DX2 stopped-flow instrument equipped with a photodiode array detector and KinetAsyst software. Experiments were performed anaerobically in the absence of an external electron acceptor for the reduced flavin as follows. Substrate and BjPutA solutions were degassed by repeated cycles of vacuum and flushing with oxygen-scrubbed nitrogen. Protocatechuate dioxygenase (PCD, 0.05 U/ml) and protocatechuic acid (PCA, 100 µM) were then added to the solutions to further remove oxygen and maintain anaerobic conditions. The solutions were transferred to anaerobic syringes under a nitrogen atmosphere in a glovebox (30). The stopped-flow mixing chambers were thoroughly washed with anaerobic buffer and incubated overnight with buffer containing PCD/PCA prior to the experiments. The temperature of the mixing chamber was maintained at 25 °C.

The observed rate constants ($k_i$) for FAD reduction (0.2 s$^{-1}$) and NAD$^+$ reduction (0.025 s$^{-1}$) with native BjPutA were obtained by fitting single exponential equations to the absorbance decrease at 451 nm ($y = Ae^{-kt} + C$) and increase at 341 nm ($y = A(1-e^{-kt})$) using non-linear least squares regression and SigmaPlot software (Figs. S7A, S7B). The concentration of oxidized FAD was calculated at each time point using $\varepsilon_{\text{ox}, 451 \text{ nm}} = 13600$ M$^{-1}$ cm$^{-1}$ and $\varepsilon_{\text{red}, 451 \text{ nm}} = 100$ M$^{-1}$ cm$^{-1}$ for the oxidized and reduced FAD species, respectively. NADH was quantitated using a molar extinction coefficient of 6220 M$^{-1}$ cm$^{-1}$ at 341 nm, which is also the isosbestic point for spectral changes that accompany FAD reduction. We note that the observed pseudo first order rate constant of 0.2 s$^{-1}$ for FAD reduction and the calculated second order rate constant of 1753 M$^{-1}$s$^{-1}$ for NADH formation ($k_{\text{obs}}/[\text{P5C}] = 0.025$ s$^{-1}$/14.26 µM from one turnover of the PRODH active site)
are consistent with the apparent turnover rate of NADH formation in steady-state channeling assays with native BjPutA under similar conditions (0.12 s\(^{-1}\) and 600 M\(^{-1}\)s\(^{-1}\) \((V/K_m(P5CDH))\)) (Fig. 3.5B).

The observed rate constants \((k_j)\) for FAD reduction (0.144 s\(^{-1}\), 90% amplitude) and NAD\(^+\) reduction (0.0163 s\(^{-1}\)) with the equimolar mixture of the two C792A and R456M variants were obtained by fitting double exponential equations to the absorbance changes at 451 nm \((y = Ae^{k_1t} + Be^{-k_2t} + C)\) and 341 nm \((y = A(1-e^{-k_1t}) + e^{k_2(t_0-t)} + C)\) (Figs. S7C, S7D). At 451 nm the second rate constant \(k_2\) represents only 10% of the FAD reduction, whereas at 341 nm the second exponential accounts for the observed lag time (~10 sec) prior to NADH formation.
References


Figure Legends

**Fig. S1.** A view of the FAD cofactor, highlighting its location relative to the strands of the (βα)₈ barrel. The map is a simulated annealing σₐ-weighted $F_o - F_c$ omit map contoured at 3.0 σ. Prior to map calculation, the flavin was omitted, and simulated annealing refinement was performed using PHENIX.

**Fig. S2.** Stereo views of the two active sites of BjPutA. (A) Superposition of the PRODH active sites of BjPutA (green) and EcPutA (white). The structure of the EcPutA active site was determined from a PRODH domain construct containing residues 86-630 of EcPutA (PutA86-630, PDB code 1tiw, (31)). The sulfate ion of BjPutA is colored cyan. The proline analog L-tetrahydro-2-furoic acid bound to PutA86-630 is colored magenta. The FAD cofactors of both enzymes are colored yellow. The dashed lines represent electrostatic interactions in BjPutA (black) and PutA86-630 (orange). (B) Superposition of the BjPutA P5CDH active site (green) and the active site of the monofunctional P5CDH from *T. thermophilus* (white, PDB code 2bhq,(32)). The sulfate ion of BjPutA is colored cyan. The bound glutamate of *T. thermophilus* is colored magenta. The dashed lines represent electrostatic interactions in BjPutA (black) and the monofunctional P5CDH (orange). Two conformations are observed for Cys792 of BjPutA, which is due to the partial occupancy of NAD⁺.

**Fig. S3.** Analytical ultracentrifugation data for BjPutA in non-deuterated (open symbols) and deuterated (filled symbols) solvent. The top panel shows a global fit of sedimentation equilibrium data to an ideal single-species model (Equation 1 of SI Text). Data were acquired at 4000 (squares), 6000 (circles), and 8000 (triangles) rpm at three
nominal loading concentrations of 4.3, 2.2, and 1.1 mg/mL. The solid lines represent the best least-squares fit to an ideal single-species model. Residuals for the optimal fit are displayed in the bottom panel. For clarity, vertical offsets have been applied, and only a subset of the data points is displayed.

**Fig. S4.** Orthogonal views of the BjPutA tetramer. (*A*) Ribbon drawing of the tetramer, viewed down the *R* molecular axis. The four chains are labeled O (red), P (blue), Q (green), and R (orange). FAD and NAD\(^+\) are shown in yellow and gray spheres, respectively. (*B-D*) Superposition of the SAXS reconstruction and the crystallographic tetramer, as viewed down the three molecular two-fold axes. The color coding is the same as in panel *A*.

**Fig. S5.** Stereographic view of the cavity. The orientation of this figure differs from that of Fig. 3.4 by a rotation of ~180° around the vertical axis. The cavity is represented as a semi-transparent surface and is colored to indicate the locations of positively (blue) and negatively (red) charged residues lining the cavity. The tubes represent pathways identified by the program MOLE. The green tube guides the eye from the PRODH active site (top) to the P5CDH active site (bottom). The orange tubes represent possible pathways leading to the bulk medium.

**Fig. S6.** Rapid-reaction kinetic data for native BjPutA (black, left vertical axes) and the mixed variants non-channeling control (red, right vertical axes). (*A*) Progress curves showing the time range of 0 – 250 sec. Note the lag in NADH production for the non-channeling control. (*B*) Expansions of the initial regions of the progress curves from
panel A. For the mixed variants \([\text{FAD}_{\text{ox}}]\) curve, the concentration of \([\text{FAD}_{\text{ox}}]\) remaining at the end of the experiment (16.5 uM) was subtracted from each data point in order to facilitate the plotting of \([\text{FAD}_{\text{ox}}]\) and \([\text{NADH}]\) on the same scale.

**Fig. S7.** Residual analysis of stopped-flow kinetic data representing (A) reduction of FAD in the native enzyme, (B) production of NADH by the native enzyme, (C) reduction of FAD in the mixture of monofunctional variants, and (D) production of NADH by the mixture of monofunctional variants. In each panel, the upper plot shows the data points (circles) and the best least-squares fit to an exponential model (solid curve), while the lower panel shows the residuals for the optimal fit.
Figure S1: A view of the FAD cofactor, highlighting its location relative to the strands of the \((\beta\alpha)_8\) barrel.
Figure S2: Stereo views of the two active sites of BjPutA.
Figure S3: Analytical ultracentrifugation data for BjPutA in non-deuterated and deuterated solvent.
Figure S4: Orthogonal views of the BjPutA tetramer.
Figure S5: Stereographic view of the cavity.
Figure S6: Rapid-reaction kinetic data for native BjPutA and the mixed variants non-channeling control.
Figure S7: Residual analysis of stopped-flow kinetic data.
<table>
<thead>
<tr>
<th></th>
<th>Crystal 1</th>
<th>Crystal 2</th>
<th>Crystal 2</th>
<th>Crystal 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beamline</strong></td>
<td>APS 19ID</td>
<td>APS 19ID</td>
<td>APS 19ID</td>
<td>APS 19ID</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P6(_2)22</td>
<td>P6(_2)22</td>
<td>P6(_2)22</td>
<td>P6(_2)22</td>
</tr>
<tr>
<td><strong>Unit cell parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Å, degrees)</td>
<td>a = 142.5</td>
<td>a = 142.1</td>
<td>a = 142.1</td>
<td>a = 142.1</td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>0.97945</td>
<td>0.97945</td>
<td>0.97962</td>
<td>0.95372</td>
</tr>
<tr>
<td><strong>Diffraction resolution (Å)</strong></td>
<td>50 – 2.75</td>
<td>50 – 2.9</td>
<td>50 – 2.9</td>
<td>50 – 3.0</td>
</tr>
<tr>
<td></td>
<td>(2.85 – 2.75)</td>
<td>(3.0 – 2.9)</td>
<td>(3.0 – 2.9)</td>
<td>(3.1 – 3.0)</td>
</tr>
<tr>
<td><strong>No. of observations</strong></td>
<td>779946</td>
<td>293686</td>
<td>268424</td>
<td>235460</td>
</tr>
<tr>
<td><strong>No. of unique reflections</strong></td>
<td>29719</td>
<td>23556</td>
<td>23406</td>
<td>21158</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>26.2 (23.5)</td>
<td>12.5 (12.4)</td>
<td>11.5 (10.9)</td>
<td>11.1 (10.3)</td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>100 (100)</td>
<td>93.6 (95.4)</td>
<td>92.8 (91.6)</td>
<td>92.6 (90.8)</td>
</tr>
<tr>
<td><strong>R(_{merge})</strong></td>
<td>0.094 (0.450)</td>
<td>0.091 (0.482)</td>
<td>0.079 (0.478)</td>
<td>0.092 (0.450)</td>
</tr>
<tr>
<td><strong>Average I/σ</strong></td>
<td>44.9 (8.6)</td>
<td>24.1 (4.4)</td>
<td>25.3 (4.1)</td>
<td>21.5 (4.0)</td>
</tr>
<tr>
<td><strong>No. Se sites</strong></td>
<td>16</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Figure of merit before DM</strong></td>
<td>0.40</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Figure of merit after DM</strong></td>
<td>0.70</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Values for the outer resolution shell of data are given in parentheses.
†The figure of merit was calculated using SOLVE before density modification.
‡The figure of merit was calculated using RESOLVE after density modification.
Table S2: Kinetic constants for native BjPutA and monofunctional BjPutA variants

<table>
<thead>
<tr>
<th></th>
<th>PRODH Activity</th>
<th>P5CDH Activity</th>
<th>Overall Pro/Glu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (mM)</td>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>$k_{cat}/K_m$ (M$^{-1}$ s$^{-1}$)</td>
</tr>
<tr>
<td>Native</td>
<td>31 ± 6</td>
<td>2.0 ± 0.1</td>
<td>65</td>
</tr>
<tr>
<td>R456M BD*</td>
<td>BD</td>
<td>BD</td>
<td>BD</td>
</tr>
<tr>
<td>C792A</td>
<td>31 ± 5</td>
<td>1.7 ± 0.1</td>
<td>55</td>
</tr>
</tbody>
</table>

*Below detection limit of assay
Chapter 4 – Structure and kinetics of human P5C dehydrogenase and its mutants

Introduction – P5C dehydrogenase, also called γ-glutamyl semialdehyde dehydrogenase or ALDH4A1, is an aldehyde dehydrogenase, which catalyzes the oxidation of γ-glutamyl semialdehyde (Pyrroline-5-carboxylate, P5C) to glutamate with concomitant reduction of NAD$^+$ to NADH. P5C is itself a product of oxidation of proline, a process catalyzed by proline dehydrogenase.

![Chemical structures and reactions]

There have been a total of nineteen aldehyde dehydrogenase genes that have been identified in humans. Most of the aldehyde dehydrogenases are either dimeric or tetrameric (Table 4.1). An exception is ALDH4 (P5C dehydrogenase) which can be dimeric (Human and Mice P5C dehydrogenase) or hexameric (Thermus thermophilus P5C dehydrogenase and Yeast Put2).

The oxidation of proline to glutamate is catalyzed by two enzymes, proline dehydrogenase (ProDH) and pyrroline-5-carboxylate dehydrogenase (P5CDH). P5C dehydrogenase is an aldehyde dehydrogenase and catalyzes the NAD$^+$ dependent oxidation of P5C to glutamate. In addition to the degradation of proline, P5C dehydrogenase is also involved in the degradation of ornithine. Ornithine aminotransferase catalyzes the interconversion of ornithine and γ-glutamyl semialdehyde.

The aldehyde dehydrogenase superfamily is classified in 11 separate families and P5C
dehydrogenase is a member of the aldehyde dehydrogenase 4 family and is subsequently also termed as ALDH4A1.

In eukaryotes, proline dehydrogenase is associated with the inner mitochondrial membrane while P5C dehydrogenase is found in the mitochondrial matrix. P5C dehydrogenase is a non-specific enzyme, which catalyzes the oxidation of widely differing substrates ranging from propionaldehyde to adipic semialdehyde. However, P5C dehydrogenase attains maximum velocity when P5C is used as the substrate. P5C and γ-glutamic semialdehyde is in equilibrium and equilibrium is towards P5C at pH 7.0. Considering the effective concentration of γ-glutamyl semialdehyde at pH 7.0, it is about 8000 times better substrate than acetaldehyde and propionaldehyde, about 30 times better than succinic semialdehyde, and about 8-10 times better than glutaric and adipic semialdehyde. P5C dehydrogenase follows an ordered bi-bi mechanism where NAD⁺ initially binds to the enzyme followed by γ-glutamyl semialdehyde. After the oxidation of γ-glutamyl semialdehyde, glutamate releases from the active site followed by the release of NADH.

The Km for NAD⁺ is 374 µM and for γ-glutamyl semialdehyde is 68 µM. Higher eukaryotes evolve to have hydroxyproline present in their collagen. As a result, these organisms have an additional enzyme for the oxidation of hydroxy-proline to Δ¹-pyrroline-3-hydroxy-5-carboxylate (3-OH-P5C), known as hydroxy-proline dehydrogenase. Based on genetic evidence there is only one P5C dehydrogenase that
oxidizes both P5C and 3-OH-P5C\(^1\). However there is no biochemical evidence supporting this hypothesis.

Deficiency in P5C dehydrogenase enzyme lead to a disease condition called type II hyperprolinemia\(^5\). Type II hyperprolinemia is a rare autosomal recessive disorder of proline catabolism characterized by 10-15 fold increase in plasma proline and increased P5C level. A clinical symptom of type II hyperprolinemia is childhood seizures\(^6\). Several mutations in human P5C dehydrogenase gene has been reported in literature. There is a frame shift mutation in codon 7, A7fs(-1), which leads to premature termination of translation and absence of P5C dehydrogenase activity. Effect of mutations P16L, S352L, and G521fs(+1) on the activity of P5C dehydrogenase has been studied by functional complementation assay in yeast\(^7\). P5C dehydrogenase deficient strain of yeast, Saccharomyces cerevisiae, expressing S352L and G521fs(+1) mutant of human P5C dehydrogenase was failed to grow on medium with proline as sole nitrogen source and did not have detectable P5C dehydrogenase activity. However P16L mutant produced fully functional P5C dehydrogenase enzyme. Recently five other mutations in P5C dehydrogenase gene have been reported\(^8\) that might have an influence on P5C (proline) metabolism, though their actual involvement in disease is not clear. These newly identified mutations are E277K, G366R, F388L, V470I, and T473A.

The structure of Thermus thermophilus P5C dehydrogenase has been solved by Tahirov et. al. (2005)\(^9\). However, the P5C dehydrogenase from Thermus thermophilus is only 31% similar to human P5C dehydrogenase and is hexameric. On the other hand human P5C dehydrogenase is dimeric.

In this work, I have solved the structure of human P5C dehydrogenase in two different crystal forms, mouse P5C dehydrogenase and Saccharomyces cerevisiae Put2.
Structure of human P5C dehydrogenase in hexagonal crystal form, solved at 2.5 Å resolution, represents the substrate bound state while structure in monoclinic crystal form, which was solved at 1.95 Å, represents the ligand free state of human P5C dehydrogenase. The structure of mouse P5C dehydrogenase, which is 95% similar to human P5C dehydrogenase, was solved at 1.3 Å resolution and serve as the high-resolution model of human P5C dehydrogenase. Structure of Put2, solved at 1.95 Å resolution, together with *Thermus thermophilus* P5C dehydrogenase helps us in understanding the oligomerization of P5C dehydrogenase. We have also done the global analysis of mammalian aldehyde dehydrogenase to find out why some aldehyde dehydrogenases are dimer while some others are tetramer. We have also compared the structure of human P5C dehydrogenase, Mouse P5C dehydrogenase, Put2 and *Thermus* P5C dehydrogenase to find out the reason of *Thermus* and Put2 being hexameric.

In this article we have provided the biochemical evidence of only a single enzyme being involved in catalyzing the oxidation of P5C and 3-OH-P5C. Based on structure and kinetics, we have tried to explain the effect of mutations on the activity of human P5C dehydrogenase and possibly in hyperprolinemia II. The high-resolution structure of mouse P5C dehydrogenase allowed us to look into the details of the active site and the water mediated hydrogen-bonding network around the active site. This structure helped in explaining the effect of the S352L mutation on the activity of human P5CDH. The structure of dimeric ALDH4A1 with hexameric TtP5CDH and Put2 and other tetrameric aldehyde dehydrogenase were compared to find out the structural basis of different oligomeric states in the aldehyde dehydrogenase superfamily.
Materials and Methods

A. Cloning, expression, and purification

1. HsP5CDH –

   a. Constructs –

(1) Parent constructs –

The gene encoding HsP5CDH (NP_003739) was obtained from ATCC and subcloned into pET28a vector as follows. To delete the mitochondrial leader sequence, the human P5CDH gene was amplified from residues 18-563 using the primer AAGTCA\text{CATATG}ACCGGGGCCGGCCTGCG and AAGTCA\text{CTCGAG}TCACTGCATGTACGCGTAGC as the sense and antisense primer respectively. The HsP5CDH gene was cloned between the NdeI and XhoI restriction sites of pET28a. S352L, S352A, F388L, V470I and T473A mutants were created by the Quick change XLII mutagenesis kit using the following primers –

S352L Mutation –

Sense Strand – 5’-CCGCGTGCG\text{CT}TGCCTCTCTACGTGCCGC-3’

Antisense Strand- 5’-CGTAGAGACGCG\text{AG}GCACGC GGACACACTTCTGG-3’

S352A Mutation –

Sense Strand – 5’-CCGCGTGCG\text{CG}CGCTCTCTACGTGCCGC-3’

Antisense Strand- 5’-CGTAGAGACGCG\text{CG}GCACGC GGACACACTTCTGG-3’

F388L Mutation –

Sense Strand - 5’-GAGGATTTTG GGACCTTCTG\text{CT}TCTGCAGTGTATGATGCCC-3’

Antisense Strand – 5’-GGCATCAATTC ACTGCGACG\text{AG}GAAGGTCCC AAAATCCTTC-3’

V470I mutation –

Sense Strand – 5’-AGACGC\text{GT}TGCAG\text{AT} TGACAG\text{CC} AACC-3’

AntiSense Strand – 5’-TGTTGCTGT\text{CA}A\text{T}CAGCTGC\text{AG}CGTCTCC-3’
T473A Mutation –

Sense Strand – 5’-GCTGGTTGACAGCGCCACCAGCTATGGCC-3’

AntiSense Strand – 5’-CATAGCTGGTGCGCTGTCAACCAGCTGC-3’

Mutations were verified by DNA sequencing at the DNA core, MU.

(2) Short construct –
To get a crystal form free of crystallographic contacts involving the N-terminal of HsP5CDH, residues 1-33 were deleted from the HsP5CDH gene. The HsP5CDH gene was amplified from residues 33-563 using the primer AAGTCAGGATCC AACGAGCCCGTCTTAGC and the T7 terminator reverse primer and cloned in to the BamHI and XhoI restriction sites of pSV281.

b. Expression and purification –

(1) Parent construct
HsP5CDH and its mutants were transformed in to BL21(DE3)pLysS competent cells. Two liters of LB media, supplemented with 50 \( \mu \)g/l kanamycine, were inoculated with 1 % overnight grown starter culture and grown at 37\(^{\circ}\)C / 250 rpm until the OD\(_{600}\)=0.6. The wild type HsP5CDH and all of its mutants, except S352A were induced with 50 \( \mu \)M of IPTG for about 20 hours at 22\(^{\circ}\) C/ 200 rpm. HsP5CDH S352A was induced with 0.5 mM IPTG at 22\(^{\circ}\) C / 200 rpm for 20 hours. The cells were harvested by centrifuging the media at 3500 rpm for 30 minutes in a Sorvall SLC 6000 rotor. Cells were resuspended into 20 mM Hepes, 60 mM NaCl, 5 % glycerol, pH 8.0, quick frozen into liquid nitrogen, and stored at -80\(^{\circ}\) C until further purification.
Cells were thawed and broken by sonication. Cell debris was cleared from the supernatant by centrifuging the cell extract at 16500 rpm for 1 hour in a SS34 rotor. The supernatant, free of cell debris, was applied to a HisTrapHP Ni-affinity column, equilibrated with buffer A (20 mM Hepes, 300 mM NaCl, 5 % glycerol, pH 8.0). The column was washed with buffer A followed by buffer A plus 30 mM midazole. Protein was eluted with buffer A plus 300 mM midazole and column was washed for next round of purification with buffer B (20 mM Hepes, 300 mM NaCl, 5 % glycerol, 1 M Imidazole, pH 8.0).

To prepare the untagged protein, the protein after Ni-affinity purification was dialyzed into thrombin cleavage buffer ( 20 mM Tris, 150 mM NaCl, 2.5 mM CaCl2, 5 % glycerol, pH 8.0). The his-tag was removed by adding 1 unit of thrombin / 4 mg of P5CDH and incubating at 4°C for one day. Thrombin and tagged protein was removed from untagged protein by passing sample through HiTrap Benzamidine FF Affinity Column followed by HisTrap HP Ni-affinity column. Purity of untagged protein was evaluated by running SDS-PAGE. Pure protein was dialyzed in to precrystallization buffer (50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, 5 % glycerol, pH 8.0).

For the His-tagged protein, eluent after Ni-affinity purification was dialyzed into 50 mM Tris, 0.5 mM EDTA, 0.5 mM THP, 5 % glycerol, pH 8.0. Sample was loaded on to a HiTrap Q anion exchange column. HsP5CDH comes out of the column in the flow through. After verifying the purity of the protein by SDS-PAGE, protein was dialyzed in to precrystallization buffer.

(2) Short construct –
HsP5CDH 33-563 was expressed and purified as the parent construct, except a few
minor changes. Induction was done with 0.5 mM IPTG and the His-tag was cleaved with TEVP. TEVP and His-tagged protein were removed from untagged protein by passing the sample through a HisTrap HP Ni-affinity column. Untagged protein eluted with 30 mM Imidazole. Purified HsP5CDH 33-563 was dialyzed in the same precrystallization buffer used for parent constructs.

(3) Se-Met derivative of human P5C dehydrogenase –
Se-Met derivative of HsP5CDH was prepared by the method of metabolic inhibition\(^\text{10}\). An overnight starter culture was grown in LB media. Cells from the starter culture were pellet down and resuspended in minimal media supplemented with 2 mM MgSO\(_4\), 0.1 mM CaCl\(_2\) and 0.4 % glucose as the carbon source. Cells were grown until the OD\(_{600}\)=0.5 and methionine biosynthetic pathway was inhibited by adding 1 mg Lysine, 1 mg threonine, 1 mg phenylalanine, 0.5 mg leucine, 0.5 mg isoleucine, 0.5 mg valine, and 0.5 mg L-selenomethionine per liter media. Temperature was lowered to 22\(^\circ\) C, 200 rpm and cells were induced with 50 \(\mu\)M IPTG 30 minutes after adding the amino acids mixture. Cells were induced for 20 hours.

2. MmP5CDH
   a. cloning
MmP5CDH gene (NP\(_{780647.3}\)) was obtained from ATCC. To delete the mitochondrial leader sequence, the gene was amplified from residues 21-562 using primers AACCAATGCCATATGCTGCGATGGAAGCACACCTC and AACTGCGAGCTCGAGCTCATGAGATAGGAGTATCGC as sense and antisense primer respectively. The gene was cloned between the NdeI and XhoI restriction sites of pET28a vector.
b. expression and purification

MmP5CDH-pET28a was transformed in to BL21(DE3)pLysS competent cells. Two liters of LB media, supplemented with 50 µg/ml kanamycine, were inoculated with 1 % overnight grown starter culture. Cells were grown at 37°C/ 250 rpm until an OD600 of 0.6 then induced with 0.5 mM IPTG for overnight at 22°C/200 rpm.

Cells were harvested by centrifugation at 3500 rpm for 30 min. and resuspended in to 20 mM Hepes, 60 mM NaCl, 5 % glycerol, pH 8.0. Cells were quick frozen into liquid nitrogen and stored at -80°C until further purification.

The stored cells were thawed and broken by sonication and cell debris was cleared by centrifugation at 16500 rpm for 1 hour. The recovered supernatant was applied to a HisTrap HP Ni-affinity column pre-equilibrated with buffer A ( 20 mM Hepes, 300 mM NaCl, 5 % glycerol, pH 8.0). The column was washed with buffer A and buffer A plus 30 mM Imidazole. Protein was eluted with buffer A plus 300 mM Imidazole. The column was washed with buffer B (20 mM Hepes, 300 mM NaCl, 1 M Imidazole, 5 % glycerol, pH 8.0) for next round of purification.

Mouse P5CDH, purified from Ni affinity column, was dialyzed into 50 mM Tris, 0.5 mM EDTA, 0.5 mM THP, 5 % glycerol, pH 7.5 and loaded onto HiTrap Q anion exchange column. Protein was eluted with continuous gradient of NaCl from 0 to 1 M.

Fractions were analyzed by SDS-PAGE, pooled and dialyzed in to pre-crystallization buffer (50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, 5 % glycerol, pH 7.5)

3. Put2

a. Cloning

The Put2 gene, cloned between the NdeI and BamHI restriction sites of the pKA8H
vector, was obtained from the Donald Becker lab. To delete the leader sequence, the
gene was amplified from residues 23-575 using
AAAACATATGAAGCCCCCCAAGCACATAAG as the forward primer and T7
terminator as the reverse primer. The Put2 gene without leader sequence was cloned
into the pKA8H vector between the NdeI and BamHI restriction sites.

b. Expression and purification
Put2-pKA8H was transformed into Bl21(DE3)pLysS competent cells. One liters of
LB media, supplemented with 100 µg/ml ampicillin, was inoculated with 1 %
overnight grown starter culture. Cells were grown until the OD$_{600}$ = 0.6 at 37$^\circ$ C/250
RPM and induced with 0.5 mM IPTG for 8 hours at 22$^\circ$ C/200 rpm.
Cells were harvested by centrifugation at 3500 rpm for 30 minutes and resuspended
in to 20 mM Hepes, 60 mM NaCl, 5 % glycerol, pH 8.0. Cells were quick frozen
into liquid nitrogen and stored at -80$^\circ$ C until further purification.
The cells were broken by sonication and cell debris was cleared by centrifugation at
16500 rpm for one hour. The supernatant was applied to HisTrap HP Ni affinity
column, pre-equilibrated with buffer A ( 20 mM Hepes, 300 mM NaCl, 5 %
glycerol, pH 8.0). The column was washed with buffer A and buffer A plus 30 mM
Imidazole. Protein was eluted with buffer A plus 300 mM Imidazole. The column
was washed with buffer B (20 mM Hepes, 300 mM NaCl, 1 M Imidazole, 5 %
glycerol, pH 8.0) for next round of purification.
Put2 purified from the Ni-affinity column was dialyzed into 50 mM Tris, 0.5 mM
EDTA, 0.5 mM THP, 5 % glycerol, pH 8.0 and loaded onto a HiTrapQ anion
exchange column. Protein was eluted with continuous gradient of NaCl from 0 to 1
M.
Fractions were analyzed by SDS-PAGE, pooled and dialyzed into pre-crystallization buffer (50 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, 5 % glycerol, pH 8.0)

**B. Crystallization**

1. HsP5CDH

a. Crystal screening

Initial crystallization conditions were identified using crystal-screening kits. Screens were performed using both the His-tagged enzyme and the tag-free enzyme at 4°C and at room temperature respectively. The sitting drop method of vapor diffusion was used with drops formed by mixing 1 µL of the enzyme and 1 µL of the reservoir solution. These experiments led to two crystal forms having hexagonal (4°C, His-tagged) and primitive monoclinic lattices (20°C, Untagged).

(1) **Hexagonal crystal form** –

Hexagonal crystal form were grown in 20-25 % PEG3350, 0.2 M (NH₄)₂SO₄, 0.1 M HEPES, pH 7.0-8.0 at 4°C. Crystals were cryoprotected with 25 % PEG3350, 0.2 M (NH₄)₂SO₄, 0.1 M HEPES, 25 % glycerol, pH 7.5 and quick frozen by plunging the looped crystals into liquid nitrogen.

(2) **Monoclinic crystal form** –

Monoclinic crystal form were grown in 20-25 % PEG3350, 0.2 M MgCl₂, 0.1 M HEPES, pH 7.0-8.0 at 20°C. Crystals were cryoprotected with 25 % PEG3350, 0.2 M MgCl₂, 0.1 M HEPES, 20 % PEG200, pH 7.5 and quick frozen by plunging the looped crystals into liquid nitrogen.

(3) **Crystallization of Se-Met derivative**

The Se-Met derivative of human P5C dehydrogenase was crystallized in the
hexagonal crystal form. The crystals were grown in 20-25 % PEG3350, 0.2 M (NH₄)₂SO₄, 0.1 M HEPES, pH 7.0-8.0 at 4⁰C. Crystals were cryoprotected with 25 % PEG3350, 0.2 M (NH₄)₂SO₄, 0.1 M HEPES, 25 % glycerol, pH 7.5 and quick frozen by plunging the looped crystals into liquid nitrogen.

(4) Crystallization of HsP5CDH S352L
S352L failed to crystallize in the forms that had been obtained for the native enzyme. However, crystal screening trials revealed a triclinic form that grew in 20-25 % PEG3350, 0.1 M HEPES, 0.2 M MgCl₂, pH 7.0-8.0 using untagged HsP5CDH S352L at 3.5 mg/ml.

2. MmP5CDH
Mouse P5CDH crystals were grown in 20-25 % PEG3350, 0.2 M Li₂SO₄, 0.1 M Bis-Tris, pH 6.0-7.0. One ml of protein was mixed with equal volume of reservoir solution in a sitting drop crystal tray and incubated at 20⁰C. Crystal typically appears within 24 hours.

MmP5CDH crystals were cryoprotected with 25 % PEG3350, 0.2 M Li₂SO₄, 0.1 M Bis-Tris, 25 % glycerol, pH 6.25. For getting the ligand bound structure, crystals were soaked with a saturating concentration of glutamate in cryosolution (300 mM), 20 mM NAD⁺, 20 mM NADH, 20 mM NADH + 300 mM glutamate, and 150 mM proline for 6-8 hours.

3. Put2
Put2 crystals were grown in 8-14 % PEG3350, 0.1 M (NH₄)₂SO₄, 0.1 M Bis-Tris, pH 5.5-6.5. One µl of protein was mixed with equal volume of reservoir solution in sitting drop crystal tray and incubated at 20⁰C. crystals were cryoprotected with 15
% PEG3350, 0.1 M (NH₄)₂SO₄, 0.1 M Bis-Tris, pH 6.0, 20 % glycerol.

C. X-ray data collection

1. Se-Met hexagonal
   Hexagonal crystals of Se-Met HsP5CDH were analyzed at the NE-CAT beamline 24-ID-C at the Advanced Photon Source using a Quantum 315 detector. The data set that was used for single-wavelength anomalous diffraction (SAD) phasing was collected at the energy corresponding to the experimentally determined peak of $f''$ (wavelength = 0.979181 Å). This data set consisted of 180 frames collected with an oscillation width of 1.0° per image, detector distance of 375 mm, exposure time of 1.0 s/image, and transmission of 2.5 %. The data were processed to 2.85 Å resolution using HKL2000, with I+ and I- treated as nonequivalent reflections during scaling. The crystal was then translated, and a second data set was collected at higher transmission for purpose of phase extension and refinement. Denoted as Se-Met-2 in Table 4.2, this data set consisted of 60 frames collected with an oscillation width of 1.0° per image, detector distance of 350 mm, exposure time of 1.0 s/image, and transmission of 10.0 %. The data were processed using HKL2000 to 2.5 Å resolution, with I+ and I- treated as equivalent reflections during scaling.

2. Native data set – Primitive monoclinic crystals of HsP5CDH were also analyzed at NE-CAT beamline 24-ID-C (Table 4.2). The data set used for refinement consisted of 180 frames collected with an oscillation width of 1° per image and detector distance of 250 mm. The data were processed to 1.95 Å resolution with HKL2000.

3. Human P5C dehydrogenase S352L mutant – Triclinic crystals of S352L were analyzed at NE-CAT beamline 24-ID-E (Table 4.2). Reflections collected
from two crystals were merged to obtain a data set with acceptable completeness and redundancy. One of the data sets consisted of 163° of reflections collected with an oscillation width of 1° and detector distance of 325 mm. The other set consisted of 180° of data (1° oscillation width and detector distance of 330 mm). The data sets were integrated separately with MOSFLM through the iMosflm graphical interface. The integrated reflections were combined with Pointless and scaled to 2.55 Å resolution with SCALA using the CCP4i interface.

d. **MmP5CDH native crystals** – Data sets from MmP5CDH crystals were collected at NE-CAT beamline 24-ID-C and beamline 4.2.2 of the Advanced Light Source (Table 4.3). For the enzyme-sulfate complex, a 120° data set was collected with an oscillation width of 1° and detector distance of 150 mm. The data were processed with HKL2000. Data sets for the other complexes were collected at 4.2.2 and consisted of 180° scans with an oscillation width of 0.5° and detector distance of 100 mm. These data sets were integrated with XDS and scaled with SCALA.

e. **Put2 native** – Put2 crystals were analyzed at the NE-CAT beamline 24-ID-C at the Advanced Photon Source using a Quantum 315 detector.

**D. Phasing and refinement**

1. **HsP5CDH**
   a. **Se-Met phasing** – Molecular replacement calculations using search models derived from the TtP5CDH structure were unsuccessful, therefore the structure of HsP5CDH was solved using SAD phasing based on data collected hexagonal crystals of Se-Met HsP5CDH. The phasing potential of each data set was analyzed with the HKL2MAP interface to the SHELXC/D/E programs. Promising data sets were
input to PHENIX AutoSol for SAD phasing, density modification, and automated building calculation\textsuperscript{19}. Thirty of the expected 36 selenium sites (i.e., four HsP5CDH molecule per asymmetric unit) were identified, which resulted in a figure of merit of 0.39 for reflections to 2.85 Å resolution. Density modification, which included 4-fold noncrystallographic symmetry averaging, increased the figure of merit to 0.69. The model from automated building included 1790 residues (10194 atoms) and had an $R$-factor of 0.48 and map-model correlation coefficient of 0.59. PHENIX AutoBuild\textsuperscript{19} was used for phase extension to 2.5 Å resolution and additional model building. The resulting model included 1751 protein residues and 505 water molecules, and had an $R$-factor of 0.32, $R$-free of 0.37, and map-model correlation coefficient of 0.72.

The model from automated building was used as the starting point for several rounds of manual building in COOT\textsuperscript{20} and simulated annealing refinement against the 2.5 Å resolution hexagonal data set in PHENIX\textsuperscript{19}. The B-factor model used during refinement consisted of an isotropic B-factor for each non-hydrogen atom plus one TLS group per chain.

b. P21 molecular replacement – The structure of monoclinic HsP5CDH was solved using molecular replacement phasing as implemented in the programs MOLREP\textsuperscript{21} and PHASER\textsuperscript{22,23}. The structure could be readily solved with either program using monomer or dimer search models. A clear solution having two dimers in the asymmetric unit was obtained in space group $P2_1$. The model from molecular replacement was improved through rounds of manual building in COOT\textsuperscript{20} and simulated annealing refinement to 1.95-Å resolution in PHENIX\textsuperscript{19}. Refinement statistics are provided in Table 4.3.

c. Human P5C dehydrogenase S352L mutant – Structure of the S352L
mutant of human P5C dehydrogenase was solved by molecular replacement phasing as described for the human P5C dehydrogenase monoclinic crystal form. The structure was solved in space group P1 using dimer as the search model. I am using strict NCS refinement as implemented in CNS and manual building in Coot to improve the model from molecular replacement.

2. **Mouse P5C dehydrogenase molecular replacement** – Structure of the mouse P5C dehydrogenase was solved by molecular replacement phasing as described for the monoclinic crystal form of human P5C dehydrogenase. Human P5C dehydrogenase dimer was used as the search model. A clear solution with one dimer in the asymmetric unit was obtained in space group P12_12_1. The model from molecular replacement was improved through rounds of manual building in COOT and simulated annealing refinement to 1.30-Å resolution in PHENIX. The coordinates of the protein part of the 1.3 Å resolution MmP5CDH structure served as the starting point for refinements of MmP5CDH complexes. Refinement statistics are provided in Table 4.3.

3. **Put2 molecular replacement** – Structure of Put2 was solved by molecular replacement phasing. Mouse P5C dehydrogenase dimer was used as the search model and structure was solved in P6_5 space group with one dimer in the asymmetric unit. I am using manual building in Coot and simulated annealing refinement in PHENIX to improve the model obtained from molecular replacement.

**E. Kinetics**

1. **Standard bi-substrate kinetics**
   Kinetics of human P5CDH and its five mutants, S352L, S352A, F388L, V470I and T473A were done in 100 mM PO_4^-, 1 mM EDTA, pH 7.0. Kinetics of wild type and mutants, except S352L, were done at different concentration of NAD^+ and P5C using 10
µg / ml enzyme. Kinetics of S352L was done using 200 µg / ml enzyme, 0.2 mM P5C, and 0.35 mM NAD⁺. Formation of NADH was followed by increase in absorption at 340 nm.

To determine the kinetic parameters, data was globally fitted to following equation.

\[
v = \frac{V_{\text{max}} [\text{NAD}][\text{P5C}]}{K_{i}^{\text{NAD}} K_{m}^{\text{P5C}} + K_{m}^{\text{P5C}} [\text{NAD}] + K_{m}^{\text{NAD}} [\text{P5C}] + [\text{NAD}][\text{P5C}]}
\]

Equation (1)

2. P5C vs. OH-P5C kinetics
To determine the relative activity of human P5C dehydrogenase for the oxidation of P5C and OH-P5C, P5C and OH-P5C were generated in situ using 30 and 70 µg/ml PutA86-630 Y540S enzyme for proline and hydroxy-proline respectively and varying amounts of proline or hydroxy-proline in 20 mM HEPES, 10 mM MgCl₂, pH 8.0. 70 µg/ml HsP5CDH was used for the assay. Formation of glutamate, or hydroxy-glutamate, was followed by fluorescence of NADH with \( \lambda_{\text{ex}}=340 \) nm and \( \lambda_{\text{em}}=460 \) nm. Rate of P5C or OH-P5C production at different proline/hydroxy-proline concentration by PutA86-630 Y540S was measured using o-Aminobenzaldehyde and following the increase in absorption at 443 nm.

F. Equilibrium analytical ultracentrifugation – The quaternary structures of Human and mouse P5C dehydrogenase and Put2 in solution were analyzed using equilibrium analytical ultracentrifugation²⁵. Data were acquired at 20 °C using a Beckman XL-I Optima analytical ultracentrifuge equipped with an An50Ti rotor.
Prior to centrifugation, a protein sample having concentration of 0.8 mg/mL was
dialed into a buffer containing 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM
THP, 5 % glycerol, pH 8.0 for human P5C dehydrogenase and pH 7.5 for mouse P5C
dehydrogenase and Put2. Absorbance data (278 nm) were collected at three protein
concentrations (0.2 mg/mL, 0.4 mg/mL, 0.8 mg/mL) and three rotor speeds (4000 rpm,
8000 rpm, 12000 rpm for human P5C dehydrogenase and 5000, 7500 and 10000 for
mouse P5C dehydrogenase and Put2).

The nine sets of data were fit globally to a single-species model (equation 2) using
Origin 8.

\[ c(r) = c(r_0) \exp \left[ \frac{M \omega^2 (1 - \bar{\nu} \rho)}{2RT} \left( r^2 - r_0^2 \right) \right] + BL \quad \text{Equation (2)} \]

In this equation, \( c_I \) is the total protein concentration at radial position \( r \); \( c(r_0) \) is the
protein concentration at an arbitrary reference position, \( r_0 \); \( M \) is the molecular weight; \( \omega \)
is the radial velocity (rad/s); \( \bar{\nu} \) is the partial specific volume of the protein; \( \rho \) is the
solvent density; \( R \) is the gas constant; \( T \) is the absolute temperature; and \( BL \) is a baseline
offset. A solvent density of 1.02 g/cm\(^3\) was employed for the calculations, and the partial
specific volume was assumed to be the value predicted from the amino acid sequence
0.74 cm\(^3\)/g.

**G – Isothermal titration calorimetry** – Wild type human P5C dehydrogenase,
S352A and S352L mutants were dialyzed in 50 mM Tris, 50 mM NaCl, 0.5 mM THP,
0.5 mM EDTA, 5 % glycerol, pH 8.0 with three exchanges of buffer. The protein
concentration was measured by using extinction coefficient of 1.1 / mg P5C
dehydrogenase at 280 nm. pH of the protein was measured, NAD$^+$ was dissolved in
dialysis bath and its pH was adjusted to that of protein solution. Concentration of NAD$^+$
was measured by absorbance at 259 nm with an extinction coefficient of 16.9 / mM.
About 1.8 mL of protein solution, at 1 mg /ml (17.6 µM) was loaded in to ITC reaction
cell and titrated with 200 µM of NAD$^+$ solution. First injection volume was 4 µl and the
rest of the injection volume was 20 µl, which was injected into the ITC reaction cell at
the interval of 240 seconds. The experiments were done at 25° C with stirring speed of
250 rpm. The data was analyzed by Origin software.

**Results**

Structure of human P5C dehydrogenase was determined at 2.5-Å resolution by
Se-Met phasing using hexagonal crystal form. There are four protein chains and two
biologically active dimers in the asymmetric unit. P5C dehydrogenase belongs to ALDH
superfamily and its structure and overall fold is similar to other ALDH superfamily
members. It has a NAD$^+$ binding domain, a catalytic domain (residue 316-514) and a
dimerization domain (residue 183-199, 542-550), which interacts with the catalytic
domain of the other subunit to form a dimer (Fig. 4.1, 4.2). NAD$^+$ binding domain has
distorted Rossmann dinucleotide binding fold, which was first discovered in aldehyde
dehydrogenase 3 (ALDH3) and is found in aldehyde dehydrogenase superfamily. It has
five parallel and two antiparallel β-strand connected by four helix. Catalytic domain have
seven parallel and two antiparallel β-strands, two of them being contributed from the
dimerization domain of other subunit. Catalysis occurs between the cleft formed by the
NAD$^+$ binding and the catalytic domain. The two subunits of biologically active dimer
are similar to each other with root mean square deviation of 0.18 Å except the very end of
the N-terminus. In this study, it was noticed that the dimerization domain is also involved in higher order oligomer formation by aldehyde dehydrogenases. Since this crystal form was grown in 20-25 % PEG3350, 0.1 M HEPES (pH 7.0-8.0), and 0.2 M (NH₄)₂SO₄, there was a sulphate ion bound to the active site. As a result, this structure is a mimic of substrate bound state.

A higher resolution structure of human P5C dehydrogenase in monoclinic crystal form was solved by using search model obtained from hexagonal crystal form. This crystal was grown in 20-25 % PEG3350, 0.1 M HEPES (pH 7.0-8.0), and 0.2 M MgCl₂. This structure represents the ligand free state of the protein. There are four protein chains and two biologically active dimers in the asymmetric unit.

Structure of mouse P5C dehydrogenase was solved at 1.3 Å resolution by molecular replacement using human P5CDH as the search model. There are two protein chains in the asymmetric unit making the biologically active dimer. The mouse P5CDH structure is very similar to human P5CDH structure with root mean square deviation of 0.22 Å for one chain and 0.24 Å for both chain (Fig. 4.3). Structure of yeast Put2 was solved by molecular replacement using mouse P5CDH as search model. Although the overall fold of yeast Put2 is very similar to human P5C dehydrogenase, but the root mean square deviation between one chain of human P5C dehydrogenase and Put2 is 1.11 Å and both chain is 1.69 Å. Visual analysis of the alignment between human or mouse P5C dehydrogenase (dimeric) and Put2 or *Thermus thermophilus* P5C dehydrogenase (hexameric) suggests that the relative orientation of one subunit relative to the other is different in dimeric and hexameric P5C dehydrogenase (Fig. 4.4)
The dimerization domain and the dimer interface – Two P5CDH subunit makes a domain swapped dimer by intermolecular β-sheet formation involving β-sheet of the catalytic domain and two β-strand of the dimerization domain from the other subunit (Fig. 4.2). Some aldehyde dehydrogenases have three β-strand from the dimerization domain involved in intermolecular β-sheet formation. Apart from the dimeric interface discussed above, an additional intermolecular interaction has been observed involving a helix (residue 288-305). This helix interacts with the same helix from the other subunit. This helix along with helix 266-281 sandwich the adenine group of NAD$^+$ and has been suggested to be involved in intersubunit communication in ALDH2. Human and mouse P5CDH have relatively long C-terminus which is also involved in intersubunit interaction. The interface area, calculated with EBI-PISA, is 9246 Å$^2$ with 77 hydrogen bonding and 31 salt bridge interaction. It has 928 atoms and 252 residues (Fig. 4.5).

Active site and conformational changes associated with ligand binding – Human P5CDH in hexagonal crystal form and mouse P5CDH was grown in a condition that has 0.2 M SO$_4^{--}$, consequently sulphate ion was bound in the active site and occupies the same site as the main chain carboxylic group of glutamate or γ-glutamyl semialdehyde (Fig. 4.6). Since mouse P5C dehydrogenase structure is very similar to its human homologue, mouse P5C dehydrogenase crystals were used for soaking experiments. Structure of mouse P5C dehydrogenase with bound glutamate was obtained by soaking the crystal with a saturated solution of glutamate (~300 mM) in cryo-solution (25 % PEG3350, 25 % glycerol, 0.2 M Li$_2$SO$_4$, 0.1 M Bis-Tris, pH 6.25) (Fig. 4.7). α-carboxylic group of glutamate makes hydrogen bonds with backbone amide group of Gly512 and Ser513 and hydroxyl group of Ser349 and Ser513. Amine group of glutamate
also makes hydrogen-bonding interaction with hydroxyl group of Ser513. Besides direct interaction of glutamate with protein atoms, there are three water mediated hydrogen-bonding interaction between protein atoms and glutamate. These are interactions between carboxyl group of glutamate and amine group of Lys347 and interaction between amine group of glutamate with hydroxyl and carbonyl group of Ser513 and carboxylic group of Glu165. Sulphate ion, which is present in the active site of human P5CDH hexagonal crystal form and mouse P5CDH structure, make the same interaction with the protein that carboxylic group of glutamate make i.e. interaction with the backbone N-H group of Gly512 and Ser513 and hydroxyl group of Ser513 and Ser349. The side chain carboxylic group of glutamate, which is a mimic of aldehyde group of γ-glutamyl semialdehyde, interacts with backbone N-H group of Cys348 and Asn211 side chain.

Structure of human P5CDH in monoclinic crystal form do not have sulphate ion bound to the active site. There are some regions of protein, disordered in this structure. There are four molecules in the asymmetric unit. Residues 514-535 are disordered in one subunit of each dimer. Part of this region except residue 515-521, is ordered in the other subunit. However, the conformation of this region is different from mouse P5CDH, human P5CDH in hexagonal form and *Thermus thermophilus* P5CDH (Fig. 4.8).

Residue 183-193 is disordered in one subunit of each dimer while only residue 188-193 is disordered in the other subunit. Conformation of residue 183-188 is different from the mouse P5CDH structure. The C-terminus end of the human P5CDH is also disordered in all the four chain.

**Cofactor binding** – NAD$^+$ and NADH complex of mouse P5CDH were obtained by soaking the crystals with 20 mM NAD$^+$ or NADH. NAD$^+$ and NADH bind to the P5C dehydrogenase in extended conformation with adenine ring sandwiched between
the two helices, helix 288-304 and helix 266-275. The ribose group adjacent to adenine ring interacts, through its hydroxyl group, with side chain amine group of Lys233 and backbone carbonyl group of Ser208 (Fig. 4.9).

However the conformation of pyrophosphate group, ribose group adjacent to nicotinamide, and nicotinamide ring are different in NAD\(^+\) and NADH complex. Because of the movement of pyrophosphate group, Nicotinamide ribose group is pulled away from the active site in NADH complex (Fig.4.10). Hydroxyl group of Nicotinic ribose group interacts with a conserved glutamate (Glu447) and Arg399 residue.

Cys315 and Gly316 have two conformations. In one conformation, Cys315 carbonyl group point towards NAD\(^+\) amide group while in the other conformation it points away from NAD\(^+\) amide group. Glu314 is disordered in all the structure of MmP5CDH.

Conformation of the loop (208-217) adjacent to the ADP-ribose group is different in monoclinic crystal form of human P5CDH and in mouse P5CDH. Conformation of this loop in monoclinic crystal form is such that NAD\(^+\) binding will create a steric clash between ADP-ribose group and Ser208 and Pro209.

**Single nucleotide polymorphism** – There are six single nucleotide polymorphism reported to be present in human P5C dehydrogenase. These are E277K, S352L, G366R, F388L, V470I and T473A. S352L is known to be responsible for the type II hyperprolinemia. The role of the other mutations in disease is not known. Effect of these mutations on the enzymatic activity of human P5CDH is unknown. Residue Glu277 and Gly366 are on the surface of the protein. However Ser352, Phe388, Val470, and Thr473 are buried inside the protein structure (Fig. 4.11). Ser352 is close to the active site and near the catalytic cysteine (Cys348). This residue is surrounded by polar residues and
is connected to the catalytic loop and backbone atom of catalytic cysteine through a hydrogen-bonding network. This hydrogen-bonding network connecting the Ser352 and catalytic cysteine is shown in Fig. 4.12 using mouse P5CDH as a higher resolution model of human P5C dehydrogenase. In addition, Fig. 4.12 also shows that residue Ser349 is involved in substrate recognition and interacts with the carboxylic group of glutamyl semialdehyde thus helping to position the aldehyde group near catalytic cysteine.

**Structure of human P5C dehydrogenase S352L mutant** – Structure of S352L mutant of human P5C dehydrogenase was solved in monoclinic crystal form at 2.6 Å resolution. There are eight molecules in the asymmetric unit making four dimer. Overall fold of the mutant enzyme and the structure of dimer are similar to wild type enzyme except the conformation of catalytic loop. There are significant conformational changes in residues 348-352, which shows shift towards NAD$^+$ binding loop (residues 445-452) (Fig. 4.13). Structural alignment between mouse P5C dehydrogenase-NAD$^+$ complex and human P5C dehydrogenase S352L mutant shows that Ser349 residue of S352L mutant occupies the place occupied by Glu447 (Fig. 4.14).

**Kinetic characterization of human P5C dehydrogenase** – Steady state kinetics of human P5C dehydrogenase and its mutants were done with NAD$^+$ as changing fixed substrate and P5C as variable substrate. At one fixed concentration of NAD$^+$, P5C was varied from 22 µM to 600 µM and the same experiment was repeated at different concentration of NAD$^+$. Kinetic parameters were obtained by globally fitting the data to following equation for compulsory ordered bi-bi reaction.

Equation (1)
\( K_i^{NAD} \) is the dissociation constant of P5CDH-NAD\(^+\) complex, \( \frac{1}{2} [NAD]^+ \) is the concentration of \( [NAD]^+ \) at saturating concentration of P5C which yield a velocity of \( \frac{1}{2} V_{\text{max}} \) and \( K_i^{P5C} \) is the concentration of P5C at saturating concentration of \( [NAD]^+ \) which yield a velocity of \( \frac{1}{2} V_{\text{max}} \). Human P5C dehydrogenase showed the phenomenon of substrate inhibition by P5C at high P5C concentration (~600 \( \mu\)M), which is more pronounced at lower \( [NAD]^+ \) concentration. So the data point showing the substrate inhibition was discarded from the analysis.

\[ K_i^{NAD} \quad K_i^{NAD} \]

The graphs of P5C concentration vs. rate of NADH production and global fit analysis of wild type enzyme and different mutants are shown in Figure 4.15 to 4.17. The kinetics parameters are shown in Table 4.4.

\[ K_i^{NAD} \quad K_i^{NAD} \]

Both the mutants have similar kinetic parameters as wild type enzyme except S352L mutant. S352L mutant of human P5C dehydrogenase is 150 fold less active then wild type under similar reaction condition even when S352L mutant was used in 20 fold excess. However, S352A mutant has comparable activity as wild type enzyme (Table 4.4).

**Thermodynamics of NAD\(^+\) binding to wild type human P5C dehydrogenase, S352A and S352L mutants** – Thermodynamic parameters for NAD\(^+\) binding to wild type human P5C dehydrogenase and two mutants, S352A and S352L were measured by using isothermal titration calorimetry (ITC). Data was fitted to single site model. NAD\(^+\) binding to wild type human P5C dehydrogenase is highly exothermic with enthalpy of NAD\(^+\) binding -16980 ± 2195 Kcal/mole, dissociation constant of 6.6 ± 1 \( \mu\)M and entropy of -33.27 Kcal/mole. The stoichiometry of NAD\(^+\) binding is 0.57 ± 0.06 suggesting half-the-sites reactivity for P5C dehydrogenase where
only one subunit of the dimer binds to the NAD$. The ITC experiment for S352A was repeated four times and data were fitted globally for four parameters; association constant, stoichiometry, enthalpy, and base line. Dissociation constant, enthalpy and entropy of NAD$^+$ binding to S352A mutant are 10.37±2.42 µM, -20539±4731 Kcal/mole and -47.89 Kcal/mole. NAD$^+$ binding to S352A has similar stoichiometry to wild type enzyme, 0.47±0.05. However there was no detectable NAD$^+$ binding to S352L when isothermal titration calorimetry (ITC) was done under the similar experimental condition.

**Oxidation of 3-OH-P5C by P5C dehydrogenase** – It has been suggested, based on genetic evidence, that P5C dehydrogenase also catalyzes the oxidation of 3-OH-P5C to hydroxy-glutamate, a degradation product of hydroxy-proline catabolism$^{26}$.

Ostrander et al.$^{27}$ have prepared a mutant of *E. coli* PutA 86-630 Y540S, which can also oxidize OH-Proline to OH-P5C, although seven times less efficiently then proline. PutA86-630 Y540S mutant was used to generate P5C and 3-OH-P5C in-situ using varying amount of proline and OH-Proline. P5C dehydrogenase present in the reaction mixture, oxidizes the P5C/OH-P5C thus produced, to glutamate/hydroxy-glutamate with the concomitant reduction of NAD$^+$ to NADH. The reaction is described in the figure below.

The oxidation of P5C/3-OH-P5C to glutamate can be followed by increase in absorbance at 340 nm due to formation of NADH. However coenzyme Q1 (Ubiquinone 1, UQ1) was used to reoxidize the FAD cofactor of PutA86-630 Y540S for further round
of catalytic cycle. Coenzyme Q1 shows a decrease in absorption at 340 nm after reduction. To overcome the problem of spectral overlap, NADH fluorescence was used to follow the oxidation of P5C/OH-P5C. Fluorescence of NADH was converted to the NADH concentration by making a NADH standard curve (Fig. 4.25). Fig. 4.24 shows the formation of NADH at three different concentration of proline/OH-proline. This figure also shows that there is no NADH formed, when P5C dehydrogenase was absent from the reaction mix. Since PutA86-630 Y540S catalyze the oxidation of hydroxy-proline less efficiently than proline, a higher amount of this enzyme was used for hydroxy-proline. Rate of P5C/OH-P5C production was measured using o-aminobenzaldehyde. o-aminobenzaldehyde react with γ-glutamyl semialdehyde to make an adduct which absorbs at 443 nm with extinction coefficient of 2.71. Figure 4-29 shows the rate of production of NADH at 150, 300, and 450 mM proline and hydroxy-proline. There is decrease in rate of NADH production with increasing concentration of proline. However rate of production of P5C at 150, 300 and 450 mM proline was 11.5, 11.3 and 11.45 µM/min. For hydroxy-proline, rate of production of NADH are similar at all the three concentration of hydroxy-proline. However the rate of production of 3-OH-P5C at 150, 300 and, 450 µM hydroxy-proline was 10.66, 15.79 and 19.5 µM/min.

\[
\begin{align*}
\text{γ-glutamyl semialdehyde} & \quad + \quad \text{o-Amino benzaldehyde} \\
\text{The adduct} & 
\end{align*}
\]
Quaternary structure of P5C dehydrogenase – Quaternary structure of P5C dehydrogenase in solution was studied by using equilibrium analytical ultracentrifugation. Molecular weights of monomer of human and mouse P5C dehydrogenase (His-tagged) and Put2 (untagged) are 62.07, 61.87, and 62.61 Kda, as calculated from sequence using ProtParam (http://expasy.org/cgi-bin/protparam). Data from equilibrium analytical ultracentrifugation was fitted to one species model by simultaneously fitting nine data from three different protein concentrations and three different rotor speeds. The fitted data for human P5C dehydrogenase, mouse P5C dehydrogenase and Put2 are shown in figure 4.26, 4.27 and 4.28 respectively. Molecular weights of human and mouse P5C dehydrogenase were 118 ± 1 Kda and 134±1 Kda respectively. Molecular weight of Put2 was 313± Kda.

Discussion:
In the present study, we have solved the crystal structure of eukaryotic P5C dehydrogenase representing first of its kind. P5C dehydrogenase from three different organisms (human, mouse and yeast) was solved using four different crystal forms. These structures shed light on the substrate recognition and associated conformational changes, molecular mechanism of type II hyperprolinemia and structural basis of different oligomeric state in P5C dehydrogenase and aldehyde dehydrogenase.

Crystallization conditions of human P5C dehydrogenase (hexagonal crystal form), mouse P5C dehydrogenase and Put2 had sulphate ion. There was a sulphate ion, bound to the active site of these structures. The sulphate ion occupies the same place as occupied by a-carboxylic group of glutamate (or γ-glutamyl semialdehyde). This sulphate ion connects Ser349 with Gly512 and Ser513 through hydrogen bonding interactions, therefore restricting the conformational flexibility of residues 514-535. Structure of
human P5C dehydrogenase in monoclinic crystal form was crystallized in a condition without sulphate ion. Consequently there is no sulphate ion bound to the active site of this structure and residues 514-535 is disordered or only partially ordered. The corresponding residues are also disordered in retinal dehydrogenase and binding of retinal to retinal dehydrogenase restrain its flexibility and order it\textsuperscript{38}. We suggest that in the absence of substrate, residues 514-535 are disordered in human P5C dehydrogenases and substrate binding induces a conformational change inducing the ordering of residues 514-535. This ligand induced conformational changes also explain why succinic semialdehyde, which is one carbon shorter than \(\gamma\)-glutamyl semialdehyde, is a poor substrate in comparison to \(\gamma\)-glutamyl semialdehyde.

Structure of mouse P5C dehydrogenase was also solved in NAD\textsuperscript{+} and NADH bound state. In the structure of mouse P5C dehydrogenase complex, nicotinamide ring of NAD\textsuperscript{+} is in the active site. Consequently, to prevent the steric clash with nicotinamide group, the sulfhydryl group of catalytic cysteine (Cys348) is pointing towards the active site. However, in the structure of P5C dehydrogenase-NADH complex, nicotinamide ring of NADH is pulled back from the active site. Glu314 is disordered in all the structure of mouse P5C dehydrogenase. However the corresponding residue in Thermus thermophilus P5C dehydrogenase\textsuperscript{9} and other aldehyde dehydrogenases show two conformations; in NAD\textsuperscript{+} bound state, Glu314 side chain points away from the active site and in NAD\textsuperscript{+} free state, it points towards the active site. This residue serves as a general base during catalysis. During catalysis in aldehyde dehydrogenase, thiolate ion formed by deprotonation of catalytic cysteine attacks the carbonyl carbon of the aldehyde group of the substrate making a thiohemiacetal intermediate\textsuperscript{29}. The second step is hydride transfer from the thiohemiacetal intermediate to NAD\textsuperscript{+} making thioester intermediate and NADH.
The last step is the hydrolysis of thioester intermediate to yield free cysteine and carboxylic acid. It requires activation of water molecule by Glu314, which must be pointing towards the active site. All the aldehyde dehydrogenases show the NADH isomerization where Nicotinamide ring (as well as nicotinic ribityl and pyrophosphate group) either completely dissociated from the enzyme while adenosine moiety still attached to the enzyme or Nicotinamide ring is pulled back from the active site due to conformational changes in the nicotinic ribityl and pyrophosphate group\textsuperscript{30}. The isomerization of NADH allows Glu314 to swing back towards the active site and activate the water molecule, which attacks on the thioester intermediate to release carboxylic acid.

S352L mutation in human P5C dehydrogenase gene causes type II hyperprolinemia. Structure of S352L mutant along with kinetics and thermodynamics was used to understand the molecular mechanism of type II hyperprolinemia caused by S352L mutation. The structure suggests that there is a significant conformational change in the catalytic loop (residues 348-352). This conformational change in catalytic loop put Ser349 in a place previously occupied by carboxylic group of Glu447, causing displacement of Glu447 due to steric clash. Glu447 is conserved among aldehyde dehydrogenase and interacts with the nicotinic ribose group of NAD\textsuperscript{+}. Mutation or any change in the conformation of this residue may affect NAD\textsuperscript{+} binding. Further, S349 interacts with the side chain carboxylic group of substrate and position the substrate in catalytically competent conformation. Due to the conformational changes in the catalytic loop, S349 will no longer be able to serve its purpose and S352L mutant will not be able to binds to the substrate. Most importantly, there is significant displacement of catalytic cysteine residue (Cys348) in S352L mutant relative to wild type enzyme.
Result of the crystal structure was supported by kinetics and ITC results. S352L mutant is inactive and ITC suggest that there is no detectable NAD$^+$ binding to S352L mutant. Structural analysis revealed that Ser352 is in a polar environment and the hydroxyl group of Ser352 interacts with backbone carbonyl group of catalytic cysteine through a network of hydrogen bonding interaction (Fig. 4.12). There are two possibilities, why S352L mutant is inactive? Since mutation of serine to leucine do not have hydroxyl group, the hydrogen-bonding network connecting serine hydroxyl group to cysteine carbonyl group will be perturbed causing the conformational changes in catalytic cysteine and thus inactive protein. Secondly, Leucine is a non-polar residue and putting a non-polar group in a polar environment could break this hydrogen-bonding network. To investigate the two possibilities, S352A mutant was made and characterized kinetically and thermodynamically. Kinetic data suggest that S352A mutant has similar activity as wild type enzyme and its kinetic parameters are similar to wild type enzyme. ITC result also supports the kinetic data. Wild type and S352A mutant has similar dissociation constant for NAD$^+$ so NAD binding is not affected by S352A mutation. These results suggest that S352L mutant is inactive because the mutation puts a non-polar residue in to a polar environment. It is further corroborated by sequence alignment of aldehyde dehydrogenase, which shows that there is always a polar residue at Ser352.

Due to the unavailability of 3-OH-P5C, oxidation of 3-OH-P5C by P5C dehydrogenase was studied in a coupled assay where 3-OH-P5C was generated in-situ by using PutA86-630 Y540S mutant. Since the efficiency of oxidation of proline and hydroxy-proline by PutA86-630 Y540S are different, different amount of enzyme was used for proline and hydroxy-proline. At 150 mM proline/hydroxy-proline, rate of production of P5C/3-OH-P5C was similar. However rate of NADH formation is 10 fold
higher when proline was used as substrate then when hydroxy-proline was used. These results suggest that the oxidation of P5C by P5C dehydrogenase is more efficient than the oxidation of 3-OH-P5C. The result also shows that there is a decrease in the rate of NADH production with increasing concentration of proline. However, rate of production of P5C was similar at all the three concentration of proline. Although there is not a significant decrease in the rate of production of NADH with increasing concentration of hydroxy-proline, but rate of production of 3-OH-P5C was increased with increasing concentration of hydroxy-proline. This decrease in rate of oxidation of P5C/3-OH-P5C with increasing concentration of proline/hydroxy-proline may be due to the inhibition of P5C dehydrogenase by proline/hydroxy-proline.
References –


<table>
<thead>
<tr>
<th>Gene</th>
<th>Enzyme</th>
<th>Substrate</th>
<th>Oligomeric state</th>
<th>PDB ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDH1</td>
<td>ALDH1</td>
<td>Retinal</td>
<td>Tetramer</td>
<td>1BXS, 1BI9</td>
</tr>
<tr>
<td>ALDH2</td>
<td>ALDH2</td>
<td>Acetaldehyde</td>
<td>Tetramer</td>
<td>1O05, 3INJ</td>
</tr>
<tr>
<td>ALDH3</td>
<td>ALDH3</td>
<td>Fatty and aromatic aldehyde</td>
<td>Dimer</td>
<td>1AD3</td>
</tr>
<tr>
<td>ALDH4</td>
<td>ALDH4</td>
<td>Glutamate γ-semialdehyde</td>
<td>Dimer or hexamer</td>
<td>1UZB</td>
</tr>
<tr>
<td>ALDH5</td>
<td>ALDH5</td>
<td>Propionaldehyde</td>
<td>Tetramer</td>
<td>2W8N</td>
</tr>
<tr>
<td>ALDH6</td>
<td>ALDH6</td>
<td>Aliphatic aldehyde, retinal</td>
<td>Not known</td>
<td></td>
</tr>
<tr>
<td>ALDH7</td>
<td>ALDH7</td>
<td>Aliphatic and aromatic aldehyde</td>
<td>Not known</td>
<td></td>
</tr>
<tr>
<td>ALDH8</td>
<td>ALDH8</td>
<td>unknown</td>
<td>Not known</td>
<td></td>
</tr>
<tr>
<td>ALDH9</td>
<td>γABDH</td>
<td>Amine aldehyde</td>
<td>Not known</td>
<td></td>
</tr>
<tr>
<td>ALDH10</td>
<td>FALDH</td>
<td>Aliphatic and aromatic aldehyde</td>
<td>Not known</td>
<td></td>
</tr>
<tr>
<td>SSDH</td>
<td>SSDH</td>
<td>Succinic semialdehyde</td>
<td>Not known</td>
<td></td>
</tr>
<tr>
<td>MMSDH</td>
<td>MMSDH</td>
<td>Methylmalonate semialdehyde</td>
<td>Not known</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.2: Data collection and refinement statistics for HsP5CDH

<table>
<thead>
<tr>
<th></th>
<th>Se-Met-1</th>
<th>Se-Met-2</th>
<th>native</th>
<th>S352L</th>
</tr>
</thead>
<tbody>
<tr>
<td>space group</td>
<td>P6\textsubscript{5}</td>
<td>P6\textsubscript{5}</td>
<td>P\textsubscript{2}\textsubscript{1}</td>
<td>P1</td>
</tr>
<tr>
<td>unit cell lengths (Å)</td>
<td>$a = 150.7,$</td>
<td>$a = 150.7,$</td>
<td>$a = 92.0$</td>
<td>$a = 66.2$</td>
</tr>
<tr>
<td></td>
<td>$c = 191.6$</td>
<td>$c = 192.0$</td>
<td>$b = 121.3$</td>
<td>$b = 94.5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$c = 93.4$</td>
<td>$c = 182.4$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\beta = 104.2$</td>
<td>$\beta = 76.9$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\gamma = 70.2$</td>
</tr>
<tr>
<td>wavelength</td>
<td>0.97918</td>
<td>0.97918</td>
<td>0.97949</td>
<td>0.97925</td>
</tr>
<tr>
<td>resolution (Å)</td>
<td>50.0 – 2.85</td>
<td>50.0 – 2.50</td>
<td>50.0 – 1.95</td>
<td>62.3 – 2.55</td>
</tr>
<tr>
<td></td>
<td>(2.95 – 2.85)</td>
<td>(2.59 – 2.50)</td>
<td>(2.02 – 1.95)</td>
<td>(2.69 – 2.55)</td>
</tr>
<tr>
<td>no. of observations</td>
<td>661829</td>
<td>321938</td>
<td>489835</td>
<td>413388</td>
</tr>
<tr>
<td>no. of unique reflections</td>
<td>113595</td>
<td>84960</td>
<td>139296</td>
<td>128074</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>100.0 (100.0)</td>
<td>99.7 (100.0)</td>
<td>96.5 (93.0)</td>
<td>97.1 (93.6)</td>
</tr>
<tr>
<td>redundancy</td>
<td>5.8 (5.8)</td>
<td>3.8 (3.8)</td>
<td>3.5 (3.3)</td>
<td>3.2 (3.1)</td>
</tr>
<tr>
<td>$R_{merge}(I)$</td>
<td>0.100 (0.524)</td>
<td>0.075 (0.539)</td>
<td>0.049 (0.453)</td>
<td>0.102 (0.487)</td>
</tr>
<tr>
<td>average $I/\sigma$</td>
<td>22.9 (4.4)</td>
<td>17.0 (2.5)</td>
<td>24.2 (2.2)</td>
<td>7.2 (2.1)</td>
</tr>
<tr>
<td>$R_{cryst}$</td>
<td></td>
<td></td>
<td>0.2035</td>
<td>0.2001</td>
</tr>
<tr>
<td>$R_{free}$\textsuperscript{b}</td>
<td></td>
<td></td>
<td>0.2422</td>
<td>0.2461</td>
</tr>
<tr>
<td>no. of protein residues</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no. of protein atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no. of water molecules</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average B-factor (Å$^2$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>water</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>active site ligand</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rmsd\textsuperscript{e}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond lengths (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond angles (deg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot\textsuperscript{d}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>favored (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>allowed (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outliers (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>coordinate error (Å)\textsuperscript{f}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDB accession code</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}Values for the outer resolution shell of data are given in parenthesis.
\textsuperscript{b}5% random test set.
\textsuperscript{c}Compared to the parameters of Engh and Huber.\textsuperscript{31}
\textsuperscript{d}The Ramachandran plot was generated with RAMPAGE.\textsuperscript{32}
Table 4.3: Data collection and refinement statistics for MmP5CDH<sup>a</sup> and Put2

<table>
<thead>
<tr>
<th></th>
<th>Sulfate</th>
<th>Glutamate 1-2</th>
<th>NAD&lt;sup&gt;+&lt;/sup&gt; 1-11</th>
<th>NADH 1-9</th>
<th>NADH/Glu 1-3</th>
<th>Put2</th>
</tr>
</thead>
<tbody>
<tr>
<td>space group</td>
<td>(P2_1,2_1,2_1)</td>
<td>(P2_1,2_1,2_1)</td>
<td>(P2_1,2_1,2_1)</td>
<td>(P2_1,2_1,2_1)</td>
<td>(P2_1,2_1,2_1)</td>
<td>(P6_3)</td>
</tr>
<tr>
<td>unit cell lengths (Å)</td>
<td>(a = 85.2, b = 94.0, c = 132.4)</td>
<td>(a = 84.8, b = 93.9, c = 132.2)</td>
<td>(a = 84.9, b = 94.0, c = 132.4)</td>
<td>(a = 84.8, b = 94.2, c = 132.2)</td>
<td>(a = 84.9, b = 94.2, c = 132.4)</td>
<td>(a = 109, b = 181)</td>
</tr>
<tr>
<td>wavelength</td>
<td>0.97949</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00</td>
</tr>
<tr>
<td>resolution (Å)</td>
<td>50.0 – 1.30 (1.32 – 1.30)</td>
<td>47.0 – 1.50 (1.58 – 1.50)</td>
<td>47.0 – 1.60 (1.69 – 1.60)</td>
<td>47.0 – 1.50 (1.58 – 1.50)</td>
<td>47.0 – 1.50 (1.58 – 1.50)</td>
<td>50.1-1.95</td>
</tr>
<tr>
<td>no. of observations</td>
<td>914888</td>
<td>1222700</td>
<td>1221816</td>
<td>1013963</td>
<td>1231354</td>
<td>579452</td>
</tr>
<tr>
<td>no. of unique reflections</td>
<td>247337</td>
<td>168017</td>
<td>168338</td>
<td>137989</td>
<td>167674</td>
<td>88348</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>94.9 (86.0)</td>
<td>99.5 (96.5)</td>
<td>99.6 (97.6)</td>
<td>98.8 (97.7)</td>
<td>99.0 (96.1)</td>
<td>99.9(100)</td>
</tr>
<tr>
<td>redundancy</td>
<td>3.7 (3.6)</td>
<td>7.3 (6.7)</td>
<td>7.3 (6.8)</td>
<td>7.3 (7.5)</td>
<td>7.3 (6.9)</td>
<td>6.6(6.5)</td>
</tr>
<tr>
<td>(R_{merge}(I))</td>
<td>0.055 (0.502)</td>
<td>0.066 (0.418)</td>
<td>0.061 (0.399)</td>
<td>0.080 (0.521)</td>
<td>0.074 (0.480)</td>
<td>0.061(0.525)</td>
</tr>
<tr>
<td>average (I/\sigma)</td>
<td>26.6 (2.4)</td>
<td>25.3 (4.8)</td>
<td>25.3 (4.9)</td>
<td>19.4 (3.6)</td>
<td>22.2 (4.3)</td>
<td>32.1(3.4)</td>
</tr>
<tr>
<td>(R_{cryst})</td>
<td>0.1316</td>
<td>0.1709</td>
<td>0.1767</td>
<td>0.1660</td>
<td>0.1665</td>
<td>0.1902</td>
</tr>
<tr>
<td>(R_{free})^b</td>
<td>0.1601</td>
<td>0.1907</td>
<td>0.1961</td>
<td>0.1909</td>
<td>0.1860</td>
<td>0.2261</td>
</tr>
<tr>
<td>no. of protein residues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no. of protein atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no. of water molecules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average B-factor (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>active site ligand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rmsd&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond lengths (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond angles (deg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>favored (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>allowed (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outliers (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>coordinate error (Å)&lt;sup&gt;f&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDB accession code</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Values for the outer resolution shell of data are given in parenthesis.
<sup>b</sup>5% random test set.
<sup>c</sup>Compared to the parameters of Engh and Huber.\(^31\)
<sup>d</sup>The Ramachandran plot was generated with RAMPAGE.\(^32\)
<sup>e</sup>Maximum likelihood-based coordinate error reported by PHENIX.
Table 4.4: Kinetic parameters of wild type and mutant HsP5CDH

<table>
<thead>
<tr>
<th>Mutant/kinetic parameter</th>
<th>Wild type</th>
<th>S352A</th>
<th>V470I</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_m(NAD^+)$</td>
<td>190±39</td>
<td>298±64</td>
<td>479±83</td>
</tr>
<tr>
<td>$K_m(P5C)$</td>
<td>174±44</td>
<td>145±43</td>
<td>278±67</td>
</tr>
<tr>
<td>$V_{max}$</td>
<td>251±24</td>
<td>159±16</td>
<td>272±30</td>
</tr>
<tr>
<td>$K_i(NAD^+)$</td>
<td>13±34</td>
<td>107±95</td>
<td>52±43</td>
</tr>
</tbody>
</table>
Figure Legend

Figure 4.1 – Protomer structure of human P5C dehydrogenase. NAD$^+$ binding domain is shown in blue, catalytic domain in cyan and dimerization domain in magenta. Catalytic cysteine residue (Cys348) is drawn in sphere.

Figure 4.2 – Structure of human P5C dehydrogenase dimer. Dimerization domain of one chain interacts with the catalytic domain of the other chain to make the dimer.

Figure 4.3 – Structural alignment of human and mouse P5C dehydrogenase. Human P5C dehydrogenase is shown in green and mouse P5C dehydrogenase is shown in magenta.

Figure 4.4 – Structural alignment of mouse and *Thermus thermophilus* P5C dehydrogenase. Mouse P5C dehydrogenase is shown in cyan and *Thermus thermophilus* P5C dehydrogenase is shown in magenta.

Figure 4.5 – Protomer structure of human P5C dehydrogenase with dimer interface residues shown in red.

Figure 4.6 – Structure of Mouse P5C dehydrogenase active site showing its direct interactions with sulphate ion. There are several water-mediated hydrogen bonding interactions between sulphate ion and protein not shown in figure. $F_o - F_c$ map superimposed on sulphate ion and contoured at 3 $\sigma$.

Figure 4.7 – Structure of Mouse P5C dehydrogenase active site showing its interactions with bound product glutamate. $F_o - F_c$ map superimposed on glutamate and contoured at 3 $\sigma$. 
Figure 4.8 – Structural alignment of human P5C dehydrogenase in monoclinic and hexagonal crystal form. The loop from hexagonal crystal form is colored in blue and from monoclinic crystal form is colored in orange.

Figure 4.9 – Structure of mouse P5C dehydrogenase complexed with NAD\(^+\). NAD\(^+\) is drawn in sticks and hydrogen bonding interactions between NAD\(^+\) and protein is shown with black dash.

Figure 4.10 – Structure of mouse P5C dehydrogenase complexed with NADH. NADH is drawn in sticks and hydrogen bonding interactions between NADH and protein is shown with black dash. Overlapped is the structure of P5C dehydrogenase-NAD\(^+\) complex showing only NAD\(^+\). NAD\(^+\) is colored yellow.

Figure 4.11 – Structure of human P5C dehydrogenase dimer showing single nucleotide polymorphism.

Figure 4.12 – Structure of Mouse P5C dehydrogenase showing hydrogen-bonding network connecting Ser352 and Cys348.

Figure 4.13 – Structural alignment of human P5C dehydrogenase (hexagonal form) with S352L mutant. Residues 348-352 of native enzyme is colored orange and S352L is shown in sticks with superimposed F\(_0\) map contoured at 1 \(\sigma\). Residues 348-352 of S352L is moved towards the loop involved in NAD\(^+\) binding (colored wheat).

Figure 4.14 – Stereo representation showing interaction of Glu447 with nicotinic ribose group and possible steric clash in S352L mutant. Mouse P5C dehydrogenase NAD\(^+\) complex (cyan) was superimposed on S352L mutant structure (green). Residues 445-452 of mouse P5C dehydrogenase are shown in
sticks representation. \( F_0 \) map, contoured at 1 \( \sigma \), is superimposed on the residues 445-552 of S352L structure. Residues 348-352 of S352L mutant are shown in sticks and colored wheat.

Figure 4.15 – Michaelis menten kinetics of native human P5CDH at different concentration of NAD\(^+\). \( \bigotimes \) - 106.5 \( \mu \)M NAD\(^+\), \( \bigcirc \) - 177.5 \( \mu \)M NAD\(^+\), \( \bigtriangleup \) - 355 \( \mu \)M NAD\(^+\), \( \Box \) - 532.5 \( \mu \)M NAD\(^+\). Kinetic parameters were obtained from global fitting at all the four concentration of NAD\(^+\).

Figure 4.16 – Michaelis menten kinetics of human P5CDH S352A mutant at different concentration of NAD\(^+\). \( \bigtriangleup \) - 648 \( \mu \)M NAD\(^+\), \( \bigcirc \) - 324 \( \mu \)M NAD\(^+\), \( \Box \) - 135 \( \mu \)M NAD\(^+\). Kinetic parameters were obtained from global fitting at all the three concentration of NAD\(^+\).

Figure 4.17 – Michaelis menten kinetics of human P5CDH V470I mutant at different concentration of NAD\(^+\). \( \bigtriangleup \) - 648 \( \mu \)M NAD\(^+\), \( \bigcirc \) - 324 \( \mu \)M NAD\(^+\), \( \Box \) - 135 \( \mu \)M NAD\(^+\). Kinetic parameters were obtained from global fitting at all the three concentration of NAD\(^+\).

Figure 4.18 – Raw ITC data of human P5C dehydrogenase S352L mutant. Each peak corresponds to the injection of 20 \( \mu \)l of 200 \( \mu \)M NAD\(^+\) in 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 1 mM THP, 10 % glycerol, pH 7.85 into 17.6 \( \mu \)M of S352L in the same buffer.

Figure 4.19 – Raw ITC data of human P5C dehydrogenase. Each peak corresponds to the injection of 20 \( \mu \)l of 200 mM NAD\(^+\) in 50 \( \mu \)M Tris, 50 mM NaCl, 0.5 mM EDTA, 1 mM THP, 5 % glycerol, pH 8.0 into 17.6 \( \mu \)M of native human P5C dehydrogenase in the same buffer.
Figure 4.20 – Binding isotherm generated by integrating each peak with respect to time after correcting for the dilution. Cumulative heat (■) is plotted on y-axis with molar ratio on x-axis. Red line is the non-linear least square fitting of the data for single site binding.

Figure 4.21 – Representative raw ITC data of human P5C dehydrogenase S352A mutant. Each peak corresponds to the injection of 20 µl of 200 µM NAD⁺ in 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 1 mM THP, 5 % glycerol, pH 8.0 into 17.6 µM of native human P5C dehydrogenase in the same buffer.

Figure 4.22 – Binding isotherm generated by integrating each peak with respect to time after correcting for the dilution. Cumulative heats of four different titration experiments are plotted on y-axis with concentration of NAD on x-axis. Global fitting of data was done by fitting four parameters simultaneously; association constant, enthalpy, stoichiometry, and baseline.

Figure 4.23 – Schematics of coupled assay designed to study the oxidation of 3-OH-P5C by P5C dehydrogenase.

Figure 4.24 – NADH formation by human P5C dehydrogenase at different concentration of proline and hydroxy-proline. □- 150 mM Proline, ⊙ - 300 mM Proline, △ - 450 mM Proline, ◊ - 150 mM OH-Proline, ○ - 300 mM OH-proline, △ - 450 mM OH-proline, × - 150 mM OH-Proline without human P5C dehydrogenase, + - 150 mM Proline without human P5C dehydrogenase. λₑₓ=340 nm, λₑₘ=460 nm, excitation and emission slit width was 10 nm.

Figure 4.25 – NADH standard curve. NADH fluorescence was measured at different concentration of NADH using λₑₓ=340 nm, λₑₘ=460 nm, excitation and emission slit width= 10 nm.
Figure 4.26 – Equilibrium analytical ultracentrifugation data for human P5C dehydrogenase showing global fit to an ideal single fit model.

Figure 4.27 – Equilibrium analytical ultracentrifugation data for mouse P5C dehydrogenase showing global fit to an ideal single fit model.

Figure 4.28 – Equilibrium analytical ultracentrifugation data for Put2 showing global fit to an ideal single fit model.

Figure 4.29 – Structural alignment of different aldehyde dehydrogenase showing conservation of Ser352 (a) and Glu447 (b).
Figure 5: Protomer structure of human P5C dehydrogenase

Figure 4.2: Structure of human P5C dehydrogenase dimer.
Figure 4.3: structural alignment of human and mouse P5C dehydrogenase.

Figure 4.4: structural alignment of mouse and Thermus thermophilus P5C dehydrogenase.
Figure 4.5: Protomer structure of human P5C dehydrogenase with dimer interface residues shown in red.

Figure 4.6: Structure of Mouse P5C dehydrogenase active site showing its direct interactions with sulphate ion.
Figure 4.7: Structure of Mouse P5C dehydrogenase active site showing its interactions with bound product glutamate.

Figure 4.8: Structural alignment of human P5C dehydrogenase in monoclinic and hexagonal crystal form.
Figure 4.9: Structure of mouse P5C dehydrogenase complexed with NAD$^+$. 

Figure 4.10: Structure of mouse P5C dehydrogenase complexed with NADH.
Figure 4.11: Structure of human P5C dehydrogenase dimer showing single nucleotide polymorphism.
Figure 4.12: Structure of Mouse P5C dehydrogenase showing hydrogen-bonding network connecting Ser352 and Cys348.
Figure 4.13: Structural alignment of human P5C dehydrogenase (hexagonal form) with S352L mutant.

Figure 4.14: Stereo representation showing interaction of Glu447 with nicotinic ribose group and possible steric clash in S352L mutant.
Figure 4.15: Michaelis menten kinetics of native human P5CDH at different concentration of NAD$^+$. 

Figure 4.16: Michaelis menten kinetics of human P5CDH S352A mutant at different concentration of NAD$^+$. 
Figure 4.17: Michaelis menten kinetics of human P5CDH V470I mutant at different concentration of NAD$^+$

Figure 4.18: Raw ITC data of human P5C dehydrogenase S352L mutant.
Figure 4.19: Raw ITC data of human P5C dehydrogenase.

Figure 4.20: Binding isotherm for wild type human P5CDH generated by integrating each peak with respect to time after correcting for the dilution.
Figure 4.21: Representative raw ITC data of human P5C dehydrogenase S352A mutant.

Figure 4.22: Binding isotherm for S352A mutant generated by integrating each peak with respect to time after correcting for the dilution.
Figure 4-23: Schematics of coupled assay designed to study the oxidation of 3-OH-P5C by P5C dehydrogenase.

Figure 4-6: NADH formation by human P5C dehydrogenase at different concentration of proline and hydroxy-proline.
Figure 4-7: NADH standard curve.

Figure 4-8: Equilibrium analytical ultracentrifugation data for human P5C dehydrogenase
Figure 4-9: Equilibrium analytical ultracentrifugation data for mouse P5C dehydrogenase

Figure 4-10: Equilibrium analytical ultracentrifugation data for Put2.
Figure 4-11: Structural alignment of different aldehyde dehydrogenase
(a) -

(b) -
Dhiraj Srivastava was born and grew up in Gorakhpur, India. He has done his undergraduate from D.D.U. Gorakhpur University and his M.S. in Biotechnology from Jawaharlal Nehru University, India. He was accepted to university of Missouri-Columbia Ph.D. program in Chemistry. He joined Jack Tanner lab and studied proline catabolic enzymes by crystallographic methods. He completed his Ph.D. in June 2011. He will start his postdoctoral position at University of California, San Diego in July 2011 where he will be studying NMDA receptor and neuroscience using crystallography, cryo-electron microscopy and cell biology techniques.