

# STRUCTURAL AND KINETIC STUDIES OF PROLINE METABOLIC ENZYME INHIBITORS AND INACTIVATORS

---

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

ALEXANDRA N. BOGNER

Dr. John J. Tanner, Dissertation Supervisor

December 2021

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

STRUCTURAL AND KINETIC STUDIES OF PROLINE  
METABOLIC ENZYME INHIBITORS AND INACTIVATORS

presented by Alexandra N. Bogner

a candidate for the degree of Doctor of Philosophy,

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

---

Dr. John J. Tanner

---

Dr. Lesa J Beamer

---

Dr. Michael S. Chapman

---

Dr. Kent S. Gates

## ACKNOWLEDGEMENTS

This research was conducted under the supervision of Dr. John J. Tanner. I thank him for instilling in me a strong sense of scientific integrity, cultivating my scientific curiosity, and developing my scientific independence. I can't imagine spending my time at Mizzou in any other lab.

I am thankful for my colleagues in the Tanner Lab, past and present. The previous members, David Korasick, Li-Kai Liu, Adrian Laciak, Emily Christensen, Gabby Tam, Ashley Campbell, Jesse Wyatt, and Kyle Stiers helped set precedents as good scientists and lab mates. The current members: Kaylen Meeks, Cole McKay, David Buckley, and Juan Ji, together we have grown to be more than colleagues and I appreciate the time we've shared.

My parents played a huge part in my success in graduate school. Over the years, my father made it abundantly clear just how proud of me he is and never doubted my ability to complete this program. My mother played a part by learning enough biochemistry to follow along in my research. She has always pushed me to explore my love of science, starting with the high school research program Project SEED at Calvin University.

Additionally, the chemistry and biochemistry professors at Calvin University deserve recognition. Without their support and guidance, I may not have fully recognized my love of research. Dr. Baker gave me an incredible foundation of scientific research knowledge that I am eternally grateful for, while Drs. Arnoys, Benson, and Tatko showed me the joys of scientific collaboration.

Finally, I would like to thank Chris Farnam for his role in my success. He has been my rock, providing laughter, friendship, and stability. You've stood by me while I

dedicated myself to my passion and I am lucky to be with someone that wishes for my success even more than I do.



## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF SCHEMES .....	vii
LIST OF FIGURES.....	viii
LIST OF TABLES .....	x
LIST OF ABBREVIATIONS .....	xi
ABSTRACT .....	xiii
 <b>Chapter 1 – Structure, Biochemistry, and Gene Expression Patterns of the Proline Biosynthetic Enzyme Pyrroline-5-Carboxylate Reductase (PYCR), An Emerging Cancer Therapy Target.....</b>	 <b>1</b>
ABSTRACT .....	2
THE ENZYMES OF PROLINE METABOLISM.....	2
THE PROLINE CYCLE .....	4
PYCR ISOFORMS.....	5
THE THREE-DIMENSIONAL STRUCTURE OF PYCR .....	7
CATALYTIC PROPERTIES OF PYCR .....	10
INHIBITOR DEVELOPMENT TARGETING PYCRs .....	12
DIFFERENTIAL GENE EXPRESSION OF <i>PYCR</i> GENES IN CANCER.... .....	13
SPLICE VARIANTS.....	19
OUTLOOK AND CHALLENGES .....	21
REFERENCES .....	25
SUPPORTING REFERENCES .....	34
SUPPORTING INFORMATION .....	51

<b>Chapter 2 – Structure-Affinity Relationships of Reversible Proline Analog Inhibitors Targeting Proline Dehydrogenase .....</b>	<b>56</b>
ABSTRACT .....	57
INTRODUCTION .....	58
MATERIALS AND METHODS .....	60
RESULTS AND DISCUSSION.....	63
CONCLUSIONS .....	67
ACKNOWLEDGEMENTS .....	69
REFERENCES .....	71

<b>Chapter 3 – Structural Basis for the Stereospecific Inhibition of the Dual Proline/Hydroxyproline Catabolic Enzyme ALDH4A1 by Trans-4-Hydroxy-L-Proline .....</b>	<b>86</b>
ABSTRACT .....	88
INTRODUCTION .....	89
RESULTS .....	91
DISCUSSION.....	95
MATERIALS AND METHODS .....	98
ACKNOWLEDGEMENTS .....	101
REFERENCES .....	102
SUPPORTING INFORMATION .....	114

<b>Chapter 4 – Preliminary Exploration of the Reaction Scope of Photoinduced Covalent Inactivation of Proline Dehydrogenase .....</b>	<b>117</b>
ABSTRACT .....	118
INTRODUCTION .....	119
MATERIALS AND METHODS .....	121

RESULTS AND DISCUSSION.....	126
CONCLUSIONS .....	130
FUTURE DIRECTIONS.....	131
ACKNOWLEDGEMENTS .....	133
REFERENCES .....	135
 <b>Vita</b> .....	 149

## LIST OF SCHEMES

### **Chapter 1**

Scheme 1.1 – <b>Two recently-discovered inhibitors of PYCR1</b> .....	42
---	----

## LIST OF FIGURES

### Chapter 1

Figure 1.1 – The enzymes and reactions of proline metabolism.....	43
Figure 1.2 – Overview of the proline cycle .....	44
Figure 1.3 – Sequence alignment of PYCR1, PYCR2, and PYCR3 isoforms .....	45
Figure 1.4 – Fold and oligomeric structure of PYCR1.....	46
Figure 1.5 – Comparison of the structures of PYCR1 and PYCR2 .....	47
Figure 1.6 – Three-dimensional structural representation of nonconservative amino acid variations in the active site of PYCR .....	48
Figure 1.7 – Differential gene expression across 28 cancer types for <i>PYCR</i> isoforms....	49
Figure 1.8 – Differential gene expression of <i>PYCR</i> splice variants in various cancer types.....	50
Supplemental Figure 1.1 – Amino acid sequence alignment of PYCR proteins generated by alternative splicing .....	55

### Chapter 2

Figure 2.1 – Enzymes and reactions of proline catabolism.....	78
Figure 2.2 – Chemical structures of the inhibitors used in this study .....	79
Figure 2.3 – Initial screening of compounds for inhibition of PRODH activity .....	80
Figure 2.4 – Inhibition of the PRODH activity of PutA86-630 by 1, 2, 3, 4, 16, 17, and 18.....	81
Figure 2.5 – Structure of PutA86-630 complexed with 2 .....	82
Figure 2.6 – Interactions and electron density for inhibitors bound to PutA86-630..	83
Figure 2.7 – Superposition of the PutA86-630-inhibitor complexes .....	84
Figure 2.8 – Chemical double mutant cycle expressing the relationship between hydrogen bonding capacity and ring size of 1 .....	85

### Chapter 3

Figure 3.1 – Enzymatic reactions and inhibitors .....	109
---	-----

Figure 3.2 – Inhibition of MmALDH4A1 activity by prolines and hydroxyprolines .....	110
Figure 3.3 – Inhibition of the GSALDH activity of SmPutA by prolines and hydroxyprolines .....	111
Figure 3.4 – Structures of the MmALDH4A1 active site inhibited by THLP and THDP .....	112
Figure 3.5 – Remote binding site of THDP .....	113
Supplemental Figure 3.1 – Double-reciprocal analysis of the THDP inhibition data.....	115
Supplemental Figure 3.2 – Analysis of SmPutA GSALDH inhibition data using the competitive inhibition model.....	116

## Chapter 4

Figure 4.1 – The proline cycle and chemical structures of compounds studied.....	140
Figure 4.2 – Proposed mechanism of photoinduced inactivation by 1.....	141
Figure 4.3 – Inhibition of the PRODH activity of PutA86-630 by 1 in the absence of light .....	142
Figure 4.4 – Comparison of SmPutA and SmPutA571 PRODH activity inhibition by 1 .....	143
Figure 4.5 – Kinetics of photoinduced covalent inactivation of PutA86-630 by 1 ..	144
Figure 4.6 – Demonstration of the reversibility of photoinduced flavin spectral changes .....	145
Figure 4.7 – Observation of the bleaching of SmPutA crystals after 90 min of blue light exposure .....	146
Figure 4.8 – Interactions of 92 noncovalently bound in the SmPutA active site.....	147
Figure 4.9 – Electron density evidence for covalent modification of the N5 atom of FAD .....	148

## LIST OF TABLES

### Chapter 1

Table 1.1 – Kinetic constants of PYCRs .....	39
Table 1.2 – Inhibition constants for PYCR inhibitors .....	41
Supplemental Table 1.1 – Summary of literature on the involvement of PYCRs in cancer .....	51

### Chapter 2

Table 2.1 – Inhibition and kinetic constants .....	76
Table 2.2 – X-ray diffraction data collection and refinement statistics.....	77

### Chapter 3

Table 3.1 – Inhibition and Kinetic Constants of ALDH4A1 .....	106
Table 3.2 – Inhibition Constants of SmPutA GSALDH from Various Models .....	107
Table 3.3 – Diffraction Data Collection and Refinement Statistics .....	108
Supplemental Table 3.1 – Inhibition and kinetic constants of ALDH4A1 from various models.....	114

### Chapter 4

Table 4.1 – Noncovalent inhibition constants compared across PRODH enzymes..	138
Table 4.2 – X-ray diffraction data collection and refinement statistics.....	139

## LIST OF ABBREVIATIONS

3-OH-P5C, $\Delta^1$ -pyrroline-3-hydroxy-5-carboxylic acid
ALDH, aldehyde dehydrogenase
C2C, tetrahydrothiophene-2-carboxylate
CHDP, cis-4-hydroxy-D-proline
CHLP, cis-4-hydroxy-L-proline
D2C, 1,3-dithiolane-2-carboxylate
DLBC, Lymphoid neoplasm diffuse large B-cell lymphoma
PutA86-630, PutA protein containing residues 86-630 of <i>Escherichia coli</i>
GSAL, L-glutamate- $\gamma$ -semialdehyde
GSALDH, L-glutamate- $\gamma$ -semialdehyde dehydrogenase
LAML, Acute myeloid leukemia
Log2FC, Log2(Fold Change)
MmALDH4A1, mouse ALDH4A1
MST, microscale thermophoresis
NPPG, <i>N</i> -propargylglycine
<i>o</i> -AB, <i>ortho</i> -aminobenzaldehyde
OAT, Ornithine- $\delta$ -aminotransferase
OH-GSAL, 4-hydroxy-L-glutamate- $\gamma$ -semialdehyde
PRODH, proline dehydrogenase
P5C, $\Delta^1$ -pyrroline-5-carboxylate
P5CS, $\Delta^1$ -pyrroline-5-carboxylate synthase
P6C, $\Delta^1$ -pyrroline-6-carboxylate
PH, hyperoxaluria
PutA, proline utilization A
PYCR, Human $\Delta^1$ -pyrroline-5-carboxylate reductase
SmPutA, proline utilization A from <i>Sinorhizobium meliloti</i>
SmPutA571, PutA protein containing residues 1-571 of <i>Sinorhizobium meliloti</i>
T2C, thiazolidine-2-carboxylate
T4C, L-thiazolidine-4-carboxylate



THDP, trans-4-hydroxy-D-proline

THFA, L-tetrahydro-2-furoic acid

THLP, trans-4-hydroxy-L-proline

THYM, Thymoma

UCS, Uterine carcinosarcoma

## ABSTRACT

Structural and kinetic studies of the proline metabolic enzymes proline dehydrogenase (PRODH),  $\Delta^1$ -pyrroline-5-carboxylate reductase (PYCR), and aldehyde dehydrogenase 4A1 (ALDH4A1) are presented in this thesis. Identifying inhibitors and inactivators of these enzymes can be purposed for chemical probes in studies elucidating the role of proline metabolism in cancer and other diseases. Chapter 1 introduces the enzymes of proline metabolism and provides a summary of literature on the involvement of PYCR isoforms in cancer. Additionally, a compilation of RNA transcript data was used to explore differential gene expression of *PYCRs* across 28 cancer types. Chapter 2 details a structural and enzyme kinetic investigation of compounds that noncovalently bind within the PRODH active site. These data provide insight into structure-affinity relationships of the best-known PRODH inhibitor, L-tetrahydro-2-furoic acid. Chapter 3 focuses on the inhibition of the proline/hydroxyproline catabolic enzyme ALDH4A1 by the stereoisomers of proline and 4-hydroxyproline. We found that hydroxyproline catabolism is subject to substrate inhibition by trans-4-hydroxy-L-proline, analogous to the known inhibition of proline catabolism by L-proline. Finally, following on our recent discovery of photoinduced covalent inactivation of PRODH by 1,3-dithiolane carboxylate, Chapter 4 preliminarily explores the scope of this reaction using seven acyclic carboxylic acids that contain an S atom bonded to the  $\alpha$ -carbon, a feature we hypothesize is important for inactivation. Crystallographic evidence for three of these compounds covalently and irreversibly inactivating the flavin of PRODH was obtained. Also, a structure of one compound noncovalently bound in the active site was determined, which represents the

initial step of the inactivation mechanism. Four compounds showed signs of reversibility, which is a new feature of this class of inactivator and remains to be explored.

# Chapter 1

## Structure, Biochemistry, and Gene Expression Patterns of the Proline Biosynthetic Enzyme Pyrroline-5-Carboxylate Reductase (PYCR), An Emerging Cancer Therapy Target

Alexandra N. Bogner<sup>1</sup>, Kyle M. Stiers<sup>1</sup>, and John J. Tanner<sup>1,2,\*</sup>

<sup>1</sup>Department of Biochemistry, University of Missouri, Columbia, Missouri 65211, United States

<sup>2</sup>Department of Chemistry, University of Missouri, Columbia, Missouri 65211, United States

**Author Contributions:** All authors wrote the paper.

**Keywords:** Proline metabolism, Pyrroline-5-carboxylate reductase, PYCR, Proline biosynthesis, Cancer

**Abbreviations:** ALDH, Aldehyde dehydrogenase; DLBC, Lymphoid neoplasm diffuse large B-cell lymphoma; GSAL, L-glutamate- $\gamma$ -semialdehyde; GSALDH, L-glutamate- $\gamma$ -semialdehyde dehydrogenase; LAML, Acute myeloid leukemia; Log<sub>2</sub>FC, Log<sub>2</sub>(Fold Change); OAT, Ornithine- $\delta$ -aminotransferase; P5C,  $\Delta^1$ -pyrroline-5-carboxylate; P6C,  $\Delta^1$ -pyrroline-6-carboxylate; P5CS,  $\Delta^1$ -pyrroline-5-carboxylate synthase; PRODH, Proline dehydrogenase; PYCR, Human  $\Delta^1$ -pyrroline-5-carboxylate reductase; T4C, L-thiazolidine-4-carboxylate; THFA, (S)-(-)-tetrahydro-2-furoic acid; THYM, Thymoma; UCS, Uterine carcinosarcoma.

This chapter was published in *Amino Acids*, Structure, Biochemistry, and Gene Expression Patterns of the Proline Biosynthetic Enzyme Pyrroline-5-Carboxylate Reductase (PYCR), an Emerging Cancer Therapy Target. Bogner, AN, Stiers, KM & Tanner, JJ, **2021**, <https://doi.org/10.1007/s00726-021-02999-5>

## ABSTRACT

Proline metabolism features prominently in the unique metabolism of cancer cells. Proline biosynthetic genes are consistently upregulated in multiple cancers, while the proline catabolic enzyme proline dehydrogenase has dual, context-dependent pro-cancer and pro-apoptotic functions. Furthermore, the cycling of proline and  $\Delta^1$ -pyrroline-5-carboxylate through the proline cycle impacts cellular growth and death pathways by maintaining redox homeostasis between the cytosol and mitochondria. Here we focus on the last enzyme of proline biosynthesis,  $\Delta^1$ -pyrroline-5-carboxylate reductase, known as PYCR in humans. PYCR catalyzes the NAD(P)H-dependent reduction of  $\Delta^1$ -pyrroline-5-carboxylate to proline and forms the reductive half of the proline metabolic cycle. We review the research on the three-dimensional structure, biochemistry, inhibition, and cancer biology of PYCR. To provide a global view of *PYCR* gene upregulation in cancer, we mined RNA transcript databases to analyze differential gene expression in 28 cancer types. This analysis revealed strong, widespread upregulation of *PYCR* genes, especially *PYCR1*. Altogether, the research over the past 20 years makes a compelling case for PYCR as a cancer therapy target. We conclude with a discussion of some of the major challenges for the field, including developing isoform-specific inhibitors, elucidating the function of the long C-terminus of PYCR1/2, and characterizing the interactome of PYCR.

## THE ENZYMES OF PROLINE METABOLISM

Proline metabolism is central to the metabolic shift that occurs in cancer cells, and certain enzymes of proline metabolism have emerged as potential cancer therapy targets (Tanner et al. 2018; Phang 2019; D'Aniello et al. 2020; Burke et al. 2020; Bergers and

Fendt 2021). The subject of this review article, the enzyme  $\Delta^1$ -pyrroline-5-carboxylate reductase, known in humans as PYCR, catalyzes the final step of proline biosynthesis. PYCR isoforms are consistently upregulated in many cancers, and recent studies have shown that lowering PYCR activity by genetic knockdown or chemical inhibition diminishes the growth of cancer cells. The accumulating evidence supports the investigation of the tractability of PYCR as a cancer therapy target, hence the time seems right to summarize our knowledge of the structure and biochemistry of PYCR in the context of cancer research.

Proline metabolism refers to the enzyme-catalyzed interconversion of L-proline and L-glutamate (Figure 1.1). The biosynthetic and catabolic arms have distinct enzymes but share the common intermediate  $\Delta^1$ -pyrroline-5-carboxylate (P5C) and its hydrolysis product, L-glutamate- $\gamma$ -semialdehyde (GSAL).

Proline biosynthesis starts from either L-glutamate or L-ornithine (Figure 1.1A). L-glutamate may be derived (via the enzyme glutaminase) from L-glutamine, a well-known nutrient used by cancer cells to increase proliferation as well as survival under metabolic stress conditions (Choi and Park 2018; Phang et al. 2015; Kuo et al. 2021; Lieu et al. 2020). The glutamate route to proline begins with the formation of GSAL via the intermediate  $\gamma$ -glutamate phosphate. This transformation is done in two distinct enzymatic steps catalyzed by ATP-dependent glutamate-5-kinase and NADPH-dependent  $\gamma$ -glutamate phosphate reductase. In later evolved eukaryotes, these two enzymatic activities are combined in the bifunctional enzyme P5C synthase (P5CS) (Perez-Arellano et al. 2010).  $\gamma$ -glutamate phosphate is highly labile, and it is likely channeled between the kinase and reductase active sites of P5CS, although this has not been studied. GSAL is also a labile intermediate,

and at physiological pH, it cyclizes with the loss of water to form P5C (Bearne and Wolfenden 1995). The ornithine route to proline also yields GSAL/P5C via the enzyme ornithine- $\delta$ -aminotransferase (OAT), thus linking proline metabolism and arginine metabolism. The glutamate and ornithine pathways converge on PYCR, which catalyzes the reduction of P5C to proline using NAD(P)H as the hydride donor.

Proline catabolism consists of two enzymes (Figure 1.1B). Proline dehydrogenase (PRODH) is a flavoenzyme that catalyzes the oxidation of proline to P5C. The enzyme-bound flavin (FAD in human PRODH, FMN in some bacterial PRODHs (Huijbers et al. 2017)) is reduced during this reaction, and the electrons stored in the flavin are transferred to the electron transport chain in the membrane to regenerate the oxidized flavin for another round of catalysis. In eukaryotes, PRODH is an inner mitochondrial membrane protein, whereas bacterial PRODH associates with the membrane in a peripheral manner. The structural basis of PRODH's interactions with membranes is not well understood. The second enzyme of proline catabolism is GSALDH (a.k.a. P5C dehydrogenase, P5CDH). GSALDH belongs to the aldehyde dehydrogenase (ALDH) structural superfamily and is known as ALDH4; human GSALDH is encoded by the *ALDH4A1* gene. GSALDH catalyzes the NAD<sup>+</sup>-dependent oxidation of GSAL to glutamate. In some bacteria, PRODH and GSALDH are combined into the bifunctional enzyme, proline utilization A (Liu et al. 2017).

## THE PROLINE CYCLE

PYCR is essential to the cycling of proline and P5C to maintain redox homeostasis between the cytosol and mitochondria, known as the proline cycle (Figure 1.2) (Phang

1985, 2019). The catabolic half-cycle is the oxidation of proline to P5C catalyzed by PRODH in mitochondria. The synthetic half-cycle is the reduction of P5C to proline catalyzed by PYCR isoforms both inside mitochondria (PYCR1/2) and in the cytosol (PYCR3). It is proposed that the cycle transfers reducing equivalents from cytosolic NAD(P)H into the mitochondrial respiratory chain, although the redox shuttling mechanism remains to be clarified. By impacting cytosolic NADP<sup>+</sup> levels, the cycle links proline metabolism to the pentose phosphate pathway and nucleotide biosynthesis (Liu et al. 2015; Phang et al. 2012; Phang 2019; Hagedorn and Phang 1986). The proline cycle has been shown to enhance oxidative phosphorylation, maintain cytosolic pyridine nucleotide levels, and generate reactive oxygen species leading to activation of numerous cell-signaling pathways (Elia et al. 2017; Liu et al. 2015; Phang 2019; Hagedorn and Phang 1986; Miller et al. 2009). The central role of the proline cycle in cell survival, proliferation, and metastasis has motivated investigations of PRODH and PYCR as potential cancer therapy targets (Tanner et al. 2018).

## **PYCR ISOFORMS**

The human genome encodes three isoforms of human PYCR: PYCR1, PYCR2, and PYCR3 (a.k.a. PYCRL) (Figure 1.3). PYCR1 is localized to the mitochondria and has 319 amino acids (UniProt P32322, chromosome 17q25.3, gene symbol *PYCR1*). PYCR2 (UniProt Q96C36, 1q42.12, gene symbol *PYCR2*) has 320 amino acids, is 85% identical to PYCR1, and is also mitochondrial. PYCR3 (UniProt Q53H96, 8q24.3, gene symbol *PYCR3*) is quite different from the other two isoforms. It has only 274 amino acids, shares just 45% identity with PYCR1/2, and is cytosolic. The shorter polypeptide length of



PYCR3 is due primarily to a shortening of the C-terminus by ~50 residues (Figure 1.3). PYCR3 functions primarily in the ornithine pathway of proline biosynthesis (De Ingeniis et al. 2012).

The N-terminal Met of all three PYCR isoforms is most likely co-translationally cleaved by the enzyme methionine aminopeptidase. This modification is suggested by the small size of the residue adjacent to the N-Met (Wingfield 2017), which is Ser in PYCR1/2 and Ala in PYCR3 (Figure 1.3).

Although PYCR1 and PYCR2 are mitochondrial proteins, the identities of their mitochondrion-targeting signals within the polypeptide chain are uncertain. Analysis of the PYCR1 amino acid sequence with TPpred2.0 (Savojardo et al. 2014), MitoProt II-V1.101 (Claros and Vincens 1996), and TargetP-2.0 (Almagro Armenteros et al. 2019) revealed no mitochondrion-targeting signal. Analysis of PYCR2 suggested N-terminal mitochondrion-targeting peptides with possible cleavage sites at residues 26, 48, or 98. It is unlikely that the suggested targeting peptide is cleaved, because this would remove a large section of the NAD(P)H-binding domain. Although the precise locations of the targeting signals are unknown, presumably PYCR1 and PYCR2 are imported into mitochondria in non-native conformations, and once inside the mitochondrial matrix, they fold into their native conformations and assemble into the appropriate quaternary structures (dimer and decamer, as discussed below).

Certain mutations in the genes encoding proline metabolic enzymes cause inherited metabolic diseases. Mutations in the *PYCR1* gene (MIM 179035) have been identified in patients with autosomal recessive cutis laxa disorders (Mohamed et al. 2011; Kariminejad et al. 2017; Reversade et al. 2009). Mutations in the *PYCR2* gene (OMIM 616420) are

associated with hypomyelinating leukodystrophy 10 (Escande-Beillard et al. 2020; Zaki et al. 2016; Patel et al. 2021b). To our knowledge, there are no reports of inherited metabolic diseases associated with mutations in the *PYCR3* gene.

## **THE THREE-DIMENSIONAL STRUCTURE OF PYCR**

The three-dimensional structure of PYCR1 has been extensively characterized by X-ray crystallography and solution biophysics. Early low resolution crystal structures (3.1 Å) provided the first information about the fold of PYCR1, but the quality of the electron density for ligands was insufficient to locate the active site (Meng et al. 2006). A decade later, high resolution structures (1.85 - 1.90 Å) of PYCR1 complexed with active site ligands confirmed the fold and furthermore provided unequivocal identification of the active site and the basis of cofactor and substrate recognition (Christensen et al. 2017). A key to obtaining high resolution crystals was the use of a truncated PYCR1 construct lacking the C-terminal 19 residues, a region of the protein that is disordered in the earlier PYCR1 structures. Apparently, removing these residues improves the crystallizability of PYCR1.

PYCR1 has a two-domain fold consisting of a Rossmann dinucleotide binding domain (residues 1-162) followed by an  $\alpha$ -helical domain (Figure 1.4A). The former domain is named for the late Michael Rossmann, an eminent structural biologist who was among the first to appreciate evolutionary structural conservation (Rossmann et al. 1974; Wu and Arnold 2019). The Rossmann domain binds the cofactor NAD(P)H. The  $\alpha$ -helical domain mediates oligomerization, both dimerization (Figure 1.4C) and the assembly of dimers into a pentamer-of-dimers decamer (Figure 1.4D). The  $\alpha$  domain also plays a key role in binding

P5C. The substrate P5C binds in a loop connecting the K and L  $\alpha$ -helices of the  $\alpha$  domain, and at first glance, P5C and NADPH appear to be very far apart from each other (Figure 1.4A). However, the active site is located in the dimer interface, where the P5C-binding loop of one protomer meets the NAD(P)H-binding site of the other protomer in the dimer (Figure 1.4B). Thus, dimerization is essential for catalysis by PYCR1.

PYCR1 exists in solution mainly as a pentamer-of-dimers decamer (Figure 1.4D) when assayed by sedimentation velocity analytical ultracentrifugation performed at the relatively high enzyme concentration of 6 mg/mL (Christensen et al. 2017). Smaller molecular weight species are observable at lower enzyme concentrations, consistent with a dynamic self-association equilibrium. The crystalline protein is decameric, as expected for the solid state, where the protein concentration is very high. We note that crystal structures of bacterial and plant homologs of PYCR1 are also decameric *in crystallo*, suggesting the pentamer-of-dimers decamer is a broadly conserved structural feature of P5C reductases (Ruszkowski et al. 2015; Nocek et al. 2005).

For PYCR2, a low resolution crystal structure (3.4 Å) of the apo enzyme has been determined and used to understand a novel missense p.Gly249Val mutation identified in a patient with hypomyelinating leukodystrophy-10 (Escande-Beillard et al. 2020). The folds of PYCR1 and PYCR2 are identical, as expected for two proteins with 85% amino acid sequence identity (Figure 1.5A). The amino acid sequence differences between PYCR1 and PYCR2 are mostly located outside of the active site and are largely solvent exposed (Figure 1.5B). The pentamer-of-dimers decamer is also observed in the crystal structure of PYCR2.

The active sites of PYCR1 and PYCR2 are 97% identical in sequence. For this

calculation, we defined active site residues as those within 4.0 Å of NADPH or the P5C/proline analog (S)-(-)-tetrahydro-2-furoic acid (THFA) in the PYCR1 ternary complex structure (PDB ID 5UAV). Given this definition, the active site consists of 32 residues, which are indicated by blue shading in the sequence alignment (Figure 1.3). The active site sequences of PYCR1 and PYCR2 differ only in one position, which is a conservative substitution: Asp36 of PYCR1 is Glu in PYCR2. This residue is located in the loop that binds the 2'-phosphoryl of NADPH. The high active site sequence identity suggests that the modes of cofactor and substrate binding to PYCR2 likely resemble PYCR1, but this should be confirmed with high resolution crystallography.

Less is known about the structure of PYCR3, as no crystal structures of the enzyme have been reported. With 45% global amino acid sequence identity to PYCR1, there is little doubt that PYCR3 has the same fold as PYCR1/2. The active site of PYCR3 is 53% identical and 72% similar to PYCR1. The PYCR3 active site contains nine residues that are not conserved in PYCR1/2 (Figure 1.6). The differences appear mainly in the NADPH binding site. In particular, the loop that binds the 2'-phosphoryl of NADPH has two variations, S34A and D36T. Also note this loop is two residues shorter in PYCR3 (Figure 1.3). These structural differences may contribute to the pronounced cofactor-dependence of the  $K_m$  for P5C (Table 1.1). Also, the Q10R and N230D sequence differences are close together in space and could result in an ion pair between Arg and Asp in the active site of PYCR3.

Another interesting feature of the PYCR3 sequence is that it lacks the  $\alpha$ -helix-breaking proline found near the active sites of PYCR1/2 (replaced with Val). Pro178 of PYCR1 introduces a kink in an otherwise helical section of polypeptide chain, and the kink forms

part of the roof of the P5C binding site (star in Figure 1.4A) (Christensen et al. 2017). This variation could result in differences in the shape of the P5C pocket.

Aside from the amino acid sequence differences in the active site, PYCR3 differs from PYCR1/2 by having a shorter C-terminus (Figure 1.3). PYCR3 lacks a ~50-residue section of polypeptide at the C-terminus compared to PYCR1/2. Interestingly, these additional PYCR1/2-specific residues are disordered in the crystal structures, so the conformation of this region of the polypeptide chain is unknown. The residues immediately N-terminal to the disordered region form a ~20-residue  $\alpha$ -helix in PYCR1/2 ( $\alpha$ M), which is involved in oligomerization (Figure 1.4). The role of the flexible C-terminus for the function of PYCR1/2, and the meaning of its absence in PYCR3, are both unknown.

## CATALYTIC PROPERTIES OF PYCR

PYCR catalyzes the NAD(P)H-dependent reduction of P5C to proline. The structure of the ternary complex of PYCR1 with NADPH and the P5C/proline analog THFA provided insight into the catalytic mechanism (Christensen et al. 2017). The ring of THFA stacks in parallel with the nicotinamide such that the C5 of THFA, which represents the hydride acceptor atom of P5C, is 3.7 Å from the C4 of the nicotinamide (Figure 1.4B). This arrangement is consistent with a direct hydride-transfer mechanism, as expected for an oxidoreductase that utilizes an obligate 2-electron cofactor.

Patel *et al.* recently studied the kinetic mechanism of PYCR2 (Patel et al. 2021a). Steady-state kinetic data suggest a sequential binding mechanism with L-P5C binding the enzyme first, followed by NAD(P)H to form a ternary complex. After the hydride transfer step, NAD(P)<sup>+</sup> dissociates before proline.

The kinetic constants of recombinant PYCRs have been measured in several studies (Table 1.1). Caution is recommended when comparing the values from different studies because the experimental details can vary with respect to the concentration of the fixed substrate, the temperature, and whether the full-length enzyme or truncated version was used.

Most of the enzymology of PYCRs has focused on the reductase reaction, using P5C as the substrate and either NADPH or NADH as the cofactor. These include studies of all three PYCR isoforms as well as a few site-directed mutant variants of PYCR1 and PYCR2. One study also investigated whether PYCR1 could play a role in lysine catabolism by catalyzing the NADH-dependent reduction of  $\Delta^1$ -pyrroline-6-carboxylate (P6C) to pipecolate (Struys et al. 2014). Not only is PYCR1 active with P6C, but the  $K_m$  for P6C is within the range found for P5C (Table 1.1). A challenge with assaying the reductase reaction is that P5C is not commercially available and must be synthesized. This has motivated investigations of the reverse reaction, in which substrates such as L-thiazolidine-4-carboxylate (T4C), L-proline, or 3,4-dehydro-L-proline are oxidized, and NAD(P)<sup>+</sup> is reduced (Meng et al. 2006; Escande-Beillard et al. 2020; Nocek et al. 2005).

Some general trends are evident in the kinetic data for the reductase reaction (i.e., P5C as the substrate). None of the isoforms stand out as being significantly more active than the others. The average apparent catalytic efficiencies ( $k_{cat}/K_m$ ) from studies performed in different labs using NADH as the cofactor are within the same order of magnitude:  $9 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$  for PYCR1,  $4 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$  for PYCR2, and  $3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$  for PYCR3. PYCRs exhibit robust catalytic activity with either NADH or NADPH as the cofactor. PYCR1 and PYCR2 tend to have higher catalytic efficiency when using NADH as the cofactor. For

PYCR1, this result reflects both a lower  $K_m$  for NADH versus NADPH, and a lower  $K_m$  for P5C when NADH is used. For PYCR2 and PYCR3 the higher efficiency with NADH is due more to  $k_{cat}$ .

Product inhibition of PYCR has been investigated (Table 2.1). L-proline inhibits PYCR1 with an inhibition constant ( $K_i$ ) in the range of 1-2 mM. For reference, the  $K_m$  for P5C when NADH is the cofactor is in the range of 0.1-0.4 mM. L-proline is a stronger inhibitor of PYCR2 (~0.1 mM) and a much weaker inhibitor of PYCR3 (8 mM).  $NAD^+$  inhibits PYCR2 with  $K_i$  of 0.8 mM, compared to a  $K_m$  for NADH of 0.3 mM.

## INHIBITOR DEVELOPMENT TARGETING PYCRs

Inhibitor discovery targeting PYCR is in an early stage. Milne et al. recently reported the results of a small-scale screening study that identified pargyline and as an inhibitor of PYCR1 (Milne et al. 2019). Subsequent optimization yielded pargyline derivatives with  $IC_{50}$  values against the purified enzyme of 9-400  $\mu M$ , such as **1** in Scheme 1.1. The kinetic mechanism by which **1** inhibits PYCR1 was not determined. Cell-based studies showed that **1** lowered intracellular proline levels in SUM159PT human breast cancer cells and inhibited proliferation of MDA-MB-321 and SUM159PT breast cancer cell lines.

Our group recently used a focused target-specific *in crystallo* screening approach to discover inhibitors of PYCR1 (Christensen et al. 2020). Focused screening (a.k.a. knowledge-based) involves selecting subsets of molecules from chemical libraries that are likely to have activity based on prior knowledge of the target protein and chemical classes that have activity at the target (Deng et al. 2006; Orry et al. 2006; Harris et al. 2011; Hughes et al. 2011). This approach leveraged the prior knowledge of ligand recognition learned

from high resolution crystal structures of PYCR1 (Christensen et al. 2017). The initial screen was performed in crystals rather than using activity assays. Crystals of apo PYCR1 were soaked with 27 commercially-available proline analogs, and the electron density in the known P5C site was used as a proxy for inhibition. Compounds that bound the enzyme *in crystallo* were studied further using enzyme assays to determine the mechanism of inhibition. One advantage of the *in crystallo* approach over the traditional inhibitor screening pipeline is that problems validating hit compounds are minimized, since the initial assay selects for compounds that bind specifically to the target.

The focused target-specific *in crystallo* screening campaign uncovered N-formyl-L-proline (**2** in Scheme 1.1) as a promising lead inhibitor of PYCR1 (Christensen et al. 2020). Compound **2** is a competitive inhibitor (competitive with P5C) with an inhibition constant of 100  $\mu$ M. The mechanism of inhibition includes an unexpected cascade of protein conformational changes emanating from the P5C site to an oligomer interface, which are needed to expand the active site to accommodate the formyl group of **2**. Furthermore, **2** was shown to phenocopy the PYCR1 knockdown in MCF10A H-RASV<sup>12</sup> breast cancer cells. To our knowledge, **2** is the only inhibitor of PYCR1 that has been thoroughly validated by demonstrating the kinetic mechanism of action against the purified enzyme, the mode of binding to the enzyme by X-ray crystallography, and activity in cancer cells.

## DIFFERENTIAL GENE EXPRESSION OF *PYCR* GENES IN CANCER

PYCR isoforms are consistently upregulated in many cancers. To provide a global view of this phenomenon, we utilized Gene Expression Profiling Interactive Analysis (GEPIA) **2** (<http://gepia2.cancer-pku.cn/#index>) to analyze differential gene expression across a



variety of cancers (Tang et al. 2019). The server provides RNA sequencing expression data of 9736 tumors and 8587 normal samples from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) projects. We compiled a list of *PYCR* gene expression in 28 cancer types using a  $\log_2(\text{Fold Change})$  ( $\log_2\text{FC}$ ) cutoff of 0.1 (equivalent to FC of 1.07), a q-value cutoff of 0.01, and the ANOVA differential method.

The clear trend in the data is that *PYCR* gene expression tends to increase in cancer cells, especially for *PYCR1* (Figure 1.7A). *PYCR1* expression increased by  $\log_2\text{FC} > 1.0$  in 22 of the 28 cancers (79%, Figure 1.7B). Among these, sixteen cancers are notable in showing large increases in *PYCR1* expression of  $\log_2\text{FC} > 2.0$ , i.e., greater than 4-fold higher. The largest increase in *PYCR1* expression occurs in uterine carcinosarcoma (UCS) ( $\log_2\text{FC} = 4.0$ ). *PYCR1* expression decreases only in acute myeloid leukemia (LAML), which has one of the largest changes in expression.

*PYCR2* is also upregulated in cancer cells, but to a lesser extent than *PYCR1* (Figure 1.7). *PYCR2* is upregulated by  $\log_2\text{FC} > 1.0$  in six cancers (21%, Figure 1.7B). The largest increases occur in thymoma (THYM) and lymphoid neoplasm diffuse large B-cell lymphoma (DLBC).

The differential expression of *PYCR3* is between those of *PYCR1* and *PYCR2* (Figure 1.7). *PYCR3* expression increases by  $\log_2\text{FC} > 1.0$  in thirteen cancers (46%, Figure 1.7B). The largest increase is in DLBC. Two cancers (7%) show a slight downregulation of the *PYCR3* gene.

We also examined correlations among the expression patterns of the three *PYCR* genes. For example, in DLBC and THYM, all three genes are substantially increased by  $\log_2\text{FC}$  of 2.0 - 3.0. Conversely, there are no cancers that show downregulation of all three

isoforms, although both *PYCR1* and *PYCR2* are substantially decreased in LAML. Considering pairwise relationships, scatter plots of the expression data suggest positive associations between all three pairs of *PYCR* genes (Figure 1.7C).

Consistent with our analysis of the transcript data, numerous studies have reported increased *PYCR1* expression in cancer (summarized in SI Table S1.1). To our knowledge, the first such report was a 2002 study using prostate cancer tissue samples (Ernst et al. 2002). Upregulation of *PYCR1* in prostate cancer was confirmed in subsequent studies (Jariwala et al. 2007; Zeng et al. 2017), and knockdown of the *PYCR1* gene was shown to inhibit prostate cancer cell growth (Zeng et al. 2017). An open mechanistic question is whether *PYCR1* mediates the action of androgen receptor signaling in prostate cancer (Jariwala et al. 2007; Zeng et al. 2017).

Phang's group has examined the involvement of PYCRs in tumor growth, stressing the importance of the proline cycle (Liu et al. 2015). They showed that the oncoprotein MYC induces proline biosynthesis from glutamine in P493 human B lymphoma cells by increasing the expression of all three PYCRs and P5CS. The knockdown of any proline biosynthetic gene markedly decreased growth of several cancer cell lines, apparently by decreasing ATP production. Paradoxically, the growth inhibitory effects of proline biosynthesis knockdown persisted even when proline was added to the cultures, suggesting that proline itself was not of primary importance. Instead, the authors concluded that proline biosynthesis contributes to tumor metabolic reprogramming through the cycling of proline and P5C (Figure 1.2).

Proline metabolism is central to the metabolic alterations that occur in breast cancer. A 2011 study using negative selection RNAi-screening identified PYCR1 among metabolic

genes associated with aggressive breast cancer and stemness (Possemato et al. 2011). Recent work confirmed the upregulation of *PYCR1* gene expression and PYCR1 protein in breast cancer cells and showed that the knockdown of *PYCR1* reduced the invasion and migration capabilities of breast cancer cell lines and increased drug sensitivity of orthotopically injected ER-positive tumors *in vivo* (Ding et al. 2017; Craze et al. 2018; Shenoy et al. 2020). Also, higher *PYCR1* expression was found to be correlated with tumor size, higher risk of tumor relapse, and poorer patient survival (Ding et al. 2017; Shenoy et al. 2020).

The landmark 2017 study from Fendt's group established the central role of the proline cycle in breast cancer metastasis (Elia et al. 2017). They found that the knockdown of either *PRODH* or *PYCR1* impaired spheroidal growth of MCF10A H-RAS<sup>V12</sup> breast cancer cells. Further, the inhibition of PRODH enzyme by the proline analog THFA impaired spheroidal cell growth and inhibited metastases in a mouse model of breast cancer. A subsequent study showed that targeting PYCR1 with a small molecule (**2** in Scheme 1.1) induced proline accumulation and impaired proline biosynthesis as well as spheroidal growth in MCF10A H-RAS<sup>V12</sup> breast cancer cells (Christensen et al. 2020). Together, these studies showed the importance of the proline cycle in breast cancer and provided clear support for investigating PRODH and PYCR1 as breast cancer therapy targets.

Several studies have also found increased expression of *PYCR1* in lung cancer (Cai et al. 2018; Guo et al. 2019; Sang et al. 2019; Wang et al. 2019; Gao et al. 2020; Lu et al. 2021). Some of these studies also found that increased *PYCR1* expression correlated with poor patient outcomes (Cai et al. 2018; Wang et al. 2019; Gao et al. 2020). In one study, the microRNA miR-328-3p was shown to target *PYCR1*, and the level of miR-328-3p was

decreased in lung adenocarcinoma cells, leading to increased *PYCR1* expression (Lu et al. 2021). Knockdown of *PYCR1* inhibited lung cancer cell proliferation and tumorigenesis, and increased cell apoptosis (Cai et al. 2018; Guo et al. 2019; Sang et al. 2019; Wang et al. 2019; Gao et al. 2020). Two recent studies have focused on the interaction of PYCR1 with Kindlin-2 in lung adenocarcinoma. They found that increased stiffening of the extracellular matrix, which is associated with the pathological progression of cancer (Lampi and Reinhart-King 2018), promotes translocation of Kindlin-2 into mitochondria where it interacts with PYCR1 (Guo et al. 2019). Apparently, this protein-protein interaction did not affect the catalytic activity of PYCR1, as measured *in vitro* by monitoring the reverse reaction at a single substrate (3,4-dehydro-L-proline) concentration using a glutathione *S*-transferase fusion of Kindlin-2. Instead, it was suggested that the interaction with Kindlin-2 enhances proline biosynthesis by preventing degradation of PYCR1. A subsequent study by the same group has linked the enhancement of proline biosynthesis by the PYCR1-Kindlin-2 interaction with PINCH-1-mediated mitochondrial dynamics (Guo et al. 2020).

*PYCR1* involvement in colorectal cancer has been investigated. One study used peptide mass fingerprinting and found increased PYCR1 protein in an *in vitro* model of the colorectal adenoma-to-carcinoma progression (Roth et al. 2010). More recently, upregulation of PYCR1 was confirmed in colorectal cancer tissues and cells, and knockdown of *PYCR1* inhibited the proliferation, drug resistance, and epithelial-mesenchymal transition of colorectal cancer (Yan et al. 2019). This study also implicated the interaction of PYCR1 enzyme with STAT3 as part of the signaling mechanism in colorectal cancer cells.

Ribosome profiling has been used to identify metabolic vulnerabilities in tumors that

could be leveraged in cancer therapy; this approach applied to kidney cancer discovered a vulnerability for proline linked to high levels of *PYCR1* (Loayza-Puch et al. 2016). A subsequent study of renal cell carcinoma also found increased expression of *PYCR1* and connected *PYCR1* levels to increased metastasis and poorer overall patient survival (Weijin et al. 2019).

The involvement of PYCR in liver cancers has received attention recently (see the review in this issue, (Ding et al. 2021)). *PYCR2* was identified as a prognostic biomarker in hepatitis B virus-related hepatocellular carcinoma (Gao et al. 2019). *PYCR1* expression is also increased in hepatocellular carcinoma, and the knockdown of *PYCR1* reduced cell proliferation of multiple hepatocellular carcinoma cell lines *in vitro* and tumor growth *in vivo* (Zhuang et al. 2019; Ding et al. 2020). Also, the combination of high *PRODH* expression, and low expression of both *PYCR1* and *P5CS* (a.k.a. *ALDH18A1*) was associated with better patient outcomes (Ding et al. 2020).

Numerous other studies have found upregulation of PYCRs in various cancers, which lends support to the idea from global analysis of transcript data (Figure 1.7) that proline biosynthesis is broadly important in the unique metabolism of cancer cells. *PYCR3* is upregulated in taxol-resistant nasopharyngeal carcinoma cell lines (Li et al. 2015). *PYCR1* expression is increased in human malignant melanoma cell lines and is an indicator of poor prognosis (Ye et al. 2018). Knockdown of *PYCR1* or *PYCR2* decreased proliferation of melanoma cells via increased apoptosis (Ye et al. 2018) or AMPK/mTOR-induced autophagy (Ou et al. 2016). Recently, the correlated upregulation of the Rho-family GTPase Rac3 and *PYCR1* has been found in bladder cancer (Cheng et al. 2020), and a genomics study uncovered *PYCR1* as a prognostic factor in bladder cancer patients (Liu et

al. 2021). Enhanced PYCR1 activity has been found in glioma cells carrying mutations in isocitrate dehydrogenase 1 (Hollinshead et al. 2018). *In situ* metabolomics has been used to identify proline as an enhanced metabolite in tissues from esophageal cancer patients, and then immunohistochemistry staining revealed increased PYCR2 protein in regions of high proline (Sun et al. 2019). Finally, in samples from gastric cancer patients, *PYCR1* expression was found to be upregulated and correlated with advanced tumor stage and aggressive histological type (Xiao et al. 2020). The same study found that the knockdown of *PYCR1* inhibited proliferation of gastric cancer cells by promoting apoptosis (Xiao et al. 2020).

Finally, in addition to the aforementioned cancer type-specific research, *PYCR* has been identified as a consistently upregulated gene in pan-cancer analyses, which look for genes that are differentially expressed in multiple cancers. These include an analysis of messenger RNA profiles of 1454 metabolic enzymes in 19 cancer types and a study of transcriptomic changes in samples representing ten major cancer types (Nilsson et al. 2014; Haider et al. 2016). Another study found an association between *PYCR1* expression and poor prognosis (Darzi et al. 2021).

## **SPLICE VARIANTS**

Alternative splicing generates multiple protein isoforms from a single gene, and several splice variants of *PYCR* genes are known. Excluding the splice variants retained as introns, we found twelve variants of *PYCR1* that are expressed as proteins, four for *PYCR2*, and three for *PYCR3* (Table 1.3). Most of the PYCR proteins generated by alternative splicing are unlikely to be active enzymes. For example, the eight PYCR1 proteins predicted to

have fewer than 290 amino acids have severe truncations that eliminate parts of the active site and/or dimer interface (see amino acid sequence alignment in SI Figure S1.1). Similarly, the three shorter variants of PYCR2 (< 320 amino acids) and the 254-residue form of PYCR3 are probably not active.

We analyzed the differential expression of *PYCR* splice variants in cancer with GEPIA2 (cutoffs of q-value < 0.01 and log2FC > 0.1). The heatmap in Figure 1.8 shows the differential expression data for the *PYCR* splice variants. Note that some data are lacking or did not meet the threshold for statistical significance (denoted by a zero in the heatmap). The expression trends for the splice variants (Figure 1.8) track those of the three *PYCR* genes (Figure 1.7), which is expected since there is overlap between the data sets used in the two analyses. Generally, the data suggest that *PYCR1* is the most consistently upregulated in cancer, followed by *PYCR3*, and then *PYCR2*.

The phenomenon of increased *PYCR* expression is broadly observed across multiple cancer types and splice variants (Figure 1.8). Twenty-two of the 28 cancer types (79%) have at least one *PYCR* variant with expression increased by  $\log_2\text{FC} \geq 1.2$ , and 16/19 splice variants (84%) are increased by  $\log_2\text{FC} \geq 1.2$  in at least one cancer type. An exception to the trend of increased *PYCR* expression in cancer is LAML, where several *PYCR* splice variants are significantly downregulated (Figure 1.8, bottom row).

As noted above, only a subset of the splice variants are predicted to encode active enzymes. It is of interest to examine the expression of these variants separately from the others, because they may provide insight into the relevance of PYCR catalytic activity to cancer metabolism. Both of the splice variants encoding the canonical PYCR1 enzyme show  $\log_2\text{FC} > 1.2$  in more than half of the 28 cancer types (GEPIA IDs PYCR1-001,

PYCR1-201). Similarly, for *PYCR2* and *PYCR3*, the variants encoding known active enzymes experience the most consistent upregulation compared to the other variants, although we note the expression data are sparse for many of the presumed inactive variants (Figure 1.8). The observation that transcripts encoding active enzymes are upregulated in cancer cells compared to normal cells is consistent with the catalytic activity of PYCRs playing a role in the metabolic transition to cancer.

Interestingly, a few splice variants predicted to encode inactive proteins are also substantially upregulated. For example, PYCR1-005, which encodes a truncated 288-residue version of the canonical enzyme, is increased by  $\log_2\text{FC} > 1.2$  in 14/28 cancers. This protein lacks a 32-residue section of the  $\alpha$  domain in the dimer interface. Because dimerization is essential for catalytic activity, this protein is probably not active. The significance of increased expression of inactive PYCR proteins in cancer remains to be determined. One possibility is that such proteins, although lacking catalytic activity, could still participate in protein-protein interactions. Another possibility is that they regulate catalytic activity by forming mixed oligomers with active PYCRs.

## OUTLOOK AND CHALLENGES

The cancer biology literature from the past 20 years makes a compelling case for PYCR, especially PYCR1, as a novel cancer drug target. *PYCR1* gene expression is consistently higher in cancerous tissue compared to normal tissue, the knockdown of *PYCR1* impairs cancer cell proliferation both *in vitro* and in animal models, and *PYCR1* gene expression is predictive of undesirable tumor characteristics and poor patient outcomes. These key results appear in dozens of independent studies and have been



observed with many different cancer types. Our analysis of *PYCR* transcript data is not only consistent with this body of research but also suggests that upregulation of *PYCR* is a highly conserved aspect of cancer cell metabolism (Figures 1.7 and 1.8).

Potent, specific inhibitors of PYCRs are needed as chemical probes to investigate the roles of PYCRs in cancer and as lead compounds for drug discovery. Currently, only one validated chemical probe of PYCR1 is known (**2** in Scheme 1.1), but its potency of 100  $\mu$ M against the purified enzyme is modest, and its isoform specificity has not been determined. The discovery of **2** from a very small library of 27 compounds showed proof-of-concept for screening proline analogs as inhibitors of PYCR1. It will be interesting to see the extent of inhibition that can be achieved from a more extensive screening of proline analogs versus compounds that access the NADPH site. High-throughput screening of large, diverse chemical libraries should also be given a high priority.

Developing isoform-specific PYCR inhibitors is a challenge for the field. As described here, the active sites of PYCR1 and PYCR2 are virtually identical in structure, which suggests that compounds directed at the active site are unlikely to discriminate between these two isoforms. A larger question is whether specificity between PYCR1 and PYCR2 is necessary for cancer therapeutic applications. The prospects for identifying PYCR3-specific inhibitors are somewhat brighter. Our analysis predicts that PYCR3 may have structural features that distinguish it from PYCR1/2, including ten non-conserved active site residues, a shorter 2'-phosphoryl-binding loop, and a substantially shorter C-terminus (by ~50 residues). A crystal structure of PYCR3 is needed to determine whether these sequence differences result in structural differences that can be leveraged for inhibitor design. It will be important to test new inhibitors against all three PYCR isoforms.

The function of the long C-termini of PYCR1 and PYCR2 is a fundamental unanswered question. The C-terminus of PYCR1/2 contains ~50 extra residues not found in PYCR3 or P5C reductases from plants and bacteria (Figure 1.3). Unfortunately, this region is disordered in crystal structures of PYCR1/2. The C-terminus is not essential for catalytic activity, since recombinant PYCR1 lacking the C-terminal 19 residues is catalytically active (Christensen et al. 2017). Nevertheless, a detailed examination of the role of the C-terminus on the catalytic properties of PYCR1 would be informative. The structures of PYCR1/2 indicate that the flexible C-terminus could access the active site, suggesting the possibility that it is a regulatory element. For example, the last resolved residue of helix  $\alpha$ M in the PYCR1 structure is only ~20 Å from the active site, and the missing stretch of ~45 residues of polypeptide chain is certainly long enough to reach into the active site. Also, helix  $\alpha$ M is in the oligomer interface, so it is possible that the flexible C-terminus mediates inter-subunit communication that could be important for regulating catalytic activity. A third possibility is that the C-terminus is involved in protein-protein interactions. Clearly, more biochemical, cellular, and structural work is needed to understand the function of the flexible C-terminus of PYCR1/2 and the impact of its absence in PYCR3.

Protein-protein interactions of PYCRs is a promising area of research where there is much to learn. Recent work has uncovered possible interacting partners of PYCR1/2, including ORAOV1 in the context of esophageal cancer (Togashi et al. 2014), ribonucleotide reductase small subunit B (Kuo et al. 2016), Kindlin-2 in lung adenocarcinoma (Guo et al. 2019; Guo et al. 2020), the SIRT3 mitochondrial deacetylase (Chen et al. 2019), STAT3 in colorectal cancer (Yan et al. 2019), Lon chaperone (Kuo et

al. 2020), Kaposi's sarcoma-associated herpesvirus K1 oncoprotein (Choi et al. 2020), and several enzymes involved in mitochondrial glycine, glutamate, and fatty acid metabolism (Escande-Beillard et al. 2020). These potential interactors of PYCR1/2 were discovered using cell-based methods, co-immunoprecipitation, and pull-downs. It will be important to validate these protein-protein interactions using biophysical experiments on the purified proteins as part of a broader investigation of the role of PYCR protein-protein interactions in cancer (Mackay et al. 2007).

In summary, PYCR has emerged over the past 20 years as a novel cancer therapy target. The major challenge for the next decade is to determine the tractability of PYCR as a drug target and translate basic science advances on proline metabolism into therapeutic reality.

## **ASSOCIATED CONTENT**

### **Supporting Information**

Table S1.1 and Figure S1.1.

## REFERENCES

1. Aken BL, Achuthan P, Akanni W, Amode MR, Bernsdorff F, Bhai J, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Gil L, Giron CG, Gordon L, Hourlier T, Hunt SE, Janacek SH, Juettemann T, Keenan S, Laird MR, Lavidas I, Maurel T, McLaren W, Moore B, Murphy DN, Nag R, Newman V, Nuhn M, Ong CK, Parker A, Patricio M, Riat HS, Sheppard D, Sparrow H, Taylor K, Thormann A, Vullo A, Walts B, Wilder SP, Zadissa A, Kostadima M, Martin FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Cunningham F, Yates A, Zerbino DR, Flicek P (2017) Ensembl 2017. *Nucleic Acids Res* 45 (D1):D635-D642. doi:10.1093/nar/gkw1104
2. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, Nielsen H (2019) Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance* 2 (5). doi:10.26508/lsa.201900429
3. Bearne SL, Wolfenden R (1995) Glutamate gamma-semialdehyde as a natural transition state analogue inhibitor of Escherichia coli glucosamine-6-phosphate synthase. *Biochemistry* 34 (36):11515-11520
4. Bergers G, Fendt SM (2021) The metabolism of cancer cells during metastasis. *Nat Rev Cancer* 21 (3):162-180. doi:10.1038/s41568-020-00320-2
5. Burke L, Guterman I, Palacios Gallego R, Britton RG, Burschowsky D, Tufarelli C, Rufini A (2020) The Janus-like role of proline metabolism in cancer. *Cell Death Discov* 6:104. doi:10.1038/s41420-020-00341-8
6. Cai F, Miao Y, Liu C, Wu T, Shen S, Su X, Shi Y (2018) Pyrroline-5-carboxylate reductase 1 promotes proliferation and inhibits apoptosis in non-small cell lung cancer. *Oncol Lett* 15 (1):731-740. doi:10.3892/ol.2017.7400
7. Chen S, Yang X, Yu M, Wang Z, Liu B, Liu M, Liu L, Ren M, Qi H, Zou J, Vucenik I, Zhu WG, Luo J (2019) SIRT3 regulates cancer cell proliferation through deacetylation of PYCR1 in proline metabolism. *Neoplasia* 21 (7):665-675. doi:10.1016/j.neo.2019.04.008
8. Cheng C, Song D, Wu Y, Liu B (2020) RAC3 Promotes Proliferation, Migration and Invasion via PYCR1/JAK/STAT Signaling in Bladder Cancer. *Front Mol Biosci* 7:218. doi:10.3389/fmolb.2020.00218
9. Choi UY, Lee JJ, Park A, Zhu W, Lee HR, Choi YJ, Yoo JS, Yu C, Feng P, Gao SJ, Chen S, Eoh H, Jung JU (2020) Oncogenic human herpesvirus hijacks proline metabolism for tumorigenesis. *Proc Natl Acad Sci U S A* 117 (14):8083-8093. doi:10.1073/pnas.1918607117
10. Choi YK, Park KG (2018) Targeting Glutamine Metabolism for Cancer Treatment. *Biomol Ther (Seoul)* 26 (1):19-28. doi:10.4062/biomolther.2017.178

11. Christensen EM, Bogner AN, Vandekeere A, Tam GS, Patel SM, Becker DF, Fendt SM, Tanner JJ (2020) In crystallo screening for proline analog inhibitors of the proline cycle enzyme PYCR1. *J Biol Chem* 295 (52):18316-18327. doi:10.1074/jbc.RA120.016106
12. Christensen EM, Patel SM, Korasick DA, Campbell AC, Krause KL, Becker DF, Tanner JJ (2017) Resolving the cofactor-binding site in the proline biosynthetic enzyme human pyrroline-5-carboxylate reductase 1. *J Biol Chem* 292 (17):7233-7243. doi:10.1074/jbc.M117.780288
13. Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* 241 (3):779-786
14. Craze ML, Cheung H, Jewa N, Coimbra NDM, Soria D, El-Ansari R, Aleskandarany MA, Wai Cheng K, Diez-Rodriguez M, Nolan CC, Ellis IO, Rakha EA, Green AR (2018) MYC regulation of glutamine-proline regulatory axis is key in luminal B breast cancer. *Br J Cancer* 118 (2):258-265. doi:10.1038/bjc.2017.387
15. D'Aniello C, Patriarca EJ, Phang JM, Minchiotti G (2020) Proline Metabolism in Tumor Growth and Metastatic Progression. *Front Oncol* 10:776. doi:10.3389/fonc.2020.00776
16. Darzi M, Gorgin S, Majidzadeh AK, Esmaeili R (2021) Gene co-expression network analysis reveals immune cell infiltration as a favorable prognostic marker in non-uterine leiomyosarcoma. *Sci Rep* 11 (1):2339. doi:10.1038/s41598-021-81952-8
17. De Ingeniis J, Ratnikov B, Richardson AD, Scott DA, Aza-Blanc P, De SK, Kazanov M, Pellecchia M, Ronai Z, Osterman AL, Smith JW (2012) Functional specialization in proline biosynthesis of melanoma. *PLoS One* 7 (9):e45190. doi:10.1371/journal.pone.0045190
18. Deng Z, Chuaqui C, Singh J (2006) Knowledge-based design of target-focused libraries using protein-ligand interaction constraints. *J Med Chem* 49 (2):490-500. doi:10.1021/jm050381x
19. Ding J, Kuo ML, Su L, Xue L, Luh F, Zhang H, Wang J, Lin TG, Zhang K, Chu P, Zheng S, Liu X, Yen Y (2017) Human mitochondrial pyrroline-5-carboxylate reductase 1 promotes invasiveness and impacts survival in breast cancers. *Carcinogenesis* 38 (5):519-531. doi:10.1093/carcin/bgx022
20. Ding Z, Ericksen RE, Escande-Beillard N, Lee QY, Loh A, Denil S, Steckel M, Haegebarth A, Wai Ho TS, Chow P, Toh HC, Reversade B, Gruenewald S, Han W (2020) Metabolic pathway analyses identify proline biosynthesis pathway as a promoter of liver tumorigenesis. *J Hepatol* 72 (4):725-735. doi:10.1016/j.jhep.2019.10.026

21. Ding Z, Ericksen RE, Lee QY, Han W (2021) Reprogramming of mitochondrial proline metabolism promotes liver tumorigenesis. *Amino Acids*. doi:10.1007/s00726-021-02961-5
22. Elia I, Broekaert D, Christen S, Boon R, Radaelli E, Orth MF, Verfaillie C, Grunewald TGP, Fendt SM (2017) Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nature communications* 8:15267. doi:10.1038/ncomms15267
23. Ernst T, Hergenhahn M, Kenzelmann M, Cohen CD, Bonrouhi M, Weninger A, Klären R, Gröne EF, Wiesel M, Güdemann C, Küster J, Schott W, Staehler G, Kretzler M, Hollstein M, Gröne HJ (2002) Decrease and gain of gene expression are equally discriminatory markers for prostate carcinoma: a gene expression analysis on total and microdissected prostate tissue. *Am J Pathol* 160 (6):2169-2180. doi:10.1016/s0002-9440(10)61165-0
24. Escande-Beillard N, Loh A, Saleem SN, Kanata K, Hashimoto Y, Altunoglu U, Metoska A, Grandjean J, Ng FM, Pomp O, Baburajendran N, Wong J, Hill J, Beillard E, Cozzone P, Zaki M, Kayserili H, Hamada H, Shiratori H, Reversade B (2020) Loss of PYCR2 Causes Neurodegeneration by Increasing Cerebral Glycine Levels via SHMT2. *Neuron* 107 (1):82-94 e86. doi:10.1016/j.neuron.2020.03.028
25. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, Huang C, Li J, Dong X, Zhou Y, Liu Q, Ma L, Wang X, Zhou J, Liu Y, Boja E, Robles AI, Ma W, Wang P, Li Y, Ding L, Wen B, Zhang B, Rodriguez H, Gao D, Zhou H, Fan J (2019) Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma. *Cell* 179 (2):561-577.e522. doi:10.1016/j.cell.2019.08.052
26. Gao Y, Luo L, Xie Y, Zhao Y, Yao J, Liu X (2020) PYCR1 knockdown inhibits the proliferation, migration, and invasion by affecting JAK/STAT signaling pathway in lung adenocarcinoma. *Mol Carcinog* 59 (5):503-511. doi:10.1002/mc.23174
27. Guo L, Cui C, Wang J, Yuan J, Yang Q, Zhang P, Su W, Bao R, Ran J, Wu C (2020) PINCH-1 regulates mitochondrial dynamics to promote proline synthesis and tumor growth. *Nature communications* 11 (1):4913. doi:10.1038/s41467-020-18753-6
28. Guo L, Cui C, Zhang K, Wang J, Wang Y, Lu Y, Chen K, Yuan J, Xiao G, Tang B, Sun Y, Wu C (2019) Kindlin-2 links mechano-environment to proline synthesis and tumor growth. *Nature communications* 10 (1):845. doi:10.1038/s41467-019-08772-3
29. Hagedorn CH, Phang JM (1986) Catalytic Transfer of Hydride ions from NADPH to Oxygen by the Interconversion of Proline to Delta-Pyrroline-5-Carboxylate. *Arch Biochem Biophys* 248:166-174

30. Haider S, McIntyre A, van Stiphout RG, Winchester LM, Wigfield S, Harris AL, Buffa FM (2016) Genomic alterations underlie a pan-cancer metabolic shift associated with tumour hypoxia. *Genome Biol* 17 (1):140. doi:10.1186/s13059-016-0999-8
31. Harris CJ, Hill RD, Sheppard DW, Slater MJ, Stouten PF (2011) The design and application of target-focused compound libraries. *Comb Chem High Throughput Screen* 14 (6):521-531
32. Hollinshead KER, Munford H, Eales KL, Bardella C, Li C, Escribano-Gonzalez C, Thakker A, Nonnenmacher Y, Kluckova K, Jeeves M, Murren R, Cuzzo F, Ye D, Laurenti G, Zhu W, Hiller K, Hodson DJ, Hua W, Tomlinson IP, Ludwig C, Mao Y, Tennant DA (2018) Oncogenic IDH1 Mutations Promote Enhanced Proline Synthesis through PYCR1 to Support the Maintenance of Mitochondrial Redox Homeostasis. *Cell Rep* 22 (12):3107-3114. doi:10.1016/j.celrep.2018.02.084
33. Hughes JP, Rees S, Kalindjian SB, Philpott KL (2011) Principles of early drug discovery. *Br J Pharmacol* 162 (6):1239-1249. doi:10.1111/j.1476-5381.2010.01127.x
34. Huijbers MM, Martinez-Julvez M, Westphal AH, Delgado-Arciniega E, Medina M, van Berkel WJ (2017) Proline dehydrogenase from *Thermus thermophilus* does not discriminate between FAD and FMN as cofactor. *Sci Rep* 7:43880. doi:10.1038/srep43880
35. Jariwala U, Prescott J, Jia L, Barski A, Pregizer S, Cogan JP, Arasheben A, Tilley WD, Scher HI, Gerald WL, Buchanan G, Coetzee GA, Frenkel B (2007) Identification of novel androgen receptor target genes in prostate cancer. *Mol Cancer* 6:39. doi:10.1186/1476-4598-6-39
36. Kariminejad A, Afroozan F, Bozorgmehr B, Ghanadan A, Akbaroghli S, Khorram Khorshid HR, Mojahedi F, Setoodeh A, Loh A, Tan YX, Escande-Beillard N, Malfait F, Reversade B, Gardeitchik T, Morava E (2017) Discriminative Features in Three Autosomal Recessive Cutis Laxa Syndromes: Cutis Laxa IIA, Cutis Laxa IIB, and Geroderma Osteoplastica. *International journal of molecular sciences* 18 (3). doi:10.3390/ijms18030635
37. Kuo CL, Chou HY, Chiu YC, Cheng AN, Fan CC, Chang YN, Chen CH, Jiang SS, Chen NJ, Lee AY (2020) Mitochondrial oxidative stress by Lon-PYCR1 maintains an immunosuppressive tumor microenvironment that promotes cancer progression and metastasis. *Cancer Lett* 474:138-150. doi:10.1016/j.canlet.2020.01.019
38. Kuo ML, Lee MB, Tang M, den Besten W, Hu S, Sweredoski MJ, Hess S, Chou CM, Changou CA, Su M, Jia W, Su L, Yen Y (2016) PYCR1 and PYCR2 Interact and Collaborate with RRM2B to Protect Cells from Overt Oxidative Stress. *Sci Rep* 6:18846. doi:10.1038/srep18846

39. Kuo MT, Chen HHW, Feun LG, Savaraj N (2021) Targeting the Proline-Glutamine-Asparagine-Arginine Metabolic Axis in Amino Acid Starvation Cancer Therapy. *Pharmaceuticals (Basel)* 14 (1). doi:10.3390/ph14010072
40. Lampi MC, Reinhart-King CA (2018) Targeting extracellular matrix stiffness to attenuate disease: From molecular mechanisms to clinical trials. *Sci Transl Med* 10 (422). doi:10.1126/scitranslmed.aao0475
41. Li W, You Y, Zhang X, Song Y, Xiang H, Peng X, Qin J, Tan G (2015) Amplification of chromosome 8q21-qter associated with the acquired paclitaxel resistance of nasopharyngeal carcinoma cells. *Int J Clin Exp Pathol* 8 (10):12346-12356
42. Lieu EL, Nguyen T, Rhyne S, Kim J (2020) Amino acids in cancer. *Exp Mol Med* 52 (1):15-30. doi:10.1038/s12276-020-0375-3
43. Liu LK, Becker DF, Tanner JJ (2017) Structure, function, and mechanism of proline utilization A (PutA). *Arch Biochem Biophys* 632:142-157. doi:10.1016/j.abb.2017.07.005
44. Liu W, Hancock CN, Fischer JW, Harman M, Phang JM (2015) Proline biosynthesis augments tumor cell growth and aerobic glycolysis: involvement of pyridine nucleotides. *Sci Rep* 5:17206. doi:10.1038/srep17206
45. Liu Z, Sun T, Zhang Z, Bi J, Kong C (2021) An 18-gene signature based on glucose metabolism and DNA methylation improves prognostic prediction for urinary bladder cancer. *Genomics* 113 (1 Pt 2):896-907. doi:10.1016/j.ygeno.2020.10.022
46. Loayza-Puch F, Rooijers K, Buil LC, Zijlstra J, Oude Vrielink JF, Lopes R, Ugalde AP, van Breugel P, Hofland I, Wesseling J, van Tellingen O, Bex A, Agami R (2016) Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. *Nature* 530 (7591):490-494. doi:10.1038/nature16982
47. Lu J, Lin J, Zhou Y, Ye K, Fang C (2021) MiR-328-3p inhibits lung adenocarcinoma-genesis by downregulation PYCR1. *Biochemical and biophysical research communications* 550:99-106. doi:10.1016/j.bbrc.2021.02.029
48. Mackay JP, Sunde M, Lowry JA, Crossley M, Matthews JM (2007) Protein interactions: is seeing believing? *Trends Biochem Sci* 32 (12):530-531. doi:10.1016/j.tibs.2007.09.006
49. Meng Z, Lou Z, Liu Z, Li M, Zhao X, Bartlam M, Rao Z (2006) Crystal structure of human pyrroline-5-carboxylate reductase. *J Mol Biol* 359 (5):1364-1377
50. Miller G, Honig A, Stein H, Suzuki N, Mittler R, Zilberstein A (2009) Unraveling delta1-pyrroline-5-carboxylate-proline cycle in plants by uncoupled expression of proline oxidation enzymes. *J Biol Chem* 284 (39):26482-26492. doi:10.1074/jbc.M109.009340



51. Milne K, Sun J, Zaal EA, Mowat J, Celie PHN, Fish A, Berkers CR, Forlani G, Loayza-Puch F, Jamieson C, Agami R (2019) A fragment-like approach to PYCR1 inhibition. *Bioorganic & medicinal chemistry letters* 29 (18):2626-2631. doi:10.1016/j.bmcl.2019.07.047
52. Mohamed M, Kouwenberg D, Gardeitchik T, Kornak U, Wevers RA, Morava E (2011) Metabolic cutis laxa syndromes. *Journal of inherited metabolic disease* 34 (4):907-916. doi:10.1007/s10545-011-9305-9
53. Nilsson R, Jain M, Madhusudhan N, Sheppard NG, Strittmatter L, Kampf C, Huang J, Asplund A, Mootha VK (2014) Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nature communications* 5:3128. doi:10.1038/ncomms4128
54. Nocek B, Chang C, Li H, Lezondra L, Holzle D, Collart F, Joachimiak A (2005) Crystal structures of delta1-pyrroline-5-carboxylate reductase from human pathogens *Neisseria meningitidis* and *Streptococcus pyogenes*. *J Mol Biol* 354 (1):91-106
55. Orry AJ, Abagyan RA, Cavasotto CN (2006) Structure-based development of target-specific compound libraries. *Drug Discov Today* 11 (5-6):261-266. doi:10.1016/S1359-6446(05)03717-7
56. Ou R, Zhang X, Cai J, Shao X, Lv M, Qiu W, Xuan X, Liu J, Li Z, Xu Y (2016) Downregulation of pyrroline-5-carboxylate reductase-2 induces the autophagy of melanoma cells via AMPK/mTOR pathway. *Tumour Biol* 37 (5):6485-6491. doi:10.1007/s13277-015-3927-8
57. Patel SM, Seravalli J, Liang X, Tanner JJ, Becker DF (2021a) Disease Variants of Human Delta-1-Pyrroline-5-Carboxylate Reductase 2 (PYCR2). *Arch Biochem Biophys* Accepted
58. Patel SM, Seravalli J, Liang X, Tanner JJ, Becker DF (2021b) Disease variants of human Delta(1)-pyrroline-5-carboxylate reductase 2 (PYCR2). *Arch Biochem Biophys*:108852. doi:10.1016/j.abb.2021.108852
59. Perez-Arellano I, Carmona-Alvarez F, Martinez AI, Rodriguez-Diaz J, Cervera J (2010) Pyrroline-5-carboxylate synthase and proline biosynthesis: from osmotolerance to rare metabolic disease. *Protein Sci* 19 (3):372-382. doi:10.1002/pro.340
60. Phang JM (1985) The regulatory functions of proline and pyrroline-5-carboxylic acid. *Curr Top Cell Reg* 25:92-132
61. Phang JM (2019) Proline Metabolism in Cell Regulation and Cancer Biology: Recent Advances and Hypotheses. *Antioxid Redox Signal* 30 (4):635-649. doi:10.1089/ars.2017.7350

62. Phang JM, Liu W, Hancock C, Christian KJ (2012) The proline regulatory axis and cancer. *Front Oncol* 2:60. doi:10.3389/fonc.2012.00060
63. Phang JM, Liu W, Hancock CN, Fischer JW (2015) Proline metabolism and cancer: emerging links to glutamine and collagen. *Curr Opin Clin Nutr Metab Care* 18 (1):71-77. doi:10.1097/MCO.0000000000000121
64. Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, Sethumadhavan S, Woo HK, Jang HG, Jha AK, Chen WW, Barrett FG, Stransky N, Tsun ZY, Cowley GS, Barretina J, Kalaany NY, Hsu PP, Ottina K, Chan AM, Yuan B, Garraway LA, Root DE, Mino-Kenudson M, Brachtel EF, Driggers EM, Sabatini DM (2011) Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* 476 (7360):346-350. doi:10.1038/nature10350
65. Reversade B, Escande-Beillard N, Dimopoulou A, Fischer B, Chng SC, Li Y, Shboul M, Tham PY, Kayserili H, Al-Gazali L, Shahwan M, Brancati F, Lee H, O'Connor BD, Schmidt-von Kegler M, Merriman B, Nelson SF, Masri A, Alkazaleh F, Guerra D, Ferrari P, Nanda A, Rajab A, Markie D, Gray M, Nelson J, Grix A, Sommer A, Savarirayan R, Janecke AR, Steichen E, Sillence D, Hausser I, Budde B, Nurnberg G, Nurnberg P, Seemann P, Kunkel D, Zambruno G, Dallapiccola B, Schuelke M, Robertson S, Hamamy H, Wollnik B, Van Maldergem L, Mundlos S, Kornak U (2009) Mutations in PYCR1 cause cutis laxa with progeroid features. *Nature genetics* 41 (9):1016-1021. doi:10.1038/ng.413
66. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42 (Web Server issue):W320-324. doi:10.1093/nar/gku316
67. Rossmann MG, Moras D, Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* 250 (463):194-199
68. Roth U, Razawi H, Hommer J, Engelmann K, Schwientek T, Müller S, Baldus SE, Patsos G, Corfield AP, Paraskeva C, Hanisch FG (2010) Differential expression proteomics of human colorectal cancer based on a syngeneic cellular model for the progression of adenoma to carcinoma. *Proteomics* 10 (2):194-202. doi:10.1002/pmic.200900614
69. Ruszkowski M, Nocek B, Forlani G, Dauter Z (2015) The structure of *Medicago truncatula* delta(1)-pyrroline-5-carboxylate reductase provides new insights into regulation of proline biosynthesis in plants. *Front Plant Sci* 6:869. doi:10.3389/fpls.2015.00869
70. Sang S, Zhang C, Shan J (2019) Pyrroline-5-Carboxylate Reductase 1 Accelerates the Migration and Invasion of Non-small Cell Lung Cancer In Vitro. *Cancer Biother Radiopharm* 34 (6):380-387. doi:10.1089/cbr.2019.2782

71. Savojardo C, Martelli PL, Fariselli P, Casadio R (2014) TPpred2: improving the prediction of mitochondrial targeting peptide cleavage sites by exploiting sequence motifs. *Bioinformatics* 30 (20):2973-2974. doi:10.1093/bioinformatics/btu411
72. Shenoy A, Belugali Nataraj N, Perry G, Loayza Puch F, Nagel R, Marin I, Balint N, Bossel N, Pavlovsky A, Barshack I, Kaufman B, Agami R, Yarden Y, Dadiani M, Geiger T (2020) Proteomic patterns associated with response to breast cancer neoadjuvant treatment. *Molecular systems biology* 16 (9):e9443. doi:10.15252/msb.20209443
73. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539. doi:10.1038/msb.2011.75
74. Struys EA, Jansen EE, Salomons GS (2014) Human pyrroline-5-carboxylate reductase (PYCR1) acts on Delta(1)-piperidine-6-carboxylate generating L-pipecolic acid. *Journal of inherited metabolic disease* 37 (3):327-332. doi:10.1007/s10545-013-9673-4
75. Sun C, Li T, Song X, Huang L, Zang Q, Xu J, Bi N, Jiao G, Hao Y, Chen Y, Zhang R, Luo Z, Li X, Wang L, Wang Z, Song Y, He J, Abliz Z (2019) Spatially resolved metabolomics to discover tumor-associated metabolic alterations. *Proc Natl Acad Sci U S A* 116 (1):52-57. doi:10.1073/pnas.1808950116
76. Tang Z, Kang B, Li C, Chen T, Zhang Z (2019) GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* 47 (W1):W556-W560. doi:10.1093/nar/gkz430
77. Tanner JJ, Fendt SM, Becker DF (2018) The Proline Cycle As a Potential Cancer Therapy Target. *Biochemistry* 57 (25):3433-3444. doi:10.1021/acs.biochem.8b00215
78. Togashi Y, Arao T, Kato H, Matsumoto K, Terashima M, Hayashi H, de Velasco MA, Fujita Y, Kimura H, Yasuda T, Shiozaki H, Nishio K (2014) Frequent amplification of ORAOV1 gene in esophageal squamous cell cancer promotes an aggressive phenotype via proline metabolism and ROS production. *Oncotarget* 5 (10):2962-2973. doi:10.18632/oncotarget.1561
79. Wang D, Wang L, Zhang Y, Yan Z, Liu L, Chen G (2019) PYCR1 promotes the progression of non-small-cell lung cancer under the negative regulation of miR-488. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* 111:588-595. doi:10.1016/j.biopha.2018.12.089
80. Weijin F, Zhibin X, Shengfeng Z, Xiaoli Y, Qijian D, Jiayi L, Qiumei L, Yilong C, Hua M, Deyun L, Jiwen C (2019) The clinical significance of PYCR1 expression in renal cell carcinoma. *Medicine (Baltimore)* 98 (28):e16384. doi:10.1097/md.00000000000016384

81. Wingfield PT (2017) N-Terminal Methionine Processing. *Curr Protoc Protein Sci* 88:614 11-16 14 13. doi:10.1002/cpps.29
82. Wu H, Arnold E (2019) Michael G. Rossmann (1930-2019). *Nat Struct Mol Biol* 26 (8):660-662. doi:10.1038/s41594-019-0271-5
83. Xiao S, Li S, Yuan Z, Zhou L (2020) Pyrroline-5-carboxylate reductase 1 (PYCR1) upregulation contributes to gastric cancer progression and indicates poor survival outcome. *Ann Transl Med* 8 (15):937. doi:10.21037/atm-19-4402
84. Yan K, Xu X, Wu T, Li J, Cao G, Li Y, Ji Z (2019) Knockdown of PYCR1 inhibits proliferation, drug resistance and EMT in colorectal cancer cells by regulating STAT3-Mediated p38 MAPK and NF- $\kappa$ B signalling pathway. *Biochemical and biophysical research communications* 520 (2):486-491. doi:10.1016/j.bbrc.2019.10.059
85. Ye Y, Wu Y, Wang J (2018) Pyrroline-5-carboxylate reductase 1 promotes cell proliferation via inhibiting apoptosis in human malignant melanoma. *Cancer Manag Res* 10:6399-6407. doi:10.2147/cmar.S166711
86. Zaki MS, Bhat G, Sultan T, Issa M, Jung HJ, Dikoglu E, Selim L, I GM, Abdel-Hamid MS, Abdel-Salam G, Marin-Valencia I, Gleeson JG (2016) PYCR2 Mutations cause a lethal syndrome of microcephaly and failure to thrive. *Ann Neurol* 80 (1):59-70. doi:10.1002/ana.24678
87. Zeng T, Zhu L, Liao M, Zhuo W, Yang S, Wu W, Wang D (2017) Knockdown of PYCR1 inhibits cell proliferation and colony formation via cell cycle arrest and apoptosis in prostate cancer. *Med Oncol* 34 (2):27. doi:10.1007/s12032-016-0870-5
88. Zhuang J, Song Y, Ye Y, He S, Ma X, Zhang M, Ni J, Wang J, Xia W (2019) PYCR1 interference inhibits cell growth and survival via c-Jun N-terminal kinase/insulin receptor substrate 1 (JNK/IRS1) pathway in hepatocellular cancer. *J Transl Med* 17 (1):343. doi:10.1186/s12967-019-2091-0

## SUPPORTING REFERENCES

1. Cai F, Miao Y, Liu C, Wu T, Shen S, Su X, Shi Y (2018) Pyrroline-5-carboxylate reductase 1 promotes proliferation and inhibits apoptosis in non-small cell lung cancer. *Oncol Lett* 15 (1):731-740. doi:10.3892/ol.2017.7400
2. Chen S, Yang X, Yu M, Wang Z, Liu B, Liu M, Liu L, Ren M, Qi H, Zou J, Vucenik I, Zhu WG, Luo J (2019) SIRT3 regulates cancer cell proliferation through deacetylation of PYCR1 in proline metabolism. *Neoplasia* 21 (7):665-675. doi:10.1016/j.neo.2019.04.008
3. Cheng C, Song D, Wu Y, Liu B (2020) RAC3 Promotes Proliferation, Migration and Invasion via PYCR1/JAK/STAT Signaling in Bladder Cancer. *Front Mol Biosci* 7:218. doi:10.3389/fmolb.2020.00218
4. Choi UY, Lee JJ, Park A, Zhu W, Lee HR, Choi YJ, Yoo JS, Yu C, Feng P, Gao SJ, Chen S, Eoh H, Jung JU (2020) Oncogenic human herpesvirus hijacks proline metabolism for tumorigenesis. *Proc Natl Acad Sci U S A* 117 (14):8083-8093. doi:10.1073/pnas.1918607117
5. Christensen EM, Bogner AN, Vandekeere A, Tam GS, Patel SM, Becker DF, Fendt SM, Tanner JJ (2020) In crystallo screening for proline analog inhibitors of the proline cycle enzyme PYCR1. *J Biol Chem* 295 (52):18316-18327. doi:10.1074/jbc.RA120.016106
6. Craze ML, Cheung H, Jewa N, Coimbra NDM, Soria D, El-Ansari R, Aleskandarany MA, Wai Cheng K, Diez-Rodriguez M, Nolan CC, Ellis IO, Rakha EA, Green AR (2018) MYC regulation of glutamine-proline regulatory axis is key in luminal B breast cancer. *Br J Cancer* 118 (2):258-265. doi:10.1038/bjc.2017.387
7. Darzi M, Gorgin S, Majidzadeh AK, Esmaeili R (2021) Gene co-expression network analysis reveals immune cell infiltration as a favorable prognostic marker in non-uterine leiomyosarcoma. *Sci Rep* 11 (1):2339. doi:10.1038/s41598-021-81952-8
8. De Ingeniis J, Ratnikov B, Richardson AD, Scott DA, Aza-Blanc P, De SK, Kazanov M, Pellicchia M, Ronai Z, Osterman AL, Smith JW (2012) Functional specialization in proline biosynthesis of melanoma. *PLoS One* 7 (9):e45190. doi:10.1371/journal.pone.0045190
9. Ding J, Kuo ML, Su L, Xue L, Luh F, Zhang H, Wang J, Lin TG, Zhang K, Chu P, Zheng S, Liu X, Yen Y (2017) Human mitochondrial pyrroline-5-carboxylate reductase 1 promotes invasiveness and impacts survival in breast cancers. *Carcinogenesis* 38 (5):519-531. doi:10.1093/carcin/bgx022
10. Ding Z, Ericksen RE, Escande-Beillard N, Lee QY, Loh A, Denil S, Steckel M, Haegebarth A, Wai Ho TS, Chow P, Toh HC, Reversade B, Gruenewald S, Han W (2020)

Metabolic pathway analyses identify proline biosynthesis pathway as a promoter of liver tumorigenesis. *J Hepatol* 72 (4):725-735. doi:10.1016/j.jhep.2019.10.026

11. Elia I, Broekaert D, Christen S, Boon R, Radaelli E, Orth MF, Verfaillie C, Grunewald TGP, Fendt SM (2017) Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nature communications* 8:15267. doi:10.1038/ncomms15267

12. Ernst T, Hergenhausen M, Kenzelmann M, Cohen CD, Bonrouhi M, Weninger A, Klären R, Gröne EF, Wiesel M, Güdemann C, Küster J, Schott W, Staehler G, Kretzler M, Hollstein M, Gröne HJ (2002) Decrease and gain of gene expression are equally discriminatory markers for prostate carcinoma: a gene expression analysis on total and microdissected prostate tissue. *Am J Pathol* 160 (6):2169-2180. doi:10.1016/s0002-9440(10)61165-0

13. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, Huang C, Li J, Dong X, Zhou Y, Liu Q, Ma L, Wang X, Zhou J, Liu Y, Boja E, Robles AI, Ma W, Wang P, Li Y, Ding L, Wen B, Zhang B, Rodriguez H, Gao D, Zhou H, Fan J (2019) Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma. *Cell* 179 (2):561-577.e522. doi:10.1016/j.cell.2019.08.052

14. Gao Y, Luo L, Xie Y, Zhao Y, Yao J, Liu X (2020) PYCR1 knockdown inhibits the proliferation, migration, and invasion by affecting JAK/STAT signaling pathway in lung adenocarcinoma. *Mol Carcinog* 59 (5):503-511. doi:10.1002/mc.23174

15. Guo L, Cui C, Wang J, Yuan J, Yang Q, Zhang P, Su W, Bao R, Ran J, Wu C (2020) PINCH-1 regulates mitochondrial dynamics to promote proline synthesis and tumor growth. *Nature communications* 11 (1):4913. doi:10.1038/s41467-020-18753-6

16. Guo L, Cui C, Zhang K, Wang J, Wang Y, Lu Y, Chen K, Yuan J, Xiao G, Tang B, Sun Y, Wu C (2019) Kindlin-2 links mechano-environment to proline synthesis and tumor growth. *Nature communications* 10 (1):845. doi:10.1038/s41467-019-08772-3

17. Haider S, McIntyre A, van Stiphout RG, Winchester LM, Wigfield S, Harris AL, Buffa FM (2016) Genomic alterations underlie a pan-cancer metabolic shift associated with tumour hypoxia. *Genome Biol* 17 (1):140. doi:10.1186/s13059-016-0999-8

18. Hollinshead KER, Munford H, Eales KL, Bardella C, Li C, Escibano-Gonzalez C, Thakker A, Nonnenmacher Y, Kluckova K, Jeeves M, Murren R, Cuzzo F, Ye D, Laurenti G, Zhu W, Hiller K, Hodson DJ, Hua W, Tomlinson IP, Ludwig C, Mao Y, Tennant DA (2018) Oncogenic IDH1 Mutations Promote Enhanced Proline Synthesis through PYCR1 to Support the Maintenance of Mitochondrial Redox Homeostasis. *Cell Rep* 22 (12):3107-3114. doi:10.1016/j.celrep.2018.02.084

19. Jariwala U, Prescott J, Jia L, Barski A, Pregizer S, Cogan JP, Arasheben A, Tilley WD, Scher HI, Gerald WL, Buchanan G, Coetzee GA, Frenkel B (2007) Identification of novel

androgen receptor target genes in prostate cancer. *Mol Cancer* 6:39. doi:10.1186/1476-4598-6-39

20. Kuo CL, Chou HY, Chiu YC, Cheng AN, Fan CC, Chang YN, Chen CH, Jiang SS, Chen NJ, Lee AY (2020) Mitochondrial oxidative stress by Lon-PYCR1 maintains an immunosuppressive tumor microenvironment that promotes cancer progression and metastasis. *Cancer Lett* 474:138-150. doi:10.1016/j.canlet.2020.01.019

21. Li W, You Y, Zhang X, Song Y, Xiang H, Peng X, Qin J, Tan G (2015) Amplification of chromosome 8q21-qter associated with the acquired paclitaxel resistance of nasopharyngeal carcinoma cells. *Int J Clin Exp Pathol* 8 (10):12346-12356

22. Liu W, Hancock CN, Fischer JW, Harman M, Phang JM (2015) Proline biosynthesis augments tumor cell growth and aerobic glycolysis: involvement of pyridine nucleotides. *Sci Rep* 5:17206. doi:10.1038/srep17206

23. Liu Z, Sun T, Zhang Z, Bi J, Kong C (2021) An 18-gene signature based on glucose metabolism and DNA methylation improves prognostic prediction for urinary bladder cancer. *Genomics* 113 (1 Pt 2):896-907. doi:10.1016/j.ygeno.2020.10.022

24. Loayza-Puch F, Rooijers K, Buil LC, Zijlstra J, Oude Vrielink JF, Lopes R, Ugalde AP, van Breugel P, Hofland I, Wesseling J, van Tellingen O, Bex A, Agami R (2016) Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. *Nature* 530 (7591):490-494. doi:10.1038/nature16982

25. Lu J, Lin J, Zhou Y, Ye K, Fang C (2021) MiR-328-3p inhibits lung adenocarcinoma-genesis by downregulation PYCR1. *Biochemical and biophysical research communications* 550:99-106. doi:10.1016/j.bbrc.2021.02.029

26. Nilsson R, Jain M, Madhusudhan N, Sheppard NG, Strittmatter L, Kampf C, Huang J, Asplund A, Mootha VK (2014) Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nature communications* 5:3128. doi:10.1038/ncomms4128

27. Ou R, Zhang X, Cai J, Shao X, Lv M, Qiu W, Xuan X, Liu J, Li Z, Xu Y (2016) Downregulation of pyrroline-5-carboxylate reductase-2 induces the autophagy of melanoma cells via AMPK/mTOR pathway. *Tumour Biol* 37 (5):6485-6491. doi:10.1007/s13277-015-3927-8

28. Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, Sethumadhavan S, Woo HK, Jang HG, Jha AK, Chen WW, Barrett FG, Stransky N, Tsun ZY, Cowley GS, Barretina J, Kalaany NY, Hsu PP, Ottina K, Chan AM, Yuan B, Garraway LA, Root DE, Mino-Kenudson M, Brachtel EF, Driggers EM, Sabatini DM (2011) Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* 476 (7360):346-350. doi:10.1038/nature10350

29. Roth U, Razawi H, Hommer J, Engelmann K, Schwientek T, Müller S, Baldus SE, Patsos G, Corfield AP, Paraskeva C, Hanisch FG (2010) Differential expression proteomics of human colorectal cancer based on a syngeneic cellular model for the progression of adenoma to carcinoma. *Proteomics* 10 (2):194-202. doi:10.1002/pmic.200900614
30. Sang S, Zhang C, Shan J (2019) Pyrroline-5-Carboxylate Reductase 1 Accelerates the Migration and Invasion of Non-small Cell Lung Cancer In Vitro. *Cancer Biother Radiopharm* 34 (6):380-387. doi:10.1089/cbr.2019.2782
31. Shenoy A, Belugali Nataraj N, Perry G, Loayza Puch F, Nagel R, Marin I, Balint N, Bossel N, Pavlovsky A, Barshack I, Kaufman B, Agami R, Yarden Y, Dadiani M, Geiger T (2020) Proteomic patterns associated with response to breast cancer neoadjuvant treatment. *Molecular systems biology* 16 (9):e9443. doi:10.15252/msb.20209443
32. Sun C, Li T, Song X, Huang L, Zang Q, Xu J, Bi N, Jiao G, Hao Y, Chen Y, Zhang R, Luo Z, Li X, Wang L, Wang Z, Song Y, He J, Abliz Z (2019) Spatially resolved metabolomics to discover tumor-associated metabolic alterations. *Proc Natl Acad Sci U S A* 116 (1):52-57. doi:10.1073/pnas.1808950116
33. Togashi Y, Arao T, Kato H, Matsumoto K, Terashima M, Hayashi H, de Velasco MA, Fujita Y, Kimura H, Yasuda T, Shiozaki H, Nishio K (2014) Frequent amplification of ORAOV1 gene in esophageal squamous cell cancer promotes an aggressive phenotype via proline metabolism and ROS production. *Oncotarget* 5 (10):2962-2973. doi:10.18632/oncotarget.1561
34. Wang D, Wang L, Zhang Y, Yan Z, Liu L, Chen G (2019) PYCR1 promotes the progression of non-small-cell lung cancer under the negative regulation of miR-488. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* 111:588-595. doi:10.1016/j.biopha.2018.12.089
35. Weijin F, Zhibin X, Shengfeng Z, Xiaoli Y, Qijian D, Jiayi L, Qiumei L, Yilong C, Hua M, Deyun L, Jiwen C (2019) The clinical significance of PYCR1 expression in renal cell carcinoma. *Medicine (Baltimore)* 98 (28):e16384. doi:10.1097/md.00000000000016384
36. Xiao S, Li S, Yuan Z, Zhou L (2020) Pyrroline-5-carboxylate reductase 1 (PYCR1) upregulation contributes to gastric cancer progression and indicates poor survival outcome. *Ann Transl Med* 8 (15):937. doi:10.21037/atm-19-4402
37. Yan K, Xu X, Wu T, Li J, Cao G, Li Y, Ji Z (2019) Knockdown of PYCR1 inhibits proliferation, drug resistance and EMT in colorectal cancer cells by regulating STAT3-Mediated p38 MAPK and NF-κB signalling pathway. *Biochemical and biophysical research communications* 520 (2):486-491. doi:10.1016/j.bbrc.2019.10.059



38. Ye Y, Wu Y, Wang J (2018) Pyrroline-5-carboxylate reductase 1 promotes cell proliferation via inhibiting apoptosis in human malignant melanoma. *Cancer Manag Res* 10:6399-6407. doi:10.2147/cmar.S166711
39. Zeng T, Zhu L, Liao M, Zhuo W, Yang S, Wu W, Wang D (2017) Knockdown of PYCR1 inhibits cell proliferation and colony formation via cell cycle arrest and apoptosis in prostate cancer. *Med Oncol* 34 (2):27. doi:10.1007/s12032-016-0870-5
40. Zhuang J, Song Y, Ye Y, He S, Ma X, Zhang M, Ni J, Wang J, Xia W (2019) PYCR1 interference inhibits cell growth and survival via c-Jun N-terminal kinase/insulin receptor substrate 1 (JNK/IRS1) pathway in hepatocellular cancer. *J Transl Med* 17 (1):343. doi:10.1186/s12967-019-2091-0

**Table 1.1** Kinetic constants of PYCRs

Enzyme	Variable Substrate	Fixed Substrate	$K_m$ ( $\mu$ M)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_m$ ( $M^{-1}s^{-1}$ )
PYCR1 <sup>a</sup>	T4C	NADP <sup>+</sup>	1260	55	$4 \times 10^4$
PYCR1 <sup>a</sup>	NAD <sup>+</sup>	T4C	151	-	-
PYCR1 <sup>a</sup>	NADP <sup>+</sup>	T4C	3060	-	-
PYCR1 E221A <sup>a</sup>	T4C	NADP <sup>+</sup>	730	13	$2 \times 10^4$
PYCR1 E221A <sup>a</sup>	NAD <sup>+</sup>	T4C	235	-	-
PYCR1 E221A <sup>a</sup>	NADP <sup>+</sup>	T4C	480	-	-
PYCR1 <sup>b</sup>	P5C	NADH	1720	70	$4 \times 10^4$
	P5C	NADPH	2150	29	$1 \times 10^4$
	NADH	P5C	260	64	$2 \times 10^5$
	NADPH	P5C	1200	46	$4 \times 10^4$
PYCR1 <sup>c</sup>	P6C	NADH	146	-	-
PYCR1 <sup>d</sup>	NADH	P5C	70	218	$3 \times 10^6$
	NADPH	P5C	283	74	$3 \times 10^5$
	P5C	NADPH	667	31	$5 \times 10^4$
PYCR1 T238A <sup>d</sup>	NADPH	P5C	159	23	$1 \times 10^5$
	P5C	NADPH	2,887	14	$5 \times 10^3$
PYCR1 <sup>e</sup>	P5C	NADH	185-374	35-69	$2 \times 10^5$
PYCR2 <sup>b</sup>	P5C	NADH	1000	149	$1.5 \times 10^5$
	P5C	NADPH	1700	85	$5 \times 10^4$
	NADH	P5C	220	219	$1 \times 10^6$
	NADPH	P5C	240	93	$4 \times 10^5$
PYCR2 <sup>f</sup>	NAD <sup>+</sup>	L-proline	1110	52.2	$5 \times 10^4$
PYCR2 G249V <sup>f</sup>	NAD <sup>+</sup>	L-proline	3280	22.2	$7 \times 10^3$
PYCR2 <sup>g</sup>	P5C	NADH	1509	61.3	$4 \times 10^4$
	P5C	NADPH	994	26.0	$3 \times 10^4$
	NADH	P5C	298	47.9	$2 \times 10^5$

PYCR2 R251C <sup>g</sup>	NADPH	P5C	216	24.0	1 x 10 <sup>5</sup>
	P5C	NADH	1315	3.2	2 x 10 <sup>3</sup>
	P5C	NADPH	1499	3.1	2 x 10 <sup>3</sup>
	NADH	P5C	953	5.9	6 x 10 <sup>3</sup>
PYCR2 R119C <sup>g</sup>	NADPH	P5C	119	1.8	1 x 10 <sup>4</sup>
	P5C	NADH	317	0.08	2 x 10 <sup>2</sup>
	P5C	NADPH	334	0.13	4 x 10 <sup>2</sup>
	NADH	P5C	34	0.015	4 x 10 <sup>2</sup>
PYCR3 <sup>b</sup>	NADPH	P5C	537	1.5	3 x 10 <sup>3</sup>
	P5C	NADH	4640	197.0	4 x 10 <sup>4</sup>
	P5C	NADPH	380	35.0	9 x 10 <sup>4</sup>
	NADH	P5C	420	196.4	5 x 10 <sup>5</sup>
	NADPH	P5C	370	24.9	7 x 10 <sup>4</sup>

<sup>a</sup>(Meng et al. 2006)

<sup>b</sup>(De Ingeniis et al. 2012)

<sup>c</sup>(Struys et al. 2014)

<sup>d</sup>(Christensen et al. 2017)

<sup>e</sup>(Christensen et al. 2020)

<sup>f</sup>(Escande-Beillard et al. 2020)

<sup>g</sup>(Patel et al. 2021a)

**Table 1.2** Inhibition constants for PYCR inhibitors

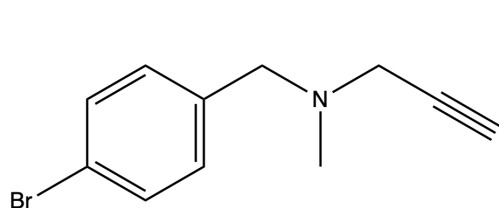
Inhibitor	Enzyme	$K_i$ ( $\mu$ M)
L-proline	PYCR1	600-1700 <sup>a,b</sup>
(S)-(-)-tetrahydro-2-furoic acid	PYCR1	2200 <sup>b</sup>
cyclopentanecarboxylate	PYCR1	1200 <sup>b</sup>
L-thiazolidine-4-carboxylate	PYCR1	600 <sup>b</sup>
L-thiazolidine-2-carboxylate	PYCR1	450 <sup>b</sup>
N-formyl-L-proline	PYCR1	100 <sup>b</sup>
L-proline	PYCR2	96 - 145 <sup>a,c</sup>
NAD <sup>+</sup>	PYCR2	800 <sup>c</sup>
L-proline	PYCR3	8500 <sup>a</sup>

<sup>a</sup>P5C was the variable substrate and NADH was fixed (De Ingeniis et al. 2012)

<sup>b</sup>P5C was the variable substrate, NADH was fixed, and truncated PYCR1 was used (Christensen et al. 2020)

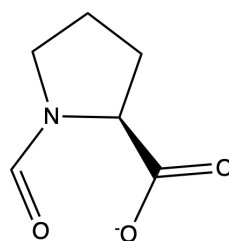
<sup>c</sup>P5C was the variable substrate and NADH was fixed (Patel et al. 2021a)

<sup>d</sup>NADH was the variable substrate and P5C was fixed (Patel et al. 2021a)



**1**

$IC_{50} = 9 \mu M$   
(mechanism unknown)

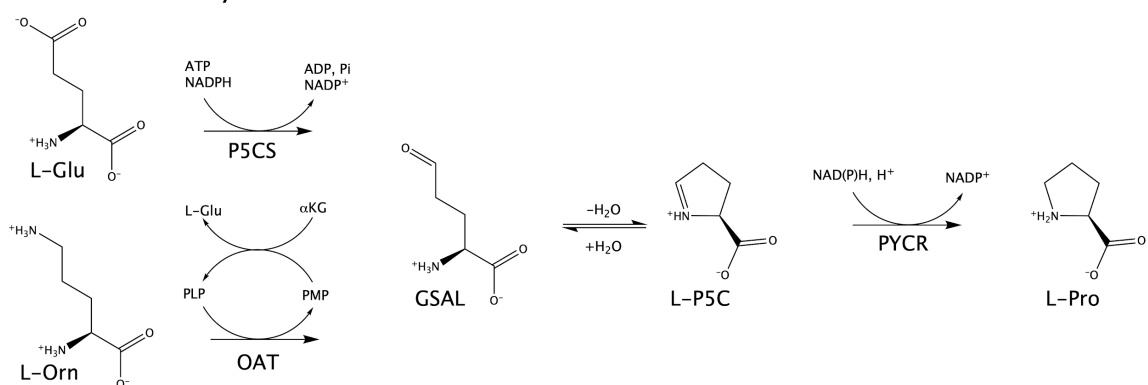


**2**

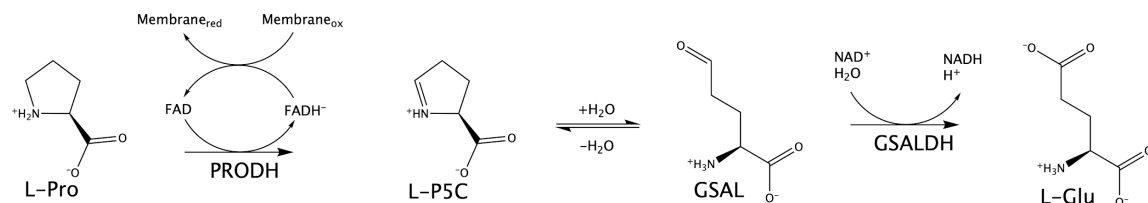
$K_i = 100 \mu M$   
(competitive with P5C)

**Scheme 1.1** Two recently-discovered inhibitors of PYCR1.

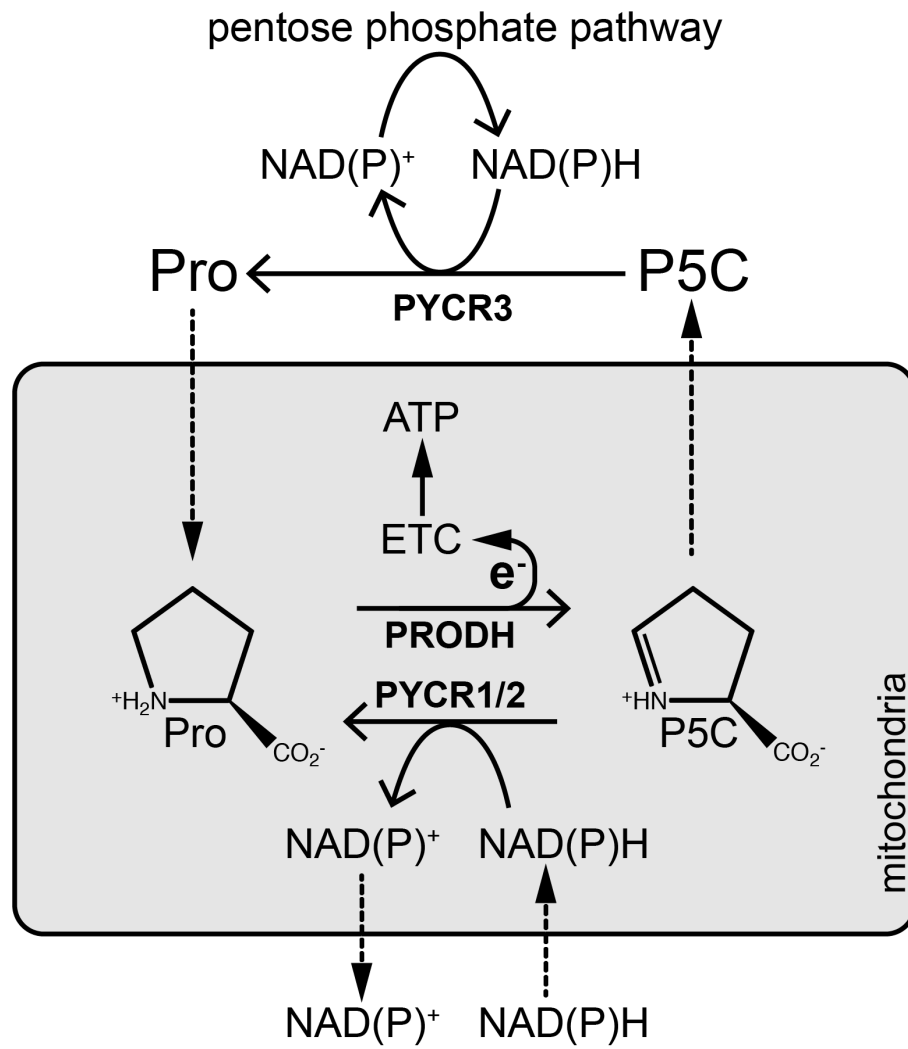
## A – Proline Biosynthesis



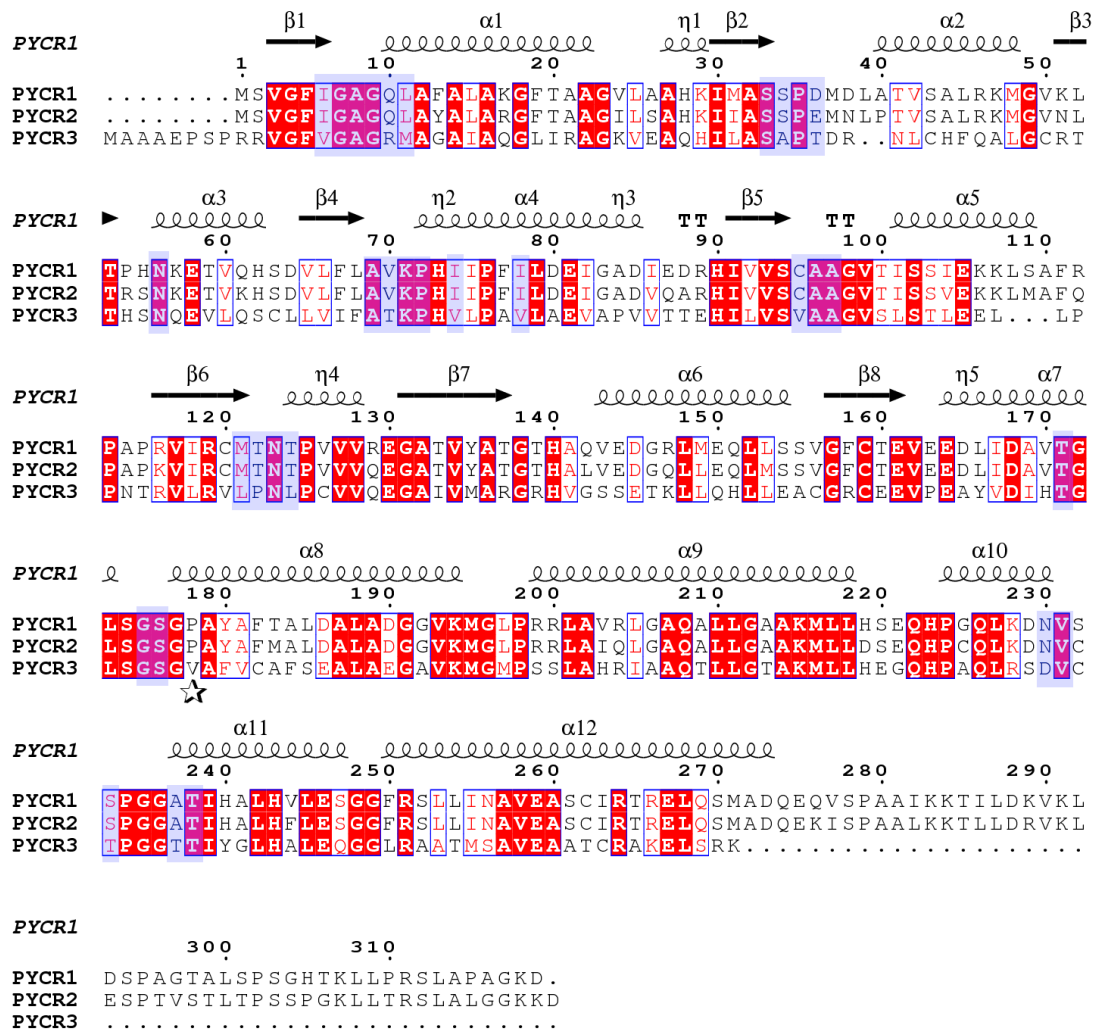
## B – Proline Catabolism



**Figure 1.1** The enzymes and reactions of proline metabolism. (A) Proline biosynthesis from glutamate and ornithine. (B) Proline catabolism. Abbreviations:  $\alpha$ -KG,  $\alpha$ -ketoglutarate; GSAL, L-glutamate- $\gamma$ -semialdehyde; GSALDH, L-glutamate- $\gamma$ -semialdehyde dehydrogenase; OAT, ornithine- $\delta$ -aminotransferase; L-Orn, L-ornithine; PLP, pyridoxal-5'-phosphate; PMP, pyridoxamine-5'-phosphate; PRODH, proline dehydrogenase; P5C,  $\Delta^1$ -pyrroline-5-carboxylate; P5CS,  $\Delta^1$ -pyrroline-5-carboxylate synthase; PYCR,  $\Delta^1$ -pyrroline-5-carboxylate reductase.

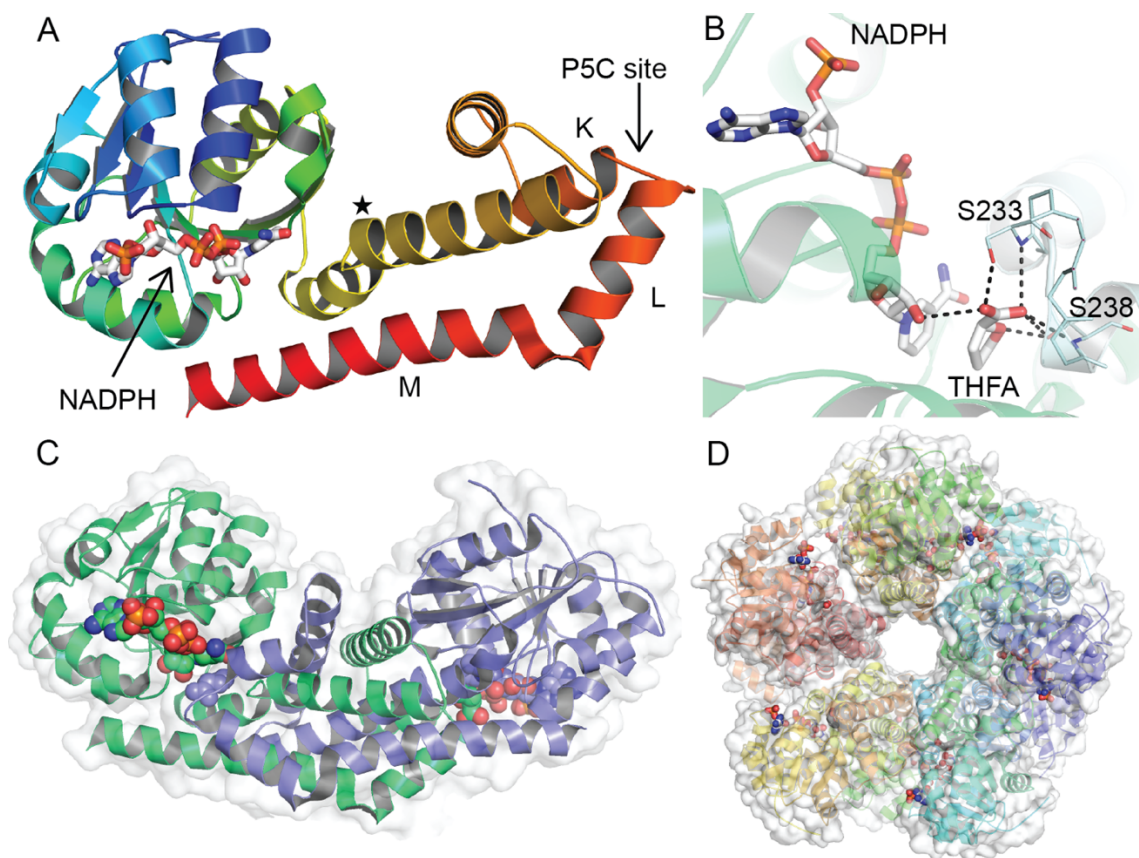


**Figure 1.2** Overview of the proline cycle.

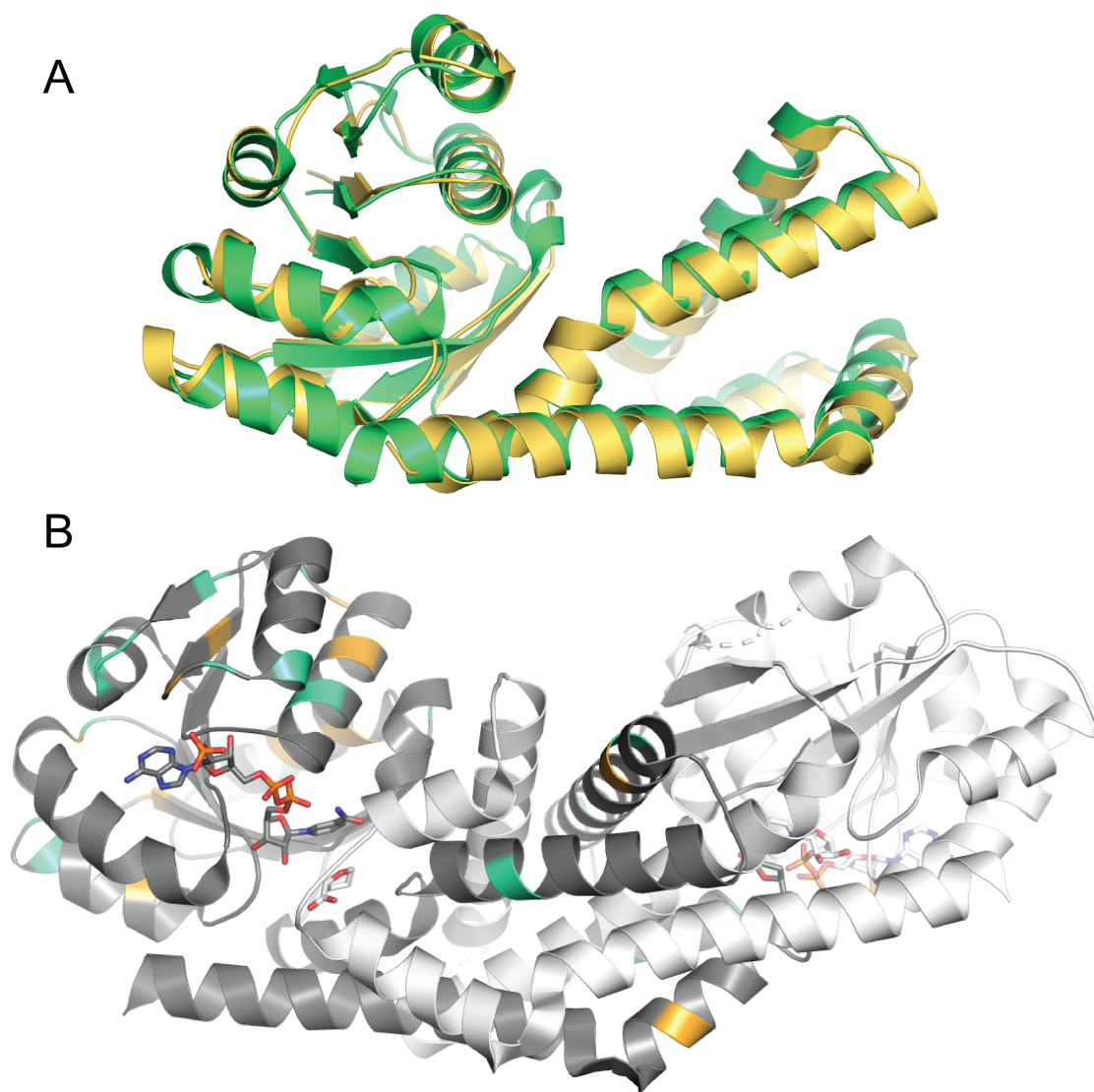


**Figure 1.3** Sequence alignment of PYCR1, PYCR2, and PYCR3 isoforms. The secondary structure elements were obtained from the structure of PYCR1 (PDB ID 5UAV). This figure was made with ESPript 3.0 (Robert and Gouet 2014) from a multiple sequence alignment calculated in Clustal Omega (Sievers et al. 2011). Blue transparent boxes represent residues in the active site of PYCR1. The  $\alpha$ -helix-breaking proline, Pro178, is denoted with a star.

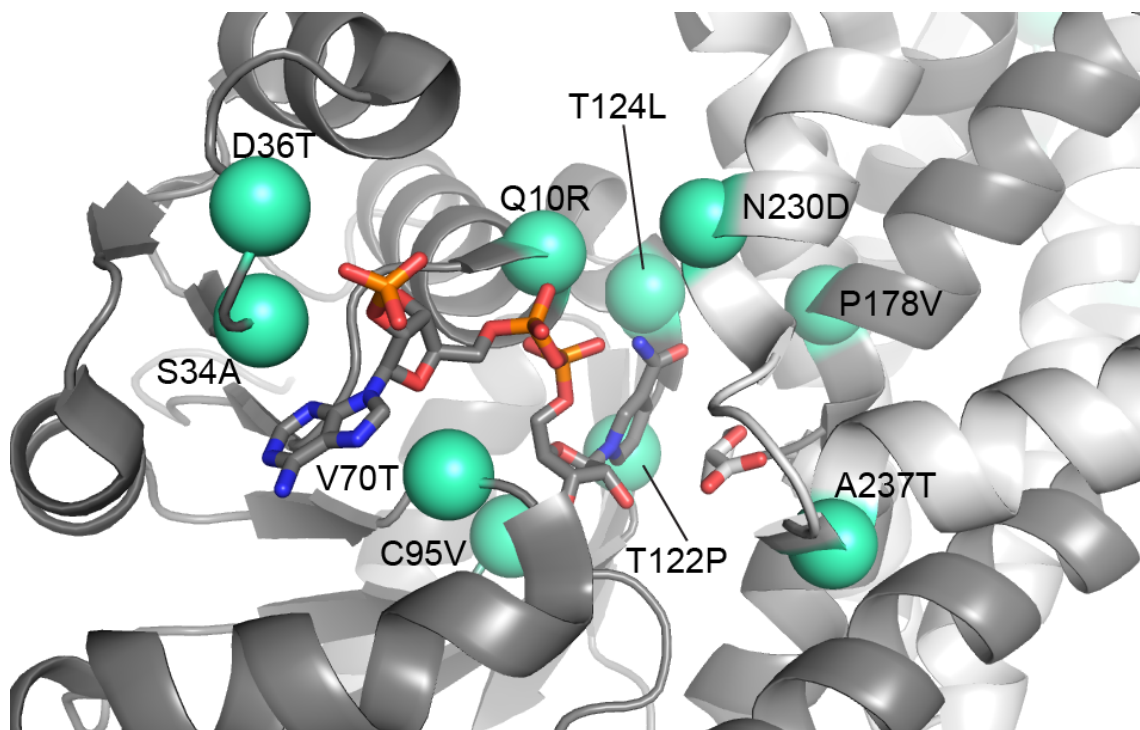




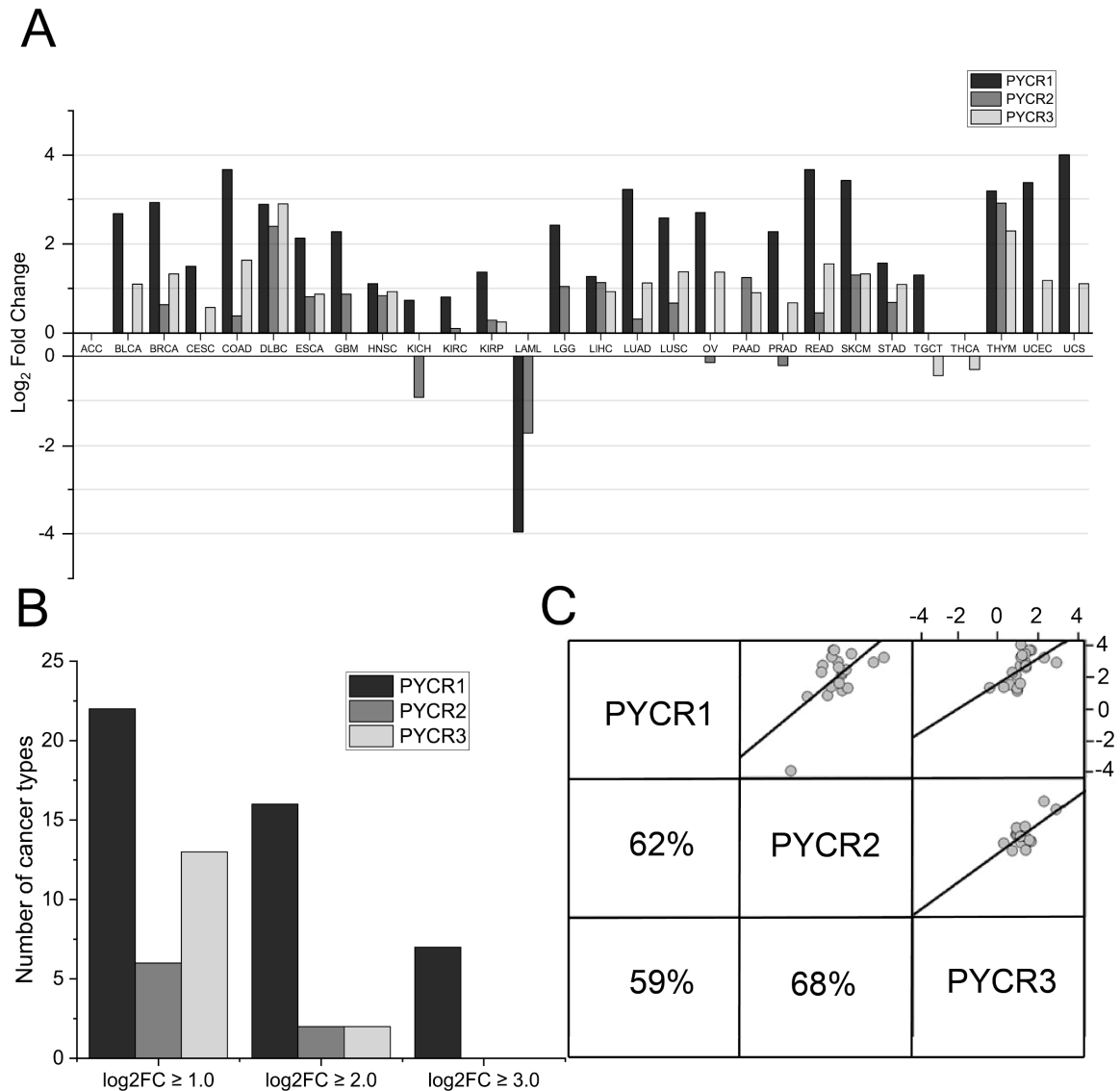
**Figure 1.4** Fold and oligomeric structure of PYCR1. (A) The structure of a PYCR1 protomer showing the protein fold (PDB ID: 5UAV). The polypeptide chain is colored in a rainbow scheme with dark blue at the N-terminus and red at the C-terminus. A black star represents the  $\alpha$ -helix-breaking Pro178 (see Figure 1.3). NADPH is shown in white sticks. Helices K, L, and M are labeled. (B) Close-up view of the active site of PYCR1 with NADPH and (S)-(-)-tetrahydro-2-furoic acid (THFA) bound. Interactions made by THFA with NADPH, Ser233, and Thr238 are shown in black dotted lines. Note the active site is fully formed by the presence of the other protomer completing the dimer. (C) The oligomeric structure of PYCR1 showing the dimer, with NADPH and THFA shown in spheres. (D) The full oligomeric assembly of PYCR1, a pentamer-of-dimers decamer.



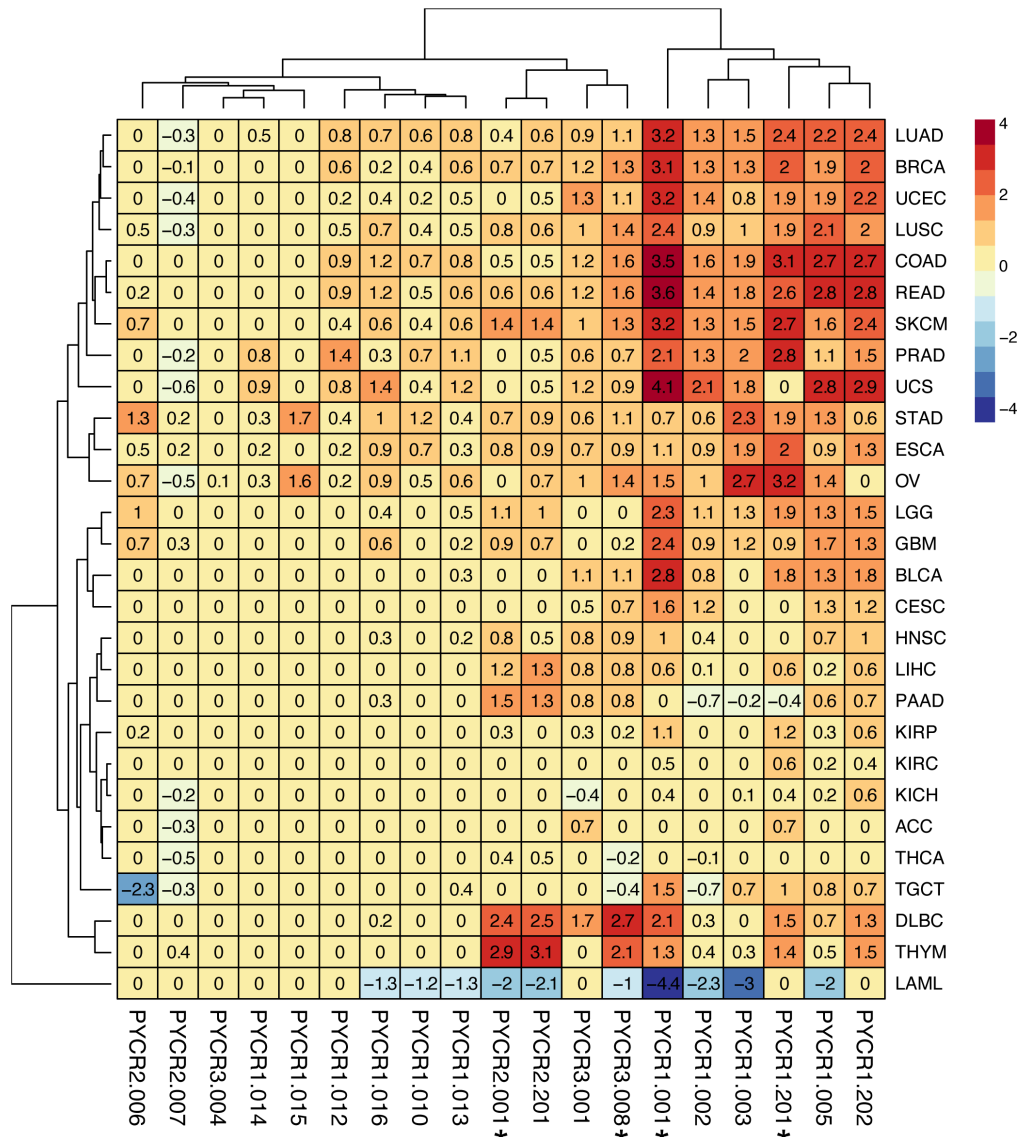
**Figure 1.5** Comparison of the structures of PYCR1 and PYCR2. (A) Cartoon representation of an alignment of the protomers of PYCR1 (PDB ID 5UAV, green) and PYCR2 (PDB ID 6LHM, yellow). (B) A dimer of PYCR1 with pairwise sequence differences from PYCR2 indicated by color. Light green coloring of the left protomer represents positions of non-conserved physicochemical changes and gold are conserved physicochemical sequence changes. The accompanying protomer is included in white to emphasize dimerization is required to form the full active site and to highlight interchain interactions between NADPH and THFA.



**Figure 1.6** Three-dimensional structural representation of nonconservative amino acid variations in the active site of PYCR. The structure of the PYCR1 dimer is shown with the two protomers colored dark gray and white. The light green spheres indicate active site residues that are not conserved in PYCR3. The notation lists the amino acid type and number in PYCR1, followed by the amino acid type in PYCR3. NADPH and the P5C/proline analog THFA are shown in sticks.



**Figure 1.7** Differential gene expression across 28 cancer types for *PYCR* isoforms. (A) The log<sub>2</sub>FC is plotted for each cancer type. (B) The data from panel A are combined to show the number of cancer types with increased expression of *PYCR* genes. (C) Pairwise scatter plot representations of the data from panel A indicating possible associations between the expression patterns of *PYCR* genes in cancer cells. Abbreviations used in panel A: ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumors; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma.



**Figure 1.8** Differential gene expression of *PYCR* splice variants in various cancer types. Numbers in the cells are log2FC values; a value of 0 indicates either that the change was not statistically significant or data were not available. The transcript labels on the horizontal axis refer to the GEPIA IDs listed in Table 1.3. The corresponding Ensembl transcript IDs and UniProt IDs for the encoded proteins are also listed in Table 1.3. Asterisks denote transcript variants encoding experimentally-verified catalytically active enzymes. Abbreviations: ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumors; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma.

## SUPPORTING INFORMATION

### Structure, Biochemistry, and Gene Expression Patterns of the Proline Biosynthetic Enzyme Pyrroline-5-Carboxylate Reductase (PYCR), An Emerging Cancer Therapy Target

**Table S1.1** Summary of literature on the involvement of PYCRs in cancer

Cancer	Synopsis	Citation
Prostate	Anti-apoptotic gene <i>PYCR1</i> newly recognized as being overexpressed	(Ernst et al. 2002)
Prostate	<i>PYCR1</i> gene stimulated by dihydrotestosterone; metastatic tumors showed overexpression; possible AR target	(Jariwala et al. 2007)
Colorectal	Upregulated PYCR1 protein in cellular cancer model	(Roth et al. 2010)
Breast	<i>PYCR1</i> gene hit in RNAi screen and tied to tumor formation	(Possemato et al. 2011)
Melanoma	<i>PYCR1/2</i> expressed in melanoma but not melanocytes; <i>PYCR3</i> expressed in melanocytes but greater in melanoma	(De Ingeniis et al. 2012)
Esophageal	ORAOV1 binds to PYCR1/2; knockdown of <i>PYCR</i> cancelled stress resistance from ORAOV1	(Togashi et al. 2014)
Cancer wide analysis	<i>PYCR1</i> is one of the most commonly overexpressed enzymes in cancer	(Nilsson et al. 2014)
Lymphoma	Increased expression of <i>PYCR1/2/3</i> by MYC; knockdowns showed decrease in cell growth, but only <i>PYCR2</i> showed small decrease in apoptosis	(Liu et al. 2015)
Nasopharyngeal	<i>PYCR3</i> mRNA expression increased in taxol resistant NPC sub-lines compared to parental cells	(Li et al. 2015)
Cancer wide analysis	<i>PYCR1</i> overexpression in 4% of tumors	(Loayza-Puch et al. 2016)

Melanoma	Silencing <i>PYCR2</i> decreased proliferation and autophagy by AMPK/mTOR in certain cell lines	(Ou et al. 2016)
Cancer wide analysis	<i>PYCR3</i> gene hit for positive correlation between somatic copy number alteration and hypoxia signature	(Haider et al. 2016)
Breast	<i>PYCR1</i> knockdown impaired spheroidal growth	(Elia et al. 2017)
Breast	<i>PYCR1</i> (but not <i>PYCR2</i> ) mRNA and protein expression tied to tumor size/grade, invasion, and patient survival	(Ding et al. 2017)
Prostate	<i>PYCR1</i> overexpressed, and knockdown inhibited cell growth and colony formation	(Zeng et al. 2017)
Lung	<i>PYCR1</i> overexpressed; knockdown inhibited cell proliferation and led to cell cycle arrest; <i>PYCR1</i> tied to regulation of cyclin D1	(Cai et al. 2018)
Glioma	Increased expression and activity of <i>PYCR1</i> in IDH1 mutant cells enhanced tumor growth	(Hollinshead et al. 2018)
Breast	Luminal B tumors associated with gain of <i>PYCR1</i> copy number and increased mRNA levels	(Craze et al. 2018)
Melanoma	<i>PYCR1</i> promoted tumor progression through AKT pathway	(Ye et al. 2018)
Lung	<i>PYCR1</i> expression negatively regulated by miR-488; <i>PYCR1</i> activates p38 MAPK pathway	(Wang et al. 2019)
Lung	<i>PYCR1</i> accelerates metastasis through EMT pathway, and knockdown suppressed migration and invasion	(Sang et al. 2019)



Breast	Acetylation of PYCR1 at K228 inhibited cell proliferation	(Chen et al. 2019)
Renal	<i>PYCR1</i> overexpression tied to metastasis and poor overall survival	(Weijin et al. 2019)
Liver	<i>PYCR2</i> identified as possible prognostic biomarker	(Gao et al. 2019)
Lung	PYCR1 regulated by kindlin-2; depleting kindlin-2 decreased PYCR1 levels and inhibited cell proliferation	(Guo et al. 2019)
Esophageal	Increased <i>PYCR2</i> expression in cancer tissue compared to normal muscle and epithelial tissue	(Sun et al. 2019)
Liver	Knockdown of <i>PYCR1</i> led to decreased cell proliferation and promoted apoptosis through regulation of JNK/IRS1 pathway	(Zhuang et al. 2019)
Colorectal	PYCR1 interacts with STAT3; knockdown of <i>PYCR1</i> inhibits proliferation, drug resistance and epithelial-mesenchymal transition of cells ; overexpression of STAT3 partially reverses these effects of <i>PYCR1</i>	(Yan et al. 2019)
Breast	Targeting PYCR1 with small molecule <b>2</b> (Scheme 1.1) impaired proline biosynthesis and spheroidal growth as well as induced proline accumulation	(Christensen et al. 2020)
Liver	<i>PYCR1</i> is one of most upregulated genes in HCC; knockdown decreased proliferation; expression correlates to tumor grade	(Ding et al. 2020)
Lung	<i>PYCR1</i> overexpressed, associated with poor prognosis, knockdown decreased proliferation, migration, and invasion of cells	(Gao et al. 2020)



Multi-cancer	PYCR1 is a client of chaperone Lon; Lon increases protein level of PYCR1 in cancer cells to induce inflammation	(Kuo et al. 2020)
Gastric	<i>PYCR1</i> upregulated, correlated to advanced tumor stage and poor outcome; knockdown stimulated apoptosis and inhibited cell growth; glucose deprivation regulated <i>PYCR1</i> expression; PI3K/Akt axis correlated with PYCR1 expression	(Xiao et al. 2020)
Bladder	RaC3 upregulates <i>PYCR1</i> and activates JAK/STAT signaling	(Cheng et al. 2020)
Lung	PINCH-1 promotes PYCR1-kindlin-2 interaction which is tied to cell proliferation	(Guo et al. 2020)
Breast	<i>PYCR1</i> tied to treatment resistance and overall survival; knockdown reduced invasion and migration capabilities	(Shenoy et al. 2020)
Multi-cancer	In KHSV associated cancers, K1 interacts with PYCR1/2 to enhance cell growth and tumorigenesis	(Choi et al. 2020)
Bladder	Identified <i>PYCR1</i> as potential biomarker and prognosis predictor	(Liu et al. 2021)
Non-uterine leiomyosarcoma	<i>PYCR1</i> expression associated with poor patient overall survival	(Darzi et al. 2021)
Lung	Found negative correlation between MiR-328–3p levels and <i>PYCR1</i> expression	(Lu et al. 2021)

```

1      10      20      30      40      50
PYCR3-004  ....M.AAAEPPSRRVGFFVGAGRMAGAIAGGLIRAGKVEAQHILASAPTDR..NLCHFQALGCRRT
PYCR3-001  .MGG.ERSG...VRGNKMAAAEPPSRRVGFFVGAGRMAGAIAGGLIRAGKVEAQHILASAPTDR..NLCHFQALGCRRT
*PYCR3-008  ....M.AAAEPPSRRVGFFVGAGRMAGAIAGGLIRAGKVEAQHILASAPTDR..NLCHFQALGCRRT
PYCR2-201  ....MSVGGFFGAGQLAYALARGFTAAGILSAHKIIASSPEMNLPTVSALRKKMGVNIT
*PYCR2-001  ....MSVGGFFGAGQLAYALARGFTAAGILSAHKIIASSPEMNLPTVSALRKKMGVNIT
PYCR2-006  ....MNLPTVSALRKKMGVNIT
PYCR2-007  ....MNLPTVSALRKKMGVNIT
PYCR1-016  ....ALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-202  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-015  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-005  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-002  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
*PYCR1-001  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-012  MVGGGRRVGRDEPVPSVGALGQGSPPDSMSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-010  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-012  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-013  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
PYCR1-014  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT
*PYCR1-201  ....MSVGGFFGAGQLAFALAKGFTAAGVLAHKIMASSPDMDLATVSALRKKMGVNIT

60      70      80      90      100     110
PYCR3-004  HSNDVLSQCLLVIFAAPKPHVLPVLAEPVAVVTTHEILVSVAAGVSLSTLEL.....LPENTAVLVLPNLCVVVEGAI
PYCR3-001  HSNDVLSQCLLVIFAAPKPHVLPVLAEPVAVVTTHEILVSVAAGVSLSTLEL.....LPENTAVLVLPNLCVVVEGAI
*PYCR3-008  HSNDVLSQCLLVIFAAPKPHVLPVLAEPVAVVTTHEILVSVAAGVSLSTLEL.....LPENTAVLVLPNLCVVVEGAI
PYCR2-201  RSNKETVKHSDVFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFOFAPKRVIRCMNTNTPVVVREGAT
*PYCR2-001  RSNKETVKHSDVFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFOFAPKRVIRCMNTNTPVVVREGAT
PYCR2-006  RSNKETVKHSDVFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFOFAPKRVIRCMNTNTPVVVREGAT
PYCR2-007  RSNKETVKHSDVFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFOFAPKRVIRCMNTNTPVVVREGAT
PYCR1-016  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-202  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-015  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-005  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-002  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
*PYCR1-001  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-003  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-010  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-012  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-013  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
PYCR1-014  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT
*PYCR1-201  PHNKETVQHSADVFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRFPAPRVIRCMNTNTPVVVREGAT

120     130     140     150     160     170     180     190
PYCR3-004  VMARGRHVGSSEETKLLQHLLEACGRCEVPEAYVDIHTGLSGSGVAFVCAFSALAEAGAVKMGMPSSLAHRIAQAQTLTLLGT
PYCR3-001  VMARGRHVGSSEETKLLQHLLEACGRCEVPEAYVDIHTGLSGSGVAFVCAFSALAEAGAVKMGMPSSLAHRIAQAQTLTLLGT
*PYCR3-008  VMARGRHVGSSEETKLLQHLLEACGRCEVPEAYVDIHTGLSGSGVAFVCAFSALAEAGAVKMGMPSSLAHRIAQAQTLTLLGT
PYCR2-201  ..AFMALDALADGGVKMGLPRLRLAIQLGAQALLGA
*PYCR2-001  VYATGTHALVEDDGLLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFMALDALADGGVKMGLPRLRLAIQLGAQALLGA
PYCR2-006  VYATGTHALVEDDGLLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFMALDALADGGV.....
PYCR2-007  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYGAAVHG.....
PYCR1-016  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-202  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-015  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-005  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-002  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
*PYCR1-001  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-003  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-010  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA
PYCR1-012  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEED.....
PYCR1-013  .....
PYCR1-014  .....
*PYCR1-201  VYATGTHAQVEDDGRLEQLLSSVGFCTEVEEDLIDAVTGLSGSGVPAYAFALDALADGGVKMGLPRLRLAVRLGAQALLGA

200     210     220     230     240     250
PYCR3-004  AKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGGLRAATMSAVEAATCRAKELSRK.....
PYCR3-001  AKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGGLRAATMSAVEAATCRAKELSRK.....
*PYCR3-008  AKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGGLRAATMSAVEAATCRAKELSRK.....
PYCR2-201  AKMLLDSEQHPCQLKDNVCSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKVKIL
*PYCR2-001  AKMLLDSEQHPCQLKDNVCSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKVKIL
PYCR2-006  .....
PYCR2-007  .....
PYCR1-016  AVHG.....
PYCR1-202  AVHG.....
PYCR1-015  AVHG.....
PYCR1-005  AKMLLHEGQHPGQLKDNVSSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKVKIL
PYCR1-002  AKMLLHEGQHPGQLKDNVSSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKDHLP
*PYCR1-001  AKMLLHEGQHPGQLKDNVSSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKVKIL
PYCR1-003  AKMLLHEGQHPGQLKDNVSSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKVKIL
PYCR1-010  AKMLLHEGQHPG.....
PYCR1-012  .....
PYCR1-013  .....
PYCR1-014  .....
*PYCR1-201  AKMLLHEGQHPGQLKDNVSSPPGGATIHALHVFLESGGFRSLLINAVEASCIRTRELQSMADQEQVSPAAIKKTLIDKVKIL

PYCR3-004  .....
PYCR3-001  .....
*PYCR3-008  .....
PYCR2-201  SPTVSTLTPSPSGK.LLTRLALGGKKD
*PYCR2-001  SPTVSTLTPSPSGK.LLTRLALGGKKD
PYCR2-006  .....
PYCR2-007  .....
PYCR1-016  .....
PYCR1-202  .....
PYCR1-015  .....
PYCR1-005  SPAGTALSPSGHTK.LLPRSLAPAGKD.
PYCR1-002  LELGSP..EGLHPLLQYQLARAPS..
*PYCR1-001  SPAGTALSPSGHTK.LLPRSLAPAGKD.
PYCR1-003  SPAGTALSPSGHTK.LLPRSLAPAGKD.
PYCR1-010  .....
PYCR1-012  .....
PYCR1-013  .....
PYCR1-014  .....
*PYCR1-201  SPAGTALSPSGHTK.LLPRSLAPAGKD.

```

**Figure S1.1** Amino acid sequence alignment of PYCR proteins generated by alternative splicing. The sequence labels refer to the GEPIA IDs listed in Table 1.3. The corresponding Ensembl transcript IDs and UniProt IDs for the encoded proteins are also listed in Table 1.3. Asterisks denote transcript variants encoding experimentally-verified catalytically active enzymes.

## Chapter 2

### Structure-Affinity Relationships of Reversible Proline Analog Inhibitors Targeting Proline Dehydrogenase

Alexandra N. Bogner<sup>1</sup> and John J. Tanner<sup>1,2\*</sup>

<sup>1</sup>Department of Biochemistry, University of Missouri, Columbia, Missouri 65211, United States

<sup>2</sup>Department of Chemistry, University of Missouri, Columbia, Missouri 65211, United States

**Corresponding author:** \*Department of Biochemistry, University of Missouri, Columbia, MO 65211. Phone: (573) 884-1280. E-mail: [tannerjj@missouri.edu](mailto:tannerjj@missouri.edu).

**Databases:** Coordinates and structure factor amplitudes have been deposited in the Protein Data Bank under accession codes 7MWT, 7MWU, and 7MWV.

**Conflict of Interest:** The authors declare no competing financial interest.

**Abbreviations:** GSAL, L-glutamate- $\gamma$ -semialdehyde; GSALDH, L-glutamate- $\gamma$ -semialdehyde dehydrogenase; *o*-AB, *ortho*-aminobenzaldehyde; P5C,  $\Delta^1$ -pyrroline-5-carboxylate; PRODH, proline dehydrogenase; PutA, proline utilization A; PutA86-630, protein containing residues 86-630 of *Escherichia coli* proline utilization A; PYCR,  $\Delta^1$ -pyrroline-5-carboxylate reductase

**Author Contributions:** A.N.B.: Conceptualization, Methodology, Investigation, Writing-Original Draft, Writing-Review & Editing, Visualization. J.J.T.: Conceptualization, Writing-Original Draft, Writing-Review & Editing, Visualization, Validation, Supervision, Project administration, Funding acquisition.

## ABSTRACT

Proline dehydrogenase (PRODH) catalyzes the first step of proline catabolism, the FAD-dependent oxidation of L-proline to  $\Delta^1$ -pyrroline-5-carboxylate. PRODH plays a central role in the metabolic rewiring of cancer cells, which has motivated the discovery and characterization of inhibitors. Here, we studied the inhibition of PRODH by 18 proline-like compounds to understand the structural and chemical features responsible for the affinity of the best-known PRODH inhibitor, S-(–)-tetrahydro-2-furoic acid (**1**). The compounds were screened for inhibition, and then six were selected for more thorough kinetic analysis: cyclobutane-1,1-dicarboxylic acid (**2**), cyclobutanecarboxylic acid (**3**), cyclopropanecarboxylic acid (**4**), cyclopentanecarboxylic acid (**16**), 2-oxobutyric acid (**17**), and (2S)-oxetane-2-carboxylic acid (**18**). These compounds are competitive inhibitors with inhibition constants in the range of 1.4 - 6 mM. Crystal structures of PRODH complexed with **2**, **3**, and **4** were determined. All three inhibitors bind in the proline substrate site, as expected. Compound **2** is unique in that the additional carboxylate displaces a structurally conserved water molecule from the active site. The binding of **3** is accompanied by compression of the active site to enable nonpolar contacts with the cyclopropane ring. The results are interpreted using a thermodynamic cycle to understand how hydrogen bonding and ring size contribute to the affinity. This analysis revealed that a water-mediated hydrogen bond with the O heteroatom of **1** contributes 2 kcal/mol of binding energy and that the size and hydrogen bonding capacity of the ring are synergistically related. These structure-affinity relationships may be useful for future inhibitor design targeting PRODH.

## INTRODUCTION

Proline catabolism, the four-electron oxidation of proline to glutamate, occurs in the mitochondria and is catalyzed by two enzymes. Proline dehydrogenase (PRODH), the first enzyme, is a flavin-dependent enzyme that catalyzes the oxidation of proline to  $\Delta^1$ -pyrroline-5-carboxylate (P5C) (Figure 2.1). P5C then undergoes nonenzymatic hydrolysis to form L-glutamate- $\gamma$ -semialdehyde (GSAL), the substrate for the second enzyme in proline catabolism, GSAL dehydrogenase (GSALDH). GSALDH uses the cofactor  $\text{NAD}^+$  in the oxidation of GSAL to L-glutamate.

Proline metabolism is central to the metabolic rewiring of cancer cells. Proline biosynthesis, catabolism, and cycling have been implicated as metabolic pathways selectively altered in cancer cells providing ATP, macromolecules, and redox cofactors.<sup>1–3</sup> PRODH acts as a tumor suppressor or an oncogene depending on the tumor type, and the environmental and metabolic context.<sup>4</sup> Its oncogenic character has been revealed in non-small cell lung cancer, where increased PRODH expression promotes tumorigenesis by inducing epithelial to mesenchymal transition and several inflammatory genes.<sup>5</sup> PRODH is also important in the rewiring of breast cancer cells leading to metastasis. Human metastasis tissue exhibits upregulated expression of PRODH compared with primary breast tumor tissue, and the inhibition of PRODH by S-(–)-tetrahydro-2-furoic acid (**1** in Figure 2.2) impairs metastasis formation in breast cancer mouse models.<sup>6</sup> PRODH's involvement in cancer metabolism is thought to manifest through the proline cycle, a substrate cycle composed of PRODH and the proline biosynthetic enzyme  $\Delta^1$ -pyrroline-5-carboxylate reductase 1 (PYCR1). PRODH forms half of the cycle by catalyzing the oxidation of proline to P5C, while PYCR1 catalyzes the reverse transformation, the NADPH-dependent

reduction of P5C to L-proline. The proline cycle has been implicated in supporting ATP production, protein and nucleotide synthesis, anaplerosis, and redox homeostasis in cancer cells.<sup>2</sup>

The involvement of the proline cycle in cancer cell metabolism has motivated the development of inhibitors targeting PRODH and PYCR1. The proline analog **1** is the best characterized reversible inhibitor of PRODH. Although the affinity for PRODH is modest ( $K_i \sim 0.2$ - $1.0$  mM),<sup>7,8</sup> **1** has proven to be useful for investigating the proline cycle in cancer cells and mouse models of cancer.<sup>6,9,10</sup> Other small carboxylic acids, such as L-lactic acid ( $K_i \sim 1$  mM) and acetic acid ( $K_i \sim 30$  mM), are weak inhibitors of PRODH.<sup>9,11</sup> Irreversible inactivators of PRODH, which covalently modify the N5 of the FAD of PRODH, have also been investigated. These include N-propargylglycine,<sup>10,12-14</sup> thiazolidine-2-carboxylate,<sup>14,15</sup> and 1,3-dithiolane-2-carboxylate.<sup>16</sup> The latter compound is unique in that the inactivation mechanism is photoinduced by blue light. Inhibitors of PYCR1 have been less studied and include the proline analog N-formyl-L-proline<sup>17</sup> and pargyline derivatives.<sup>18,19</sup>

Motivated by the observation that certain small carboxylic acids inhibit PRODH, we screened a series of 18 cycloalkyl carboxylic acids and related compounds as probes to explore the active site (Figure 2.2). The inhibition constants for six of the compounds were determined and found to be in the range of 1 - 6 mM. The crystal structures of three of the compounds in complex with PRODH were determined to understand the basis for inhibition. Our results provide insight into the tolerance of the active site for inhibitor ring size and demonstrate the importance of hydrogen bonding to a conserved active site water molecule for potency.

## MATERIALS AND METHODS

**Materials.** The following compounds were bought from Sigma: (S)-(-)-tetrahydro-2-furoic acid (**1**) product number 527890, cyclobutane-1,1-dicarboxylic acid (**2**) product number C95803, cyclobutanecarboxylic acid (**3**) product number C95609, cyclopropanecarboxylic acid (**4**) product number C116602, 2-methylcyclopropanecarboxylic acid (**5**) product number 209759, cyclopropane-1,1-dicarboxylic acid (**6**) product number 343412, transcyclobutane-1,2-dicarboxylic acid (**7**) product number 28684, 2,2,3,3-tetramethylcyclopropanecarboxylic acid (**8**) product number 301566, 1-(aminomethyl)cyclopropanecarboxylic acid (**9**) product number CDS015451, 3-oxocyclobutanecarboxylic acid (**10**) product number CDS004694, cyclopropylacetic acid (**11**) product number CDS000888, 2-thiophenecarboxylic acid (**12**) product number T32603, (S)-(-)-tetrahydrofuran-2-carboxamide (**13**) product number 573310, cis-cyclobutane-1,2-dicarboxylic acid (**14**) product number 28682, tetrahydropyran-2-carboxylic acid (**15**) product number CDS015473, cyclopentanecarboxylic acid (**16**) product number C112003, 2-oxobutyric acid (**17**) product number 68217, (2S)-oxetane-2-carboxylic acid (**18**) product number SY3H6E416CAC.

**Protein Expression and Purification.** The PRODH domain of the bifunctional PRODH-GSALDH enzyme proline utilization A (PutA) from *Escherichia coli* was used for crystallization and kinetic assays. A PRODH domain construct containing residues 86-630 of *E. coli* PutA (PutA86-630) with a C-terminal His tag was expressed and purified using a protocol developed from previous studies of *E. coli* PutA PRODH domain constructs.<sup>9,20-22</sup> PutA86-630 was overexpressed in *E. coli* strain BL21(DE3) pLysS. Cells

were grown in TB media for 3 hours at 37 °C and induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside for 3 hours at 37 °C. Harvested cells were resuspended in a buffer containing 20 mM Tris pH 7.9, 5 mM imidazole, 500 mM NaCl, and 10% w/v glycerol (buffer A) and then lysed via sonication in the presence of 1 mM FAD and EDTA-free protease inhibitor tablets (Thermo Fisher). Centrifugation at 16000 rpm for 1 hour at 4 °C was performed to remove insoluble material. The lysate was then purified by gravity-flow chromatography on a column containing Ni<sup>2+</sup>-NTA resin (Qiagen) pre-equilibrated with buffer A. The column was washed with buffer A and eluted with buffer A supplemented with 500 mM imidazole. Elution fractions containing PutA86-630 were identified on SDS-PAGE, pooled, spiked with 1 mM FAD, and dialyzed into 70 mM Tris pH 8.1, 2 mM EDTA, and 10% w/v glycerol. Anion exchange with a 5 mL HiTrap Q column was performed the next day using a gradient of KCl (0-0.5 M over 20 column volumes). PutA86-630 was collected in the flow-through, while contaminating proteins were retained by the resin. PutA86-630 was dialyzed into 70 mM Tris pH 7.5, 5 mM EDTA, and 10% w/v glycerol. The protein concentration was estimated with a Bradford assay and confirmed with absorbance at 280 nm. The purified enzyme was distributed in 50  $\mu$ L aliquots in 1.5 mL microcentrifuge tubes and stored at -80 °C.

**Enzyme Activity Assays.** Kinetic measurements were performed in a 96-well plate in a BioTek Epoch 2 microplate spectrophotometer at room temperature in a buffer containing 20 mM MOPS pH 7.5 and 10 mM MgCl<sub>2</sub>. The assay monitors the production of P5C as an adduct formed with *ortho*-aminobenzaldehyde (*o*-AB), which is detected by absorbance at 443 nm ( $\epsilon_{443} = 2.59 \text{ mM}^{-1}\text{cm}^{-1}$ ).<sup>23</sup> Menadione was used as the electron acceptor to reoxidize the reduced FAD of PROD<sub>H</sub>, enabling catalytic cycling. An initial assessment of enzyme



inhibition by a panel of compounds was performed at one substrate concentration using 200 mM L-proline, 4 mM *o*-AB, 0.15 mM menadione, 63 nM PutA86-630, and 5 mM of the potential inhibitor. For compounds **1** - **4** and **16-18**, kinetic measurements were performed at varied L-proline (0-500 mM) and inhibitor (0-10 mM) concentrations. L-proline and inhibitors were spotted on the plate and a master mix including enzyme, menadione, *o*-AB, and buffer was added to the plate by multichannel pipette to initiate the reaction. The initial rate was determined from linear regression of the first 10 minutes of the progress curve using Origin v9.7.0.188 software. The initial rate data as functions of both substrate and inhibitor concentrations were fit globally to a competitive inhibition model using Origin. Kinetic constants from fitting are listed in Table 2.1.

**Crystallization and Cryoprotection.** Crystals were grown at 20 °C in a hanging drop setup with 20 mg/mL protein. A drop ratio of 2 uL:2 uL for protein: reservoir solution was used. The reservoir solution contained 50-175 mM sodium citrate pH 6.2 and 20-26% PEG 3350. Crystals were harvested, soaked in reservoir solution supplemented with 50 mM **2** for a few minutes, cryoprotected with reservoir solution supplemented with 20% PEG 200, and flash-cooled in liquid nitrogen. Crystals for the complexes with **3** and **4** were grown using a reservoir solution containing 100 mM sodium citrate pH 5.6, 5.75, 5.9 or 6.0 and 18-28% PEG 3000. Crystals were harvested, soaked with reservoir solution supplemented with 100 mM **3** or 50 mM **4** for a few minutes, and then cryoprotected with reservoir solution supplemented with 20% PEG 200 and flash-cooled in liquid nitrogen.

**X-ray Crystal Structure Determination.** X-ray diffraction data were collected at Advanced Photon Source beamline 24-ID-E using an Eiger-16 M detector and 24-ID-C using an Eiger2 X 16 M detector. The data were processed with XDS<sup>24</sup> and AIMLESS<sup>25</sup>.

The space group is *I*222, and the asymmetric unit contains one PutA86-630 chain. We note this is the same crystal form used previously for structural studies of the *E. coli* PutA PRODH domain.<sup>13,20,26,27</sup> Data processing statistics are summarized in Table 2.2.

A 1.85 Å resolution structure of a PutA86-630 variant complexed with **1** (PDB ID 3E2R)<sup>26</sup> was used as a starting model for crystallographic refinement of the enzyme complexed with **2** in PHENIX.<sup>28,29</sup> Refinements of the complexes with **3** and **4** were started from the finalized model of the complex with **2**. The *B*-factor model consisted of one TLS group per protein chain and isotropic *B*-factors for all non-hydrogen atoms. Interactive model building was performed with COOT.<sup>30</sup> SMILES strings for inhibitors were used as the input to ELBOW<sup>31</sup> to generate the coordinates and restraint files used during refinement. The structures were validated using MolProbity and the wwPDB validation service.<sup>32,33</sup> Modeling of ligands was validated with polder omit maps.<sup>34</sup> Refinement statistics are summarized in Table 2.2.

## RESULTS AND DISCUSSION

**Compound Screening.** For this study, we used the PRODH domain of the bifunctional PRODH-GSALDH enzyme proline utilization A (PutA) from *E. coli*, which we used previously to investigate the reversible inhibition of PRODH by **1**, L-lactic acid, and acetic acid.<sup>9</sup> The PutA PRODH domain construct used here (residues 86-630, PutA86-630) is a good model system for inhibitor testing because of its crystallizability (routinely better than 2.0 Å resolution)<sup>26,27</sup> and the high sequence conservation of PRODH active sites across bacteria and eukaryotes.<sup>2</sup> We note that human PRODH is an inner mitochondrial membrane protein and is challenging to isolate for structural and kinetic study.

The inhibition of PutA86-630 by several cycloalkyl carboxylic acids and related compounds was investigated to provide insight into the nature of chemical structures that bind the PRODH active site (Figure 2.2). The need for a deeper understanding of inhibitory effects of **1** drove the design of the panel of compounds. Starting with **1** and making iterative alterations to the chemical structures, we formulated a list of low molecular weight compounds to explore the active site of PRODH (Figure 2.2).

An initial assessment was performed with the inhibitor at a concentration of 5 mM and the substrate L-proline at 200 mM ( $\sim 2K_m$ ). Under these conditions, the enzyme activity in the presence of **1** is 12% of the no-inhibitor control (Figure 2.3). Although none of the compounds achieved a level of inhibition similar to **1**, the results provide structure-activity relationships that help explain the activity of **1**. For example, removal of the O atom in the ring of **1**, as in **16**, or increasing the ring size to six, as in **15**, severely decreases inhibition. Substituting an amide for the carboxylate (**13**) also decreases inhibition. These results show that interactions with both the tetrahydrofuran and carboxylate group are essential to the activity of **1**.

The compounds tested included carboxylates of cyclopropane and cyclobutane to explore the potential of inhibitors with rings smaller than the substrate L-proline. Generally, the cyclobutyl compounds exhibited better inhibition than the cyclopropyl derivatives. In particular, **2** as well as **3** and **18**, reduced enzyme activity by over 2-fold relative to the no-inhibitor control (Figure 2.3). In contrast, none of the cyclopropyl compounds achieved better than 50% inhibition. Including a second carboxylate group (**2**, **6**, **7**, **14**) did not improve inhibition compared to the mono-carboxylate parents (**3**, **4**).

Cyclobutane carboxylate (**3**) inhibited PRODH almost twice as much as cyclopentane carboxylate (**16**), which led us to hypothesize that **18** had potential to be a better inhibitor than **1**. Surprisingly, **18** was no more effective than **3** in the single-point assay (Figure 2.3). Thus, tetrahydrofuran is preferred over oxetane.

Compound **17** was included to test an acyclic analog of **18** and for its structural similarity to L-lactate, a known inhibitor of PRODH. **17** was less effective than **18**, which perhaps reflects a larger conformational entropy penalty for binding **17**.

We tested **10** as an example of a compound with the same atom inventory as **1** but with a different three-dimensional structure (i.e., functional group isomer). **10** showed very little inhibition, suggesting the oxo group does not engage hydrogen bonding partners as effectively as the heteroatom O of **1**.

**Estimation of Inhibition Constants.** Six compounds were selected for additional kinetic analysis to estimate the inhibition constants for comparison to our reference compound **1**. Compounds **2**, **3**, **17**, and **18** were chosen because they caused significant inhibition in the single-point assay (>50% inhibition, Figure 2.3). **4** and **16** were included for comparison to **3** to investigate the influence of cycloalkyl ring size on ligand recognition. For each compound, the initial rate as functions of L-proline and inhibitor concentration were fit globally to the competitive model (Figure 2.4). Table 2.1 lists the  $K_i$  values from global fitting.

Compound **1** has a  $K_i$  of 0.3 mM, consistent with previous studies.<sup>7,8</sup> In the comparison of cycloalkyl ring size, cyclobutane carboxylate (**3**) is preferred with a  $K_i$  of 1.9 mM, followed by cyclopropane carboxylate (**4**,  $K_i$  = 3.1 mM), and cyclopentane carboxylate (**16**,  $K_i$  of 6.2 mM). Incorporation of an O atom into the ring of the preferred cycloalkane

carboxylate (i.e., **18**) or the addition of a second carboxylate (**2**) lowered  $K_i$  by about 20% from 1.9 mM to ~1.5 mM. The acyclic compound **17** inhibits comparably to **3**. These results are generally consistent with the single-point survey data (Figure 2.3).

**Crystal Structures of PRODH-Inhibitor Complexes.** Crystal structures of PutA86-630 inhibited by **2**, **3**, and **4** were determined (Table 2.2). All the inhibitors occupy the known proline substrate site, between the FAD isoalloxazine and the  $\alpha 8$  helix, as shown for **2** (Figure 2.5). Helix  $\alpha 8$  is notable for containing the conserved sequence motif YXXRRXXET/N, which provides residues that interact with L-proline (Figure 2.5B). Electron density was strong for all the inhibitors (Figure 2.6), and all were modeled at an occupancy of 1.0 with reasonable  $B$ -factors, i.e., similar to that of the FAD (Table 2.2).

The binding poses of the inhibitors are very similar (Figure 2.6). In each structure, the carboxylate group of the inhibitor ions pairs with Arg555 and Arg556 of  $\alpha 8$ , as well as Lys329. Additional residues surround the aliphatic part of the inhibitor, including Leu513, Tyr552, and Tyr540. This mode of binding is very similar to that of **1** (Figure 2.6A).

A water molecule hydrogen bonded to Tyr437 is present in all the structures except the complex with **2** (Figure 2.6). This water molecule is present in ligand-free PRODH structures, implying it is a conserved structural feature of the active site (e.g., PDB IDs 6X9A<sup>35</sup> and 4NM9<sup>36</sup>). The conserved water molecule plays a key role in the recognition of **1** by hydrogen bonding with the O heteroatom of the inhibitor (Figure 2.6A). Obviously, **3** and **4** do not have a hydrogen bonding group available to engage the conserved water molecule. Interestingly, the conserved water is absent in the complex with **2**. The extra carboxylate **2** occupies the conserved water site and forms a hydrogen bond with Tyr437 (Figure 2.7A). These results suggest that the binding of **2** is accompanied by the

dissociation of the conserved water molecule. The increase in entropy associated with release of the water molecule into the bulk solvent presumably contributes to the affinity of **2**.

The active site accommodates the different ring sizes with minimal structural change. The only perceptible difference is that Tyr540 tilts 0.4 Å toward the ligand binding site in the complex with the smallest compound, **4** (Figure 2.7B). This slight deflection allows Tyr540 to pack against the cyclopropane ring, diminishing open space around the inhibitor.

The poses of **3** and **4** differ somewhat from that of the reference compound, **1**. Compounds **3** and **4** appear to be rotated by ~45° away from the conserved water molecule and toward Leu513 (Figure 2.7B). The apparent rotation enables the nonpolar rings of these compounds to form favorable interactions with Leu513 while avoiding an unfavorable interaction with the conserved water molecule.

## CONCLUSIONS

The growing evidence for the proline cycle enzymes as cancer therapy targets motivates the discovery of chemical probes of PRODH and PYCR1.<sup>2,3,37</sup> The exploration of how their active sites recognize chemical matter may inform inhibitor design. Here, we performed a focused screen of compounds related to **1**, the best characterized reversible inhibitor of PRODH. The panel of compounds enabled us to dissect the contributions of the structural and chemical features of **1** to affinity, including the carboxylate group, the heteroatom O, and ring size.

The carboxylate of **1** is clearly important for potency. Substituting the carboxylate of **1** with an amide (**13**) markedly reduced affinity and confirms the importance of the ionic

interactions of **1** with the conserved arginine residues of helix  $\alpha 8$ . Indeed, all known reversible inhibitors and covalent inactivators of PRODH contain a carboxylate group. This result suggests that the investigation of compounds with carboxylate mimics, such as phosphonates and sulfonates, may be an avenue to discovering new inhibitors.

Our data provide insight into the contributions of the O heteroatom and ring size of **1** for potency at the enzyme active site, which is represented by a chemical double mutant thermodynamic cycle (Figure 2.8). This type of cycle is analogous to the double mutant thermodynamic cycles used to analyze site-directed mutagenesis of proteins.<sup>38</sup> In this case, the two “mutagenesis” steps are removing the O heteroatom from **1** to generate **16** and reducing the ring size of **1** by one carbon to compound **18**. The double mutation thus converts **1** to **3**. The  $\Delta\Delta G$  values for each leg of the cycle are obtained from the inhibition  $K_i$  values.

Starting from **1**, removing the O heteroatom to generate **16** resulted in a 20-fold decrease in  $K_i$ , implying that the hydrogen bond with the conserved water molecule contributes almost 2 kcal/mol to the binding energy (Figure 2.8). In contrast, removing the O heteroatom from the 4-membered ring of **18** to generate **3** resulted in only a modest loss of binding energy of 0.18 kcal/mol. Decreasing the ring size of **1** by one C atom to **18** increased  $K_i$  5-fold, which amounts to a loss of 0.95 kcal/mol of binding energy. In contrast, decreasing the cyclopentane of **16** to cyclobutane of (**3**) *improved* affinity by 0.71 kcal/mol. These results suggests that the strength of the hydrogen bond between the O heteroatom and the conserved water molecule depends on ring size, i.e., that these two elements of molecular recognition interact thermodynamically. This interaction may be obtained from

the thermodynamic cycle as -1.7 kcal/mol (Figure 2.8), which indicates that the hydrogen bond is stronger in the context of the 5-membered ring than in the 4-membered ring.

Five appears to be the optimal ring size for the O-heterocyclic inhibitors. Both expanding the ring size to six (**15**) or decreasing it to four (**18**) decreased affinity. In contrast, we found that four is the optimal ring size for the cycloalkane carboxylate inhibitors. For example, decreasing the ring size from five in cyclopentane carboxylate (**16**) to four in cyclobutane carboxylate (**3**) improved affinity by 0.7 kcal/mol; however, decreasing the ring size further to cyclopropane carboxylate (**4**) decreased affinity by 0.3 kcal/mol. Thus, the presence of the O heteroatom allows the PRODH active site to accommodate a larger ring. These results are consistent with the contributions of the size and hydrogen bonding capacity of the ring being nonadditive. These structure-affinity relationships may be useful for future inhibitor design targeting PRODH.

## ACKNOWLEDGMENTS

We thank Jonathan Schuermann for help with X-ray diffraction data collection. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by a NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.



**Funding Information**

This work was supported by NIGMS, National Institutes of Health, Grant R01GM132640.

A.N.B. is the recipient of a Wayne L. Ryan Fellowship through The Ryan Foundation.

## REFERENCES

1. Phang JM (2019) Proline Metabolism in Cell Regulation and Cancer Biology: Recent Advances and Hypotheses. *Antioxid Redox Signal* [Internet] 30:635–649. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28990419>
2. Tanner JJ, Fendt SM, Becker DF (2018) The Proline Cycle As a Potential Cancer Therapy Target. *Biochemistry* 57:3433–3444.
3. Bogner AN, Stiers KM, Tanner JJ (2021) Structure, biochemistry, and gene expression patterns of the proline biosynthetic enzyme pyrroline-5-carboxylate reductase (PYCR), an emerging cancer therapy target. *Amino Acids* [Internet]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/34003320>
4. Burke L, Guterman I, Palacios Gallego R, Britton RG, Burschowsky D, Tufarelli C, Rufini A (2020) The Janus-like role of proline metabolism in cancer. *Cell Death Discov* [Internet] 6:104. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/33083024>
5. Liu Y, Mao C, Wang M, Liu N, Ouyang L, Liu S, Tang H, Cao Y, Liu S, Wang X, et al. (2020) Cancer progression is mediated by proline catabolism in non-small cell lung cancer. *Oncogene* [Internet] 39:2358–2376. Available from: <http://dx.doi.org/10.1038/s41388-019-1151-5>
6. Elia I, Broekaert D, Christen S, Boon R, Radaelli E, Orth MF, Verfaillie C, Grünewald TGP, Fendt S-M (2017) Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nature Communications* [Internet] 8:15267. Available from: <http://www.nature.com/doi/10.1038/ncomms15267>
7. Zhu W, Gincher Y, Docherty P, Spilling CD, Becker DF (2002) Effects of proline analog binding on the spectroscopic and redox properties of PutA. *Arch. Biochem. Biophys.* [Internet] 408:131–136. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12485611](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12485611)
8. White TA, Krishnan N, Becker DF, Tanner JJ (2007) Structure and kinetics of monofunctional proline dehydrogenase from *Thermus thermophilus*. *J. Biol. Chem.* [Internet] 282:14316–14327. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17344208](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17344208)
9. Zhang M, White TA, Schuermann JP, Baban BA, Becker DF, Tanner JJ (2004) Structures of the *Escherichia coli* PutA proline dehydrogenase domain in complex with competitive inhibitors. *Biochemistry* [Internet] 43:12539–12548. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15449943](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15449943)

10. Scott GK, Yau C, Becker BC, Khateeb S, Mahoney S, Jensen MB, Hann B, Cowen BJ, Pegan SD, Benz CC (2019) Targeting Mitochondrial Proline Dehydrogenase with a Suicide Inhibitor to Exploit Synthetic Lethal Interactions with p53 Upregulation and Glutaminase Inhibition. *Mol Cancer Ther* [Internet] 18:1374–1385. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/31189611>
11. Kowaloff EM, Phang JM, Granger AS, Downing SJ (1977) Regulation of proline oxidase activity by lactate. *Proc. Natl. Acad. Sci. U S A* [Internet] 74:5368–5371. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=271958](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=271958)
12. White TA, Johnson Jr. WH, Whitman CP, Tanner JJ (2008) Structural basis for the inactivation of *Thermus thermophilus* proline dehydrogenase by N-propargylglycine. *Biochemistry* [Internet] 47:5573–5580. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=18426222](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18426222)
13. Srivastava D, Zhu W, Johnson Jr. WH, Whitman CP, Becker DF, Tanner JJ (2010) The 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. *Biochemistry* [Internet] 49:560–569. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=19994913](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19994913)
14. Scott GK, Mahoney S, Scott M, Loureiro A, Lopez-Ramirez A, Tanner JJ, Ellerby LM, Benz CC (2021) N-Propargylglycine: a unique suicide inhibitor of proline dehydrogenase with anticancer activity and brain-enhancing mitohormesis properties. *Amino Acids* [Internet]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/34089390>
15. Campbell AC, Becker DF, Gates KS, Tanner JJ (2020) Covalent Modification of the Flavin in Proline Dehydrogenase by Thiazolidine-2-Carboxylate. *ACS Chem Biol* [Internet] 15:936–944. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/32159324>
16. Campbell AC, Prater AR, Bogner AN, Quinn TP, Gates KS, Becker DF, Tanner JJ (2021) Photoinduced Covalent Irreversible Inactivation of Proline Dehydrogenase by S-Heterocycles. *ACS Chemical Biology* [Internet]. Available from: <https://pubmed.ncbi.nlm.nih.gov/34542291/>
17. Christensen EM, Bogner AN, Vandekeere A, Tam GS, Patel SM, Becker DF, Fendt SM, Tanner JJ (2020) In crystallo screening for proline analog inhibitors of the proline cycle enzyme PYCR1. *J Biol Chem* [Internet] 295:18316–18327. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/33109600>

18. Milne K, Sun J, Zaal EA, Mowat J, Celie PHN, Fish A, Berkers CR, Forlani G, Loayza-Puch F, Jamieson C, et al. (2019) A fragment-like approach to PYCR1 inhibition. *Bioorg Med Chem Lett* [Internet] 29:2626–2631. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/31362921>
19. Forlani G, Sabbioni G, Ragno D, Petrollino D, Borgatti M (2021) Phenyl-substituted aminomethylene-bisphosphonates inhibit human P5C reductase and show antiproliferative activity against proline-hyperproducing tumour cells. *J Enzyme Inhib Med Chem* [Internet] 36:1248–1257. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/34107832>
20. Nadaraia S, Lee YH, Becker DF, Tanner JJ (2001) Crystallization and preliminary crystallographic analysis of the proline dehydrogenase domain of the multifunctional PutA flavoprotein from *Escherichia coli*. *Acta Crystallogr.* [Internet] D57:1925–1927. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=11717519](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11717519)
21. Zhang W, Zhang M, Zhu W, Zhou Y, Wanduragala S, Rewinkel D, Tanner JJ, Becker DF (2007) Redox-induced changes in flavin structure and roles of flavin N(5) and the ribityl 2'-OH group in regulating PutA-membrane binding. *Biochemistry* [Internet] 46:483–491. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17209558](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17209558)
22. Moxley MA, Tanner JJ, Becker DF (2011) Steady-state kinetic mechanism of the proline:ubiquinone oxidoreductase activity of proline utilization A (PutA) from *Escherichia coli*. *Arch. Biochem. Biophys.* [Internet] 516:113–120. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22040654>
23. Tanner JJ (2019) Structural biology of proline catabolic enzymes. *Antioxidants and Redox Signaling* 30:650–673.
24. Kabsch W (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 66:125–132. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=20124692](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20124692)
25. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 69:1204–1214. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23793146>
26. Ostrander EL, Larson JD, Schuermann JP, Tanner JJ (2009) A conserved active site tyrosine residue of proline dehydrogenase helps enforce the preference for proline over hydroxyproline as the substrate. *Biochemistry* [Internet] 48:951–959. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=19140736](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19140736)

27. Zhu W, Haile AM, Singh RK, Larson JD, Smithen D, Chan JY, Tanner JJ, Becker DF (2013) Involvement of the beta3-alpha3 loop of the proline dehydrogenase domain in allosteric regulation of membrane association of proline utilization A. *Biochemistry* [Internet] 52:4482–4491. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23713611>
  
28. Adams PD, Afonine P v, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 66:213–221. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=20124702](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20124702)
  
29. Afonine P v, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 68:352–367. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22505256>
  
30. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Cryst. D Biol. Crystallogr.* [Internet] 66:486–501. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20383002>
  
31. Moriarty NW, Grosse-Kunstleve RW, Adams PD (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 65:1074–1080. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19770504>
  
32. Chen VB, Arendall 3rd WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] D66:12–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20057044>
  
33. Gore S, Sanz Garcia E, Hendrickx PMS, Gutmanas A, Westbrook JD, Yang H, Feng Z, Baskaran K, Berrisford JM, Hudson BP, et al. (2017) Validation of Structures in the Protein Data Bank. *Structure* [Internet] 25:1916–1927. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/29174494>
  
34. Liebschner D, Afonine P v, Moriarty NW, Poon BK, Sobolev O v, Terwilliger TC, Adams PD (2017) Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol* [Internet] 73:148–157. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28177311>
  
35. Campbell AC, Bogner AN, Mao Y, Becker DF, Tanner JJ (2021) Structural analysis of prolines and hydroxyprolines binding to the l-glutamate-gamma-semialdehyde

dehydrogenase active site of bifunctional proline utilization A. Arch Biochem Biophys [Internet] 698:108727. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/33333077>

36. Singh H, Arentson BW, Becker DF, Tanner JJ (2014) Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-channeling tunnel and the quinone-binding site. Proceedings of the National Academy of Sciences of the United States of America [Internet] 111:3389–3394. Available from: <https://pubmed.ncbi.nlm.nih.gov/24550478/>

37. Geng P, Qin W, Xu G (2021) Proline metabolism in cancer. Amino Acids 2021 [Internet]:1–9. Available from: <https://link.springer.com/article/10.1007/s00726-021-03060-1>

38. Cockroft SL, Hunter CA (2007) Chemical double-mutant cycles: Dissecting non-covalent interactions. Chemical Society Reviews [Internet] 36:172–188. Available from: <https://pubmed.ncbi.nlm.nih.gov/17264921/>

**Table 2.1.** Inhibition and kinetic constants

Compound	$K_i$ (mM)	$K_m$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_m$ ( $M^{-1}s^{-1}$ )
<b>1</b>	$0.28 \pm 0.02$	$86.6 \pm 7.84$	$1.16 \pm 0.03$	$13.0 \pm 1.3$
<b>2</b>	$1.5 \pm 0.2$	$123.8 \pm 20.3$	$0.90 \pm 0.05$	$7.2 \pm 1.6$
<b>3</b>	$1.9 \pm 0.2$	$121.0 \pm 11.6$	$1.47 \pm 0.05$	$12.2 \pm 1.6$
<b>4</b>	$3.1 \pm 0.3$	$87.5 \pm 8.3$	$1.20 \pm 0.03$	$13.7 \pm 1.7$
<b>16</b>	$6.3 \pm 0.7$	$118.6 \pm 10.4$	$1.35 \pm 0.04$	$11.4 \pm 1.3$
<b>17</b>	$1.8 \pm 0.1$	$84.5 \pm 6.1$	$1.14 \pm 0.02$	$13.5 \pm 0.1$
<b>18</b>	$1.4 \pm 0.1$	$136.2 \pm 10.7$	$1.42 \pm 0.04$	$10.4 \pm 0.9$

**Table 2.2.** X-ray diffraction data collection and refinement statistics

	Compound 2	Compound 3	Compound 4
Space group	<i>I</i> 222	<i>I</i> 222	<i>I</i> 222
Unit cell parameters (Å)	<i>a</i> = 72.80 <i>b</i> = 141.57 <i>c</i> = 144.97	<i>a</i> = 73.08 <i>b</i> = 141.82 <i>c</i> = 146.64	<i>a</i> = 72.99 <i>b</i> = 141.74 <i>c</i> = 146.30
Wavelength (Å)	0.97918	0.97918	0.97918
Resolution (Å)	101.26 - 2.31 (2.31 - 2.19)	101.94 - 1.71 (1.71 - 1.68)	73.15 - 1.72 (1.72 - 1.69)
Observations <sup>a</sup>	183109 (26444)	631755 (23468)	639699 (29360)
Unique reflections <sup>a</sup>	38519 (5542)	85566 (3890)	85319 (4112)
<i>R</i> <sub>merge</sub> ( <i>I</i> ) <sup>a</sup>	0.110 (2.344)	0.076 (2.843)	0.119 (2.989)
<i>R</i> <sub>meas</sub> ( <i>I</i> ) <sup>a</sup>	0.139 (2.764)	0.082 (3.105)	0.127 (3.218)
<i>R</i> <sub>pim</sub> ( <i>I</i> ) <sup>a</sup>	0.063 (1.247)	0.030 (1.224)	0.045 (1.170)
Mean <i>I</i> /σ <sup>a</sup>	10.3 (1.0)	15.0 (0.6)	15.2 (0.6)
CC <sub>1/2</sub> <sup>a</sup>	0.997 (0.234)	0.998 (0.333)	0.999 (0.200)
Completeness (%) <sup>a</sup>	99.4 (99.7)	99.4 (88.8)	99.7 (94.6)
Multiplicity <sup>a</sup>	4.8 (4.8)	7.4 (6.0)	7.5 (7.1)
No. of protein residues	499	499	499
No. of atoms			
Protein	3708	3740	3732
FAD	53	53	53
Inhibitor	10	7	6
Water	57	219	247
<i>R</i> <sub>cryst</sub> <sup>a</sup>	0.223 (0.386)	0.207 (0.645)	0.213 (0.658)
<i>R</i> <sub>free</sub> <sup>a,b</sup>	0.257 (0.404)	0.234 (0.574)	0.240 (0.653)
rmsd bonds (Å)	0.007	0.007	0.006
rmsd angles (°)	0.874	0.869	0.859
Ramachandran plot <sup>c</sup>			
Favored (%)	96.75	98.78	97.57
Outliers (%)	0.00	0.00	0.00
Clashscore (PR) <sup>c</sup>	3.92 (99)	1.86 (99)	1.60 (99)
Molprobity score (PR) <sup>c</sup>	1.76 (94)	1.22 (98)	1.11 (99)
Average <i>B</i> (Å <sup>2</sup> )			
Protein	69.1	43.7	40.3
FAD	48.2	26.8	23.5
Inhibitor	53.3	30.1	31.4
Water	50.7	40.3	37.9
Coord. error (Å) <sup>d</sup>	0.38	0.32	0.34
PDB code	7MWT	7MWU	7MWV

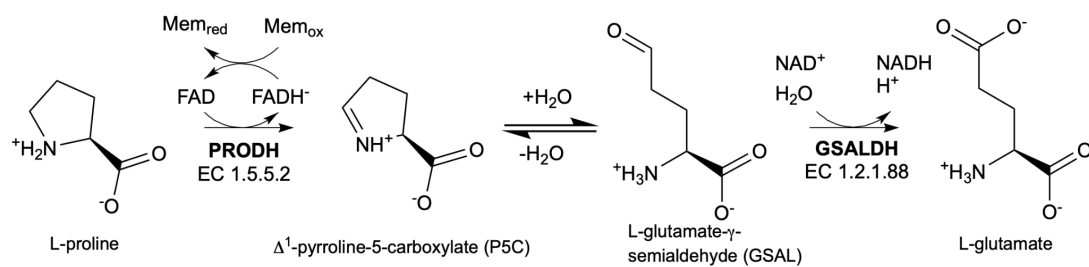
<sup>a</sup>Values for the outer resolution shell of data are given in parenthesis.

<sup>b</sup>5% test set.

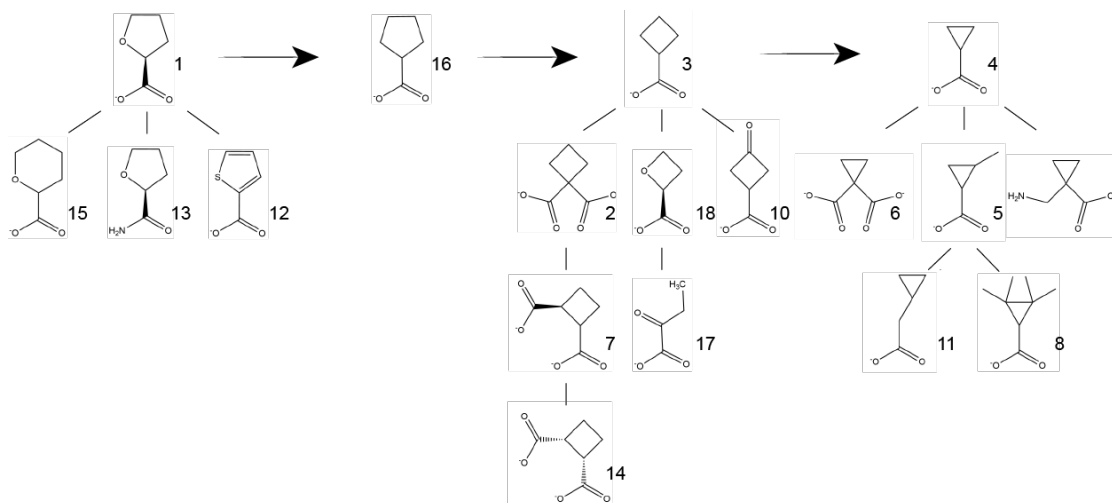
<sup>c</sup>From MolProbity. The percentile ranks (PR) for Clashscore and MolProbity score are given in parentheses.

<sup>d</sup>Maximum likelihood-based coordinate error estimate from PHENIX.

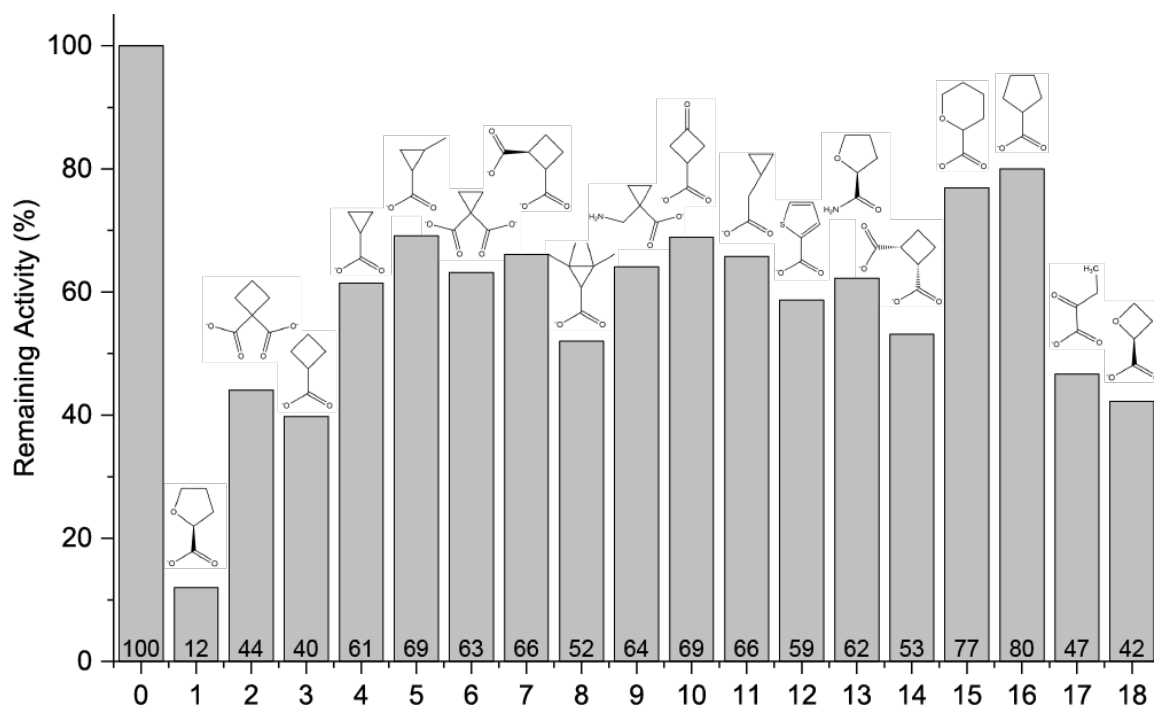




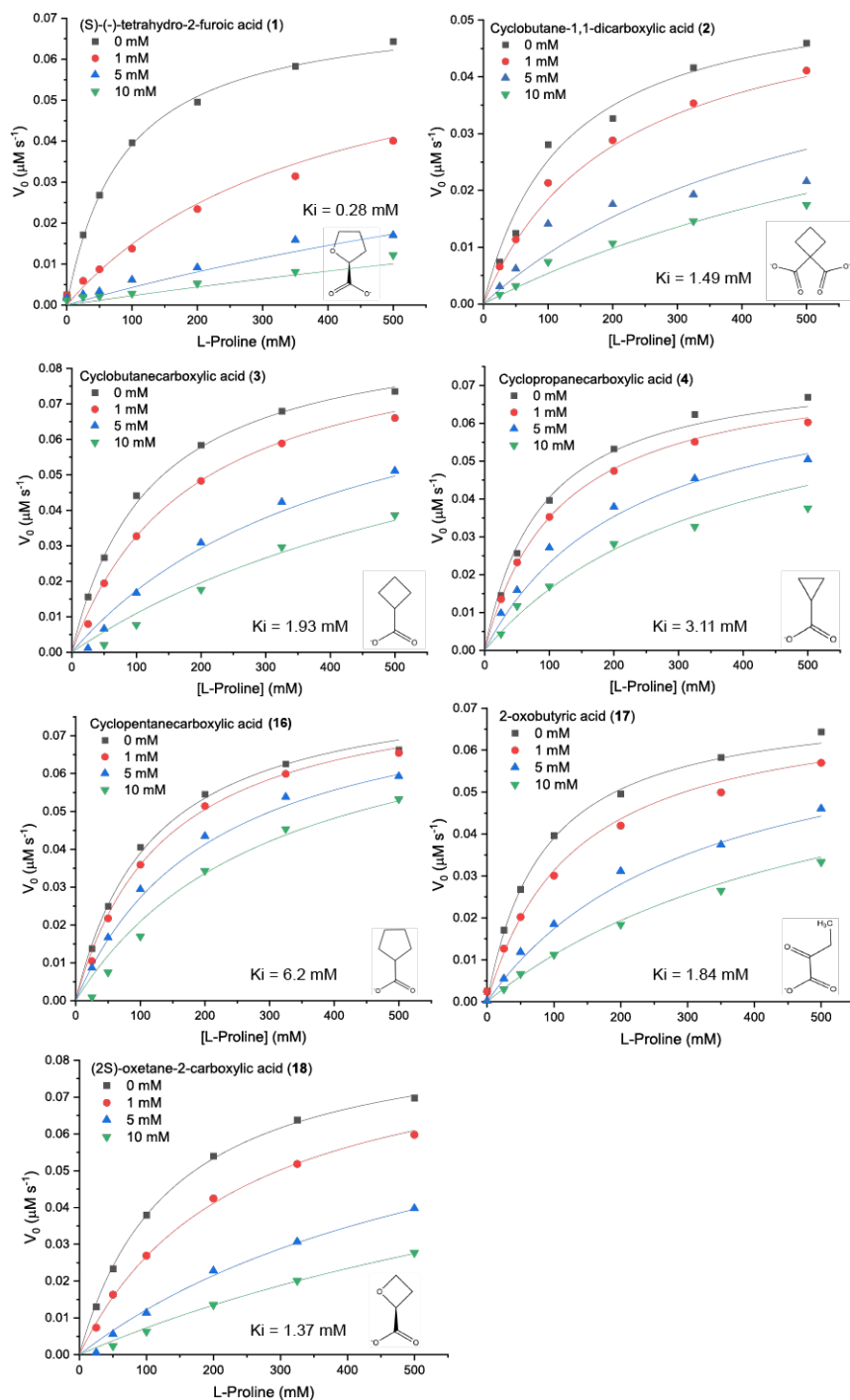
**Figure 2.1.** Enzymes and reactions of proline catabolism.



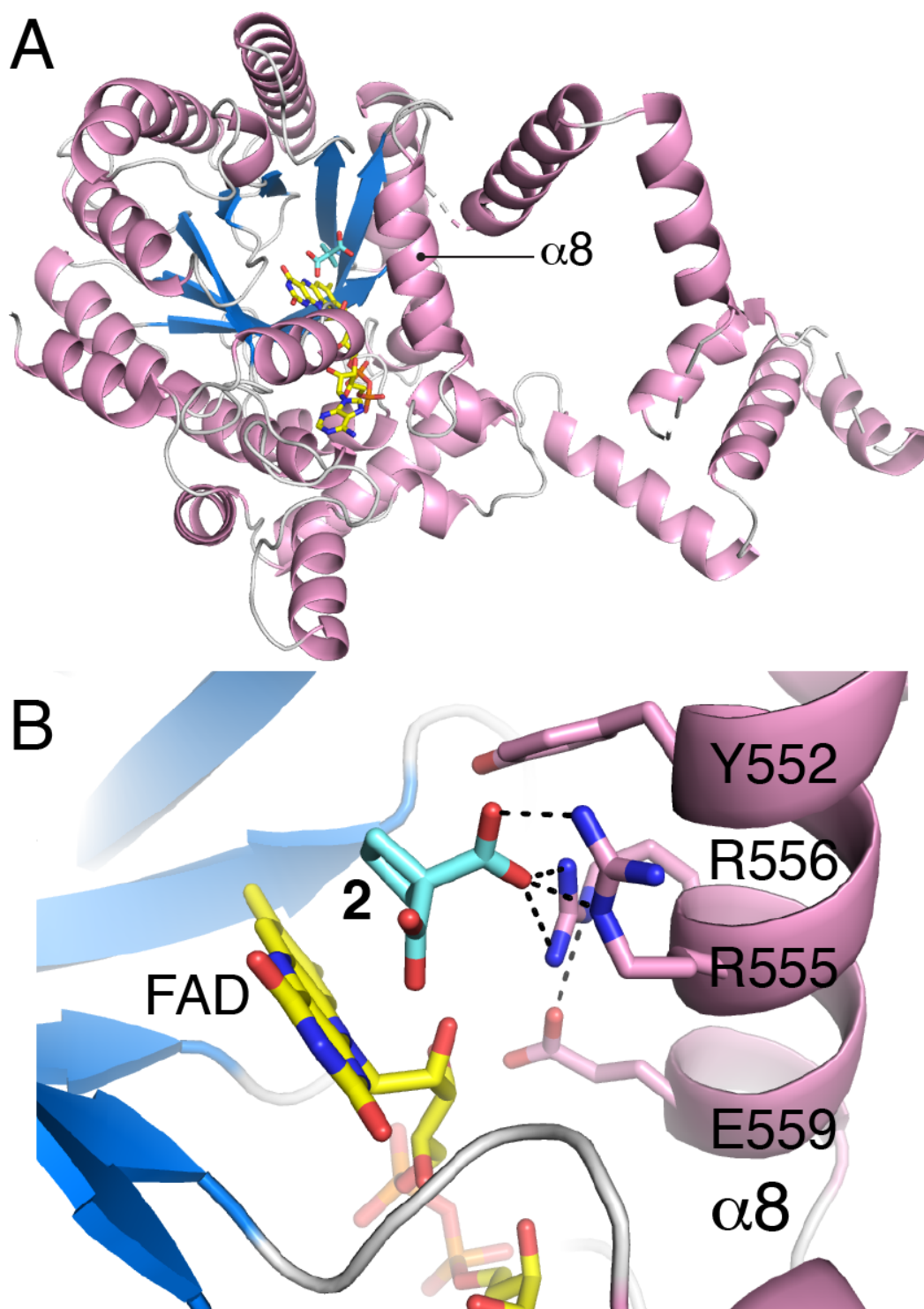
**Figure 2.2.** Chemical structures of the inhibitors used in this study.



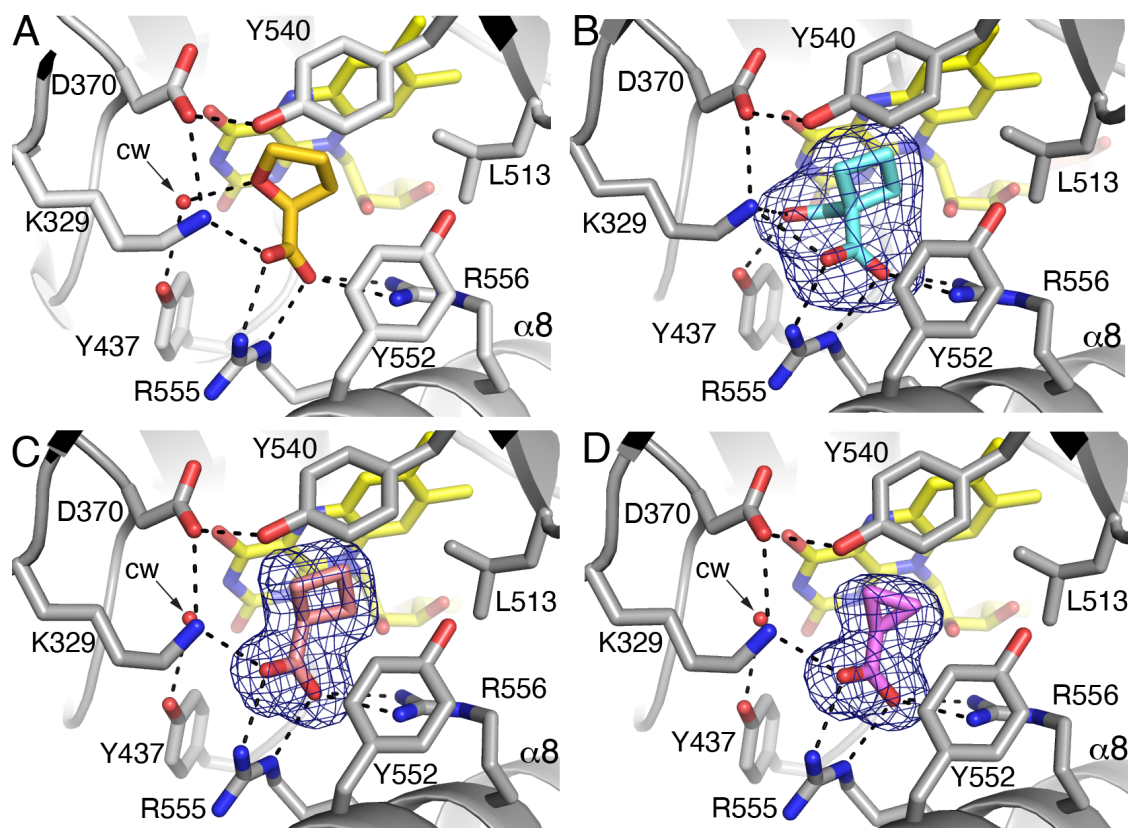
**Figure 2.3.** Initial screening of compounds for inhibition of PRODH activity. The bars represent the enzyme activity of PutA86-630 measured at a single L-proline substrate concentration (200 mM) and the compound present at 5 mM. The data are normalized to the activity in the absence of any inhibitor (compound 0). Compound 1 was included as a positive control. The percent activity relative to the no-inhibitor control is labeled above the x-axis.



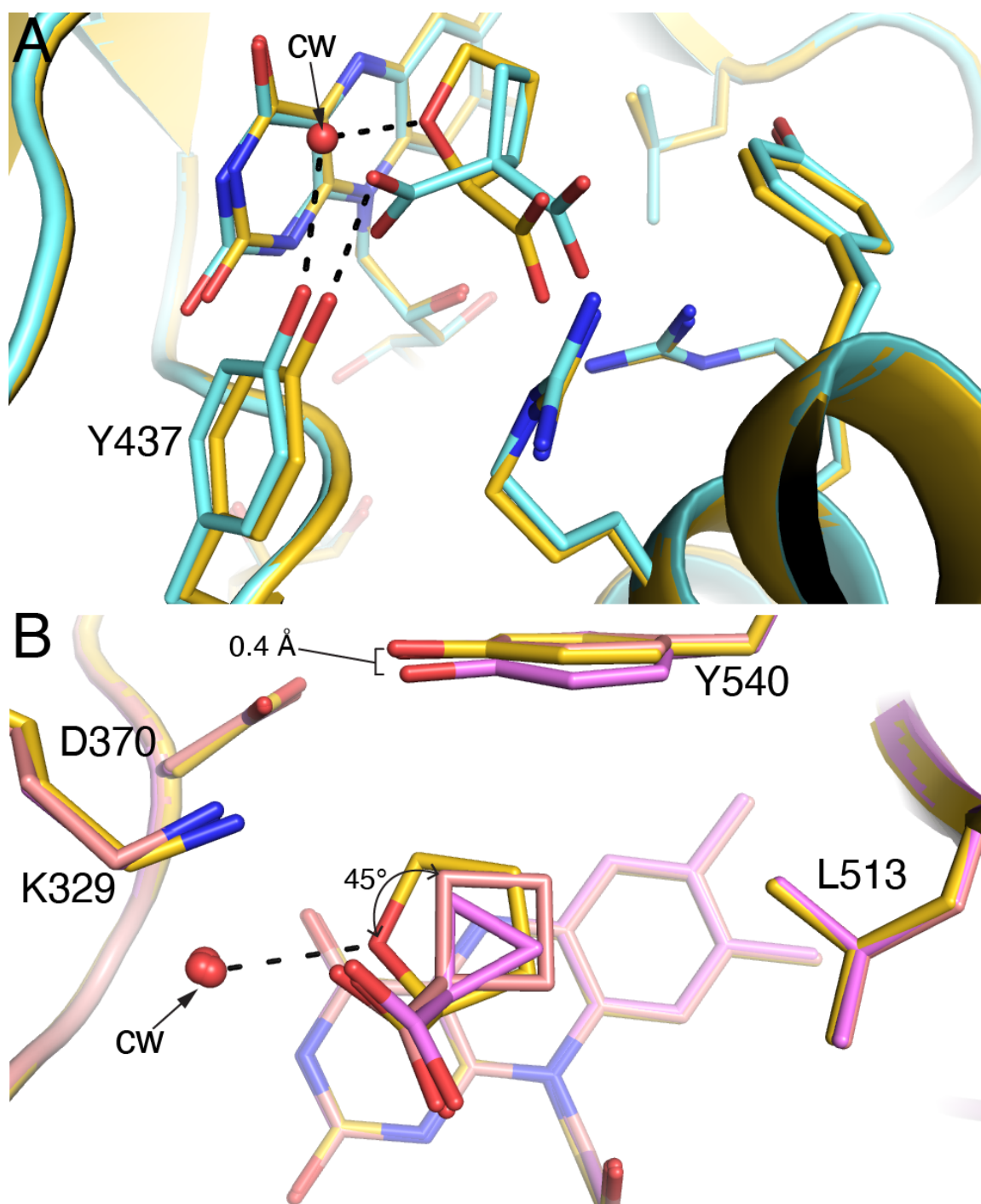
**Figure 2.4.** Inhibition of the PROD H activity of PutA86-630 by **1**, **2**, **3**, **4**, **16**, **17**, and **18**. The assays were performed at room temperature with 0-500 mM L-proline, 4 mM *o*-AB, 0.15 mM menadione, and 63 nM PutA86-630 in a buffer containing 20 mM MOPS pH 7.5 and 10 mM  $\text{MgCl}_2$ . The data for each inhibitor were analyzed by global fitting to the competitive inhibition model using Origin software. Kinetic constants are listed in Table 2.1.



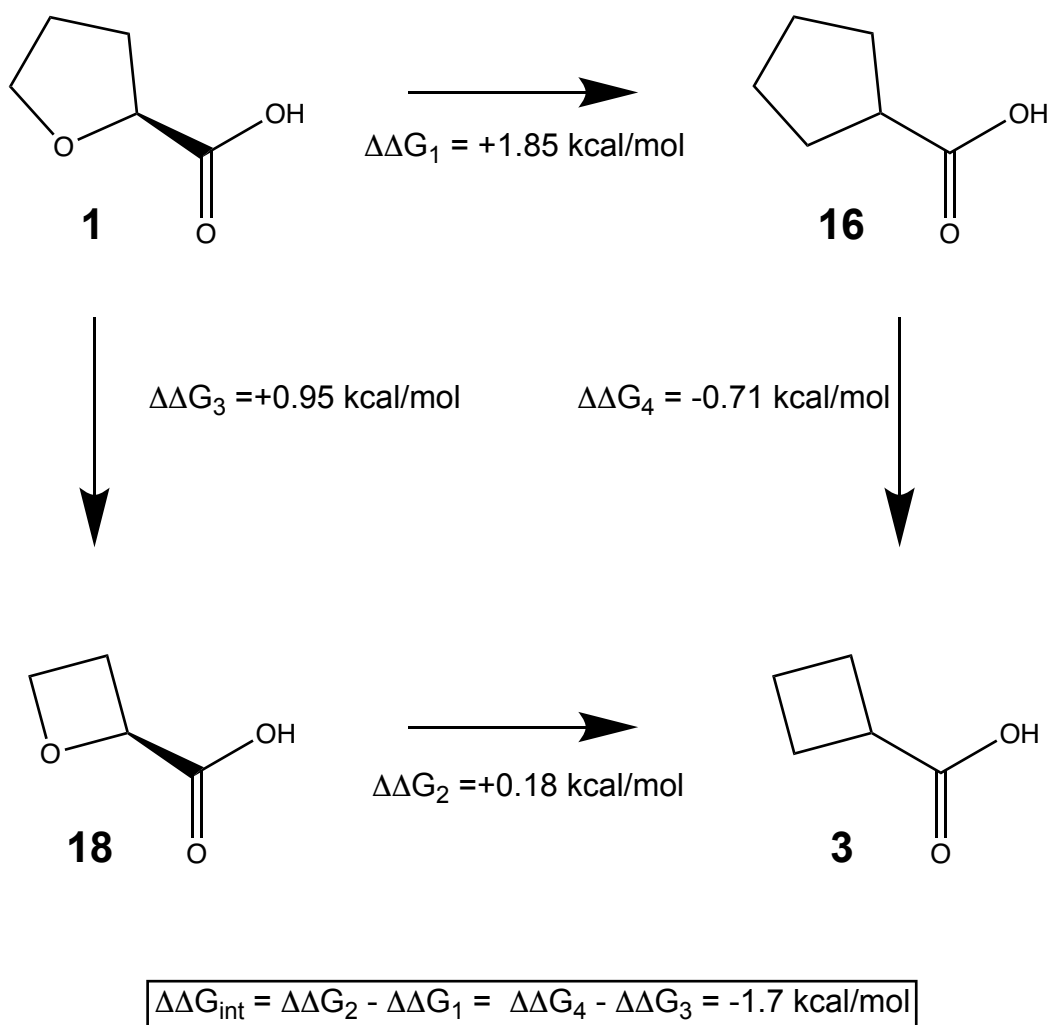
**Figure 2.5.** Structure of PutA86-630 complexed with 2. (A) Cartoon drawing showing the location of the active site within the protein fold. (B) Close-up view of the active site showing 2 and side chains of the YXXRRXXET/N motif of  $\alpha 8$ .



**Figure 2.6.** Interactions and electron density for inhibitors bound to PutA86-630. (A) **1**, from PDB ID 1TIW<sup>9</sup>. (B) Compound **2**. (C) Compound **3**. (D) Compound **4**.



**Figure 2.7.** Superposition of the PutA86-630-inhibitor complexes. (A) The active site complexed with **1** (gold, PDB ID 1TIW<sup>9</sup>) and **2** (light blue). (B) The active site complexed with **1** (gold, PDB ID 1TIW<sup>9</sup>), **3** (salmon), and **4** (pink). The curved arrow denotes the difference in orientation of the rings of **3/4** relative to compound **1**.



**Figure 2.8.** Chemical double mutant cycle expressing the relationship between hydrogen bonding capacity and ring size of **1**. The  $\Delta\Delta G$  values for the chemical mutations were calculated from the  $K_i$  values of the inhibitors.



## Chapter 3

### Structural Basis for the Stereospecific Inhibition of the Dual Proline/Hydroxyproline Catabolic Enzyme ALDH4A1 by Trans-4- Hydroxy-L-Proline

Alexandra N. Bogner,<sup>1</sup> Kyle M. Stiers,<sup>1</sup> Cole McKay,<sup>1</sup> Donald F. Becker,<sup>2</sup> and John J. Tanner<sup>1,3</sup>

<sup>1</sup>Department of Biochemistry, University of Missouri, Columbia, Missouri 65211, United States

<sup>2</sup>Department of Biochemistry, Redox Biology Center, University of Nebraska, Lincoln, Nebraska

<sup>3</sup>Department of Chemistry, University of Missouri, Columbia, Missouri 65211, United States

#### Author Contributions

A.N.B.: Conceptualization; investigation; methodology; visualization; writing-original draft; writing- review & editing. K.M.S.: Conceptualization; investigation; methodology; writing-original draft; writing- review & editing. C.M.M.: Investigation. D.F.B.: Funding acquisition; project administration; writing-review & editing. J.J.T.: Conceptualization; funding acquisition; project administration; supervision; validation; visualization; writing-original draft; writing-review & editing.

**Keywords:** aldehyde dehydrogenase, ALDH4A1, enzyme inhibition, hydroxyproline catabolism, L-glutamate- $\gamma$ -semialdehyde dehydrogenase, proline metabolism, X-ray crystallography

**Accession Codes:** Coordinates and structure factor amplitudes have been deposited in the Protein Data Bank under accession codes 7MER and 7MES.

**Abbreviations:** 3-OH-P5C,  $\Delta^1$ -pyrroline-3-hydroxy-5-carboxylic acid; ALDH, aldehyde dehydrogenase; CHDP, cis-4-hydroxy-D-proline; CHLP, cis- 4-hydroxy-L-proline; GSAL, L-glutamate- $\gamma$ -semialdehyde; GSALDH, L-glutamate- $\gamma$ -semialdehyde dehydrogenase; MmALDH4A1, mouse ALDH4A1; OH-GSAL, 4-hydroxy-L-glutamate- $\gamma$ -semialdehyde; P5C,  $\Delta^1$ -pyrroline-5-carboxylate; PH, hyperoxaluria; PRODH, proline dehydrogenase; THDP, trans-4-hydroxy-D-proline.; THLP, trans-4-hydroxy-L-proline.

This chapter was adapted from our publication in *Protein Science*, Structural Basis for the Stereospecific Inhibition of the Dual Proline/Hydroxyproline Catabolic Enzyme ALDH4A1 by Trans-4-Hydroxy-L-Proline, Bogner AN, Stiers KM, McKay CM, Becker

DF, Tanner JJ, **2021**, (8):1714-1722. doi: 10.1002/pro.4131 and our publication in *Archives of Biochemistry and Biophysics*, Structural Analysis of Prolines and Hydroxyprolines Binding to the L-Glutamate- $\gamma$ -Semialdehyde Dehydrogenase Active Site of Bifunctional Proline Utilization A, Campbell AC, Bogner AN, Mao Y, Becker DF, Tanner JJ, **2021**, 15;698:108727. doi: 10.1016/j.abb.2020.108727

## ABSTRACT

Aldehyde dehydrogenase 4A1 (ALDH4A1) catalyzes the final steps of both proline and hydroxyproline catabolism. It is a dual substrate enzyme that catalyzes the NAD<sup>+</sup>-dependent oxidations of L-glutamate- $\gamma$ -semialdehyde to L-glutamate (proline metabolism), and 4-hydroxy-L-glutamate- $\gamma$ -semialdehyde to 4-erythro-hydroxy-L-glutamate (hydroxyproline metabolism). Here we investigated the inhibition of mouse ALDH4A1 by the six stereoisomers of proline and 4-hydroxyproline using steady-state kinetics and X-ray crystallography. Trans-4-hydroxy-L-proline is the strongest of the inhibitors studied, characterized by a competitive inhibition constant of 0.7 mM, followed by L-proline (1.9 mM). The other compounds are very weak inhibitors (approximately 10 mM or greater). Insight into the selectivity for L-stereoisomers was obtained by solving crystal structures of ALDH4A1 complexed with trans-4-hydroxy-L-proline and trans-4-hydroxy-D-proline. The structures suggest that the 10-fold greater preference for the L-stereoisomer is due to a serine residue that hydrogen bonds to the amine group of trans-4-hydroxy-L-proline. In contrast, the amine group of the D-stereoisomer lacks a direct interaction with the enzyme due to a different orientation of the pyrrolidine ring. These results suggest that hydroxyproline catabolism is subject to substrate inhibition by trans-4-hydroxy-L-proline, analogous to the known inhibition of proline catabolism by L-proline. Also, drugs targeting the first enzyme of hydroxyproline catabolism, by elevating the level of trans-4-hydroxy-L-proline, may inadvertently impair proline catabolism by the inhibition of ALDH4A1.

## INTRODUCTION

Aldehyde dehydrogenase 4A1 (ALDH4A1) is a dual-substrate enzyme that catalyzes the final steps of both proline- and hydroxyproline catabolism in mammals (Figures 3.1A, 3.1B).<sup>1, 2</sup> The first steps of these pathways are catalyzed by distinct proline dehydrogenase (PRODH) enzymes. The proline-specific PRODH (PRODH1) catalyzes the FAD-dependent oxidation of L-proline to  $\Delta^1$ -pyrroline-5-carboxylate (P5C), while the hydroxyproline-specific PRODH (PRODH2, 39% identical to PRODH1) analogously catalyzes the oxidation of trans-4-hydroxy-L-proline (THLP) to  $\Delta^1$ -pyrroline-3-hydroxy-5-carboxylate (3-OH-P5C). The presumed nonenzymatic hydrolysis of the imine products of PRODH1/2 generates the semialdehyde substrates for ALDH4A1, which catalyzes the NAD<sup>+</sup>-dependent oxidations of both L-glutamate- $\gamma$ -semialdehyde (GSAL) to L-glutamate, and 4-hydroxy-L-glutamate- $\gamma$ -semialdehyde (OH-GSAL) to 4-erythro-hydroxy-L-glutamate.

Interest in ALDH4A1 stems from the involvement of proline and hydroxyproline metabolism in many aspects of human health and disease. Proline metabolism plays a central role in the altered metabolism of cancer cells,<sup>3-7</sup> and both the *PRODH1* and *ALDH4A1* genes are regulated by the tumor suppressor p53.<sup>8, 9</sup> Inherited mutations in the *ALDH4A1* gene cause hyperprolinemia II, a metabolic disorder that can result in neurological problems, including intellectual disability.<sup>10, 11</sup> Ligands of ALDH4A1 could potentially be used as pharmacological chaperones to stabilize the misfolded variants of ALDH4A1 produced in patients with hyperprolinemia II.<sup>12, 13</sup> A recent study discovered that ALDH4A1 was significantly elevated in the plasma of atherosclerosis-prone *Ldlr*<sup>-/-</sup> mice as well as atherosclerotic tissue from humans, and the administration of an anti-

ALDH4A1 antibody delayed atherosclerosis progression in *Ldlr*<sup>-/-</sup> mice.<sup>14</sup> These results suggest that anti-ALDH4A1 antibodies, and potentially ALDH4A1 inhibitors, may have therapeutic value in cardiovascular disease. Another recent study showed that disruption of *ALDH4A1* in *Caenorhabditis elegans* decreased the quality and quantity of sperm in males, which was linked to a buildup of P5C and aberrant ROS homeostasis.<sup>15, 16</sup> Finally, hydroxyproline catabolism is a promising target for the development of drugs to treat primary hyperoxaluria (PH), an autosomal recessive disorder associated with excess oxalate production and increased risk of calcium oxalate stone formation.<sup>17, 18</sup>

ALDH4A1 belongs to the ALDH structural superfamily<sup>19</sup> and has been characterized structurally and biochemically. In the literature, it is also known as L-glutamate- $\gamma$ -semialdehyde dehydrogenase (GSALDH) and  $\Delta^1$ -pyrroline-5-carboxylate dehydrogenase, particularly in reference to the bacterial enzymes. In some bacteria, GSALDH is combined with PRODH into the bifunctional enzyme known as proline utilization A (PutA).<sup>20</sup> Several crystal structures of ALDH4A1 from eukaryotic and bacterial sources have been determined, including structures of the enzyme complexed with the product L-glutamate, cofactor NAD<sup>+</sup>, L-proline, and several aliphatic dicarboxylic acids.<sup>21-23</sup> Also, the kinetic mechanism of human ALDH4A1 has been investigated and found to follow a compulsory ordered mechanism in which NAD<sup>+</sup> binds before GSAL, and L-glutamate dissociates before NADH.<sup>21, 24</sup>

Here we investigated the inhibition of mouse ALDH4A1 (MmALDH4A1) by D,L-proline and the four stereoisomers of 4-hydroxyproline (Figure 3.1C). Trans-4-hydroxy-L-proline (THLP) is the strongest of the inhibitors studied, characterized by a competitive (with P5C) inhibition constant of 0.7 mM, followed by L-proline ( $K_i$  = 1.9 mM). The other

compounds are very weak inhibitors ( $K_i$  of 10 mM or greater). Insight into the selectivity for L-stereoisomer was obtained by solving crystal structures of MmALDH4A1 complexed with THLP and trans-4-hydroxy-D-proline (THDP). The 10-fold greater preference for THLP is attributed to hydrogen bonding of the inhibitor amine group with a serine residue in the substrate anchor loop. For comparison, we also tested inhibition of the four stereoisomers of 4-hydroxyproline against the GSALDH domain of the bifunctional enzyme proline utilization A from *Sinorhizobium meliloti* (SmPutA). Interestingly, the four stereoisomers of 4-hydroxyproline showed weaker affinity (millimolar range) and different stereoselectivity for SmPutA. In contrast to ALDH4A1, for SmPutA the D-isomer is preferred over the L-isomer and cis configuration is preferred over trans. The difference in stereoselectivity of inhibition is consistent with the absence of a serine residue in the anchor loop of SmPutA.

## RESULTS

**Inhibition of ALDH4A1 by the Stereoisomers of Proline and 4-Hydroxyproline.** The inhibition of MmALDH4A1 by D,L-proline and the four stereoisomers of 4-hydroxyproline was investigated with kinetics assays using L-P5C as the variable substrate and  $\text{NAD}^+$  fixed at the saturating concentration of 1 mM (Figure 3.2). The data were fit to the competitive, uncompetitive, and mixed models of inhibition (Table S3.1). For all but THDP, the data could be fit satisfactorily to the competitive inhibition model. The uncompetitive model yielded an improved fit for the THDP data, based on inspection of the adjusted  $R^2$  value (Table S3.1). Double-reciprocal plot analysis of the THDP data yielded a set of lines that did not intersect at a common point on the vertical axis,

confirming a deviation from classical competitive inhibition (Figure S3.1). Table 3.1 lists the results from the best model for each inhibitor.

Two of the L-stereoisomers, THLP and L-proline, produced the highest apparent inhibition, characterized by inhibition constants ( $K_i$ ) of 0.7 mM and 1.9 mM, respectively (Table 3.1). THDP is a much weaker inhibitor, with  $K_i$  of 9 mM. The other three compounds produced no significant reduction of enzyme activity, even at the highest inhibitor concentration tested (10 mM). Thus, with the exception of CHLP, ALDH4A1 appears to show a preference for binding L-stereoisomers of proline and 4-hydroxyproline over the D- stereoisomers.

**Inhibition of the GSALDH Activity of SmPutA by the Stereoisomers of Proline and 4-Hydroxyproline.** The GSALDH activity of SmPutA was measured with L-P5C as the variable substrate and  $\text{NAD}^+$  at fixed concentration. Since the proline compounds bind to the P5C/GSAL binding site,<sup>29</sup> the data were fit initially to the competitive model (Figure S3.2). This analysis suggested that CHLP does not inhibit the GSALDH activity of SmPutA, and the other compounds are weak inhibitors with inhibition constants in the range of 2 – 11 mM (Table 3.2). For reference, the  $K_m$  for L-P5C is  $0.3 \pm 0.1$  mM. The quality of the fits to the competitive model was not optimal, so other inhibition models were tested. Improved fit quality was obtained when applying the uncompetitive inhibition model for D-proline and L-proline, and the mixed inhibition model for THDP, CHDP, and THLP (Figure 3.3). The resulting inhibition parameters are in the millimolar range, consistent with these compounds being weak inhibitors (Table 3.2). We note that  $\alpha$  is close to one for the mixed inhibitors, implying the uncompetitive component is significant for these compounds (Table 3.2).

**Structural Basis of the Preference for L-Stereoisomers.** Crystal structures of MmALDH4A1 complexed with THLP ( $K_i = 0.7$  mM) and THDP ( $K_i = 9$  mM) were determined to understand the basis for the apparent 10-fold preference for the former ligand. The structures were determined at high resolution limits of 1.74 Å (THLP) and 1.37 Å (THDP) (Table 3.3). Attempts to obtain crystal structures with the other compounds were unsuccessful, presumably because of their very low affinities. We note that a structure of MmALDH4A1 complexed with L-proline was previously reported by our lab.<sup>23</sup>

The electron density maps clearly defined the poses of THLP and THDP in the active site (Figure 3.4). The occupancy of THLP refined to 0.78 and 0.85 in chains A and B of the dimer, respectively. Those of THDP refined to 0.86 and 0.82. The THDP complex also has electron density for NAD<sup>+</sup>, which was added during crystallization and cryoprotection. The density was strong for the AMP portion of NAD<sup>+</sup> but weak and diffuse for the nicotinamide mononucleotide portion; therefore, only the AMP part was included in the final model. Density for NAD<sup>+</sup> in the THLP complex was weak and the cofactor was not included in the model. Disorder in NAD<sup>+</sup> bound to ALDH4A1 is common.<sup>21, 25</sup>

The hydroxyproline ligands occupy the GSAL site, whose location is known from a previous structure of ALDH4A1 complexed with the product, L-glutamate (Figure 3.3).<sup>21</sup> The recognition elements of the site include the anchor loop (residues 511-513), which binds the amine and carboxylate groups of the substrate, an aromatic box<sup>26</sup> consisting of Phe212 and Phe520, which clamps the aliphatic chain of GSAL, and the charged residues, Glu165 and Lys347, which provide electrostatic compensation for the substrate amine and carboxylate groups, respectively.



The poses of the two inhibitors share some similarities (Figure 3.4A, 3.4B). For example, the carboxylate groups of THLP and THDP form direct hydrogen bonds with backbone amine groups of the anchor loop, Ser513, and Ser349, as well as a water-mediated hydrogen bond with Lys347 of the catalytic loop. And in both cases, the 4-hydroxyl group hydrogen bonds to Glu165.

The poses of THLP and THDP differ in the orientation of the pyrrolidine ring in the aromatic box, which results in different hydrogen bonding opportunities for the amine group (Figure 3.4C). In THLP, the amine is directed toward the anchor loop and forms hydrogen bonds with Ser513 and Glu165 (Figure 3.4A). In contrast, the amine of THDP faces away from the anchor loop and lacks a hydrogen bond with Ser513 (Figure 3.4B). Also, the carboxylate of THLP has more hydrogen bonds with the backbone amine groups of the anchor loop (three for THLP, versus one for THDP). Thus, the main difference between the two complexes is that the amine and carboxylate groups of THLP enjoy greater interactions with the protein. These extra interactions may explain the greater affinity for THLP.

**Remote Binding Site for THDP.** The THDP structure revealed a secondary binding site for the inhibitor. The remote site is located on the rim of the tunnel to the GSAL binding site, which is presumably the path travelled by aldehyde substrates and competitive inhibitors on their way to the active site (Figure 3.5A, 3.5B). The remote THDP is 15 Å from the THDP bound in the active site. The refined occupancies of THDPs in the remote site are 0.68 and 0.87.

Several noncovalent interactions stabilize THDP in the remote site (Figure 3.5C). The hydrogen bonding potential of the amine and carboxylate groups of THDP are satisfied by

Glu342, Gln157, and Thr154. The C4 and C5 carbon atoms of the pyrrolidine ring pack tightly against the phenyl ring of Phe387. In contrast, the 4-hydroxyl group lacks interactions with the protein. Interestingly, we found no electron density evidence for THLP in the remote site. It may be that steric clash of the 4-hydroxyl group of THLP with Phe387 prevents binding.

Binding to the remote site is accompanied by a large conformational change of the C-terminus of the protein. In the absence of THDP, residues 557-562 block the remote site. In particular, Tyr559 and Met562 invade the space occupied by remote THDP (Figure 3.5D). In the presence of THDP, electron density for the C-terminal nine residues (555-563) is very weak, implying conformational disorder. Although the electron density for these residues is strong in all other structures of MmALDH4A1, apparently the C-terminus is flexible enough to allow the binding of THDP.

## **DISCUSSION**

We showed that the substrate of the first enzyme of hydroxyproline catabolism inhibits the second (and final) enzyme of the pathway. This is a form of substrate inhibition and is analogous to the better-known inhibition of proline catabolism by L-proline. For example, substrate inhibition of proline catabolism has been studied in the bifunctional PRODH-GSALDH enzyme, PutA. The basis for inhibition of the coupled PRODH-GSALDH reaction of PutA is the binding of L-proline in the GSALDH active site,<sup>27</sup> similar to what we described here with THLP. PutAs are present only in bacteria, and the inhibition of PutA by proline may be advantageous during osmotic stress, when bacteria need to accumulate high levels of proline rather than catabolizing it. Whether substrate inhibition

of hydroxyproline catabolism is physiologically relevant is not known. The concentration of free THLP in humans typically is very low (e.g., <50  $\mu$ M in plasma<sup>11</sup>) compared to the  $K_i$  of THLP for MmALDH4A1 (0.7 mM), suggesting substrate inhibition may not be important. However, in people with hyperhydroxyprolinemia, caused by genetic mutations that impair PRODH2 function, the concentration of THLP can reach 0.5 mM.<sup>11</sup> In these individuals, the inhibition of ALDH4A1 by THLP could be relevant.

MmALDH4A1 was found to have a marked preference for binding L-proline and THLP over D-proline and THDP, respectively, indicating the enzyme specifically recognizes the stereochemistry of the  $\alpha$ -carbon. The structures suggest that Ser513 of the anchor loop plays a role in this aspect of molecular recognition. In THLP, the amine group is directed toward the anchor loop and forms a hydrogen bond with Ser513. This hydrogen bond is also present in the structures of MmALDH4A1 complexed with L-proline and L-glutamate (PDB IDs 3V9K<sup>21</sup> and 4E3X<sup>23</sup>). In contrast, the amine of THDP faces in the opposite direction and lacks the hydrogen bond with Ser513 (Figure 3.4C). We note that the serine at this position in the anchor loop is conserved in mammalian ALDH4A1, whereas it is replaced by alanine in other organisms, including bacteria, fungi, reptiles, and birds, as well as plants, where GSALDH is known as ALDH12.<sup>28</sup> Thus, the stereospecific inhibition observed here is likely a common feature of mammalian ALDH4A1.

The preferred recognition of L-amino acid inhibitors observed with MmALDH4A1 contrasts our experience with a bacterial GSALDH.<sup>29</sup> In fact, the bacterial GSALDH exhibited the opposite trend of the D-stereoisomers of proline and 4-hydroxyproline being better inhibitors than their respective L-stereoisomers. The different outcome likely reflects the replacement of Ser513 with Ala in bacterial GSALDHs, which supports our conclusion

that Ser513 is crucial for the preferred recognition of L-amino acid inhibitors by MmALDH4A1.

A curious result is that the kinetic data for THDP fit better to the uncompetitive model than the competitive model. Uncompetitive inhibition is classically explained by a model in which the inhibitor binds to an enzyme-substrate complex. The structure of the THDP complex revealed two binding sites, the expected one in the GSAL site, and another one in the GSAL tunnel (Figure 3.5). It is possible that the remote site underlies the deviation from classical competitive inhibition kinetics.

Hydroxyproline catabolism is a target for the development of drugs to treat PH,<sup>17, 18</sup> and our results could inform drug design efforts. The product of hydroxyproline catabolism is a precursor for the production of oxalate via glyoxylate metabolism. Certain mutations in the genes encoding the glyoxylate metabolic enzymes alanine:glyoxylate aminotransferase, glyoxylate reductase, and 4-hydroxy-2-oxoglutarate aldolase disable these enzymes and cause oxalate levels to rise, leading to increased susceptibility to calcium oxalate kidney stones, renal inflammation, and urinary tract infections.<sup>30</sup> Inhibition of PRODH2 is being explored as a possible treatment for PH, and hydroxyprolines have shown promise in inhibiting PRODH2 activity<sup>31</sup> and lowering calcium oxalate crystal formation in a fly model of PH.<sup>32</sup> Our results suggest that the therapeutic inhibition of PRODH2, by increasing THLP levels, could also impair ALDH4A1. This may be advantageous for the treatment of PH, because inhibitors of PRODH2 would, in effect, disable the entire hydroxyproline catabolism pathway. Also, the dual inhibition of PRODH2 and ALDH4A1 could impact the therapeutic window, allowing lower drug doses to achieve efficacy. However, because ALDH4A1 also functions in proline catabolism,

PRODH2 inhibitors may also cause P5C levels to increase, which could have negative consequences. These issues should be considered in drug discovery efforts targeting PRODH2.

In conclusion, we have developed structure-activity relationships for the recognition of proline and 4-hydroxyproline stereoisomers by ALDH4A1. Ser513 of the anchor loop endows a preference for binding the L-configuration, and Glu165 forms stabilizing hydrogen bonds with the 4-hydroxyl group, regardless of the stereochemistry of the  $\alpha$ -carbon. These two factors contribute to THLP having the highest affinity of the inhibitors tested. These results provide insight into the nature of chemical structures that bind ALDH4A1, which could aid the development of chemical probes targeting proline and hydroxyproline metabolic enzymes.

## MATERIALS AND METHODS

**Enzyme Kinetics.** Steady-state kinetics assays were used to study the inhibition of mouse ALDH4A1 (MmALDH4A1, Q8CHT0) by D,L-proline and the four stereoisomers of 4-hydroxyproline. MmALDH4A1 (93% identical to human ALDH4A1) was expressed in *Escherichia coli* and purified as described previously.<sup>21</sup> The following compounds were purchased from Sigma: L-proline (product number P0380), D-proline (product number 858919), trans-4-hydroxy-L-proline (THLP, product number H54409), trans-4-hydroxy-D-proline (THDP, product number 702501), cis-4-hydroxy-D-proline (CHDP, product number H5877), cis-4-hydroxy-L-proline (CHLP, product number H1637).

The activity of MmALDH4A1 in the presence of inhibitors was measured by monitoring NADH production at 340 nm with L-P5C as the variable substrate (0 - 180  $\mu$ M) and NAD<sup>+</sup>

as the fixed substrate (1 mM) in a reaction buffer of 100 mM sodium phosphate pH 7 and 10 mM EDTA. D,L-P5C was synthesized from D,L-5-hydroxylysine-HCl as described previously.<sup>29</sup> We note the concentration of NAD<sup>+</sup> in the assays is much greater than the  $K_m$  of 100  $\mu$ M for human ALDH4A1.<sup>21</sup> D,L-P5C was neutralized to pH 7.5 immediately prior to enzyme assays using 1 M Tris (pH 7.5) and 6 M NaOH. The concentration of L-P5C was assumed to be half the total D,L-P5C concentration added to the assays. The data were acquired at room temperature in 96-well plates using a BioTek Epoch 2 microplate spectrophotometer. The final MmALDH4A1 concentration in each assay was 5  $\mu$ g/mL. The total volume of the assay was 200  $\mu$ L. Inhibitor and P5C were spotted on the plate and the reaction was initiated by the addition of a master mix containing enzyme, NAD<sup>+</sup>, and assay buffer. The initial rates were estimated from linear regression of the first 5 - 6 minutes of the progress curve. The initial rate data for each inhibitor were fit globally with Origin software to various inhibition models, including competitive, uncompetitive, and mixed. The results of these calculations are summarized in Table S3.1. The results from the best-fitting model for each compound are listed in Table 3.1.

The GSALDH activity of SmPutA in the presence of inhibitors was measured by monitoring NADH production at 340 nm with L-P5C as the variable substrate (0 - 1.25 mM) and NAD<sup>+</sup> as the fixed substrate (0.2 mM). SmPutA was expressed and purified by Ashley Campbell.<sup>29</sup> The data were acquired at room temperature in 96-well plates using a BioTek Epoch 2 microplate spectrophotometer. The assay buffer contained 100 mM sodium phosphate pH 7 and 1 mM EDTA. The SmPutA concentration in each assay was 170 nM. The initial rates were estimated from linear regression of the first 5 - 6 minutes of the progress curve. The initial rate data for each inhibitor were fit globally with Origin

software to various inhibition models, including competitive, uncompetitive, mixed, and competitive with substrate inhibition.<sup>26</sup>

**Crystal Structure Determination.** Crystallization experiments were set up with MmALDH4A1 (6 mg/mL) in a buffer containing 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM TCEP, and 5% glycerol at pH 7.5. MmALDH4A1 was co-crystallized with 20 mM NAD<sup>+</sup> at 20 °C using the sitting drop vapor diffusion method, combining 1 µL each of the protein and reservoir solutions. The reservoir contained 0.1 M Bis Tris pH 6-7, 15-25% (w/v) PEG 3350, and 0.2 M Li<sub>2</sub>SO<sub>4</sub>. The THLP complex was formed *in crystallo* by soaking crystals in a solution containing 300 mM THLP and 20% (v/v) PEG 200 for five minutes and then flash-cooling in liquid nitrogen. The THDP complex was obtained by soaking crystals in a solution containing 180 mM THDP, 20 mM NAD<sup>+</sup>, and 32% (v/v) PEG 200 for five minutes and then flash-cooling in liquid nitrogen.

X-ray diffraction data were collected at the Advanced Photon Source beamline 24-ID-E using an Eiger-16 M detector. The data were processed with XDS<sup>33</sup> and AIMLESS.<sup>34</sup> The space group is  $P2_12_12_1$  and the asymmetric unit contains a dimer of MmALDH4A1. We note this is the same crystal form used in previous structural studies of MmALDH4A1.<sup>21, 23, 25</sup> Data processing statistics are summarized in Table 3.3.

The starting model for crystallographic refinement in PHENIX<sup>35, 36</sup> was obtained from a 1.30 Å resolution structure of MmALDH4A1 (PDB ID 3V9J<sup>21</sup>). The *B*-factor model consisted of one TLS group per protein chain and isotropic *B*-factors for all non-hydrogen atoms. Iterative model building and manual adjustments were performed using COOT.<sup>37</sup> The restraint files for ligands were generated in PHENIX eLBOW from the three-digit chemical component code and employing AM1 optimization.<sup>38</sup> The structures were

validated using MolProbity and the wwPDB validation service.<sup>39, 40</sup> Modeling of ligands was validated with polder omit maps.<sup>41</sup> Refinement statistics are summarized in Table 3.3.

## **ACKNOWLEDGMENTS**

We thank Dr. J. Schuermann for help with X-ray diffraction data collection and processing. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on the 24-ID-E beam line is funded by a NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## **Funding Information**

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award numbers R01GM065546 and R01GM132640. A.N.B. was supported by a Wayne L. Ryan Fellowship through The Ryan Foundation.

## **ASSOCIATED CONTENT**

### **Supporting Information**

Table S3.1 and Figure S3.1



## REFERENCES

- 1.Adams, E.; Frank, L., Metabolism of proline and the hydroxyprolines. *Annu. Rev. Biochem.* **1980**, *49*, 1005-1061.
- 2.Tanner, J. J., Structural Biology of Proline Catabolic Enzymes. *Antioxid Redox Signal* **2019**, *30* (4), 650-673.
- 3.Tanner, J. J.; Fendt, S. M.; Becker, D. F., The Proline Cycle As a Potential Cancer Therapy Target. *Biochemistry* **2018**, *57* (25), 3433-3444.
- 4.Phang, J. M., Proline Metabolism in Cell Regulation and Cancer Biology: Recent Advances and Hypotheses. *Antioxid Redox Signal* **2019**, *30* (4), 635-649.
- 5.Burke, L.; Guterman, I.; Palacios Gallego, R.; Britton, R. G.; Burschowsky, D.; Tufarelli, C.; Rufini, A., The Janus-like role of proline metabolism in cancer. *Cell Death Discov* **2020**, *6*, 104.
- 6.Pranzini, E.; Pardella, E.; Paoli, P.; Fendt, S. M.; Taddei, M. L., Metabolic Reprogramming in Anticancer Drug Resistance: A Focus on Amino Acids. *Trends Cancer* **2021**.
- 7.Bergers, G.; Fendt, S. M., The metabolism of cancer cells during metastasis. *Nat Rev Cancer* **2021**, *21* (3), 162-180.
- 8.Yoon, K. A.; Nakamura, Y.; Arakawa, H., Identification of ALDH4 as a p53-inducible gene and its protective role in cellular stresses. *J. Hum. Genet.* **2004**, *49* (3), 134-40.
- 9.Polyak, K.; Xia, Y.; Zweier, J. L.; Kinzler, K. W.; Vogelstein, B., A model for p53-induced apoptosis. *Nature* **1997**, *389* (6648), 300-5.
- 10.Geraghty, M. T.; Vaughn, D.; Nicholson, A. J.; Lin, W. W.; Jimenez-Sanchez, G.; Obie, C.; Flynn, M. P.; Valle, D.; Hu, C. A., Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia. *Hum. Mol. Genet.* **1998**, *7* (9), 1411-5.
- 11.Phang, J. M.; Hu, C. A.; Valle, D., Disorders of proline and hydroxyproline metabolism. In *Metabolic and molecular basis of inherited disease*, Scriver, C. R.; Beaudet, A. L.; Sly, W. S.; Valle, D., Eds. McGraw Hill: New York, 2001; pp 1821-1838.
- 12.Convertino, M.; Das, J.; Dokholyan, N. V., Pharmacological Chaperones: Design and Development of New Therapeutic Strategies for the Treatment of Conformational Diseases. *ACS Chem Biol* **2016**, *11* (6), 1471-89.
- 13.McCorvie, T. J.; Yue, W. W., Chapter 13 - Structure-guided discovery of pharmacological chaperones targeting protein conformational and misfolding diseases. **2020**, 281-308.

14. Lorenzo, C.; Delgado, P.; Busse, C. E.; Sanz-Bravo, A.; Martos-Folgado, I.; Bonzon-Kulichenko, E.; Ferrarini, A.; Gonzalez-Valdes, I. B.; Mur, S. M.; Roldan-Montero, R.; Martinez-Lopez, D.; Martin-Ventura, J. L.; Vazquez, J.; Wardemann, H.; Ramiro, A. R., ALDH4A1 is an atherosclerosis auto-antigen targeted by protective antibodies. *Nature* **2021**, 589 (7841), 287-292.
15. Yen, C. A.; Curran, S. P., Incomplete proline catabolism drives premature sperm aging. *Aging Cell* **2021**, 20 (2), e13308.
16. Yen, C. A.; Ruter, D. L.; Turner, C. D.; Pang, S.; Curran, S. P., Loss of flavin adenine dinucleotide (FAD) impairs sperm function and male reproductive advantage in *C. elegans*. *Elife* **2020**, 9.
17. Buchalski, B.; Wood, K. D.; Challa, A.; Fargue, S.; Holmes, R. P.; Lowther, W. T.; Knight, J., The effects of the inactivation of Hydroxyproline dehydrogenase on urinary oxalate and glycolate excretion in mouse models of primary hyperoxaluria. *Biochim Biophys Acta Mol Basis Dis* **2020**, 1866 (3), 165633.
18. Coulter-Mackie, M. B., 4-Hydroxyproline metabolism and glyoxylate production: A target for substrate depletion in primary hyperoxaluria? *Kidney Int.* **2006**, 70 (11), 1891-3.
19. Vasiliou, V.; Thompson, D. C.; Smith, C.; Fujita, M.; Chen, Y., Aldehyde dehydrogenases: from eye crystallins to metabolic disease and cancer stem cells. *Chem. Biol. Interact.* **2013**, 202 (1-3), 2-10.
20. Liu, L. K.; Becker, D. F.; Tanner, J. J., Structure, function, and mechanism of proline utilization A (PutA). *Arch Biochem Biophys* **2017**, 632, 142-157.
21. Srivastava, D.; Singh, R. K.; Moxley, M. A.; Henzl, M. T.; Becker, D. F.; Tanner, J. J., The three-dimensional structural basis of type II hyperprolinemia. *J Mol Biol* **2012**, 420 (3), 176-89.
22. Pemberton, T. A.; Tanner, J. J., Structural basis of substrate selectivity of Delta(1)-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): Semialdehyde chain length. *Arch. Biochem. Biophys.* **2013**, 538 (1), 34-40.
23. Pemberton, T. A.; Still, B. R.; Christensen, E. M.; Singh, H.; Srivastava, D.; Tanner, J. J., Proline: Mother Nature's cryoprotectant applied to protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, 68 (Pt 8), 1010-8.
24. Forte-McRobbie, C.; Pietruszko, R., Human glutamic-gamma-semialdehyde dehydrogenase. Kinetic mechanism. *Biochem. J.* **1989**, 261 (3), 935-43.
25. Pemberton, T. A.; Srivastava, D.; Sanyal, N.; Henzl, M. T.; Becker, D. F.; Tanner, J. J., Structural Studies of Yeast Delta(1)-Pyrroline-5-carboxylate Dehydrogenase

(ALDH4A1): Active Site Flexibility and Oligomeric State. *Biochemistry* **2014**, 53 (8), 1350-9.

26.Riveros-Rosas, H.; Gonzalez-Segura, L.; Julian-Sanchez, A.; Diaz-Sanchez, A. G.; Munoz-Clares, R. A., Structural determinants of substrate specificity in aldehyde dehydrogenases. *Chem. Biol. Interact.* **2013**, 202 (1-3), 51-61.

27.Korasick, D. A.; Pemberton, T. A.; Arentson, B. W.; Becker, D. F.; Tanner, J. J., Structural Basis for the Substrate Inhibition of Proline Utilization A by Proline. *Molecules* **2017**, 23 (1), 32.

28.Korasick, D. A.; Koncitikova, R.; Kopečna, M.; Hajkova, E.; Vigouroux, A.; Morera, S.; Becker, D. F.; Sebela, M.; Tanner, J. J.; Kopečný, D., Structural and Biochemical Characterization of Aldehyde Dehydrogenase 12, the Last Enzyme of Proline Catabolism in Plants. *J Mol Biol* **2019**, 431 (3), 576-592.

29.Campbell, A. C.; Bogner, A. N.; Mao, Y.; Becker, D. F.; Tanner, J. J., Structural analysis of prolines and hydroxyprolines binding to the l-glutamate-gamma-semialdehyde dehydrogenase active site of bifunctional proline utilization A. *Arch Biochem Biophys* **2021**, 698, 108727.

30.Howles, S. A.; Thakker, R. V., Genetics of kidney stone disease. *Nat Rev Urol* **2020**, 17 (7), 407-421.

31.Summitt, C. B.; Johnson, L. C.; Jonsson, T. J.; Parsonage, D.; Holmes, R. P.; Lowther, W. T., Proline dehydrogenase 2 (PRODH2) is a hydroxyproline dehydrogenase (HYPDH) and molecular target for treating primary hyperoxaluria. *Biochem J* **2015**, 466 (2), 273-81.

32.Yang, H.; Male, M.; Li, Y.; Wang, N.; Zhao, C.; Jin, S.; Hu, J.; Chen, Z.; Ye, Z.; Xu, H., Efficacy of Hydroxy-L-proline (HYP) analogs in the treatment of primary hyperoxaluria in *Drosophila Melanogaster*. *BMC Nephrol* **2018**, 19 (1), 167.

33.Kabsch, W., XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66 (Pt 2), 125-32.

34.Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **2013**, 69 (Pt 7), 1204-14.

35.Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66 (Pt 2), 213-21.

36.Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D.,

Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, 68 (Pt 4), 352-67.

37.Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Cryst. D Biol. Crystallogr.* **2010**, 66 (Pt 4), 486-501.

38.Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D., electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, 65 (Pt 10), 1074-80.

39.Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, D66 (Pt 1), 12-21.

40.Gore, S.; Sanz Garcia, E.; Hendrickx, P. M. S.; Gutmanas, A.; Westbrook, J. D.; Yang, H.; Feng, Z.; Baskaran, K.; Berrisford, J. M.; Hudson, B. P.; Ikegawa, Y.; Kobayashi, N.; Lawson, C. L.; Mading, S.; Mak, L.; Mukhopadhyay, A.; Oldfield, T. J.; Patwardhan, A.; Peisach, E.; Sahni, G.; Sekharan, M. R.; Sen, S.; Shao, C.; Smart, O. S.; Ulrich, E. L.; Yamashita, R.; Quesada, M.; Young, J. Y.; Nakamura, H.; Markley, J. L.; Berman, H. M.; Burley, S. K.; Velankar, S.; Kleywegt, G. J., Validation of Structures in the Protein Data Bank. *Structure* **2017**, 25 (12), 1916-1927.

41.Liebschner, D.; Afonine, P. V.; Moriarty, N. W.; Poon, B. K.; Sobolev, O. V.; Terwilliger, T. C.; Adams, P. D., Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol* **2017**, 73 (Pt 2), 148-157.

**Table 3.1.** Inhibition and Kinetic Constants of ALDH4A1<sup>a</sup>

Compound	Model	$K_i$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (μM)
L-Pro	Competitive	$1.9 \pm 0.3$	$0.23 \pm 0.01$	$32 \pm 4$
THLP	Competitive	$0.7 \pm 0.1$	$0.25 \pm 0.01$	$30 \pm 2$
CHLP	Competitive	$> 10$	$0.43 \pm 0.01$	$45 \pm 2$
D-Pro	Competitive	$> 10$	$0.43 \pm 0.01$	$47 \pm 2$
THDP	Uncompetitive	$9 \pm 1$	$1.8 \pm 0.07$	$84 \pm 7$
CHDP	Competitive	$> 10$	$0.27 \pm 0.004$	$34 \pm 2$

<sup>a</sup>P5C was the variable substrate with NAD<sup>+</sup> fixed at 1 mM. One trial was performed for each P5C concentration, and the uncertainties were obtained from global nonlinear curve fitting in Origin.

**Table 3.2**Inhibition Constants of SmPutA GSALDH from Various Models<sup>a</sup>

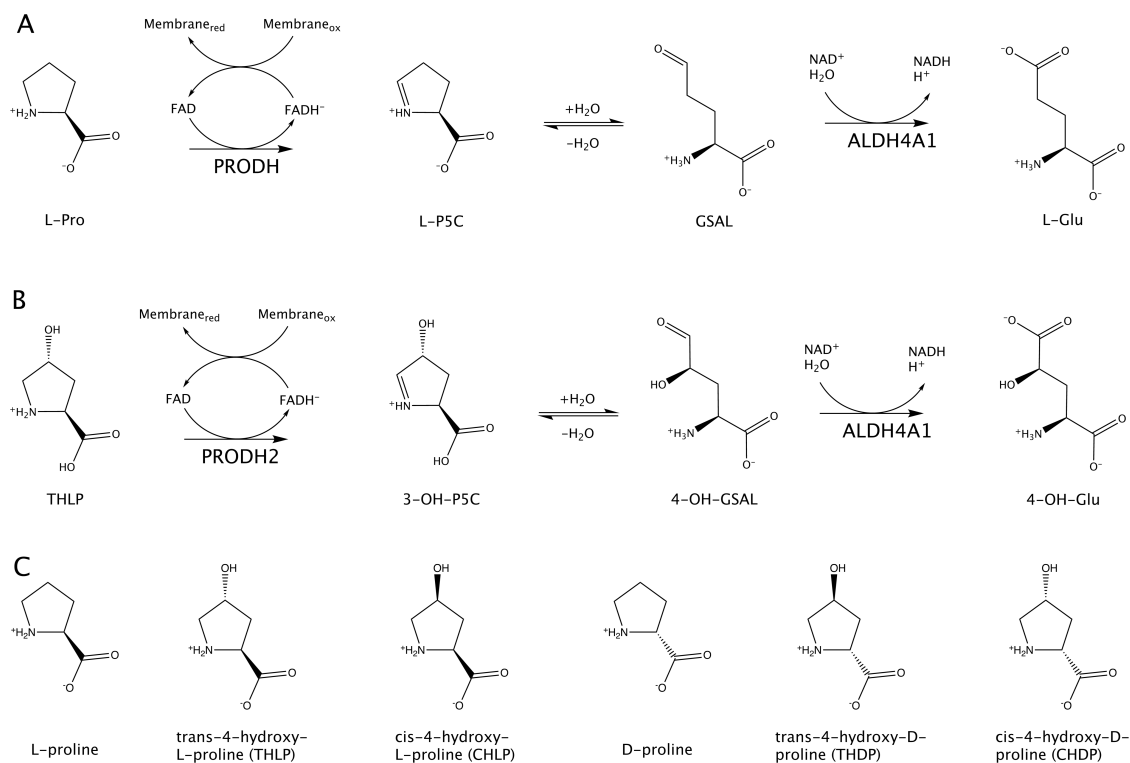
Compound	Model	$K_i$ (mM)	$\alpha$	Adjusted $R^2$
D-Pro	Competitive	$1.5 \pm 0.5$	N/A	0.889
THDP	Competitive	$4.5 \pm 0.5$	N/A	0.984
CHDP	Competitive	$8. \pm 1.$	N/A	0.968
L-Pro	Competitive	$11. \pm 2.$	N/A	0.941
THLP	Competitive	$7. \pm 1.$	N/A	0.974
D-Pro	Uncompetitive	$4.5 \pm 0.2$	N/A	0.992
THDP	Mixed	$10. \pm 1.$	$1.6 \pm 0.4$	0.996
CHDP	Mixed	$25. \pm 9.$	$0.9 \pm 0.5$	0.985
L-Pro	Uncompetitive	$19. \pm 2.$	N/A	0.974
THLP	Mixed	$27. \pm 4.$	$0.6 \pm 0.1$	0.997

<sup>a</sup>P5C was the variable substrate with NAD<sup>+</sup> fixed at 0.2 mM.

**Table 3.3** Diffraction Data Collection and Refinement Statistics

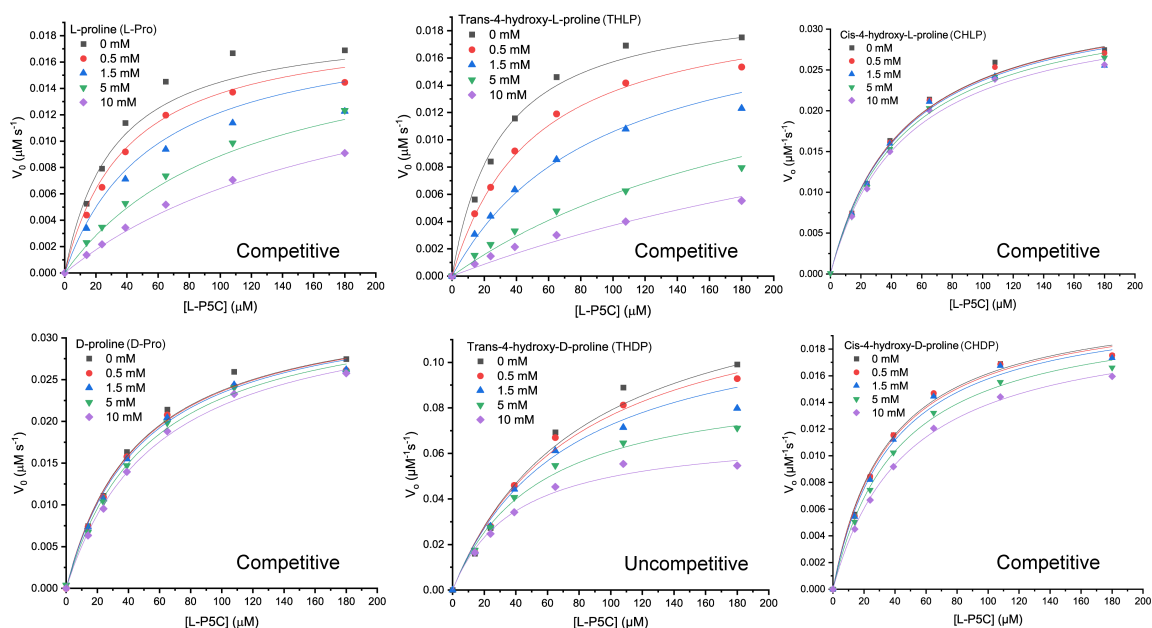
	THLP	THDP + NAD <sup>+</sup>
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell parameters (Å)	<i>a</i> = 85.06 <i>b</i> = 94.60 <i>c</i> = 132.59	<i>a</i> = 84.74 <i>b</i> = 94.11 <i>c</i> = 131.64
Wavelength (Å)	0.97918	0.97918
Resolution (Å)	132.59 - 1.74 (1.77 - 1.74)	131.64-1.37 (1.39–1.37)
Observations <sup>a</sup>	904402 (37591)	1874978 (71818)
Unique reflections <sup>a</sup>	110402 (5207)	221058 (9660)
<i>R</i> <sub>merge</sub> ( <i>I</i> ) <sup>a</sup>	0.131 (1.078)	0.048 (0.614)
<i>R</i> <sub>meas</sub> ( <i>I</i> ) <sup>a</sup>	0.139 (1.160)	0.051 (0.656)
<i>R</i> <sub>pim</sub> ( <i>I</i> ) <sup>a</sup>	0.048 (0.418)	0.017 (0.226)
Mean <i>I</i> /σ <sup>a</sup>	11.9 (1.9)	23.7 (2.8)
CC <sub>1/2</sub> <sup>a</sup>	0.998 (0.662)	1.000 (0.870)
Completeness (%) <sup>a</sup>	99.7 (95.6)	99.4 (88.6)
Multiplicity <sup>a</sup>	8.2 (7.2)	8.5 (7.4)
No. of protein residues	1067	1050
Protein	8224	8041
Pro ligand	18	36
NAD <sup>+</sup>	N/A	46
Water	757	637
<i>R</i> <sub>cryst</sub>	0.1639 (0.3358)	0.1606 (0.2198)
<i>R</i> <sub>free</sub> <sup>b</sup>	0.1896 (0.3668)	0.1717 (0.2454)
rmsd bonds (Å)	0.006	0.005
rmsd angles (°)	0.777	0.817
Favored (%)	98.21	98.09
Outliers (%)	0.00	0.00
Clashscore (PR) <sup>c</sup>	1.10 (99)	0.86 (99)
MolProbity score (PR) <sup>c</sup>	0.82 (100)	0.76 (100)
Average <i>B</i> -factor		
Protein	17.0	19.5
Pro ligand	22.1	20.3
NAD <sup>+</sup>	N/A	19.6
Water	24.0	25.7
Coord. error (Å) <sup>d</sup>	0.22	0.12
PDB code	7MER	7MES

<sup>a</sup>Values for the outer resolution shell of data are given in parenthesis.<sup>b</sup>5% test set.<sup>c</sup>From MolProbity. The percentile ranks (PR) for Clashscore and MolProbity score are given in parentheses.<sup>d</sup>Maximum likelihood-based coordinate error estimate from PHENIX.

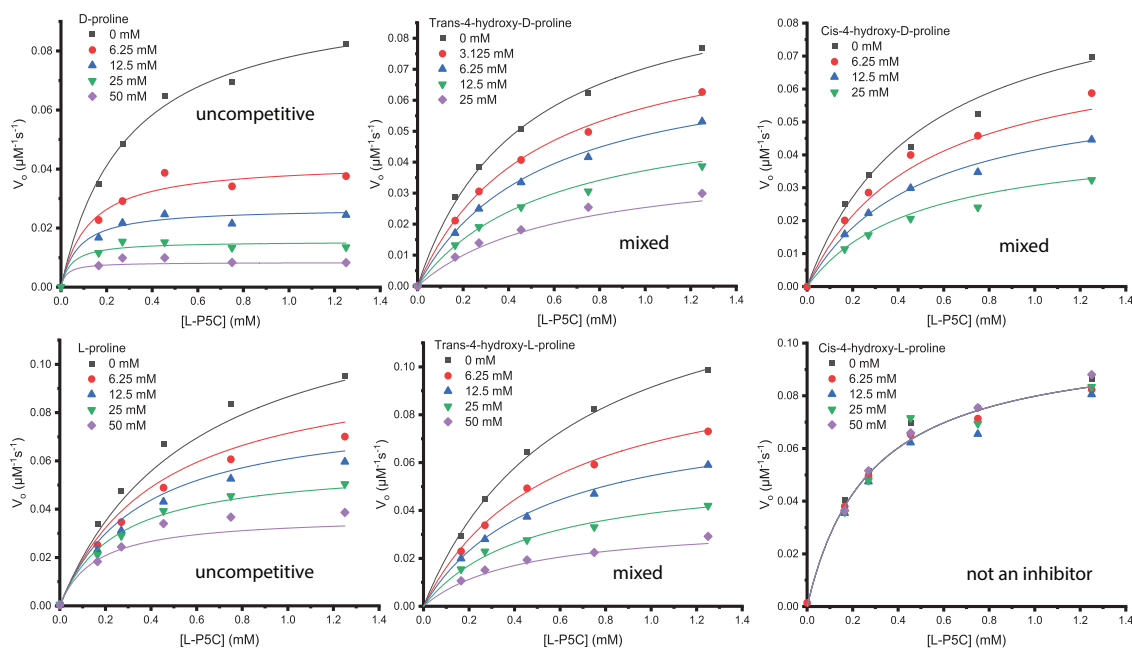


**Figure 3.1.** Enzymatic reactions and inhibitors. (A) Reactions and enzymes of proline catabolism. (B) Reactions and enzymes of hydroxyproline catabolism. (C) Inhibitors used in this study.

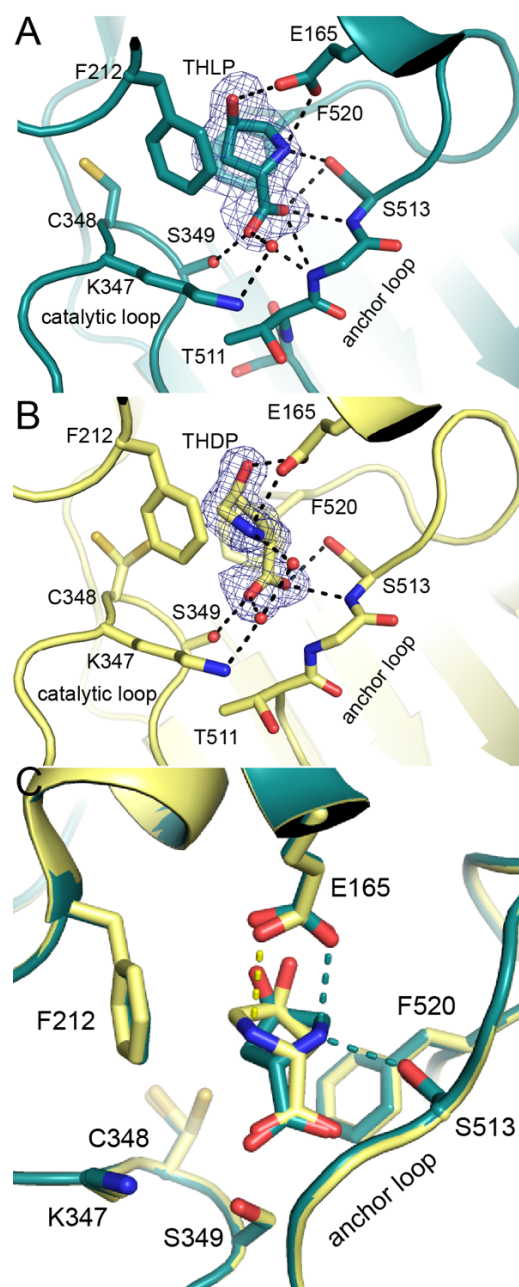




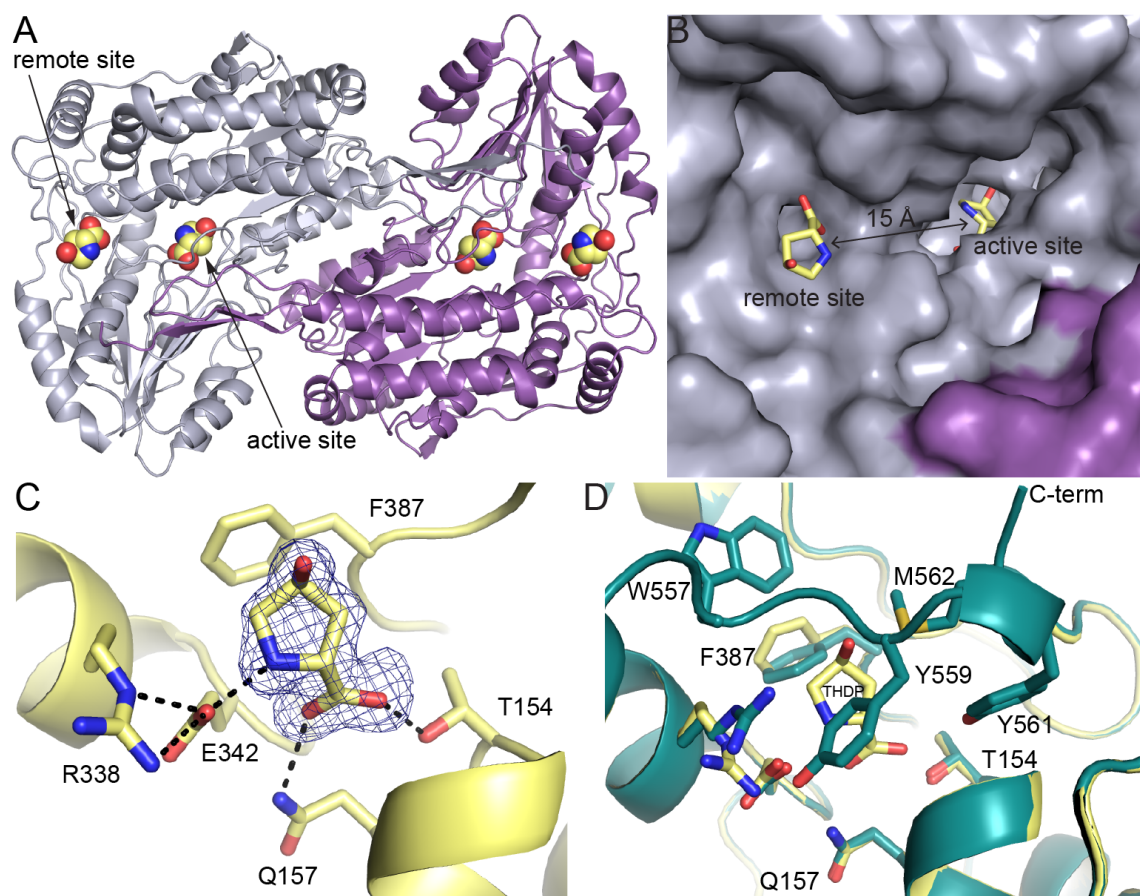
**Figure 3.2** Inhibition of MmALDH4A1 activity by prolines and hydroxyprolines. The assays were performed at room temperature with  $\text{NAD}^+$  at 1 mM and MmALDH4A1 at 5  $\mu\text{g/mL}$  in a buffer containing 100 mM sodium phosphate pH 7.0 and 10 mM EDTA. The data for each inhibitor were analyzed by global fitting to either the competitive inhibition model (L-proline, THLP, CHLP, D-proline, and CHDP) or the uncompetitive inhibition model (THDP) using Origin software.



**Figure 3.3** Inhibition of the GSALDH activity of SmPutA by prolines and hydroxyprolines. The assays were performed at room temperature with  $\text{NAD}^+$  at 0.2 mM and SmPutA at 170 nM in a buffer containing 100 mM sodium phosphate pH 7 and 1 mM EDTA. The data for each inhibitor were analyzed by global fitting to the either the uncompetitive inhibition model (D-proline, L-proline) or the mixed inhibition model (THDP, CHDP, THDP) using Origin software. See Table 3.2 for inhibition constants.



**Figure 3.4** Structures of the MmALDH4A1 active site inhibited by THLP and THDP. (A) Electron density and interactions for THLP. (B) Electron density and interactions for THDP. (C) Superposition of the THLP (teal) and THDP (yellow) complexes highlighting the difference in the orientations of the pyrrolidine rings. The mesh in panels A and B represents polder omit maps (4.5σ).



**Figure 3.5** Remote binding site of THDP. (A) Dimer of MmALDHA1 viewed down the two-fold symmetry axis showing the locations of the THDP sites. (B) Close-up view of the GSAL entrance tunnel showing the two THDP binding sites. (C) Electron density and interactions for THDP in the remote site (polder omit,  $4\sigma$ ). (D) Superposition of the THDP (yellow) and THLP (teal) structures showing how the C-terminus is ordered and blocks the remote site in the absence of THDP.

## SUPPORTING INFORMATION

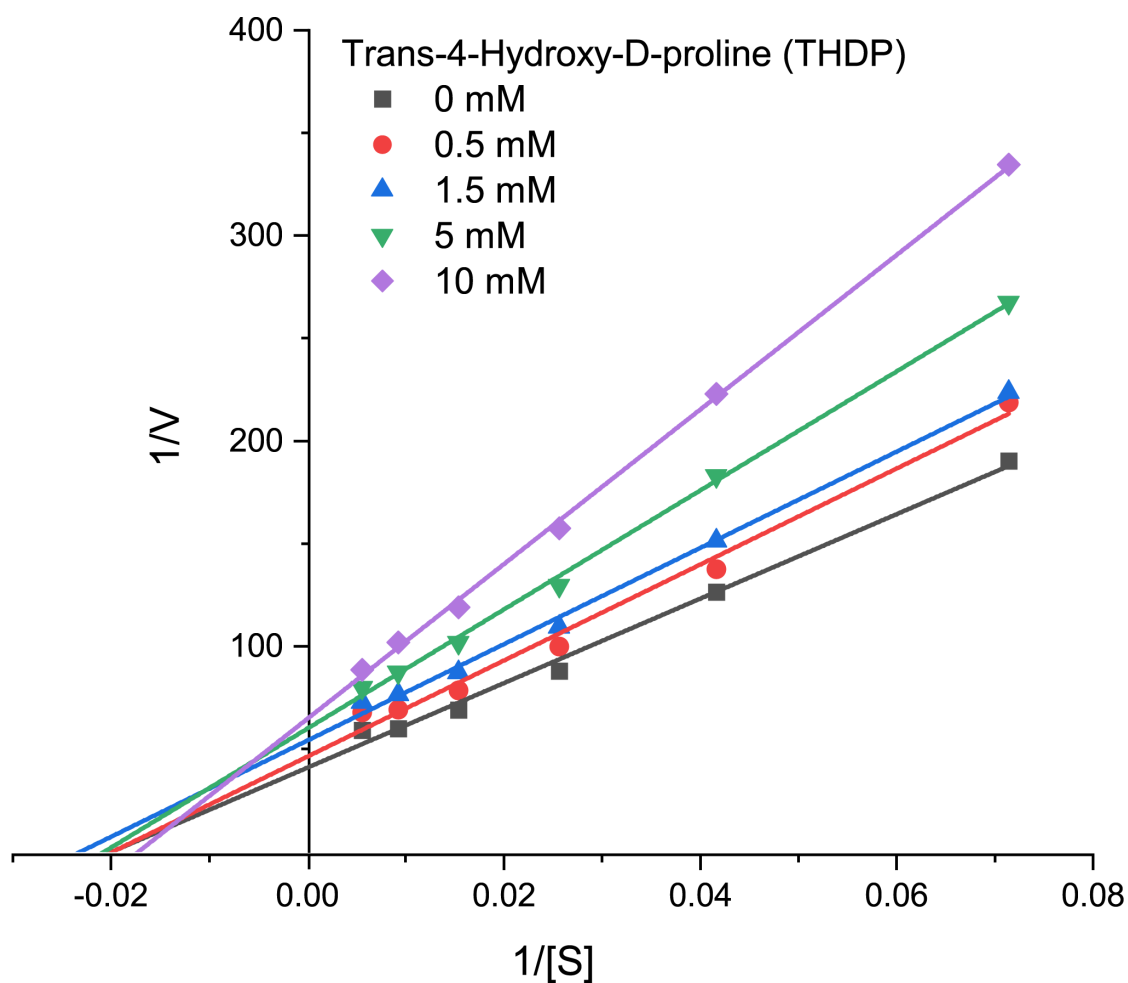
### Structural Basis for the Stereospecific Inhibition of the Dual

#### Proline/Hydroxyproline Catabolic Enzyme ALDH4A1 by Trans-4-Hydroxy-L-Proline

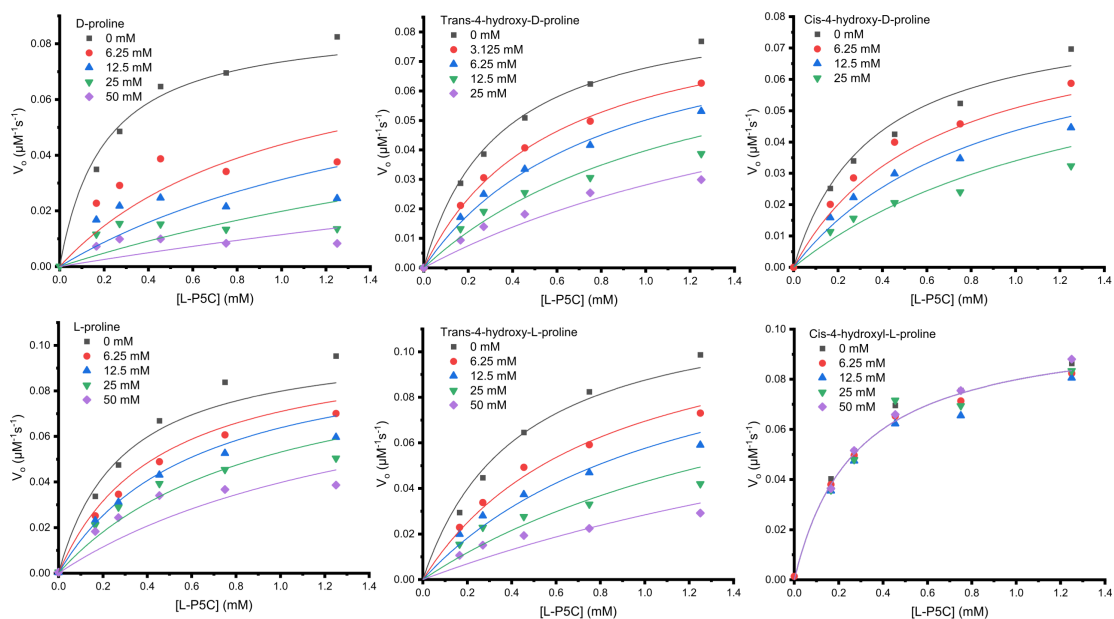
**Table S3.1** Inhibition and Kinetic Constants of ALDH4A1 from Various Models<sup>a</sup>

Compound	Model	$K_i$ (mM)	$\alpha$	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (μM)	Adjusted $R^2$
L-Pro	Competitive	$1.9 \pm 0.3$	N/A	$0.23 \pm 0.01$	$32 \pm 4$	0.976
THLP	Competitive	$0.7 \pm 0.1$	N/A	$0.25 \pm 0.01$	$30 \pm 2$	0.992
CHLP	Competitive	$59 \pm 21$	N/A	$0.43 \pm 0.01$	$45 \pm 2$	0.993
D-Pro	Competitive	$39 \pm 9$	N/A	$0.43 \pm 0.01$	$47 \pm 2$	0.994
THDP	Competitive	$8 \pm 2$	N/A	$1.5 \pm 0.09$	$59 \pm 8$	0.962
CHDP	Competitive	$17 \pm 2$	N/A	$0.27 \pm 0.004$	$34 \pm 2$	0.994
L-Pro	Mixed	$2.0 \pm 0.5$	$54 \pm 164$	$0.24 \pm 0.01$	$32 \pm 5$	0.975
THLP	Mixed	$1.1 \pm 0.1$	$6.3 \pm 1.6$	$0.26 \pm 0.005$	$35 \pm 2$	0.996
CHLP	Mixed	$102 \pm 99$	$2 \pm 4$	$0.43 \pm 0.01$	$47 \pm 3$	0.993
D-Pro	Mixed	$38 \pm 15$	0	$0.43 \pm 0.01$	$47 \pm 3$	0.993
THDP	Uncompetitive	$9 \pm 1$	N/A	$1.8 \pm 0.07$	$84 \pm 7$	0.986
CHDP	Mixed	$23 \pm 6$	$5.8 \pm 4.8$	$0.27 \pm 0.01$	$36 \pm 2$	0.995

<sup>a</sup>P5C was the variable substrate with NAD<sup>+</sup> fixed at 1 mM. One trial was performed for each P5C concentration, and the uncertainties were obtained from global nonlinear curve fitting in Origin.



**Figure S3.1** Double-reciprocal analysis of the THDP inhibition data. The assays were performed at room temperature with  $\text{NAD}^+$  at 1 mM and MmALDH4A1 at 5  $\mu\text{g/mL}$  in a buffer containing 100 mM sodium phosphate pH 7.0 and 10 mM EDTA.



**Figure S3.2** Analysis of SmPutA GSALDH inhibition data using the competitive inhibition model. The assays were performed at room temperature with  $\text{NAD}^+$  at 0.2 mM and SmPutA at 170 nM in a buffer containing 100 mM sodium phosphate pH 7 and 1 mM EDTA. The data for each inhibitor were analyzed by global fitting to the competitive inhibition model using Origin software. The apparent  $K_i$  values are (mM):  $1.5 \pm 0.5$ , D-proline;  $4.5 \pm 0.5$ , THDP;  $8. \pm 1.$ , CHDP;  $11. \pm 2.$ , L-proline;  $7. \pm 1.$ , THLP; CHLP, no inhibition observed.

## Chapter 4

### Preliminary Exploration of the Reaction Scope of Photoinduced Covalent Inactivation of Proline Dehydrogenase

Alexandra N. Bogner<sup>1</sup>, Juan Ji<sup>1</sup>, and John J. Tanner<sup>1,2</sup>

<sup>1</sup>Department of Biochemistry, University of Missouri, Columbia, Missouri 65211, United States

<sup>2</sup>Department of Chemistry, University of Missouri, Columbia, Missouri 65211, United States

**Author Contributions:** A.N.B.: Conceptualization, Methodology, Investigation, Writing-Original Draft, Writing-Review & Editing, Visualization. J.J.: Investigation. J.J.T.: Conceptualization, Writing-Review & Editing, Visualization, Validation, Supervision, Project administration, Funding acquisition.

**Keywords:** X-ray crystallography, flavoenzymes, proline metabolism, mechanism-based inactivation, covalent modification of FAD, photopharmacology.

**Abbreviations:** GSALDH, L-glutamate- $\gamma$ -semialdehyde dehydrogenase; MST, microscale thermophoresis; NFLP, N-formyl-L-proline; NPPG, *N*-propargylglycine; *o*-AB, *ortho*-aminobenzaldehyde; PRODH, proline dehydrogenase; P5C,  $\Delta^1$ -pyrroline-5-carboxylate; PutA86-630, protein containing residues 86-630 of *Escherichia coli* proline utilization A; PYCR,  $\Delta^1$ -pyrroline-5-carboxylate reductase; SmPutA, proline utilization A from *Sinorhizobium meliloti*; SmPutA571, PutA protein containing residues 1-571 of *Sinorhizobium meliloti*; THFA, L-tetrahydro-2-furoic acid.



## ABSTRACT

The proline cycle consists of the catabolic enzyme proline dehydrogenase (PRODH) and the biosynthetic enzyme  $\Delta^1$ -pyrroline-5-carboxylate reductase (PYCR). L-proline is oxidized to  $\Delta^1$ -pyrroline-5-carboxylate (P5C) by PRODH and P5C is reduced back to L-proline by PYCR1. The proline cycle has emerged as a potential cancer therapy target, so inhibitors of the enzymes are sought. Recently, our lab reported the photoinduced covalent inactivation of PRODH by 1,3-dithiolane-2-carboxylic acid (**1**). Here we report additional kinetic characterization of **1**, as well as preliminary data on the scope of the inactivation reaction using acyclic S-containing carboxylic acids. Kinetic assays show that **1** is a time-dependent inhibitor characterized by an apparent second-order rate constant for enzyme inactivation of  $400 \text{ M}^{-1}\text{s}^{-1}$ . Activity measurements to assess the noncovalent potency of **1** yielded variable results depending on the type of assay and PRODH enzyme used and warrants clarification. Crystal structures of PRODH treated with 2-(methylsulfanyl)acetic acid (**90**), 2-(cyclopropylsulfanyl)acetic acid (**91**), or 2-(propylsulfanyl)acetic acid (**92**) and blue light reveal apparent decarboxylation of the compound and covalent modification of the FAD N5, indicating these compounds inactivate via the proposed mechanism. Visual inspection of flavin bleaching *in crystallo* suggests that inactivation by these compounds is irreversible. A structure of PRODH complexed with **92** determined under low-light conditions reveals the noncovalent complex prior to inactivation. Flavin bleaching studies suggest that the effects of 2-(methylthio)propanoic acid (**86**), 2-sulfanylacetic acid (**87**), 2-sulfanylpropanoic acid (**88**), and 2-(ethylsulfanyl)propanoic acid (**89**) may be reversible. These results represent the initial phase of determining the scope of photoinduced covalent inactivation of PRODH.

## INTRODUCTION

Two enzymes in proline metabolism, proline dehydrogenase (PRODH) and  $\Delta^1$ -pyrroline-5-carboxylate reductase (PYCR) are of increasing interest due to their newly found implications in cancer. Combined, the catabolic enzyme PRODH and the biosynthetic enzyme PYCR form the proline cycle, and a better understanding of the implications of the cycle in cancer is becoming abundantly important. The alterations of proline metabolism in cancer can be attributed to its role in redox stress, energy production, and apoptosis.<sup>1-3</sup> Both enzymes in the proline cycle are differentially expressed in cancerous versus normal cells. PYCR tends to be upregulated while PRODH expression can be either upregulated or downregulated depending on the cancer type.<sup>4-6</sup> PRODH catalyzes the oxidation of L-proline to  $\Delta^1$ -pyrroline-5-carboxylate (P5C) through the reduction of an FAD cofactor (Figure 4.1A, adapted from Bogner et al., 2021).<sup>2</sup> This reaction relates to cell proliferation through the generation of ATP and redox cofactors (Figure 4.1A).<sup>7,8</sup> PYCR catalyzes the reduction of L-P5C back to L-proline through the oxidation of NAD(P)H (Figure 4.1A). A product of this reaction, NAD(P)<sup>+</sup>, is fed through the oxidative pentose phosphate pathway and glycolysis, which are known to promote cellular proliferation.<sup>3,6,9</sup>

Targeting enzymes that are upregulated in cancer with inhibitors is a potential route for cancer therapy. Studies have shown that cancer cells treated with L-tetrahydro-2-furoic acid (THFA), a well-studied noncovalent inhibitor of PRODH, impaired lung metastasis of 4T1 and EMT6.5 mouse models.<sup>9</sup> Additionally, a study of the role of PRODH in non-small cell lung cancer showed that treatment with THFA decreased cell proliferation and migration *in vitro* and tumor growth *in vivo*.<sup>10</sup> A mechanism based inactivator of PRODH, N-propargylglycine (NPPG), was shown to induce rapid decay of PRODH without

detriment to the overall health of breast tumor (MCF7)-xenografted mice.<sup>11,12</sup> N-formyl-L-proline (NFLP) was studied recently for its inhibitory effects as the first validated chemical probe against PYCR1. When MCF10A H-RAS<sup>V12</sup> breast cancer cells were treated with NFLP, proline levels increased, and spheroid growth decreased in a manner that phenocopied a *PYCR1* knockout.<sup>13</sup> Both PRODH and PYCR1 have validated chemical probes, but increasing potency, specificity, and duration of inhibition can always be improved.

Two compounds were recently discovered by our lab to covalently inactivate PRODH in a light-dependent manner: 1,3-dithiolane-2-carboxylate (**1** in Figure 4.1B) and tetrahydrothiophene-2-carboxylate (**2**).<sup>14</sup> The first step of the proposed mechanism for **1** is the formation of the noncovalent inactivator-enzyme complex in the absence of light (Figure 4.2).<sup>14</sup> Next, exposing the complex to blue light produces an excited flavin state, FAD\*. The inactivator then transfers an electron to FAD\*, creating a reduced FAD semiquinone and either a carboxyl- or sulfur-centered radical. The covalently inactivated enzyme is then formed by the combination of a carbon-centered radical through decarboxylation of the inactivator and FAD semiquinone.

Here we report a preliminary study of the scope of photoinduced covalent inactivation of PRODH using a series of acyclic carboxylic acid compounds containing an S atom bonded to the  $\alpha$ -carbon (Figure 4.1B). Bacterial PRODH enzymes were used in the place of human PRODH due to the active sites of the bacterial enzymes sharing a high sequence conservation with eukaryotes.<sup>3</sup> Also, human PRODH is located in the inner mitochondrial membrane, making it difficult to express and purify, therefore the bacterial enzymes are useful in structural and kinetic studies. All seven compounds showed initial promise as

photoinduced inactivators, but upon further study, **86**, **87**, **88**, and **89** demonstrated signs of reversibility in reaction. **90**, **91**, and **92** covalently inactivated PRODH and crystal structures of the covalently inactivated enzyme complex with these compounds, and noncovalent complex with **92**, were solved. By comparing the reversibility or irreversibility of compounds with their chemical structures, our results provide insight into the scope of photoinduced inactivation of PRODH.

## MATERIALS AND METHODS

**Materials.** The following compounds were purchased through Molport from individual suppliers: 2-(methylthio)propanoic acid (**86**) Molport-004-346-067; ChemBridge, 2-sulfanylacetic acid (**87**) Molport-003-935-930; AK Scientific, Inc., 2-sulfanylpropanoic acid (**88**) Molport-003-939-593; Vitas-M Laboratory, Ltd., 2-(ethylsulfanyl)propanoic acid (**89**) Molport-004-346-226; ChemBridge, 2-(methylsulfanyl)acetic acid (**90**) Molport-000-157-541; BLD Pharmatech Ltd., 2-(cyclopropylsulfanyl)acetic acid (**91**) Molport-029-001-860; Enamine, and 2-(propylsulfanyl)acetic acid (**92**) Molport-002-472-611; ChemBridge. 1,3-Dithiolane-2-carboxylic acid (**1**) was purchased through Enamine with product number EN300-107439.

**Protein Purification.** The bifunctional PRODH - L-glutamate- $\gamma$ -semialdehyde dehydrogenase (GSALDH) enzyme proline utilization A from *Sinorhizobium meliloti* (SmPutA) was purified as previously described.<sup>15</sup> A new PRODH domain construct containing residues 1-571 of SmPutA with an N-terminal His tag and TEV cleavage site (SmPutA571) was synthesized by GenScript Biotech Corp. The following transformation, expression, and purification were performed by Juan Ji. pNIC28-SmPutA571 was

transformed into BL21 (DE3) competent cells and plated onto LB agar plates with 50 µg/ml kanamycin. A starter culture of 10 mL LB, shaken overnight at 37 °C and 250 rpm, was used to inoculate a 1 L culture, all containing 50 µg/ml kanamycin. The 1 L culture shook at 37 °C and 250 rpm and once the OD<sub>600</sub> reached above 0.8 the culture was induced with 1 mM isopropyl β -D-1- thiogalactopyranoside and grown with continued shaking at 18 °C overnight. Centrifugation at 5000 rpm for 30 min at 4 °C was used to collect the cells and the pellet was stored at -80 °C until ready for purification. The pellet was then resuspended in 50 mM Tris, 300 mM NaCl, 10 mM imidazole, and 5% glycerol at pH 7.5 (buffer A), with the inclusion of 0.15mg/ml lysozyme and 2 tablets of EDTA-free protease inhibitor. FAD was added into the lysis buffer until the buffer turned pale yellow. The cells were lysed with sonication and centrifuged at 16,500 rpm for 1 hour at 4 °C. The lysate was then purified by gravity-flow chromatography on a column containing Ni<sup>2+</sup>-NTA resin (Qiagen) pre-equilibrated with buffer A. The column was washed with buffer A supplemented with 20-, 40-, and 250mM imidazole. SmPutA571 did not stick to the column, so the flow and 10 mM imidazole wash were precipitated with ammonium sulfate and centrifuged at 16,500 rpm for 1 hour at 4 °C. The precipitated protein pellet was resuspended with buffer A and incubated at RT with TEV protease to remove the His Tag at a ratio of 1 mg of protease: 10 mg of SmPutA571. The protein was then dialyzed in 50 mM Tris pH 7.8, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM TECP, and 5% glycerol (buffer B) and purified further with anion exchange using a HiTrap-Q column. SmPutA571 was collected in the flow-through, while contaminating proteins were retained by the resin and washed with a gradient of 0-1M NaCl. The protein was then aliquoted and stored at -80 °C.

A PRODH domain construct containing residues 86-630 of *Escherichia coli* PutA with a C-terminal His tag (PutA86-630) was expressed and purified as described in Chapter 2.

**Characterization of Reversible Inhibition.** Reversible inhibition was studied using kinetic measurements performed under low-light conditions. Kinetic measurements were performed using the *ortho*-aminobenzaldehyde (*o*-AB) assay in a 96-well plate in a BioTek Epoch 2 microplate spectrophotometer at room temperature. The assay monitors the production of P5C as an adduct formed with *o*-AB which is detected by absorbance at 443 nm ( $\epsilon_{443} = 2.59 \text{ mM}^{-1}\text{cm}^{-1}$ ). The electron acceptor menadione was used to reoxidize the reduced FAD of PRODH, enabling catalytic cycling. The assay buffer contained 20 mM MOPS pH 7.5 and 10 mM  $\text{MgCl}_2$ . Michaelis-Menten kinetic measurements were performed at varied L-proline (0-500 mM) and **1** (0-0.3125 mM or 0-1.25 mM) concentrations, 4 mM *o*-AB, 0.15 mM menadione, and 63 nM of enzyme (PutA86-630, SmPutA, or SmPutA571). L-proline and inhibitors were spotted on the plate and a master mix including enzyme, menadione, *o*-AB, and buffer was added to the plate by multichannel pipette to initiate the reaction. The initial rate was determined from linear regression of the first 5 minutes of the progress curve using Origin v9.7.0.188 software. The initial rate data as functions of both substrate and inhibitor concentrations were fit globally to a competitive inhibition model using Origin.

**Characterization of Photoinduced Covalent Inactivation.** To measure the photoinduced covalent inactivation of PRODH, 5  $\mu\text{L}$  aliquots of PutA86-630 and varying concentrations of **1** were incubated together in a 96-well plate and exposed to a 700 lumen LED bulb from a distance of 20 cm for 0-21 min. At 21 min, 190  $\mu\text{L}$  of a master mix containing L-proline, *o*-AB, menadione, and *o*-AB assay buffer was added to the plate to

initiate the PRODH activity assay. The final concentration of reagents in the PRODH assay (total volume of 200  $\mu$ L) were 100 mM L-proline, 4 mM *o*-AB, 0.15 mM menadione, 63 nM PutA86-630, and 0-20  $\mu$ M **1**. Initial rates were measured, expressed as percent activity (normalized by the rate of PutA86-630 that had not been treated with **1**) vs time, and fit to an exponential decay function using Origin software. The time constants from the exponential decay plot were plotted against the reciprocal of the concentration of **1** on a Kitz and Wilson replot.

**Assessment of the Reversibility of Covalent Inactivation by Absorbance Spectroscopy.** Spectra of PutA86-630 (3 mg/mL) in the presence of 25 mM **87** or **88** were measured in a quartz cuvette using a Nanodrop2000c. Spectra were acquired at 1 min intervals for a total of 5 or 8 min. The cuvette was removed from the holder between spectral acquisitions and exposed to a 700 lumen LED bulb from a distance of 1 cm for 50 sec between readings. Then, the samples were stored in the dark at 4 °C for 21 h. The next day, each sample was transferred to the spectrophotometer under low-light conditions, and the spectrum was measured.

**Crystallization.** SmPutA was co-crystallized with 50 mM ligand (**86**, **89**, **90**, **91**, or **92**) and 10 mM NAD<sup>+</sup> at 13 °C using the sitting-drop vapor diffusion method. Crystallization experiments were set up with SmPutA (6 mg/mL) in a buffer containing 50 mM Tris (pH 8.0), 50 mM NaCl, 5% (w/v) glycerol, and 0.5 mM Tris(2-carboxyethyl)phosphine. Crystals were grown using a reservoir solution containing 10-20% PEG-3350, 0.1-0.25 M ammonium sulfate, 0.1 M magnesium chloride, 0.1 M HEPES (pH 8.0), and 0.1 M sodium formate. A drop ratio of 2  $\mu$ L:2  $\mu$ L for protein: reservoir solution was used. The crystals were harvested in low-light conditions for noncovalent complex structures or exposed to

blue light for covalent complex structures. The noncovalent complex crystals of SmPutA and **92** were cryoprotected with reservoir solution supplemented with 20% PEG 200, and flash-cooled in liquid nitrogen. Blue light was used to induce a covalent N5 modification of the FAD in the covalent complex structures. The SmPutA crystals complexed with **90**, **91**, or **92** were exposed to 90 min of blue light (470 nm) from a 283-lumen source held 20 cm from the crystal, cryoprotected with reservoir solution supplemented with 20% PEG 200, and flash-cooled in liquid nitrogen. SmPutA crystals complexed with **86**, **89**, **90**, **91**, or **92** were exposed to 90 min of blue light and observed for an assessment of the reversibility of covalent inactivation.

**X-ray Diffraction Data Collection and Refinement.** X-ray diffraction data were collected at Advanced Photon Source beamline 24-ID-E using an Eiger-16 M detector and 24-ID-C using an Eiger2 X 16 M detector. The data were processed with XDS<sup>16</sup> and AIMLESS<sup>17</sup>. All the data sets are in space group  $P2_1$  and have similar unit cell dimensions listed in Table 4.2 The asymmetric unit contains a dimer of SmPutA. We note this is the same crystal form used for previous crystallographic studies of SmPutA.<sup>14,15,18</sup>

PHENIX<sup>19</sup> was used for refinement, and Coot<sup>20</sup> was used for model building. The starting model for refinement of the covalent complex structures was derived from a minimal model of SmPutA complexed with trans-4-hydroxy-L-proline (PDB ID: 6X9A).<sup>21</sup> This model omitted the most variable regions of the structure (residues 13, 810-813, 1001-1004, 884, 1230-1233 and the  $\alpha 8$  helix) to reduce model bias in the electron density maps. The starting model for the noncovalent complex structure was derived from the structure of SmPutA complexed with THFA (PDB ID: 5KF6)<sup>18</sup>. The SMILES string for **92** was used



as the input to ELBOW<sup>22</sup> to generate the ligand coordinates and restraint files used during refinement. The covalent modifications have not been modeled yet.

## RESULTS AND DISCUSSION

**Noncovalent Inhibition of PRODH by 1.** The noncovalent inhibition of PRODH by **1** was measured using an *o*-AB assay in low-light conditions. The assay was performed at room temperature with 0-500 mM L-proline, 0-0.3125 mM **1**, 4 mM *o*-AB, 0.15 mM menadione, and 63 nM PutA86-630 in a buffer containing 20 mM MOPS pH 7.5 and 10 mM MgCl<sub>2</sub>. Inhibition kinetics with the initial rate as functions of the concentrations of L-proline and **1** were fit globally to the competitive model (Figure 4.3). An inhibition constant ( $K_i$ ) of 11.8  $\mu$ M was calculated (Table 4.1). Previous studies demonstrated the noncovalent inhibition of SmPutA by **1** using a coupled assay where proline is provided as the substrate and the generation of NADH by the L-glutamate- $\gamma$ -semialdehyde dehydrogenase (GSALDH) active site is measured. A modest IC<sub>50</sub> value of 2.3 mM was reported.<sup>14</sup> The juxtaposition of these results led us to do further kinetic studies to compare the effects of **1** on different PutA constructs. Thus, the inhibition of the PRODH activity of full-length SmPutA and a new PRODH domain construct of SmPutA (SmPutA571) with the *o*-AB assay were performed. L-proline was varied from 0-350 mM, **1** was varied from 0-1.25 mM, 63 nM SmPutA or SmPutA571 was used, and all other *o*-AB assay reagents were at the previously described concentrations. A  $K_i$  of 363  $\mu$ M was calculated for SmPutA (Figure 4.4A) and SmPutA571 gave a  $K_i$  almost identical to that of PutA86-630, 11.5  $\mu$ M (Figure 4.4B). The resulting  $K_i$  values for all three enzymes are summarized in Table 4.1.

Both enzyme construct and activity assay affected  $K_i$  value. Measurements of PutA86-630 and SmPutA571 using the *o*-AB assay shared the lowest  $K_i$  value, SmPutA in combination with the *o*-AB assay produced an intermediate value, and SmPutA with the coupled assay had the highest  $K_i$  value. One apparent increase in potency could be related to PutA86-630 and SmPutA571 lacking a GSALDH domain. Furthermore, the *o*-AB assay is a more direct measurement of the PRODH domain activity than the coupled assay that measures the combined activity of the bifunctional enzyme, possibly explaining another discrepancy. The size of the PutA construct seems to be more important than the type of bacterial enzyme, as both PRODH domain constructs gave similar results. Overall, this data shows **1** to be a comparable, and possibly more potent, inhibitor than THFA ( $K_i \sim 0.2$ – $1$  mM).<sup>23,24</sup> Additionally, the decrease in  $K_i$ , indicating an increase in affinity of noncovalent binding to PRODH, from originally  $2.3$  mM<sup>14</sup> to  $11.5$ – $363$   $\mu$ M may redirect our thinking of **1** as a photoinduced inactivator, but further studies must be done to elucidate this.

**Covalent Inactivation of PutA86-630 by 1.** Based off our new findings of PutA86-630 inhibition by **1**, we decided to calculate the parameters of PutA86-630 photoinduced inactivation by **1**. PutA86-630 and **1** were incubated together, exposed to white light for 0–21 min, and then diluted (40-fold) into an *o*-AB activity assay. Final concentrations of 100 mM L-proline, 4 mM *o*-AB, 0.15 mM menadione, 63 nM PutA86-630, and 0–20  $\mu$ M **1** were used. Inactivation profiles of PutA86-630 with three different concentrations of **1** (2.5, 10, 20  $\mu$ M) were used to estimate inactivation parameters. An inactivation constant ( $k_{\text{inact}}$ ) and  $K_i$  were obtained from plotting the percent activity as a function of illumination time, fitting to an exponential decay function to determine time constants ( $t_{1/2}$ ), and finally

plotting  $t_{1/2}$  values versus the inverse of inactivator concentration (Kitz and Wilson replot). The parameters for enzyme inactivation are  $k_{\text{inact}} = 0.26 \pm 0.02 \text{ min}^{-1}$  and  $K_i = 11.01 \pm 1.40 \text{ }\mu\text{M}$  (Figure 4.5, inset). The apparent second-order rate constant for enzyme inactivation ( $k_{\text{inact}}/K_i$ ) is  $393.41 \pm 59.26 \text{ M}^{-1}\text{s}^{-1}$ . Seemingly, the  $K_i$  value in the inactivation experiment closely matches that of the inhibition experiment.

**Reversibility of Covalent Inactivation.** The goal of this study was to identify other nonreversible photoinduced inactivators of PRODH, therefore the compounds were tested in-solution or *in-crystallo* for reversibility. Using in-solution absorbance spectroscopy, the reduction and N5 covalent modification of FAD can be assessed through a decrease in the 450 nm peak and an increase in the 320 nm peak of a flavin spectra.<sup>14,25</sup> 25 mM of **87** or **88** were individually added to 3 mg/mL PutA86-630, exposed to white light, and the FAD absorbance was measured in quartz cuvettes in a Nanodrop2000c (Figure 4.6). For both **87** and **88**, an apparent decrease in the 450 nm peak was observed, as well as a bleaching of the solution. However, an increase in the 320 nm peak indicating a modification to the N5 atom of the FAD was not observed. Furthermore, when the bleached solution (after 5 or 8 minutes of white light exposure for **87** and **88** respectively) was stored in the dark and measured again 21 hours later, there was an increase in absorbance for both compounds. This indicates that whatever caused the apparent reduction of the FAD was reversible. Interestingly, the absorbance curve around 320 nm did not follow the previous absorbance as the day before, with slightly lower absorbance values from 300-350 nm, while the absorbances from 350-600 nm compared to the previous measurements. This may indicate that the occurring reaction is not fully reversible. **88** showed a higher increase in absorbance than **87** after 21 hours which may be related to their structural differences, the

addition of a methyl group at the  $\alpha$ -carbon. We note that both compounds contain a free thiol, which complicates the interpretation of the results.

**86** and **89-92** were tested for reversibility of covalent inactivation *in crystallo*. Bleaching of yellow color of PRODH indicates reduction of the FAD, a necessary but not sufficient condition for a covalent modification of the N5 atom of FAD.<sup>14,15,25</sup> SmPutA crystals were grown in the presence of 50 mM **86**, **89**, **90**, **91**, or **92**, and 10 mM NAD<sup>+</sup>. The crystals were then exposed to 90 minutes of blue light, as this type of light was determined to be preferred for the photoinduced reaction.<sup>14</sup> The color of each crystal started as yellow and was bleached over the course of 90 minutes (Figure 4.7). Crystals complexed with **92**, **91**, or **90** remained colorless after spending 15 hours in the dark (Figure 4.7A, B, C), whereas crystals with **89** or **86** regained their yellow appearance (Figure 4.7D, E). A week later the colorless crystals remained colorless (data not shown). A common feature of the compounds that are seemingly irreversible is that they are linear, whereas the reversible compounds have an methyl group on the  $\alpha$ -carbon. Interestingly, **88** is also branched like **86** and **89**, and it showed a faster reversibility than **87**, indicating **88**. Further exploration of the reversibility and time dependence of light exposure should be done in-solution to monitor flavin spectral changes at 320- and 450 nm peaks.

**Crystal Structures of Noncovalently-Inhibited and Covalently-Inactivated PRODH.** Crystals from the same preparation for the reversibility test *in-crystallo* were harvested before the addition of light for the purpose of solving a noncovalent complex. A preliminary 1.7 Å structure was solved of **92** noncovalently bound in the active site of SmPutA (Table 4.2 and Figure 4.8). The carboxylate of **92** binds in the same manner as the proline analogs studied by our group.<sup>14,15,26</sup> Ion pairing of the inhibitor carboxylate with

Arg488, Arg489, and Lys265 can be seen (Figure 4.8). The conserved water molecule that is typically seen in the active site is not present in this structure, possibly due to the increased size of the sulfur atom, where an oxygen or nitrogen normally sits, and thus displaces the water. The inhibitor has a bent conformation in which the hydrocarbon tail packs across the *si* face of the FAD isoalloxazine ring. The density showed no evidence of covalent modification to the FAD, consistent with the yellow color of the harvested crystal. Further model building and refinements must be done to understand the implications of the noncovalent binding of **92** on the PRODH active site.

Crystals from the same preparation for the reversibility test *in-crystallo* were harvested after blue light exposure for the purpose of solving a covalent complex. Preliminary structures of SmPutA inactivated by **90**, **91**, and **92** were solved at 1.55-1.79 Å resolution (Table 4.2). A minimal model of 6X9A<sup>21</sup> was used as the starting model and no additional model building or waters were included in the structure determination. Strong electron density connected to the FAD N5 is evidence for covalent modification by **90**, **91**, and **92** (Figure 4.9). The density feature appears smaller than the compounds used in co-crystallization, consistent with decarboxylation as occurred with **1**. Further model building and refinement must be done to understand the full effects of the N5 modification and how the structure compares to that of SmPutA inactivated by **1** (PDB ID 7MYA).<sup>14</sup>

## CONCLUSIONS

Preliminary studies show that differences in PRODH enzyme constructs and activity assays influence the estimation of the inhibition constant for noncovalent inhibition by **1**. When using a PRODH domain construct (PutA86-630 or SmPutA571) compared to the

full length SmPutA, the apparent affinity for **1** increases ~ 30-fold. Use of different activity assays with the same enzyme (SmPutA) results in a 6-fold change in  $K_i$ . Combining the two - a PRODH domain construct and *o*-AB assay - alters the  $K_i$  by ~200-fold from our published estimations.<sup>14</sup> Additionally, we have potentially identified other photoinduced covalent-inactivators of PRODH. **90**, **91**, and **92** when co-crystallized with SmPutA and exposed to blue light cause a modification to the N5 atom of the flavin. Further model building and refinement of the structures must be done to elucidate the full reaction observed and chemical structure of the attachment to the N5 atom. Reversibility of a photoinduced reaction became apparent in this study and may be related to the presence of a methyl group on the  $\alpha$ -carbon of **86**, **88**, and **89**. The exploration done here on chemical structure of S-containing compounds may eventually inform on usage of light-activated chemical probes in cancer biology research.

## FUTURE DIRECTIONS

**Discrepancy in the Potency of **1** as a Noncovalent Inhibitor.** Further work should be done to resolve the discrepancies in the noncovalent potency of **1**. Inhibition of SmPutA could be compared to other PutAs to further ensure variations in bacterial enzymes is not an issue. An orthogonal technique of microscale thermophoresis (MST) can be used to determine a binding constant ( $K_d$ ) between SmPutA and **1** and compared to that of SmPutA and THFA. Utilizing existing inhibitors of PRODH like THFA can help identify if this problem is specific to **1**, or the inhibition of PRODH/PutA more generally.

**Compounds **86** and **89-92** as Noncovalent Inhibitors.** Calculating a consistent inhibition constant for the compounds in the absence of light has proven to be difficult.  $K_i$

values fluctuate with experiment indicating the possibility of the compounds changing over time (Table 4.1). Identifying potential chemical reactions within our experiments or stock solutions could be beneficial. Solving crystal structures for the remaining compounds could elucidate any changes that are occurring in the chemical structures of the compounds. MST can also be used in this instance to determine a  $K_d$  in place of a  $K_i$ .

Interestingly, **92** was originally determined to be the compound with the weakest affinity (noncovalent) for PRODH in this study (Table 4.1) but was the only noncovalent complex structure solved. We note that structure determination of noncovalent complexes of **86**, **89**, **90**, and **91** were attempted using compound concentrations of 50 mM but the maps lacked electron density for the inhibitor. This conundrum needs to be solved. It is possible to make predictions about the poses of the other compounds based on the structure of the **92**-complex. For example, the presence of the methyl group on the  $\alpha$ -carbon could change the entire positioning of the molecules, as it may clash with Ala372. However, having a crystal structure is always more satisfying. Solidifying the inhibition constants for the compounds may aid in solving the noncovalent complex structures.

**Compounds 86 and 89-92 as Covalent Inactivators.** In-solution experiments measuring flavin spectra should be done to further validate whether these compounds have reversible activity. For **87** and **88**, a 320 nm peak did not increase as the 450 nm peak decreased, and when the compounds were tested after being stored in the dark, the FAD absorbance spectra varied from the previous measurements slightly. Similar experiments for **86** and **89** should be done to observe flavin spectral changes and to determine the speed of reversibility in comparison to **87** and **88**. Moreover, the previously studied weaker

inactivator of PRODH, **2**,<sup>14</sup> has not been tested for reversibility and could help clarify the connection between the chemical structure and ability to reverse the reaction.

Understanding the mechanism of reversibility could be revealed in a crystal structure, so further experiments should be done to capture intermediate states of the compounds in crystals. We hypothesize there could potentially be three different states: (1) the compound sitting noncovalently in the active site, (2) the compound reacting with the flavin in the presence of light, and (3) a partial reverse of this reaction with some modification of the flavin and the compound. It may be possible to capture these states using kinetic crystallography approaches.

In-solution assessment of covalent inactivation of PRODH should be done to validate the presence of an N5 covalent modification of the FAD with **90**, **91**, and **92**. An initial test of all seven compounds was done to show the reduction of the flavin, and all compounds showed a decrease in the 450 nm peak. Whether or not **90**, **91**, or **92**, show a clear increase in the 320 nm peak indicating a modification of the N5 atom of the FAD in-solution is still unknown. Once in-solution experiments have been performed and irreversibility is further validated, inactivation assays to calculate  $k_{\text{inact}}$  and  $K_i$  should be performed.

## ACKNOWLEDGMENTS

We thank Jonathan Schuermann for help with X-ray diffraction data collection. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by a NIH-ORIP HEI grant (S10OD021527). This research used resources of the



Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

### **Funding Information**

This work was supported by NIGMS, National Institutes of Health, Grant R01GM132640.

A.N.B. is the recipient of a Wayne L. Ryan Fellowship through The Ryan Foundation.

## REFERENCES

1. Phang JM (2019) Proline Metabolism in Cell Regulation and Cancer Biology: Recent Advances and Hypotheses. *Antioxid Redox Signal* [Internet] 30:635–649. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28990419>
2. Bogner AN, Stiers KM, Tanner JJ (2021) Structure, biochemistry, and gene expression patterns of the proline biosynthetic enzyme pyrroline-5-carboxylate reductase (PYCR), an emerging cancer therapy target. *Amino Acids* [Internet]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/34003320>
3. Tanner JJ, Fendt SM, Becker DF (2018) The Proline Cycle As a Potential Cancer Therapy Target. *Biochemistry* 57:3433–3444.
4. Pandhare J, Donald SP, Cooper SK, Phang JM (2009) Regulation and function of proline oxidase under nutrient stress. *Journal of Cellular Biochemistry* 107:759–768.
5. Burke L, Guterman I, Palacios Gallego R, Britton RG, Burschowsky D, Tufarelli C, Rufini A (2020) The Janus-like role of proline metabolism in cancer. *Cell Death Discov* [Internet] 6:104. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/33083024>
6. D’Aniello C, Patriarca EJ, Phang JM, Minchiotti G (2020) Proline Metabolism in Tumor Growth and Metastatic Progression. *Frontiers in Oncology* 10:776.
7. Liu W, Phang JM (2012) Proline dehydrogenase (oxidase) in cancer. *BioFactors* 38:398–406.
8. Liu W, Glunde K, Bhujwalla ZM, Raman V, Sharma A, Phang JM (2012) Proline oxidase promotes tumor cell survival in hypoxic tumor microenvironments. *Cancer Research* 72:3677–3686.
9. Elia I, Broekaert D, Christen S, Boon R, Radaelli E, Orth MF, Verfaillie C, Grünewald TGP, Fendt SM (2017) Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nature Communications* 8.
10. Liu Y, Mao C, Wang M, Liu N, Ouyang L, Liu S, Tang H, Cao Y, Liu S, Wang X, et al. (2020) Cancer progression is mediated by proline catabolism in non-small cell lung cancer. *Oncogene* [Internet] 39:2358–2376. Available from: <http://dx.doi.org/10.1038/s41388-019-1151-5>
11. Scott GK, Mahoney S, Scott M, Loureiro A, Lopez-Ramirez A, Tanner JJ, Ellerby LM, Benz CC (2021) N-Propargylglycine: a unique suicide inhibitor of proline dehydrogenase with anticancer activity and brain-enhancing mitohormesis properties. *Amino Acids* [Internet]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/34089390>

12. Scott GK, Yau C, Becker BC, Khateeb S, Mahoney S, Jensen MB, Hann B, Cowen BJ, Pegan SD, Benz CC (2019) Targeting Mitochondrial Proline Dehydrogenase with a Suicide Inhibitor to Exploit Synthetic Lethal Interactions with p53 Upregulation and Glutaminase Inhibition. *Molecular Cancer Therapeutics* [Internet] 18:1374–1385. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31189611>
13. Christensen EM, Bogner AN, Vandekeere A, Tam GS, Patel SM, Becker DF, Fendt SM, Tanner JJ (2020) In crystallo screening for proline analog inhibitors of the proline cycle enzyme PYCR1. *J Biol Chem* [Internet] 295:18316–18327. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/33109600>
14. Campbell AC, Prater AR, Bogner AN, Quinn TP, Gates KS, Becker DF, Tanner JJ (2021) Photoinduced Covalent Irreversible Inactivation of Proline Dehydrogenase by S-Heterocycles. *ACS Chemical Biology* [Internet]. Available from: <https://pubmed.ncbi.nlm.nih.gov/34542291/>
15. Campbell AC, Becker DF, Gates KS, Tanner JJ (2020) Covalent Modification of the Flavin in Proline Dehydrogenase by Thiazolidine-2-Carboxylate. *ACS Chem Biol* [Internet] 15:936–944. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/32159324>
16. Kabsch W (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 66:125–132. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=20124692](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20124692)
17. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 69:1204–1214. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23793146>
18. Luo M, Gamage TT, Arentson BW, Schlasner KN, Becker DF, Tanner JJ (2016) Structures of Proline Utilization A (PutA) Reveal the Fold and Functions of the Aldehyde Dehydrogenase Superfamily Domain of Unknown Function. Available from: <http://wwpdb.org/>
19. Afonine P v, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 68:352–367. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22505256>
20. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Cryst. D Biol. Crystallogr.* [Internet] 66:486–501. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20383002>
21. Campbell AC, Bogner AN, Mao Y, Becker DF, Tanner JJ (2021) Structural analysis of prolines and hydroxyprolines binding to the l-glutamate-gamma-semialdehyde

dehydrogenase active site of bifunctional proline utilization A. *Arch Biochem Biophys* [Internet] 698:108727. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/33333077>

22. Moriarty NW, Grosse-Kunstleve RW, Adams PD (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* [Internet] 65:1074–1080. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19770504>

23. Zhu W, Gincherman Y, Docherty P, Spilling CD, Becker DF (2002) Effects of proline analog binding on the spectroscopic and redox properties of PutA. *Arch. Biochem. Biophys.* [Internet] 408:131–136. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12485611](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12485611)

24. White TA, Krishnan N, Becker DF, Tanner JJ (2007) Structure and kinetics of monofunctional proline dehydrogenase from *Thermus thermophilus*. *J. Biol. Chem.* [Internet] 282:14316–14327. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17344208](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17344208)

25. White TA, Johnson Jr. WH, Whitman CP, Tanner JJ (2008) Structural basis for the inactivation of *Thermus thermophilus* proline dehydrogenase by N-propargylglycine. *Biochemistry* [Internet] 47:5573–5580. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=18426222](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18426222)

26. Zhang M, White TA, Schuermann JP, Baban BA, Becker DF, Tanner JJ (2004) Structures of the *Escherichia coli* PutA proline dehydrogenase domain in complex with competitive inhibitors. *Biochemistry* [Internet] 43:12539–12548. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15449943](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15449943)

**Table 4.1** Noncovalent inhibition constants compared across PRODH enzymes.  $K_i$  values are reported by experiment, in some cases multiple experiments were performed.

Compound	$K_i$ ( $\mu\text{M}$ )		
	PutA86-630	SmPutA	SmPutA571
<b>1</b>	$11.88 \pm 0.01$	$363 \pm 54$	$11.54 \pm 1.05$
<b>86</b>	$52 \pm 4$ $3000 \pm 1600$	$51 \pm 19$	-
<b>89</b>	$786 \pm 80$ $780 \pm 94$	$2100 \pm 1500$	-
<b>90</b>	$260 \pm 17$	$900 \pm 375$	$97 \pm 20$
<b>91</b>	$5600 \pm 400$ $3500 \pm 500$	$2750 \pm 2500$	-
<b>92</b>	$450 \pm 50$	$1030 \pm 450$	$12300 \pm 2000$

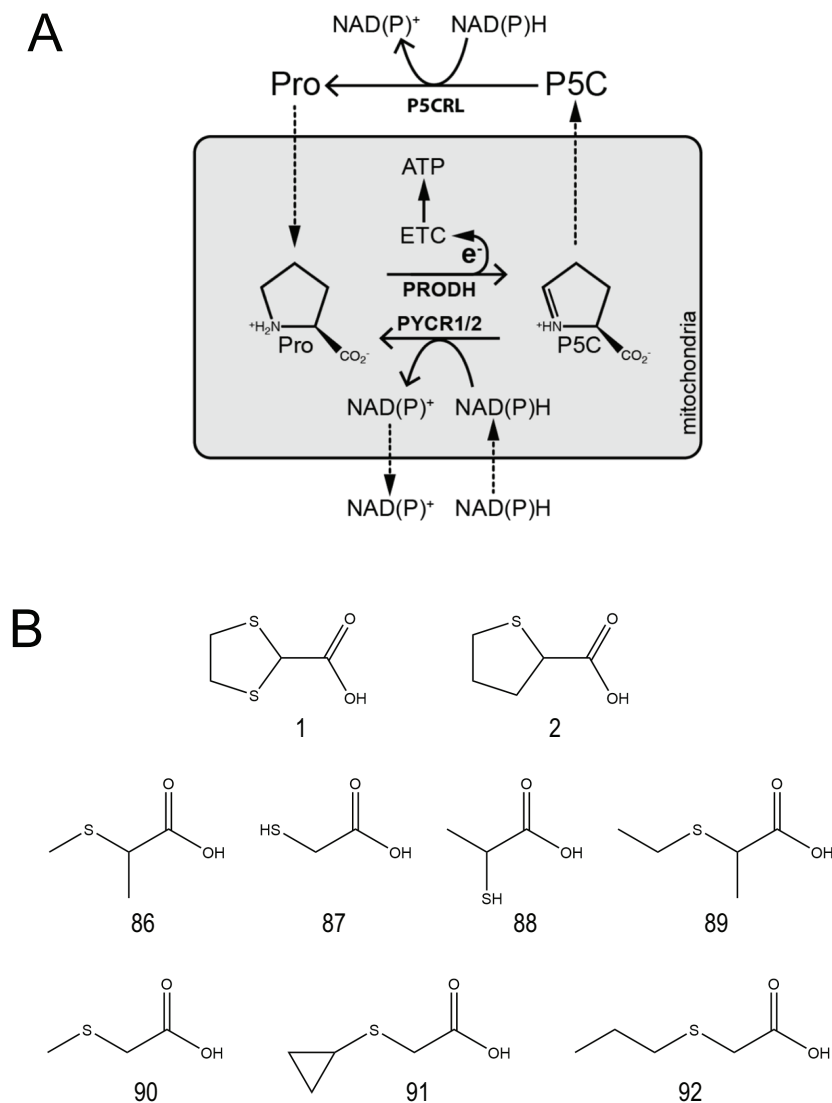
**Table 4.2** X-ray diffraction data collection and refinement statistics

	Compound			
	90 (Covalent Complex)	91 (Covalent Complex)	92 (Covalent Complex)	92 (Noncovalent Complex)
Beamline	APS (24-ID-C)	APS (24-ID-C)	APS (24-ID-C)	APS (24-ID-E)
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit cell parameters (Å)	$a = 101.38$ $b = 102.70$ $c = 127.25$ $\beta = 106.46$	$a = 101.55$ $b = 102.65$ $c = 127.08$ $\beta = 106.49$	$a = 101.49$ $b = 102.62$ $c = 126.63$ $\beta = 106.54$	$a = 101.57$ $b = 102.82$ $c = 127.17$ $\beta = 106.41$
Wavelength (Å)	0.97911	0.97911	0.97911	0.97918
Resolution (Å)	102.70-1.53 (1.55-1.53)	59.85-1.65 (1.67-1.65)	89.33-1.76 (1.79-1.76)	121.99-1.70 (1.73-1.70)
Observations <sup>a</sup>	1408841 (61774)	1130558 (49975)	1118750 (46402)	1039007 (42211)
Unique reflections <sup>a</sup>	373473 (17138)	295995 (13610)	244007 (11261)	263234 (11106)
$R_{\text{merge}}(I)^a$	0.039 (0.844)	0.062 (0.966)	0.180 (1.411)	0.039 (0.253)
$R_{\text{meas}}(I)^a$	0.045 (0.991)	0.073 (1.118)	0.203 (1.608)	0.045 (0.293)
$R_{\text{pim}}(I)^a$	0.023 (0.512)	0.036 (0.556)	0.094 (0.758)	0.022 (0.144)
Mean $I/\sigma^a$	16.1 (1.3)	12.8 (1.2)	7.3 (0.9)	20.5 (4.1)
$CC_{1/2}^a$	0.999 (0.550)	0.998 (0.506)	0.991 (0.244)	0.999 (0.943)
Completeness (%) <sup>a</sup>	98.9 (91.8)	98.3 (91.5)	98.7 (92.2)	95.6 (81.4)
Multiplicity <sup>a</sup>	3.8 (3.6)	3.8 (3.7)	4.6 (4.1)	3.9 (3.8)
$R_{\text{cryst}}^{a,c}$	0.237 (0.319)	0.232 (0.325)	0.237 (0.334)	0.227 (0.236)
$R_{\text{free}}^{a,b,c}$	0.253 (0.345)	0.256 (0.340)	0.267 (0.369)	0.254 (0.272)

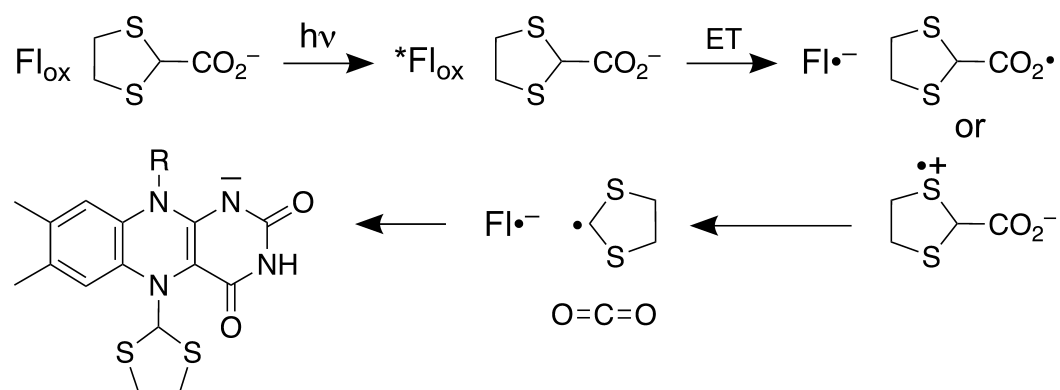
<sup>a</sup>Values for the outer resolution shell of data are given in parenthesis.

<sup>b</sup>5% test set.

<sup>c</sup> $R$ -factors are from the first round of refinement without extensive model building or inclusion of solvent.

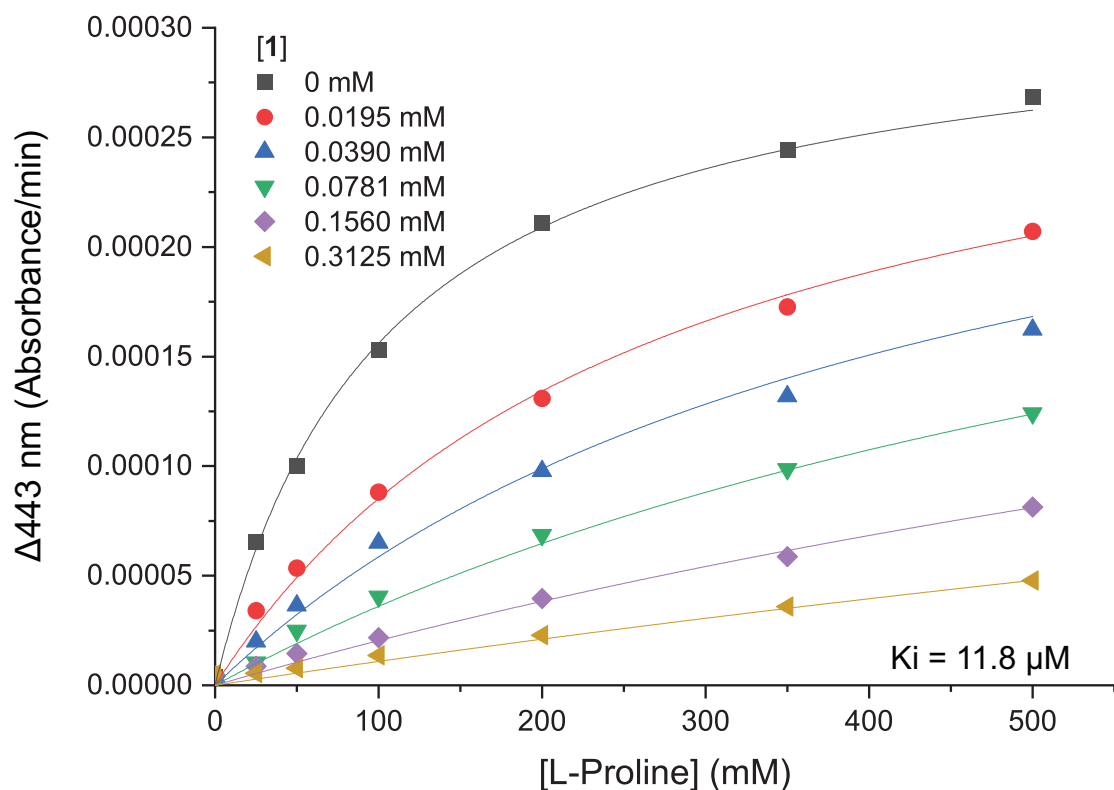


**Figure 4.1** The proline cycle and chemical structures of compounds studied. (A) Enzymes and reactions of the proline cycle. Figure taken from Bogner et al., 2021.<sup>2</sup> (B) Chemical structures of compounds selected for this study based off their similarity in sulfur placement to **1**. The compounds are numbered as: 2-(methylthio)propanoic acid (**86**), 2-sulfanylacetic acid (**87**), 2-sulfanylpropanoic acid (**88**), 2-(ethylsulfanyl)propanoic acid (**89**), 2-(methylsulfanyl)acetic acid (**90**), 2-(cyclopropylsulfanyl)acetic acid (**91**), and 2-(propylsulfanyl)acetic acid (**92**).

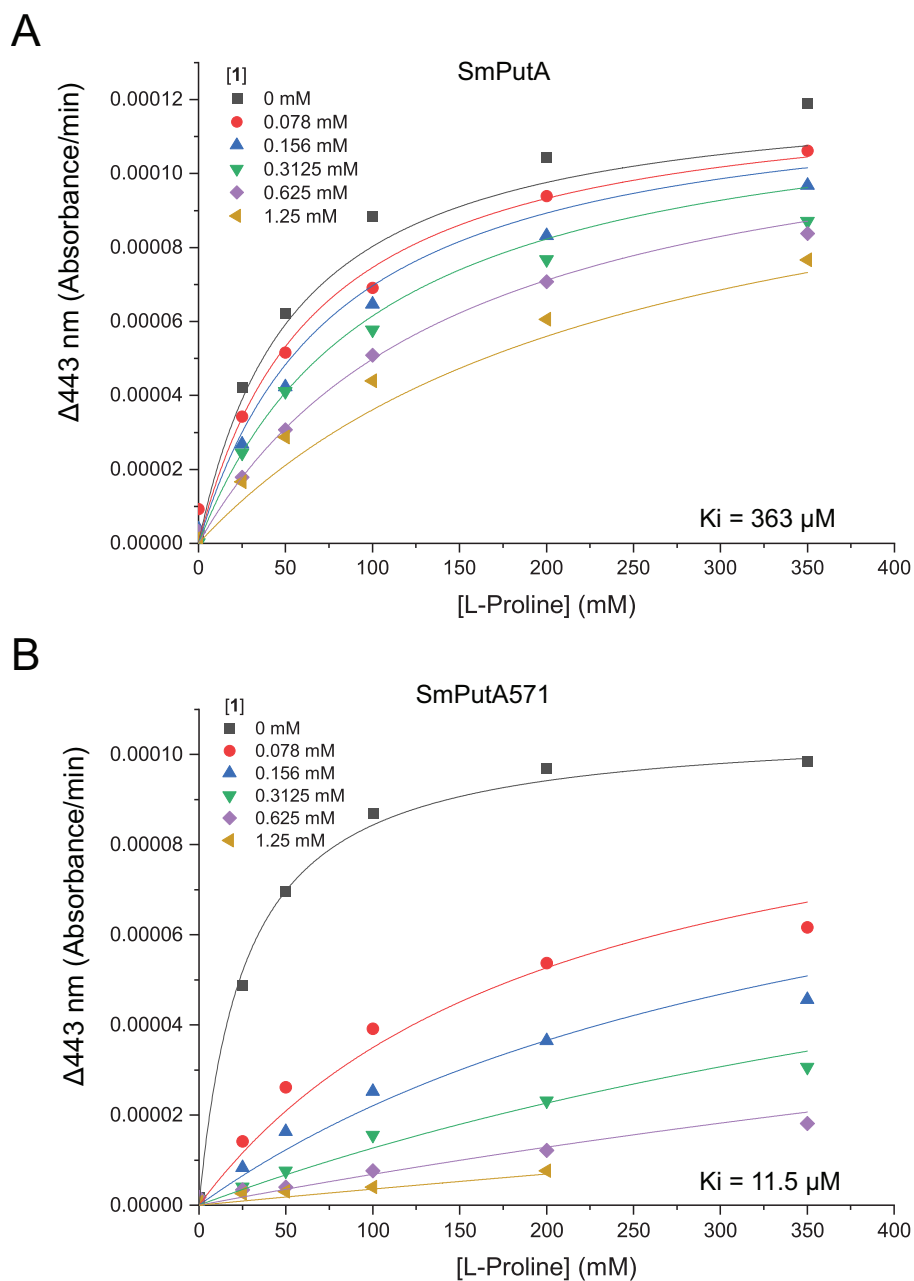


**Figure 4.2** Proposed mechanism of photoinduced inactivation by **1**. Figure taken from Campbell et al., 2021.<sup>14</sup>

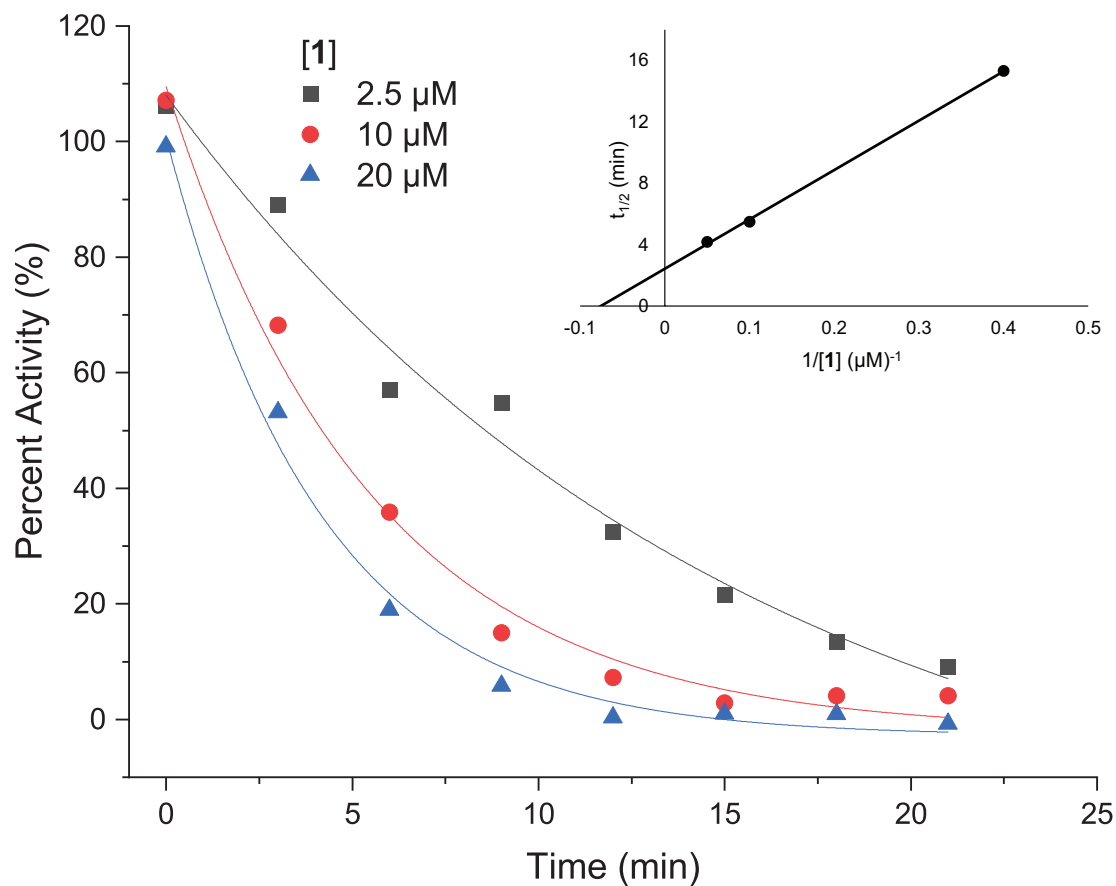




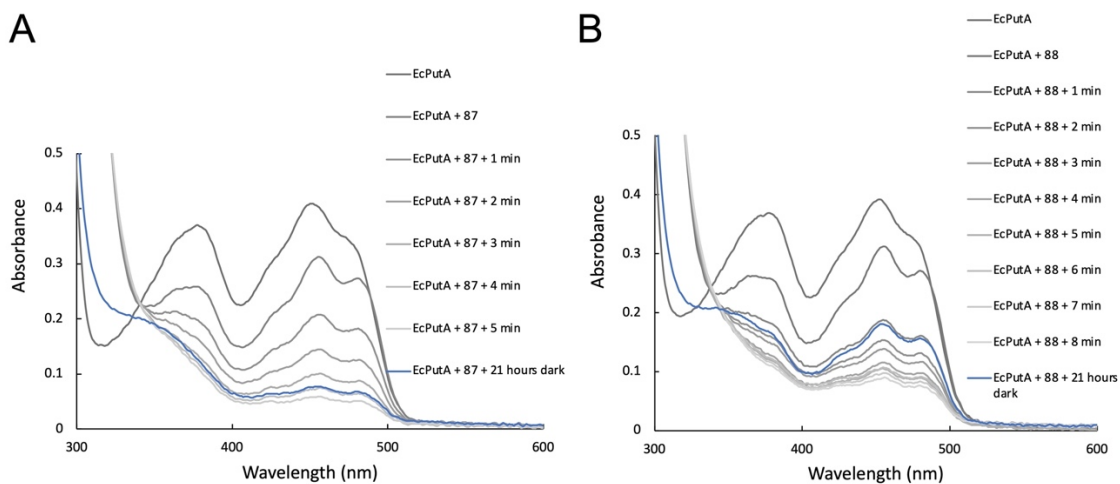
**Figure 4.3** Inhibition of the PROD H activity of PutA86-630 by **1** in the absence of light. The assays were performed at room temperature with 0-500 mM L-proline, 0-0.3125 mM **1**, 4 mM *o*-AB, 0.15 mM menadione, and 63 nM PutA86-630 in a buffer containing 20 mM MOPS pH 7.5 and 10 mM MgCl<sub>2</sub>. The data for each inhibitor were analyzed by global fitting to the competitive inhibition model using Origin software.



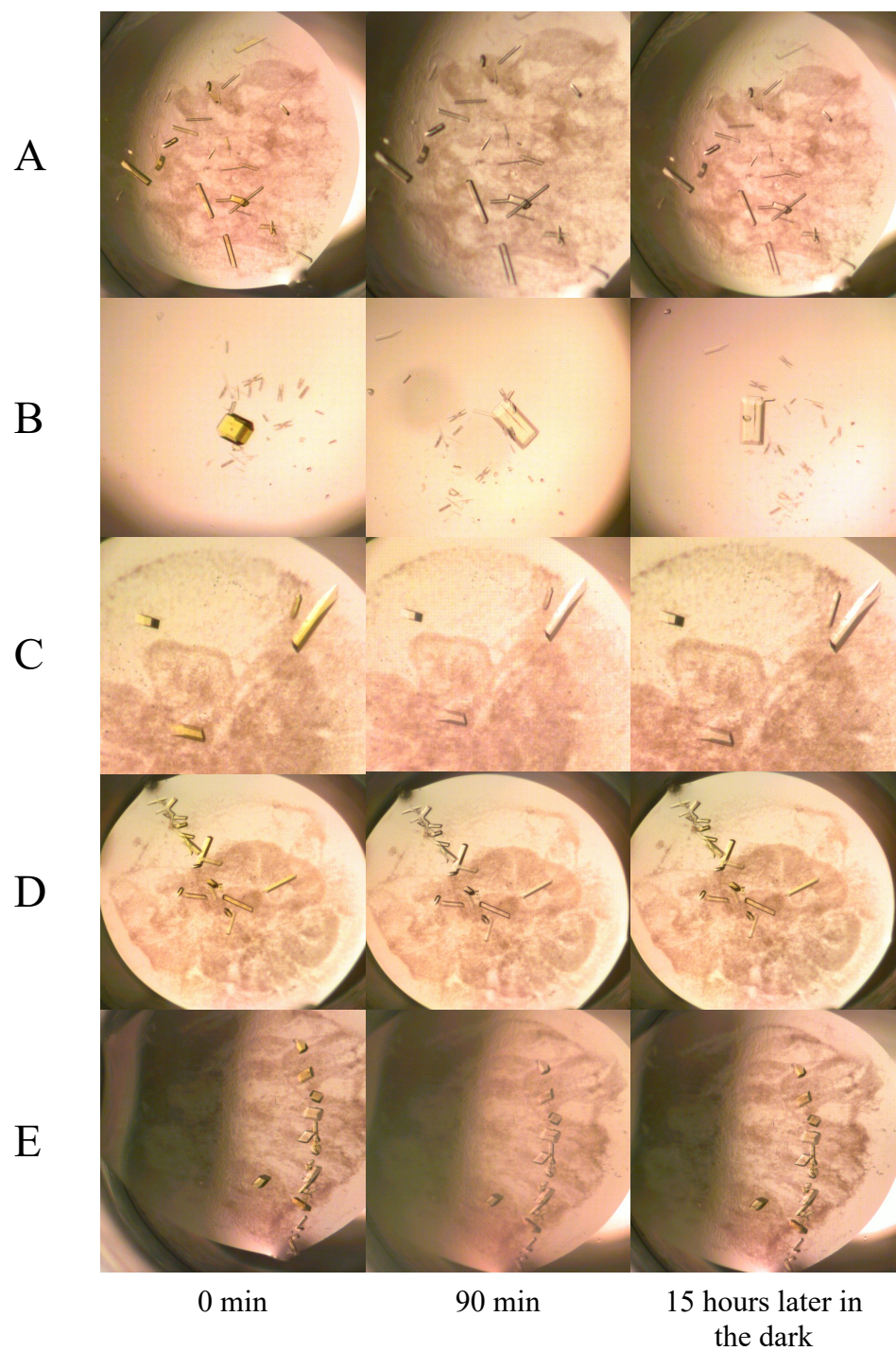
**Figure 4.4** Comparison of SmPutA and SmPutA571 PRODH activity inhibition by **1**. The assays were performed at room temperature with 0-350 mM L-proline, 0-1.25 mM **1**, 4 mM *o*-AB, 0.15 mM menadione, and 63 nM SmPutA (A) or 63 nM SmPutA571 (B) in a buffer containing 20 mM MOPS pH 7.5 and 10 mM MgCl<sub>2</sub>. The data for each inhibitor were analyzed by global fitting to the competitive inhibition model using Origin software.



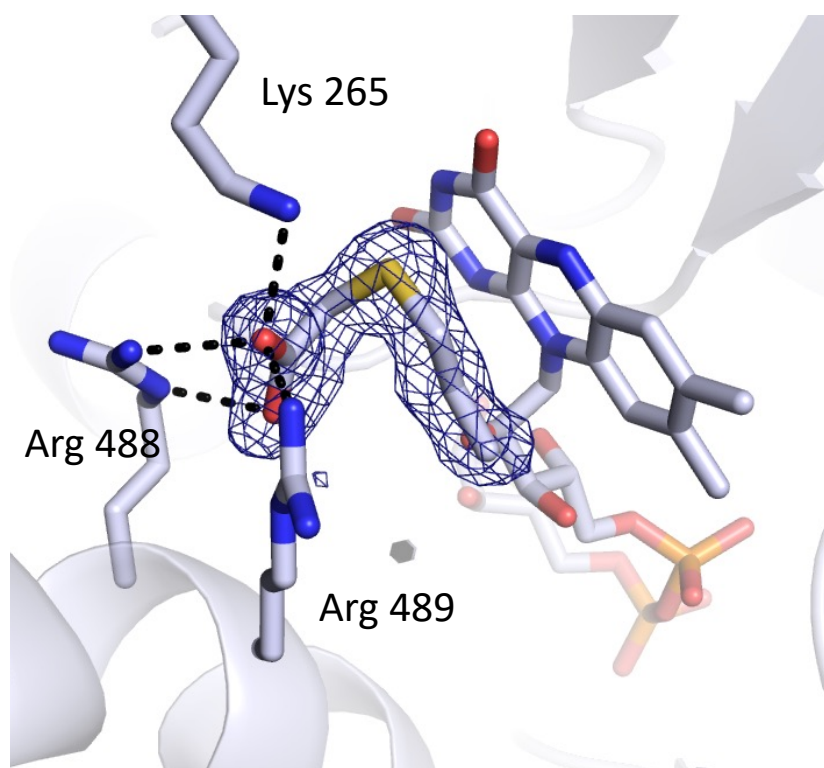
**Figure 4.5** Kinetics of photoinduced covalent inactivation of PutA86-630 by **1**. Using the *o*-AB assay, PRODH activity remaining after incubation of enzyme and varying concentrations of **1** is plotted as percent activity as a function of time for three inactivator concentrations (2.5, 10, and 20  $\mu\text{M}$ ). The inset shows the replot of the time constant of inactivation ( $t_{1/2}$ ) as a function of the reciprocal concentration of **1**. The following inactivation parameters were obtained from the fitting:  $k_{\text{inact}} = 0.26 \pm 0.02 \text{ min}^{-1}$  and  $K_i = 11.0 \pm 1.4 \text{ }\mu\text{M}$ .



**Figure 4.6** Demonstration of the reversibility of photoinduced flavin spectral changes. PutA86-630 (3 mg/mL) was combined with 25 mM **87** (A) or 25 mM **88** (B) and exposed to 0-5 min of white light or 0-8 min of white light respectively. Samples were then kept in the dark for 21 hours and remeasured (blue curve).

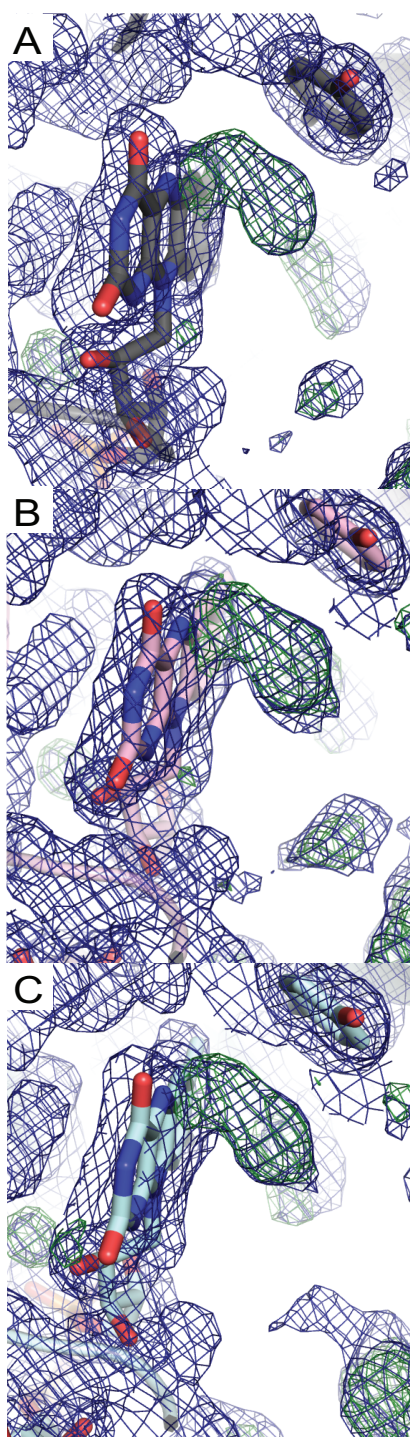


**Figure 4.7** Observation of the bleaching of SmPutA crystals after 90 min of blue light exposure. The enzyme was co-crystallized with 50 mM (A) **92** (B) **91** (C) **90** (D) **86** (E) **89**. In the case of (D) and (E), crystals returned to a yellow color after 15 hours spent in the dark.



**Figure 4.8** Interactions of **92** noncovalently bound in the SmPutA active site. The blue cage represents a polder omit map ( $4\sigma$ ).





**Figure 4.9** Electron density evidence for covalent modification of the N5 atom of FAD. SmPutA was co-crystallized with 50 mM of (A) **90**, (B) **91**, or (C) **92** and the crystals were exposed to 90 min of blue light. The blue and green cages represent the  $2F_o-F_c$  ( $1\sigma$ ) and  $F_o-F_c$  ( $3\sigma$ ) maps after the initial round of refinement.

## **VITA**

Alexandra N. Bogner was born in Grand Rapids, Michigan to Gary and Rhonda Bogner. Alexandra attended Calvin University for her B.S. in Biochemistry and B.A. in French. She began her PhD at the University of Missouri – Columbia in 2017 and joined Dr. John J. Tanner’s laboratory in the spring of 2018 and completed her graduate work in the fall of 2021.