UNIVERSITY OF CALIFORNIA
RIVERSIDE

NMR Crystallography and Integrative Structural Biology of Enzyme Complexes

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Biochemistry and Molecular Biology

by

Rittik Kumar Ghosh

March 2023

Dissertation Committee:
Dr. Leonard J Mueller, Chairperson
Dr. Michael F Dunn
Dr. Russ C Hille

The Dissertation of Rittik Kumar Ghosh is approved:

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE DISSERTATION

NMR Crystallography and Integrative Structural Biology of Enzyme Complexes

by

Rittik Kumar Ghosh

Doctor of Philosophy, Graduate Program in Biochemistry and Molecular Biology
University of California, Riverside, March 2023
Dr. Leonard J. Mueller, Chairperson

This dissertation aims to provide a full mechanistic understanding of enzyme catalysis by investigating the active site protonation states and chemical dynamics. Accurate knowledge of proton positions and sidechain orientations is crucial for establishing the mechanism of acid-base catalysis and allostery in enzymes These finite details are necessary for developing various computational techniques such as molecular dynamics simulations, molecular docking routines, and structure-based drug design. While high-resolution X-ray crystal structures provide information about protein residues and cofactors interacting with the substrate, they often fail to determine protonation states due to insufficient resolution. Currently, neutron crystallography is making strides in defining heavy and hydrogen atom locations but is limited by the need for large, preferably perdeuterated, crystals and multi-week-long acquisition times. Cryo-EM is also making progress in hydrogen atom detection but is far from routine.

Nuclear magnetic resonance (NMR) spectroscopy has been proven to be a sensitive tool for detecting chemical environment in enzyme active sites and can be combined with Cryo-EM and diffraction methods for atomic-resolution descriptions of structure and function. However, to delineate the chemistry of the active site, NMR and diffraction are

more powerful when combined with first-principles computational chemistry. This dissertation presents a comprehensive investigation of enzyme catalysis by combining X-ray Crystallography, NMR Spectroscopy and First Principles Calculations in a synergistic approach of NMR Crystallography to determine enzyme mechanism and establish a full understanding of the catalytic pathway.

Acknowledgement

Graduate school had its ups and downs, but overall, it was a fantastic and productive experience for me. There are many people who contributed to my success, and while this list is just the beginning, I am incredibly grateful for their support.

I would like to express my gratitude to the NIH for funding our research, and to the National High Magnetic Field Laboratory and its staff for providing us with an outstanding facility and for spending countless hours running experiments on our behalf. I would like to thank Joanna Long, Maria Luiza Caldas Nogueira, Faith Scott and Fred Mentink-Vigier for their contributions and whose hard work was instrumental in producing the DNP data. Their assistance was invaluable, and I am deeply appreciative of all the time and effort they put into helping us. I would also like to thank our Organic Chemistry collaborator Dr. Kevin Kou for providing the organic synthesis expertise that is very important to the research in this dissertation.

I am grateful to my committee members, Len Mueller, Mike Dunn, and Russ Hille, for their continuous support and valuable feedback throughout my academic journey. Their genuine interest in my research and guidance has made me a better scientist, and I owe a great deal of my success to their contributions.

I would also like to express my appreciation to all the staff members who have made my time in graduate school enjoyable and memorable. Your assistance and support have been invaluable.

There are many faculty members who have played a significant role in my academic development, including Len Mueller, Mike Dunn, Russ Hille, Stephanie Dingwall, Gregor Blaha, Sean O' Leary, Jikui Song, Chia-en Chang and Maitree Bhattacharyya. Your ideas and feedback have been greatly appreciated, and you have set great examples for

me as both a researcher and an educator. Thank you for your mentorship and for contributing to my growth as a PhD scholar.

I am deeply grateful to all my lab mates and friends throughout the years for their unwavering support, assistance, humor, and affection. Manpreet Kaler, Jacob Holmes, Chris Williams, Dave Amarasinghe, Jennifer Romero, and former lab members Dr. Kevin Chalek and Dr. Viktoriia Liu, the memories we have made together, from in lab discussions to conferences and lunch outings, will always hold a special place in my heart. You have made my time here joyful, and it was a great ride!

I would like to extend my gratitude to Dimitri Niks for his invaluable troubleshooting tips and insightful conversations. Our discussions on proteins and enzyme kinetics in the hallway were some of the most cherished moments of my graduate school experience.

But more than that, I treasure our friendship and the time we have spent together. Thank you for everything, Dimitri.

I would like to express my thankfulness to some of the amazing individuals who have been my dear friends, family and support system throughout my journey: Nilanjan Sinha, Aritra Sarkar, Subhamita Sengupta, Pritha Ghosh, Chhaya Srivastava, Utkarsh Srivastava, Rounak Sinha, Debarati Barman, Somraj Ganguly, Rohan Rakshit, Dipanjan Chakraborty, Ayan Das, Varun Raviprolu, Pradeep Srivastava, Arnab Datta, Sayanti Acharjee, Kishore Maji, Dripta Sinha, Sandip Biswas, Mainak Roy, Agni Roy, Utsav Chakraborty, Soham Sarkar and many others. I apologize if I missed anyone, but each of you has played a significant role in providing me with unwavering support and encouragement. Thank you for being there for me during the ups and downs. I couldn't have asked for better friends.

I would like to express my deepest gratitude to my parents for their unwavering support, love, and belief in me. Your encouragement has been the driving force behind my academic success, and I am forever grateful for your faith in me.

Even though I may find your strict fullness slightly insane, I love you all the more for it. Thank you for always being there for me through the good times and the bad.

To my dear late sister Mamon, you have always been a pillar of strength in my life. I am certain that you would be overjoyed to witness this dissertation today. When nobody else had faith in me, you did. I miss you dearly and am grateful to have had you as my sister, even though we were not related by blood. I am confident that your blessings will accompany me as I continue my life journey.

To my PhD advisor Len Mueller, I am incredibly fortunate to have you as my PhD boss, mentor, ally, and a go to person for me whenever I have needed help. Len, I doubt I will ever be able to repay you for all that you have done for my career. I still remember the day when I was struggling to find a lab in the Spring quarter of my first year, and you took me in, gave me a platform to shine. I cannot imagine having had a better graduate school experience other than the Mueller lab, a place which has become a home away from home for me for the last five plus years.

You have always been there for me, providing unwavering support and guidance whenever I needed it. I will miss our daily interactions dearly, and now that I know I am graduating, it feels like a void that is not going to be filled anytime soon. Your impact on my life has been immeasurable, and I love and respect you from the bottom of my heart, now and always. I hope to live up to the example you have set and make you proud in all that I do. Thank you so much for everything.

I am incredibly grateful to have had the opportunity to work with two exceptional faculty members and individuals during my PhD journey, and my mentor Mike Dunn is one of them. I consider myself incredibly lucky to have had such a wonderful mentor and a giant in the field of enzymology who took me under his wing and provided me with invaluable knowledge and guidance.

Mike, your wisdom and expertise have been a constant source of support throughout my PhD journey, and I cannot thank you enough for everything you have done for me. Without your guidance and support, this thesis would not have been possible.

Your mentorship has played a vital role in my academic success, and I will forever cherish the memories of our time together. You and Linda have become like family to me, and I will truly miss our Friday afternoon mechanistic interactions and of course the Dunn family Thanksgiving party. Thank you so much for everything, Mike.

I express my deepest gratitude to Harshita Srivastava, my life partner, for being my constant source of support and showing incredible patience throughout our journey together. My achievements and successes would not have been possible without you by my side. You are an exceptional person with qualities that are unmatched, and I am lucky to have you in my life. As we move forward together, I look forward to deepening our bond and being there for you as you have been for me. I love you with all my heart.

The work in this thesis comes entirely or in part from the following publications:

Bosken, Y. K., Ai, R., Hilario, E., **Ghosh, R. K.,** Dunn, M. F., Kan, S. H., Niks, D., Zhou, H., Ma, W., Mueller, L. J., Fan, L., & Chang, C. E. A. (2021). Discovery of antimicrobial agent targeting tryptophan synthase. Protein Science, 31(2), 432-442. https://doi.org/10.1002/pro.4236

Appears in Chapter 6

> *Author Contributions:* Yuliana K. Bosken: Conceptualization (equal); formal analysis (equal); investigation (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). Rizi Ai: Conceptualization (equal); formal analysis (equal); investigation (equal); methodology (equal); validation (equal); visualization (equal); writing – original draft (equal). Eduardo Hilario: Formal analysis (equal); investigation (equal); validation (equal); visualization (equal); writing – original draft (supporting). Rittik K. Ghosh: Formal analysis (supporting); investigation (supporting); validation (supporting); visualization (supporting); writing – original draft (supporting). Michael F. Dunn: Investigation (supporting); methodology (supporting); resources (supporting); supervision (supporting); validation (supporting); writing – review and editing (supporting). Shih-Hsin Kan: Formal analysis (supporting); investigation (supporting); resources (supporting); validation (supporting). Dimitri Niks: Formal analysis (supporting); investigation (supporting); resources (supporting); validation (supporting). Huanbin Zhou: Formal analysis (supporting); investigation (supporting); resources (supporting); validation (supporting). Wenbo Ma: Investigation (supporting); resources (supporting); supervision (supporting); validation (supporting). Leonard J. Mueller: Formal analysis (supporting); funding acquisition (equal); investigation (supporting); project administration (supporting); resources (supporting); supervision (supporting); validation (supporting); writing – review and editing (supporting). Li Fan: Formal analysis (supporting); funding acquisition (equal); resources (equal). Chia-En A. Chang: Conceptualization (lead); formal analysis (equal); funding acquisition (lead); investigation (lead); methodology (lead); project administration (lead); resources (equal); supervision (lead); validation (equal); visualization (lead); writing – original draft (lead); writing – review and editing (lead).

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., **Ghosh, R. K.,** Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., Bogie, P. M., Paulino, J., Wang, X., Riviere, G., Bosken, Y. K., Struppe, J., Hassan, A., Guidoulianov, J., Perrone, B., Mentink-Vigier, F., Chang, C. A., Long, J. R., Hooley, R. J., Mueser, T. C., Dunn, M. F., & Mueller, L. J. (2022). Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. Proceedings of the National Academy of Sciences, 119(2), e2109235119. https://doi.org/10.1073/pnas.2109235119

Appears in Chapter 3, Chapter 6, Chapter 9.

> *Author Contributions:* M.F.D. and L.J.M. designed research; J.B.H., V.L., B.G.C., E.H., R.K.G., V.N.D., R.P.Y., J.A.R., J.P., X.W., G.R., Y.K.B., J.S., A.H., J.G., B.P.,

and F.M.-V. performed research; A.D.G., P.M.B., and R.J.H. contributed new reagents/analytic tools; C.-e.A.C., J.R.L., R.J.H., T.C.M., M.F.D., and L.J.M. analyzed data; and J.B.H., V.L., B.G.C., R.K.G., M.F.D., and L.J.M. wrote the paper.

Ghosh, R. K., Hilario, E., Chia-en, A. C., Mueller, L. J., & Dunn, M. F. (2022). Allosteric regulation of substrate channeling: Salmonella typhimurium tryptophan synthase. *Frontiers in Molecular Biosciences*, *9*.
(doi: https://doi.org/10.3389%2Ffmolb.2022.923042)

Appears in Chapter 3, Chapter 4.

Author Contributions: MD and LM conceived and designed the content of this review and MD wrote the final version of the manuscript. RG prepared the figures and RG and C-eC prepared sections of the text. EH solved many of the x-ray structures discussed in this manuscript. All authors contributed to the article and approved the submitted version.

Ghosh, R. K., Hilario, E., Liu, V., Wang, Y., Niks, D., Holmes, J. B., Sakhrani, V. V., Mueller, L. J., & Dunn, M. F. (2021). Mutation of βGln114 to Ala alters the stabilities of allosteric states in tryptophan synthase catalysis. *Biochemistry*, *60*(42), 3173-3186. https://doi.org/10.1021/acs.biochem.1c00383

Appears in Chapter 5.

Author Contributions: All authors have contributed to this work with primary contributions in the following areas. Overall research direction: M.F.D. and L.J.M. Mutation and X-ray crystallography: E.H. Protein purification: E.H., R.K.G., and J.B.H. UV–vis and SSNMR: R.K.G., J.B.H., D.N., V.V.S., M.F.D., and L.J.M. Computational Chemistry: R.K.G., V.L., Y.W., L.J.M., and M.F.D. The final drafts of the manuscript were written by M.F.D and L.J.M.

Sakhrani, V. V., **Ghosh, R. K.,** Hilario, E., Weiss, K. L., Coates, L., & Mueller, L. J. (2021). Toho-1 β-lactamase: backbone chemical shift assignments and changes in dynamics upon binding with avibactam. *Journal of biomolecular NMR*, *75*, 303-318. (doi: https://doi.org/10.1007/s10858-021-00375-9)

Appears in Chapter 7.

Author Contributions: All authors have contributed to this work with primary contributions in the following areas. Overall research direction: L.J.M and L.C. Solution State NMR Spectroscopy: V.VS, R.K.G. Protein purification: K.L.W., V.V.S., E.H., and R.K.G.

Table of Contents

List of Figures and Figure Captions

List of Tables

List of Schemes

Chapter 1

Introduction to the Thesis

The primary focus of this thesis centers on the significance of comprehending the protonation states present at the active site of enzymes to understand the mechanisms of enzyme catalysis. Accurate proton positions and the dynamics of the active site are crucial to establish mechanisms of acid-base catalysis and to provide valuable input for molecular dynamics simulations, molecular docking routines, and the design of drugs based on the structure of the enzyme (Mueller & Dunn, 2013).

Enzymology has traditionally focused on elucidating enzyme catalysis at the chemical and structural level, using high-resolution X-ray crystal structures to identify interacting protein residues and cofactors with the substrate (Michael F Dunn, 2012; Michael F Dunn et al., 2008; Limbach et al., 2011). However, determining the protonation states of acid-base functional groups at the catalytic site has been challenging. It is typically inferred from hydrogen-bonding patterns and pKa values of small molecules in aqueous solution due to the limited resolution of protein structures (Jacob B. Holmes et al., 2022). Recent advancements in neutron crystallography have enabled the precise determination of heavy atom and light hydrogen atom locations when combined with X-ray crystallography (Kovalevsky et al., 2008; Michalczyk et al., 2015; Tomanicek et al., 2013). However, this technique requires large, preferably perdeuterated crystals and multi-week-long acquisition times at room temperature, which currently limits its ability to analyze reactive intermediates and fully understand the mechanism of enzyme function. Cryo-electron microscopy (cryo-EM) has achieved remarkable resolution, but it still faces challenges in routine detection of hydrogen atoms (Perilla et al., 2017).

An alternative method for studying the chemical environment of enzyme active sites is through nuclear magnetic resonance (NMR) spectroscopy. NMR, when combined with cryo-EM and diffraction techniques, has led to atomic-resolution descriptions of enzyme structure and function. The power of NMR and diffraction methods in delineating the chemistry of active sites can be enhanced by integrating them with first-principles computational chemistry (Schütz, 2021).

Thus, my research focuses on developing and utilizing the "NMR Crystallography" methodology, which combines X-ray crystallography, NMR spectroscopy, and first-principles calculations to uncover the chemical environment and reaction mechanism of enzymes. This approach allows for the precise determination of the positions of all atoms in the active site, including hydrogen atoms, providing an invaluable tool for integrated structural biology.

Integrating NMR crystallography with enzyme kinetics can provide a more complete understanding of enzyme function. NMR can provide information on the conformational dynamics of enzymes, while crystallography can provide high-resolution structures of the enzyme in different states. By combining this information with enzyme kinetic data, we can gain insights into how enzyme conformational changes relate to catalytic function.

For instance, NMR can detect subtle changes in protein dynamics that may be relevant to enzyme activity. Crystallography can then be used to identify the specific structural changes that occur upon substrate binding or during the reaction. Enzyme kinetics can provide information on the rate of the reaction and the effects of inhibitors or mutations on enzyme activity.

Integrating these approaches can help us understand the complex interplay between enzyme structure, dynamics, and function. It can also provide a basis for designing more

effective enzyme inhibitors or developing enzymes with improved catalytic activity. Ultimately, the insights gained from this research may pave the way for designing more effective enzyme inhibitors or developing enzymes with improved catalytic activity.

References

Dunn, M. F. (2012). Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Archives of biochemistry and biophysics*, *519*(2), 154-166.

Dunn, M. F., Niks, D., Ngo, H., Barends, T. R., & Schlichting, I. (2008). Tryptophan synthase: the workings of a channeling nanomachine. *Trends in biochemical sciences*, *33*(6), 254-264.

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., Ghosh, R. K., Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., Bogie, P. M., Paulino, J., Wang, X., Riviere, G., Bosken, Y. K., Struppe, J., Hassan, A., Guidoulianov, J., Perrone, B., Mentink-Vigier, F., . . . Mueller, L. J. (2022). Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *Proceedings of the National Academy of Sciences*, *119*(2), e2109235119. https://doi.org/doi:10.1073/pnas.2109235119

Kovalevsky, A. Y., Katz, A. K., Carrell, H. L., Hanson, L., Mustyakimov, M., Fisher, S. Z., Coates, L., Schoenborn, B. P., Bunick, G. J., Glusker, J. P., & Langan, P. (2008). Hydrogen Location in Stages of an Enzyme-Catalyzed Reaction: Time-of-Flight Neutron Structure of d-Xylose Isomerase with Bound d-Xylulose. *Biochemistry*, *47*(29), 7595-7597. https://doi.org/10.1021/bi8005434

Limbach, H.-H., Chan-Huot, M., Sharif, S., Tolstoy, P. M., Shenderovich, I. G., Denisov, G. S., & Toney, M. D. (2011). Critical hydrogen bonds and protonation states of pyridoxal 5′-phosphate revealed by NMR. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1814*(11), 1426-1437.

Michalczyk, R., Unkefer, C. J., Bacik, J.-P., Schrader, T. E., Ostermann, A., Kovalevsky, A. Y., McKenna, R., & Fisher, S. Z. (2015). Joint neutron crystallographic and NMR solution studies of Tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proceedings of the National Academy of Sciences*, *112*(18), 5673-5678. https://doi.org/doi:10.1073/pnas.1502255112

Mueller, L. J., & Dunn, M. F. (2013). NMR Crystallography of Enzyme Active Sites: Probing Chemically Detailed, Three-Dimensional Structure in Tryptophan Synthase. *Accounts of chemical research*, *46*(9), 2008-2017. https://doi.org/10.1021/ar3003333

Perilla, J. R., Zhao, G., Lu, M., Ning, J., Hou, G., Byeon, I.-J. L., Gronenborn, A. M., Polenova, T., & Zhang, P. (2017). CryoEM Structure Refinement by Integrating NMR Chemical Shifts with Molecular Dynamics Simulations. *The Journal of Physical Chemistry B*, *121*(15), 3853-3863. https://doi.org/10.1021/acs.jpcb.6b13105

Schütz, A. K. (2021). Solid-state NMR approaches to investigate large enzymes in complex with substrates and inhibitors. *Biochemical Society Transactions*, *49*(1), 131-144.

Tomanicek, S. J., Standaert, R. F., Weiss, K. L., Ostermann, A., Schrader, T. E., Ng, J. D., & Coates, L. (2013). Neutron and X-ray Crystal Structures of a Perdeuterated Enzyme Inhibitor Complex Reveal the Catalytic Proton Network of the Toho-1 β-Lactamase for the Acylation Reaction *. *Journal of Biological Chemistry*, *288*(7), 4715-4722. https://doi.org/10.1074/jbc.M112.436238

Chapter 2

Brief Overview of NMR Spectroscopy

2.1 Basic NMR Concepts and Theory

Nuclear Magnetic Resonance (NMR) is a highly effective technique for investigating local chemical structure, which can provide information on a variety of properties such as local geometry, hybridization, ionization, and protonation states. NMR parameters can also reveal dynamic processes such as conformational and chemical exchange, with proton transfers being a notable example of the latter. This allows for the determination of general acid-base chemistry associated with the system under study. (Claridge, 2016; Reif et al., 2021)

In addition to the conventional NMR techniques, solid-state NMR (ssNMR) can provide valuable information about the anisotropic nature of the local electronic environment. This section aims to briefly introduce the fundamental concepts and terminology related to the NMR phenomenon which will be utilized throughout this dissertation. (Hodgkinson, 2020)

Nuclei that are NMR active possess a magnetic dipole moment (μ) and an intrinsic magnetic spin angular momentum, which is characterized by the nuclear spin quantum number (*I*). The *I* values can be either higher or half-integer, but those with *I* = 0 are not NMR active (Keeler, 2010; Sieklucka & Pinkowicz, 2017). The magnitude of the spin angular momentum, which is a vector quantity, can be calculated using Eq (1).

$$|I| = [I.I]^{1/2} = \hbar[I(I+1)]^{1/2} \quad (1)$$

The magnetic moment is proportional to the magnitude of the spin angular momentum by

$$\mu = \gamma\hbar[I(I+1)]^{1/2} \quad (2)$$

where γ is the isotope specific gyromagnetic ratio most often reported in units of radian per second per Tesla (rad s^{-1} T^{-1}). When an external magnetic field is applied, the degeneracy of magnetic energy in the nuclear ground state is lifted (Zeeman splitting) resulting in 2I + 1 equally spaced energy levels (Hore, 2015; Levitt, 2013).

When photons with frequency ν (Hz) are in resonance with the energy differences, the transition between energy levels $\upsilon_{m\to m'} = h^{-1}\,\Delta E_{m\to m'} = \gamma B_o/2\pi$ can be expressed in angular frequency units (rad s^{-1}) as $\omega_{m\to m'} = \gamma B_o$. The energy difference between these transitions is relatively small, as seen with a ^{1}H nucleus having a gyromagnetic ratio of $\gamma = 2.68 \times 10^8\ rad\ s^{-1}T^{-1}$ (where *T* is the magnetic field strength in Tesla), and the energy difference:

$$\Delta E_{m\to m'} = \hbar\,\gamma\,B_o(\Delta m = m \to m' = \pm 1) \quad (3)$$

in a $B_o = 21.4\,T$ magnetic field being only 6×10^{-25} J. This value is smaller than the thermal energy product k_bT (T = temperature), where $k_b$is the Boltzmann constant. When a static magnetic field B_o is applied, the magnetic moments of the nuclei precess about the z-axis of the laboratory frame with a characteristic frequency, $\omega_o = -\gamma B_o\ (rad\ s^{-1})$ or $\nu_o = -\frac{\gamma B_o}{2\pi}\,(Hz)$, which is referred to as the Larmor frequency (Hore, 2015; Hore et al., 2015). The energy difference corresponding to the orientation of the magnetic dipole moment in the static field causes a bias for the low energy alignment, resulting in a population difference between $m_i = +\frac{1}{2}\ and\ m_i = -\frac{1}{2}$. Over time, this leads to a net, or bulk magnetization in the sample at thermal equilibrium that is aligned with the static field and is denotated as M_z^{eq}. For spin – ½ nuclei, the total number of spins in the low energy level, $N^{+1/2}$, versus those in the higher level $N^{-1/2}$ follows the Boltzmann distribution, which is given by the following Eq (4).

$$\frac{N^-}{N^+} = e^{-\Delta E/k_b T} = e^{-\gamma \hbar B_o/k_b T} \quad (4)$$

When a radiofrequency pulse perpendicular to the static magnetic field is applied in the x-y plane of the lab frame, it creates a new field known as the B1. If this field is resonant with the Larmor frequency, the magnetic dipoles start precessing around it. A 90º pulse equalizes the population difference and creates phase coherence that leads to the generation of a net magnetization vector, M_{xy}, in the transverse plane (x-y), which denotes the magnetization in that plane. The components of this magnetization precess around the z-axis once the pulse is completed. In the absence of any other factors, a bare nucleus would precess at the Larmor frequency in the transverse plane. However, in molecules and atoms, the electrons' movement in the magnetic field induces local fields that slightly shield the nuclei. As a result, the effective field at a given nuclear site is given by Eq (5), which depends on the shielding constant, σ. decay

$$B_{eff} = (1 - \sigma)B_o \quad (5)$$

This constant is a rank two tensor and varies with the magnetic field's orientation. However, in solution, σ is averaged to a single observable value. This shielding leads to the precession of magnetic moments at slightly different frequencies depending on the electronic and chemical environment, as described by Eq (6).

$$\nu_i(Hz) = \frac{-\gamma B_{eff}}{2\pi} \quad (6)$$

This phenomenon is known as chemical shift and is detectable despite being relatively small, on the order of parts per million (ppm). When the magnetic moments precess in the transverse plane, they induce a current in NMR instrument's detection coil, which generates the NMR signal.

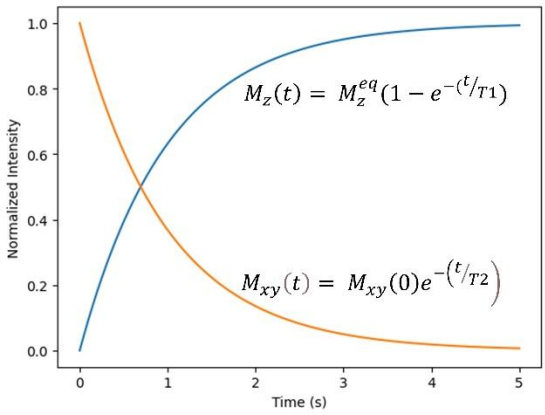


Fig 2.1: Plot of the intensities vs. time of the transverse, M_{xy}, and longitudinal, M_z, magnetization following a 90° pulse.

The NMR signal, which is generated by the precession of magnetic moments in the transverse plane induced by the pulse, diminishes over time due to the loss of coherence resulting from the variations in local magnetic fields at each nucleus caused by relative motion of the molecules. This decay process in characterized by a decay constant called, T_2, which is associated with the transverse relaxation. In addition to T_2, there is another decay constant called the T_1 that governs the restoration of the pre-pulse thermal equilibrium populations of the magnetic moments. Fig 2.1 illustrates the decay of transverse magnetization (M_{xy}) and the restoration of the equilibrium magnetization (M_z^{eq}) through longitudinal relaxation. The NMR signal obtained after the pulse is known as a free-induction decay (FID) in the time domain.

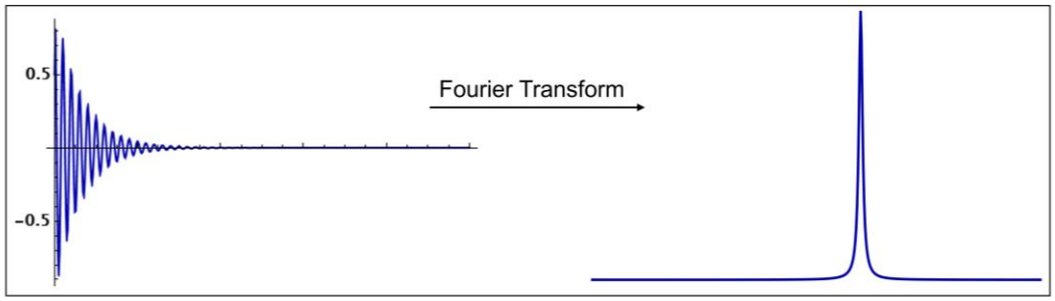


Fig 2.2 Fourier transform of the real component of the FID (time domain) gives the NMR spectrum in the frequency domain in which the signal is an absorptive Lorentzian shown here at zero offset with the carrier frequency.

When the Fourier transform is applied to the FID, it produces real and imaginary spectra that vary with frequency. The absorption Lorentzian component of the spectrum is given by Eq (7) and is shown in Fig 2.2.

$$S(\omega) = \frac{R_2}{R_2^2+4\pi^2+(\nu-\nu_0)^2} \tag{7}$$

The peak offset frequency is represented by ν_0 and $R_2 = 1/T_2$. Fig 2.2 displays the NMR spectrum in the frequency domain, where the signal is an absorptive Lorentzian at zero offset with the carrier frequency. There are various factors contributing to transverse relaxation, but only those relevant to the discussion will be addressed explicitly. The fluctuations in the local fields, B_{loc}, which affect both longitudinal and transverse relaxation, can be described by the complex spectral density, $J(\omega) + iQ(\omega)$. The real component, $J(\omega)$, reflects the density of motions at specific frequencies, particularly $J(0)$, the zero frequency, $J(\omega_0)$, the Larmor frequency, and the $J(2\omega_0)$, twice the Larmor frequency.

$$J_n = \frac{\tau_c}{1+(\eta\omega_0-\tau_c)^2}, \eta = 0, 1, or\ 2 \tag{8}$$

The spectral density at specific frequencies of interest, namely zero, Larmor, and twice the Larmor frequencies, are denoted by J_0, J_1, J_2 respectively, and their reduced forms

are given in Eq (8). The rotational correlation time, τ_c, which represents the time for an arbitrary point on a molecule to traverse approximately 1 rad, is crucial for understanding the effects of spectral density on longitudinal and transverse relaxation rates. The J_0, J_1, J_2 components affect transverse relaxation, characterized by T_2, whereas J_1, J_2 are responsible for longitudinal relaxation, characterized by T_1. The relationship between rotational correlation time and relaxation rates is demonstrated in Figure 2.3, where T_1 and T_2 are plotted against τ_c, namely fast isotropic reorientational motion for $\omega_0\tau_c << 1$, intermediate motion for $\omega_0\tau_c \sim 1$, and slow isotropic motion for $\omega_0\tau_c >> 1$, with extreme narrowing being the limit of fast isotropic reorientational motion.

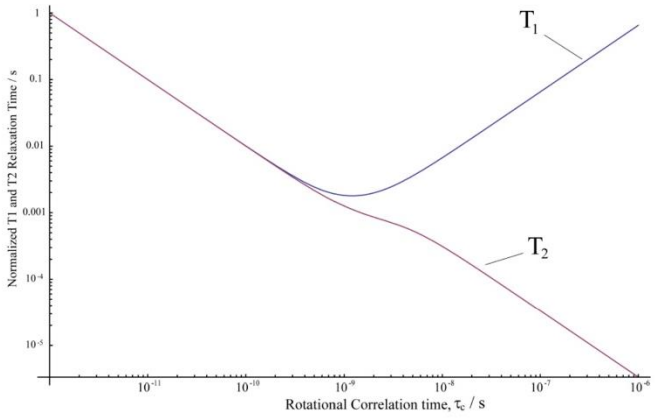


Fig 2.3. Relaxation time constants T1 and T2 as a function of the rotational correlation time, τc . (Redrawn from (Claridge, 2016; Young, 2016)).

In non-viscous solvents, small molecules exhibit short rotational correlation times (pico-seconds), whereas large molecules like proteins and enzymes display longer timescales (nano-seconds), as illustrated in Fig 2.3. The resulting NMR spectrum depends on the frequency, which is converted to a chemical shift spectrum to normalize measurements made on spectrometers with different field strengths. This conversion is achieved by referencing the frequency to a nucleus in an accepted reference compound, such as the

methyl protons in tetramethylsilane (TMS) for a ^{1}H, as shown in Fig 2.2. Although shielding is orientation dependent, rapid isotropic motion in the solution state averages it to a single observable value. The conversion from frequency to isotropic chemical shift is provided in Eq (9).

$$\delta_{iso}\ (ppm) = \frac{\nu_i(MHz) - \nu_{ref}(MHz)}{\nu_{ref}(MHz)} \times 10^6 \quad (9)$$

In ssNMR, access to the anisotropic shielding tensor provides even more information regarding the nucleus's immediate electronic environment. However, in practice, it is the chemical shift tensor, δ, that is experimentally measured (Duer, 2008; Levitt, 2013). The chemical shielding and shift are related, with δ being equal to the negative of the shielding (δ = - σ). The rank-2 chemical shift anisotropy (CSA) tensor, which is experimentally measured is given by Eq (10).

$$\delta_{ij}\ (ppm) = \begin{bmatrix} \delta_{xx} & \delta_{xy} & \delta_{xz} \\ \delta_{yx} & \delta_{yy} & \delta_{yz} \\ \delta_{zx} & \delta_{xy} & \delta_{zz} \end{bmatrix} \quad (10)$$

Diagonalization of the CSA tensor yields 3 eigen values, namely the principal components, δ_{11}, δ_{22}, and δ_{33} in the CSA principal axis system. These components are typically ordered as $\delta_{11} \geq \delta_{22} \geq \delta_{33}$ in Mehring notation (Mehring & Waugh, 1972). The isotropic chemical shift, δ_{iso}, is equal to the 1/3-rd of the trace of the CSA tensor and represents the average of the principal components, as given in equation (11).

$$\delta_{iso}\ (ppm) = \frac{1}{3}\ (\delta_{11} + \delta_{22} + \delta_{33}) \quad (11)$$

The isotropic chemical shift value is often sufficient to determine important details such as hybridization or protonation state, indicating the local chemical environment. Numerous approaches exist to further characterize the local environment using NMR spectroscopy, which are discussed later.

2.2 Magic Angle Spinning Cross Polarization (CPMAS) NMR

Magic Angle Spinning Cross Polarization (CPMAS) NMR is a powerful analytical technique used in bio-physical chemistry and materials sciences to investigate structure and function relationships of biomolecules and materials at an atomistic level. It combines two important NMR methods: MAS and CP, resulting in a highly sensitive and informative spectroscopic tool (Pines et al., 1972).

In MAS, a sample is placed inside a spinning rotor (~3.5 mm, 4 mm), which is rotated at different speeds around a specific axis. Usually, the sample is spun at high speed and at a “magic angle” of 54.74º, relative to the external magnetic field. This results in sharp lines with high resolution NMR spectra, giving us information about the isotropic chemical shifts, sensitive to the structure and microenvironment of the enzyme active site dynamics (Mehring & Waugh, 1972; Pines et al., 1972; Pines et al., 1973).

CP is a technique that utilizes the dipolar coupling interaction to transfer magnetization from one spin type to another and is commonly used to observe dilute spins like ^{13}C and ^{15}N with higher sensitivity. To initiate CP, an initial 90º pulse is applied on the ^{1}H, and the magnetization is then transferred to the ^{13}C nuclei by reintroducing the dipolar coupling between them with simultaneous spin locks on both nuclei.(Schaefer & Stejskal, 1976) One of the key benefits of CP measurements is the increase in signal sensitivity in a $^{13}C/^{15}N$ spectrum, and as a result, it speeds up the rate at which experiments can be repeated by setting the proton relaxation time, which is typically shorter than that for the carbon nuclei (Pines et al., 1972). However, magnetization can only be transferred between nuclei that are dipolarly coupled and are able to transfer or receive magnetization effectively in the solid state. However, this makes it difficult to observe signals from molecular structures that are undergoing rapid motions and exchanging

conformations; a problem which is alleviated to some extent by using Dynamic Nuclear Polarization (DNP) enhanced CPMAS NMR spectroscopy. Thus, CP, when coupled with MAS, yields a solid-state spectrum with resolution and sensitivity comparable to, and sometimes, surpassing, solution state experiments (Hu et al., 2004).

2.3 DNP-CPMAS NMR

DNP is a technique that transfers polarization from electrons to nuclei, resulting in a significant enhancement in NMR signal intensities. DNP-CPMAS NMR involves placing a sample in a microwave cavity at low temperatures (~ 100 K), where DNP can be used to transfer polarization from unpaired electrons to nuclei. The sample is then transferred to a MAS NMR probe where it is spun at a magic angle, and high-resolution spectra are obtained (Afeworki et al., 1992; Rosay et al., 2010).

One of the main advantages of DNP-CPMAS NMR is the ability to obtain high signal-to-noise ratios, which is particularly important for the study of materials with low natural abundance isotopes or low concentration samples. The sensitivity enhancement is achieved by the transfer of polarization from the electron spins to nuclei, resulting in signal enhancements of several orders of magnitude (Kobayashi et al., 2017).

DNP enhanced CPMAS NMR has been successfully used to study a wide range of materials, including biomolecules, pharmaceuticals, and polymers. The method has enabled the investigations of materials and dynamic systems that were previously too difficult to study using room temperature ssNMR experiments. This has provided important insights into the structure and dynamics of these systems.

This is still a developing field of NMR spectroscopy and in the coming years, this method has the potential to greatly expand the range of dynamic systems and materials that can

be studied using ssNMR, providing information rich high-resolution spectra with greatly enhanced signal-to-noise ratio.

2.4 Conclusions

This brief introduction to common NMR topics presents a number of ideas that will be referenced in the following chapters, specifically the technique of CPMAS and the associated theory behind it. Subsequent chapters will show established ssNMR methods in determining chemical structure and relate it to structure function relationship in enzyme systems.

2.5 References

Afeworki, M., McKay, R. A., & Schaefer, J. (1992). Selective observation of the interface of heterogeneous polycarbonate/polystyrene blends by dynamic nuclear polarization carbon-13 NMR spectroscopy. *Macromolecules*, *25* (16), 4084-4091.

Claridge, T. D. (2016). *High-resolution NMR techniques in organic chemistry* (Vol. 27). Elsevier.

Duer, M. J. (2008). *Solid state NMR spectroscopy: principles and applications*. John Wiley & Sons.

Hodgkinson, P. (2020). NMR crystallography of molecular organics. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *118*, 10-53.

Hore, P. J. (2015). *Nuclear magnetic resonance*. Oxford University Press, USA.

Hore, P. J., Jones, J. A., & Wimperis, S. (2015). *NMR: The toolkit: How pulse sequences work* (Vol. 92). Oxford University Press, USA.

Hu, K.-N., Yu, H.-h., Swager, T. M., & Griffin, R. G. (2004). Dynamic nuclear polarization with biradicals. *Journal of the American Chemical Society*, *126* (35), 10844-10845.

Keeler, J. (2010). *Understanding NMR spectroscopy*. John Wiley & Sons.

Kobayashi, T., Perras, F. A., Chaudhary, U., Slowing, I. I., Huang, W., Sadow, A. D., & Pruski, M. (2017). Improved strategies for DNP-enhanced 2D 1H-X heteronuclear correlation spectroscopy of surfaces. *Solid State Nuclear Magnetic Resonance*, *87*, 38-44.

Levitt, M. H. (2013). *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley & Sons.

Mehring, M., & Waugh, J. (1972). Magic-angle NMR experiments in solids. *Physical Review B*, *5* (9), 3459.

Pines, A., Gibby, M., & Waugh, J. (1972). Proton-enhanced nuclear induction spectroscopy. A method for high resolution NMR of dilute spins in solids. *The Journal of chemical physics*, *56* (4), 1776-1777.

Pines, A., Gibby, M. G., & Waugh, J. (1973). Proton-enhanced NMR of dilute spins in solids. *The Journal of chemical physics*, *59* (2), 569-590.

Reif, B., Ashbrook, S. E., Emsley, L., & Hong, M. (2021). Solid-state NMR spectroscopy. *Nature Reviews Methods Primers*, *1* (1), 2.

Rosay, M., Tometich, L., Pawsey, S., Bader, R., Schauwecker, R., Blank, M., Borchard, P. M., Cauffman, S. R., Felch, K. L., & Weber, R. T. (2010). Solid-state dynamic nuclear polarization at 263 GHz: spectrometer design and experimental results. *Physical Chemistry Chemical Physics*, *12*(22), 5850-5860.

Schaefer, J., & Stejskal, E. (1976). Carbon-13 nuclear magnetic resonance of polymers spinning at the magic angle. *Journal of the American Chemical Society*, *98*(4), 1031-1032.

Sieklucka, B., & Pinkowicz, D. (2017). Molecular magnetic materials: concepts and applications.

Chapter 3

Integrative NMR Spectroscopy to probe Enzyme Mechanisms

3.1 Introduction

The focus of this chapter is to lay the foundation of the method - integrative structural biology of enzyme active sites, where chemistry is of paramount importance. The role of structure is crucial in shaping our comprehension of the chemistry that occurs within enzyme active sites, and a combination of solid-state NMR, diffraction, and computational chemistry techniques has demonstrated their ability to elucidate enzyme active site structures at an atomic level (B. G. Caulkins et al., 2016a; Jacob B. Holmes et al., 2022; Jinfeng Lai et al., 2011). These techniques have emphasized the importance of determining the location of hydrogen atoms and finding the optimized conformation of atoms and molecular bonding arrangements, which are challenging to locate using primary structural methods such as X-ray diffraction and cryo-EM yet are essential for understanding active site chemistry (Guo & Rubinstein, 2018; Kovalevsky et al., 2008). Precise knowledge of proton locations is critical for establishing the mechanism of acid-base catalysis, the function of active site waters, and as input for molecular dynamics simulations, molecular docking routines, and structure-based drug design (Jacob B. Holmes et al., 2022; Jinfeng Lai et al., 2011; Perilla et al., 2021).

Integrative techniques for enzyme active sites assemble components of structure across length scales that vary from the dimensions of the protein crystal unit cell down to the protonation states of the active site species (Aumiller et al., 2018). The global structural framework is typically determined using X-ray crystallography or cryo-EM (Perilla et al., 2021; Perilla et al., 2017). This broad overview of the protein structure provides essential information on the catalytic mechanism, such as identifying the protein residues that

interact with cofactors and reacting substrate(s). However, the resolution of most structures is generally insufficient to locate hydrogen atoms and determine protonation states (Michalczyk et al., 2015). NMR spectroscopy provides a window into the chemical level structure, as the interactions of chemical shift, dipolar coupling, and quadrupolar coupling are extremely sensitive probes of the chemical microenvironment and can distinguish, for example, protonation states of an ionizable group or the hybridization of a reacting substrate (Bethany G. Caulkins et al., 2014; B. G. Caulkins et al., 2016a; Ghosh et al., 2021; Eduardo Hilario et al., 2016; Jacob B. Holmes et al., 2022; Jinfeng Lai et al., 2011). Yet, the empirical correlation of NMR parameters with structure can be limiting, providing chemical structure without three-dimensional context (Mueller & Dunn, 2013). Here, computational chemistry can provide the high-resolution link that integrates the two. For example, chemically-detailed models of the active site can be built upon the coarse framework structure, and these models can be tested and ranked based on how well their first-principles predicted properties reproduce the observable NMR spectroscopic parameters (Joshua D. Hartman et al., 2015). This approach merges the global and local images of the active site to give the highest possible resolution. It takes advantage of the fact that the spectroscopic parameters are sensitive not just to the chemical state, but to the larger three-dimensional chemical microenvironment in ways that can be captured computationally. This general integrative approach is often referred to as NMR crystallography (Bryce & Taulelle, 2017; Harris et al., 2009), from here on, referred to as NMRX which was initially developed and now extended to (i) *de novo* structure determination using NMR data, (ii) structure refinement using NMR chemical shifts restraints and (iii) cross validation of structural models using NMR data.

This chapter takes as a detailed example the application of integrated structural biology to the enzyme active site of *S. typhimurium* tryptophan synthase, examining each of the requisite components in turn and how they come together to build a chemically rich picture of structure and mechanism. By chemically rich, we mean structures in which the position of every atom, including hydrogen, is specified. Examples of other enzyme complexes are then presented and highlight the diversity of this integrated approach and recognize the significant variation in the components. Yet all share the theme of merging global and local chemical structures from diffraction and NMR using computational methods. First, this chapter sets the stage for the focus on these highly optimized biomolecular nanomachines.

3.2 Enzyme Catalyzed Reactions

Enzymes are proteins that act as biological catalysts in chemical reactions. They play a crucial role in the functioning of cells and organisms, by increasing the rate of chemical reactions that occur in living systems. Enzyme catalyzed reactions are essential for life, as they allow cells to carry out metabolic processes that are necessary for growth, reproduction, and survival (Voet et al., 2016).

Enzymes work by lowering the activation energy required for a reaction to occur. They do this by providing an environment that is conducive to the reaction and by bringing the reactants together in the correct orientation and protonation states for the reaction to occur (Walsh, 1979).

Enzymes are highly specific in their action, meaning that each enzyme is designed to catalyze a specific reaction. This specificity is due to the unique shape of the enzyme's active site, which is the part of the enzyme that interacts with the substrate, or the molecule that the enzyme acts upon. The shape of the active site is determined by the

sequence of amino acids that make up the enzyme's structure, and any change in this sequence can alter the shape of the active site, which in turn can affect the enzyme's activity (Koshland Jr, 1958).

Enzyme catalyzed reactions are influenced by a variety of factors, including temperature, pH, substrate concentration, and the presence of inhibitors or activators. Each enzyme has an optimal temperature and pH range at which it functions best, and any deviation from this range can affect the enzyme's activity. Substrate concentration also plays a role in enzyme activity, as the rate of the reaction increases with increasing substrate concentration up to a certain point, after which the reaction rate levels off.

3.3. Case Study: inhibition of tryptophan synthase by benzimidazole

In NMRX, it is essential to have a structural framework as a basis for constructing, refining, and validating three-dimensional chemical models of the enzyme active site. In this chapter, we examine the evolution of crystal structures in tryptophan synthase and the crucial role that these advancements have played in elucidating the mechanisms and allostery of this enzyme. Furthermore, we highlight how these developments inspire the utilization of even more precise mechanistic probes to gain deeper insights into the functioning of this enzyme.

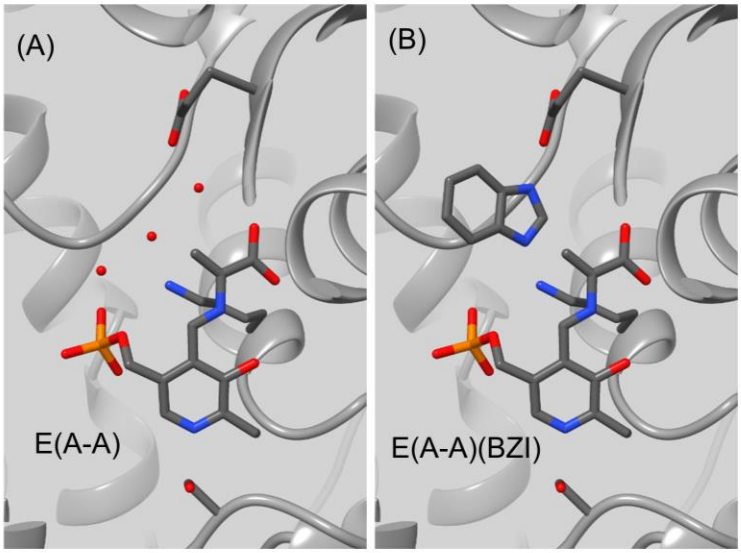
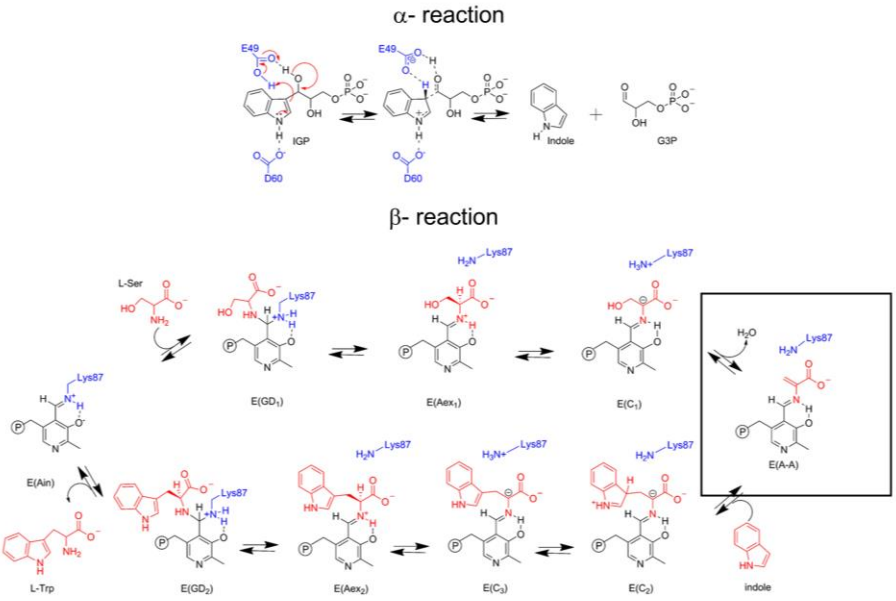


Fig 3.1 (A) Active site depicting the three water molecules adjacent to the E(A-A) intermediate. (B) Active site of the E(A-A)(BZI) complex.

We take as an example our recent work on the inhibited complex between benzimidazole and the α-aminoacrylate intermediate in tryptophan synthase β-subunit active site. We start by placing the X-ray crystal structure of the inhibited complex within the larger structural and mechanistic context of catalysis in tryptophan synthase. We discuss the strategy for obtaining ^{13}C and ^{15}N chemical shifts of key sites within the enzyme-bound intermediate under conditions of active catalysis and their implications for the chemical structure of the intermediate. Finally, we detail the strategy to determine the atomic-resolution, three-dimensional structure of the active site using computational models that are consistent with both the X-ray and NMR data and the novel structural and mechanistic hypothesis that results. For the α-aminoacrylate intermediate, NMRX highlighted the reaction coordinate for the loss of the substrate L-Ser hydroxyl and the protonation states that are critical to why the inhibitor benzimidazole (BZI) does not react while the natural substrate indole does (Fig 3.1).

3.3.1 Framework Structure: X-Ray Crystallography



Scheme 3.1 *(top)* α-Reaction with catalytic residues αGlu49 and αAsp60 shown in blue. *(bottom)* β-Reaction with the reacting substrates L-Ser and indole and product L-Trp shown in red. Catalytic residue βLys87 is shown in blue. PLP species are shown in black and reacting substrate species are shown in red. The focus of this chapter is going to be the E(A-A) intermediate where it has lost the water molecule and the indole is yet to make the nucleophilic attack on the Cβ of the L-Ser moiety (Redrawn from (Ghosh et al., 2022)).

The biosynthesis of L-tryptophan is completed by the Tryptophan synthase, which catalyzes the last two steps of the process (Scheme 3.1). It has been established through early research that the substrates for the bienzyme complex are 3-indole-D-glycerol-3′-phosphate (IGP) and L-serine (L-Ser), with the subunit composition being $\alpha_2\beta_2$. It has also been found that the β-subunits require pyridoxal 5′-phosphate (PLP) for catalytic activity, and that indole is a channeled intermediate (Ghosh et al., 2022; Hyde et al., 1988). The first X-ray structures provided significant details regarding the catalytic mechanism of the α- and β-subunits (Hyde et al., 1988). Prior to the determination of

these structures, there was no consensus on the mechanism for the channeling of the intermediate, indole, between the α- and β-sites, and many protein residues were considered potential candidates for involvement in catalysis. Additionally, the protein folds and the αββα linear arrangement of the subunits were not known, and there was no clear explanation for the dependence of the activities of the subunits on the state of aggregation (Miles, 1995; Miles et al., 1989). Subsequent X-ray structures determined over the next 20 years have shed light on many of these details (Michael F Dunn, 2012; Michael F Dunn et al., 2008).

The determination of the structure of the wild-type enzyme with closed β-subunits in 2007 (Barends, Domratcheva, et al., 2008) revealed the key residues involved in catalysis in the β-site and how the transition between the open and closed states of the subunits coordinates the reactions of the α- and β-sites during the overall catalytic cycle. The closed subunit structures also provided insight into the α- and β-site transformations in Scheme 3.1. The closed conformation of the α-subunit brings αGlu49 and αAsp60 into the correct bonding positions to catalyze the cleavage of IGP to indole. It is postulated that αAsp60 provides coulombic charge-charge interactions that stabilize partial positive charge on the indolyl ring as bond scission at the ring C-3 position occurs, and that αGlu49 plays an acid-base catalytic role in facilitating the C-C bond scission step. In the β-subunit, the switch from the open to the closed conformation rearranges the interactions within the β-site, facilitating the conversion of $E(Aex_1)$ to $E(C_1)$ and then $E(C_1)$ to E(A-A) and release of a water molecule in Stage I of the β-reaction in Scheme 3.1. The rearrangement of βAsp305 side chain carboxylate allows formation of a salt bridge between βArg141 and βAsp305 which locks the β-subunit into the closed conformation and blocks access to the β-site from solution. Within this closed

conformation, the active site acid-base catalytic groups, βLys87 and βGlu109, are correctly positioned to facilitate the chemical transformations of the reacting substrate (Fig 3.1).

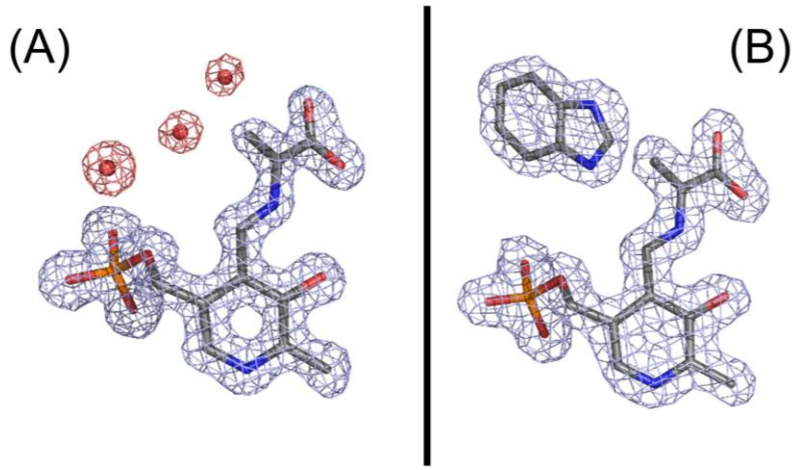


Fig 3.2 Electron density maps of the (A) E(A-A), (B) E(A-A)(BZI) complexes.

The X-ray crystal structures of E(A-A) and E(A-A)(BZI) serves as a valuable reference point for gaining mechanistic insights into catalysis in tryptophan synthase (TS). In the crystal structures of E(A-A), there are three crystallographic water molecules situated adjacent to the serine substrate in the β-subunit active site, forming a hydrogen-bonded chain that extends from the carboxylate of the catalytically essential βGlu109. The position of the central water molecule is especially noteworthy as it is in close proximity to both the substrate Cβ (dCO = 3.2 Å) and the βLys87 ε-amino group (dNO = 3.0 Å), suggesting that it is near the site where the substrate β-hydroxyl is expected to be located before elimination (Jacob B. Holmes et al., 2022; Dimitri Niks et al., 2013). Binding of BZI (Fig 3.2) displaces these water molecules and causes slight perturbations of the α-aminoacryloyl group. Furthermore, binding induces minor movements of the βGlu109 carboxylate and the βLys87 ε-amino groups so that they are within hydrogen

bonding distances of the BZI nitrogen atoms (Fig 3.1). These interactions, combined with the structural similarity of BZI to indole, make the BZI complex an excellent model for the anticipated positioning of indole for nucleophilic attack at Cβ. Notably absent from Fig 3.1 and Fig 3.2, however, are the hydrogen atoms positions and orientations on electronegative atoms, which play a crucial role in identifying hydrogen bond donors and acceptors and their mechanistic functions.

Obtaining chemically detailed structural models of intermediates in the enzyme active site is a challenging task. It requires combining various experimental techniques, including X-ray crystallography, spectroscopy, and kinetic measurements, with computational simulations such as quantum mechanics/molecular mechanics (QM/MM) calculations and molecular dynamics (MD) simulations (Facelli & Grant, 1993). By combining these techniques, it is possible to determine the protonation states and the dynamics of the reacting substrate, PLP coenzyme, and catalytic residues, as well as the structures and energetics of the intermediates formed during the catalytic cycle.

Recent advances in X-ray crystallography, Neutron crystallography and cryo-electron microscopy (cryo-EM) have allowed for the determination of high-resolution structures of large enzyme complexes, including those involved in tryptophan biosynthesis. Moreover, computational simulations such as QM/MM and MD have become more powerful, allowing for the study of large systems with high accuracy. Despite these successes, significant questions remain regarding the chemical mechanism for the substrates' transformation remain as the resolution of the crystal and the Cryo-EM structures does not allow for protonation states to be established on the reacting substrate. In E(A-A), this becomes more challenging as it is quickly converted to pyruvate and ammonium ion (Scheme 3.1), making it difficult to study the chemistry of the enzyme under conditions of

active catalysis. Although room temperature serial X-ray crystallography has ameliorated this challenge to some extent (Heymann et al., 2014), it is far from routine. Critical dynamic and chemical information on the PLP cofactor atoms and the residues' sidechain of βLys87 and βGlu109 remain elusive (Bethany G. Caulkins et al., 2014; B. G. Caulkins et al., 2016a; Jinfeng Lai et al., 2011; Mueller & Dunn, 2013). The catalytic activity of the β subunit in TS is also dependent on the protonation states of the ionizable groups on the PLP moiety (the PLP phenolic hydroxyl, pyridine ring nitrogen, and phosphoryl group) and on the reacting substrate (the Schiff base nitrogen linked to the PLP C4′ carbon, and the carboxylate of the reacting substrate (Fig 3.1). Therefore, to fully understand the mechanism, chemically detailed structural models of the intermediates in the enzyme active site is a pre-requisite.

Some of the other long-standing challenges in the TS mechanism remain understanding the reaction mechanism of indole with E(A-A) forming the new C-C bond at the β carbon of the L-Ser moiety of the E(A-A) intermediate. Studies from the Dunn group here at University of California, Riverside, using indole isostere BZI laid the foundation of the discussion of mechanistic enzymology in chapter. They reported that BZI, although a great nucleophile on its own, failed to give carbanion complexes with E(A-A) and that BZI was a strong inhibitor of carbanion formation (Peter S Brzovic et al., 1992; Rodney M Harris et al., 2005). It was proposed that owing to the underlying differences with which indole and BZI react with nucleophiles, where indole reacts with C3 carbon and the BZI with N3 atom, both these small molecules should bind with similar geometries in the indole subsite of the β subunit. Now, if this were the case, the BZI complex could be incorrectly bound for the forward reaction owing to the fundamental difference in nucleophilic reaction mechanism of indole and BZI (Melinda Roy et al., 1988).

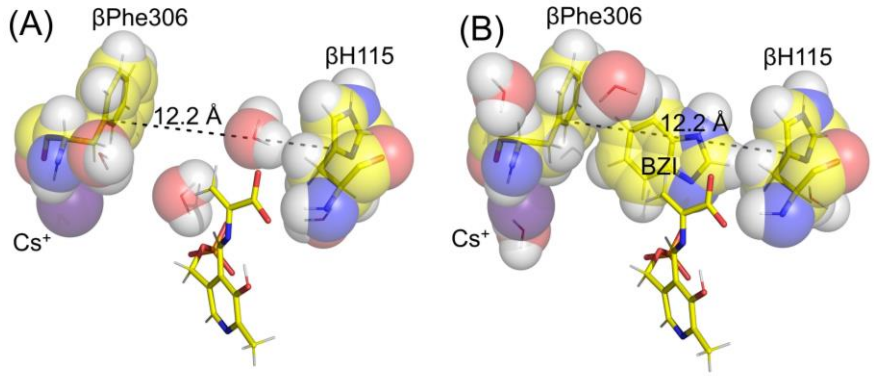


Fig 3.3. VDW contact structural detail of the β-subunit indole sub-site in (A) E(A-A) (PDB ID: 4HN4) with hydrogens (white) modeled onto the heavy atoms. (B) E(A-A)(BZI) complex (PDB ID: 4HPX) showing structure detail of the indole subsite occupied by the indole isostere BZI including the VDW contacts of BZI with residues βPhe306 and βH115. The subsite distance of 12.2 Å is just right to bind BZI. The monovalent cation sites in these structures are occupied by Cs^+. Coloring scheme: carbons, yellow and in VDW spheres overlapping the stick representations, while the Cs^+ ions are colored purple. Redrawn from (Ghosh et al., 2022).

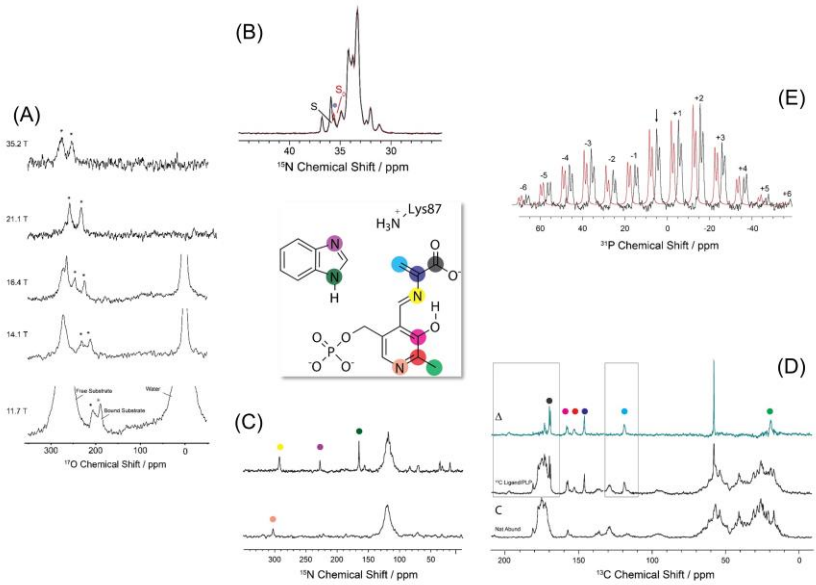
Close inspection of the X-ray crystal structures of E(A-A) reacted with BZI, 2AP and indoline (see chapter 4 and chapter 5) show that all of them occupy the same indole subsite. Small differences in the orientation of the aromatic rings could reflect the chemical interactions these molecules make with subsite residues (Fig 3.3). From X-ray crystal structures, it was not clear whether there were other factors such as protonation states of the βLys87 or electrostatics interactions around the PLP moiety could be a major factor in understanding why BZI doesn't react while indole does.

This is where the challenge to NMRX comes in the picture, where we are trying to achieve atomic resolution and structural characterization to answer some of the most fundamental mechanistic questions regarding enzyme mechanism. In this chapter, we demonstrate how this can be accomplished using the example of our successful

determination of the NMRX assisted structure of the E(A-A)(BZI) complex resulting from the reaction of E(A-A) and BZI. While the kinetic behavior of the E(A-A) has been well characterized and numerous crystal structures have been solved. Our approach builds upon these results to provide a more comprehensive understanding of the intermediate's chemical structure.

3.3.2 Chemical Structure: NMR Spectroscopy

Our research has mainly focused on utilizing isotropic chemical shifts to enable NMRX in tryptophan synthase. We acknowledge that anisotropic NMR interactions can provide additional restraints for refining models, if available. NMRX relies on isotropic chemical shifts in two specific ways. Firstly, the chemical shift is used to directly report on the chemical state of a probe atom, providing information on its protonation or hybridization state (Bethany G. Caulkins et al., 2014; B. G. Caulkins et al., 2016a; Jacob B. Holmes et al., 2022; Jinfeng Lai et al., 2011; Mueller & Dunn, 2013). For example, McDowell et al. used solid-state ^{13}C NMR spectroscopy to determine the change in hybridization at the beta carbon of enriched L-[3-^{13}C]Ser when forming the E(A-A) intermediate in tryptophan synthase. Secondly, chemical shifts are used to report on the chemical and structural environment surrounding a probe nucleus. In the structural refinement of organic and inorganic molecular crystals using NMRX, the chemical state of the probe atom is known, but its shift also depends on its three-dimensional molecular conformation and the precise location and chemical state of nearby atoms in the surrounding crystal lattice. Interpreting the chemical shift of a probe nucleus in terms of the immediate chemical and structural environment can be challenging, but atomic-resolution structural models can be uniquely determined when chemical shifts for multiple probe nuclei are measured and analyzed collectively within the context of NMRX, as illustrated below.



Chemical Shifts / ppm

	E(A-A)
N1	302.4
C2	153.1
C2'	18.2
C3	158
N(SB)	292.3
C^{α}	146
C'	169.8
C^{β}	118.7
O1	258
O2	286
N^{ζ}	35.6
N1	165.5
N3	227.8

Fig 3.4 (A) ^{17}O NMR Spectroscopy of E(A-A)(BZI). Peaks marked with (*) shows the resonances of the oxygen atoms of the carboxylate group; (B) ^{15}N-observe, ^{31}P-dephased Rotational Echo Double Resonance (REDOR) experiments of the E(A-A)(BZI) complex prepared with U-^{15}N-Lys TS. Spectra acquired at 21.1 T, -10 °C, and 8 kHz MAS. Both the black and red spectra were acquired with a 25 ms echo period on ^{15}N before detection, but the black spectrum includes the application of dipolar dephasing pulses on ^{31}P that attenuate signals from 15N atoms that are dipolar coupled to ^{31}P atoms. The dipolar coupling falls off as the inverse cube of the interatomic distance, and the $^{15}N(^{31}P)$-REDOR editing used here is selective for nitrogen atoms within ~3-4 Å of the PLP ^{31}P atoms. There is resolution of many individual ε-amino group nitrogen sites, but only the signal at 35.6 ppm dephases. Based on proximity to the PLP phosphate group, the ε-amino group of βLys87 is assigned to this resonance. (C) ^{15}N ssNMR CPMAS spectra of microcrystalline TS E(A-A)(BZI) prepared with the following isotopic labeling: (*bottom*) ^{15}N-labeled on the substrate L-Ser; selectively (*top*) ^{15}N enriched on the PLP cofactor, L-Ser and BZI; (D) ^{13}C ssNMR CPMAS spectra of microcrystalline TS E(A-A)(BZI) U-$^{13}C_3$-labeled L-Ser substrate; and selectively ^{13}C,^{15}N-enriched on the PLP cofactor and Cβ of the substrate L-Ser. The top spectra in (D) are formed as the difference between the E(A-A)(BZI) spectra with various cofactor/ligand isotopic labels and the same spectra acquired at natural abundance, highlighting the resonances for the specific site labels. The large peak at 63.1 ppm is free serine. Spectra acquired at 9.4 T, -10 °C, and 8 kHz MAS;
(E) The E(A-A)(BZI) phosphorus chemical shift tensors. Slow spinning ^{31}P SSNMR CPMAS spectrum of the TS E(A-A)(BZI) complex formed upon reaction TS microcrystals with L-Ser+BZI. The PLP phosphate isotropic peak at 4.9 ppm is indicated by the arrow. The peak at 3.7 ppm was previously assigned to the α-site ligand F9 (B. G. Caulkins et al., 2016a). The fit (red) to the sideband manifold in BrukerTopspin 3.6 allows for the extraction of the CSA principal axis components (δ_{11}, δ_{22}, δ_{33}) = (62.4±0.4, -8.2±0.4, -39.6±0.4) ppm for the PLP phosphate group. Both the isotropic and anisotropic chemical shifts indicate that the phosphate group is dianionic. The order of the spinning sidebands is given above each peak. (Redrawn from (Jacob B. Holmes et al., 2022)).

Solid-state NMR spectra can be obtained on microcrystalline protein samples under conditions similar to those used for X-ray crystallography, making it the preferred method for NMRX. The sensitivity of solid-state NMR is not necessarily lower than that of solution-state NMR, and in larger systems such as tryptophan synthase, ^{13}C and ^{15}N sensitivities are often better under CP MAS conditions in solid-state NMR. Biological relevance is also not compromised by acquisition in the solid-state, as data on microcrystalline samples of enzymes such as tryptophan synthase are collected under conditions of active catalysis.

Although there may be unique challenges in acquiring and assigning solid-state NMR spectra for each system, obtaining ^{13}C and ^{15}N chemical shifts for the E(A-A)(BZI) complex in tryptophan synthase can be achieved by supplying isotopically-enriched L-serine and BZI substrates to microcrystals of the enzyme complex **(**Fig 3.4). Despite the crystal lattice containing over 50% water by mass, the substrates can diffuse freely within lattice channels to establish steady-state concentrations of the intermediates during enzyme catalysis. By following the reaction progress with NMR, it is possible to detect the depletion of the enriched substrate in the mother liquor and the formation of the product, pyruvate (^{13}C) and ammonium ion (^{15}N), using direct $^{13}C/^{15}N$ excitation with low-power ^{1}H decoupling. CP MAS with high-power ^{1}H decoupling can be used to select bound substrates, allowing for the assignment of the ^{13}C and ^{15}N resonances for the intermediate fragment derived from serine, BZI and βLys87 (Table 3.1). The Cα and Cβ chemical shifts derived from serine and the nitrogen chemical shifts derived from L-serine, BZI and βLys87 can be resolved in 1D CP-MAS experiments. The solid-state NMR spectra acquired under these conditions can provide valuable structural information for the determination of atomic-resolution models using NMRX.

One of the main objectives of NMRX is to determine the protonation states of ionizable groups within the active site. Fig 3.4B illustrates potential sites of protonation located on or near the intermediate complex, such as the phosphoryl group, pyridoxal phenolic oxygen, Schiff base nitrogen, both carboxylate oxygens derived from L-Ser, and the neighboring Nε group of βLys87. The Schiff base nitrogen is the only ionizable group that has a directly measured chemical shift. Previous measurements of PLP-Schiff base complexes in model compounds indicate that protonated Schiff bases typically fall near 190 ppm, while neutral Schiff base linkages fall above 315 ppm (referenced to liq-NH_3) (Chan-Huot et al., 2010; Limbach et al., 2011). The experimental chemical shift of this intermediate's Schiff base nitrogen is at 286.5 ppm, which falls between the range expected for either fully protonated or deprotonated Schiff bases, suggesting that the observed shift is due to a fast exchange equilibrium between the two forms, similar to what has been observed for carbanion complexes with 2AP and indole. The partner structures and the protonation states of the other ionizable groups cannot be easily determined from the empirical correlation of shift with structure. More importantly, there is one unique difference in the chemical shift measurements of the E(A-A) and E(A-A)(BZI) complex, in the ^{15}N chemical shift of the Nε group of βLys87 (Table 3.1)**.** In the following section, we demonstrate how combining *ab initio* computational chemistry and X-ray crystallography with the chemical shifts can identify with high certainty the equilibrium between two specific tautomeric forms and give chemically rich information about the active site chemistry.

Table 3.1 E(A-A) and E(A-A)(BZI) experimental and first-principles chemical shifts (ppm) for the phenolic (Phen) and protonated Schiff base (PSB) species and their two-site exchange with the following populations: E(A-A) 89.3% Phen,10.7% PSB; E(A-A)(BZI) 89.4% Phen, 10.6% PSB

E(A-A)	Atom	Phen	PSB	Two-Site	Expt	E(A-A) (BZI)	Atom	Phen	PSB	Two-Site	Expt
PLP	N1	302.9	303.3	303.0	297.6	PLP	N1	310.2	307.3	309.9	302.4
	C2	148.7	159.1	149.8	151.2		C2	152.1	162.7	153.2	153.1
	C2'	18.8	19.8	18.9	17.5		C2'	19.4	20.7	19.6	18.2
	C3	154.9	173.2	156.9	158.1		C3	155.2	174.0	157.2	158
L-Ser	SB N	301.4	153.6	285.5	286.7	L-Ser	SB N	307.6	152.0	291.2	292.3
	Cα	144.4	133.1	143.2	145.6		Cα	147.2	136.3	146.0	146
	C'	171.2	168.7	171.0	170.9		C'	168.9	166.5	168.7	169.8
	Cβ	123.3	121.3	123.0	118.8		Cβ	120.6	118.2	120.3	118.7
	O1	262.1	261.4	262.0	257		O1	261.4	260.5	261.3	258
	O2	302.4	294.3	301.5	289		O2	301.4	291.8	300.4	286
βK87	Nζ	20.5	21.1	20.6	24.2	βK87	Nζ	35.6	36.4	35.6	35.6
						BZI	N1	167.5	167.9	167.5	165.5
							N3	231.4	231.7	231.4	227.8
	red-χ^2	3.21	106.1	1.85			red-χ^2	2.19	97.8	0.90	

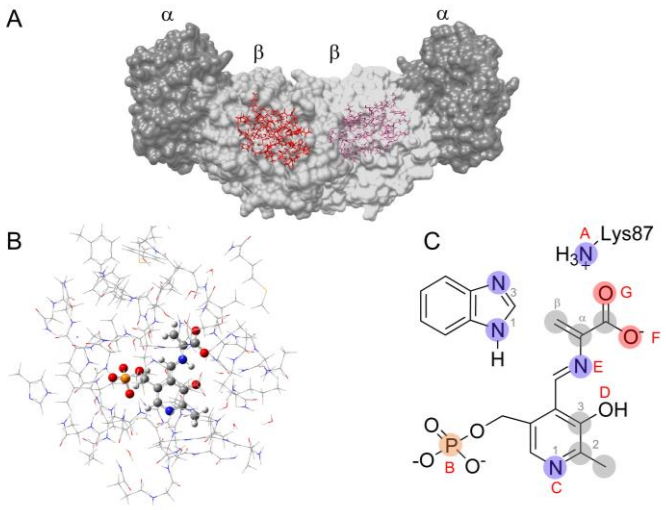


Fig 3.5. Cluster model of the E(A-A) active site. (A) X-ray crystal structure of the TS $\alpha_2\beta_2$ heterodimer with the β-subunit active site in red. (B) Cluster model of the active site for first-principles geometry optimization and chemical shift calculations; protein side chains displayed in wireframe and cofactor and substrate in ball-and-stick. (C) Protonation sites on and near the cofactor/substrate complex: A: the βLys87 side chain; B: the PLP phosphate group; C: the PLP pyridine ring nitrogen; D: the PLP phenolic oxygen; E: the Schiff-base nitrogen; and F/G: the substrate carboxylate. Shaded nuclei indicate sites for which experimental NMR chemical shifts are reported. Redrawn from (Jacob B. Holmes et al., 2022).

3.3.3 Chemically-Rich Structure: Adding First-Principles Computational Chemistry

X-ray diffraction methods are unmatched in their ability to determine framework structures for systems that have long-range crystalline order, while NMR spectroscopy is unmatched in its ability to determine local chemical structure. Together, the two can provide chemically detailed, three-dimensional structures. The sophistication with which NMR and diffraction methods have been combined took a remarkable leap forward when they were paired with *ab initio* computational chemistry. In one of the first examples of this, Facelli and Grant (Facelli & Grant, 1993) demonstrated that structural models based upon the X-ray crystal structure framework could be distinguished by the comparison of experimental chemical shift tensor components with those predicted for competing structural models developed using *ab initio* computational chemistry. This work laid the foundation for the development of NMRX in organic and inorganic molecular crystals, and it was suggested this approach might also be attractive for the study of active sites in biomolecules Indeed, the atomic resolution of NMRX allows changes in protonation and hybridization states during enzymatic transformations to be directly observed, and in this chapter we detail some of my thesis work on the development and application of NMRX to a functioning enzyme active site.

Model building and testing are often key components of the NMRX method, placing a premium on the integration of computational approaches such as molecular dynamics (MD) and first-principles computational chemistry. These methods use quantum mechanics to calculate the electronic structure of a molecule or solid-state material, and then use this information to calculate NMR properties such as chemical shifts, spin-spin coupling constants, and relaxation rates. The accuracy of these calculations can be

benchmarked against experimental NMR data and improved using more sophisticated theoretical models and larger basis sets.

The combination of *ab initio* calculations with experimental NMR data allows for the determination of the protonation and tautomeric states of ionizable groups in enzyme active sites, as well as the identification of specific hydrogen bonding networks and other interactions that stabilize these states. In the case of the tryptophan synthase E(A-A) intermediate, for example, the comparison of experimental chemical shifts with those calculated for different tautomers of the Schiff base intermediate allowed the identification of an equilibrium between two specific tautomeric forms and major and minor species of the intermediate relevant to catalysis. Overall, the use of computational chemistry in NMRX has greatly expanded the scope and accuracy of this technique, allowing for the determination of chemically-detailed, three-dimensional structures of complex biomolecular systems.

In NMRX, the X-ray crystal structure provides the scaffolding upon which chemically detailed structures are constructed. The approach of fixing the backbone and side chain atoms of the enzyme while optimizing the non-peptide components allows for more accurate models of enzyme active sites to be constructed. By using computational chemistry to refine the non-peptide components, the chemical details of the active site can be determined with high precision (B. G. Caulkins et al., 2016a; Joshua D. Hartman et al., 2015). However, it is important to keep in mind that the accuracy of the models is limited by the quality of the X-ray crystal structure and the accuracy of the computational methods used for geometry optimization and NMR parameter calculations. In other words, while the computational optimization of non-peptide components such as coenzymes and substrates may introduce some degree of uncertainty in their structural

coordinates, the use of weaker constraints can help to maintain overall structural features and compensate for potential limitations in the computational modeling. Additionally, the expected changes in bond lengths (<0.2 Å) and angles (~5^{o}) during geometry optimization are relatively minor, indicating that the overall structural integrity of the enzyme active site should remain largely intact.

Molecular mechanics methods use simplified descriptions of molecules, such as classical force fields, to calculate the potential energy of a system based on the interactions between atoms and molecules. The accuracy of these force fields depends on the parameters used to describe the chemical bonds, angles, and torsions of the atoms in the system. Parameterization of the force fields is typically done using experimental data and quantum-mechanical calculations, and the accuracy of the force fields can be improved by incorporating more sophisticated quantum-mechanical calculations. However, these methods are still limited by the accuracy of the force fields and the approximations made in the quantum-mechanical calculations, which can lead to inaccuracies in the calculated energies and structures.

In our application of NMRX to the E(A-A) intermediate in TS, we first built a truncated model of the active site consisting of residues located within 7 Å of the intermediate (PLP + reacting substrate) (Fig 3.5). With the truncated model of the active site, the intermediate was optimized using a hybrid layer-rewise method, which treats the PLP-ligand complex at a higher level of theory (density-functional level of theory, B3LYP/6-31G(d,p)) and the amino acid residues in the catalytic pocket at a lower level of theory (B3LYP/6-31G). This allowed for accurate modeling of the interactions between the intermediate and the surrounding amino acid residues, while keeping the computational cost manageable. The optimized intermediate structure was then used to calculate NMR

chemical shifts using the GIAO method in the gauge-invariant atomic orbital (GIAO) approximation. These calculated chemical shifts were compared with the experimental chemical shifts obtained from solid-state NMR spectra of the tryptophan synthase crystals, allowing for validation of the proposed intermediate structure and resolve previously unanswered mechanistic questions. The success of this approach demonstrates the power of combining experimental NMR data, X-ray crystallography, and computational chemistry to obtain detailed, chemically-specific structural information of enzyme active sites.

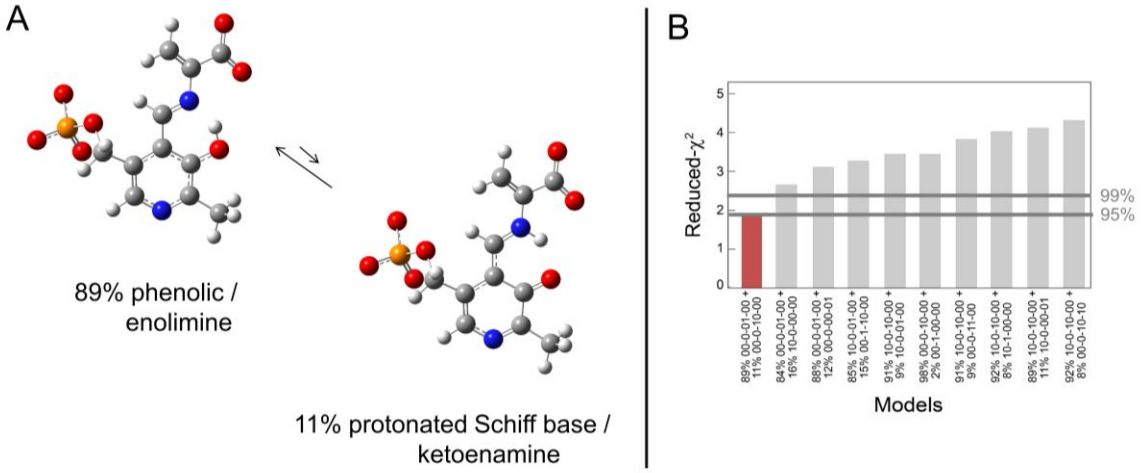


Fig 3.6 (A) Fast Exchange Equilibrium Model in E(A-A). (B) Ranking of the best tautomeric exchange models based on the χ^2_{red}, As expected the major species is the phenolic form of the intermediate, and it's exchange partner is PSB form in which the proton is transferred from the phenolic oxygen to the Schiff base nitrogen. (Redrawn from (Jacob B. Holmes et al., 2022)).

In the next steps, candidate structures for the intermediate in the active site were systematically generated with varying protonation states for the six ionizable sites on or near the PLP-ligand complex and ranked based on reduced χ2 statistic**.** The reduced χ2 statistic is a measure of how well the calculated chemical shifts match the experimental chemical shifts. The closer the value of reduced χ2 is to 1, the better the agreement

between the calculated and experimental shifts. Weightings corresponding to root-mean-square deviations of 1.4 ppm for ^{13}C and 4.2 ppm for ^{15}N and 7.2 ppm for ^{17}O are used to account for the expected experimental uncertainties (Joshua D. Hartman et al., 2015). The candidate structures for the intermediate in the active site were generated with varying protonation states for the six ionizable sites on or near the PLP-ligand complex, and models with more than a single proton placed at certain sites were not considered. The chemical shieldings of each refined model were calculated (using B3LYP/ 6-311++G (d,p) theory level) and converted to chemical shifts referenced to TMS for ^{13}C and to liquid ammonia for ^{15}N. The agreement between the calculated chemical shifts for the model structures and the experimental shifts was then ranked using the reduced $\chi 2$ statistic.

3.4 Mechanistic Implications

For the E(A-A) and the E(A-A)(BZI) analogue complex, we found that a fast-exchange equilibrium model was necessary to account for the large discrepancy in chemical shift values at the Schiff base nitrogen of both the E(A-A) and E(A-A)(BZI) (Fig 3.6). In this model, exchange partners that differed by the position of a single proton were systematically paired and the relative populations of the tautomers were optimized for best agreement with the experimental chemical shifts. This approach allowed us to account for the dynamic behavior of the isostere complex and find a better fit to the experimental data. The winning structural model in both the intermediates were found to be a tautomeric exchange model between the SB N and PO with a reduced $\chi 2$ of 1.4 and 0.7 for E(A-A) and E(A-A)(BZI) complex. But a major difference in characterizing the winning structure was observed in terms of the protonation state of the Nε group of βLys87, where in E(A-A), the Nζ is deprotonated (24.7 ppm) while in E(A-A)(BZI), the Nζ

is protonated (35.1 ppm) determined via ^{15}N-^{31}P dephasing REDOR experiments taking advantage of the unique distance of Nε group of βLys87 to the PLP phosphate group (~3.6 Å). The calculated chemical shifts for the PSB and PO structures, in addition to their fast-exchange average, are summarized in Table 3.1.

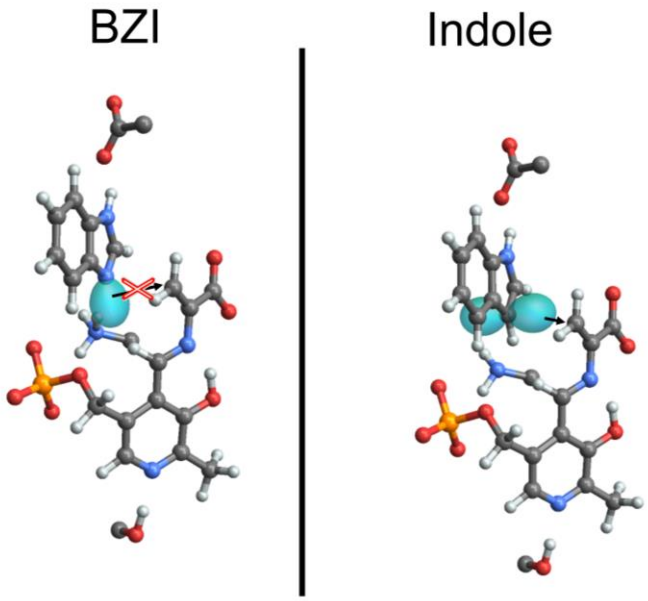


Fig 3.7 BZI reacts as a nucleophile via a nitrogen lone pair at N3 but is held in the wrong orientation by the additional hydrogen bond to the charged ε-amino group of βLys87, Indole reacts via the pi system at C3 and when modeled into the BZI binding pocket is properly aligned to form the new C-C bond. Figures drawn using TensorView (Young et al., 2019).

Just like the E(A-A), the E(A-A)(BZI) complex shows a similarly detailed view of the active site chemistry for the subsequent mechanistic step, nucleophilic addition at the substrate C^{β}. The view from NMR-assisted crystallography shows BZI bound in the active site, displacing the three waters, and poised for nucleophilic attack. Despite being more nucleophilic than indole, however, BZI does not react to form a covalent bond. Thus, this intermediate analog appears to be poised close to the transition state between two reactive intermediates in the TS catalytic cycle. That BZI is not a substrate for TS

and does not covalently react with E(A-A) can be attributed to two important factors: the different nucleophilic reaction mechanisms of indole and BZI, and the tight packing and hydrogen bonding interactions within the indole subsite. The mechanisms of nucleophilic reaction for indole and BZI are fundamentally distinct, as BZI reacts via a nitrogen lone pair, and indole reacts via the pi system at C3. NMR-assisted crystallography identifies that it is the nucleophilic N3 that is adjacent to the substrate C^{β} in the E(A-A)(BZI) complex. Importantly, this structure shows that the BZI lone pair does not point toward the E(A-A) C^{β} p-orbital. The tight packing of atoms within the indole subsite, the hydrogen bond between N3 and the charged ε-amino group of βLys87, and the hydrogen bond between the βGlu109 carboxylate and N1 of BZI preclude rearrangement of BZI within the site to allow the reaction.

In contrast, when indole was modeled in place of BZI, the C3 carbon of the 5-membered ring is perfectly aligned to form the new C-C bond. The attacking p-orbital on indole points toward the electron deficient C^{β} carbon and is poised to make orbital overlap as the complex moves along the reaction coordinate to the sp^3 geometry of the product. At the same time, N1 of indole is positioned to hydrogen bond to the carboxylate of βGlu109, which stabilizes positive charge development on N1 as the transition state is approached. The C3 carbon of indole is relatively electron rich, but still requires assistance in the attack on C^{β}. The critical role played by βGlu109 can be seen in the βGlu109Asp mutation, which reduces the β reaction rate 27-fold We posit that the E(A-A)(BZI) complex models how indole is bound to E(A-A) just prior to C-C bond formation.

The application of SSNMR spectroscopy, X-ray crystallography, and first principles computational chemistry in combination provides a detailed understanding of the TS E(A-A) and E(A-A)(BZI) active sites. By determining heavy atom positions, protonation

states, and the placement and orientation of active site waters, a comprehensive picture of structure and reactivity emerges. The E(A-A) intermediate, for example, is characterized by a central water molecule that is ideally positioned for the reverse nucleophilic attack on Cβ, with an orientation that points back to the active site acid-base catalytic residue, βLys87. This highlights the reaction coordinate for the elimination of the substrate β-hydroxyl, and both X-ray crystallography and SSNMR indicate variable occupancy for this site.

In the next step, when E(A-A) reacts with BZI, the previously observed waters are displaced, and NMR-assisted crystallography reveals BZI occupying the indole binding pocket, but unable to react. Protonation states complete the chemical picture for why BZI is unable to initiate the next step in the reaction. Despite being a good nucleophile, BZI is held in the wrong orientation by hydrogen bonds to βGlu109 and the charged ε-amino group of βLys87. Therefore, the combined application of SSNMR spectroscopy, X-ray crystallography, and first principles computational chemistry provides unparalleled insights into the complex chemistry of the TS E(A-A) and E(A-A)(BZI) active sites.

3.5 A Survey of Integrative Structural Biology in Other Enzyme Systems

3.5.1 ^{51}V NMR Crystallography of Vanadium Chloroperoxidase and Its Directed Evolution P395D/L241V/T343A Mutant: Protonation Environments of the Active Site.

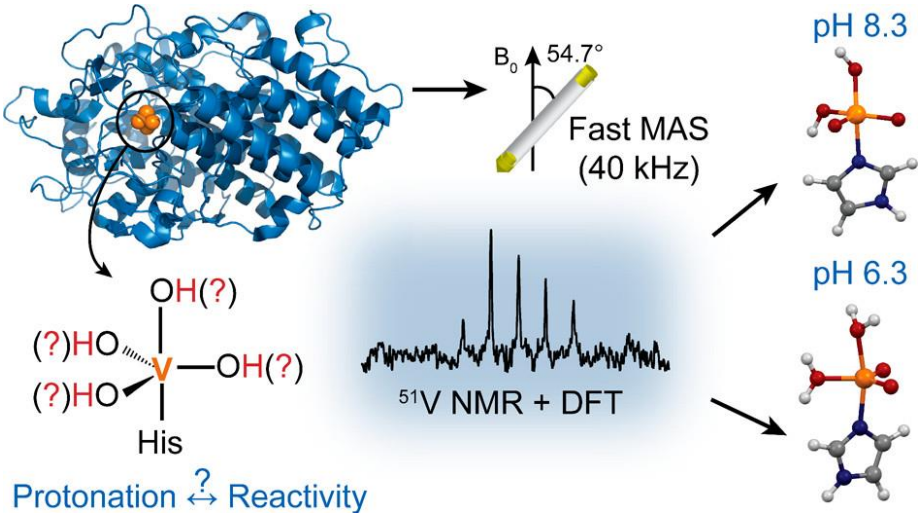


Fig 3.8 Integrative NMRX approach to understanding chemical transformation in Vanadium dependent Haloperoxidases. (Figure redrawn from (Gupta et al., 2015)).

NMRX studies from the Polenova group on Vanadium dependent Haloperoxidases (VHPOs) showed critical protonation states of the Vanadate cofactor in the wild-type and directed evolution P395D/L241V/T343A mutants, shedding light on two important mechanistic questions related to VHPO mechanism. First, what determines the substrate specificity, where a particular VHPO enzyme will be able to oxidize a chloride ion? Second, what is the pH dependence of the catalytic activity? as the reaction rates are significantly lowered under basic conditions both in the wild-type and in the P395D/L241V/T343A mutants of VCPO. The synergistic approach of NMRX led to the assignment of the protonation states on the vanadate cofactor for the first time. Their

studies concluded that the protonation states change with the pH of the solution. The NMR data combined with DFT calculations showed that the triple mutant has a doubly protonated cofactor with an axial water molecule. This provided insights into the structural differences in the enzyme active site of the wild type and the mutant. The experimental work established by Gupta et al., enabled the characterization of inactive and active peroxo states of the vanadate cofactor in vanadium haloperoxidases. and has broader applications in the investigations of geometric and electronic structure of diamagnetic metal ion sites in different enzyme and protein microenvironments (Gupta et al., 2015; Pooransingh-Margolis et al., 2006; Waller et al., 2007).

3.5.2 Fast Microsecond Dynamics of the Protein−Water Network in the Active Site of Human Carbonic Anhydrase II Studied by Solid-State NMR Spectroscopy

The work done by Singh et al., have combined ssNMR and solution NMR studies with First Principles DFT calculations and MD simulations to probe chemical and active site dynamics within the conserved protein-water network of human carbonic anhydrase II.

This integrative approach coherently points towards a fast µs dynamics in the water network in the absence of small molecule inhibitor. Strong chemical shift / dipolar coupling fluctuations in amide nitrogen relaxation dispersion experiments shed light on this µs scale dynamics, an information that was not directly available from the analysis of cryogenic B factors of X-ray crystal structures. In addition, DFT calculations shed light on H-bonding mediated changes on the chemical shifts. Taken together, this approach provided evidence of temporal divergence from the H-bonding donor and acceptor architecture (Singh et al., 2019).

3.5.3 Unexplored Nucleotide Binding Modes for the ABC Exporter MsbA

Integrative approaches have united the strengths of solid-state NMR, X-ray and neutron crystallography and Cryo-EM in answering some of the complex structural biology problems in large enzyme complexes.

One such demonstration includes the work of Kaur et al. (Kaur et al., 2018). who have used a concerted structural biology approach to investigate the nucleotide-protein complex in MsbA, combining high field and DNP-enhanced MAS NMR, EPR and MD simulations studies. Integrating these techniques, they have successfully shown that the highly conserved Q-loop in ATP binding cassette (ABC) transporter plays a critical role in mediating reverse adenylate kinase reaction. The integrated approach successfully illustrates that the Q loop is an important component for nucleotide coordination at the adenylate kinase site. In addition, DNP enhanced NMR spectroscopy enabled identification of nucleotide-protein contacts. This study highlights the capability of integrative structural biology work to explore stable as well as intermediate conformation states which are not captured via other techniques alone, such as Cryo-EM and X-ray crystallography. This study shows the importance of combination of DNP-enhanced NMR, EPR and Molecular Dynamics in obtaining insights into catalytic mechanism of membrane proteins directly within lipid bilayer. The illustration of novel nucleotide binding modes from this work extends the paradigm of working mechanism of ABC transporters which can be further applied to the whole ABC superfamily of proteins (Kaur et al., 2018).

3.5.4 Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex

The work by Gauto et al. using ssNMR and Cryo-EM have revolutionized the field on integrative structural biology where in a first, they have resolved the structure of a megadalton protein complex, dodecameric aminopeptidase TET2 to a precision below 1 Å using NMR structural restraints and a 4.1 Å Cryo-EM electron density map. In a stepwise manner, after assignments of the α helices and β sheets, they used CYANA structure calculations and ranked the helix assignments after which helix-to-density assignments and goodness of fit provided another round of structural refinement. This methodology improved the final resolution of the structure drastically and could be a routine protocol for high resolution structure determination via integrative structural biology methods (Gauto et al., 2019).

3.5.5 Characterization of PEGylated Asparaginase: New Opportunities from Integrative NMR Analysis of Large PEGylated Therapeutics

To characterize large PEGylated proteins and assemblies structurally, they used an integrated approach utilizing solid-state and solution NMR and X-ray crystallography. The strategy involved recording SSNMR spectra on crystals of the same form as the native protein crystal structure (ANSII crystal) and on the PEGylated protein in the form of a sediment (PEG-ANSII sediment), and then comparing the two datasets. Resonance assignments were done using a combination of solid-state and solution NMR experiments. Structural restraints obtained by NMR were used to confirm the fold of both the native protein in solution and the sedimented PEGylated protein. Finally, a structural model was obtained through restraints-driven docking calculations, which were validated

against residual dipolar couplings (RDCs) measured on the native protein (Cerofolini et al., 2019).

To summarize, in recent years, integrative and hybrid approaches that combine the strengths of various biophysical techniques, such as solution and solid-state NMR, crystallography, and cryo-EM, have become essential in studying the structural biology of large enzyme complexes. The combination of multiple magnetic resonance techniques, such as electron paramagnetic resonance, can yield complementary information, as demonstrated in studies of MsbA and DnaB. Integrative biophysics was used to study caseinolytic protease and non-ribosomal peptide synthetase, where crystallography, NMR, molecular dynamics simulations, and isothermal titration calorimetry provided insights into allosteric transitions and subunit assembly. Signal enhancement techniques, such as dynamic nuclear polarization (DNP) and high-speed MAS, are actively being developed for ssNMR, while magnetic resonance in the cell is gaining momentum. By joining radical moieties to substrates and inhibitors, investigations of enzymes even against a strong background from cellular components become feasible (Schütz, 2021). The future of NMR as an integral building block in the structural biology of large enzyme complexes is bound to be shaped by the union of spectroscopic method development and advances in biochemical sample preparation.

3.6 Conclusions

This chapter illustrates how Integrated NMR structural biology is important for enzyme systems as this technique allows for a comprehensive understanding of the structure and function of enzymes at the atomic level. As shown in this chapter, by using a combination of solid-state and solution NMR spectroscopy, X-ray crystallography, First Principles DFT calculations, cryo-EM, and other biophysical techniques, research groups can obtain detailed structural information on large enzyme complexes and their interactions with substrates, cofactors, and inhibitors.

Integrative NMR structural biology allows for obtaining complementary information from different techniques, which can provide a more comprehensive picture of the enzyme's structure and function. For example, as discussed before solid-state NMR can provide information on the structure and dynamics of enzymes in their native environment, while solution NMR can provide information on the same enzyme in solution. X-ray crystallography can provide high-resolution structural information on the enzyme and its interactions with ligands, while cryo-EM can provide information on the overall architecture of large enzyme complexes. All these structural methods can be integrated with First Principles DFT calculations to build testable models and decipher active site chemistry of enzyme complexes with no size limitations.

By combining these different techniques, researchers can obtain a more detailed understanding of the mechanism of enzyme catalysis, identify potential drug targets, and design new inhibitors or modulators of enzyme activity. This information can be crucial for the development of new therapeutics for the treatment of diseases caused by enzyme dysregulation.

3.7 References

Aumiller, W. M., Uchida, M., & Douglas, T. (2018). Protein cage assembly across multiple length scales [10.1039/C7CS00818J]. *Chemical Society Reviews*, *47*(10), 3433-3469. https://doi.org/10.1039/C7CS00818J

Barends, T. R., Domratcheva, T., Kulik, V., Blumenstein, L., Niks, D., Dunn, M. F., & Schlichting, I. (2008). Structure and mechanistic implications of a tryptophan synthase quinonoid intermediate. *ChemBioChem*, *9*(7), 1024-1028.

Bryce, D. L., & Taulelle, F. (2017). NMR crystallography. In (Vol. 73, pp. 126-127): International Union of Crystallography 5 Abbey Square, Chester, Cheshire CH1 ….

Brzovic, P. S., Kayastha, A. M., Miles, E. W., & Dunn, M. F. (1992). Substitution of glutamic acid 109 by aspartic acid alters the substrate specificity and catalytic activity of the. beta.-subunit in the tryptophan synthase bienzyme complex from Salmonella typhimurium. *Biochemistry*, *31*(4), 1180-1190.

Caulkins, B. G., Bastin, B., Yang, C., Neubauer, T. J., Young, R. P., Hilario, E., Huang, Y.-m. M., Chang, C.-e. A., Fan, L., Dunn, M. F., Marsella, M. J., & Mueller, L. J. (2014). Protonation States of the Tryptophan Synthase Internal Aldimine Active Site from Solid-State NMR Spectroscopy: Direct Observation of the Protonated Schiff Base Linkage to Pyridoxal-5′-Phosphate. *Journal of the American Chemical Society*, *136*(37), 12824-12827. https://doi.org/10.1021/ja506267d

Caulkins, B. G., Young, R. P., Kudla, R. A., Yang, C., Bittbauer, T. J., Bastin, B., Hilario, E., Fan, L., Marsella, M. J., Dunn, M. F., & Mueller, L. J. (2016). NMR Crystallography of a Carbanionic Intermediate in Tryptophan Synthase: Chemical Structure, Tautomerization, and Reaction Specificity. *Journal of the American Chemical Society*, *138*(46), 15214-15226. https://doi.org/10.1021/jacs.6b08937

Cerofolini, L., Giuntini, S., Carlon, A., Ravera, E., Calderone, V., Fragai, M., Parigi, G., & Luchinat, C. (2019). Characterization of PEGylated Asparaginase: new opportunities from NMR analysis of large PEGylated therapeutics. *Chemistry–A European Journal*, *25*(8), 1984-1991.

Chan-Huot, M., Sharif, S., Tolstoy, P. M., Toney, M. D., & Limbach, H.-H. (2010). NMR studies of the stability, protonation states, and tautomerism of 13C-and 15N-labeled aldimines of the coenzyme pyridoxal 5′-phosphate in water. *Biochemistry*, *49*(51), 10818-10830.

Dunn, M. F. (2012). Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Archives of biochemistry and biophysics*, *519*(2), 154-166.

Dunn, M. F., Niks, D., Ngo, H., Barends, T. R., & Schlichting, I. (2008). Tryptophan synthase: the workings of a channeling nanomachine. *Trends in biochemical sciences*, *33*(6), 254-264.

Eigen, M. (1964). Proton transfer, acid-base catalysis, and enzymatic hydrolysis. Part I: elementary processes. *Angewandte Chemie International Edition in English*, *3*(1), 1-19.

Facelli, J. C., & Grant, D. M. (1993). Determination of molecular symmetry in crystalline naphthalene using solid-state NMR. *Nature*, *365*(6444), 325-327. https://doi.org/10.1038/365325a0

Gauto, D. F., Estrozi, L. F., Schwieters, C. D., Effantin, G., Macek, P., Sounier, R., Sivertsen, A. C., Schmidt, E., Kerfah, R., & Mas, G. (2019). Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nature communications*, *10*(1), 2697.

Ghosh, R. K., Hilario, E., Chang, C.-e. A., Mueller, L. J., & Dunn, M. F. (2022). Allosteric regulation of substrate channeling: Salmonella typhimurium tryptophan synthase [Review]. *Frontiers in Molecular Biosciences*, *9*. https://doi.org/10.3389/fmolb.2022.923042

Ghosh, R. K., Hilario, E., Liu, V., Wang, Y., Niks, D., Holmes, J. B., Sakhrani, V. V., Mueller, L. J., & Dunn, M. F. (2021). Mutation of βGln114 to Ala alters the stabilities of allosteric states in tryptophan synthase catalysis. *Biochemistry*, *60*(42), 3173-3186.

Guo, H., & Rubinstein, J. L. (2018). Cryo-EM of ATP synthases. *Current Opinion in Structural Biology*, *52*, 71-79. https://doi.org/https://doi.org/10.1016/j.sbi.2018.08.005

Gupta, R., Hou, G., Renirie, R., Wever, R., & Polenova, T. (2015). 51V NMR crystallography of vanadium chloroperoxidase and its directed evolution P395D/L241V/T343A mutant: protonation environments of the active site. *Journal of the American Chemical Society*, *137*(16), 5618-5628.

Harris, R. K., Wasylishen, R. E., & Duer, M. J. (2009). *NMR crystallography* (Vol. 4). John Wiley & Sons.

Harris, R. M., Ngo, H., & Dunn, M. F. (2005). Synergistic effects on escape of a ligand from the closed tryptophan synthase bienzyme complex. *Biochemistry*, *44*(51), 16886-16895.

Hartman, J. D., Neubauer, T. J., Caulkins, B. G., Mueller, L. J., & Beran, G. J. O. (2015). Converging nuclear magnetic shielding calculations with respect to basis and system size in protein systems. *Journal of Biomolecular NMR*, *62*(3), 327-340. https://doi.org/10.1007/s10858-015-9947-2

Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., & Karplus, M. (2007). Intrinsic motions along an enzymatic reaction trajectory. *Nature*, *450*(7171), 838-844.

Herschlag, D., & Natarajan, A. (2013). Fundamental challenges in mechanistic enzymology: Progress toward understanding the rate enhancements of enzymes. *Biochemistry*, *52*(12), 2050-2067.

Heymann, M., Opthalage, A., Wierman, J. L., Akella, S., Szebenyi, D. M., Gruner, S. M., & Fraden, S. (2014). Room-temperature serial crystallography using a kinetically optimized microfluidic device for protein crystallization and on-chip X-ray diffraction. *IUCrJ*, *1*(5), 349-360.

Hilario, E., Caulkins, B. G., Huang, Y.-M. M., You, W., Chang, C.-E. A., Mueller, L. J., Dunn, M. F., & Fan, L. (2016). Visualizing the tunnel in tryptophan synthase with crystallography: Insights into a selective filter for accommodating indole and rejecting water. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1864*(3), 268-279.

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., Ghosh, R. K., Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., Bogie, P. M., Paulino, J., Wang, X., Riviere, G., Bosken, Y. K., Struppe, J., Hassan, A., Guidoulianov, J., Perrone, B., Mentink-Vigier, F., . . . Mueller, L. J. (2022). Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *Proceedings of the National Academy of Sciences*, *119*(2), e2109235119. https://doi.org/doi:10.1073/pnas.2109235119

Hyde, C., Ahmed, S., Padlan, E., Miles, E. W., & Davies, D. (1988). Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *Journal of Biological Chemistry*, *263*(33), 17857-17871.

Jencks, W. P. (1975). Binding energy, specificity, and enzymic catalysis: the circe effect. *Advances in enzymology and related areas of molecular biology*, *43*, 219-410.

Jencks, W. P. (1987). *Catalysis in chemistry and enzymology*. Courier Corporation.

Kaur, H., Abreu, B. r., Akhmetzyanov, D., Lakatos-Karoly, A., Soares, C. u. M., Prisner, T., & Glaubitz, C. (2018). Unexplored nucleotide binding modes for the ABC exporter MsbA. *Journal of the American Chemical Society*, *140*(43), 14112-14125.

Klinman, J. P. (2015). Dynamically achieved active site precision in enzyme catalysis. *Accounts of chemical research*, *48*(2), 449-456.

Klinman, J. P., Offenbacher, A. R., & Hu, S. (2017). Origins of enzyme catalysis: Experimental findings for C–H activation, new models, and their relevance to prevailing theoretical constructs. *Journal of the American Chemical Society*, *139*(51), 18409-18427.

Knowles, J. R., & Albery, W. J. (1977). Perfection in enzyme catalysis: the energetics of triosephosphate isomerase. *Accounts of chemical research*, *10*(4), 105-111.

Kovalevsky, A. Y., Katz, A. K., Carrell, H. L., Hanson, L., Mustyakimov, M., Fisher, S. Z., Coates, L., Schoenborn, B. P., Bunick, G. J., Glusker, J. P., & Langan, P. (2008). Hydrogen Location in Stages of an Enzyme-Catalyzed Reaction: Time-of-Flight Neutron Structure of d-Xylose Isomerase with Bound d-Xylulose. *Biochemistry*, *47*(29), 7595-7597. https://doi.org/10.1021/bi8005434

Lai, J., Niks, D., Wang, Y., Domratcheva, T., Barends, T. R. M., Schwarz, F., Olsen, R. A., Elliott, D. W., Fatmi, M. Q., Chang, C.-e. A., Schlichting, I., Dunn, M. F., & Mueller, L. J. (2011). X-ray and NMR Crystallography in an Enzyme Active Site: The Indoline Quinonoid Intermediate in Tryptophan Synthase. *Journal of the American Chemical Society*, *133*(1), 4-7. https://doi.org/10.1021/ja106555c

Limbach, H.-H., Chan-Huot, M., Sharif, S., Tolstoy, P. M., Shenderovich, I. G., Denisov, G. S., & Toney, M. D. (2011). Critical hydrogen bonds and protonation states of pyridoxal 5′-phosphate revealed by NMR. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1814*(11), 1426-1437.

Michalczyk, R., Unkefer, C. J., Bacik, J.-P., Schrader, T. E., Ostermann, A., Kovalevsky, A. Y., McKenna, R., & Fisher, S. Z. (2015). Joint neutron crystallographic and NMR solution studies of Tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proceedings of the National Academy of Sciences*, *112*(18), 5673-5678. https://doi.org/doi:10.1073/pnas.1502255112

Miles, E. W. (1995). Tryptophan synthase: structure, function, and protein engineering. *Proteins: Structure, Function, and Engineering*, 207-254.

Miles, E. W., Kawasaki, H., Ahmed, S. A., Morita, H., & Nagata, S. (1989). The β Subunit of Tryptophan Synthase: Clarification of the roles of histidine 86, lysine 87, arginine 148, cysteine 170, and cysteine 230. *Journal of Biological Chemistry*, *264*(11), 6280-6287.

Mueller, L. J., & Dunn, M. F. (2013). NMR Crystallography of Enzyme Active Sites: Probing Chemically Detailed, Three-Dimensional Structure in Tryptophan Synthase. *Accounts of chemical research*, *46*(9), 2008-2017. https://doi.org/10.1021/ar3003333

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T. J., Fan, L., Mueller, L. J., & Dunn, M. F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: Evidence for two subunit conformations and four quaternary states. *Biochemistry*, *52*(37), 6396-6411.

Page, M. I., & Jencks, W. P. (1971). Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proceedings of the National Academy of Sciences*, *68*(8), 1678-1683.

Pauling, L. (1956). The future of enzyme research. *Henry Ford Hospital Medical Journal*, *4*(1), 1-4.

Perilla, J. R., Hadden-Perilla, J. A., Gronenborn, A. M., & Polenova, T. (2021). Integrative structural biology of HIV-1 capsid protein assemblies: combining experiment and computation. *Current Opinion in Virology*, *48*, 57-64. https://doi.org/https://doi.org/10.1016/j.coviro.2021.03.005

Perilla, J. R., Zhao, G., Lu, M., Ning, J., Hou, G., Byeon, I.-J. L., Gronenborn, A. M., Polenova, T., & Zhang, P. (2017). CryoEM Structure Refinement by Integrating NMR Chemical Shifts with Molecular Dynamics Simulations. *The Journal of Physical Chemistry B*, *121*(15), 3853-3863. https://doi.org/10.1021/acs.jpcb.6b13105

Pooransingh-Margolis, N., Renirie, R., Hasan, Z., Wever, R., Vega, A. J., & Polenova, T. (2006). 51V solid-state magic angle spinning NMR spectroscopy of vanadium chloroperoxidase. *Journal of the American Chemical Society*, *128*(15), 5190-5208.

Roy, M., Keblawi, S., & Dunn, M. F. (1988). Stereoelectronic control of bond formation in Escherichia coli tryptophan synthase: substrate specificity and enzymatic synthesis of the novel amino acid dihydroisotryptophan. *Biochemistry*, *27*(18), 6698-6704.

Schütz, A. K. (2021). Solid-state NMR approaches to investigate large enzymes in complex with substrates and inhibitors. *Biochemical Society Transactions*, *49*(1), 131-144.

Singh, H., Vasa, S. K., Jangra, H., Rovó, P., Päslack, C., Das, C. K., Zipse, H., Schäfer, L. V., & Linser, R. (2019). Fast Microsecond Dynamics of the Protein–Water Network in the Active Site of Human Carbonic Anhydrase II Studied by Solid-State NMR Spectroscopy. *Journal of the American Chemical Society*, *141*(49), 19276-19288.

Waller, M. P., Bühl, M., Geethalakshmi, K. e. R., Wang, D., & Thiel, W. (2007). 51V NMR chemical shifts calculated from QM/MM models of vanadium chloroperoxidase. *Chemistry–A European Journal*, *13*(17), 4723-4732.

Warshel, A. (1998). Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *Journal of Biological Chemistry*, *273*(42), 27035-27038.

Wolfenden, R. (1969). Transition state analogues for enzyme catalysis. *Nature*, *223*, 704-705.

Wu, Y., Fried, S. D., & Boxer, S. G. (2020). A preorganized electric field leads to minimal geometrical reorientation in the catalytic reaction of ketosteroid isomerase. *Journal of the American Chemical Society*, *142*(22), 9993-9998.

Young, R. P., Lewis, C. R., Yang, C., Wang, L., Harper, J. K., & Mueller, L. J. (2019). TensorView: A software tool for displaying NMR tensors. *Magnetic Resonance in Chemistry*, *57*(5), 211-223.

Chapter 4
A General Introduction to Tryptophan Synthase and Tyrosine Phenol Lyase

4.1 Introduction

Pyridoxal-5'-Phosphate (PLP)-dependent enzymes are found widely throughout nature and play important roles in various biological pathways, including natural product synthesis, amino acid metabolism, and glucose metabolism (Schneider et al., 2000). The first PLP-dependent enzyme structure was reported more than four decades ago, and since then, a wealth of structural and functional information has been discovered for a diverse range of these enzymes. Protein allostery is a functional mechanism that is increasingly appreciated in drug design, whereby the binding of a protein or ligand at a distant site can affect the structure, organization, and function of the enzyme's active site. This chapter briefly reviews some of the key allosteric features and global structural organization of two important PLP enzymes. Understanding the catalytic mechanisms could help identify key similarities and differences among enzyme families that could be leveraged for therapeutic development.

PLP, the active form of vitamin B6, is a coenzyme that is essential for the activity of many proteins found in bacteria and humans. These enzymes perform various functions, such as amino acid metabolism and the synthesis of antibiotics (Schneider et al., 2000). PLP-dependent enzymes usually bind the cofactor covalently through a conserved catalytic lysine residue, forming an internal aldimine. After binding the amino acid substrate, the internal aldimine is replaced by an external aldimine Schiff base with the amino group of the substrate and the PLP aldehyde, which releases the free lysine (Schneider et al., 2000; Thomas R Schneider et al., 1998). The reaction then progresses

through a quinonoid intermediate, with the PLP cofactor serving as an electron sink to stabilize the transient reaction intermediates (Eliot & Kirsch, 2004). After these initial steps, PLP-dependent enzymes facilitate a wide range of reactions, including decarboxylation, transamination, and racemization (Liang et al., 2019). In fact, PLP-dependent enzymes have been reported to catalyze over 140 distinct activities (Percudani & Peracchi, 2009).

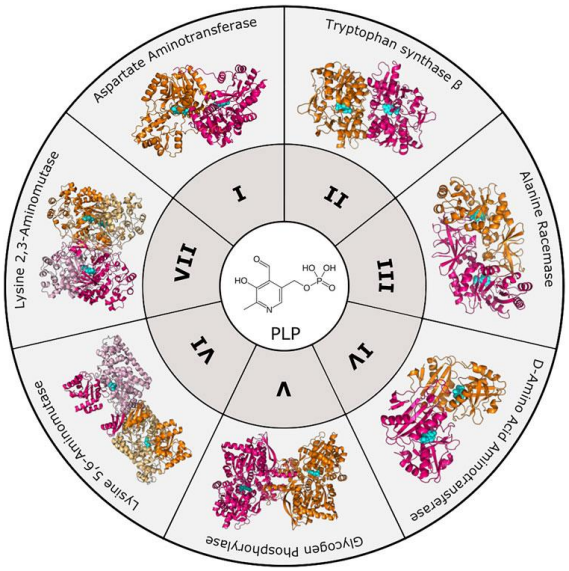


Fig 4.1 PLP-dependent enzymes are currently classified into seven distinct families based on their structural similarity to a prototype enzyme. These families include Aspartate Aminotransferase (Fold Type I, represented by PDB 8AAT), Tryptophan Synthase β-subunit (Fold Type II, represented by PDB 1BKS), Alanine Racemase (Fold Type III, represented by PDB 1SFT), D-Amino Acid Aminotransferase (Fold Type IV, represented by PDB 1DAA), Glycogen Phosphorylase (Fold Type V, represented by PDB 1GPB), Lysine 5,6-Aminomutase (Fold Type VI, represented by PDB 1XRS), and Lysine 2,3-Aminomutase (Fold Type VII, represented by PDB 2A5H). In these structures, the protomers are colored magenta and orange, while the PLP cofactor is shown in cyan. (Redrawn from (Tran & Brown, 2022)).

Currently, seven distinct families of PLP-dependent proteins have been identified based on their 3-dimensional structures (Fig 4.1). Despite limited overall sequence similarity, each family displays unique structural features. Notably, a particular fold type does not necessarily dictate a specific reaction as each fold can facilitate multiple types of

reactions [for a comprehensive review, refer to (Eliot & Kirsch, 2004)]. Fold Type I is the largest family, exemplified by aspartate aminotransferase, and comprises homodimers with each protomer consisting of a large and small subdomain. However, the PLP-binding sites are formed by residues from both subunits. Fold Type II, or the tryptophan synthase β family, has a similar structure to Type I, but the active sites usually comprise residues from only one protomer. Additionally, this family often contains regulatory domains and typically facilitates β-elimination, β-replacement, and racemization reactions. Fold Type III, exemplified by alanine racemase, is comprised of α/β-barrel domains connected to β-strand domains and functions as dimers. This family also includes several amino-acid decarboxylases. Fold Type IV, which includes D-amino acid aminotransferase and a few other enzymes, functions as homodimers but utilizes a distinct PLP-binding mode compared to Folds I and II. The unique Fold Type V group, which includes glycogen phosphorylase, uses the PLP phosphate group for catalysis, distinguishing it from other families (Eliot & Kirsch, 2004). Recently, two new PLP-dependent families, Fold Type VI (lysine 5,6-aminomutase) and Fold Type VII (lysine 2,3-aminomutase), were discovered to have structural differences from the other families (Percudani & Peracchi, 2009).

PLP-dependent enzymes exhibit not only structural diversity but also a wide range of allosteric mechanisms. Allostery refers to the ability of ligand or protein binding at one site to influence the binding or activity of another site on the protein. This phenomenon is critical for the modulation of enzyme activity within a specific cellular context or in the presence of binding partners. The Bohr effect, where the binding affinity of hemoglobin for oxygen changes in response to factors affecting blood pH, is a classic example of allostery. Since the discovery of allostery, it has been identified in many proteins that

perform various functions within the cell. The methods used to study allostery have also evolved along with our understanding of the phenomenon, and it can even be exploited for drug design. This review will explore the different allosteric mechanisms and their structural bases utilized by each PLP-dependent enzyme family. Although specific examples will be highlighted, it's essential to note that PLP-dependent enzymes employ a broad range of allosteric mechanisms, and further research in therapeutic development could benefit from a comprehensive understanding of these principles.

This dissertation is going to be focused on the understanding the structure function relationship in two different PLP enzyme systems: (i) Tryptophan Synthase, and (ii) Tyrosine Phenol Lyase belonging to Fold Type II and Fold Type I family of PLP enzymes respectively. Before introducing the enzymes and describing the mechanistic details, the Fold type I and Fold type II systems are described below in brief.

4.1.1 Fold Type I

Fold type I encompasses aminotransferases, except for aminotransferase class IV, as well as decarboxylase groups II and III and certain enzymes that exhibit α-, β-, or γ-elimination activity (Nick V Grishin et al., 1995; Kappes et al., 2011; Steffen-Munsberg et al., 2015). The Schiff base Lys residue of these fold type I enzymes is typically located near the C-terminus, with a glycine-rich loop involved in binding the PLP phosphate group and a hydrophobic β-strand preceding the Lys residue. A conserved aspartate residue interacts with the PLP ring N atom and is located 20-50 amino acids before the Lys residue (Nick V Grishin et al., 1995). Fold type I enzymes are typically homodimers or homotetramers, with each subunit containing a PLP molecule. The active site of these enzymes is located at the interface between subunits and consists of residues from two subunits, with most residues originating from one subunit (Han et al., 2010; Milano et al.,

2013; Schneider et al., 2000). Although dimers are typically the minimum assembly required for catalytic activity, the active site of Escherichia coli L-threonine aldolase tetramer contains residues from three monomers (Di Salvo et al., 2014). The subunit of fold type I enzymes consists of a large domain and a small domain, with the former comprising a seven-stranded β-sheet at the N-terminal and the latter folding into a 3- or 4-stranded β-sheet covered with helices on one side (Han et al., 2010; Schneider et al., 2000). Based on the structure of the N-terminal part, structural alignment further divides fold type I enzymes into six subclasses (Nick V Grishin et al., 1995; Käck et al., 1999; MEHTA et al., 1993; Schneider et al., 2000).

4.1.2 Fold Type II

The Fold Type II (also known as the tryptophan synthase β family) encompasses a diverse group of allosteric enzymes that play critical roles in various metabolic processes (Nick V Grishin et al., 1995; Griswold & Toney, 2011). These enzymes, including tryptophan synthase $\alpha_2\beta_2$ complex, threonine deaminase, threonine synthase, O-acetylserine sulfhydrylase, and cystathionine β-synthase (CBS), all share a conserved structural fold that includes a central beta-sheet surrounded by alpha-helices on either side, as well as a PLP-binding domain that interacts with PLP cofactor (Burkhard et al., 2000; Gallagher et al., 1998; Hyde et al., 1988). Enzymes in this fold are active in various oligomeric states, usually as dimers or tetramers, and accommodate allosteric regulation (Barends, Dunn, et al., 2008; Michael F Dunn, 2012; Michael F Dunn et al., 2008; Fatmi & Chang, 2010; Gallagher et al., 1998; Garrido-Franco et al., 2002). For instance, plant threonine synthase functions as a homodimer and is activated by AdoMet, while fungal threonine synthase functions as a monomer and is not subject to allosteric regulation (Garrido-Franco et al., 2002).

Cystathionine β-synthase (CBS) is an enzyme involved in the initial step of sulfur-containing amino acid biosynthesis, where it catalyzes the condensation of serine and homocysteine to produce cystathionine (Gerritsen & Waisman, 1964). CBS is unique among family members in how both the N-terminal and C-terminal domains participate in allostery (Taoka et al., 2002). CBS assembles as a tetramer, with each subunit consisting of a catalytic N-terminal domain that binds PLP and heme and a regulatory C-terminal domain (Meier et al., 2001).

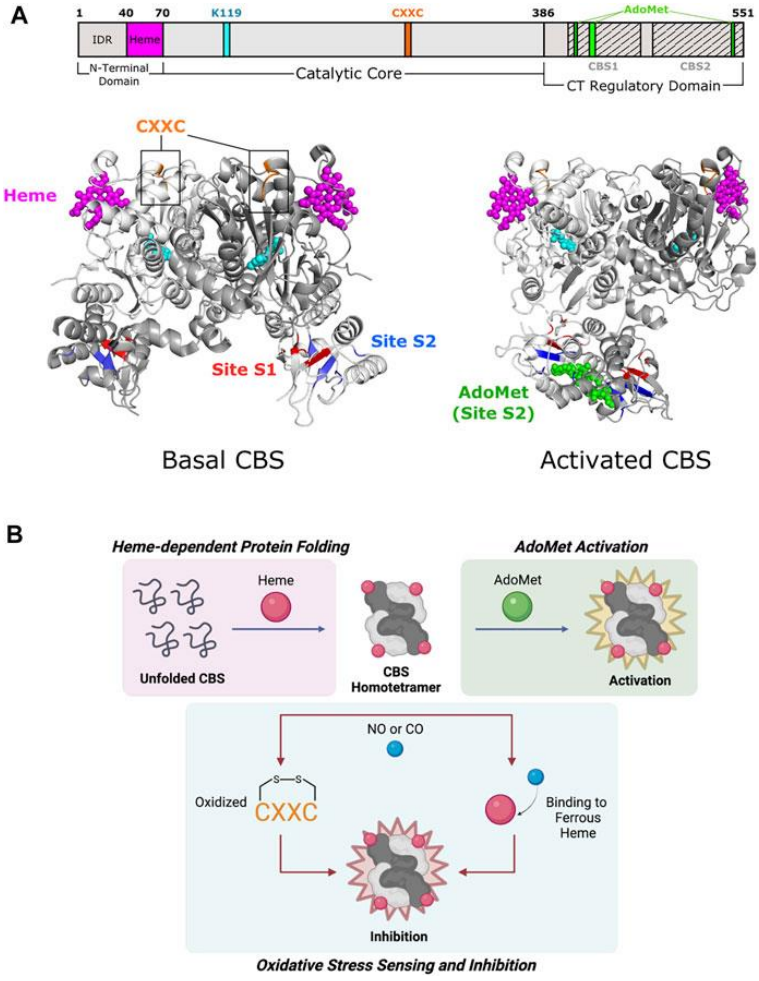


Fig 4.2 Cystathionine β-synthase of Fold Type II involves each of its domains in allostery. (A) Human CBS contains three domains that are all implicated in allostery. Heme binding occurs in a shallow pocket (aa 40–70; magenta) at the N-terminal domain. The catalytic core contains a CXXC oxidoreductase motif (aa 272–275; orange). The C-terminal regulatory domain has a Bateman module consisting of two tandem CBS motifs, CBS1 (aa 412–471) and CBS2 (aa 477–551), which fold to form Site S1 (M458, V459, Y484, F487, F508, and A509; red) and Site S2 (P422, L423, F443, A446, P447, V448, V533, and V534; blue) for AdoMet binding. (B) Formation of the CBS homotetramer is mediated by heme binding, in which the heme porphyrin scaffold facilitates protein folding. AdoMet binds and activates the CBS homotetramer. CBS inhibition takes place via oxidation of the CXXC motif and/or gaseous signaling molecule (i.e., NO, CO) binding to ferrous heme. (Redrawn from (Tran & Brown, 2022)).

In addition to PLP, CBS uses heme as a cofactor and is further activated by AdoMet (see Fig 4.2). The allostery in CBS involves both the binding of AdoMet to the regulatory C-terminal domain and the communication between the N-terminal and C-terminal domains. The allosteric regulation of CBS is critical for maintaining the balance of sulfur-containing amino acids in the body and for preventing the accumulation of toxic homocysteine (Meier et al., 2001; Taoka et al., 2002).

PLP enzymes in the fold type II class are characterized by a conserved structural fold that includes a central β-sheet surrounded by α-helices on either side, as well as a PLP-binding domain that is typically located at the interface between the beta-sheet and the alpha-helices (Catazaro et al., 2014). The PLP-binding domain is responsible for the recognition and binding of PLP, as well as the stabilization of the PLP cofactor in its active form (Miles, 1995; Shaw et al., 1997). This domain typically includes a conserved lysine residue that forms a Schiff base linkage with the PLP cofactor, as well as several other conserved residues that interact with the PLP pyridine ring and phosphate group (Shaw et al., 1997). PLP enzymes in the fold type II class play critical roles in a variety of metabolic processes, including the biosynthesis and catabolism of amino acids, neurotransmitters, and other biomolecules (Tran & Brown, 2022). The catalytic mechanism of PLP enzymes typically involves the formation of a Schiff base intermediate between PLP and a substrate molecule, followed by a series of proton transfer, elimination, and addition reactions that lead to the formation of a product molecule and the release of PLP (Michael F Dunn et al., 2008). Overall, the fold type II class of PLP enzymes represents an important and diverse group of biomolecules that play critical roles in a wide range of metabolic processes in living organisms (Riccardo Percudani & Alessio Peracchi, 2003).

4.2 Tryptophan Synthase (TS)

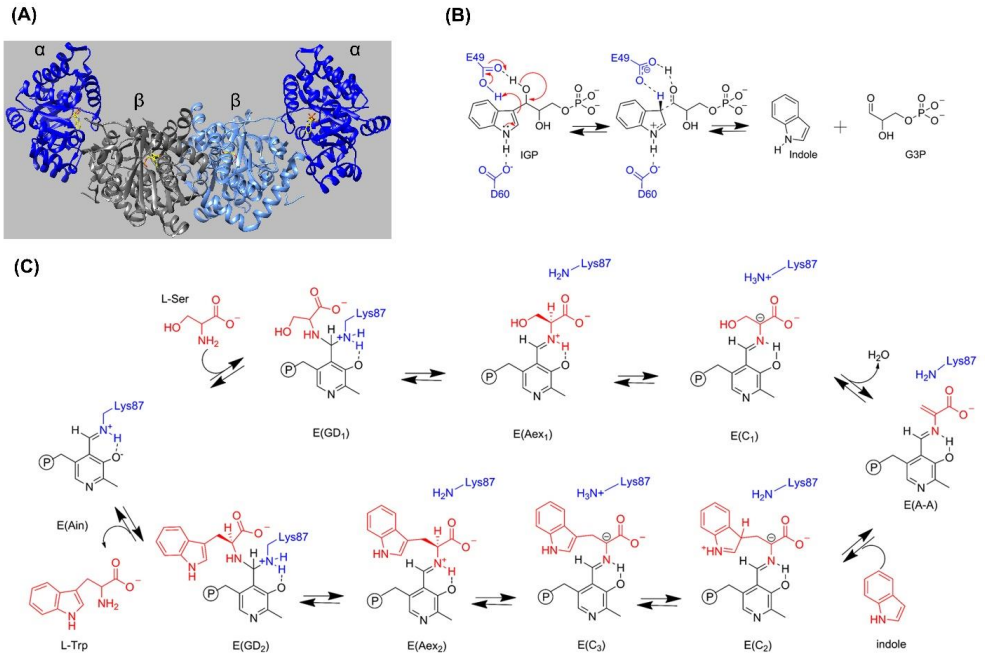


Fig 4.3. (A) Three-dimensional structural representation of the α2β2 tryptophan synthase hetero-tetrameric multienzyme complex from S. typhimurium. α-Subunits dark blue, β-subunits gray and light blue. The substrate for the α site, 3-indole D-glyceraldehyde 3'-phosphate, and the PLP cofactor as the internal aldimine covalently attached to βLys87 at the β-site are shown with yellow carbons. PDB ID: 2RHG(Barends, Domratcheva, et al., 2008). (B) α-Reaction with catalytic residues αGlu49 and αAsp60 shown in blue. (C) β-Reaction with the reacting substrates L-Ser and indole and product L-Trp shown in red. Catalytic residue βLys87 is shown in blue. PLP species are shown in black and reacting substrate species are shown in red (Redrawn from (Ghosh et al., 2022)).

The tryptophan synthase (TS, EC 4.2.1.20) enzyme is a fold type II PLP-dependent enzyme that catalyzes a β-elimination and replacement reaction (Yanofsky & Crawford, 1972). This reaction involves the removal of the β-hydroxyl group of L-serine and its replacement with indole to synthesize tryptophan. The discovery of the connection between L-tryptophan and indole was made during the search for new antibiotics. Scientists soon revealed the ability of LSer and indole to produce L-Trp (Platten et al., 2019). The PLP-dependency of the enzyme was discovered when researchers observed

that the purified enzyme could not convert indole and serine to tryptophan without the addition of PLP, which also gave the enzyme a yellow color (Umbreit et al., 1946). It was later found that the enzyme, which is only present in plants, bacteria, yeast, and fungi and not in higher mammals, consists of two different subunits, the α- and β-subunits, that catalyze different reactions (Crawford & Yanofsky, 1958). When the subunits were associated, the protein activity increased 30-100-fold (Michael F Dunn et al., 2008).

By identifying the individual reactions performed by the α- and β-subunits, researchers were able to determine that the α-subunit converts IGP to indole and G3P, while the β-subunit catalyzes the conversion of indole and serine to tryptophan and a water molecule (Miles, 1995). The use of indole produced in the α-reaction as a substrate in the β-reaction, and the lack of free indole as a substrate in the β-reaction, suggested that the active sites of each subunit are located in close proximity to allow for easy diffusion of the substrate between them (Michael F Dunn et al., 1990; William H Matchett, 1974; Yanofsky, 1955). Further investigations revealed that the $\alpha_2\beta_2$ complex could be dissociated into the individual subunits without affecting their catalytic activity (Andrew R Buller et al., 2015; Miles et al., 1989). Additionally, it was found that the dissociated β-subunits could still bind PLP and serine and form the active holoenzyme (Miles, 1995). This suggested that the α-subunit does not directly interact with the PLP cofactor and that the β-subunit is responsible for cofactor binding and activation (Kulik et al., 2005).

The $\alpha_2\beta_2$ complex also displays allosteric regulation, with tryptophan acting as an inhibitor of the enzyme (Michael F Dunn, 2012; Michael F Dunn et al., 1990; Michael F Dunn et al., 2008; Ghosh et al., 2022). Binding of tryptophan to the β-subunit causes a conformational change that reduces the affinity of the α-subunit for the β-subunit, leading

to decreased enzyme activity. This feedback inhibition mechanism ensures that the synthesis of tryptophan is tightly regulated and responds to changes in cellular concentrations of the amino acid (Mozzarelli & Bettati, 2006).

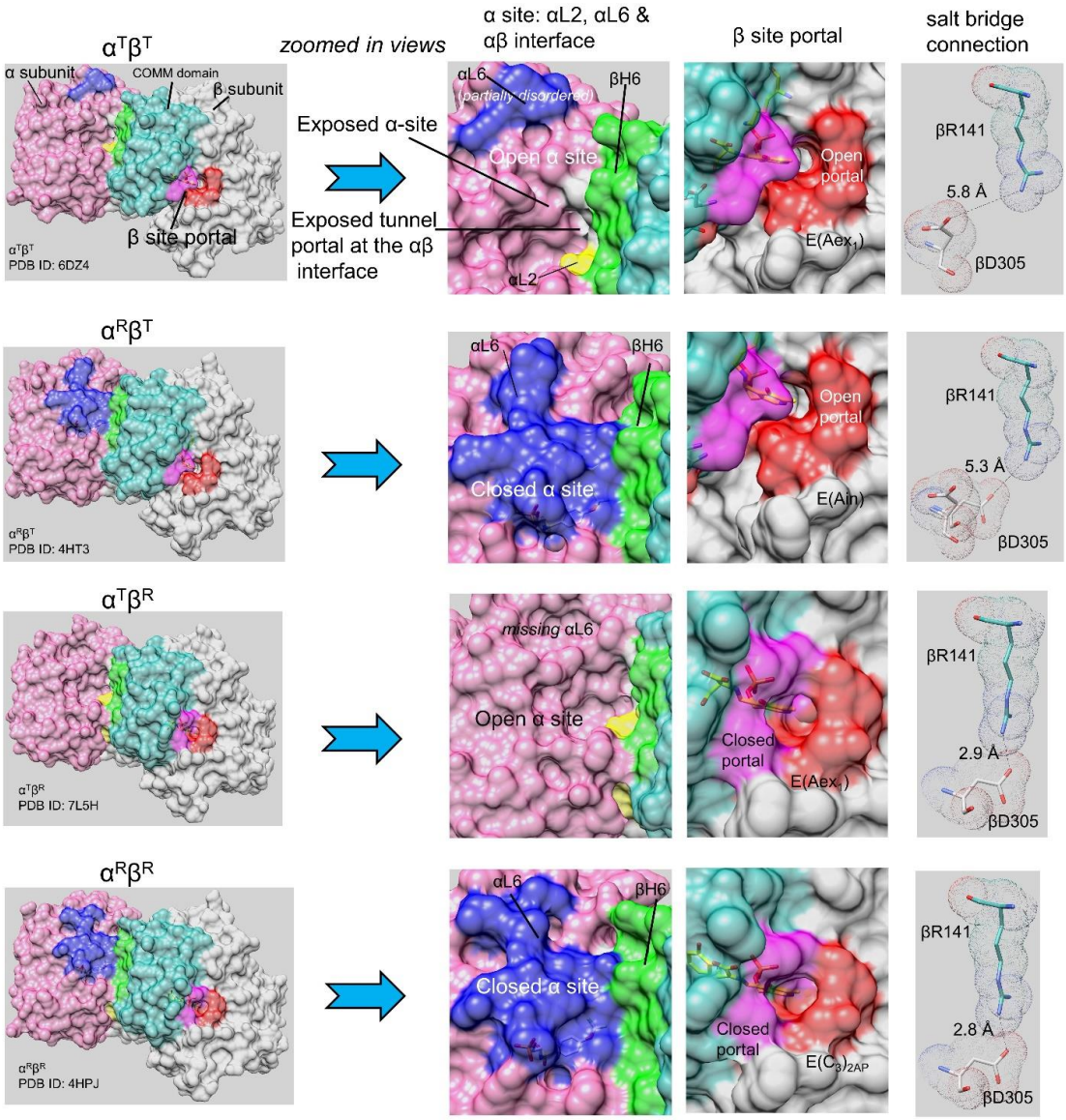

αTβT
zoomed in views
α site: αL2, αL6 & αβ interface
β site portal
salt bridge connection
α subunit
COMM domain
β subunit
β site portal
αTβT
PDB ID: 6DZ4
Exposed α-site
Exposed tunnel portal at the αβ interface
αL6
(partially disordered)
βH6
Open α site
αL2
Open portal
E(Aex1)
βR141
5.8 Å
βD305
αRβT
PDB ID: 4HT3
αL6
βH6
Closed α site
Open portal
E(Ain)
βR141
5.3 Å
βD305
αTβR
PDB ID: 7L5H
missing αL6
Open α site
Closed portal
E(Aex1)
βR141
2.9 Å
βD305
αRβR
PDB ID: 4HPJ
αL6
βH6
Closed α site
Closed portal
E(Cβ)2AP
βR141
2.8 Å
βD305

Fig 4.4 The TS allostery model consists of two subunit conformations, T and R, and four quaternary states, $\alpha^T\beta^T$, $\alpha^R\beta^T$, $\alpha^T\beta^R$ and $\alpha^R\beta^R$ (Dimitri Niks et al., 2013). Surface models are shown for the four quaternary states (left column) for each heterodimeric unit. The left column provides an overview of each quaternary state: each panel shows the α-subunit in light purple with loop αL2 (residues 53-60) yellow, loop αL6 (residues 179-193) blue; the β-subunit is shown in gray with the COMM domain (residues 102-189) in light blue-green and helix βH6 in green. The β-site portal is shown in dark purple (COMM domain residues) and orange. The central two columns show expanded views of the α- and β-subunits focusing on the catalytic sites, the α-β subunit interface, and with the portals into the interconnecting tunnel. Ligands bound to the α- and β-sites are shown as sticks. The last column shows structural detail for βArg141 and βAsp305 in each quaternary state; when the β-subunit is closed, these residues form an H-bonded salt bridge. Notice that the loop αL6 residues are either partially missing or completely missing in the α^T conformation, and that the β-subunit portal switches between open and closed states depending on whether the β-subunit is in the β^T or β^R conformation. Redrawn from (Ghosh et al., 2022).

4.2.1 Allosteric Switching in TS

As introduced in the earlier section, TS complex is composed of two α-subunits and two β-subunits, and the assembly of these subunits into the $\alpha_2\beta_2$ TS complex is essential for full activity (Hyde et al., 1988). The isolated α-subunit shows a catalytic activity that is significantly lower compared to the holo α2β2 complex, while the activity of the PLP-requiring β2 dimer is also diminished. The α-subunits catalyze the cleavage of substrate 3-indole D-glyceraldehyde 3'-phosphate (IGP), yielding indole and D-glyceraldehyde 3-phosphate (G3P), while the β-subunits catalyze replacement of the L-Ser β-hydroxyl by indole yielding L-Trp and a water molecule (Fig 4.3). During the catalytic cycle, indole formed at the α-site is transferred to the β-site via a 25 Å-long tunnel (Michael F Dunn et al., 1990). The channeling of the common metabolite, indole, is a key feature of the allosteric control mechanism for the synthesis of L-Trp. Within $\alpha_2\beta_2$, αβ dimers form allosteric units that work independently of each other while the allosteric interactions within each αβ unit are essential to the efficient synthesis of L-Trp.

The αβ reaction cycle relies on the swapping of subunits between two global conformational states: open and closed. Assuming each subunit has only two global conformations, there are four possible conformations of the αβ dimeric unit during each chemical step of the β-reaction. Fig 4.4 outlines these global conformations as they relate to the chemical state of the β-site and ligand-bound species along the tryptophan synthase αβ reaction pathway (Ghosh et al., 2022; Dimitri Niks et al., 2013). It is hypothesized that switching between global conformational states during the β-site's progression from one intermediate to the next provides the allosteric control that synchronizes the catalytic cycles of the α- and β-sites, preventing the escape of indole, which is the channeled intermediate (Michael F Dunn, 2012; Michael F Dunn et al.,

1990). Early research by the Dunn group found that the α-sites of certain forms of the enzyme (E(Ain), E(GD), and E(Aex)) reside in a low activity state, while conversion of $E(Aex_1)$ to E(A-A) activates the α-site by approximately 28-fold, and conversion of $E(C_3)$ to $E(Aex_2)$ deactivates the α-site by a similar factor (Ngo, Kimmich, et al., 2007). The binding of IGP analogues activates the conversion of $E(Aex_1)$ to E(A-A) by around 10-fold during Stage I of the β-reaction (Ngo, Harris, et al., 2007; Ngo, Kimmich, et al., 2007). Available structural and kinetic information suggests that the open conformations of the α- and β-subunits have little to no catalytic activity, while the closed conformations are active. When conformational state is correlated with catalytic activity, it appears that chemical transformations occur via closed global conformations, while the transfer of substrates/products between solution and the catalytic sites occurs via open global conformations. Indole transfer between the α- and β-sites occurs within αβ-dimeric units with the closed conformations to prevent indole escape. Under physiological conditions, bacterial cells have a significant pool of L-Ser, so the (quasi-stable) α-aminoacrylate species, $\alpha^T\beta^R$(A-A), produced in Stage I of the β-reaction is the predominant form of the bienzyme complex (Ghosh et al., 2022; Dimitri Niks et al., 2013). When the cell requires L-Trp, enzymes of the Trp operon convert phosphoenolpruvate and erythrose 4-phosphate through a series of metabolites to produce IGP. IGP binds to the α-site of TS where it is cleaved to indole and G3P, initiating Stage II of the β-reaction. After indole transfer from the α-site to the β-site via the interconnecting tunnel, the reaction proceeds in Stage II via the closed conformation of the αβ-dimeric unit to produce the L-Trp carbanion species, $(G3P)\alpha^R\beta^R(E_{2/3})$. Conversion of the carbanion species to $(G3P)\alpha^R\beta^R(Aex_2)$ and the switch to the quasi-stable $\alpha^R\beta^R(Aex_2)$ releases G3P. Conformational interconversions between inactive and active forms then allow

conversion of the L-Trp external aldimine to the L-Trp gem diamine, followed by the release of L-Trp and the formation of the internal aldimine.

This switching between open and closed subunit conformations is crucial for the allosteric regulatory mechanism of tryptophan synthase (TS) and is essential for the efficient synthesis of L-Trp. There are numerous X-ray crystal structures of TS in both open and closed states, available in the Protein Data Bank archive (Barends, Domratcheva, et al., 2008; Kulik et al., 2005; Thomas R Schneider et al., 1998), and many studies have demonstrated that ligand binding at the α-site and the covalent state of intermediates bound to the β-site can induce a switch between these states. The T-state and R-state designations for open and closed conformations, respectively, are used in accordance with early allosteric nomenclature (Forman, 2012; Ghosh et al., 2021). Rapid kinetic experiments have shown that ligands bind and dissociate rapidly in open conformations (T-state), while they bind and dissociate slowly in closed conformations (R-state). Mutations in the β-subunit can reverse the relative stabilities of these allosteric states as will be shown later (see Chapter 5). This switching between open and closed states of subunits is critical for the allosteric regulation of TS and for its efficient catalytic activity (Ghosh et al., 2022).

4.2.2 Substrate Channeling in TS

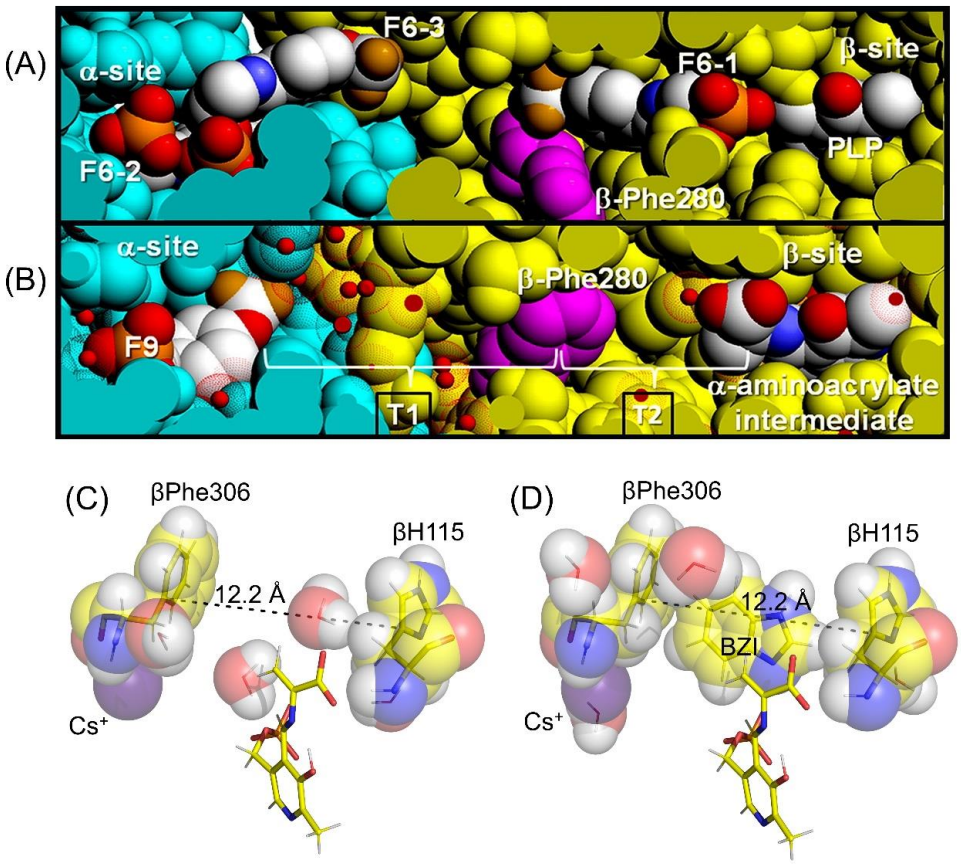


Fig 4.5. Tunnel views and Indole sub-site views at the VDW radii. (A) View showing F6 bound to the three sites identified in PDB ID: 4WX2 and designated as F6-1, F6-2, and F6-3 (CPK colors) compared with (B) the structure of the α-aminoacrylate intermediate (PDB ID: 4HN4). In these VDW views, some of the amino acid residues forming the channel have been cut away (slabbed) to expose the tunnel interior and reveal the locations of the three F6 molecules bound within the tunnel and the α-site. Water molecules are shown as small red balls inside dot surfaces at the VDW radius of oxygen. Color scheme: α-subunit residues, powder blue, β-subunit residues gold. The location of βPhe280 (magenta) is also shown. The brackets indicate the hydrophilic (T1) and hydrophobic (T2) regions of the tunnel. Figure redrawn from (Eduardo Hilario et al., 2016). (C) VDW contact structural detail of the β-subunit indole sub-site in E(A-A) (PDB ID: 4HN4) with hydrogens (white) modeled onto the heavy atoms. (D) E(A-A)(BZI) complex (PDB ID: 4HPX) showing structure detail of the indole subsite occupied by the indole isostere BZI including the VDW contacts of BZI with residues βPhe306 and βH115. The subsite distance of 12.2 Å is just right to bind BZI. The monovalent cation sites in these structures are occupied by Cs^+. Coloring scheme: carbons, yellow and in VDW spheres overlapping the stick representations, while the Cs^+ ions are colored purple. Redrawn from (Ghosh et al., 2022).

In biological systems, the selective transfer of small molecules and ions among cellular compartments and across cell membranes is achieved through the channeling of these molecules via tunnels within macromolecular assemblages (Michael F Dunn, 2012; Michael F Dunn et al., 1990; Galdiero et al., 2012; Eduardo Hilario et al., 2016; Shaffer et al., 2009; Ziervogel & Roux, 2013). This strategy is a key aspect of biological systems, as it allows for high selectivity and control over the transfer of these molecules(Raushel et al., 2003). The macromolecular protein structures responsible for these transfers form channels that act as molecular filters, restricting passage based on various factors such as charge, molecular cross-section, and hydrophobicity. Enzyme assemblages within certain metabolic pathways also utilize substrate channeling to prevent side reactions, enhance catalytic efficiency, and prevent the loss of labile small molecules into hydrophobic environments (Michael F Dunn et al., 2008).

The direct transfer of a substrate between the two catalytic sites of a bienzyme complex is an example of enzyme channeling. The enzyme system TS was the first to demonstrate this phenomenon, by channeling a common intermediate, indole, through an intermolecular tunnel that extends from the α-catalytic site near the α-β subunit interface to the β-catalytic site, a distance of approximately 30 Å (Michael F Dunn et al., 1990; Hyde et al., 1988). The tunnel is composed of two sections: T1, which is a relatively hydrophilic region filled with a hydrogen-bonded network of water molecules, and T2, which is a very hydrophobic, dewetted nanotube (Fig 4.5). The tunnel is lined with specific residues, such as βTyr279 and βPhe280, that restrict the passage of molecules based on their charge, molecular cross-section, and hydrophobicity. This enables TS to achieve high selectivity for indole and prevent the loss of labile small molecules into hydrophobic environments (Eduardo Hilario et al., 2016).

4.2.3 Monovalent Cation Activation in TS

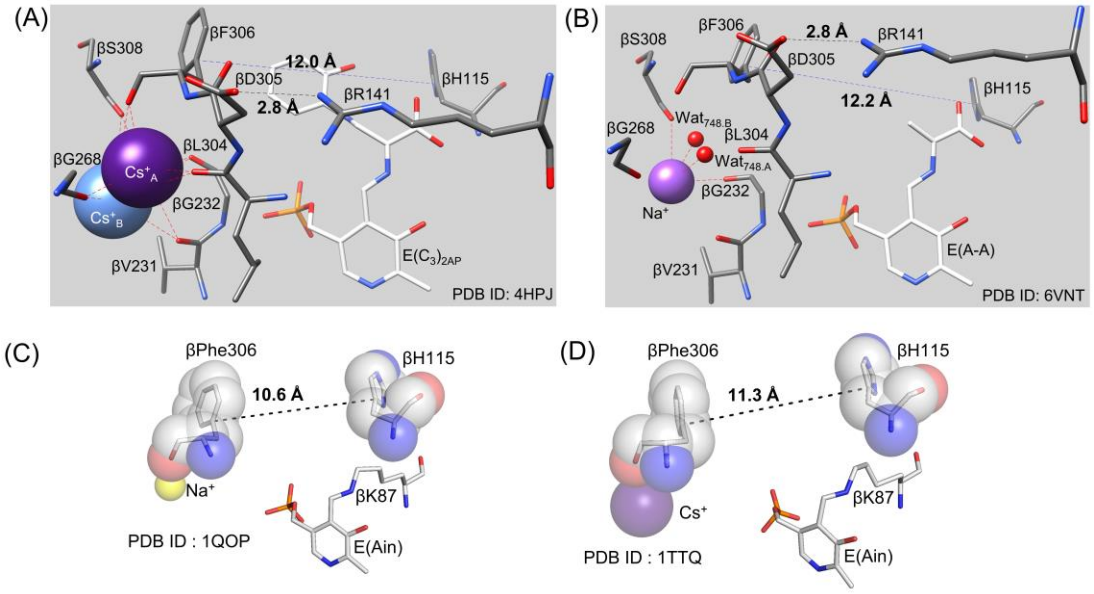


Fig 4.6. (A) and (B), Comparisons of the pre-formed indole binding sites and the monovalent cation sites of the $E(C_3)_{2AP}(Cs^+)$ complex (A) and the E(A-A)(Na^+) complex (B). The distances spanning the indole sub-site cavity between the Cɣ atoms of βPhe306 and βHis115 (12.0 Å and 12.2 Å, respectively) are just right to match the VDW surface of indole. In the E(Ain) monovalent cation complex with Na^+ (C), the subsite distance between βPhe306 and βH115 is 10.6 Å and is too small to accommodate indole. In the E(Ain)Cs^+ complex (D), the βPhe306 and βH115 distance is only slightly too small. (PDB IDs: 4HPJ, 6VNT, 1QOP and 1TTQ.). Redrawn from (Ghosh et al., 2022).

The TS enzyme's catalytic activity is significantly impaired in the absence of monovalent cations (MVCs), as shown by various studies (Adam T Dierkers et al., 2009; Mozzarelli, Peracchi, Rovegno, Dale, et al., 2000; Peracchi et al., 1995; Weber-Ban et al., 2001; Woehl & Dunn, 1999; Woehl & Dunn, 1995a). Specifically, the enzyme's activity is 45-fold less in the absence of Na+, which is one of the MVCs that can activate the enzyme (Michael F Dunn, 2012; Mozzarelli, Peracchi, Rovegno, Dale, et al., 2000). However, other MVCs, including K^+, Cs^+, $NH4^+$, and guanidinium ion, can also activate the enzyme. The structural information available suggests that a binding site for MVCs is located in the β-subunit of the $\alpha_2\beta_2$ bienzyme complex. This site can accommodate Na^+,

K^+, NH_4^+, or Cs^+ (Adam T Dierkers et al., 2009; Fan et al., 2000; Rhee et al., 1996). Thus, MVCs have a strong emphasis on TS catalytic activity (Fig 4.6).

According to kinetic studies, monovalent cation binding is crucial for both catalyzing the β-reaction and transmitting allosteric signals between the β- and α-sites of the TS enzyme. The pioneering work done in the Dunn laboratory (Adam T Dierkers et al., 2009; Weber-Ban et al., 2001; Woehl & Dunn, 1999; Woehl & Dunn, 1995a) provide evidence for this finding. Notably, the Na^+, K^+, and $NH4^+$ forms of the TS enzyme increase the catalytic activity of the β-reaction for L-Ser by about 30-fold, 26-fold, and 40-fold, respectively. However, even though the monovalent cation-free TS enzyme forms an α-aminoacrylate species, it does not activate the α-reaction. These observations were made in experiments cited in (Weber-Ban et al., 2001).

The TS enzyme's various forms, depending on the bound monovalent cation, lead to different distributions of intermediates along the β-reaction path. Specifically, Na+ promotes the formation of $E(Ain)^T$ and $E(Aex_1)^T$, while Cs^+ favors the formation of $E(A\text{-}A)^R$ and $E(C)^R$ (Adam T Dierkers et al., 2009). The available x-ray crystal structures suggest that the variation in ionic radii of the monovalent cations leads to changes in coordination number and geometry (Woehl & Dunn, 1995b), which in turn affects the dimensions of the β-subunit indole sub-site. The x-ray crystal structures of various TS enzymes with different bound monovalent cations identify the coordination site(s) as a cavity bounded by the backbone carbonyls of βVal231, βGly232, βGly268, βLeu304, βPhe306, and βSer308. Notably, these residues are not part of the COMM domain.

The different sizes of monovalent cations have an impact on their coordination with the carbonyls in the TS enzyme. Na^+ coordinates with three carbonyls and two waters, while Cs^+ coordinates with five or six carbonyls, almost filling the cavity. K^+ coordinates with

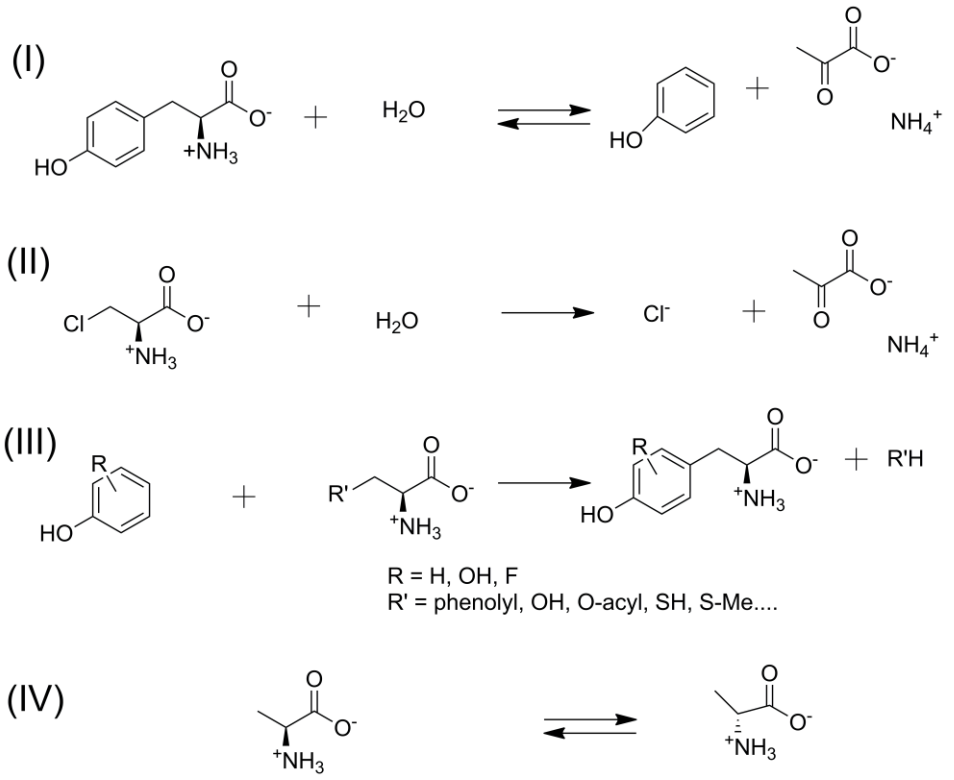
three carbonyls and one water. The residue βPhe306 plays a crucial role in stabilizing the R state conformation of the β-subunit indole sub-site, and its phenyl ring forms a component of the sub-site. The coordination of Cs^+ with βPhe306 affects the positioning of the βAsp305 side chain carboxylate, which stabilizes the salt bridge with βArg141 and the positioning of the βPhe306 phenyl group. This creates a preformed sub-site with dimensions that match the VDW surface of indole. The Na+ complex stabilizes the T-state, while Cs+ favors complexes with sub-sites that match or nearly match the dimensions of indole. The movement of the COMM domain causes the expansion of the indole sub-site, which matches the VDW surface of indole, allowing for its binding.

4.3 Tyrosine Phenol Lyase

Tyrosine phenol-lyase (TPL, EC 4.1.99.2; systematic name: L-tyrosine phenol-lyase (deaminating; pyruvate-forming)) is an enzyme that relies on pyridoxal 5'-phosphate (PLP) for its function (Kumagai et al., 1975). Its discovery dates to 1953 when Y. Kakihara and K. Ichihara reported the direct cleavage of L-tyrosine and its derivatives into phenol by bacterial cultures of *Escherichia coli* (Kakihara & Ichihara, 1953). M. Uchida et al. (M. Uchida, 1953) later isolated the enzyme responsible for this cleavage, which they named "β-tyrosinase," and demonstrated that it is a PLP-dependent enzyme that converts L-tyrosine into phenol and ammonium pyruvate.

TPL was initially identified in enterobacteria, but it has also been found in other bacteria and potentially some arthropods. The genes encoding TPL have been cloned and their primary structures deduced from DNA sequences of various bacterial species, including *Citrobacter intermedius*, *Citrobacter freundii*, *Enterobacter agglomerans,* and *Symbiobacterium thermophilum* (Duffey et al., 1977). With the advent of complete bacterial genome sequencing, TPL sequences from over 50 bacterial species have been

identified. The polypeptide sequences are highly similar, with identities no less than 43%, and TPL is in the cytoplasmic space of bacterial cells (Fleischmann et al., 1995). Its synthesis is induced by L-tyrosine and regulated by the TyrR protein and the cyclic AMP receptor protein (Iwamori et al., 1991).



Scheme 4.1 (I) Reversible β-elimination reaction of l-tyrosine catalyzed by TPL; (II) β-Elimination reaction of β-chloro-L-alanine catalyzed by TPL.; (III) β-Substitution reactions catalyzed by TPL.; (IV) Racemization of alanine catalyzed by TPL.

4.3.1 β-elimination Reactions.

Under physiological conditions, TPL catalyzes the reversible hydrolytical cleavage (αβ-elimination or β-elimination reaction) of L-tyrosine to produce phenol and ammonium pyruvate (Scheme 4.1) (Chen & Phillips, 1993). TPL can also catalyze the β-elimination reaction of several β-substituted L-amino acids and even some D-amino acids with good

leaving groups. However, amino acids and their analogs without good leaving groups, such as L-alanine, L-phenylalanine, L-methionine, and dicarboxylic amino acids, act as competitive inhibitors of wild-type TPL. TPL also catalyzes the β-elimination reaction for 3-chloro-L-alanine, while L-serine, L-cysteine, and their alkyl and aryl derivatives undergo practically irreversible decomposition catalyzed by TPL. However, the relative rates of these reactions are significantly lower than that for L-tyrosine (H. Chen et al., 1995; Chen & Phillips, 1993; H. Y. Chen et al., 1995; Phillips, 1987).

S-(o-nitrophenyl)-L-cysteine (SOPC) is a substrate that works well for the β-elimination reaction which is why it is commonly used to measure the enzymatic activity of TPL spectroscopically (Chen & Phillips, 1993; Milic et al., 2011). TPL can catalyze the irreversible β-elimination reaction of O-acyl-L-serine derivatives as well, with O-benzoyl-L-serine being a better substrate than O-acetyl-L-serine (Phillips, 1987).

The presence of a hydroxyl group at the para-position is crucial for substrate molecules to interact with the enzyme and progress to β-elimination. Hence, TPL cannot act on substrates that lack a hydroxyl group at the para-position, like L-phenylalanine, o-L-tyrosine, m-L-tyrosine, 4-fluoro-L-phenylalanine, and 4-chloro-L-phenylalanine. Although substrates with hydroxyl, halogen, methyl, or methoxy groups at the ortho- or meta-position can be degraded by TPL, they do so at a slower rate than L-tyrosine. However, TPL can still degrade substrates with bulky groups, such as 3-iodo-l-tyrosine and 3,5-diiodo-L-tyrosine (Chen & Phillips, 1993; H. Y. Chen et al., 1995; Phillips, 1987).

4.3.2 β-Substitution reactions

TPL also catalyzes β-substitution reactions between substrates for β-elimination and phenol derivatives, leading to the formation of various products including L-DOPA, 2,3,4-trihydroxyphenyl-L-alanine, and fluorinated derivatives of L-tyrosine (Milic et al., 2008;

Milic et al., 2011). The efficiency and selectivity of these reactions depend on the specific substrates used, with some amino acid derivatives being better substrates than others. The products formed from these reactions may have potential applications in the fields of medicine and materials science.

4.3.3 Alanine Racemization reactions

The racemization reaction catalyzed by TPL involves the transfer of a proton from the α-carbon to the enzyme's active site, followed by rotation of the α-carbon around the C-N bond (Milic et al., 2011). This results in the formation of an intermediate with a planar sp2-hybridized carbon atom, which can undergo attack by a water molecule from either side to form the racemic mixture of the amino acid. Although the racemization of alanine is not a biologically relevant reaction, it has been studied in detail as a model reaction for enzyme-catalyzed racemization and to investigate the mechanism of TPL-catalyzed reactions. The reaction has also been used to prepare racemic mixtures of amino acids for various applications in organic chemistry, materials science, and drug discovery.

4.3.4 Global Structural Organization of TPL

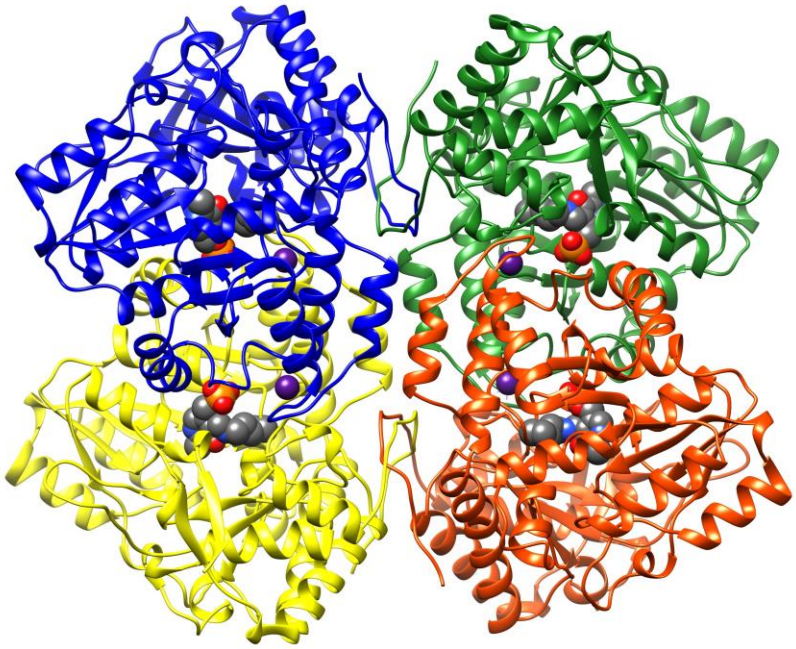


Fig 4.7 Quaternary structure of TPL molecule. (a) Schematic representation of overall organization of the TPL homotetramer. Different subunits are shown in different colors, and the two domains of each subunit are shown in different color tones. One catalytic dimer consists of the subunits shown in blue and yellow, while the other is formed from the subunits shown in green and red. Ribbon diagram of the TPL molecule (PDB code: 2TPL). Cofactor PLP, the side chains of the PLP binding Lys257 residues, and monovalent cations (Cs+) are represented by spheres. The view is along the molecular 2-fold axis P.

TPL is a homotetrameric enzyme, meaning that it is composed of four identical subunits. Each subunit has a molecular mass of around 51 kDa, and the tetramer has a total molecular mass of around 204 kDa. The subunits are arranged in the corners of a flattened tetrahedron, giving the tetramer a D2 point group symmetry with approximate dimensions of *80 x 60 x105 Å*^{3}. This arrangement is observed both in solution and in all known crystal structures of the enzyme. The homotetrameric structure of TPL is important for its catalytic activity, as it allows for the cooperative binding and activation of substrates at the active sites of the enzyme (Milić et al., 2006).

Each subunit of TPL can bind one molecule of PLP, which is the cofactor required for the enzyme's catalytic activity. The PLP binding residue is Lys257. The active site of the enzyme is situated between two protein subunits that are related by a 2-fold molecular axis, forming a catalytic dimer. The architecture of the TPL catalytic dimer is very similar to that of a dimeric molecule of aspartate aminotransferase (AAT) (Toney, 2014). Two catalytic dimers in a tetrameric TPL molecule are related by a molecular 2-fold axis, forming a dimer of dimers. This arrangement is important for the cooperative binding and activation of substrates at the active site of the enzyme.

The Fold Type I class of PLP-dependent enzymes, to which TPL belongs, is characterized by a conserved (β/α) 8-barrel fold (Borisov et al., 1980; Ford et al., 1980). This fold consists of eight parallel β-strands surrounded by eight α-helices. PLP is covalently bound to a lysine residue located at the C-terminus of the eighth α-helix, forming the internal aldimine. The active site is located at the interface between the N-terminal and C-terminal domains of the protein subunit. The N-terminal domain consists of the first four β-strands and the first two α-helices, while the C-terminal domain consists of the remaining β-strands and α-helices. The active site of TPL contains two residues, Asp214 and Lys257, that are essential for catalysis (Antson et al., 1994; Antson et al., 1992).

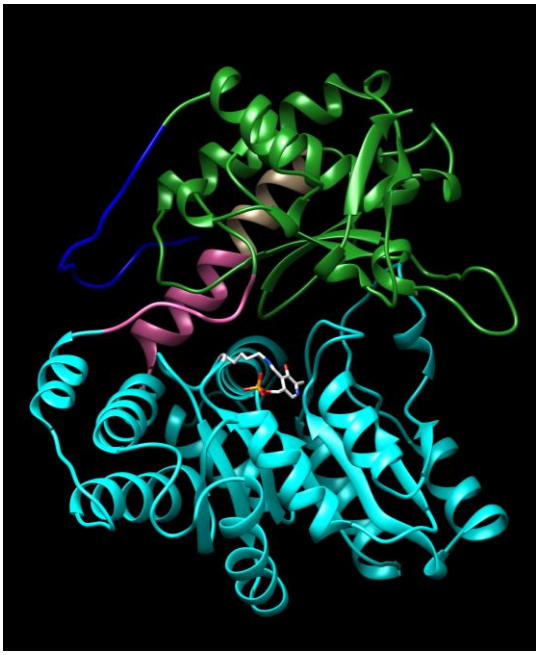


Fig 4.8 Structural and functional parts of the TPL subunit: the N-terminal arms (*blue*), the small domain (*green*), the large domain (*cyan*) and the connecting regions between domains (*pink*). PLP and the PLP-binding side chain of Lys257 are depicted as *sticks*.

Each protein subunit can be divided into four structural and functional parts (Fig 4.8):

• the N-terminal arms (residue: 19-48),

• the small domain (residue: 19-48, 333-456),

• the large domain (residue: 56-310), and

• the connecting regions between domains (residue: 49-56, 311-332).

4.3.5 Intersubunit contacts and monovalent cation binding site of the TPL

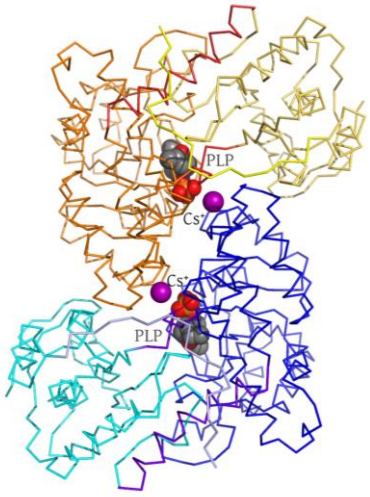


Fig 4.9 TPL catalytic dimer. Parts of each subunit are highlighted by *different colors*: the large domains – *blue* and *orange*; the small domains – *cyan* and *yellow-orange*; the connecting parts – *purple-blue* and *red*; the N-terminal arms – *light blue* and *yellow*. Monovalent cations (Cs+ in this case) and non-hydrogen atoms in PLP molecules are shown as *spheres*. Redrawn from (MILIĆ, 2010; Milić et al., 2006).

The catalytic dimer is formed by two subunits of TPL and is stabilized by hydrogen bonds and ionic interactions at the interface of the two subunits. When PLP molecules bind to the active sites of the subunits, as well as monovalent cations in their proximity, new interactions between monomers are formed, which further stabilizes the dimeric structure.

The TPL homotetramer, on the other hand, is stabilized by a hydrophobic cluster located in the center of the molecule. This hydrophobic cluster is formed by the side chains of Met56, Trp61, Met64, and Met65, each from a different protein subunit. The sulfur atoms of Met65 and indole rings of Trp61 are in the core of the hydrophobic cluster.

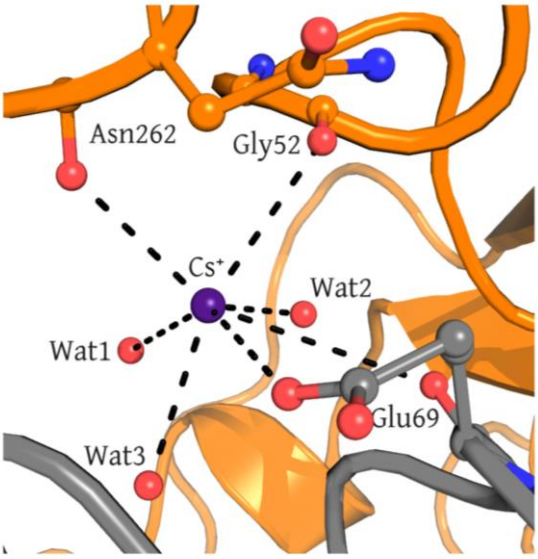


Fig 4.10 Monovalent cation binding site in TPL. Each cation is coordinated by seven oxygen atoms. Different subunits are shown in *different colors*. Redrawn from (MILIĆ, 2010; Milić et al., 2006).

Based on a survey of literature, it is observed that the presence of K^+, Rb^+, Cs^+, and NH_4^+ will enhance the β-elimination of L-tyrosine, while Li^+ will have no effect. Na^+ will act as a non-competitive inhibitor, meaning that it will bind to a site on the enzyme other than the active site, thereby reducing the enzyme's overall activity. The fact that the enzyme is mainly in the K^+-bound state in bacterial systems suggests that the presence of K^+ is important for the proper functioning of the enzyme. This is consistent with the observation that K^+ and other similar ions can activate the enzyme.

4.3.6 Brief description of the active site of the TPL

PLP forms an aldimine bond with the e-amino group of Lys257 in the enzyme, and its phosphate group is hydrogen bonded to several other residues in the active site.

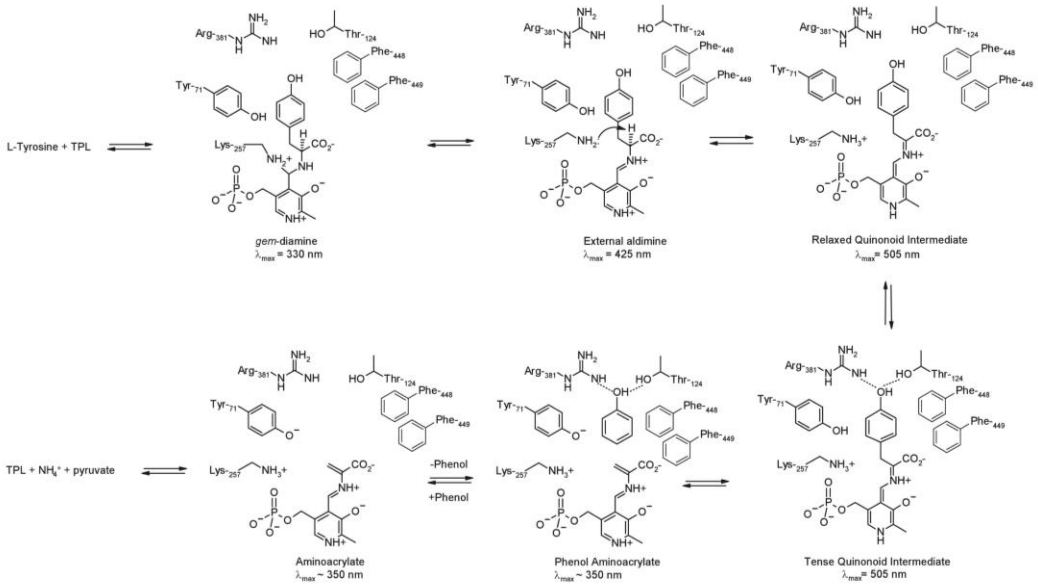
In the absence of PLP, sulfate anions can bind in the same site as the phosphate group of PLP in TPL. When PLP binds to the enzyme, it causes a displacement of the small domain towards the large one, resulting in a narrowing of the active site cleft. These structural changes presumably stabilize the enzyme and enhance its catalytic activity. The phosphate group is hydrogen bonded to the main chain NH group of the Gly99 and Arg100 and also to the sidechains of Gln98 (Asp98 in *E.herbicola* TPL), Arg100, and Ser254 (Milic et al., 2008; Milic et al., 2011; Milić et al., 2006).

The nitrogen (N1) atom of the PLP ring is hydrogen bonded to the sidechain of Asp214, whereas the PLP O3' atom is hydrogen bonded to the sidechain of Arg217. Parallel phenyl groups of the Phe123 stacks the pyridine ring. These interactions presumably stabilize the protein structure on a global scale. In addition, the structures of *E.herbicola* apo and holo TPL revealed displacement of the small domain towards the large one (by 1.5 Å on an average) and narrowing of the active site cleft upon PLP binding (Milić et al., 2006).

The pH dependence of the absorption and circular dichroism spectra of *C. freundii* TPL suggests that, in the presence of activating K^+ cations, the internal aldimine of the enzyme occurs in a cationic tautomeric form (Demidkina & Myagkikh, 1989). The predominant form (70%) is ketoenamine, which has hydronated nitrogen atoms and a dehydronated O3 atom. This form is in tautomeric equilibrium with enolimine, which has a dehydronated nitrogen atom of the aldimine bond and a dehydronated O3 atom (H. Y. Chen et al., 1995; Phillips et al., 2019; Phillips et al., 2016).

Overall, these observations provide insight into the molecular interactions between PLP and TPL and suggest a mechanism by which PLP binding enhances the catalytic activity of the enzyme.

4.3.7 Proposed β-elimination reaction mechanism in TPL



Scheme 4.2 Proposed Mechanism of TPL (Redrawn from (Phillips et al., 2016)).

The reaction mechanism for the β-elimination of L-tyrosine catalyzed by TPL involves the formation of a quinonoid intermediate, which is stabilized by delocalization into the z orbitals of the pyridoxal-P iminium ion. The proposed mechanism proceeds via several main steps, including the formation of the external aldimine with the substrate molecule, abstraction of the Cα hydrogen of the substrate yielding the quinonoid intermediate, hydronation of Cγ and cleavage of the Cβ–Cγ bond, resulting in the removal of phenol and formation of the α-aminoacrylate intermediate. The latter is attacked further by the ε-amino group of the active site lysine residue that binds the coenzyme (Lys257), resulting in regeneration of the holoenzyme and a release of iminopyruvate. Finally,

iminopyruvate is non-enzymatically hydrolyzed to give ammonium pyruvate(Phillips et al., 2016).

One of the challenges taken on in this thesis is determining the protonation states of the PLP cofactor in the catalytic pathway of TPL, as this information is currently unknown. This chapter lays the groundwork for the studies described in the latter half of the thesis, which aim to answer key mechanistic questions using NMR Crystallography (*see Chapter 3 for description*).

4.4 References

Antson, A., Dodson, G., Wilson, K., Pletnev, S., Harutyunyan, E., & Demidkina, T. (1994). Crystallographic studies of tyrosine phenol-lyase. Biochemistry of Vitamin B 6 and PQQ,

Antson, A. A., Strokopytov, B. V., Murshudov, G. N., Isupov, M. N., Harutyunyan, E. H., Demidkina, T. V., Vassylyev, D. G., Dauter, Z., Terry, H., & Wilson, K. S. (1992). The polypeptide chain fold in tyrosine phenol-lyase, a pyridoxal-5′-phosphate-dependent enzyme. *FEBS letters*, *302*(3), 256-260.

Barends, T. R., Domratcheva, T., Kulik, V., Blumenstein, L., Niks, D., Dunn, M. F., & Schlichting, I. (2008). Structure and mechanistic implications of a tryptophan synthase quinonoid intermediate. *ChemBioChem*, *9*(7), 1024-1028.

Barends, T. R., Dunn, M. F., & Schlichting, I. (2008). Tryptophan synthase, an allosteric molecular factory. *Current opinion in chemical biology*, *12*(5), 593-600.

Borisov, V., Borisova, S., Sosfenov, N., & Vainshtein, B. (1980). Electron density map of chicken heart cytosol aspartate transaminase at 3.5 Å resolution. *Nature*, *284*(5752), 189-190.

Buller, A. R., Brinkmann-Chen, S., Romney, D. K., Herger, M., Murciano-Calles, J., & Arnold, F. H. (2015). Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proceedings of the National Academy of Sciences*, *112*(47), 14599-14604.

Burkhard, P., Tai, C.-H., Jansonius, J. N., & Cook, P. F. (2000). Identification of an allosteric anion-binding site on O-acetylserine sulfhydrylase: structure of the enzyme with chloride bound. *Journal of molecular biology*, *303*(2), 279-286.

Catazaro, J., Caprez, A., Guru, A., Swanson, D., & Powers, R. (2014). Functional evolution of PLP-dependent enzymes based on active-site structural similarities. *Proteins: Structure, Function, and Bioinformatics*, *82*(10), 2597-2608.

Chen, H., Gollnick, P., & Phillips, R. S. (1995). Site-Directed Mutagenesis of His343→ Ala in Citrobacter freundii Tyrosine Phenol-Lyase: Effects on the Kinetic Mechanism and Rate-Determining Step. *European Journal of Biochemistry*, *229*(2), 540-549.

Chen, H., & Phillips, R. S. (1993). Binding of phenol and analogs to alanine complexes of tyrosine phenol-lyase from Citrobacter freundii: Implications for the mechanisms of. alpha.,. beta.-elimination and alanine racemization. *Biochemistry*, *32*(43), 11591-11599.

Chen, H. Y., Demidkina, T. V., & Phillips, R. S. (1995). Site-directed mutagenesis of tyrosine-71 to phenylalanine in Citrobacter freundii tyrosine phenol-lyase:

evidence for dual roles of tyrosine-71 as a general acid catalyst in the reaction mechanism and in cofactor binding. *Biochemistry*, *34*(38), 12276-12283.

Crawford, I. P., & Yanofsky, C. (1958). On the separation of the tryptophan synthetase of Escherichia coli into two protein components. *Proceedings of the National Academy of Sciences*, *44*(12), 1161-1170.

Demidkina, T. V., & Myagkikh, I. V. (1989). The activity and reaction specificity of tyrosine phenol-lyase regulated by monovalent cations. *Biochimie*, *71*(4), 565-571.

Di Salvo, M. L., Remesh, S. G., Vivoli, M., Ghatge, M. S., Paiardini, A., D'Aguanno, S., Safo, M. K., & Contestabile, R. (2014). On the catalytic mechanism and stereospecificity of E scherichia coli l-threonine aldolase. *The FEBS journal*, *281*(1), 129-145.

Dierkers, A. T., Niks, D., Schlichting, I., & Dunn, M. F. (2009). Tryptophan synthase: structure and function of the monovalent cation site. *Biochemistry*, *48*(46), 10997-11010.

Duffey, S., Aldrich, J., & Blum, M. (1977). Biosynthesis of phenol and guaiacol by the hemipteran Leptoglossus phyllopus. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *56*(2), 101-102.

Dunn, M. F. (2012). Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Archives of biochemistry and biophysics*, *519*(2), 154-166.

Dunn, M. F., Aguilar, V., Brzovic, P., Drewe Jr, W. F., Houben, K. F., Leja, C. A., & Roy, M. (1990). The tryptophan synthase bienzyme complex transfers indole between the. alpha.-and. beta.-sites via a 25-30. ANG. long tunnel. *Biochemistry*, *29*(37), 8598-8607.

Dunn, M. F., Niks, D., Ngo, H., Barends, T. R., & Schlichting, I. (2008). Tryptophan synthase: the workings of a channeling nanomachine. *Trends in biochemical sciences*, *33*(6), 254-264.

Eliot, A. C., & Kirsch, J. F. (2004). Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annual review of biochemistry*, *73*(1), 383-415.

Fan, Y.-X., McPhie, P., & Miles, E. W. (2000). Regulation of tryptophan synthase by temperature, monovalent cations, and an allosteric ligand. Evidence from Arrhenius plots, absorption spectra, and primary kinetic isotope effects. *Biochemistry*, *39*(16), 4692-4703.

Fatmi, M. Q., & Chang, C.-e. A. (2010). The role of oligomerization and cooperative regulation in protein function: the case of tryptophan synthase. *PLoS computational biology*, *6*(11), e1000994.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., & Merrick, J. M. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science*, *269*(5223), 496-512.

Ford, G. C., Eichele, G., & Jansonius, J. N. (1980). Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proceedings of the National Academy of Sciences*, *77*(5), 2559-2563.

Forman, S. A. (2012). Monod-Wyman-Changeux allosteric mechanisms of action and the pharmacology of etomidate. *Current opinion in anaesthesiology*, *25*(4), 411.

Galdiero, S., Falanga, A., Cantisani, M., Tarallo, R., Elena Della Pepa, M., D'Oriano, V., & Galdiero, M. (2012). Microbe-host interactions: structure and role of Gram-negative bacterial porins. *Current Protein and Peptide Science*, *13*(8), 843-854.

Gallagher, D. T., Gilliland, G. L., Xiao, G., Zondlo, J., Fisher, K. E., Chinchilla, D., & Eisenstein, E. (1998). Structure and control of pyridoxal phosphate dependent allosteric threonine deaminase. *Structure*, *6*(4), 465-475.

Garrido-Franco, M., Ehlert, S., Messerschmidt, A., Marinkovic, S., Huber, R., Laber, B., Bourenkov, G. P., & Clausen, T. (2002). Structure and function of threonine synthase from yeast. *Journal of Biological Chemistry*, *277*(14), 12396-12405.

Gerritsen, T., & Waisman, H. A. (1964). Homocystinuria: Absence of cystathionine in the brain. *Science*, *145*(3632), 588-588.

Ghosh, R. K., Hilario, E., Chang, C.-e. A., Mueller, L. J., & Dunn, M. F. (2022). Allosteric regulation of substrate channeling: Salmonella typhimurium tryptophan synthase [Review]. *Frontiers in Molecular Biosciences*, *9*. https://doi.org/10.3389/fmolb.2022.923042

Ghosh, R. K., Hilario, E., Liu, V., Wang, Y., Niks, D., Holmes, J. B., Sakhrani, V. V., Mueller, L. J., & Dunn, M. F. (2021). Mutation of βGln114 to Ala alters the stabilities of allosteric states in tryptophan synthase catalysis. *Biochemistry*, *60*(42), 3173-3186.

Grishin, N. V., Phillips, M. A., & Goldsmith, E. J. (1995). Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Science*, *4*(7), 1291-1304.

Griswold, W. R., & Toney, M. D. (2011). Role of the pyridine nitrogen in pyridoxal 5′-phosphate catalysis: activity of three classes of PLP enzymes reconstituted with deazapyridoxal 5′-phosphate. *Journal of the American Chemical Society*, *133*(37), 14823-14830.

Han, Q., Ding, H., Robinson, H., Christensen, B. M., & Li, J. (2010). Crystal structure and substrate specificity of Drosophila 3, 4-dihydroxyphenylalanine decarboxylase. *PLoS One*, *5*(1), e8826.

Hilario, E., Caulkins, B. G., Huang, Y.-M. M., You, W., Chang, C.-E. A., Mueller, L. J., Dunn, M. F., & Fan, L. (2016). Visualizing the tunnel in tryptophan synthase with crystallography: Insights into a selective filter for accommodating indole and rejecting water. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1864*(3), 268-279.

Hyde, C., Ahmed, S., Padlan, E., Miles, E. W., & Davies, D. (1988). Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *Journal of Biological Chemistry*, *263*(33), 17857-17871.

Iwamori, S., Yoshino, S., Ishiwata, K.-I., & Makiguchi, N. (1991). Structure of tyrosine phenol-lyase genes from Citrobacter freundii and structural comparison with tryptophanase from Escherichia coli. *Journal of fermentation and bioengineering*, *72*(3), 147-151.

Käck, H., Sandmark, J., Gibson, K., Schneider, G., & Lindqvist, Y. (1999). Crystal structure of diaminopelargonic acid synthase: evolutionary relationships between pyridoxal-5′-phosphate-dependent enzymes. *Journal of molecular biology*, *291*(4), 857-876.

Kakihara, Y., & Ichihara, K. (1953). Studies of phenol formation. I. Method of the determination of phenol and its microbial formation from tyrosine and tyrosine derivatives. *Med J Osaka Univ*, *3*, 497-507.

Kappes, B., Tews, I., Binter, A., & Macheroux, P. (2011). PLP-dependent enzymes as potential drug targets for protozoan diseases. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1814*(11), 1567-1576.

Kulik, V., Hartmann, E., Weyand, M., Frey, M., Gierl, A., Niks, D., Dunn, M. F., & Schlichting, I. (2005). On the structural basis of the catalytic mechanism and the regulation of the alpha subunit of tryptophan synthase from Salmonella typhimurium and BX1 from maize, two evolutionarily related enzymes. *Journal of molecular biology*, *352*(3), 608-620.

Kumagai, H., Utagawa, T., & Yamada, H. (1975). Studies on tyrosine phenol lyase. Modification of essential histidyl residues by diethylpyrocarbonate. *Journal of Biological Chemistry*, *250*(5), 1661-1667.

Liang, J., Han, Q., Tan, Y., Ding, H., & Li, J. (2019). Current advances on structure-function relationships of pyridoxal 5′-phosphate-dependent enzymes. *Frontiers in Molecular Biosciences*, *6*, 4.

M. Uchida, Y. T., Y. Kakihara, and K. Ichihara. (1953). *Med. J. Osaka Univ.*, *3*, 509–519. .

Matchett, W. H. (1974). Indole channeling by tryptophan synthase of neurospora. *Journal of Biological Chemistry*, *249*(13), 4041-4049.

MEHTA, P. K., HALE, T. I., & CHRISTEN, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. *European Journal of Biochemistry*, *214*(2), 549-561.

Meier, M., Janosik, M., Kery, V., Kraus, J. P., & Burkhard, P. (2001). Structure of human cystathionine β-synthase: a unique pyridoxal 5′-phosphate-dependent heme protein. *The EMBO journal*, *20*(15), 3910-3916.

Milano, T., Paiardini, A., Grgurina, I., & Pascarella, S. (2013). Type I pyridoxal 5 [variant prime]-phosphate dependent enzymatic domains embedded within multimodular nonribosomal peptide synthetase and polyketide synthase assembly lines. *BMC Structural Biology*, *13*.

Miles, E. W. (1995). Tryptophan synthase: structure, function, and protein engineering. *Proteins: Structure, Function, and Engineering*, 207-254.

Miles, E. W., Kawasaki, H., Ahmed, S. A., Morita, H., & Nagata, S. (1989). The β Subunit of Tryptophan Synthase: Clarification of the roles of histidine 86, lysine 87, arginine 148, cysteine 170, and cysteine 230. *Journal of Biological Chemistry*, *264*(11), 6280-6287.

Milic, D., Demidkina, T. V., Faleev, N. G., Matkovic-Cˇalogovic, D., & Antson, A. A. (2008). Insights into the Catalytic Mechanism of Tyrosine Phenol-lyase from X-ray Structures of Quinonoid Intermediates∗. *Journal of Biological Chemistry*, *283*(43), 29206-29214.

Milic, D., Demidkina, T. V., Faleev, N. G., Phillips, R. S., Matkovic-Calogovic, D., & Antson, A. A. (2011). Crystallographic snapshots of tyrosine phenol-lyase show that substrate strain plays a role in C–C bond cleavage. *Journal of the American Chemical Society*, *133*(41), 16468-16476.

Milić, D., Matković-Čalogović, D., Demidkina, T. V., Kulikova, V. V., Sinitzina, N. I., & Antson, A. A. (2006). Structures of apo-and holo-tyrosine phenol-lyase reveal a catalytically critical closed conformation and suggest a mechanism for activation by K+ ions. *Biochemistry*, *45*(24), 7544-7552.

Mozzarelli, A., & Bettati, S. (2006). Exploring the pyridoxal 5′-phosphate-dependent enzymes. *The Chemical Record*, *6*(5), 275-287.

Mozzarelli, A., Peracchi, A., Rovegno, B., Dale, G., Rossi, G. L., & Dunn, M. F. (2000). Effect of pH and monovalent cations on the formation of quinonoid intermediates of the tryptophan synthase α2β2 complex in solution and in the crystal. *Journal of Biological Chemistry*, *275*(10), 6956-6962.

Ngo, H., Harris, R., Kimmich, N., Casino, P., Niks, D., Blumenstein, L., Barends, T. R., Kulik, V., Weyand, M., & Schlichting, I. (2007). Synthesis and characterization of allosteric probes of substrate channeling in the tryptophan synthase bienzyme complex. *Biochemistry*, *46*(26), 7713-7727.

Ngo, H., Kimmich, N., Harris, R., Niks, D., Blumenstein, L., Kulik, V., Barends, T. R., Schlichting, I., & Dunn, M. F. (2007). Allosteric regulation of substrate channeling in tryptophan synthase: modulation of the L-serine reaction in stage I of the β-reaction by α-site ligands. *Biochemistry*, *46*(26), 7740-7753.

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T. J., Fan, L., Mueller, L. J., & Dunn, M. F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: Evidence for two subunit conformations and four quaternary states. *Biochemistry*, *52*(37), 6396-6411.

Peracchi, A., Mozzarelli, A., & Rossi, G. L. (1995). Monovalent cations affect dynamic and functional properties of the tryptophan synthase. alpha. 2. beta. 2 complex. *Biochemistry*, *34*(29), 9459-9465.

Percudani, R., & Peracchi, A. (2003). A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO reports*, *4*(9), 850-854.

Percudani, R., & Peracchi, A. (2009). The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. *BMC bioinformatics*, *10*, 1-8.

Phillips, R. S. (1987). Reactions of O-acyl-L-serines with tryptophanase, tyrosine phenol-lyase, and tryptophan synthase. *Archives of biochemistry and biophysics*, *256*(1), 302-310.

Phillips, R. S., Craig, S., Kovalevsky, A., Gerlits, O., Weiss, K., Iorgu, A. I., Heyes, D. J., & Hay, S. (2019). Pressure and temperature effects on the formation of aminoacrylate intermediates of tyrosine phenol-lyase demonstrate reaction dynamics. *ACS Catalysis*, *10*(3), 1692-1703.

Phillips, R. S., Vita, A., Spivey, J. B., Rudloff, A. P., Driscoll, M. D., & Hay, S. (2016). Ground-state destabilization by Phe-448 and Phe-449 contributes to tyrosine phenol-lyase catalysis. *ACS Catalysis*, *6*(10), 6770-6779.

Platten, M., Nollen, E. A., Röhrig, U. F., Fallarino, F., & Opitz, C. A. (2019). Tryptophan metabolism as a common therapeutic target in cancer, neurodegeneration and beyond. *Nature reviews Drug discovery*, *18*(5), 379-401.

Raushel, F. M., Thoden, J. B., & Holden, H. M. (2003). Enzymes with molecular tunnels. *Accounts of chemical research*, *36*(7), 539-548.

Rhee, S., Parris, K. D., Ahmed, S. A., Miles, E. W., & Davies, D. R. (1996). Exchange of K+ or Cs+ for Na+ induces local and long-range changes in the three-

dimensional structure of the tryptophan synthase α2β2 complex. *Biochemistry*, *35*(13), 4211-4221.

Schneider, G., Käck, H., & Lindqvist, Y. (2000). The manifold of vitamin B6 dependent enzymes. *Structure*, *8*(1), R1-R6.

Schneider, T. R., Gerhardt, E., Lee, M., Liang, P.-H., Anderson, K. S., & Schlichting, I. (1998). Loop closure and intersubunit communication in tryptophan synthase. *Biochemistry*, *37*(16), 5394-5406.

Shaffer, P. L., Goehring, A., Shankaranarayanan, A., & Gouaux, E. (2009). Structure and Mechanism of a Na-Independent Amino Acid Transporter. *Science*, *325*(5943), 1010-1014. https://doi.org/doi:10.1126/science.1176088

Shaw, J. P., Petsko, G. A., & Ringe, D. (1997). Determination of the structure of alanine racemase from Bacillus stearothermophilus at 1.9-Å resolution. *Biochemistry*, *36*(6), 1329-1342.

Steffen-Munsberg, F., Vickers, C., Kohls, H., Land, H., Mallin, H., Nobili, A., Skalden, L., van den Bergh, T., Joosten, H.-J., & Berglund, P. (2015). Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. *Biotechnology advances*, *33*(5), 566-604.

Taoka, S., Lepore, B. W., Kabil, Ö., Ojha, S., Ringe, D., & Banerjee, R. (2002). Human cystathionine β-synthase is a heme sensor protein. Evidence that the redox sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme. *Biochemistry*, *41*(33), 10454-10461.

Toney, M. D. (2014). Aspartate aminotransferase: an old dog teaches new tricks. *Archives of biochemistry and biophysics*, *544*, 119-127.

Tran, J. U., & Brown, B. L. (2022). Structural Basis for Allostery in PLP-dependent Enzymes [Review]. *Frontiers in Molecular Biosciences*, *9*. https://doi.org/10.3389/fmolb.2022.884281

Umbreit, W. W., Wood, W., & Gunsalus, I. (1946). The activity of pyridoxal phosphate in tryptopbane formation by cell-free enzyme preparations. *Journal of Biological Chemistry*, *165*, 731-732.

Weber-Ban, E., Hur, O., Bagwell, C., Banik, U., Yang, L.-H., Miles, E. W., & Dunn, M. F. (2001). Investigation of allosteric linkages in the regulation of tryptophan synthase: the roles of salt bridges and monovalent cations probed by site-directed mutation, optical spectroscopy, and kinetics. *Biochemistry*, *40*(12), 3497-3511.

Woehl, E., & Dunn, M. F. (1999). Mechanisms of monovalent cation action in enzyme catalysis: the tryptophan synthase α-, β-, and αβ-reactions. *Biochemistry*, *38*(22), 7131-7141.

Woehl, E. U., & Dunn, M. F. (1995a). Monovalent metal ions play an essential role in catalysis and intersubunit communication in the tryptophan synthase bienzyme complex. *Biochemistry*, *34*(29), 9466-9476.

Woehl, E. U., & Dunn, M. F. (1995b). The roles of Na+ and K+ in pyridoxal phosphate enzyme catalysis. *Coordination chemistry reviews*, *144*, 147-197.

Yanofsky, C. (1955). [30] Tryptophan synthetase from neurospora.

Yanofsky, C., & Crawford, I. P. (1972). 1 tryptophan synthetase. In *The enzymes* (Vol. 7, pp. 1-31). Elsevier.

Ziervogel, B. K., & Roux, B. (2013). The binding of antibiotics in OmpF porin. *Structure*, *21*(1), 76-87.

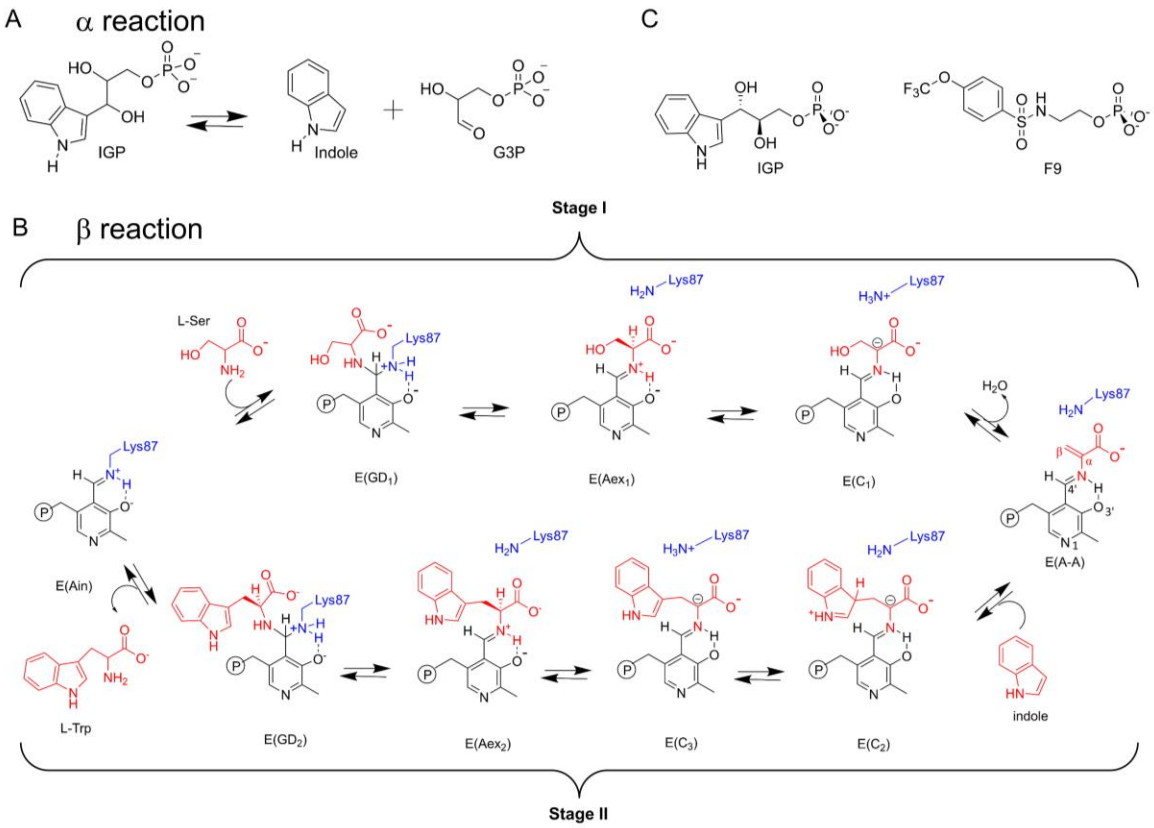
Chapter 5

Mutation of βGln114 to Ala Alters the Stabilities of Allosteric States in Tryptophan Synthase Catalysis.[*]

5.1 Introduction

Pyridoxal-5'-Phosphate (PLP) is a versatile cofactor that facilitates diverse chemical transformations of biological amines. This cofactor plays an essential role in the catalytic mechanisms of ~140 different enzymatic transformations broadly categorized as elimination, replacement, racemization, phosphorylation, decarboxylation, aldol cleavage, and transamination reactions (R. Percudani & A. Peracchi, 2003).

[*] The texts that appear in this chapter was published in *Biochemistry* journal (https://doi.org/10.1021/acs.biochem.1c00383). Rittik, K. Ghosh wrote the initial drafts and prepared the tables and figures, which then were subjected to editing by Michael F. Dunn and Leonard J. Mueller. The solution experimental work was carried out by Rittik, K. Ghosh and the x-ray crystal structures were solved by Eduardo Hilario and analyzed by Rittik K Ghosh.



Scheme 5.1: (A) α-reaction; (B) β-reaction; (C) comparison of the structures of IGP and the analog F9(Ngo, Harris, et al., 2007) that bind to the α-site of *St*TS.

The *Salmonella typhimurium* tryptophan synthase (*St*TS) αββα complex (Miles, 1979; Yanofsky & Crawford, 1972) catalyzes the formation of L-tryptophan from 3-indole-D-glycerol 3'-phosphate (IGP) and L-serine (Scheme 5.1). The α-subunit catalyzes cleavage of IGP into indole and glyceraldehyde-3-phosphate (G3P) (Yanofsky & Crawford, 1972) (Scheme 5.1A), while the β-subunit catalyzes the replacement of the substrate serine hydroxyl with indole to form L-tryptophan. Allosteric interactions regulate catalysis by switching the α and β subunits between open (T state) and closed (R state) conformations (Monod et al., 1965). This switching synchronizes catalysis and controls the channeling of indole between the α and β sites. The structural data base and ^{19}F NMR studies (D. Niks et al., 2013) provide evidence that TS allostery is

consistent with two subunit conformations designated as open (T) and closed (R) and four quaternary states $\alpha^T\beta^T$, $\alpha^T\beta^R$, $\alpha^R\beta^T$ and $\alpha^R\beta^R$ of the allosteric unit. The observation of only two β-subunit conformations in *St*TS contrasts with the observation of a third, partially closed conformation (R') in the stand alone β subunit (TrpB) from *Pyrococcus furiousus* (*Pf*TrpB) (A. R. Buller et al., 2015; Maria-Solano et al., 2019). β-subunit structures from *St*TS and *Pf*TrpB are compared in Fig 5.10.

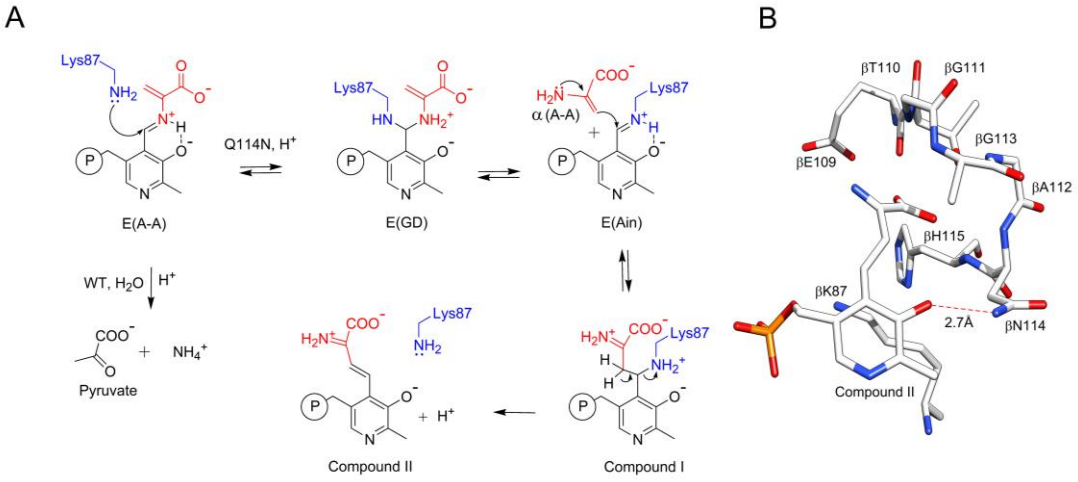
To further investigate the roles played by the amino acid residues in the vicinity of the β-active site that interact with the PLP moiety and the reacting substrate during the β-reaction cycle, we have undertaken a detailed investigation of the βQ114A TS variant. In this study we have selected experimental conditions where the α-subunits are stabilized in the R-state by the binding of the IGP analog, F9 (Scheme 5.1(C)), while the β-subunits switch between T and R states as dictated by the thermodynamic stability of the specific complex. We designate β subunit conformations in $\alpha_2\beta_2$ TS throughout this report as E^T or E^R for the T and R states, respectively.

Within the β active site, the catalytic cycle begins when bound substrate L-Ser makes a nucleophilic attack on the C4' carbon of the internal aldimine form of the PLP cofactor, E^T(Ain), and in subsequent steps is converted via gem diamine, $E^T(GD_1)$ (M. Roy et al., 1988), external aldimine, $E^T(Aex_1)$, and the L-Ser carbanion, $E^R(C_1)_{L\text{-}Ser}$, intermediates to the α-aminoacrylate intermediate, E^R(A-A), with the release of a water molecule (Stage I of the β-reaction) (Drewe Jr & Dunn, 1985, 1986). The β-subunit conformational cycle begins in Stage I with E^T(Ain), $E^T(GD_1)$, and $E^T(Aex_1)$ in the open conformation(M. F. Dunn, 2012). As $E^T(Aex_1)$ reacts, allosteric signaling triggered by formation of E^T(A-A) switches the α-subunit to the R state complex with IGP and activates the α-site ~30-fold (P. S. Brzovic et al., 1992; Michael F. Dunn et al., 1990; Ngo, Harris, et al., 2007),

catalyzing the retro aldol cleavage of IGP to give indole and G3P (Scheme 5.1(A)) (M. F. Dunn, 2012; Michael F. Dunn et al., 1990; Yanofsky & Crawford, 1972). The indole is trapped within the confines of the α-site and tunnel by formation of the closed conformations of the αβ-dimeric unit (M. F. Dunn, 2012; Michael F. Dunn et al., 1990). Thus, indole directly diffuses into the ~25 Å hydrophobic tunnel (E. Hilario et al., 2016) extending to the β-active site and binds to the indole subsite (Michael F. Dunn et al., 1990; W. H. Matchett, 1974). Then in Stage II of the β-reaction, indole makes a nucleophilic attack at the β-carbon of the E^R(A-A) to give $E^R(C_2)$ and forms $E^R(C_3)$ following proton extraction from the indole ring C-3 (Drewe Jr & Dunn, 1985, 1986). Protonation of $E^R(C_3)$ at the β-carbon gives the L-Trp external aldimine, $E^T(Aex_2)$. Nucleophilic attack of βLys87 at the PLP C4' carbon of $E^T(Aex_2)$ gives the L-Trp gem-diamine intermediate, $E^T(GD_2)$ and finally L-Trp is released along with regeneration of E^T(Ain) (Scheme 5.1B).

The β-subunit of tryptophan synthase is classified as a fold type II PLP enzyme (Miles, 2001; R. Percudani & A. Peracchi, 2003). This family of enzymes includes O-acetylserine-sulfhydrase, cystathione-β-synthase, threonine deaminase, serine dehydratase, and aminocyclopropane-1-carboxylate deaminase (N. V. Grishin et al., 1995; J. N. Jansonius, 1998). In most PLP-requiring enzymes, the PLP O3' oxygen is hydrogen-bonded to the Schiff base nitrogen (B. G. Caulkins et al., 2016b). Tautomers with the Schiff base nitrogen protonated are expected to be the catalytically active species when bond scission/formation at C-α or nucleophilic attack at C4' of PLP is required (Jencks, 1987; Walsh, 1979). In many fold type I and II PLP enzymes, the side chain of an Asn residue is H-bonded to the PLP O3' oxygen of the cofactor (Schiroli & Peracchi, 2015), presumably acting as a hydrogen bond donor and biasing protonation

of the Schiff base nitrogen (Chan-Huot et al., 2013). In TS, the carboxamide side chain of βQ114 is located close to the PLP O3'oxygen. However, most published TS structures show this carboxamide group rotated away from the cofactor while the β and γ methylene groups block access of potential H-bond donors to the PLP O3'. Molecular dynamics studies indicate that the βQ114 side chain can rotate into a position where the carboxamide group H-bonds to PLP O3' (Huang et al., 2016), although there is a lack of direct structural support for this in the protein data bank. Thus, TS appears to be quite unique amongst PLP-dependent enzymes. The role(s) played by βQ114 in the structure and function of TS and how the unique nature of these interactions in the active site of TS may be tied to both the allosteric control and catalysis are thus still not entirely clear.



Scheme 5.2. (A) Proposed βQ114N reaction mechanism (Blumenstein et al., 2007). (B) X-ray crystal structure (PDB ID: 2J9Y (Blumenstein et al., 2007)) of the modified cofactor in the βQ114N TS β-site showcasing the βL3 loop residues (109-115) and the H-bonding between βAsn114 and PLP O3' oxygen.

When βGln114 is mutated to Asn, a second chemistry competes with Stage II of the β-reaction (Scheme 5.2A). The competing side reaction involves the nucleophilic attack of βLys87 at the PLP C4' carbon of the α-aminoacrylate Schiff base intermediate, forming

E^T(Ain) and releasing the 3-carbon enamine, α-aminoacrylate, designated as α(A-A). Then α(A-A) makes an enamine-like nucleophilic attack on the PLP C4' carbon of E^T(Ain) to ultimately give Compounds I and II (Scheme 5.2A), inactivating the β-site. Interestingly, the X-ray crystal structure of Compound II bound to the β-site of βQ114N (PDB ID 2J9Y (Blumenstein et al., 2007); Scheme 5.2B) shows ND2 of the N114 carboxamide side chain hydrogen-bonded to the PLP O3'oxygen. The β-subunit of this structure has the closed conformation. Solid-state NMR (SSNMR) studies have established that the Schiff base nitrogen of the dominating tautomer of WT E^R(A-A) and $E^R(C_3)_{2AP}$ intermediates are predominantly neutral (B. G. Caulkins et al., 2016b). The α-aminoacrylate side reaction (Scheme 5.2A) exhibited by βQ114N previously has been reported for tryptophan synthase and several other PLP-requiring enzymes with catalytic mechanisms that involve α-aminoacrylate Schiff base intermediates (Roise et al., 1984). Indeed, it has been reported that the biological function of the RidA (Flynn & Downs, 2013) family of proteins is to deaminate 3- or 4-carbon enamines produced *in vivo* as a result of the side reactions of PLP-dependent serine/threonine dehydratases, thus preventing additional side reactions with PLP enzymes.

βQ114 does appear to play an important structural and functional role in TS. This residue is part of the substrate α-carboxylate binding loop (βL3, residues βT110 - βH115) (Barends, Domratcheva, et al., 2008; Kulik et al., 2002; T. R. Schneider et al., 1998), and is part of the communication (COMM) domain (residues β102-β189) (T. R. Schneider et al., 1998). The COMM domain (A. R. Buller et al., 2015) is involved in the switching of the β-subunits between T and R conformations and is essential both to the allosteric regulation of the catalytic activities of the α- and β-subunits of the $\alpha_2\beta_2$ oligomer and to interactions with the reacting substrate that facilitate catalysis. In both the T and R

β-subunit conformations (Barends, Domratcheva, et al., 2008; Blumenstein et al., 2007; M. F. Dunn, 2012; Miles, 2001; T. R. Schneider et al., 1998), the βQ114 side chain hydrogen-bonds to the carbonyl oxygen of βGly83. When the COMM domain is switched to the R conformation, the βQ114 side chain additionally hydrogen-bonds to βN145 and βR148 and becomes part of a network of hydrogen-bonds within the COMM domain that connect to the substrate carboxylate. In the T conformation βQ114 is too far away from βN145 and βR148 to form hydrogen-bonds and, except for βGly83, there is no other hydrogen-bonding linkage between the βQ114 side chain and other components of the β-subunit. During catalysis, the backbone amide N-H of βQ114 remains hydrogen-bonded to the carboxylate oxygen of the reacting substrate.

To test the extent to which βQ114 is involved in control of reaction specificity via selective protonation of active site atoms in TS and to what extent this residue is important for stabilization of closed (R) allosteric states in the β subunit, we present here a mechanistic investigation of the effects resulting from replacement of βGln114 by Ala on the spectroscopic, structural, and catalytic properties of the β-site. These comparisons consist of steady-state kinetic analyses of the β-reaction, static UV/Vis, and SSNMR spectroscopic studies of β-site complexes together with single crystal X-ray diffraction structures of the WT and βQ114A mutant for the E^T(Ain), $E^T(Aex_1)$, and E^R(A-A) complexes (PDB ID: 6C73, 6DZO, 6D0V), and for the $E^R(C_3)_{2AP}$ species formed with 2-aminophenol (2AP) (PDB ID: 6O1H), a nucleophilic analog of indole (Barends, Domratcheva, et al., 2008; Blumenstein et al., 2007; M. F. Dunn, 2012). As will be shown here, the replacement of Gln by Ala alters the catalytic functioning and

spectroscopic properties of the TS β-site with important ramifications for understanding the role(s) played by βQ114 in the stabilization of TS allosteric states during catalysis.

5.2 Experimental Section

5.2.1 Materials. PLP, L-Ser, L-His, L-Trp, indole, indoline, benzimidazole (BZI), CsCl, and 2-AP were purchased from Sigma-Aldrich. Bicine was bought from Thomas Scientific. Unless mentioned otherwise, all solution studies were conducted at 22 ± 3°C and maintained at pH 7.8 in 0.05 M Bicine CsCl buffer containing 1 mM EDTA.

5.2.2 Site-Directed Mutagenesis. All molecular biology steps were performed as described in Molecular Cloning: A Laboratory Manual (Sambrook et al., 1989). The translational coupling gene encoding the ***α***- and ***β***-chains of WT TS was previously cloned into NcoI and SacI restriction sites to generate the construct pEBA-10 expression vector. This construct was used as the DNA template during two-steps of polymerase chain reaction (PCR) to introduce specific point-mutations alongside the ***β***-chain polynucleotide sequence of WT TS as described previously (Hilario et al., 2020), and as outlined in the SI section.

5.2.3 X-ray Crystal Structures. TS samples in complex with Cs^+ bound to the monovalent cation site at the enzyme β-subunit and the IGP analog F9 (Ngo, Harris, et al., 2007) bound to the enzyme α-subunit were prepared with different combinations of β-site substrate and substrate analogs. These crystals diffract to high resolution. Therefore, most of the WT and the βQ114A variant X-ray crystal structures discussed in this report are F9 complexes of the Cs^+ enzymes. These include the internal aldimine, L-Ser external aldimine, α-aminoacrylate, and the 2AP carbanion complexes.

5.2.4 Site-Directed Mutagenesis. The translational coupling gene encoding the ***α***- and ***β***-chains of WT TS was previously cloned into NcoI and SacI restriction sites to generate

the construct pEBA-10 expression vector. This construct was used as the DNA template during two-steps of polymerase chain reaction (PCR) to introduce specific point-mutations alongside the ***β***-chain polynucleotide sequence of wild-type TS. Oligonucleotide TS-NcoI-FW and TS-SacI-REV, respectively, anneals upstream and downstream of the WT TS polynucleotide sequence in the pEBA-10 expression vector. The first round of PCR reactions was performed using the pairs of nucleotide primers TS-NcoI-FW (5'- CAA TTT CAC ACA GGA AAC AGA *CCA TGG* -3') and TS-βQ114A-REV (5'- C AGA GGC GAC GCC GTG C**GC** ACC GGC GCC GGT TTC -3') to generate fragment A, and TS-βQ114A-FW (5'- GAA ACC GGC GCC GGT **GC**G CAC GGC GTC GCC TCT G -3') and TS-SacI-REV (5'- TTA *TGC GCG* GCT GGC GGC TTT CAT GGC TGA G -3') to generate fragment B, where an underlined and italic sequence corresponds to a restriction site, and an underlined and bold sequence corresponds to a mutation. The purification of PCR products A and B from agarose gel slices was carried out by using Ultra-Sep gel extraction kit (Omega Bio-Tek). Fragments A and B were mixed in equimolar amounts, heat denatured, annealed, and the recombinant strands were extended. Then, oligo primers TS-NcoI-FW and TS-SacI-REV were added, and a second PCR was performed. The fragment encoding the βQ114A TS mutant was digested with NcoI and SacI, gel purified, and ligated with pEBA-10 vector, previously digested with the same enzymes and gel purified. The recombinant plasmid construct βQ114A TS was DNA sequenced to confirm the polynucleotide sequence. All molecular biology steps were followed as described in Molecular Cloning: A Laboratory Manual.

5.2.5 Enzymes. The wild-type $\alpha_2\beta_2$ TS was purified as previously described. The βQ114A mutant enzyme was purified by a modification of the procedure used for wild-type $\alpha_2\beta_2$. Cells containing the mutant plasmid were grown in LB media with IPTG

induction. The cell extracts were dissolved in Tris-HCl buffer and then adjusted to 45% Ammonium Sulfate saturation by stepwise addition. The yellow pellet was re-dissolved in a minimal volume of buffer and applied to a Sephadex S-200 gel filtration column (60 x 2.5 cm), eluting at 1 mL/min. The yellow fractions containing βQ114A TS were pooled and concentrated in a centrifugal concentrator (Vivaspin, Sartorius) with a 30000 MW cutoff. An extra dialysis step in Bicine buffer containing EDTA and Cesium Chloride (CsCl), pH 7.8, after size exclusion column separation, was added to ensure Cs^+ ion incorporation in the enzyme. Enzyme active site concentrations were estimated from the UV/Vis absorbance at 412 nm based on ε_{412} = 7.3 $mM^{-1}cm^{-1}$. Yields of both WT and βQ114A TS were >100 mg/L of culture.

5.2.6 Crystallization Assays. TS samples were centrifuged at 10,000 g for 10 min to clarify the solution before initiating any crystallization trials. Equal volumes (4 µL) of protein solution (15 mg ml^{-1}) and reservoir solution were mixed and equilibrated against a 500 µL reservoir solution consisting of 50 mM bicine-CsOH, pH 7.8, 50 mM CsCl, 6-10% PEG 8000, 2-8 mM spermine, 1 mM 2-({[4-(trifluoromethoxy)phenyl]sulfonyl}amino)ethyl dihydrogen phosphate (PDB ID = F9F) (E. Hilario et al., 2016). The crystal quality was optimized by screening different concentrations of PEG 8000 (6-10%) and spermine (2-8 mM) in the reservoir solution. Crystals of tryptophan synthase were grown using a 24-well sitting-drop vapor-diffusion plate (Cryschem M plate, HR1-002 (Hampton Research) at 25°C. Large TS crystals (80-200 µm) were used for X-ray diffraction experiments.

5.2.7 X-ray Data Collection and Analysis. Prior to X-ray diffraction data collection, crystals were briefly immersed in cryoprotectant consisting of the reservoir solution containing 30% dimethyl sulfoxide (DMSO) and mounted in a nylon-fiber loop (Hampton

Research, USA). Other TS βQ114A crystals were soaked in cryoprotectant buffer containing specific compounds such as L-Serine, and 2-aminophenol (PDB ID: 2AF), F9 (PDB ID: F9F). Crystals were flash-frozen in liquid nitrogen and shipped to the Northeastern Collaborative Access Team beamline 24ID-C at the Advanced Photon Source (APS), Argonne National Laboratory, for X-ray diffraction experiments. All X-ray diffraction data were collected at a wavelength of 0.9792 Å with a pixel array detector Pilatus-6MF. X-ray diffraction intensities were processed using iMOSFLM and them scaled with Scala, implemented in the CCP4 package. The molecular replacement step was performed in Phaser using PDB ID: 4HT3 as search model.

5.2.8 Microcrystalline Protein Samples for Solid-State NMR. Microcrystalline samples of βQ114A TS were prepared by diluting the enzyme solution 1:1 with 50 mM Cs^+ bicine buffer, pH 7.8, containing 20% PEG-8000 and 2.0 mM spermine as previously described (B. G. Caulkins et al., 2014). Yellow colored microcrystals were collected and washed with 50 mM Cs^+-bicine buffer, pH 7.8, containing 10% PEG-8000, 1 mM spermine, and 2 mM F9. The crystals were packed at 10 000 g into a Bruker 4 mm magic-angle spinning (MAS) rotor with an approximate volume of 50 µL; each rotor contained 60-70 mg of protein. For making the $E(C_3)_{2AP}$ carbanion, L-Ser was introduced by direct addition of 3-5 µL of 1.2 M L-Ser to the packed MAS rotor, while 2AP was introduced by addition of 4 µL of a concentrated stock solution of 2AP (~250 mM) in acetonitrile.

5.2.9 X-ray Structure Determination. The water content varies considerably in TS structures. In this work, the Matthews coefficient (Matthews, 1968) with units of $Å^3$/Da and the water content are 2.60 (52.64% water content), 2.50 (50.83%), 2.64 (53.38%), 2.63 (53.30%) and 2.52 (51.24%) for crystal structures 6DZ4, 6DZO, 6D0V, 601H and

6C73 respectively. These values fall within the range typically observed for 70 kDa crystal structures (Kantardjieff & Rupp, 2003). During refinement, the 2Fo-Fc map was kept at counter level cut off of 1.00 σ during manual modeling and fitting. The crystals were prepared in buffer containing either sodium chloride or cesium chloride. Na^+ and Cs^+ bind to a well-described monovalent cation binding site within the β-subunit of TS and cation binding to this site is important for the biological functioning of TS (A. T. Dierkers et al., 2009). At the β subunit of TS, there is a monovalent cation site. Based on the structure factors and B-factor analysis, the electron densities at this site were easily distinguished from water, ammonium ion or chloride ion in the reported structures. The average water B-factor for each of the new crystal structures 6DZ4, 6DZO, 6D0V, 601H and 6C73 respectively are 38.8, 32.3, 35.5, 45.1, 32.5 and 34.5 $Å^2$. While modelling the electron density maps, care was taken to avoid interpreting noise as water molecules at the reported resolution.

Based on the validation report and refinement statistics as is shown in Table 5.1, the 0.2% disallowed residues for the Ramachandran plot belong to the residue αPhe212. This is due to the presence of the α-site ligand F9, which disturbs the phi/psi angles of the αPhe212. To our best of knowledge, all the crystal structures with F9 bound to the α-site reported in the pdb database have this feature in common.

Table 5.1. X-Ray Crystal Structure Data Collection and Refinement Statistics for New Complexes Presented in this Study (* Values in parentheses are for the highest-resolution shell.)					
	PDB ID 6DZ4	PDB ID 6DZO	PDB ID 6D0V	PDB ID 6O1H	PDB ID 6C73
Data collection					
Space group	*C* 1 2 1	*C* 1 2 1	*C* 1 2 1	*C* 1 2 1	*C* 1 2 1
Cell dimensions					
a, b, c (Å)	182.51, 59.30, 67.33	181.35, 58.38, 67.30	183.73, 60.88, 67.32	183.75, 60.72, 67.21	181.61, 58.98, 67.15
α, β, γ (deg)	90.00, 94.68, 90.00	90.00, 94.89, 90.00	90.00, 94.87, 90.00	90.00, 94.63, 90.00	90.00, 94.40, 90.00
Resolution (Å)	30.00−1.45	30.00−1.64	40.00−1.64	39.31−1.64	20.00−1.65
R_{pim}	0.039 (0.289)	0.032 (0.110)	0.055 (0.211)	0.031 (0.195)	0.039 (0.170)
I/σ(I)	9.9 (2.2)	16.1 (6.2)	10.5 (3.8)	11.5 (2.4)	15.1 (4.7)
$CC_{1/2}$	0.998 (0.859)	0.998 (0.985)	0.997 (0.943)	0.999 (0.951)	0.998 (0.961)
Completeness (%)	99.5 (98.1)	94.1 (90.2)	90.0 (83.6)	99.9 (99.5)	99.9 (99.6)
Multiplicity	5.6 (5.3)	3.6 (3.6)	3.6 (3.6)	3.5 (3.4)	3.5 (3.4)
Refinement					
Resolution (Å)	28.94−1.45	29.18−1.64	37.32−1.64	39.31−1.64	19.82−1.65
No. reflections	120118	76612	77328	85508	80889
R_{work} /R_{free}	0.14/0.18	0.17/0.19	0.13/0.17	0.15/0.19	0.14/0.18
No. atoms					
Protein	4922	4986	4996	4950	5005
Ligand/ion	80	99	168	109	149
Water	738	684	744	709	863
Ramachandran					
Favored (%)	96.3	95.8	95.8	96.6	96.6
Allowed (%)	3.6	4.0	4.0	3.2	3.2
Disallowed (%)	0	0.2	0.2	0.2	0.2
B factors					
Protein	21.2	19.9	20.2	27.6	15.9
Ligand/ion	39.8	31.8	31.5	32.2	27.6
Water	38.8	32.3	35.5	45.1	32.5
r.m.s. deviations					
Bond lengths (Å)	0.007	0.007	0.011	0.006	0.013
Bond angles (deg)	1.376	1.466	1.456	1.240	1.535

5.2.10 UV/Vis Studies: Static spectra and steady state reaction rates were measured using a Hewlett-Packard 8452A diode array spectrophotometer. Steady state reactions

comparing the turnover number of the mutant and the WT TS were determined using substrates indole and L-serine. Rate measurements were quantified using the difference in extinction coefficients between indole and L-Trp at 290 nm ($\Delta\varepsilon$ = 1.89 $mM^{-1}cm^{-1}$) (Creighton, 1970; Woehl & Dunn, 1999). The k_{cat} values for L-Ser were measured using standard methods (Lane & Kirschner, 1983). The UV/Vis plots were generated in Mathematica (Caprio, 2005; Wolfram Research, 2020). All the experiments were performed in at least triplicates.

5.2.11 UV/Vis Calculations. UV/Vis excitation energies and transition probabilities were calculated using the time-dependent density functional theory (TDDFT) (Jacquemin et al., 2011) in Gaussian 09 (M. Frisch et al., 2009). The calculations were performed on model cofactor complexes excised from the optimized 7Å clusters of the intermediate species in the active site of corresponding crystal structures. The clusters were constructed as described previously (B. G. Caulkins et al., 2016b; J. D. Hartman et al., 2015b; J. Lai et al., 2011). TDDFT calculations were carried out using the B3LYP functional and the 6-311++G(d,p) basis set (Fabian, 2010). The IEFPCM solvent model (Klamt et al., 2015; Tomasi et al., 1999) (dielectric constant = 10.0; Dichloroethane) was used to coarsely approximate the surrounding protein environment.

5.2.12 NMR Spectroscopy. ^{15}N cross-polarization (CP) magic-angle-spinning (MAS) experiments of catalytically active TS micro-crystals were performed at 9.4 T (400.37 MHz ^{1}H, 40.58 MHz ^{15}N) on a Bruker AVIII spectrometer equipped with a double resonance, 4 mm MAS probe, spinning at MAS rates of 8 KHz (^{15}N acquisition); the bearing gas was cooled to -10 ˚C. Cross-polarization was accomplished at a ^{1}H spin-lock field of 44.3 kHz and a ^{15}N spin-lock of 29.5 kHz (^{15}N) (ramped $\pm$ 5 kHz); 83 kHz Spinal64 ^{1}H decoupling was used throughout. ^{15}N spectra consist of the sum of 49,152

transients acquired with a relaxation delay of 4 s, for a total acquisition time of 55h 30m. ^{15}N chemical shifts were referenced indirectly to liquid-NH_3 (25 °C) via an external solid-state sample of NH_4Cl, in which the resonance frequency was set to 39.3 ppm (Harris et al., 2007). The acquisition of solid-state NMR spectra was interleaved with single pulse, low-power decoupling experiments (128 scans) reporting predominantly on free ligand and reaction products in solution (mother liquor). Acquisition of solid-state NMR spectra for the intermediate was halted before reactant concentrations in solution fell to zero.

5.2.13 RSSF and SWSF Spectroscopy

RSSF (Rapid Scanning Stopped Flow) measurements were determined using an instrument assembled from components consisting of either a Tungsten-Iodide (50 w 12 v) lamp or a 75 w Xe lamp, quartz fiber optics, a Hi-Tech rapid mixing unit, and a 512-element photodiode array (Andor) with programmable exposure and scan rates. Typically, 1000 scans per kinetic run were collected with scan rates ranging from 10 ms/scan to 10 s/scan. Cut-off and cut on filters were used to isolate transmission wavelength windows (typically of ~200 nm appropriate to the wavelength region of interest and to attenuate or block light of unwanted wavelengths. Wavelength resolution was 2-3 nm. Single-wavelength (SWSF) transient kinetic studies were performed using a Tungsten-Iodide (50 w 12 v) lamp, the monochromator and pre-amplifier from a Durrum Model D-I10 stopped-flow spectrophotometer chasse, quartz fiber optics, a Hi-Tech rapid mixing unit, and a photomultiplier (PM) tube. Signals from the PM tube were captured and digitized with a high-speed A/D converter, converted to absorbance units and stored with customized software. Time courses were fitted using the Levenberg-Marquardt algorithm utilizing the following general equation,

$$A = A_\infty \pm \sum A_n \exp\left(\frac{-t}{\tau_n}\right) \qquad \text{Eq 5.1}$$

where A and A_∞ are the absorbance values at time t and at time infinity, respectively, and A_n is the amplitude of the nth relaxation, τ_n. All spectral analyses were performed by generating raw ASCII files from the RSSF and HP8425A instrumentation for processing using Mathematica 11.0 and plots were generated using custom Mathematica codes.

5.3 Results

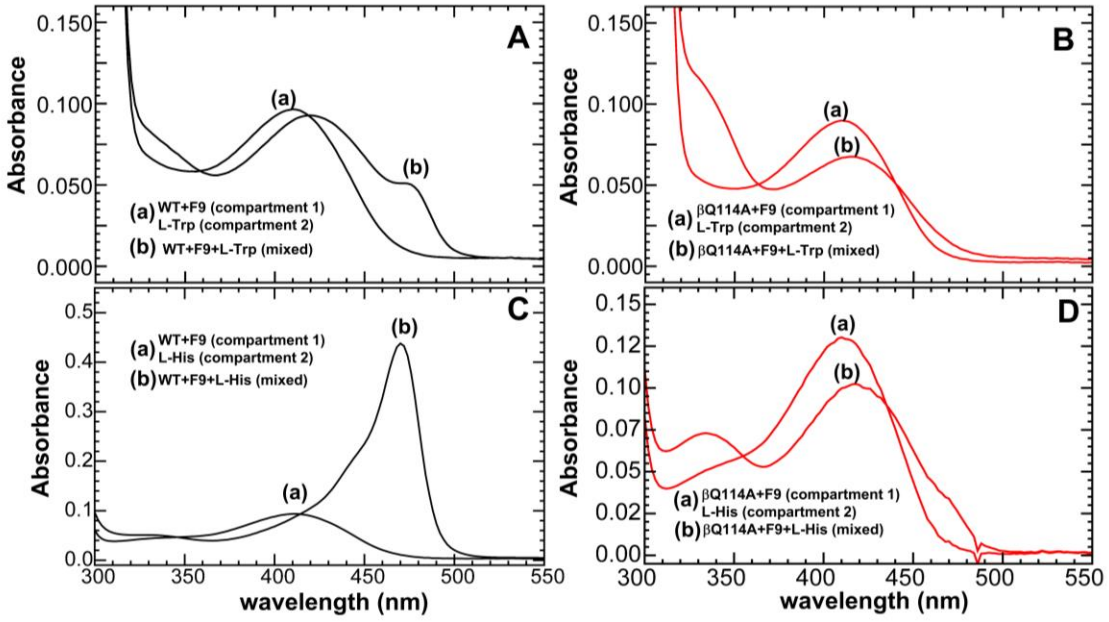


Fig 5.1: Reactions of WT and βQ114A TS with L-Trp (panels A and B) and with L-His (panels C and D) respectively. A double-difference quartz cuvette was used to acquire the spectra. In each panel, the solutions in the two compartments of the cuvette are designated as follows: (a), the spectra before mixing. (b), Spectra following mixing. Color code: WT (black), βQ114A (red). All the reactions were carried out in the presence of F9 (1 mM), with [WT TS], = 17 μM and [βQ114A TS] = 17 μM (Panel B) and 23.5 μM (Panel D); when added, [L-Trp] = 50 mM and [L-His] = 200 mM.

5.3.1 Static UV/Vis Spectral Changes in Substrate and Substrate Analog Reactions with TS.

The complexes formed when the internal aldimines of WT (A. T. Dierkers et al., 2009; Houben et al., 1989) and βQ114A TS react with L-Trp or L-His give remarkably different spectra as shown in Fig 5.1. In each panel, (a) and (b) designate spectra before and

after mixing, respectively. Spectrum (a) is the spectrum of the internal aldimine of either WT (Panels A and C) or βQ114A (Panels B and D); there is almost no difference in the internal aldimine spectral band (λ_{max} = 412 nm) of the PLP cofactor for the WT (B. G. Caulkins et al., 2014; Miles, 1979; Ngo, Kimmich, et al., 2007; Yanofsky & Crawford, 1972) and mutant. In Panel A, spectrum (b) shows that upon mixing with L-Trp the E^T(Ain) 412 nm absorption band is replaced by bands at 424 nm and 476 nm and a shoulder at ~ 330 nm, previously assigned to $E^T(Aex_2)_{L\text{-}Trp}$, $E^R(C_3)_{L\text{-}Trp}$ and $E^T(GD_2)_{L\text{-}Trp}$ respectively (Drewe Jr & Dunn, 1986; Ngo, Kimmich, et al., 2007; M. Roy et al., 1988). Panel B spectrum (b) shows that upon mixing with L-Trp, the 412 nm absorption band of βQ114A is replaced by a spectral band that is broadened and shifted to ~ 420 nm and a new band appears with a λ_{max} at 330 - 340 nm. The product mixture likely consists of $E^T(Aex_2)_{L\text{-}Trp}$, E(Ain)(L-Trp), $E^T(GD_2)_{L\text{-}Trp}$, and trace amounts of $E^R(C_3)_{L\text{-}Trp}$ at approximately 476 nm. Panel C spectrum (b) shows that reaction of WT TS with L-His gives the carbanionic species, $E^R(C_3)_{His}$ (λ_{max} 470 nm, shoulder ~ 440 nm) as the major component (A. T. Dierkers et al., 2009; Houben et al., 1989). However, as found for L-Trp, reaction of L-His with βQ114A (Panel D) gives a very different spectrum exhibiting bands at ~ 330 nm (shoulder), ~ 422 nm and ~ 470 nm.

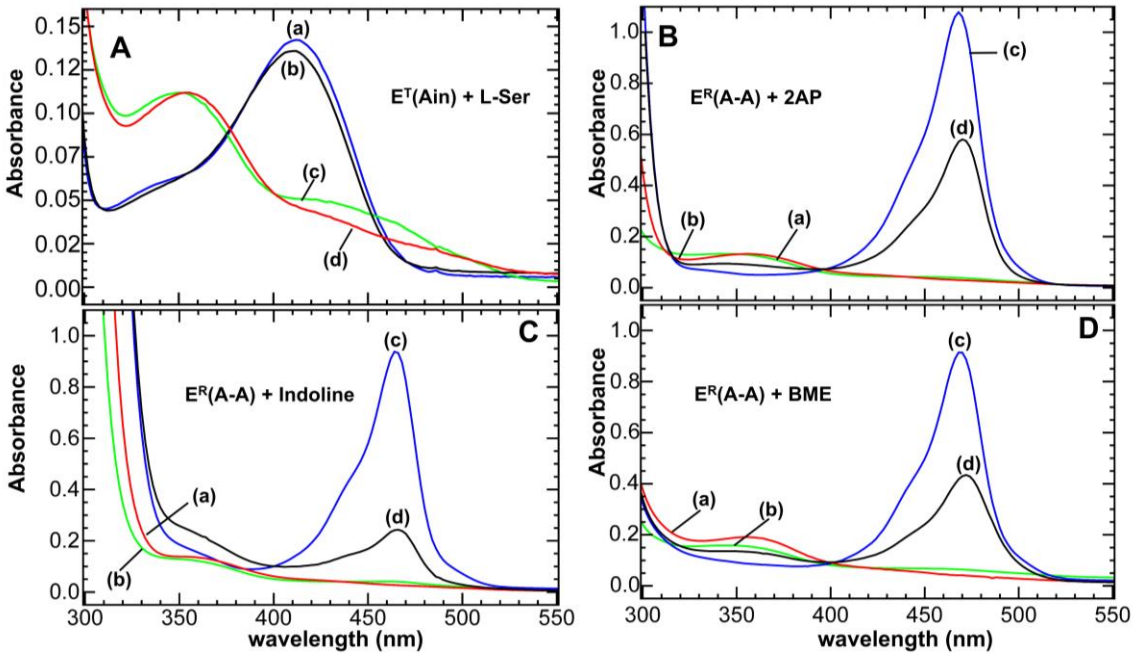


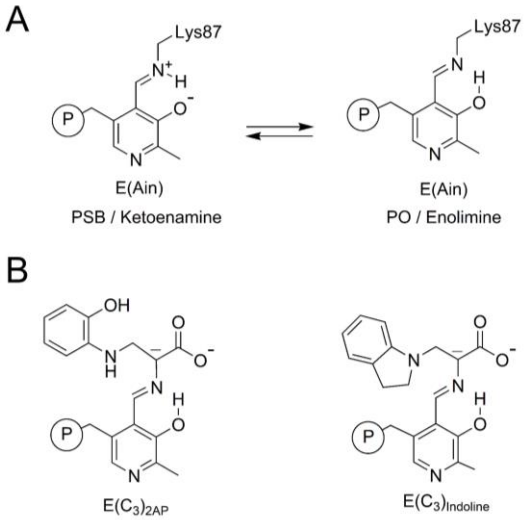
Fig 5.2: Static UV/Vis spectra of substrate and substrate analog reaction products of WT and βQ114A TS. Panel (A): reaction of the WT (a) (blue) and βQ114A mutant (b) (black) internal aldimines with L-Ser to give product mixtures dominated by the corresponding α-aminoacrylate species WT (c) (green) and mutant (d) (red). Panels (B-C): reaction of the WT (a) and βQ114A mutant (b) α-aminoacrylate reaction mixtures with the nucleophiles 2AP (Panel B); indoline (Panel C); and β-mercaptoethanol (BME) (Panel D). In each panel, the reaction product spectra for the WT enzyme are colored blue and designated by (c), while the product spectra for the βQ114A mutant are colored black and designated by (d). All the spectra were acquired using a double-difference quartz cuvette. Final concentrations: WT $\alpha_2\beta_2$, 15 μM; Q114A $\alpha_2\beta_2$, 15 μM; F9, 1 mM; when present, L-Ser 50 mM; 2AP, 4 mM; indoline, 25 mM; and BME, 100 mM.

Fig 5.2A shows the spectra for the reaction of both the WT and βQ114A mutant internal aldimines (spectra (a, b)) with substrate L-Ser (product spectra (c, d)); this reaction gives mixtures of species dominated by the α-aminoacrylate intermediate, E^R(A-A)(Drewe Jr & Dunn, 1985; Woehl & Dunn, 1999). The spectra for the internal aldimines establish that the absorption bands which characterize these species are only slightly altered by the βQ114A mutation. When reacted with L-Ser, the WT enzyme (spectrum c) shows two E^R(A-A) tautomers, one with a peak at 352 nm the other with a peak at ~ 460 nm, and trace amounts of $E^T(Aex_1)$ with a peak at 422 nm (Drewe Jr & Dunn, 1985; Hur et al.,

2002; Ngo, Kimmich, et al., 2007). Whereas the βQ114A, spectrum (d), is comprised of species corresponding to an α-aminoacrylate tautomer that absorbs mostly at ~354 nm, red shifted by ~ 2 nm, and species absorbing at ~ 460 nm. There is very little accumulation of the ~ 460 nm E^R(A-A) tautomer in the mutant enzyme.

5.3.2 Reaction of Nucleophiles (Nu) with E^R(A-A). Reaction of indole with WT E^R(A-A) in Stage II of the β-reaction gives a transiently formed intermediate absorbing at 476 nm that has been identified as the L-Trp carbanionic species, $E^R(C_3)_{L\text{-}Trp}$ (Drewe Jr & Dunn, 1986; Lane & Kirschner, 1983). Fig 5.2, panels B, C, and D compare the spectral changes observed in the reactions of nucleophilic analogs of indole with the WT and βQ114A E^R(A-A) to give quasi-stable carbanionic species (Barends, Domratcheva, et al., 2008; M. F. Dunn, 2012; Houben et al., 1989; J. Lai et al., 2011; Mozzarelli, Peracchi, Rovegno, Dalè, et al., 2000). These analogs react with WT E^R(A-A) to give intense UV/Vis spectral bands with λ_{max} 465-480 nm and ε_{max}~ 40,000 $M^{-1}cm^{-1}$. The $E^R(C)_{Nu}$ species (Nu = indole, 2AP, indoline, or β-mercaptoethanol) formed with βQ114A (spectra (d) in each panel) show λ_{max} red-shifted by ~ 2 nm. In the WT system, the indoline and 2AP quasi-stable carbanions slowly convert to the corresponding amino acid analogs of L-Trp (B. G. Caulkins et al., 2016b; M. F. Dunn et al., 2008; J. Lai et al., 2011), while the β-mercaptoethanol (BME) carbanion turns over relatively rapidly to the new amino acid(Ahmed et al., 1986). Assuming the extinction coefficients are similar for the WT TS and βQ114A $E^R(C)_{Nu}$, the spectra in Panels B-D show that at saturating nucleophile concentrations the yields of carbanionic species that accumulate with βQ114A are significantly less than those observed with the WT enzyme; thus, the steady-state yields of $E^R(C)_{Nu}$ with indoline, 2AP and BME are approximately 50 -60% of the amount obtained with WT TS. In addition to the $E^R(C)_{Nu}$ peaks at the longer wavelengths, the

βQ114A spectra in Panels B-D exhibit new features consisting of broad shoulders at 340-350 nm that we assign here as the $E^T(GD)_{Nu}$ species derived from the new amino acids derived from these nucleophiles (see Discussion).



Scheme 5.3. (A): The protonated Schiff base (PSB)/ketoenamine and phenolic (PO)/enolimine tautomeric states of E(Ain). (B) The structures of the $E(C_3)$ intermediate formed with indole analog 2AP and indoline in the tryptophan synthase catalytic cycle.

5.3.3 UV/Vis Simulations. TDDFT calculations were carried out to predict the absorption maxima and extinction coefficients of the intermediates in the TS catalytic cycle so that reasonable assignments of the experimental spectra to chemical structures could be made. This analysis assumes that the various species detected in Figs 5.1 and 5.2 correspond to the ketoenamine (or protonated Schiff base (PSB)) and enolimine (or phenolic (PO)) tautomers of the intermediates (Schemes 5.1 and 5.3). Table 5.2 shows there is a good correlation between TDDFT calculated UV/Vis spectra with the experimental values of λ_{max} and ε_{max}. For E(Ain), the experimental spectrum with λ_{max} of 412 nm corelates well with the calculated ~ 394 nm value for the PSB form of E^T(Ain)(B. G. Caulkins et al., 2014). For E^R(A-A), the experimental spectra show multiple potential species present, and the UV/Vis simulations show that the PSB tautomer can be

assigned to the more intense experimental transition at 447 nm, whereas the PO tautomer is expected to give rise to the smaller band at 362 nm. Reaction of the nucleophiles with E^R(A-A) results in carbanionic species with red shifted spectra and large extinction coefficients, which are successfully replicated in the TDDFT calculations (Table 5.2). The PSB tautomer of these intermediates gives calculated λ_{max} of ~ 460 nm while those for PO fall at λ_{max} ~ 430 nm, consistent with the observed experimental peaks at ~ 465-476 nm with shoulders at 440 nm. Finally, the calculated transition for the PSB form of E^T(GD) at 313 nm is consistent with the observed experimental transitions at 330 nm.

Table 5.2: Comparison of Observed and TDDFT Calculated Spectra for PSB and PO Tautomers				
PLP Chromophore	Observed $\lambda_{max}(\varepsilon_{max})$	Calculated PSB $\lambda_{max}^{\#}(\varepsilon_{max})^{\#}$	Calculated PO $\lambda_{max}^{\#}(\varepsilon_{max})^{\#}$	Experimental References
E^{T}(Ain)	412 nm (5,800)	394 nm (7000)	334 nm (5400)	(B. G. Caulkins et al., 2014; Drewe Jr & Dunn, 1986)
E^{R}(A-A)	352 nm (5,000) 460 nm (6,000)	447 nm (11,000)	362 nm (11,000)	(Woehl & Dunn, 1999; Woehl et al., 1996)
$E^{R}(C_3)_{2AP}$	468 nm (48,000) 440 nm (30,000) (Shoulder)	460 nm (28,000)	428 nm (35,000)	(B. G. Caulkins et al., 2016b)
$E^{R}(C_3)_{Indoline}$	466 nm (40,000) 440 nm (30,000) (Shoulder)	464 nm (28,000)	430 nm (32,000)	(A. T. Dierkers et al., 2009; Mozzarelli, Peracchi, Rovegno, Dalè, et al., 2000)
E^{T}(GD)	330 nm (~7000)	313 nm (~8500)	n.d.	(Karsten & Cook, 2009; C. M. Metzler et al., 1980; M. Roy et al., 1988)
#TDDFT calculations were carried out on model PLP cofactor complexes geometry optimized using a cluster model of the active site;(B. G. Caulkins et al., 2016b) the number of transitions was specified at 6. The reported λ_{max} and ε_{max} values are for the transitions with the highest absorption intensity.				

5.3.4 SSNMR Studies of E^{R}(A-A) and $E^{R}(C_3)_{2AP}$. Fig 5.3 shows ^{15}N solid-state NMR spectra for the reaction of WT and βQ114A TS microcrystals with (A,B) ^{15}N-L-Ser and (C,D) ^{15}N-L-Ser and 2AP to give the respective E^{R}(A-A) and $E^{R}(C_3)_{2AP}$ complexes. For the βQ114A mutant, the Schiff base nitrogen chemical shifts fall at 288.2 ppm and 300.5 ppm for E^{R}(A-A) and $E^{R}(C_3)_{2AP}$, respectively, only slightly perturbed from the corresponding chemical shifts of 286.0 and 298.6 in the WT TS complexes. In all cases, the ^{15}N chemical shift predicts that the PO population is the dominating tautomeric species under these experimental conditions.

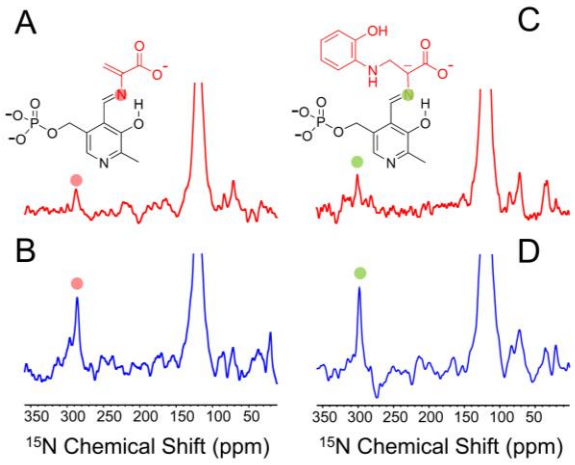


Fig 5.3: ^{15}N CPMAS spectra of the reaction of (A,B) E^R(A-A) and (C,D) $E^R(C_3)_{2AP}$ of WT (blue) and βQ114A (red) TS microcrystals with (A, B) 50-100 mM L-[^{15}N] Ser and (C, D) 50-100 mM L-[^{15}N] Ser and 8-10 mM 2AP in the presence of 3 mM F9 and 50 mM CsCl. Spectra acquired at 9.4 T and 9 kHz magic angle spinning (MAS) at -10 ˚C. Resonances assigned to the Schiff base nitrogen are indicated by the red and green dots for the E^R(A-A) and $E^R(C_3)_{2AP}$ intermediates and fall at approximately 288.2 ppm and 300.5 ppm, respectively, for the mutant, and 286.0 ppm and 298.6 ppm for WT. The differences in relative intensity reflect a true decrease in the amount of the intermediate formed.

5.3.5 X-ray Structures of WT and βQ114A Mutant TS Complexes. Figs 5.4 and 5.5 compare the crystal structures of WT and βQ114ATS with F9 bound to the α-sites and Cs^+ at the monovalent cation (MVC) binding site for the following complexes: E^T(Ain), $E^T(Aex_1)$, E^R(A-A), and $E^R(C_3)_{2AP}$ with PDB identifier codes for WT structures 4HT3 (D. Niks et al., 2013), 6DZ4, 4HN4 (D. Niks et al., 2013), and 4HPJ (D. Niks et al., 2013), respectively, and for βQ114A structures 6C73, 6DZO, 6D0V and 6O1H, respectively. All structures are of relatively high resolution (1.30 - 1.65Å). Unless mentioned otherwise, all structures with F9 bound to the α-site have the completely closed α-subunit conformation. Notice the E^T(Ain) and $E^T(Aex_1)$ structures have open β-subunit conformations, while the E^R(A-A) and $E^R(C_3)_{2AP}$ structures have the completely closed

conformation of the $\alpha_2\beta_2$ complex (Barends, Dunn, et al., 2008; A. R. Buller et al., 2015; M. F. Dunn, 2012; Ferrari et al., 2003) (Figs 5.4 and 5.5).

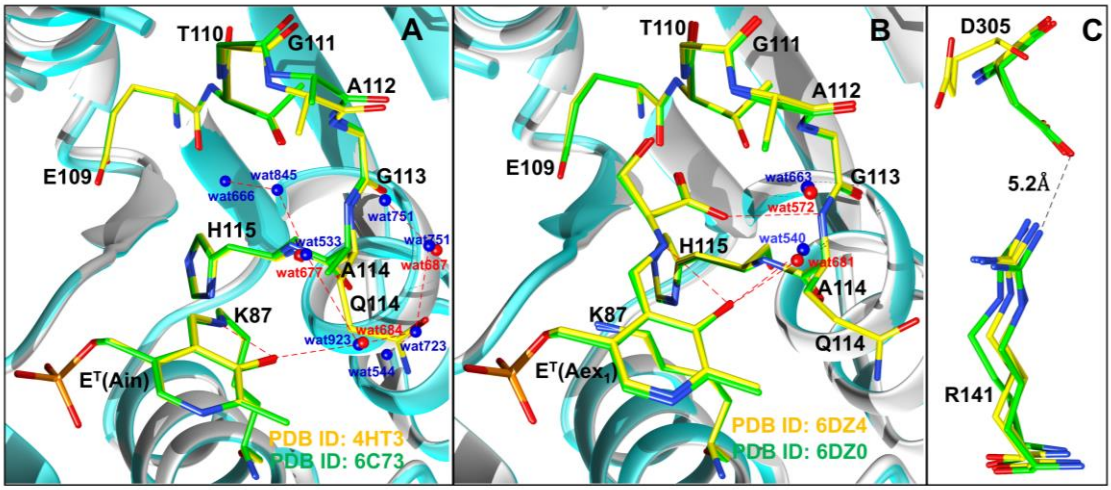


Fig 5.4: Comparison of WT (PDB ID: 4HT3(D. Niks et al., 2013), 6DZ4) and βQ114A (PDB ID: 6C73, 6DZO) internal aldimine and external aldimine complexes of the TS bienzyme complex. Coloring scheme: WT structures, yellow carbons and red waters; βQ114A structures, green carbons and blue waters. Ribbons for the WT structures are gray, ribbons for the βQ114A structures are teal. Panel (A): Active site comparisons of the WT and βQ114A internal aldimine forms. Panel (B): Active sites of the WT and βQ114A TS external aldimines. Panel (C): The βR141 and βD305 residues in the T state (open) complex of the β subunit are too distant to form an H-bonded salt bridge (M. F. Dunn, 2012; M. F. Dunn et al., 2008; Ferrari et al., 2003).

5.3.6 Comparison of the WT and βQ114A Structures. The X-ray crystal structures of the WT and βQ114A internal aldimine complexes (Fig 5.4) show differences in the β active site of the PLP cofactor. The two waters ($wat666_{\beta Q114A}$, $wat845_{\beta Q114A}$) adjacent to the PLP O3' oxygen, Schiff base nitrogen, and the βL3 residues in the βQ114A internal aldimine complex are not observable in the WT structure; these two waters are part of a hydrogen bonding network with $wat533_{\beta Q114A}$, which itself aligns with WT water $wat677_{WT}$. Water molecules $wat923_{\beta Q114A}$ and $wat684_{WT}$ also nearly overlap and both weakly hydrogen bond to the PLP O3'oxygen (~3.6Å). In βQ114A, three water molecules

occupy the cavity created by the Gln to Ala mutation. wat751$_{\beta Q114A}$ and wat687$_{WT}$ occupy the same positions in the crystal structure, but based on the electron density map, water molecule wat751$_{\beta Q114A}$ is split between two positions. In βQ114A, wat751$_{\beta Q114A}$ is strongly hydrogen bonded to wat723$_{\beta Q114A}$ (2.0Å), whereas wat687$_{WT}$ is the sole occupant of the space in the WT internal aldimine (which hydrogen bonds to the carboxamide group of βQ114).

When L-Ser reacts with E^T(Ain) to form E^T(Aex_1), water molecules in and near the active site of the enzyme are displaced (Fig 5.4). There is a water molecule in both the WT and βQ114A external aldimine complexes (wat540$_{\beta Q114A}$ and wat681$_{WT}$, respectively) hydrogen bonded to the PLP O3' oxygen. An additional water molecule in βQ114A occupies the space which was previously occupied by the βQ114 side chain. In both the internal aldimine and external aldimine complexes, the carboxylate group of βAsp305 is rotated away from βR141, a signature of the open conformation of the TS enzyme (Fig 5.4, Panel C). The comparison of the WT internal and external aldimines (PDB ID: 4HT3 (D. Niks et al., 2013) and PDB ID: 6DZO, respectively) using various methods for superimposing the two structures in Chimera (Pettersen et al., 2004a) and PyMol (Schrodinger, 2010) indicate that there are only small changes in conformation that accompany this chemical transformation. These changes appear to be on the order of 0.3 to 0.5 Å for COMM domain residues, while the carboxylate binding loop (residues 109 - 115) shows slightly larger changes (~1.0Å). There is little or no change in the remainder of the β-subunit.

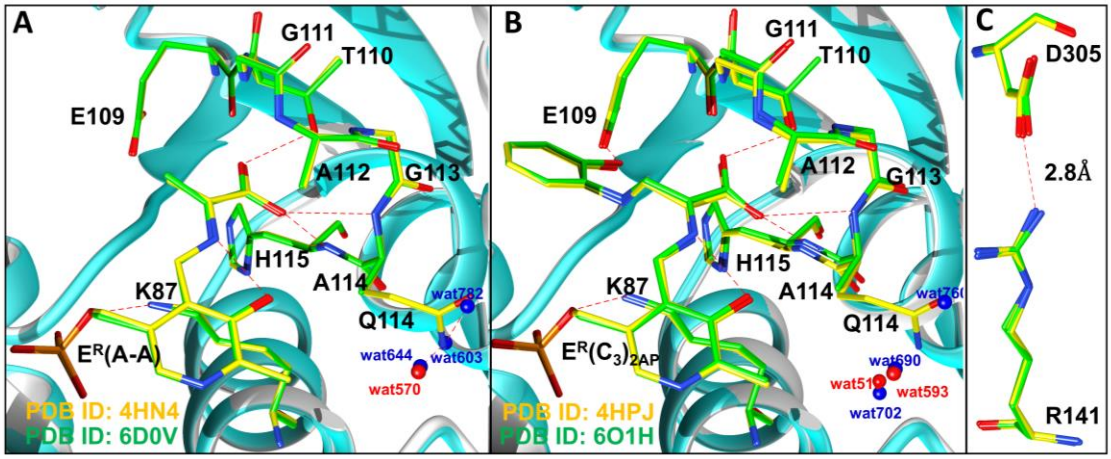


Fig 5.5: Comparison of E^R(A-A) (PDB ID: 4HN4(D. Niks et al., 2013), 6D0V) and $E^R(C_3)_{2AP}$ (PDB ID: 4HPJ(D. Niks et al., 2013), 6O1H) complexes of TS bienzyme complexes. Panel (A): Active sites of the WT and βQ114A α-aminoacrylate structures. Panel (B): Active sites of the WT and βQ114A $E(C_3)_{2AP}$ structures. Panel (C): Salt bridges between βR141 and βD305 a signature of the R state (closed) complex of the β-subunit. Coloring scheme: WT structures are represented in yellow with red waters while the βQ114A structures are represented in green with blue waters. Ribbons for the WT structure are shown in gray whereas teal ribbons are used to depict the βQ114A structures.

5.3.7 The E^R(A-A) and $E^R(C_3)_{2AP}$ Carbanion Complexes of Tryptophan Synthase. In Stage II of the β-reaction (Scheme 5.1), the enzyme transitions from E^R(A-A) to two subsequent L-Trp carbanion complexes, $E^R(C_2)$ and $E^R(C_3)$, before release of the product. The crystal structures of the WT and βQ114A mutant E^R(A-A) species are presented in Fig 5.5(A), and the two respective carbanion analogs of $E^R(C_3)$ formed by reaction of E^R(A-A) with 2AP are shown in Fig 5.5(B). The only significant structural changes arise from the differences among the covalent structures of the cofactor moieties and the replacement of Gln by Ala at β114 amino acid residue. The much larger differences between the open and closed conformations of the β-subunit in Figs 5.4 and 5.5 result from the motion of the COMM domain during the TS β-subunit allosteric transition (Barends, Domratcheva, et al., 2008; Barends, Dunn, et al., 2008). During the

conformation change, β-helix6 undergoes a one turn rotation (Barends, Dunn, et al., 2008) which closes the entrance into the α-site from solution, trapping the reacting substrates inside. At the same time, we anticipate that the side chain carboxylate OE1 of βE109, acting as a hydrogen bond acceptor, rotates to form an H-bond with the reacting substrate in $E^R(C_1)$; this conclusion is based on the observations that in the E^R(A-A)(BZI) and $E(C_3)_{2AP}$ structures an analogous H-bond forms with the hydroxyl group of 2AP (PDB IDs: 4HPX (D. Niks et al., 2013), 4HPJ (D. Niks et al., 2013)), and in the $E^R(C_3)_{indoline}$ structure the distance between the Glu109 E1 oxygen and the C3 carbon of indoline is 2.7 Å (PDB IDs: 3PR2 (J. Lai et al., 2011), 3CEP (Barends, Domratcheva, et al., 2008)). Furthermore, the βD305 carboxylate rotates into the correct position to make the salt bridge connection with βR141, a structural signature of the closed complex of the TS (Barends, Dunn, et al., 2008; M. F. Dunn et al., 2008; Maria-Solano et al., 2019). The Gln to Ala mutation results in the introduction of two water molecules ($wat603_{\beta Q114A}$ and $wat782_{\beta Q114A}$) and a single water molecule ($wat760_{\beta Q114A}$) in the α-aminoacrylate and 2AP carbanion complexes, respectively.

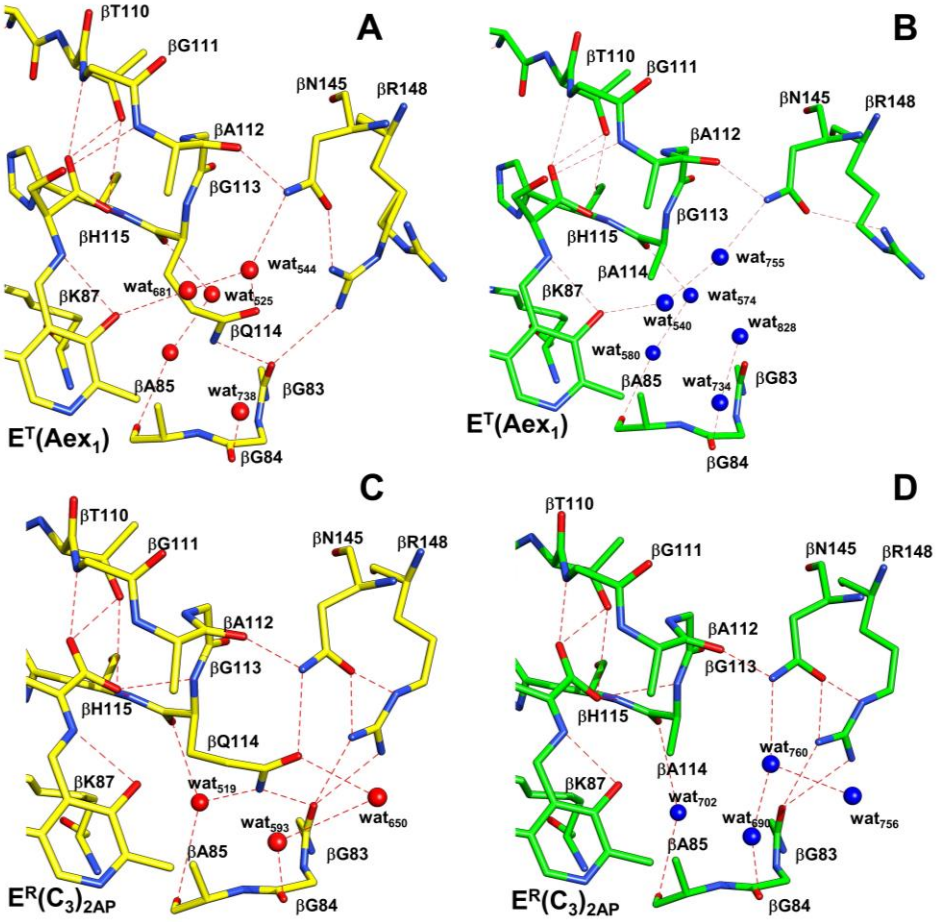


Fig 5.6: X-ray structural details of the T and R β-site conformations. Coloring scheme: WT structures, yellow carbons and red waters; βQ114A structures, green carbons and blue waters. Plausible H-bonding interactions between atoms are shown as red dashes. (A) and (B), $E^{T}(Aex_1)$ T (open) complexes of WT and βQ114A (PDB ID: 6DZ4, 6DZO respectively). (C) and (D), $E^{R}(C_3)_{2AP}$ R (closed) complexes of WT and βQ114A (PDB ID: 4HPJ (D. Niks et al., 2013), 6O1H). The Q to A mutation in T complexes (A, B) causes minor disruptions of the H-bonding network of waters and neighboring protein residues. The R complexes (C, D) show a larger disruption, in particular the H-bonding of βQ114 to βN145 and βR148 is lost. These changes are essentially the same in the closed structures of the E^{R}(A-A) complexes (PDB ID: 4HN4 (D. Niks et al., 2013), 6D0V), not shown.

5.3.8 Comparisons of the Open and Closed Complexes of Intermediates in WT and βQ114A TS. TS switches between open and closed conformations during catalysis (M. F. Dunn, 2012) (Figs 5.4-5.6). Comparisons of the WT and βQ114A structures (Fig 5.6)

shows that in the open conformations there appear to be almost no positional changes in β-subunit protein atoms due to the mutation, except for βR148. The βR148 side chain has dual conformations in the WT (Fig 5.6A), whereas there is only a single conformation of βR148 in the βQ114A $E^T(Aex_1)$ structure (Fig 5.6B). However, the comparison of the crystal structures of the closed complexes of WT and βQ114A TS shows that the mutation of βQ114 to Ala in loop 3 of the COMM domain eliminates three hydrogen bonding interactions that contribute to the stabilization of the closed complexes. As a result of the mutation, the H-bonding network between the βQ114 carboxamide side chain and residues βG83, βN145 and βR148 are disrupted. Whereas a new water molecule (wat760, compare Fig 5.6C, 5.6D) occupies the cavity created by the shortened side chain of Ala114 and becomes part of an H-bonded water lattice, that is not present in the WT enzyme.

Table 5.3: Steady state rates for the β-reaction of TS using Indole and L-Ser as substrates (Scheme 5.1B).

Enzyme	$K_{m(apparent)}$	k_{cat} (s^{-1})	$k_{cat}/K_{m(apparent)}$
WT (0.8 μM)	2.16 ± 0.26	4.23 ± 0.4	1.95
βQ114A (1.4 μM)	0.85 ± 0.13	0.67 ± 0.07	0.79

#[L-Ser] varied from 0 – 8 mM, while [Indole] kept fixed at 0.4 mM.

5.3.9 Steady State Kinetic Comparisons Between WT and βQ114A TS. Steady state turnover rates were measured for the β-reaction and show that βQ114A retains significant activity (Table 5.3). Nevertheless, the steady-state rate of L-Trp formation for the reaction of L-Ser with indole catalyzed by βQ114A is reduced by ~ 6.5-fold compared

to the WT TS, and the comparison of the steady state parameters ($K_{m(apparent)}$ and $k_{cat}/K_{m(apparent)}$) shows that the catalytic efficiency for L-Trp production is reduced by ~2.5 fold in the βQ114A mutant compared to the WT enzyme under the conditions used in this study (closed α-site and Cs^+ bound at the MVC in the β-subunit).

Table 5.4: $K_{d,app}$ for Reactions of L-Ser and Indole Analogs in WT and βQ114A TS. All the reactions were carried out in the presence of 1 mM F9. Concentrations: WT enzyme 20 μM[(a,c)] and 10 μM[(b)]; βQ114A 20 μM[(a,c)] or 30 μM[(b)].

Substrate(s)	WT $K_{d,app}$	βQ114A $K_{d,app}$	Equation for $K_{d,app}$ calculation
(1) L-Ser[(a)]	1.1 mM	0.27 mM	$E^T(\text{Ain}) + \text{L-Ser} = \text{E}^R(\text{A-A})$
(2) 2AP[#(b)]	20.8 μM	135.6 μM	$E^R(\text{A-A}) + \text{Nu} = \text{E}^R(C)_{\text{2AP}}$
(3) Indoline[# (b)]	5.2 mM	n.d.	$E^R(\text{A-A}) + \text{Nu} = \text{E}^R(C)_{\text{indoline}}$
(4) BME[# (b)]	6.4 mM	39.6 mM	$E^R(\text{A-A}) + \text{Nu} = \text{E}^R(C)_{\text{BME}}$
(5) L-His [(c)]	31.9 mM	n.d.	$E^T(\text{Ain}) + \text{L-His} = \text{E}^R(C)_{\text{L-His}}$

#Calculation of $K_{d,app}$ of nucleophiles in (2), (3) and (4) are performed under saturating [L-Ser] of 50 mM.

5.3.10 Mutation of Glutamine to Alanine at β114 Impairs the Synthesis of L-Tryptophan. Table 5.4 summarizes the apparent dissociation constants for the formation of E^R(A-A) and E^R(A-A) complexes with nucleophilic indole analogs in the presence of the α-site ligand F9. The $K_{d,app}$ estimates for these processes were calculated using the equations in Table 5.4.

As measured by the apparent $K_{d,app}$ values, the WT enzyme binds L-Ser less tightly than does βQ114A TS ($K_{d,app}$ = 0.27 mM for βQ114A compared to 1.1 mM for the WT enzyme). However, there are larger differences in the dissociation constants for formation of the carbanion complexes. The WT E^R(A-A) shows higher affinities for the nucleophiles, ~ 6.5-fold for 2AP and 5-fold for BME. A value for indoline binding to βQ114A could not be determined due to limitations imposed by the low solubility of indoline and extreme low yields of $E^R(C_3)_{indoline}$ in βQ114A TS (viz. Fig 5.2C). The dissociation constant for L-His binding to βQ114A also could not be determined due to the apparent low affinity for L-His and to solubility limitations (viz Fig 5.1D).

5.3.11 Stopped Flow Measurements

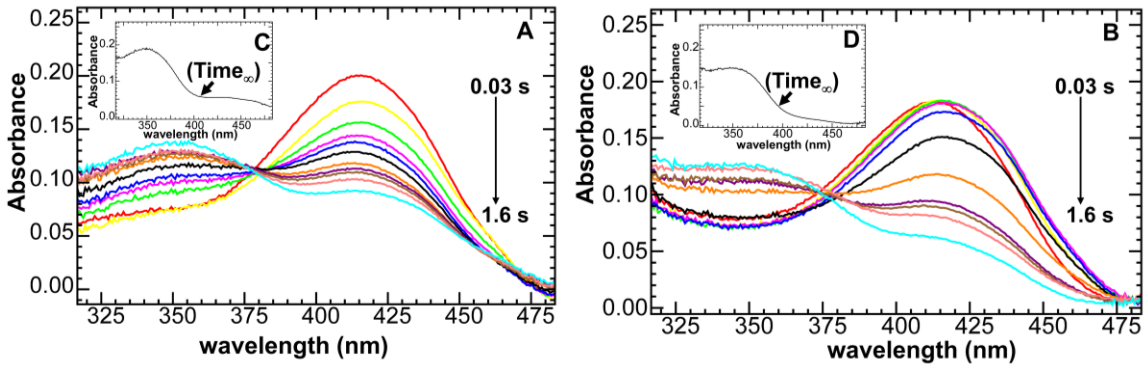


Fig 5.7: Rapid-scanning stopped-flow spectra showing the pre-steady state time courses for the conversion of L-Ser to the α-aminoacrylate intermediate during reaction with the internal aldimine forms of wild-type (A) and βQ114A mutant (B) TS. Insets (C) and (D) shows the time$_\infty$ traces for both wt and βQ114A spectra. The scans were collected at 0.0336 s (red), 0.0420 s (yellow), 0.0504 s (green), 0.0588 s (magenta), 0.0672 s (blue), 0.1092 s (black), 0.2436 s (orange), 0.4116 s (purple), 0.4956 s (brown), 0.8316 s (pink), 1.6716 s (cyan). Concentrations after mixing: [wild-type E(Ain)] = 37 μM, [βQ114A E(Ain)] = 33μM, [L-Ser] = 50 mM, [F9] = 1mM.

5.3.12 Stopped Flow Kinetic Analysis of Wild-Type and βQ114A TS Systems in the β-Reaction. The RSSF UV/Vis spectra shown in Fig 5.7 compare the reactions of L-Ser with wild-type and βQ114A mutant E(Ain) with the IGP analogue F9 bound to the α-site and Cs^+ bound to the monovalent cation site of the β-subunit. The spectral changes that accompany the L-Ser reaction previously have been described in detail for the wild-type enzyme (Drewe Jr & Dunn, 1985, 1986). The wild-type E(Ain) is characterized by an absorption band with λ_{max} 412 nm arising from the protonated nitrogen of the internal aldimine Schiff base linking the PLP cofactor to the ε-amino group of Lys 87. In wild-type TS, reaction with L-Ser results in the rapid formation (rate $1/\tau_1$) of a new spectral band with λ_{max} = 422 nm previously identified as the protonated SB N tautomer of the external

aldimine, $E(Aex)_1$). The 422 nm band subsequently decays in a triphasic process ($1/\tau_2 > 1/\tau_3 > 1/\tau_4$) to yield a spectrum characterized by a peak at 352 nm with a broad shoulder extending to 500 nm. This spectrum has been shown to be comprised of a steady state mixture dominated by the E(A-A) species with a deprotonated SB N (λ_{max} 352 nm) and traces of the protonated SB N (λ_{max}~460 nm) with contributions from small amounts of E(Ain) and $E(Aex)_1$ (respectively λ_{max} 412 nm and 422 nm).

The transient spectral changes that occur during the reaction of L-Ser with βQ114A are different. The mutant reacts more slowly and, in comparison to the wild-type system, the spectrum at steady state shows an altered final distribution of intermediates with a greater proportion of E(A-A) species with λ_{max} ~354 nm and an absorbance shoulder extending to 460 nm and beyond.

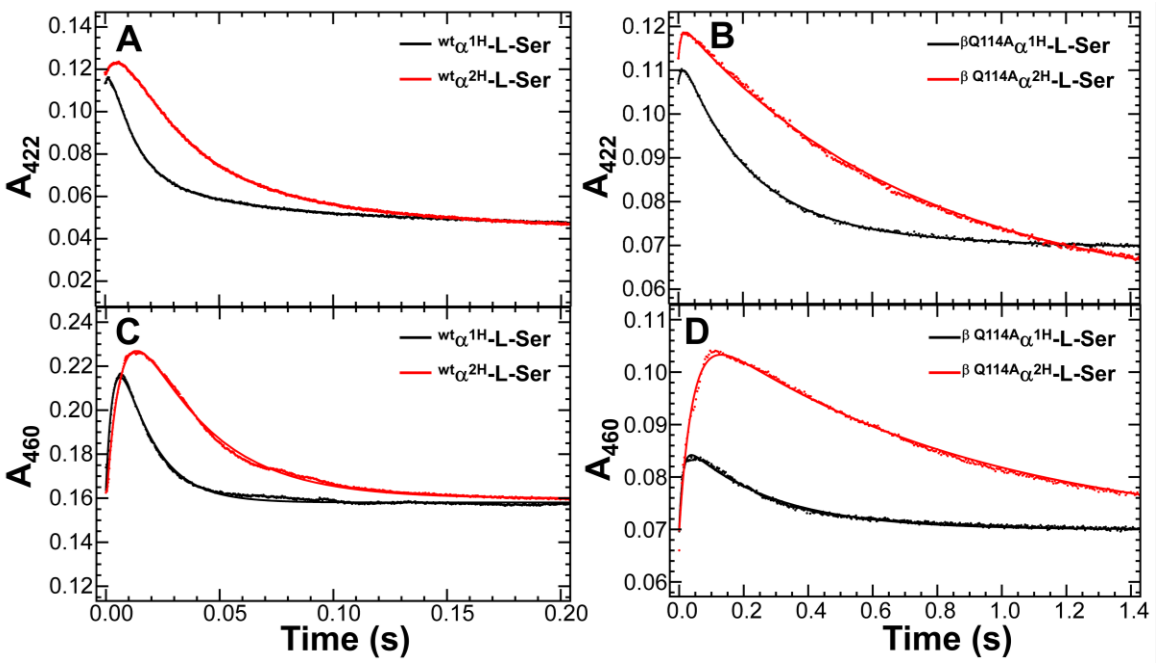
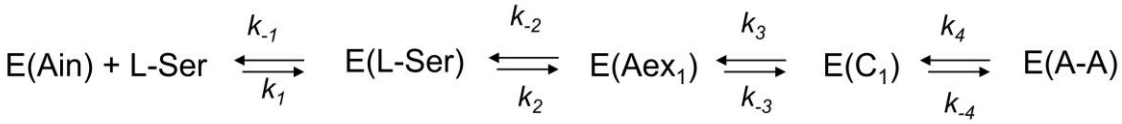


Fig 5.8: Single wavelength time courses at 422 and 460 nm for the reaction of $\alpha^{1H/2H}$L-Ser with wild-type (A, C) and βQ114A (B, D). The least-squares best fit of the data to a sum of exponentials is superimposed on each time course. Concentrations after mixing: [L-Ser] = 50 mM;[βQ114A] = 50 μM; [wild-type] = 35μM in presence of 1mM F9.

Table 5.5(A): Summary of the relaxation rates, Kinetic Isotope Effects (KIE) in Stage I of β reaction at 422 nm.

Enzyme	α^{1H} $1/\tau_1$ (s^{-1})	α^{1H} $1/\tau_2$ (s^{-1})	α^{1H} $1/\tau_3$ (s^{-1})	α^{1H} $1/\tau_4$ (s^{-1})	α^{2H} $1/\tau_1$ (s^{-1})	α^{2H} $1/\tau_2$ (s^{-1})	α^{2H} $1/\tau_3$ (s^{-1})	α^{2H} $1/\tau_4$ (s^{-1})	KIE (1H/2H) – $1/\tau_1$ ratio	KIE (1H/2H) – $1/\tau_2$ ratio	KIE (1H/2H) – $1/\tau_3$ ratio	KIE (1H/2H) – $1/\tau_4$ ratio
wild-type	n.d.	43.6± 6.4	3.6± 0.4	0.5± 0.03	n.d.	21.02± 3.01	1.2± 0.2	0.4± 0.08	-	2.07	3.0	1.2
βQ114A	132.5 ± 2.05	9.4± 0.01	0.5± 0.001	n.d.	65.6 ± 25.7	2.2 ± 0.4	0.3 ± 0.2	n.d.	2.02	4.3	1.6	



Scheme 5.4 Rate dependencies in Stage I of the β reaction.

Table 5.5(B): Summary of the relaxation rates, Kinetic Isotope Effects (KIE) in Stage I of β reaction at 460 nm.

Enzyme	α^{1H} $1/\tau_1$ (s^{-1})	α^{1H} $1/\tau_2$ (s^{-1})	α^{2H} $1/\tau_1$ (s^{-1})	α^{2H} $1/\tau_2$ (s^{-1})	KIE (1H/2H) – $1/\tau_1$ ratio	KIE (1H/2H) – $1/\tau_2$ ratio
wild-type	223.5 ± 56.6	46.2 ± 13.7	71.3 ± 4.8	35.02 ± 3.4	3.1	1.3
βQ114A	56.9± 2.5	7.1 ± 0.2	24.6 ± 2.3	1.16 ± 0.07	2.3	6.2

The SWSF time courses for the reaction of L-Ser with βQ114A measured at 422 nm are well described by a sum of three relaxations; $1/\tau_1$ is characterized by a rapid rate and an increasing absorbance with a tiny amplitude due to $E(Aex)_1$ formation, followed by two relaxations, $1/\tau_2 > 1/\tau_3$, both with decreasing absorbance as conversion of $E(Aex)_1$ to E(A-A) unfolds in Stage I of the β-reaction. Under the experimental conditions of Fig 7 and Table 3 the 422 nm time course for the wild-type enzyme is well fit by relaxation rates of $1/\tau_2 = 43.6\ s^{-1}$, $1/\tau_3 = 3.6\ s^{-1}$, and$1/\tau_4 = 0.5\ s^{-1}$. The relaxation rates determined for βQ114A are $1/\tau_1 = 132.5\ s^{-1}$, $1/\tau_2 = 9.3\ s^{-1}$, and $1/\tau_3 = 0.5\ s^{-1}$. Thus, the values of $1/\tau_2$ and $1/\tau_3$ are 4-, 6 - and 7.7-fold slower, respectively, than for the wild-type enzyme. In

the wild-type reaction time course, the formation of $E(Aex)_1$ from E(Ain) is very rapid and has a much smaller amplitude, prohibiting accurate fitting to an exponential. However, in the βQ114A system, the amplitude for this phase is increased and can be fitted to an increasing exponential function. To obtain more accurate fits of the relaxations for the formation of $E(Aex)_1$ from E(Ain), time courses were measured at 460 nm. At this wavelength only two relaxations are observed, one increasing and one decreasing. Under the conditions shown in Fig 8 Panels C and D, the wild-type enzyme gives relaxation rates for $1/\tau_1$ and $1/\tau_2$ of 223.5 s^{-1} and 46.2 s^{-1}, while βQ114A gives relaxation rates for $1/\tau_1$ and $1/\tau_2$ of 56.9 s^{-1} and 7.2 s^{-1}. Thus τ_1 and τ_2 for βQ114A are slowed by 3.9 - and 6.4-fold respectively.

Primary Kinetic Deuterium Isotope Effects in Stage I of the β-Reaction. As shown in Fig 5.8 (Panels A and B) and Table 3, both $1/\tau_2$ and $1/\tau_3$, are subject to apparent primary kinetic isotope effects when the C^{α} proton of L-Ser is substituted by deuterium. At 422 nm, the wild-type TS time course gives apparent KIEs of 2.0, 3.0 and 1.2 for $1/\tau_2$, $1/\tau_3$ and $1/\tau_4$. The βQ114A mutant gives time courses which are well fitted with three exponential phases with apparent KIEs of 2.0, 4.2 and 1.5 for $1/\tau_1$, $1/\tau_2$ and $1/\tau_3$ respectively.

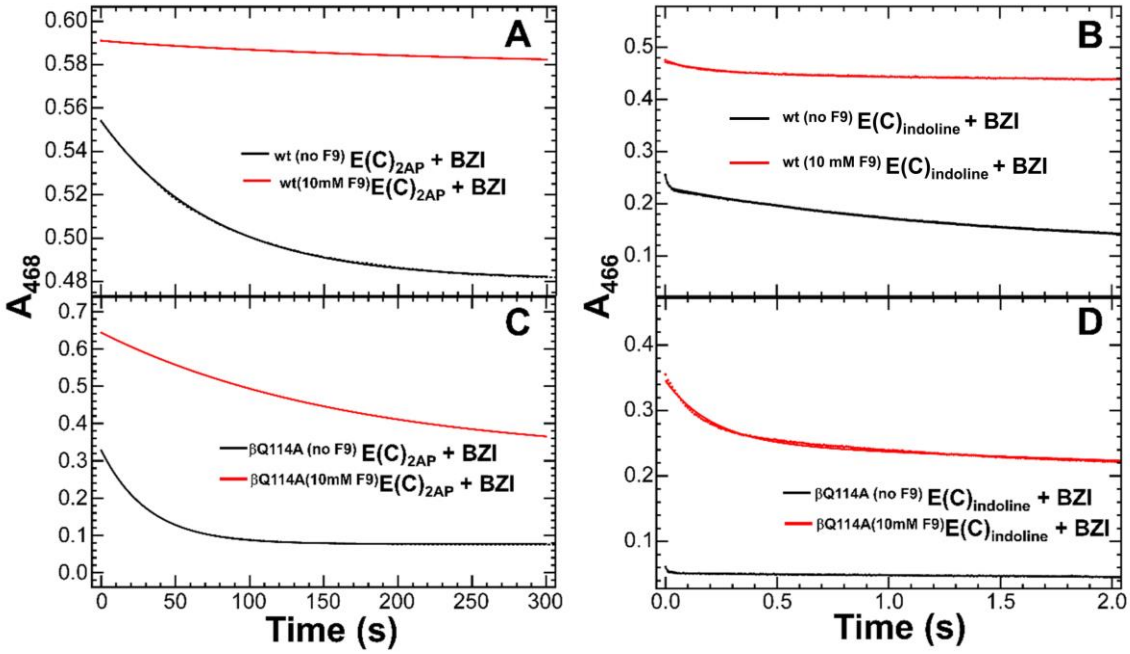


Fig 5.9. Time courses for BZI displacement of 2AP and indoline in (A, C) wild-type and (B, D) βQ114A TS in presence (red line) and absence (black line) of ASL F9. For the 2AP time courses in A,B. Concentrations: wt, 12 μM, βQ114A; 25 μM, 2AP 4 mM, Indoline, 25 mM, BZI, when present 2 mM, F9, when present, 10 mM.

Table 5.6 – Displacement of nucleophiles from $E(C)_{indoline}$ and $E(C)_{2AP}$ by BZI in wild type and βQ114A.					
Syringe 1	Syringe 2	$E(C)_{nu}$ decay $1/\tau_1$ (s^{-1})	$E(C)_{nu}$ decay $1/\tau_2$ (s^{-1})	ΔA_1	ΔA_2
$E(C\text{-wt})_{indoline}$+ no ASL	2 mM BZI	70.1 ± 11.9	0.7 ± 0.001	0.03 ± 0.0006	0.11 ± 0.0008
$E(C\text{-wt})_{indoline}$+ 10 mM F9	2 mM BZI	5.3 ± 0.4	0.2 ± 0.053	0.02 ± 0.002	0.034 ± 0.003
$E(C\text{-}\beta Q114A)_{indoline}$ + no ASL	2 mM BZI	150.9 ±17.2	0.3 ± 0.011	0.007 ±0.001	0.0125 ±0.0007
$E(C\text{-}\beta Q114A)_{indoline}$ + 10 mM F9	2 mM BZI	4.7 ± 0.1	0.1 ± 0.005	0.09 ±0.001	0.18 ±0.001
$E(C\text{-wt})_{2AP}$ + no ASL	2 mM BZI	0.09	n.d.	0.005	n.d.
$E(C\text{-wt})_{2AP}$ + 10 mM F9	2 mM BZI	0.07	0.004	0.0003	0.0139
$E(C\text{-}\beta Q114A)_{2AP}$ + no ASL	2 mM BZI	0.05	n.d.	0.09	n.d.
$E(C\text{-}\beta Q114A)_{2AP}$ + 10 mM F9	2 mM BZI	0.005	n.d.	0.3	n.d.

5.3.13 Benzimidiazole (BZI) Competes with Indole and Indole Analogues at the β-Site of the α-Aminoacrylate Intermediate. BZI is a potent reversible inhibitor of the formation of carbanionic species in the β-reaction (Melinda Roy et al., 1988). Because $E(C)_{Nu}$ formation is rapid and reversible, the binding of BZI causes a relatively rapid displacement of nucleophiles from the $E(C)_{Nu}$ complexes. The time course for the BZI displacement of indoline from wild-type $E(C)_{Indoline}$ is biphasic with ($1/\tau_1$ = 5.3 $s^{-1}$and $1/\tau_2$ = 0.3 s^{-1}). The results of these kinetic and spectroscopic studies establish that BZI also displaces indoline from the βQ114A $E(C)_{Indoline}$ complex via the same mechanism in a time course that also is biphasic ($1/\tau_1$ = 4.7 s^{-1} and $1/\tau_2$ = 0.1s^{-1}). Investigation of the reaction of 2AP with E(A-A) and the displacement of 2AP from the $E(C)_{2AP}$ complex by BZI shows that 2AP binds significantly more tightly than does indoline to E(A-A) (Fig 5.9, Table 5.6). The X-ray structures of the $E(C)_{Indoline}$ and $E(C)_{2AP}$ complexes show that the aromatic rings of indoline and 2AP occupy the same space and make similar van der Waals contacts within the subsite. However, the 2AP hydroxyl H-bonds to one

carboxylate oxygen of βGlu109 (Fig 4(b)), while the 2AP amino group forms an N-C covalent bond to C-β of E(A-A) and H-bonds to the βLys87 nitrogen. Indoline also reacts to make the covalent bond to C-β but lacks the H-bonding interaction with βGlu109. Thus, the greater affinity of 2AP likely results from the two additional H-bonding interactions. The time course for BZI displacement of 2AP from wild-type enzyme is a biphasic process with $1/\tau_1 = 0.067\ s^{-1}$ and $1/\tau_2 = 0.004\ s^{-1}$. BZI displaces 2AP from the βQ114A $E(C)_{2AP}$ complex, with a single relaxation, $1/\tau_1 = 0.005\ s^{-1}$.

The βQ114A mutation causes subtle differences in the properties of the carbanions formed in the reactions of indoline and 2AP with the E(A-A) species. 2AP binds more tightly ($^{2AP}K_{d,app}$ = 135.8 μM) than indoline. No saturation in the indoline binding isotherm was obtained even at high concentrations of indoline at its solubility limit in aqueous solutions. Furthermore, the static UV/Vis spectra of the two carbanions species is more stable than the $E(C)_{Indoline}$ species. Both nucleophiles, show shoulders at 340-350 nm possibly appearing due to the E(GD) species derived from the new amino acids formed with these nucleophiles or from PO tautomer of the carbanions.

5.4 Discussion

5.4.1 The βGln114 Side Chain is Important to the Functioning of Tryptophan Synthase. In βQ114A, the mutation reverses the relative ground stabilities of carbanionic, external aldimine and gem diamine species. Thus, the static spectra presented in Fig 5.1 reveal that carbanionic species derived either from L-Trp or analog L-His are destabilized by the Gln to Ala mutation, and the spectra presented in Fig 5.2 demonstrate that nucleophilic analogs of indole (indoline, 2AP and BME) are significantly less effective at forming carbanionic species via reaction with the E^R(A-A) βQ114A intermediate. From the UV/Vis spectral investigations, at ligand concentrations that saturate binding, only 40-60% of the β-sites are converted to $E^R(C)_{Nu}$. The remainder of the β-sites consist of species absorbing in the 330-380 nm region. Accordingly, based on the assignment of electronic transitions for PLP chemical intermediates and the calculated ε_{max} and λ_{max} values presented in Table 5.2, we assign the electronic transitions to an equilibrating mixture of $E^R(C)_{Nu}$ and $E^T(GD)_{Nu}$ species and possibly an inactive Michaelis complex of the E^R(A-A) with nucleophile/substrate (Figs 5.1 and 5.2, Table 5.2). The spectra presented in Fig 5.2, (panel A) also show that the mutation destabilizes the neutral SB N tautomer of E^R(A-A), causing a decrease in the UV/Vis spectrum at ~460 nm. However, the relative distributions of the carbanion tautomers (Fig 5.2, panel (B-D)) are unaffected by the mutation, hence the shape of the spectral envelope at long wavelengths (~470 nm peak, ~440 nm shoulder) due to the carbanion electronic transitions is essentially unchanged, while the yields of the carbanion species are significantly decreased in the mutant.

Fig 5.3 shows the ^{15}N chemical shift of the SB N in the βQ114A E^R(A-A) and $E^R(C_3)_{2AP}$ species is shifted 2.2 and 1.9 ppm to higher frequency relative to the WT ^{15}N chemical

shifts for the same intermediates. The chemical shift of the SB N has in previous work been interpreted as a population-weighted, fast-exchange average between the PSB and PO tautomers (B. G. Caulkins et al., 2016b) (Scheme (5.3)). Accordingly, these SSNMR data reveal that in the βQ114A E(A-A) and $E(C_3)_{2AP}$ complexes, as in the WT TS, the more stable tautomeric species is the PO intermediate, but it is in fast-exchange with the PSB protonated Schiff base form with a population of 10%-20%. We speculate that the switching between the PSB and PO tautomers of the cofactor make possible a finetuning of the reaction specificity. This tautomeric shift, then, is guided by the presence of the nucleophile (indole or analog) and the allosteric structural changes taking place in the active site of the enzyme. It is expected that formation of the PSB is aided by hydrogen bonding interactions to the phenolic oxygen that lower its pK_a and stabilize the phenolate form (C. M. Metzler et al., 1980). Although MD simulations have shown that the side chain of βQ114 is able to rotate and hydrogen bond to the PLP O3' (Huang et al., 2016), the perturbations in the solid-state NMR ^{15}N chemical shifts indicate negligible change in the relative populations of the PSB and PO tautomers, suggesting that this mechanism likely plays no substantial role in establishing the protonation states on the reacting E^R(A-A) and $E^R(C_3)_{2AP}$ cofactor-substrate complexes.

5.4.2 Correlations Between the Static Solution Spectra and Theoretical UV/Vis Spectra. The UV/Vis spectra presented in Figs 5.1 and 5.2 provide important information about the chemical structures and physio-chemical environments of the reacting species along the β-reaction path. When models based on crystal structures are used, our TDDFT-based approach to the calculation of UV/Vis spectra for the PLP chromophores gives results that correlate well with the experimentally observed spectra. Thus, we are able to predict the absorption energies and transition probabilities for

various protonation states of those PLP-derivatives for which there are crystal structures. The reaction of L-Ser with E^T(Ain) gives a complicated spectrum with λ_{max} at 352 nm and a broad shoulder at ~ 460 nm previously proposed to consist of the PO (352 nm) and PSB (~ 460 nm shoulder) tautomers of E^R(A-A) along with minor species (Drewe Jr & Dunn, 1985, 1986). This finding agrees with the UV/Vis calculations, which predict λ_{max} = 362 nm for the PO species and λ_{max}= 447 nm for the Zwitterionic PSB species. The TDDFT calculations also predict distinct UV/Vis transitions with λ_{max} at ~ 460 nm and ~ 430 nm, respectively, for the PSB and PO populations for both the $E^R(C_3)_{2AP}$ and $E^R(C_3)_{indoline}$ carbanions. This prediction is in good agreement with the observed λ_{max} for $E^R(C_3)_{2AP}$ at 468 nm with a ~ 440 nm shoulder, and the observed λ_{max} for $E^R(C_3)_{indoline}$ at 466 nm with a ~ 440 nm shoulder (Fig 5.2, Table 5.2). These results indicate that the PSB and PO tautomers for the α-aminoacrylate and the carbanion intermediates of TS are both present in appreciable quantities, and, based on the NMR results above, undergoing exchange that is fast on the NMR time scale, a finding consistent with the H-bonded structures seen in the X-ray structures of these species (Figs 5.4-5.6).

Reactions of L-Trp and L-His with WT and βQ114A TS result in an equilibrating mixture of intermediates that includes the external aldimine and carbanionic species, plus a third species that absorbs at ~330 nm. The latter has been tentatively assigned to the gem-diamine intermediate (Karsten & Cook, 2009; M. Roy et al., 1988). TDDFT calculations on a model gem-diamine complex provide support for this assignment, and computationally predicted UV/Vis spectra for the gem-diamine model compound predict a transition at 313 nm, consistent with the observed spectral transition.

5.4.3 UV/Vis and SSNMR Spectra of the Carbanionic Species. The static absorption spectra of the carbanions formed in the reactions of indole and indole analogs (indoline,

2AP and BME) with the E(A-A) species, and the carbanions formed in reactions of L-Trp and L-His with E(Ain) show the most striking differences between WT and βQ114A TS (Figs 5.1 and 5.2). In solution, Both L-His (A. T. Dierkers et al., 2009; Houben & Dunn, 1990; Houben et al., 1989) and L-Trp (Drewe Jr et al., 1989; Phillips & Harris, 2021) react with WT TS to give significantly greater populations of carbanionic species (λ_{max} 470 nm and 476 nm, respectively) than they do with the βQ114A mutant. Instead, the mutant gives species with λ_{max} ~ 330 nm (Fig 5.1), which are assigned to $E^T(GD_2)_{L\text{-}Trp}$ and $E^T(GD_2)_{L\text{-}His}$ as discussed above. Thus, reaction of the WT E^T(Ain) F9 complex with L-His gives spectra that are dominated by $E^R(C_3)_{L\text{-}His}$ (A. T. Dierkers et al., 2009), (Fig 5.1C, black traces), whereas, the βQ114A spectrum consists of equilibrating mixtures dominated by $E^T(Aex_2)_{L\text{-}His}$ (λ_{max} 422 nm) and $E^T(GD_2)_{L\text{-}His}$ (λ_{max} 330 nm) with only a small amount of $E^R(C_3)_{L\text{-}His}$ (λ_{max} 470 nm) (Fig 5.1D, red traces). Although WT TS gives a carbanionic species when reacted with L-Trp (Fig 5.1A), βQ114A gives only a trace amount of $E^R(C_3)_{L\text{-}Trp}$ (Fig 5.1B). Consequently, the mutation shifts the relative ground state stabilities of these species in favor of E^T(Aex) and E^T(GD) rather than E^R(C).

Close inspection of the UV/Vis absorption spectrum of WT $E^R(C_3)_{2AP}$ (Fig 5.2B, trace (c)) suggests that most, perhaps nearly all, of the β-sites have been converted to carbanionic species, consistent with the absence of electronic transitions in the 315 - 380 nm region arising from the PLP chromophore. The SSNMR analysis of $E^R(C_3)_{2AP}$ shows that in the crystalline state at -10°C, the intermediate is best modeled as a rapidly equilibrating mixture of two tautomeric forms consisting of PO (81%) and PSB (19%) species (B. G. Caulkins et al., 2016b). A similar trend is followed in the $E^R(C_3)_{2AP}$ formed by the

βQ114A, where the chemical shift of 300.5 ppm for the Schiff base N is also consistent with a mixture of PO and PSB species (Scheme 5.3).

Owing to the similarity of the SSNMR spectra of the carbanions presented herein we propose that the UV/Vis spectra are dominated by carbanionic structures with a neutral (deprotonated) PLP N1 (B. G. Caulkins et al., 2016b) and a mixture of PO (~ 440 nm) and PSB tautomers (~ 460-470 nm) in aqueous solution. The carbanionic species formed with βQ114A (Fig 5.2) appear to have the same protonation states as WT carbanion species. However, the yields of carbanion species in the reactions of nucleophiles with the E(A-A) species are typically no more than 50% of the sites; the remainder appear to be PLP species exhibiting one or more absorbance bands at ~ 350 nm (Fig 5.2). This is consistent with the lower absolute intensities observed in the SSNMR experiments for the carbanionic intermediates in the mutant compared to WT (Fig 5.3).

5.4.4 βQ114A Structural Changes. The striking effects of the βQ114A mutation on the reactions of L-Ser, indole, L-Trp and the α-site substrate analogs (Figs 5.1 and 5.2) establish that βQ114 has heretofore unsuspected roles in the structure and function of the TS β-subunit. Comparison of WT and mutant X-ray crystal structures (Figs 5.4-5.6) show there are small structural perturbations in the vicinity of the β114 side chain caused by the Gln to Ala point mutation. These small changes are localized to the vicinity of the mutated residue and occur in response to the structural changes resulting from the chemical transformation and include the incorporation of water molecules into the cavity created by the mutation. Although these structural changes are small, the relative magnitudes of the activation energies for catalytic steps in the β-reaction are significantly altered by the mutation. One consequence is that the β-reaction turnover

rate is reduced by ~ 6.5-fold, another is that the apparent affinities of nucleophilic analogs for the E(A-A) complex are significantly decreased (Table 5.4).

5.4.5 βQ114A and Comparison of other Mutants in TS. The formation of a salt bridge between the side chains of Arg141 and Asp305 is a clear signature of the conformational switch of the β-subunit to the closed state in WT TS. While there are no x-ray structures for the R141A or D305A mutants, it seems likely that these mutations also alter the relative stabilities of the open and closed states by destroying the salt bridge in the closed conformation (Ferrari et al., 2003). Indeed, these mutations were constructed to test this prediction. Unlike the βD305A mutant, our preliminary UV/Vis spectroscopic data suggests that the βQ114A TS is not activated towards the formation of the unnatural amino acid dihydro-iso-L-tryptophan (DIT) from the reaction of the indole analog indoline with the E(A-A).

5.4.6 PLP O3'-Water Hydrogen Bonds in TS Complexes. The structures presented in Fig 5.4 for E^T(Ain) and $E^T(Aex_1)$ complexes of WT and βQ114A TS show that some waters within the site are hydrogen bonded to the PLP O3' oxygen. These hydrogen bonding interactions are expected to stabilize the PSB tautomer. The UV/Vis spectra and ^{15}N and ^{13}C chemical shifts measured via SSNMR establish that the E^T(Ain) complexes are predominantly protonated at the Schiff base nitrogen (B. G. Caulkins et al., 2014). Protonation of the Schiff base nitrogen is anticipated to activate both E^T(Ain) for the nucleophilic attack of the L-Ser α-NH_2 at the PLP C4' carbon to give $E^T(GD_1)$, and for proton abstraction from the C-α of $E^T(Aex_1)$ to give $E^R(C_1)$ (Scheme 5.1B). The X-ray structures presented in Fig 5.4 establish that the mutation of Gln to Ala at β114 does not interfere with PLP O3'-water hydrogen bonding and therefore should exert little effect on the stability of the PSB tautomer in these intermediates. In contrast to the E^T(Ain) and

$E^T(Aex_1)$ complexes, the structures of WT and mutant E^R(A-A) complexes (Fig 5.5) indicate that the side chain β-carbons of β114 and βA112 shield the PLP O3' oxygen from hydrogen bonding with water molecules, and the UV/Vis spectra (Fig 5.1) and ^{15}N chemical shifts measured via SSNMR establish that the predominating tautomer of E^R(A-A) is the PO form.

5.4.6 Implications of the Neutral State of E^R(A-A) in TS. We speculate that, owing to the presence of a significant pool of L-Ser in bacterial cells, it is likely that *in vivo* TS resides in an inactive (resting) state as the PO tautomer of the α-aminoacrylate intermediate. This inactive form prevents nucleophilic attack of the α-amino N of βLys87 at the PLP C4' carbon and the ensuing side reaction of α(A-A) with E^T(Ain) (Blumenstein et al., 2007), as detailed in Scheme 5.2. Consequently, TS evolved to stabilize the PO tautomer of E^R(A-A), thus preventing this deleterious side reaction by rendering infrequent the H-bonding interactions that stabilize the PSB form. While it is tempting based on analogy to other PLP enzymes to suggest that βQ114 might transiently provide such stabilizing interactions for a given species, the tautomeric equilibrium is not affected by the mutation of βQ114 to alanine. Nevertheless, the UV/Vis λ_{max} and tautomer ratios for E^R(C) species are significantly perturbed by the βQ114N mutation (Blumenstein et al., 2007).

5.4.7 Implications for Allosteric Regulation and Energies of Intermediates in WT and βQ114A TS. While the crystal structures presented in Figs 5.4-5.6 establish that the βQ114A mutation does not significantly alter the conformations of the T and R structures, the UV/Vis spectra shown in Figs 5.1 and 5.2 dramatically establish that the mutation changes the relative distributions of intermediates/analogs. Thus, the relative stabilities of the T and R conformational states are altered by the βQ114A mutation. The

UV/Vis spectra presented in Fig 5.1 establish that reaction of L-His with E^T(Ain) gives an equilibrating mixture of E^T(GD), E^T(Aex) and E^R(C) species. The TS structural database shows that E^T(Aex) complexes have the T subunit conformation, while E^R(C) have the R conformation. Thus, while WT TS stabilizes the R conformation of the carbanionic complexes, E^R(C), the βQ114A TS mutant destabilizes the R conformation relative to the gem-diamine and external aldimine complexes, E^T(GD) and E^T(Aex). Fig 5.1 indicates L-Trp behaves similarly.

5.4.8 Implications from the RSSF Spectra and SWSF Time Courses. The RSSF spectral changes and difference spectra and the single wavelength time courses presented in Figs 5.7 - 5.8, define the nature of the UV/Vis absorbance changes that occur in the reactions catalyzed by wild-type (Fig 5.2 A,B) and βQ114A TS (Fig 5.2 C,D).The time resolved spectral changes for both show a rapid conversion of E(Ain) to E(Aex_1) species followed by multiphasic transformations to quasi-stable mixtures dominated by the deprotonated E(A-A) species. Drew and Dunn, correctly identified the transient species detected in Stage I as the intermediates depicts Scheme 5.1. While rates and amplitudes of the RSSF time course are altered by the βQ114A mutation, the mutant appears to follow the steps outlined in Scheme 5.1. What seems surprising is that the Gln to Ala mutation, situated 6-8 Å distant from the loci of covalent changes during catalysis, exerts substantial changes in catalysis while causing only minor changes in the 3D structure of the protein.

5.4.9 Implications of the Kinetic Isotope Effects in Stage I of the β-Reaction. The time courses measured at 460 nm for each of the observed relaxations in both the wild-type and βQ114A TS reactions with L-Ser show small kinetic isotope effects (KIEs) in each kinetic phase. The observation of a KIE in each of the relaxations (Fig 5.7 A, B

Table 5.5A,B) is rationalized by inspection of the mathematical expressions for the relaxation equations (Bernasconi, 2012). According to relaxation kinetic theory, in the general case, the relaxations describing linked equilibrating processes are complex expressions comprised of the specific rate constants for the forward and reverse rate constants of all the intermediate chemical steps involved in the interconversion of reactants and products. In special cases, a multi-step process such as Stage I of the β-reaction, will simplify if the rate of disappearance of one or more intermediates is much faster than the rate of formation. Thus, only when the C-H bond scission step is much slower than the other rate processes, will the observed KIE approach the value of the intrinsic KIE. In the more general case, if the rate of cleavage of the C-H bond in Stage I of the β-reaction is comparable to the rates of the other steps, then the observed relaxation rates will each exhibit a KIE that is smaller than the intrinsic KIE because the mathematical expressions for each relaxation contain terms with the specific rate constants for the C-H bond scission step together with terms for the other rate processes. Thus, the KIEs observed for βQ114A and wild-type enzymes are smaller than the expected intrinsic KIEs, pointing to a mechanism where, along with scission of the C^{α}-H-bond, additional steps are involved in formation of E(A-A) that contribute to rate limitation.

5.4.10 Implications of the RSSF Time Courses for BZI Displacement Reactions. In the TS αβ-reaction (Scheme 5.1), the switching between open and closed conformations of the $\alpha_2\beta_2$ complex is essential for the synchronization of the catalytic activities of the α- and β-sites and the channeling of indole. To determine whether or not the Gln to Ala mutation perturbs this allosteric transition, the kinetics of the displacement of Nu: from $E(C)_{Indoline}$ and $E(C)_{2AP}$ were examined (Fig 5.9). The binding of the IGP analogue F9 to

the α-site and reaction of L-Ser at the β-site to give E(A-A) switches TS to the completely closed conformation of $\alpha_2\beta_2$. With F9 bound, $E(C)_{Nu}$ complexes retain the closed conformation (viz. Figs 5.3 and 5.4). In the absence of F9, displacement of Nu: species from $E(C)_{Nu}$ is relatively rapid (Rodney M Harris et al., 2005). When F9 is bound displacement of Nu: becomes very slow, a finding consistent with the conclusion that the allosteric switching mechanism remains intact in βQ114A.

The binding and reaction of Nu: with E(A-A) to give $E(C)_{Nu}$ is rapid and reversible. BZI also binds rapidly and reversibly to E(A-A) and competes for the same subsite where Nu: species bind and react with E(A-A).

$$E(C)_{Nu} \rightleftarrows E(A\text{-}A) + Nu: \qquad \text{Eq 5.2}$$

$$E(A\text{-}A) + BZI \rightleftarrows E(A\text{-}A)(BZI) \qquad \text{Eq 5.3}$$

Therefore, provided the chemical potential of BZI is sufficient to overcome the stability of the $E(C)_{Nu}$ species reaction of BZI with $E(C)_{Nu}$ causes a net displacement of Nu:. While the equations for displacement (Eq. 5.2 and 5.3) are those for a simple competitive process, the molecular pathway for the binding and exchange of ligands at the indole subsite of the β-subunit has been shown to occur as follows:

When the α-subunit has the open conformation, BZI enters the β-site via the α-site and the tunnel unimpeded.

When the α-site has the closed conformation, then entry of BZI into the tunnel via the α-site is blocked and the rate of displacement of Nu: from $E(C)_{Nu}$ is limited by the rate of switching between closed and open conformation states of the β-subunit.

As is shown in Fig 5.9, BZI is able to displace indoline and 2AP from their respective $E(C)_{Nu}$ complexes in both the wild-type and βQ114A TS systems. However, the displacement rates for βQ114A are considerably greater compared to the wild-type system (Table 5.6), indicating BZI binds with lower affinity to the mutant E(A-A) than to the wild-type enzyme.

5.4.11 Conformational States in *St*TS β-subunit and *Pf*TrpB: Comparison of the COMM Domain Positions

A careful alignment of the *St*TS crystal structures 4HT3 (E(Ain)), 6DZ4 ($E(Aex_1)$), and 4HN4 (E(A-A)) in Chimera shows that except for small changes introduced by substrate bonding with the PLP cofactor there is no significant displacement of the COMM domain when E(Ain) is converted to $E(Aex_1)$; both conformations are designated as open (T). However, upon formation of the α-aminoacrylate, there is a sizable conformational change, with a displacement of the COMM domain of ~3.0 Å as the subunit adopts the closed (R) conformation. Analysis of the engineered *Pf*TrpB structures 5DVZ (E(Ain)) and 5DW0 ($E(Aex_1)$) shows there is a significant displacement of the COMM domain (~2.0 Å) in the chemical transformation of E(Ain) to $E(Aex_1)$, with the latter designated as a partially closed (R') structure (A. R. Buller et al., 2015; Maria-Solano et al., 2019). The interconversion of *Pf*TrpB $E(Aex_1)$ to E(A-A) (PDB ID: 5VM5) results in the formation of a closed structure that is consistent with the closed β-subunit structure in *St*TS.

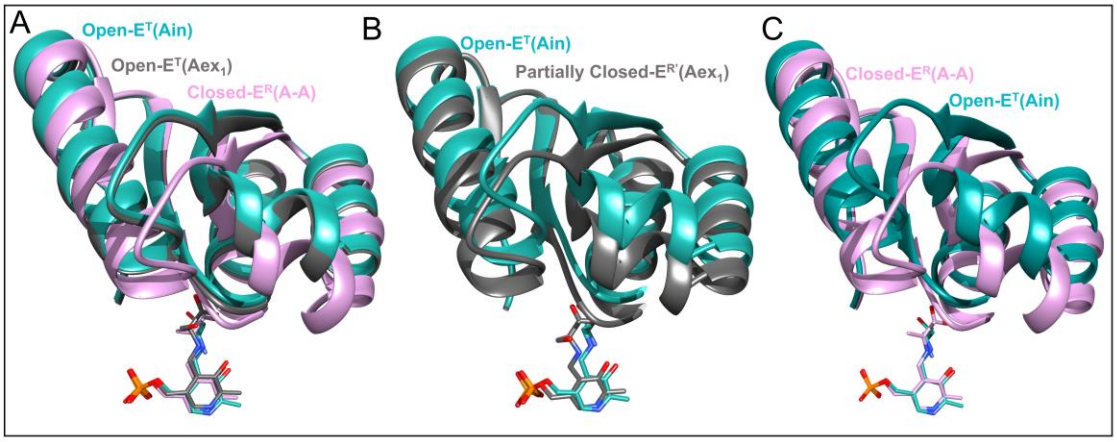


Fig 5.10: Comparisons of the β-subunit COMM domain and PLP cofactor positions for aligned crystal structures of *St*TS and *Pf*TrpB. (A) *Sf*TS: Open (T) complexes E^T(Ain) (teal) and $E^T(Aex_1)$ (grey), and the closed (R) complex E^R(A-A) (pink). (B) *Pf*TrpB: Open (T) complex E^T(Ain) (teal) and partially closed (R') complex $E^{R'}(Aex_1)$ (grey). (C) *Pf*TrpB: Open (T) complex E^T(Ain) (teal) and closed (R) complex E^R(A-A) (pink). Coloring Scheme: Ribbons and carbons are colored as: teal-blue in the E(Ain) structures, dark grey in the E(Aex_1) structures, and pink in E(A-A) structures. Stick structures: Oxygen, nitrogen, and phosphorous atoms are colored red, blue and orange respectively and carbons retain the colors of the corresponding ribbons. Images rendered in Chimera.

Previous work has established that the switching between open and closed conformation states is essential to the allosteric regulation of substrate channeling in the α β-catalytic cycle (M. F. Dunn, 2012; Michael F. Dunn et al., 1990; M. F. Dunn et al., 2008; Ngo, Kimmich, et al., 2007; D. Niks et al., 2013). The β-reaction catalytic cycle begins with the binding of L-Ser to the open conformation of the β-subunit. Reaction ensues via the T conformations of the $E^T(GD_1)$ and E^T(Aex) complexes (Fig 5.4). As the conversion of $E^T(Aex_1)$ to $E^R(C_1)$ occurs, the β-subunit switches to the R conformation (Figs 5.5 and 5.6) and remains closed as E^R(A-A), $E^R(C_2)$ and $E^R(C_3)$ are formed. Then the conversion to $E^T(Aex_2)$ switches the β-subunit back to the T state as $E^T(Aex_2)$ is formed and converts to $E^T(GD_2)$. The functional consequences of switching from the T to the R state includes the binding and release of substrates and product, activation of the α- and β-sites, the prevention of the escape of the channeled indole intermediate, and

synchronization of the α- and β-reactions (Harris & Dunn, 2002; R. M. Harris et al., 2005).

The allosteric interactions responsible for switching the β-site between T and R conformations have their origins in the weak bonding and covalent bonding changes that constitute the chemical transformations during the catalytic cycle. To modulate function via allosteric interactions in this system, the interconverting chemical-conformational ground states must be thermodynamically comparable. Thus, the switch to the R state conformation is dependent on interactions within the α β-dimeric unit triggered by the reacting substrate that slightly favor the R state. These studies show that the interactions between the βQ114 carboxamide side chain with βGly83, βN145 and βR148 are important in stabilizing the R conformation and that loss of these H-bonding interactions (Fig 5.6) is sufficient to shift the stability of the β-subunit conformation to the T state such that the energies of the E^T(GD), E^T(Aex) and E^R(C) species of βQ114A become comparable (Figs 5.1, 5.2, 5.4-5.6).

The hydrogen bonding interaction between the βQ114N side chain carboxamide group and the PLP O3' indicates the deleterious side reaction takes place in the TS system as a result of a shift in the distribution species from PO to PSB in E^R(A-A) (Scheme 5.2), thus facilitating the nucleophilic attack of the βLys87 at the PLP C4' carbon and the subsequent release of α(A-A). This interaction is disrupted in βQ114A TS, just as it is in the WT enzyme with the Q114 carboxamide sidechain rotated away from the PLP O3' (Figs 5.4-5.6). This structural paradigm is a key step in conferring reaction specificity in TS.

5.5 Conclusion

Previous investigations established that the mutation of βQ114 to Asn facilitates an unusual side reaction when TS reacts with substrate L-Ser. This side reaction generates α-aminoacrylate which then inactivates the enzyme. The detailed investigations of the Gln to Ala mutation at β114 presented here reveal that weak bonding interactions between the WT βQ114 side chain carboxamide and other elements of the β-subunit COMM domain are critically important for maintaining a balance between the energies of the open and closed conformations of the β subunit during the catalytic cycle. We conclude that the mutation of βQ114 to Ala impairs the catalytic behavior of TS by disrupting the H-bonding network between βQ114, βN145 and βR148. Furthermore, this disruption does not alter either the global folding of the β subunit, the α-β subunit contacts, or the bonding interactions between PLP and the β-site. Rather, the βQ114A mutant alters the relative stabilities of the open and closed conformations of the β subunit, such that the open conformations of E^T(GD) and E^T(Aex) complexes become more stable than the closed conformation of the E^R(C) species while protecting TS from undergoing the deleterious side reaction induced by the Gln to Asn mutation.

5.6 Accession Codes

WT TS E(Aex_1): 6DZ4; βQ114A TS E(Ain): 6C73; βQ114A TS E(Aex_1): 6DZO; βQ114A TS E(A-A): 6D0V; βQ114A TS E$(C_3)_{2AP}$: 6O1H.

5.7 References

Ahmed, S. A., Martin, B., & Miles, E. W. (1986). β-Elimination of Indole from L-Tryptophan Catalyzed by Bacterial Tryptophan Synthase: A Comparison between Reactions Catalyzed by Tryptophanase and Tryptophan Synthase. *Biochemistry*, *25*(15), 4233-4240. https://doi.org/DOI 10.1021/bi00363a010

Barends, T. R., Domratcheva, T., Kulik, V., Blumenstein, L., Niks, D., Dunn, M. F., & Schlichting, I. (2008). Structure and Mechanistic Implications of a Tryptophan Synthase Quinonoid Intermediate. *Chembiochem*, *9*(7), 1024-1028. https://doi.org/10.1002/cbic.200700703

Barends, T. R., Dunn, M. F., & Schlichting, I. (2008). Tryptophan synthase, an allosteric molecular factory. *Curr Opin Chem Biol*, *12*(5), 593-600. https://doi.org/10.1016/j.cbpa.2008.07.011

Bernasconi, C. (2012). *Relaxation kinetics*. Elsevier.

Blumenstein, L., Domratcheva, T., Niks, D., Ngo, H., Seidel, R., Dunn, M. F., & Schlichting, I. (2007). βQ114N and βT110V mutations reveal a critically important role of the substrate α-carboxylate site in the reaction specificity of tryptophan synthase. *Biochemistry*, *46*(49), 14100-14116. https://doi.org/10.1021/bi7008568

Brzovic, P. S., Ngo, K., & Dunn, M. F. (1992). Allosteric Interactions Coordinate Catalytic Activity between Successive Metabolic Enzymes in the Tryptophan Synthase Bienzyme Complex. *Biochemistry*, *31*(15), 3831-3839. https://doi.org/DOI 10.1021/bi00130a014

Buller, A. R., Brinkmann-Chen, S., Romney, D. K., Herger, M., Murciano-Calles, J., & Arnold, F. H. (2015). Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proc Natl Acad Sci U S A*, *112*(47), 14599-14604. https://doi.org/10.1073/pnas.1516401112

Caprio, M. A. (2005). LevelScheme: A level scheme drawing and scientific figure preparation system for Mathematica. *Computer Physics Communications*, *171*(2), 107-118. https://doi.org/10.1016/j.cpc.2005.04.010

Caulkins, B. G., Bastin, B., Yang, C., Neubauer, T. J., Young, R. P., Hilario, E., Huang, Y. M., Chang, C. E., Fan, L., Dunn, M. F., Marsella, M. J., & Mueller, L. J. (2014). Protonation states of the tryptophan synthase internal aldimine active site from solid-state NMR spectroscopy: direct observation of the protonated Schiff base linkage to pyridoxal-5'-phosphate. *J Am Chem Soc*, *136*(37), 12824-12827. https://doi.org/10.1021/ja506267d

Caulkins, B. G., Young, R. P., Kudla, R. A., Yang, C., Bittbauer, T. J., Bastin, B., Hilario, E., Fan, L., Marsella, M. J., Dunn, M. F., & Mueller, L. J. (2016). NMR Crystallography of a Carbanionic Intermediate in Tryptophan Synthase: Chemical

Structure, Tautomerization, and Reaction Specificity. *J Am Chem Soc*, *138*(46), 15214-15226. https://doi.org/10.1021/jacs.6b08937

Chan-Huot, M., Dos, A., Zander, R., Sharif, S., Tolstoy, P. M., Compton, S., Fogle, E., Toney, M. D., Shenderovich, I., & Denisov, G. S. (2013). NMR studies of protonation and hydrogen bond states of internal aldimines of pyridoxal 5′-phosphate acid–base in alanine racemase, aspartate aminotransferase, and poly-l-lysine. *Journal of the American Chemical Society*, *135*(48), 18160-18175.

Creighton, T. E. (1970). A steady-state kinetic investigation of the reaction mechanism of the tryptophan synthetase of Escherichia coli. *Eur J Biochem*, *13*(1), 1-10. https://doi.org/10.1111/j.1432-1033.1970.tb00892.x

Dierkers, A. T., Niks, D., Schlichting, I., & Dunn, M. F. (2009). Tryptophan synthase: structure and function of the monovalent cation site. *Biochemistry*, *48*(46), 10997-11010. https://doi.org/10.1021/bi9008374

Drewe Jr, W. F., & Dunn, M. F. (1985). Detection and identification of intermediates in the reaction of L-serine with Escherichia coli tryptophan synthase via rapid-scanning ultraviolet-visible spectroscopy. *Biochemistry*, *24*(15), 3977-3987.

Drewe Jr, W. F., & Dunn, M. F. (1986). Characterization of the reaction of L-serine and indole with Escherichia coli tryptophan synthase via rapid-scanning ultraviolet-visible spectroscopy. *Biochemistry*, *25*(9), 2494-2501.

Drewe Jr, W. F., Koerber, S. C., & Dunn, M. F. (1989). Application of rapid-scanning, stopped-flow spectroscopy to the characterization of intermediates formed in the reactions of l-and d-tryptophan and β-mercaptoethanol with Escherichia coli tryptophan synthase. *Biochimie*, *71*(4), 509-519.

Dunn, M. F. (2012). Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Arch Biochem Biophys*, *519*(2), 154-166. https://doi.org/10.1016/j.abb.2012.01.016

Dunn, M. F., Aguilar, V., Brzovic, P., Drewe Jr, W. F., Houben, K. F., Leja, C. A., & Roy, M. (1990). The tryptophan synthase bienzyme complex transfers indole between the α-and β-sites via a 25-30 Å long tunnel. *Biochemistry*, *29*(37), 8598-8607.

Dunn, M. F., Niks, D., Ngo, H., Barends, T. R., & Schlichting, I. (2008). Tryptophan synthase: the workings of a channeling nanomachine. *Trends Biochem Sci*, *33*(6), 254-264. https://doi.org/10.1016/j.tibs.2008.04.008

Fabian, J. (2010). TDDFT-calculations of Vis/NIR absorbing compounds. *Dyes and Pigments*, *84*(1), 36-53. https://doi.org/10.1016/j.dyepig.2009.06.008

Ferrari, D., Niks, D., Yang, L. H., Miles, E. W., & Dunn, M. F. (2003). Allosteric communication in the tryptophan synthase bienzyme complex: Roles of the β-

subunit aspartate 305-arginine 141 salt bridge. *Biochemistry*, *42*(25), 7807-7818. https://doi.org/10.1021/bi034291a

Flynn, J. M., & Downs, D. M. (2013). In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. *J Bacteriol*, *195*(16), 3603-3609. https://doi.org/10.1128/JB.00463-13

Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., Scalmani, G., Barone, V., Mennucci, B., & Petersson, G. (2009). Gaussian 09, revision A. 02, vol 19. Gaussian. *Inc., Wallingford*, 227-238.

Grishin, N. V., Phillips, M. A., & Goldsmith, E. J. (1995). Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci*, *4*(7), 1291-1304. https://doi.org/10.1002/pro.5560040705

Harris, R. K., Becker, E. D., Cabral De Menezes, S. M., Granger, P., Hoffman, R. E., & Zilm, K. W. (2007). Further conventions for NMR shielding and chemical shifts (IUPAC Recommendations 2008). *Pure and Applied Chemistry*, *80*(1), 59-84.

Harris, R. M., & Dunn, M. F. (2002). Intermediate trapping via a conformational switch in the Na(+)-activated tryptophan synthase bienzyme complex. *Biochemistry*, *41*(31), 9982-9990. https://doi.org/10.1021/bi0255672

Harris, R. M., Ngo, H., & Dunn, M. F. (2005). Synergistic effects on escape of a ligand from the closed tryptophan synthase bienzyme complex. *Biochemistry*, *44*(51), 16886-16895.

Hartman, J. D., Neubauer, T. J., Caulkins, B. G., Mueller, L. J., & Beran, G. J. (2015). Converging nuclear magnetic shielding calculations with respect to basis and system size in protein systems. *J Biomol NMR*, *62*(3), 327-340. https://doi.org/10.1007/s10858-015-9947-2

Hilario, E., Caulkins, B. G., Huang, Y. M. M., You, W. L., Chang, C. E. A., Mueller, L. J., Dunn, M. F., & Fan, L. (2016). Visualizing the tunnel in tryptophan synthase with crystallography: Insights into a selective filter for accommodating indole and rejecting water. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, *1864*(3), 268-279. https://doi.org/10.1016/j.bbapap.2015.12.006

Hilario, E., Fan, L., Mueller, L. J., & Dunn, M. F. (2020). PCR Mutagenesis, Cloning, Expression, Fast Protein Purification Protocols and Crystallization of the Wild Type and Mutant Forms of Tryptophan Synthase. *J Vis Exp*, *163*(163). https://doi.org/10.3791/61839

Houben, K. F., & Dunn, M. F. (1990). Allosteric effects acting over a distance of 20-25 Å in the Escherichia coli tryptophan synthase bienzyme complex increase ligand affinity and cause redistribution of covalent intermediates. *Biochemistry*, *29*(9), 2421-2429.

Houben, K. F., Kadima, W., Roy, M., & Dunn, M. F. (1989). L-Serine Analogs Form Schiff-Base and Quinonoidal Intermediates with Escherichia-Coli Tryptophan Synthase. *Biochemistry*, *28*(10), 4140-4147. https://doi.org/DOI 10.1021/bi00436a003

Huang, Y. M., You, W., Caulkins, B. G., Dunn, M. F., Mueller, L. J., & Chang, C. E. (2016). Protonation states and catalysis: Molecular dynamics studies of intermediates in tryptophan synthase. *Protein Sci*, *25*(1), 166-183. https://doi.org/10.1002/pro.2709

Hur, O., Niks, D., Casino, P., & Dunn, M. F. (2002). Proton transfers in the β-reaction catalyzed by tryptophan synthase. *Biochemistry*, *41*(31), 9991-10001. https://doi.org/10.1021/bi025568u

Jacquemin, D., Mennucci, B., & Adamo, C. (2011). Excited-state calculations with TD-DFT: from benchmarks to simulations in complex environments. *Phys Chem Chem Phys*, *13*(38), 16987-16998. https://doi.org/10.1039/c1cp22144b

Jansonius, J. N. (1998). Structure, evolution and action of vitamin B6-dependent enzymes. *Curr Opin Struct Biol*, *8*(6), 759-769. https://doi.org/10.1016/s0959-440x(98)80096-1

Jencks, W. P. (1987). *Catalysis in chemistry and enzymology*. Courier Corporation.

Kantardjieff, K. A., & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci*, *12*(9), 1865-1871. https://doi.org/10.1110/ps.0350503

Karsten, W. E., & Cook, P. F. (2009). Detection of a gem-diamine and a stable quinonoid intermediate in the reaction catalyzed by serine-glyoxylate aminotransferase from Hyphomicrobium methylovorum. *Biochim Biophys Acta*, *1790*(6), 575-580. https://doi.org/10.1016/j.bbagen.2009.02.012

Klamt, A., Moya, C., & Palomar, J. (2015). A Comprehensive Comparison of the IEFPCM and SS(V)PE Continuum Solvation Methods with the COSMO Approach. *J Chem Theory Comput*, *11*(9), 4220-4225. https://doi.org/10.1021/acs.jctc.5b00601

Kulik, V., Weyand, M., Seidel, R., Niks, D., Arac, D., Dunn, M. F., & Schlichting, I. (2002). On The Role of αThr183 in the Allosteric Regulation and Catalytic Mechanism of Tryptophan Synthase. *J Mol Biol*, *324*(4), 677-690. https://doi.org/10.1016/s0022-2836(02)01109-9

Lai, J., Niks, D., Wang, Y., Domratcheva, T., Barends, T. R., Schwarz, F., Olsen, R. A., Elliott, D. W., Fatmi, M. Q., Chang, C. E., Schlichting, I., Dunn, M. F., & Mueller, L. J. (2011). X-ray and NMR crystallography in an Enzyme Active Site: The Indoline Quinonoid Intermediate in Tryptophan Synthase. *J Am Chem Soc*, *133*(1), 4-7. https://doi.org/10.1021/ja106555c

Lane, A. N., & Kirschner, K. (1983). The Catalytic Mechanism of Tryptophan Synthase from Escherichia coli. Kinetics of the Reaction of Indole with the Enzyme-L-serine Complexes. *Eur J Biochem*, *129*(3), 571-582. https://doi.org/10.1111/j.1432-1033.1983.tb07087.x

Maria-Solano, M. A., Iglesias-Fernandez, J., & Osuna, S. (2019). Deciphering the Allosterically Driven Conformational Ensemble in Tryptophan Synthase Evolution. *J Am Chem Soc*, *141*(33), 13049-13056. https://doi.org/10.1021/jacs.9b03646

Matchett, W. H. (1974). Indole channeling by tryptophan synthase of neurospora. *J Biol Chem*, *249*(13), 4041-4049. https://www.ncbi.nlm.nih.gov/pubmed/4277534

Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol*, *33*(2), 491-497. https://doi.org/10.1016/0022-2836(68)90205-2

Metzler, C. M., Cahill, A., & Metzler, D. E. (1980). Equilibria and Absorption-Spectra of Schiff-Bases. *Journal of the American Chemical Society*, *102*(19), 6075-6082. https://doi.org/DOI 10.1021/ja00539a017

Miles, E. W. (1979). Tryptophan synthase: Structure, Function, and Subunit interaction. *Adv Enzymol Relat Areas Mol Biol*, *49*, 127-186. https://doi.org/10.1002/9780470122945.ch4

Miles, E. W. (2001). Tryptophan synthase: A multienzyme complex with an intramolecular tunnel. *Chem Rec*, *1*(2), 140-151. https://doi.org/10.1002/tcr.4

Monod, J., Wyman, J., & Changeux, J. P. (1965). On the Nature of Allosteric Transitions: A Plausible Model. *J Mol Biol*, *12*(1), 88-118. https://doi.org/10.1016/s0022-2836(65)80285-6

Mozzarelli, A., Peracchi, A., Rovegno, B., Dalè, G., Rossi, G. L., & Dunn, M. F. (2000). Effect of pH and Monovalent Cations on the Formation of Quinonoid Intermediates of the Tryptophan Synthase α2β2 Complex in Solution and in the Crystal. *Journal of Biological Chemistry*, *275*(10), 6956-6962.

Ngo, H., Harris, R., Kimmich, N., Casino, P., Niks, D., Blumenstein, L., Barends, T. R., Kulik, V., Weyand, M., Schlichting, I., & Dunn, M. F. (2007). Synthesis and characterization of allosteric probes of substrate channeling in the tryptophan synthase bienzyme complex. *Biochemistry*, *46*(26), 7713-7727. https://doi.org/10.1021/bi700385f

Ngo, H., Kimmich, N., Harris, R., Niks, D., Blumenstein, L., Kulik, V., Barends, T. R., Schlichting, I., & Dunn, M. F. (2007). Allosteric regulation of substrate channeling in tryptophan synthase: modulation of the L-serine reaction in stage I of the β-reaction by α-site ligands. *Biochemistry*, *46*(26), 7740-7753. https://doi.org/10.1021/bi7003872

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T. J., Fan, L., Mueller, L. J., & Dunn, M. F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: evidence for two subunit conformations and four quaternary states. *Biochemistry*, *52*(37), 6396-6411. https://doi.org/10.1021/bi400795e

Percudani, R., & Peracchi, A. (2003). A Genomic Overview of Pyridoxal-Phosphate-Dependent Enzymes. *EMBO Rep*, *4*(9), 850-854. https://doi.org/10.1038/sj.embor.embor914

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, *25*(13), 1605-1612. https://doi.org/10.1002/jcc.20084

Phillips, R. S., & Harris, A. P. (2021). Structural Basis of the Stereochemistry of Inhibition of Tryptophan Synthase by Tryptophan and Derivatives. *Biochemistry*, *60*(3), 231-244. https://doi.org/10.1021/acs.biochem.0c00635

Roise, D., Soda, K., Yagi, T., & Walsh, C. T. (1984). Inactivation of the Pseudomonas striata broad specificity amino acid racemase by D and L isomers of β-substituted alanines: kinetics, stoichiometry, active site peptide, and mechanistic studies. *Biochemistry*, *23*(22), 5195-5201. https://doi.org/10.1021/bi00317a017

Roy, M., Keblawi, S., & Dunn, M. F. (1988). Stereoelectronic control of bond formation in Escherichia coli tryptophan synthase: substrate specificity and enzymatic synthesis of the novel amino acid dihydroisotryptophan. *Biochemistry*, *27*(18), 6698-6704.

Roy, M., Miles, E. W., Phillips, R. S., & Dunn, M. F. (1988). Detection and Identification of Transient Intermediates in the Reactions of Tryptophan Synthase with Oxindolyl-L-alanine and 2,3-Dihydro-L-tryptophan. Evidence for a Tetrahedral (gem-diamine) Intermediate. *Biochemistry*, *27*(23), 8661-8669. https://doi.org/10.1021/bi00423a023

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold spring harbor laboratory.

Schiroli, D., & Peracchi, A. (2015). A subfamily of PLP-dependent enzymes specialized in handling terminal amines. *Biochim Biophys Acta*, *1854*(9), 1200-1211. https://doi.org/10.1016/j.bbapap.2015.02.023

Schneider, T. R., Gerhardt, E., Lee, M., Liang, P. H., Anderson, K. S., & Schlichting, I. (1998). Loop closure and intersubunit communication in tryptophan synthase. *Biochemistry*, *37*(16), 5394-5406. https://doi.org/10.1021/bi9728957

Schrodinger, L. L. C. (2010). The PyMOL molecular graphics system. *Version*, *1*(5), 0.

Tomasi, J., Mennucci, B., & Cances, E. (1999). The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. *Journal of Molecular Structure-Theochem*, *464*(1-3), 211-226. https://doi.org/Doi 10.1016/S0166-1280(98)00553-3

Walsh, C. (1979). *Enzymatic reaction mechanisms*. WH Freeman.

Woehl, E., & Dunn, M. F. (1999). Mechanisms of monovalent cation action in enzyme catalysis: the tryptophan synthase α-, β-, and αβ-reactions. *Biochemistry*, *38*(22), 7131-7141.

Woehl, E. U., Tai, C.-H., Dunn, M. F., & Cook, P. F. (1996). Formation of the α-aminoacrylate intermediate limits the overall reaction catalyzed by O-acetylserine sulfhydrylase. *Biochemistry*, *35*(15), 4776-4783.

Wolfram Research, I. (2020). *Mathematica*. In (Version 12.1) Wolfram Research, Inc.

Yanofsky, C., & Crawford, I. P. (1972). 1 tryptophan synthetase. In *The enzymes* (Vol. 7, pp. 1-31). Elsevier.

Chapter 6

^{17}O and ^{19}F Nuclei as NMR Probes for Mechanistic Enzymology

6.1. Introduction

6.1.1 A brief introduction on ^{17}O NMR Spectroscopy

NMR is a powerful tool for probing tautomerization and acid-base chemistry at atomic resolution. Tracking the equilibrium involved in tautomerization can be accomplished by interrogating the heteronuclei involved in the proton exchange, yet to date this has largely been limited to ^{15}N chemical shift measurements in the solid state (Hayashi, 1995; Johan N Jansonius, 1998; Limbach et al., 2011; Carol M Metzler et al., 1980). Oxygen, the other key atomic species often involved in tautomerism, is not yet a standard nuclear probe in biological NMR spectroscopy, despite the potential wealth of chemical insight it can provide (Lemaıtre et al., 2004; Wu, 2019). Here this chapter reports the application of ^{17}O quadrupole central transition (QCT) NMR to interrogate kinetically competent species under conditions of active catalysis, probing two quasi-stable intermediates (E(A-A) and E(A-A)(BZI)) in the catalytic cycle of the 143 kDa, PLP-dependent enzyme tryptophan synthase (TS) (Rodney M Harris et al., 2005; Hyde et al., 1988). This chapter highlights the use of ^{17}O QCT NMR to mechanistic enzymology extending from the work done by former lab member (Young, 2016). ^{17}O is the only NMR active isotope of oxygen and possesses a quadrupolar, spin I = 5/2 nucleus with a gyromagnetic ratio comparable to ^{15}N and low natural abundance (0.037%) (Gerothanassis, 2010; Young et al., 2016). NMR studies of oxygen in biological macromolecules, whether in the solution or solid state, are extraordinarily rare(Lee et al., 1988; Oldfield et al., 1991; Zhu et al., 2010). ^{17}O QCT NMR spectroscopy in solution takes advantage of the unique relaxation properties of the ^{17}O

central transition; in the limit of slow isotropic motion the linewidth narrows with increasing rotational correlation time, τc, an unusual feature compared to spin 1/2 nuclei (Bull et al., 1979; Zhu & Wu, 2011). Applied initially by Lee and Oldfield (Lee et al., 1988; Oldfield et al., 1991) to carbon monoxide bound to hemeproteins, ^{17}O QCT NMR spectroscopy was recently demonstrated by Zhu and Wu (Zhu & Wu, 2011) to be applicable even when the quadrupolar coupling constant of the substrate is in the MHz range typical of many organic compounds. A second unusual feature of QCT NMR is that the peak position of the central transition is perturbed by a 2nd order magnetic-field dependent frequency shift; to extract the isotropic chemical shift requires measurements at multiple magnetic fields (Zhu & Wu, 2011). The field-dependent line shapes also allow for the extraction of quadrupole and chemical shift anisotropy product parameters, PQ and PSA, that may also be diagnostic of chemical state. Protonation states and tautomerization in the active site of TS have been studied using ^{13}C, ^{15}N, and ^{31}P solid-state NMR for a number of intermediates in the catalytic cycle (Bethany G Caulkins et al., 2016; Young et al., 2016). This chapter focuses on the application of ^{17}O solution state NMR spectroscopy to the E(A-A) intermediate and the BZI-E(A-A) intermediate to provide insights into the H-bonding interactions and ionic state of the carboxyl group in this complex.

6.1.2 Using ^{19}F NMR as a Probe to investigate the allosteric interactions in TS.

Multidrug-resistant pathogens are an increasing danger to public health, with more than 2.8 million infections reported annually in the United States, resulting in more than 35 000 deaths (Hunter, 2020). The biochemical mechanisms by which bacterial cells resist treatment have been classified into four general types, including modification of the drug, alteration of the target site, decreased permeability and overall structural or

functional adaptations (Hawkey, 1998; Munita & Arias, 2016). Clinically resistant strains of Salmonella have been identified by the US Center for Disease Control and Prevention as the primary concern among the 10 major public health threats. These strains have displayed a significant rise in resistance to commonly prescribed antibiotics. Additionally, there is apprehension regarding drug-resistant strains of Escherichia coli due to their prevalent occurrence in both humans and animals (Munita & Arias, 2016; Rasheed et al., 2014). Therefore, the development of novel drugs targeting vital bacterial pathways is essential to successfully fight bacterial infections.

One critical pathway in microorganisms is L-tryptophan (L-Trp) biosynthesis. In bacteria, yeast, molds, and plants, the tryptophan operon encodes enzymes involved in metabolism of this amino acid (Lott, 2020; Wellington et al., 2017). Auxotrophic mutants of pathogenic bacteria with defects in L-Trp biosynthesis lose virulence within a host organism (Smith et al., 2001). Antigenic and inflammatory signals lead to activation of the indoleamine 2,3 dioxygenase pathway responsible for L-Trp catabolism, thus resulting in L-Trp depletion upon microbial infection. With the bacterial survival capacity under L-Trp starvation, the limited supply of the amino acid in host cells, and the lack of comparable biosynthetic machinery in higher animals and humans, components of the L-Trp biosynthetic pathway are a highly specific drug target (Zelante et al., 2009). This work presented in this chapter is done in collaboration with the Chang group here at University of California, Riverside focused on inhibiting the function of the TS bienzyme complex. In a biological setting, the TS exists as a tetramer, and two dimers connect linearly with two β subunits attached to each other (Hyde et al., 1988). Protein conformational changes play an important role in ligand binding; therefore, four conformations were selected from molecular dynamics (MD) simulations and three

experimental structures for docking screening targeting IGP binding site in α subunit of the enzyme (Bosken et al., 2022). Ranking of compounds was based on the lowest binding energy from docking results and scores from their absorption, distribution, metabolism, and excretion (ADME) properties. A whole-cell minimum inhibitory concentration (MIC) assay was used to evaluate the inhibiting capacity of the selected 28 compounds, and one compound, 3-amino-3-imino-2-phenyldiazenylpropanamide (Compound 1), was identified to significantly prevent bacterial growth. We performed solution state ^{19}F NMR studies to verify binding of Compound 1 to the α subunit. Here, we present ^{19}F NMR studies to further characterize the allosteric linkages among ligand binding, chemical reaction, and conformation change in the α-subunit of tryptophan synthase bienzyme complex. To accomplish this objective, we synthesized and characterized substrate analogue (N-(4-trifluoromethyoxylbenzenesulfonyl)-2-amino-1-ethylphosphate (F9) for IGP (Ghosh et al., 2021; Ngo, Harris, et al., 2007; Dimitri Niks et al., 2013). As will be shown, the CF3 groups of the F9 complexes provide ^{19}F signals which sample the channel microenvironment at the α−β subunit interface as well as the IGP binding site. It will be shown herein that the chemical shifts of these probes provide signatures for competitive binding for the IGP binding site in the α subunit of TS and correlate to active site chemistry in the α-subunit of TS, missed by single crystal X-ray studies.

In addition to the IGP binding site assessed in the docking screen and NMR experiments, crystallization of the protein in complex with Compound 1 showed an alternative binding site for the ligand at the αβ interface of the enzyme. MD simulations were performed to identify specific ligand-protein interactions at both sites and provide prospects for improving this lead compound. Based on our results, the proposed

compound is a promising lead for further development of a strong antimicrobial agent targeting the L-Trp metabolic pathway of several bacterial species due to significant structural conservation of TS.

6.2 Experimental Section

6.2.1 ^{17}O L-Ser Synthesis

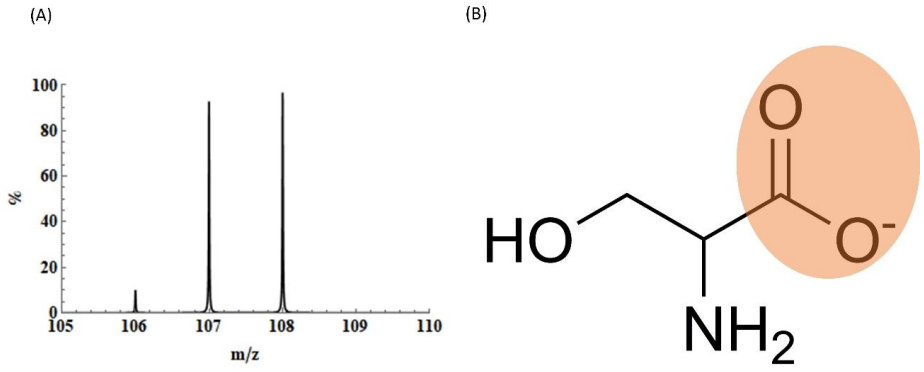


Fig 6.1 (A) Mass Spectrogram showing incorporation of ^{17}O at the carboxylic oxygen site(s) done here using electrospray ionization mass spectrometry (ESI-MS) (Micromass Q-TOF Micro) in positive ion mode. (B) carboxylic sites of ^{17}O enrichment in L-Ser.

Acid-catalyzed exchange was employed to enrich L-serine at the carboxylic oxygen site(s). 210.6 mg of solid L-serine (99%; Alfa-Aesar) was added to 1000 µL of $H_2{}^{17}O$ (90 atom %, Cambridge Isotope Laboratories, Inc.) in a 3 mL conical vial. 35 µL of concentrated ACS Grade HCl(aq) (37%, EMD Millipore) was added to the resulting solution to lower the pH to approximately 2, below the pKa of carboxylic acid group. The vial was capped with a self-sealing Teflon cap, evacuated, and placed under an atmosphere of Ar gas. The vial was placed in an oil bath maintained at a temperature of 55 °C for approximately 2-3 weeks after which the solution was neutralized with NaOD (99 atom %, conc. 40% in D_2O, Isotec Inc.), attached to a high vacuum manifold and dried. Incorporation of ^{17}O at the carboxylic oxygen site(s) was confirmed using

electrospray ionization mass spectrometry (ESI-MS) (Micromass Q-TOF Micro) in positive ion mode which indicated enrichment of more than 88%, see Figure 6.1.A for the ESI-MS spectrum of the resulting [^{17}O]-L-Ser. No enrichment of the beta sidechain hydroxyl group was detected indicating that enrichment only occurred at the carboxylic oxygen sites. ESI-MS mass spectrum of [^{17}O]L-serine following enrichment with H2 17O (90 atom %) stirred at 55°C and pH 2 for 16 days. The peaks at 106 m/z, 107 m/z, and 108 m/z correspond to unenriched, singlesite, and two-site enrichment respectively (Fig 6.1).

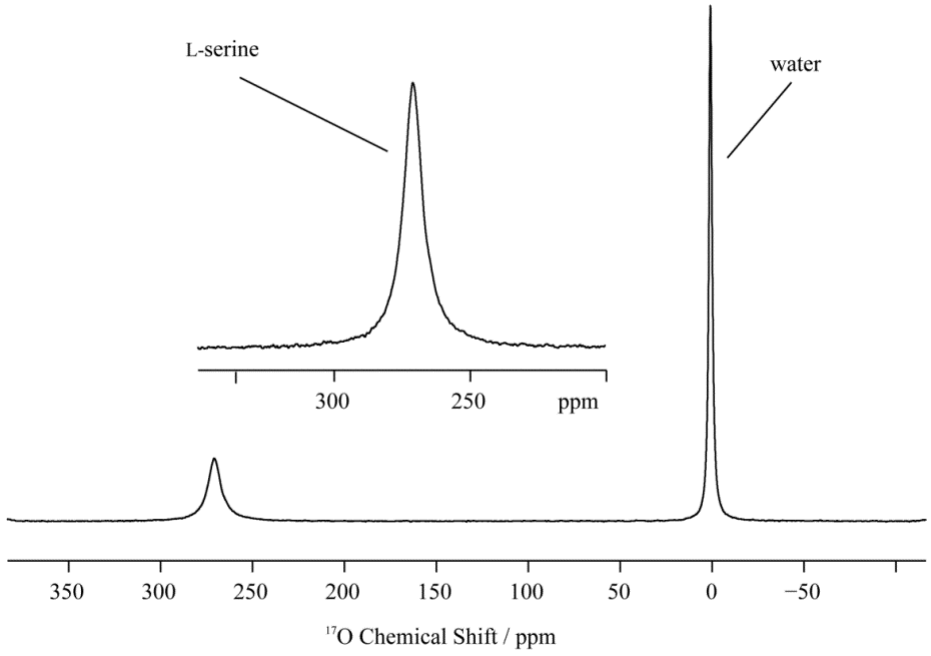


Fig 6.2. ^{17}O-NMR spectrum of L-serine (10 mM) enriched at the carboxylic site (expanded view) measured at 5°C in 50 mM triethanolamine buffer (pH = 7.8), 50 mM CsCl, and 10% D_2O/90% H_2O at a field strength of 14.1 T. 64 (TEA) buffer, pH 7.8, with 50 mM CsCl and 10% D_2O/90% H_2O in a standard 5 mm glass NMR tube at 14.1 T. The L-serine signal, shown in the expanded view, was observed at 271.2 ppm, as referenced to an external sample of pure H_2O measured at 25 °C. (Adapted from (Young, 2016))

6.2.2 ^{17}O QCT NMR Experiments and the Series Connected Hybrid

^{17}O QCT NMR experiments on E(A-A) and E(A-A)(BZI) were conducted at temperatures between 2 °C and 5 °C at field strengths of 11.7 T, 14.1 T, 16.4 T, 21.1T, and 35.2T. The 35.2T is the Series Connected Hybrid (SCH) magnet at the National High Magnetic Field Laboratory (NHMFL) in Tallahassee. All spectra were referenced to the natural abundance ^{17}O signal in an external sample of pure water measured at 25°C and assigned the value of 0 ppm. 90° pulse width calibrations were first determined using the natural abundance ^{17}O signal of water; the nutation frequencies were 9 kHz, 30 kHz, 46 kHz, 60 kHz, and 60 kHz for field strengths of 11.7 T, 14.1 T, 16.4 T (BZI sample only), 21.1T, and 35.2T respectively. Simultaneous mode quadrature detection was used to acquire ^{17}O spectra at all fields. A total of 2048 or 4096 points were collected and zero-filled to two times the number of real points, spectral widths ranged from 41 kHz to 100 kHz, relaxation delays were between 5 and 10 ms, and acquisition times were between 10 and 25 ms. A 90°- τ -180°- τ -acquire pulse sequence was used to measure the spectra of all intermediates. A 90°-τ-180°-2τ-180°-2τ-180°-τ -acquire (triple-echo, 3π) was additionally used to measure the spectra of the E(A-A) and E(A-A)(BZI) complexes at 14.1T and higher to improve observation of bound resonances by suppressing signal from the free substrate and reaction product (Young, 2016; Young et al., 2016). This is based on the selective excitation of the central transition in ^{17}O QCT NMR which leads to a nutation rate that is three times faster for protein-bound substrates than that of the broadband excitation felt by free substrate in solution. The power levels for the 180° refocusing pulses were set to those of the 90° pulses for 11.7T, 14,1T, and 16.4T, but were lowered to 10 kHz for 21.1T and 35.2T. In each case the 180° pulse length was set to 1/3 the nominal pulse for water;

for 11.7T, 14.1T, and 16.4T, the initial excitation pulse was set to 1/3 the water 90° pulse, while at 21.1T and 35.2T it was set to 1/2 the water 90°. The τ delay times were between 5 and 50 μs. The pulse diagrams and phase cycling lists are provided in our previous publication. All spectra were ultimately processed using Bruker's Topspin 3.6 or 4.0 software. Field specific details including hardware are listed below.

11.7 T (500.05 MHz ^{1}H, 67.79 MHz ^{17}O): Varian ^{UNITY}Inova spectrometer equipped with a 5 mm SW PFG probe. Approximately 350 μL of sample was contained in 5 mm Shigemi restricted volume NMR tubes magnetic susceptibility matched to D_2O (Shigemi Inc.).

14.1 T (600.01 MHz ^{1}H, 81.34 MHz ^{17}O): Bruker Avance I spectrometer equipped with a 5 mm BBO Z-grad probe. Approximately 350 μL of sample was contained in 5 mm Shigemi restricted volume NMR tubes magnetic susceptibility matched to D_2O.

16.4 T (699.69 MHz ^{1}H, 94.85 MHz ^{17}O): Bruker Avance I spectrometer equipped with a modified Alderman-Grant coil probe (MAGC) built and tuned to ^{17}O by C. V. Grant at the Center for NMR Spectroscopy and Imaging of Proteins, University of California, San Diego. Approximately 200 μL of sample was contained in 5 mm x 1 cm flat bottom glass NMR tubes.

21.1 T (900 MHz ^{1}H, 121.56 MHz ^{17}O): Bruker Avance III spectrometer equipped with a low-E static double resonance probe with a 5 mm round coil. Both magnet and probe were built in-house at the National Magnetic Field Laboratory in Tallahassee, FL.. Approximately 200 μL of sample was contained in 5 mm x 1 cm flat bottom glass NMR tubes.

35.2 T (1600 MHz ^{1}H, 203.36 MHz ^{17}O): Bruker NEO spectrometer equipped with a low-E static double resonance probe containing a 4 mm solenoid coil built in-house at the National Magnetic Field Laboratory in Tallahassee, FL. Approximately 75 µL of sample was contained in 4 mm OD cylindrical holder (Gan et al., 2017).

6.2.3 Data fitting

The spectra at multiple fields were fit simultaneously to extract the isotropic chemical shift (δiso (ppm)), quadrupole product parameter (PQ (MHz)), and shielding anisotropy product parameter (PSA (ppm)) as described previously (Gan et al., 2017). The procedure was modified to include additional higher order contributions to the linewidth following Wu and coworkers (Gan et al., 2017; Wu, 2019; Young et al., 2016). In this analysis, the peak position δobs and linewidth $\Delta\nu_{1/2}$ are given as:

$$\delta_{obs}(\text{ppm}) = \delta_{iso} - 6000\left(\frac{P_Q}{\nu_0}\right)^2 \quad (12)$$

$$\Delta\nu_{\frac{1}{2}}(\text{Hz}) = \frac{R_{obs}^2}{\pi}$$
$$= 9.6\times10^{-4}\frac{P_Q^4}{\nu_0^2}\tau_c + 7.2\times10^{-3}\frac{P_Q^2}{\nu_0^2}\frac{1}{\tau_c} + 1.1P_{SA}^2\nu_0^2\tau_c + 0.029P_Q^2P_{SA}\tau_c(3\cos^2\beta - 1) \quad (13)$$

These equations apply in the limit of slow isotropic motion and are specific to spin-5/2 nuclei. β is the angle between the principal axis systems of the chemical shift and quadrupolar coupling tensors under the assumption of axial symmetry; here we take β=π/2 (Gan et al., 2017; Wu, 2019), which is supported by modeling of the tensors using first principles methods. Following our earlier protocol (Young et al., 2016), τc was set at 232 ns based on QCT experiments at the same 5 fields for the considerably more stable $E(C_3)^{2AP}$ intermediate.

At each field, the free substrate/product signals were fit and subtracted from the spectrum prior to global fitting of the bound signals. For the SCH E(A-A)(BZI) data only, a reference spectrum for the free substrate/product was first subtracted.

6.2.4 ^{19}F NMR

Salmonella typhimurium tryptophan synthase (*St*TS) and N-(4-trifluoromethyoxylbenzenesulfonyl)-2-amino-1-ethylphosphate (F9) were prepared as previously described (Ghosh et al., 2021; Ngo, Harris, et al., 2007). A stock solution of 100 mM Compound 1 was prepared in triethanolamine (TEA) buffer at pH 7.8 in the presence of 50 mM CsCl and 20% DMSO (v/v).

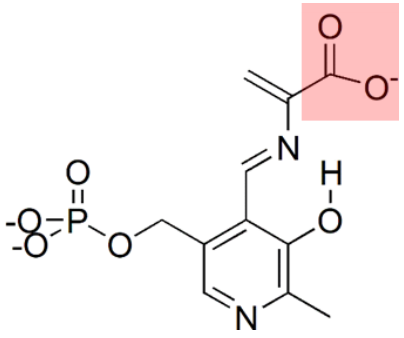
^{19}F NMR experiments (Dimitri Niks et al., 2013) were performed on a Bruker 600 MHz NMR spectrometer equipped with a 5 mm ^{1}H/^{19}F (BBFO) double resonance probe operating at a ^{1}H frequency of 599.75 MHz. All samples were prepared in TEA buffer adjusted to pH 7.8 containing 50 mM CsCl and 10 % D_2O (v/v) (no pH correction was made for the D_2O content). Data were acquired at 25°C with a relaxation delay of 1.0 s and a spectral width of 177.21 ppm. For each experiment, 4096 transients were accumulated and averaged. All the spectra were processed with 6 Hz line broadening.

6.3 Results and Discussion

6.3.1 ^{17}O NMR Spectroscopy

Table 6.1: ^{17}O NMR parameters extracted from QCT NMR

	δ_{iso}/ppm	P_Q/MHz	P_{SA}/ppm
E(A-A)			
O1	256.6 ± 0.1	7.40 ± 0.01	341 ± 3
O2	288.6 ± 0.1	8.00 ± 0.01	279 ± 3
E(A-A)(BZI)			
O1	257.8 ± 0.1	7.23 ± 0.01	505 ± 3
O2	285.8 ± 0.1	7.89 ± 0.01	492 ± 4



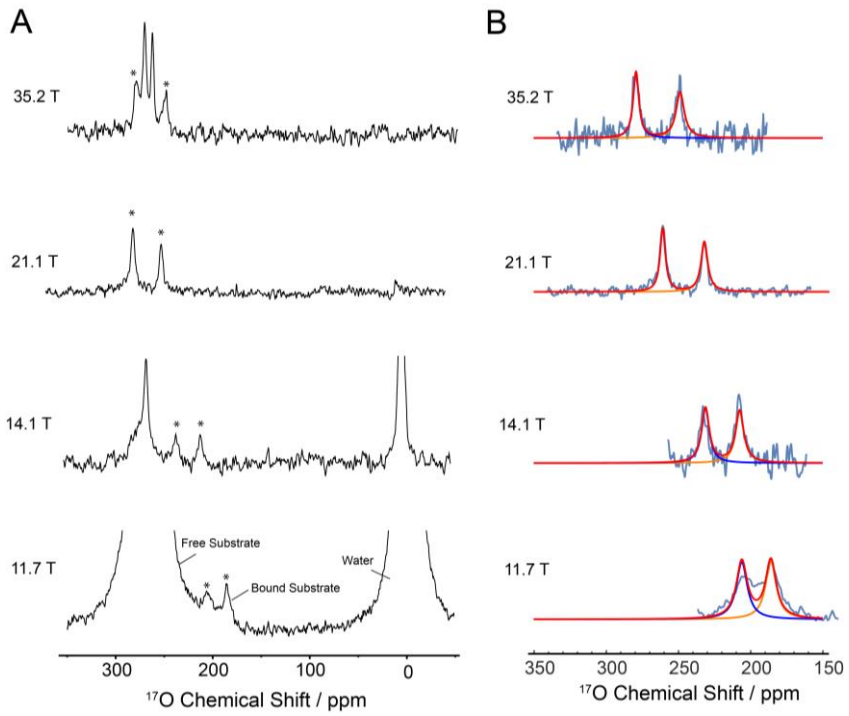


Fig 6.3: ^{17}O QCT NMR spectra of the E(A-A) intermediate from 11.7 – 35.2 T. At 11.7 T the spectrum was acquired with a single echo before detection and displays considerable intensity from the natural abundance $H_2{}^{17}O$ water and free $^{17}O_2$-L-Ser substrate in solution. At 14.1T and higher, the spectra were acquired using the triple echo sequence to greatly suppress the free substrate and water. (B) Traces of the best-fit component line shapes O1 (blue), O2 (orange), and their sum O1 + O2 (red), overlaid on experimental spectra (after background peak subtraction). 100 Hz of line broadening was applied to all spectra shown in the figure.

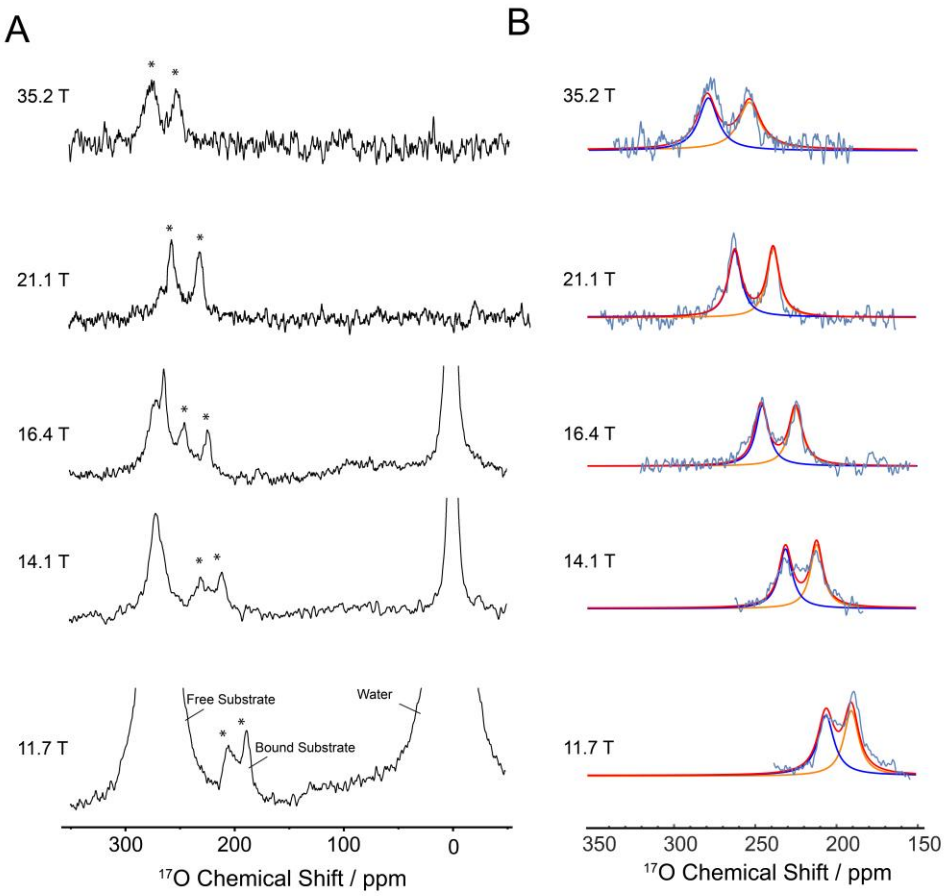


Fig 6.4. ^{17}O QCT NMR spectra of the E(A-A)(BZI) intermediate from 11.7 – 35.2 T. At 11.7 T the spectrum was acquired with a single echo before detection and displays considerable intensity from the natural abundance $H_2{}^{17}O$ water and free $^{17}O_2$-L-Ser substrate in solution. At 14.1T and higher, the spectra were acquired using the triple echo sequence to greatly suppress the free substrate and water. (B) Traces of the best-fit component line shapes O1 (blue), O2 (orange), and their sum O1 + O2 (red), overlaid on experimental spectra (after background peak subtraction). 100 Hz of line broadening was applied to all spectra shown in the figure.

The ^{17}O chemical shifts of the substrate carboxylate group were measured using ^{17}O quadrupole central transition (QCT) NMR in solution (Fig 6.3 and Fig 6.4). These were the only shifts measured for the intermediate in solution. ^{17}O NMR is not yet considered a standard high-resolution probe for biomolecular NMR, but QCT NMR takes advantage of the unique property of the NMR central transition to narrow as the size of the protein-substrate complex increases. We previously reported preliminary isotropic ^{17}O shifts for both E(A-A) and E(A-A)(BZI). Here advantage is taken of additional measurements on the 35.2 T series connect hybrid magnet at the National High Magnetic Field Laboratory to improve the accuracy of the extracted parameters. The ^{17}O chemical shifts of 257 ppm and 289 ppm are consistent with an ionized carboxylate group. The upfield shifted signal is tentatively assigned to the oxygen hydrogen bonded to the side chain of βThr110, which forms the closest contact (Jacob B. Holmes et al., 2022).

6.3.2 Characteristics of Compound 1 in inhibition of TS in bacterial systems.

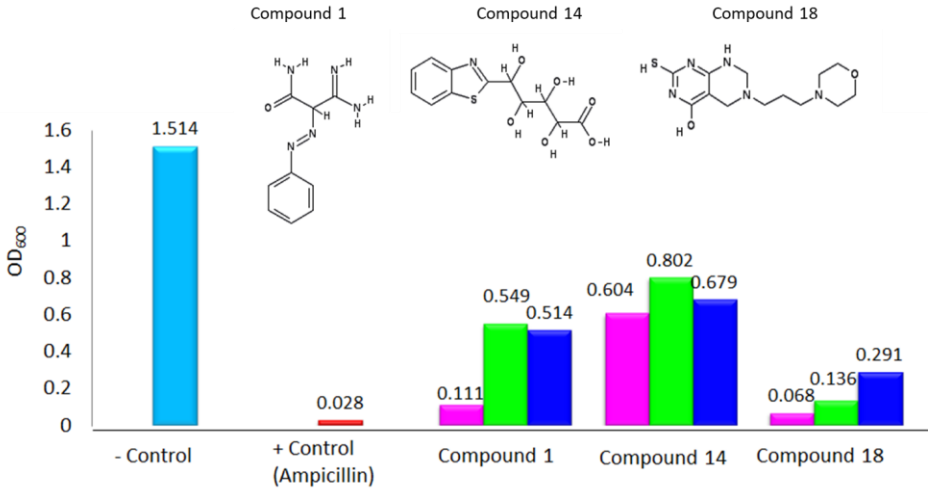


Fig 6.5**.** Bacterial growth assays. A chemical compound concentration 50 μg/ml was used in each minimum inhibitory concentration (MIC) assay. Cell density of *E. coli.* OD_{600} was measured after 24-h incubation at 37°C (magenta), after the addition of 0.45 mg/ml L-Trp (green) and 0.82 mg/ml L-Trp (blue). (Redrawn from (Bosken et al., 2022)).

6.3.3 Competitive Binding of Compound 1 to the IGP Binding Site from ^{19}F NMR.

To validate the binding of Compound 1 to the α subunit IGP binding site, we performed a ^{19}F NMR competitive binding assay between Compound 1 and the high-affinity α-site ligand, F9. F9 is known to bind to the IGP binding site (see, for example, X-ray crystal structure 4HT3 and contains a trifluormethyl group that gives a strong and unique ^{19}F NMR resonance for the free and TRPS-bound forms. Figure 6.6 displays one-dimensional ^{19}F NMR spectra of the complex between TRPS and F9 with increasing concentration of Compound 1. The initial spectrum (blue trace) shows the presence of two ^{19}F resonances: the bound species at -57.11 ppm, which is well described by a Lorentzian line shape, and a slightly sharper resonance for free F9 at -57.44 ppm. Upon subsequent additions of Compound 1, there is an increase in the relative ratio of the free to bound F9 peak. The observed shift in the population of F9 from bound to free as the concentration of Compound 1 increase is consistent with competitive binding of

Compound 1 to the IGP binding site (Fig 6.6), in which the binding of Compound 1 results in fewer available sites for F9 binding. We note, however, that although competition assays are widely used methods to examine ligand binding, the indirect evidence does not completely rule out the alternate possibility that Compound 1 binds to a different site that results in an allosteric effect preventing F9 from binding.

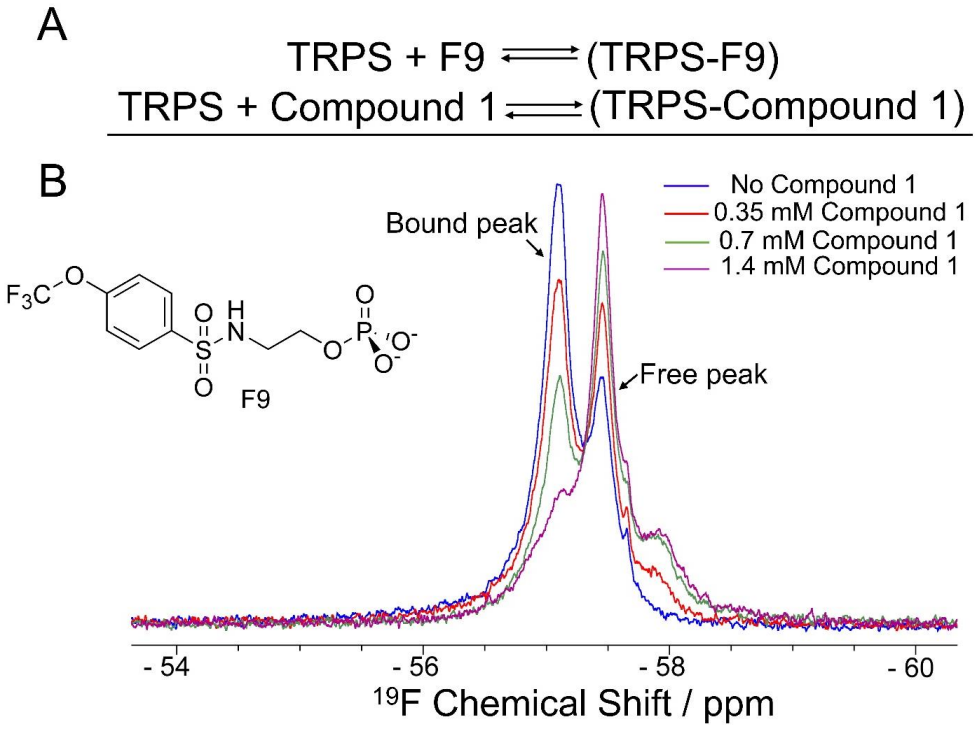


Fig 6.6. Competitive binding between Compound 1 and F9 observed using ^{19}F NMR. (A) Schematic for competitive binding of F9 and Compound 1 to the α subunit of *St*TS. (B) Titration of compound 1 into the TRPS-F9 sample shows a gradual decrease in the bound peak intensity (-57.11 ppm) with a concomitant increase in the free peak intensity (-57.44 ppm). Conditions: [$\alpha\beta$] = 0.376 mM; compound 1 concentrations, 0 mM (blue), 0.35 mM (red), 0.7 mM (green) and 1.4 mM (magenta); [F9] = 2 mM.

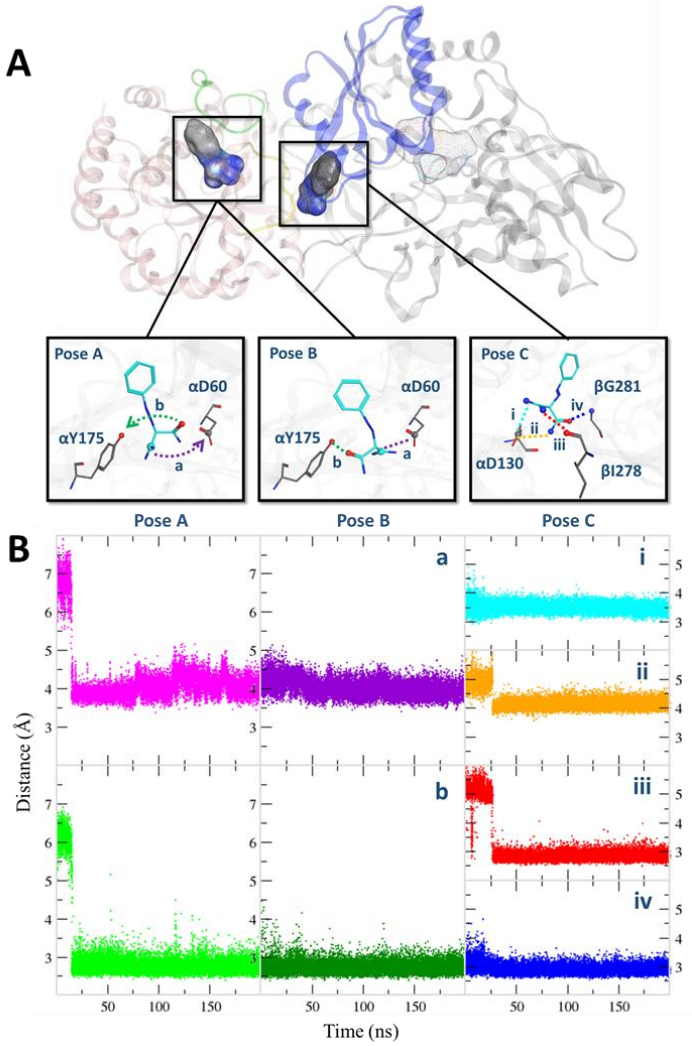


Fig 6.7. Compound 1 conformations and interactions with protein in the α-subunit active site and interdomain surface. (A) Three selected initial conformations of Compound 1 for molecular dynamics (MD) simulations. Compound 1 can bind to two distinct binding sites: the active site (poses A and B) of the α subunit, and that in the αβ inter-domain interface (pose C). The two major conformations in the active site are characterized by a 180° rotation of the amido group resulting in stable hydrogen bonds with the nearby residues. (B) Intermolecular hydrogen bonds between Compound 1 and TRPS during 200-ns MD runs. Hydrogen bonds between Compound 1 and residues D60 (purple, a) and Y175 (green, b) in the active site of α subunit. At the αβ inter-domain interface, hydrogen bonds are observed between residues D130 of α subunit (cyan, i and orange, ii), I278 of β subunit (red, iii) and G281 of β subunit (blue, iv). Redrawn from (Bosken et al., 2022).

6.3.4 ^{17}O NMR Spectroscopy as a probe for mechanistic enzymology

Two conclusions can be drawn. First, the ^{17}O oxygen chemical shifts predict the carboxylate group of the L-Ser moiety in the E(A-A) and E(A-A)(BZI) complexes to be deprotonated with minor differences in H-bonding interactions with βT110. Second, as the nitrogen chemical shift of the Schiff base in E(A-A) and E(A-A)(BZI) complexes indicate that it is (mostly) deprotonated, the phenolic tautomer must predominate. The ^{17}O NMR data combined with First Principles calculations help in identification of correct tautomeric pairs as the winning models of structures relevant to TS catalysis.

^{17}O QCT NMR spectroscopy holds great promise for the characterization of biomolecular systems, and perhaps no field is as well-poised to take advantage of it as chemical enzymology, where proton locations and electrostatic environments are key to understanding mechanisms. This is particularly true of the acid–base chemistry that takes place within the active sites of PLP-dependent enzymes such as tryptophan synthase, where proton transfers and changes in electron density could previously only be inferred from the shifts of other nuclei in the active site. What is remarkable is that such measurements can be performed under conditions of active catalysis, catching kinetically competent intermediates in the act of chemical transformation.

6.3.5 ^{19}F NMR Spectroscopy as a tool for understanding allosteric interactions in TS.

In vitro MIC testing showed that the compound 1 significantly reduced bacterial growth, and equally important, adding L-Trp to the medium recovered bacterial growth, indicating that the compound specifically targets the L-Trp biosynthetic pathway. Data from the solution state NMR assays shows that Compound 1 also binds to the α subunit active site and competes with F9 binding, resulting in reduced signals of the TRPS-F9

complex. In addition to the screened binding pocket in the α subunit, crystallization of the TRPS-Compound 1 complex revealed a binding site of Compound 1 in the $\alpha\beta$-inter-domain interface. The ^{19}F NMR data combined with MD runs for three different binding poses showed that TRPS was stable when Compound 1 was bound to either binding site. TRPS showed similar dynamic fluctuations when its natural substrate IGP or compound 1 (pose B) was bound to the active site. In summary, this work showed that Compound 1 is a good lead for further improvement and development of a potent antimicrobial drug.

6.4 Conclusions

In brief conclusion, this chapter demonstrates that although far from routine, the ^{17}O and ^{19}F NMR spectroscopy can be used as a probe to investigate structure-function relationship in biological systems.

6.5 References

Bosken, Y. K., Ai, R., Hilario, E., Ghosh, R. K., Dunn, M. F., Kan, S.-H., Niks, D., Zhou, H., Ma, W., Mueller, L. J., Fan, L., & Chang, C.-E. A. (2022). Discovery of antimicrobial agent targeting tryptophan synthase. *Protein Science*, *31*(2), 432-442. https://doi.org/https://doi.org/10.1002/pro.4236

Bull, T., Forsén, S., & Turner, D. L. (1979). Nuclear magnetic relaxation of spin 5/2 and spin 7/2 nuclei including the effects of chemical exchange. *The Journal of Chemical Physics*, *70*(6), 3106-3111.

Caulkins, B. G., Young, R. P., Kudla, R. A., Yang, C., Bittbauer, T. J., Bastin, B., Hilario, E., Fan, L., Marsella, M. J., & Dunn, M. F. (2016). NMR crystallography of a carbanionic intermediate in tryptophan synthase: chemical structure, tautomerization, and reaction specificity. *Journal of the American Chemical Society*, *138*(46), 15214-15226.

Gan, Z., Hung, I., Wang, X., Paulino, J., Wu, G., Litvak, I. M., Gor'kov, P. L., Brey, W. W., Lendi, P., & Schiano, J. L. (2017). NMR spectroscopy up to 35.2 T using a series-connected hybrid magnet. *Journal of Magnetic Resonance*, *284*, 125-136.

Gerothanassis, I. P. (2010). Oxygen-17 NMR spectroscopy: Basic principles and applications (Part I). *Progress in Nuclear Magnetic Resonance Spectroscopy*, *56*(2), 95.

Ghosh, R. K., Hilario, E., Liu, V., Wang, Y., Niks, D., Holmes, J. B., Sakhrani, V. V., Mueller, L. J., & Dunn, M. F. (2021). Mutation of βGln114 to Ala alters the stabilities of allosteric states in tryptophan synthase catalysis. *Biochemistry*, *60*(42), 3173-3186.

Harris, R. M., Ngo, H., & Dunn, M. F. (2005). Synergistic effects on escape of a ligand from the closed tryptophan synthase bienzyme complex. *Biochemistry*, *44*(51), 16886-16895.

Hawkey, P. M. (1998). The origins and molecular basis of antibiotic resistance. *Bmj*, *317*(7159), 657-660.

Hayashi, H. (1995). Pyridoxal enzymes: mechanistic diversity and uniformity. *The journal of biochemistry*, *118*(3), 463-473.

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., Ghosh, R. K., Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., Bogie, P. M., Paulino, J., Wang, X., Riviere, G., Bosken, Y. K., Struppe, J., Hassan, A., Guidoulianov, J., Perrone, B., Mentink-Vigier, F., . . . Mueller, L. J. (2022). Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *Proceedings of the National Academy of Sciences*, *119*(2), e2109235119. https://doi.org/doi:10.1073/pnas.2109235119

Hunter, P. (2020). A war of attrition against antibiotic resistance: current strategies try to keep antibiotic resistance at bay and further encourage research to produce genuinely novel antibacterials. *EMBO reports*, *21*(6), e50807.

Hyde, C., Ahmed, S., Padlan, E., Miles, E. W., & Davies, D. (1988). Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *Journal of Biological Chemistry*, *263*(33), 17857-17871.

Jansonius, J. N. (1998). Structure, evolution and action of vitamin B6-dependent enzymes. *Current opinion in structural biology*, *8*(6), 759-769.

Lee, H. C., Cummings, K., Hall, K., Hager, L., & Oldfield, E. (1988). Oxygen-17 nuclear magnetic resonance spectroscopic studies of carbonmonoxyperoxidases. *Journal of Biological Chemistry*, *263*(31), 16118-16124.

Lemaıtre, V., Smith, M., & Watts, A. (2004). A review of oxygen-17 solid-state NMR of organic materials—towards biological applications. *Solid state nuclear magnetic resonance*, *26*(3-4), 215-235.

Limbach, H.-H., Chan-Huot, M., Sharif, S., Tolstoy, P. M., Shenderovich, I. G., Denisov, G. S., & Toney, M. D. (2011). Critical hydrogen bonds and protonation states of pyridoxal 5′-phosphate revealed by NMR. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1814*(11), 1426-1437.

Lott, J. S. (2020). The tryptophan biosynthetic pathway is essential for Mycobacterium tuberculosis to cause disease. *Biochemical Society Transactions*, *48*(5), 2029-2037.

Metzler, C. M., Cahill, A., & Metzler, D. E. (1980). Equilibriums and absorption spectra of Schiff bases. *Journal of the American Chemical Society*, *102*(19), 6075-6082.

Munita, J. M., & Arias, C. A. (2016). Mechanisms of antibiotic resistance. *Virulence mechanisms of bacterial pathogens*, 481-511.

Ngo, H., Harris, R., Kimmich, N., Casino, P., Niks, D., Blumenstein, L., Barends, T. R., Kulik, V., Weyand, M., & Schlichting, I. (2007). Synthesis and characterization of allosteric probes of substrate channeling in the tryptophan synthase bienzyme complex. *Biochemistry*, *46*(26), 7713-7727.

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T. J., Fan, L., Mueller, L. J., & Dunn, M. F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: Evidence for two subunit conformations and four quaternary states. *Biochemistry*, *52*(37), 6396-6411.

Oldfield, E., Lee, H. C., Coretsopoulos, C., Adebodun, F., Park, K. D., Yang, S., Chung, J., & Phillips, B. (1991). Solid-state oxygen-17 nuclear magnetic resonance spectroscopic studies of [17O2] picket fence porphyrin, myoglobin, and hemoglobin. *Journal of the American Chemical Society*, *113*(23), 8680-8685.

Rasheed, M. U., Thajuddin, N., Ahamed, P., Teklemariam, Z., & Jamil, K. (2014). Antimicrobial drug resistance in strains of Escherichia coli isolated from food sources. *Revista do Instituto de Medicina Tropical de São Paulo*, *56*, 341-346.

Smith, D. A., Parish, T., Stoker, N. G., & Bancroft, G. J. (2001). Characterization of auxotrophic mutants of Mycobacterium tuberculosis and their potential as vaccine candidates. *Infection and immunity*, *69*(2), 1142-1150.

Wellington, S., Nag, P. P., Michalska, K., Johnston, S. E., Jedrzejczak, R. P., Kaushik, V. K., Clatworthy, A. E., Siddiqi, N., McCarren, P., & Bajrami, B. (2017). A small-molecule allosteric inhibitor of Mycobacterium tuberculosis tryptophan synthase. *Nature chemical biology*, *13*(9), 943-950.

Wu, G. (2019). 17O NMR studies of organic and biological molecules in aqueous solution and in the solid state. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *114*, 135-191.

Young, R. P. (2016). *Biological Structure and Dynamics at the Interface of Experiment and Theory* UC Riverside].

Young, R. P., Caulkins, B. G., Borchardt, D., Bulloch, D. N., Larive, C. K., Dunn, M. F., & Mueller, L. J. (2016). Solution-State 17O Quadrupole Central-Transition NMR Spectroscopy in the Active Site of Tryptophan Synthase. *Angewandte Chemie International Edition*, *55*(4), 1350-1354.

Zelante, T., Fallarino, F., Bistoni, F., Puccetti, P., & Romani, L. (2009). Indoleamine 2, 3-dioxygenase in infection: the paradox of an evasive strategy that benefits the host. *Microbes and infection*, *11*(1), 133-141.

Zhu, J., & Wu, G. (2011). Quadrupole central transition 17O NMR spectroscopy of biological macromolecules in aqueous solution. *Journal of the American Chemical Society*, *133*(4), 920-932.

Zhu, J., Ye, E., Terskikh, V., & Wu, G. (2010). Solid-State 17 O NMR Spectroscopy of Large Protein–Ligand Complexes. *Angewandte Chemie International Edition*, *45*(49), 8399-8402.

Chapter 7

Toho-1 β-Lactamase Solution NMR Backbone Dynamics Measurements

7.1 Introduction

β-lactam antibiotics inhibit the synthesis of bacterial cell walls by targeting penicillin-binding proteins (PBPs). The binding of β-lactam antibiotics to PBPs makes them chemically inert, leading to bacterial cell death. β-lactam antibiotics remain the most widely used in antimicrobial chemotherapy due to their low toxicity, as there are no human counterparts of their bacterial targets (Kong et al. 2010). To counter these potent antimicrobials, bacteria have evolved to produce β-lactamase enzymes that cleave the amide bond within the β-lactam ring via a general base hydrolysis mechanism (Bebrone et al. 2010; Bush 2010). This family of hydrolytic enzymes can be divided into four distinct classes based on sequence homology, with classes A, C, and D utilizing a catalytic serine residue to hydrolyze the β-lactam ring, while class B are metalloenzymes that utilize active-site zinc ion(s) to perform the hydrolysis (Bush 2013). Representative families of class A β-lactamases are TEM, SHV, and the class A extended-spectrum β-lactamases (ESBL) CTX-M (Jacoby 2006); the latter includes the Toho-1 β-lactamase studied here (also classified as CTX-M-44). ESBLs were found in the 1980s and exhibit increased hydrolytic activity against first, second, and third generation extended-spectrum cephalosporins and monobactam antibiotics.

Along with the development of β-lactamase-resistant antibiotics, the strategy towards combatting resistance includes combining β-lactam antibiotics with mechanism-based β-lactamase inhibitors (Tooke et al. 2019). Clavulanic acid, sulbactam, and tazobactam are examples of β-lactam-based inhibitors active against class A β-lactamases that were introduced into clinical practice four decades ago (Drawz and Bonomo 2010). These

inhibitors form a stable acyl-enzyme intermediate with the catalytic serine (S70), but eventually the bound inhibitor hydrolytically deacetylates and regenerates the catalytically active enzyme (Helfand et al. 2003).

The increasing abundance of ESBL-producing pathogens has driven increased use of and reliance on carbapenems, a class of highly effective antibiotic agents reserved for the treatment of severe or high-risk bacterial infections (Shirley 2018). The emergence and spread of carbapenemase-producing pathogens are particularly concerning and have highlighted the urgent need for new antimicrobial agents. Classical β-lactamase inhibitors such as clavulanic acid, tazobactam, and sulbactam lack activity against many classes of β-lactamases, meaning that many first-generation β-lactam/β-lactamase inhibitor combinations are frequently ineffective against multidrug-resistant pathogens (Shirley 2018).

The emergence of resistant class A β-lactamases (Bebrone et al. 2010) continues to motivate the development of novel inhibitors that are capable of impeding antibiotic resistance by being structurally different from β-lactam ring containing substrates. Avibactam, a member of the class of azabicycloalkanes, is one such non-β-lactam β-lactamase inhibitor that is approved by the FDA as part of combination therapies for the treatment of complicated intra-abdominal and urinary tract infections (Lagacé-Wiens et al. 2014). Ceftazidime-avibactam, sold under several brand names, is an intravenously administered combination of the third-generation cephalosporin ceftazidime and avibactam. It has excellent in vitro activity against many Gram-negative pathogens, including many ESBLs. However, it is not active against metallo-β-lactamase (Class B) producing bacterial strains (Shirley 2018).

Avibactam does not have a β-lactam core, and a sulfate moiety replaces the carboxyl group, as shown in Fig 7.1. Avibactam forms a unique carbamyl covalent linkage with the catalytic serine, which does not hydrolytically dissociate. Instead, the decarbamylation leads to the recyclization of the azabicyclooctane ring structure, reforming the intact inhibitor capable of further inhibition (Ehmann et al. 2012).

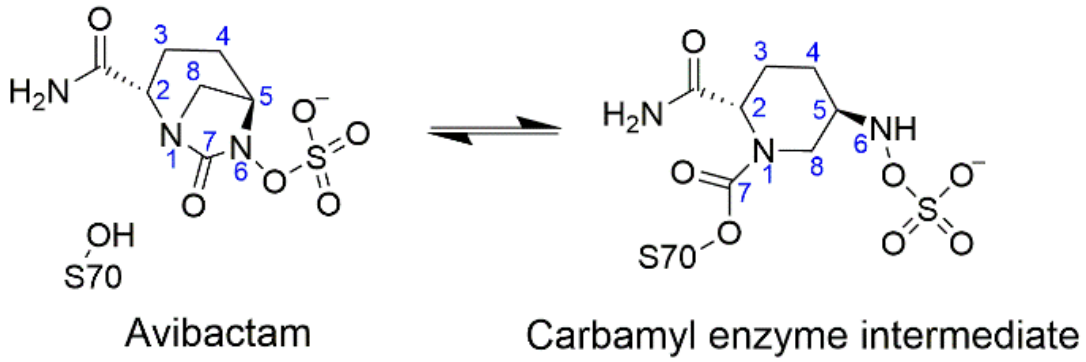


Fig 7.1 The reversible avibactam-mediated inhibition of class A β-lactamases. During the inhibition reaction, the N6-C7 bond is broken, and N6 is protonated, resulting in the opening of the diazobicyclooctane (Langan et al. 2020).

Previous kinetic, mutagenesis, spectroscopic, and computational studies have tried to elucidate the molecular details related to the activation of the catalytic serine residue during catalysis (Damblon et al. 1996; Tomanicek et al. 2013; Langan et al. 2016; Vandavasi et al. 2016; Langan et al. 2018) and during inhibition by avibactam (Ehmann et al. 2012; Lahiri et al. 2013; King et al. 2016; Lizana and Delgado 2018; Pemberton et al. 2020). There is still considerable debate over the exact roles of the active site lysine (K73), serine (S130), and glutamic acid (E166) residues, as well as the active site water molecule.

The dynamical properties of β-lactamases are linked to their biological function and are thought to play an essential role in governing the evolution of this family of enzymes (Petrovic et al. 2018). Recent computational modeling studies have addressed the role

of dynamics towards conferring resistance in the evolution of Class-A β-lactamases (Verma et al. 2013; Zou et al. 2015; Cortina and Kasson 2016; Hart et al. 2016). These studies concentrate primarily on sequence comparisons, but it appears that sequence and structure alone are not enough to classify the physicochemical properties across the class-A β-lactamases: detailed comparisons of electrostatic and dynamical properties are necessary to reflect evolutionary trends. This highlights the need for empirical data on dynamics for modern class-A members (Verma et al. 2013). To date, the dynamics of a small number of class-A β-lactamases including TEM-1, PSE-4, and BlaC have been characterized by NMR relaxation experiments to obtain residue-specific generalized order parameters and local correlation times (Savard and Gagné 2006; Morin and Gagné 2009; Fisette et al. 2010; Fisette et al. 2012; Elings et al. 2020). These experiments show well-ordered structures, with generalized order parameters on the picosecond-to-nanosecond timescale of ~ 0.85-0.9, with microsecond-to-millisecond motions in the vicinity of the active site. This combination of high rigidity on short timescales and active site flexibility on longer timescales is consistent with hypotheses for achieving both high catalytic efficiency and broad substrate specificity (Fisette et al. 2010).

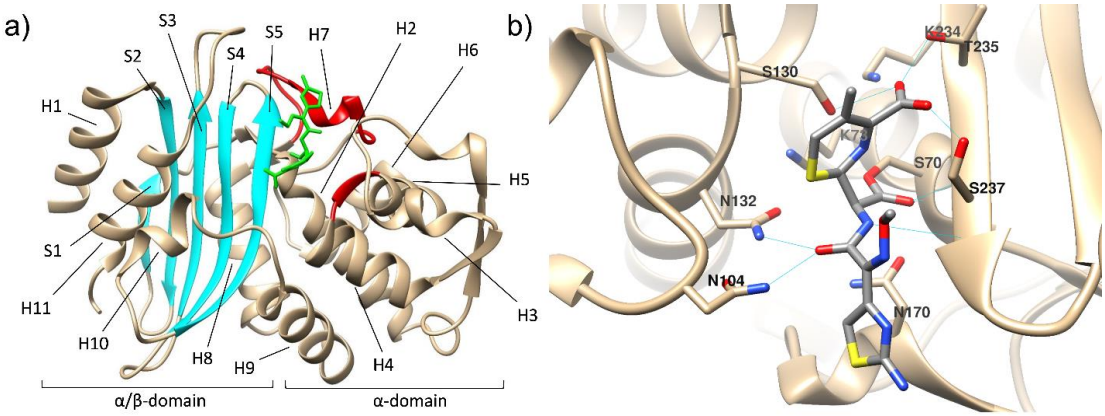


Fig 7.2 (a) Ribbon model of the Toho-1 β-lactamase from X-ray crystallography (PDB ID: 1IYO) (Shimamura et al. 2002). The β-strands forming the β-sheet in the alpha-beta domain are shown in blue. The Ω loop and SDN loop are shown in red. The antibiotic cefotaxime is shown in green bound in the active site. (b) The active site highlighting sidechains that interact with cefotaxime; blue lines indicate hydrogen bonds. Catalytic residue E166, mutated to A166 in this structure, is not shown. Figure prepared using Chimera (Pettersen et al. 2004).

Here we examine the dynamics of the Toho-1 β-lactamase in the absence and presence of the non-β-lactam β-lactamase inhibitor avibactam. Based on sequence homology and activity against extended-spectrum cephalosporins, Toho-1 β-lactamase is classified as a CTX-M-type extended-spectrum β-lactamase (ESBL) (Ishii et al. 1995). This modern member of the class A enzymes is a 263-residue protein and consists of two highly conserved domains (α/β and α) with the active site cavity formed at their interface (Fig 2). Helix H2 forms the active site's base/bottom and provides the catalytic residues S70 and K73. Three structural motifs lying over the base form a cavity for the active site. The first of these is the segment consisting of residues S130, D131, and N132 (SDN loop), between helices H4 and H5, of which S130 is the catalytic residue. Strand S5 provides catalytic residues K234 and T235 and forms the second structural motif. The third motif consists of the N-terminal portion of the Ω loop and provides the catalytic residue E166. The 16 residue-long Ω loop (consisting of residues 164 to179 and devoid of any extended secondary structure other than helix H7) and the 9 residue-long SDN loop are

standard, defining features of class A β-lactamases. The presence of a network of hydrogen bonds and salt bridges between the catalytic residue sidechains rigidifies the active site in the absence of a substrate, as seen in the crystal structure of the free enzyme (PDB ID: 2ZQ8) (Shimamura et al. 2009).

Within the Ω loop of class A β-lactamases members, the backbone amide and side chain of R164 form salt bridges to the side chain of D179 to maintain a “bottleneck” structure having mobility essential for substrate recognition and catalysis (Banerjee et al. 1998). Reduced Ω loop flexibility correlates with increased hydrolysis of antibiotic cefotaxime (Hart et al. 2016). The loop contains the active site residue El66, which is highly conserved in class A β-lactamases and has been shown to participate in the deacylation and acylation steps of the catalytic cycle via a water molecule, potentially competing with K73 (Lobkovsky et al. 1993; Langan et al. 2020). Mutations within the Ω loop affect catalytic specificity and mechanism and play a role in governing the evolution of this class of enzymes (Petrosino and Palzkill 1996; Stojanoski et al. 2015; Patel et al. 2017). In contrast to the other ESBL’s, the Ω loop in Toho-1 β-lactamase is shifted to (rests on) helix H5 instead of H6, which avoids steric interactions with the bulky side chain of oxyimino-cephalosporins, resulting in improved affinities for these substrates (Shimamura et al. 2002).

Complexes of the wild-type β-lactamases with their target substrates such as cefotaxime cannot typically be characterized by solution-state NMR because of the rapid turnover rate (the crystal structure shown in Fig 7.2 is for the E166A mutant, which is unable to hydrolyze the substrate). Inhibitors such as avibactam, however, remain covalently bound for the duration of the experiment (Elings et al. 2020). Here, we make use of this extended stability to characterize both the Toho-1 β-lactamase free in solution and in

complex with avibactam. Backbone chemical shift assignments are reported for the R274N/R276N double mutant of Toho-1 β-lactamase based on triple resonance solution-state NMR experiments on a uniformly ^{2}H,^{13}C,^{15}N-labeled sample. The non-active-site R274N/R276N mutations prevent the protein's aggregation in solution. Based on these assignments and the protein sequence, the secondary structure is predicted. Perturbations in the assigned chemical shifts upon titration with avibactam allows for mapping of the inhibitor binding site. ^{15}N relaxation data obtained for ^{2}H,^{15}N-labeled samples at a single field are analyzed by model-free analysis (MFA) to obtain residue-specific dynamical parameters, including the generalized order parameter (S^2), local correlation time (τ_e), and residual exchange contribution (R_{ex}). Complementary relaxation dispersion experiments indicate that binding of the inhibitor avibactam leads to enhanced dynamics on the millisecond timescale in and around the active site. The combination of high rigidity on short timescales and active site flexibility on longer timescales is proposed to play an important role in effecting both high catalytic efficiency and broad substrate specificity.

7.2 Experimental Section

7.2.1 Expression and Purification of Recombinant Protein

Uniformly ^{2}H-^{15}N labeled and ^{2}H-^{13}C-^{15}N labeled Toho-1 β-lactamase were expressed and purified as described previously with some modifications (Tomanicek et al. 2010; Langan et al. 2016). Briefly, BL21 *E. coli* cells harboring the Toho-1 β-lactamase expression plasmid were cultured and progressively adapted to $^{2}H_2O$ in minimal medium containing D-glucose in a shaking incubator set to 250 rpm and 303 K. Once adapted to $^{2}H_2O$, the cells were transferred to a modified $^{2}H_2O$ minimal medium containing 5.4 g/l ^{2}H-^{13}C-D-glucose (Cambridge Isotope Labs) 2 g/l $^{15}NH_4SO_4$ (Sigma Aldrich), 15.8 g/L

Na_2HPO_4, and 4.8 g/l KH_2PO_4. After growth in the triple-labeled medium, the culture was scaled up to 500 ml split equally between two 2.8 l baffled Fernbach flasks for sufficient aeration. After the culture reached OD_{600}~1.9, IPTG was added to 0.5 mM. After 6 hours of induction, the cells were harvested by centrifugation, and the pellets were stored at 193 K until further use. For purification, the frozen cells were thawed and resuspended in lysis buffer (50 mM MES pH 6.5, 1 mM EDTA) containing SigmaFast™ EDTA-free protease inhibitor cocktail tablets (Sigma Aldrich). The cells were then lysed on ice by sonication with a Branson 450D Digital Sonifier (Emerson Industrial Automation, St. Louis, MO, USA) and clarified by centrifugation at 34000 x g for 45 min at 277 K. Following centrifugation, the supernatant was filtered through 1.6 µm and 0.45 µm GD/X syringe filters (GE Healthcare, Pittsburgh, PA, USA) as well as diluted with 20 mM MES, pH 6.5 buffer. The crude protein solution was then loaded onto a 5 ml HiTrap SP Sepharose FF column (GE Healthcare) that had been previously equilibrated with Buffer A (20 mM MES, pH 6.5). A linear gradient of Buffer B (20 mM MES, 300 mM NaCl pH 6.5) was used to elute the protein at approximately 20-30 mM NaCl at room temperature. Using UV absorbance measurements and SDS-PAGE analysis, fractions containing pure protein were selected, pooled, and concentrated using a 10 K MWCO Vivaspin 15R concentrator (Sartorius, Gottingen, Germany). Concentrated protein was loaded onto a 120-ml HiPrep Sephacryl S-100 HR gel filtration column (GE Healthcare) pre-equilibrated with Buffer A. Using UV absorbance and SDS-PAGE analysis, fractions containing pure protein were again selected, pooled, and concentrated using a 10 K MWCO Vivaspin 15R concentrator (Sartorius). The protein concentration was estimated using UV absorbance.

7.2.2 NMR Relaxation Experiments

^{15}N-backbone amide longitudinal relaxation rates (^{15}N-R_1) and in-phase transverse relaxation rates (^{15}N-R_2) of the isolated and inhibitor-bound ([inhibitor]/[protein]=1.2) Toho-1 β-lactamase were measured on the uniformly ^{2}H-^{15}N labeled protein at 700 MHz (^{1}H) and 25 ºC using gradient-enhanced phase-sensitive HSQC pulse sequences containing a water flip-back pulse. Delays of 20, 60, 100, 200, 400, 600, 800, and 1200 ms were used for ^{15}N-R_1 experiments, and Carr-Purcell-Maiboom-Gill (CPMG) loop lengths of 16, 32, 64, 128, 160, 192, 224, and 256 ms were used for ^{15}N-R_2 measurements. Within the interleaved CPMG loop, the delay between the ^{15}N 180º pulses was 900 µs, and the ^{15}N pulse width was 100 µs. The ^{15}N-R_1 and ^{15}N-R_2 experiments were conducted with 8 transients per increment and an interscan delay of 2 s and 1 s, respectively. Best-fit relaxation rates were calculated by non-linear least-squares fitting of ^{15}N-^{1}H cross-peak intensities to a two-parameter single-exponential decay in NMRFAM-SPARKY.

{^{1}H}-^{15}N steady-state NOEs were obtained by measuring the percent change in intensity for spectra obtained with and without ^{1}H pre-saturation. Saturation was obtained by applying cycles consisting of a 120º pulse (pulse width 14.2 µs) followed by a 5 ms delay during the 10 s recycle (relaxation) delay. The spectral widths were identical to those for ^{15}N-R_1 and R_2 experiments, and all relaxation experiments were performed with interleaved acquisition.

CPMG relaxation dispersion experiments were performed as detailed by Tollinger (Tollinger et al. 2001), with blocks of ^{15}N 180° pulses resulting in effective fields of frequencies (υ_{CPMG}) of 25, 50, 100, 250, 500, 750, and 1000 s^{-1} in a 40-ms constant time echo period. The carrier was set at 117 ppm and data were acquired in an interleaved

manner with a recycle delay of 2 s. Effective transverse relaxation rates (R_2') were calculated from the peak heights measured at the end of the constant CPMG period, and R_{ex} was calculated as the difference between R_2' at υ_{CPMG} = 25 s^{-1} and R_2' at υ_{CPMG} = 1,000 s^{-1}. Uncertainty in R_2' was calculated based on the noise measured in a reference spectrum acquired with no CPMG block (Ishima and Torchia 2005).

7.2.3 CSP Titrations and Estimation of K_d

Chemical shift perturbations were used to estimate K_d by recording a series of ^{1}H-^{15}N HSQC experiments with varying concentrations of the inhibitor avibactam and a total protein concentration of 850 µM. After the initial protein-only experiments, avibactam was titrated to a final concentration of 10%, 20%, 80%, 120%, 150%, and 200% of the total protein concentration. The sample was equilibrated at 298 K for 5 minutes after each inhibitor addition. All data were processed and analyzed using the Mestrelab NMR software package (Mestrelab Research) and NMRLineGuru GUI (Feng et al. 2019; Pemberton et al. 2020).

7.3 Results and Discussions

7.3.1 Backbone Chemical Shift Assignments

The protein's modest α-helix content (approximately 40%) resulted in relatively well-dispersed ^{15}N-^{1}H correlations in the HSQC spectrum. The HSQC backbone assignments were used to perform the classic solution state relaxation dispersion experiments and calculate R_1, R_2 and NOE values with and without the incorporation of Avibactam. The relaxation measurements are detailed in the following sections.

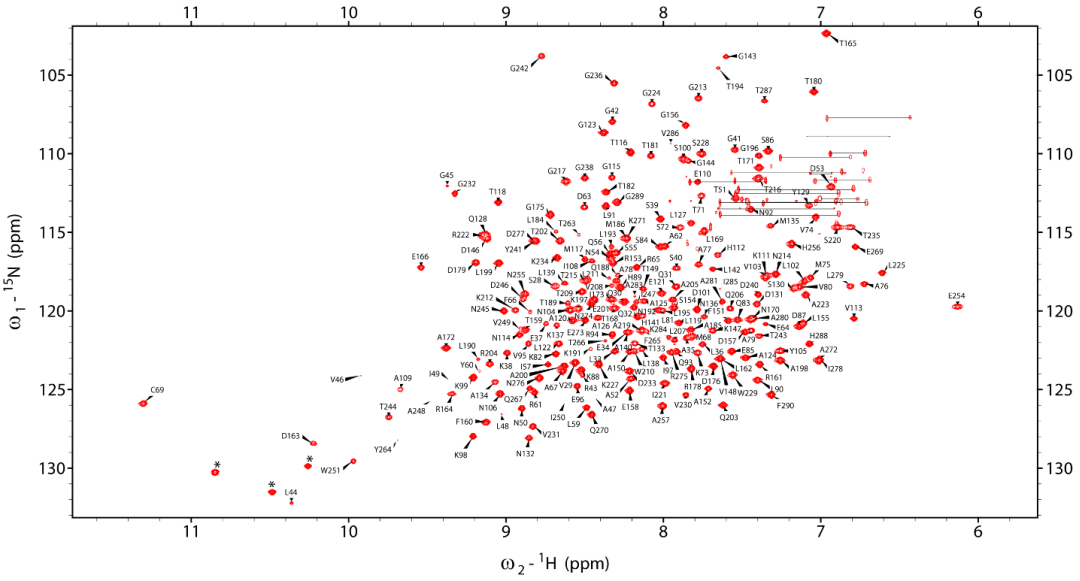


Fig 7.3 Two-dimensional ^{15}N HSQC spectrum of U-^{2}H-^{13}C-^{15}N Toho-1 β-lactamase at 700 MHz (^{1}H), 298K, pH 6.5, and a protein concentration of 850 µM. The 237 labels indicate backbone amide ^{1}H and ^{15}N resonances that could be assigned with probability > 99%. Peaks connected by thin lines (upper right) correspond to correlations from the asparagine and glutamine side chain -NH_2 groups; peaks marked with an asterisk (bottom and bottom left) correspond to the sidechain correlations of arginine (folded) and to H^{ε}-N^{ε} correlations of tryptophan. (Adapted from Sakhrani et al., 2021).

N-H correlations for the conserved active site residues S70 and S237 were not observed in the ^{15}N HSQC spectrum of the ligand-free enzyme. This was also true in the related Class-A β-lactamases PSE-4 and TEM-1 (Morin and Gagné 2009). Signal broadening due to either conformational dynamics inherent to the enzyme's active site or the absence of stabilizing interactions with a substrate could be responsible for the missing resonances. A complete absence of assignments was also observed for residues 258 to 262, presumably due to insufficient ^{1}H back exchange as these residues have modeled accessibility measures of zero (Esque et al. 2011).

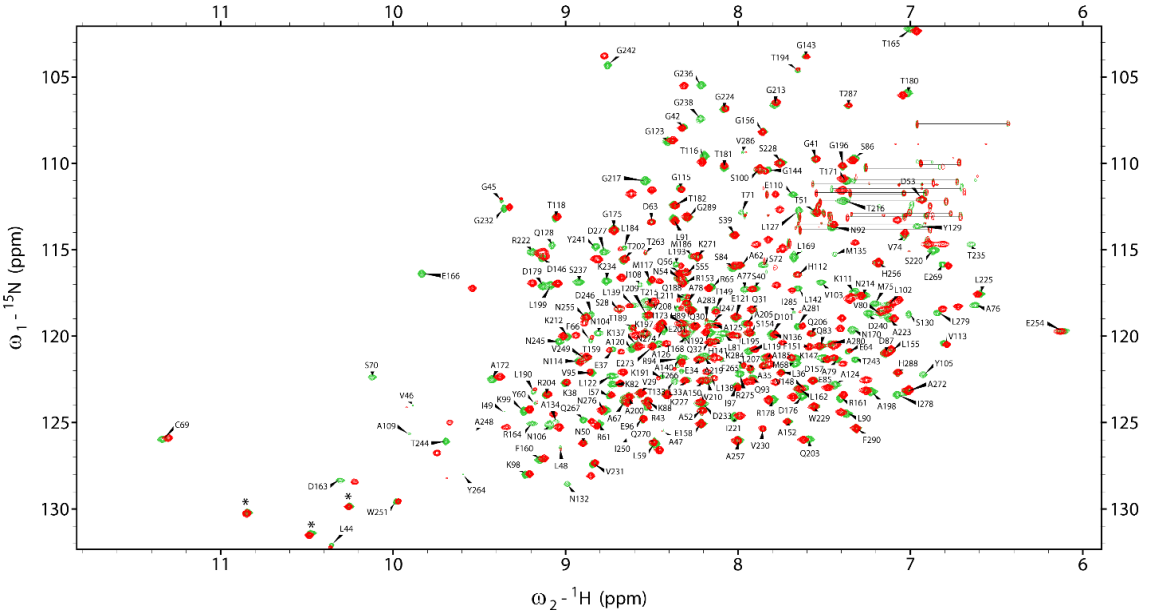


Fig 7.4 Overlap of the two-dimensional ^{1}H-^{15}N HSQC spectrum of 850 µM, uniformly ^{2}H-^{13}C-^{15}N labeled Toho-1 β-lactamase at 700 MHz, 298K and pH 6.5 before (red) and after (green) the addition of avibactam to a final concentration of ~ 1020 µM, giving [ligand]/[protein] ratio = 1.2. The 237 labels indicate backbone amide ^{1}H and ^{15}N correlations that could be sequentially assigned with probability >99%. Peaks at the same nitrogen but different hydrogen ppm values in the upper right-hand corner correspond to the asparagine and glutamine side chain -NH_2 groups and are shown connected by thin lines. Peaks marked with an asterisk in the lower half of the spectrum are the peaks corresponding to arginine and tryptophan side chain H^{ε}-N^{ε} correlations (folded). The folded peaks have their actual nitrogen shifts at about 87 ppm. (Adapted from Sakhrani et al., 2021).

In the presence of avibactam, the HSQC peak intensities for the active site residues K73 and D131 reduced to zero with increasing concentration of avibactam; at a 20% excess of avibactam, resonances for these residues were not observed. We ascribe this intensity loss to exchange broadening in the active site, which is supported by both neutron and X-ray crystal structures that show dual conformations of several key active site residues when a substrate is bound, including K73 and Y105 (Langan et al. 2018). Residues K73 and Y105 form several close contacts with residues Y129, S130, and N132, which dictate conformational changes of the amide group of D131. At the same time, strong signals for the ^{15}N-^{1}H correlations for residue S70 and S237 appeared in the

spectrum for the bound protein with the amide proton (H^N) shift of S70 significantly downfield (10.12 ppm). Avibactam covalently linked to the side chain of S70 via a carbamoyl linkage presumably stabilizes this residue and eliminates exchange broadening, while the downfield shift is attributed to hydrogen bonding between the backbone amide and the strongly electronegative carbamoyl group. In the X-ray structure of avibactam bound to CTX-M-15 (PDB code: 4HBU), a closely related class A member, residue S237 forms a hydrogen bond to the carboxamide of avibactam and interacts with the sulfate group (Lahiri et al. 2013).

7.3.2 Chemical Shift Perturbations and the Ligand Binding Site

The addition of avibactam led to chemical shift perturbations (CSPs) (Williamson 2013) in the ^{15}N HSQC spectrum of Toho-1 β-lactamase, as shown in Fig 7.4. The CSPs were quantified as the average chemical shift perturbation $\Delta\delta_{NH}$ (in ppm) based on the weighted root-mean-square of the proton and nitrogen chemical shift perturbations ($\Delta\delta_H$ and $\Delta\delta_N$, respectively), according to

$$\Delta\delta_{NH} = \sqrt{\frac{\Delta\delta_H^2 + 0.14\Delta\delta_N^2}{2}} \quad (4)$$

where the factor of 0.14 accounts for differences in sensitivity of the amide proton and nitrogen backbone resonances (Schumann et al. 2007). These are summarized in Fig 7.5.

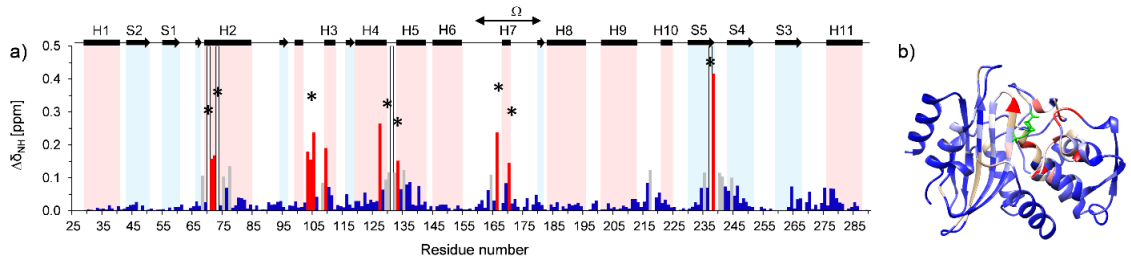


Fig 7.5 (a) Average chemical shift perturbations ($\Delta\delta_{NH}$) calculated for residues of Toho-1 β-lactamase at a protein concentration of 850 μM and a [inhibitor]/[protein] ratio of 1.2. Red indicates residues that show relatively large CSPs, gray indicates residues with moderate CSPs, while blue indicates residues with low CSPs; hollow bars represent residues whose correlations either become indiscernible or appear upon the addition of avibactam at the final concentration of 1.2 times that of the protein. Asterisks are placed over active-site residues S70, K73, N104, S130, N132, E166, N170, and S237 that participate in catalysis and display relatively large CSPs themselves or in their sequential neighborhood. The secondary structure elements from the X-ray crystal structure are indicated across the top (bars α-helices, arrows β-strands). (b) Mapping of the avibactam binding site onto the Toho-1 β-lactamase crystal structure 2ZQ8 from CSPs. The coordinates of avibactam were extracted from crystal structure 4HBU after alignment with 2ZQ8.

The most significant CSPs were found for residues S70, K73, N104, S130, N132, E166, N170, and S237, marked with asterisks in Fig 7.5a. The relatively large CSPs observed for these residues allowed for the mapping of the ligand-binding site (Fig 7.5b). The CSPs can be rationalized from ligand atom interactions with catalytic residues, as observed in the X-ray crystal structure of the related protein CTX-M-15 complexed with avibactam (PDB ID: 4HBU) and 100 ns MD simulation of the CTX-M-15 covalent complex (Lahiri et al. 2013; Lizana and Delgado 2019). Analysis of the MD trajectory showed that residues C69, N104, N132, E166, and N170 make contact with the carbonyl oxygen of the acylamide side chain of avibactam and that residues K234, S237, T235, and R274 make polar contacts to the negatively charged sulfate group. The side chain of S130 hydrogen bonds with the amine linkage to the sulfate group, and the piperidine ring forms a hydrophobic stack with the Y105 side chain. The range of CSPs observed for the binding of avibactam to Toho-1 β-lactamase was small, possibly because the sulfate

group of avibactam can adopt a range of bound conformations, as observed in the MD trajectory, and therefore the perturbations in chemical shifts are averaged.

Ligand-titration experiments, in which the ^{15}N-HSQC chemical shift perturbations were followed as a function of the total added inhibitor, were also performed to quantify the dissociation constant (K_d) as described in the ESM. The majority of the CSPs showed a characteristic very tight binding titration curve. These binding curves signified slow exchange, 1:1 binding, and allowed only an upper limit of the binding constant to be estimated, K_d < 1.0 μM (Fielding 2007; Williamson 2013). These results are consistent with an essentially irreversible (on the NMR timescale) covalent bond being formed between inhibitor and protein. Isothermal titration calorimetry measurements for the related TEM-1with avibactam reported $K_{d,}$ to be in the lower nanomolar range (Ehmann et al. 2012).

7.3.3 Backbone Dynamics from Analysis of Relaxation Experiments

The high quality of the NMR spectra allowed the measurement of site-specific ^{15}N-R_1, R_2, and {^{1}H}-^{15}N NOE for 236 of the 263 residues before and after the addition of avibactam; these initial experiments were performed at a single field of 16.44 T. The data are summarized in Fig 7.6 and provide a window into the backbone dynamics of the protein scaffold (Kay 1998). As is typical, the loop regions and the N- and C-termini were found to be more dynamic on the pico- to nanosecond timescale than the rest of the protein, displaying increased ^{15}N-R_1 relaxation rates and decreased ^{15}N-R_2 relaxation rates. At the same time, increased dynamics on the micro- to millisecond timescale could be identified for residues in β-strands S1, S2, and S3, and α-helices H4, H8, and H11 based on increased ^{15}N-R_2 relaxation rates as highlighted in the plot of the ^{15}N-R_2/R_1 ratio (Fig 7.6d). Residues exhibiting large deviations in the ^{15}N-R_2/R_1 plot due to

anisotropic motion could be separated from those undergoing residual exchange broadening by plotting the product of the relaxation rates, $R_2 \times R_1$, (Fig 7.6e) as motional anisotropy affects ^{15}N-R_1 and ^{15}N-R_2 in opposite ways (Kneller et al. 2002).

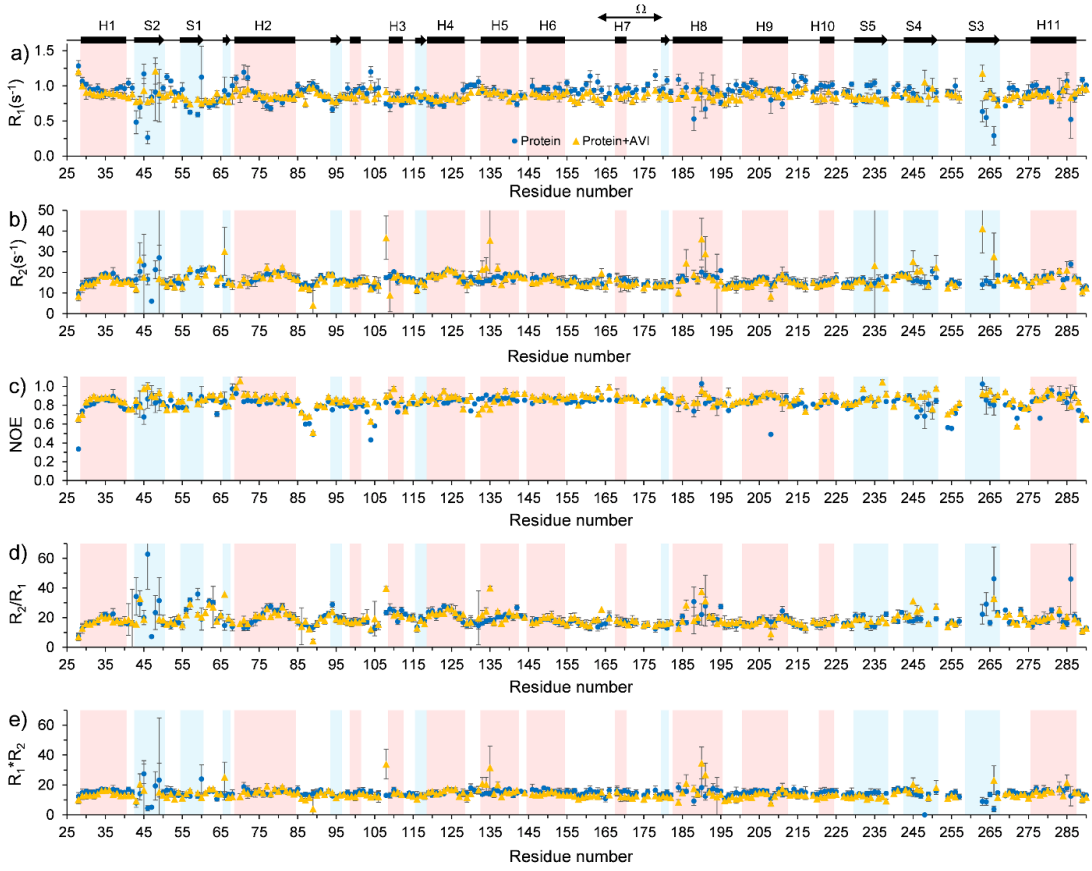


Fig 7.6 Experimental (a) ^{15}N-R_1 and (b) ^{15}N-R_2 relaxation times and (c){^{1}H}-^{15}N NOE for residues of Toho-1 β-lactamase with (orange triangles) and without (blue circles) avibactam at 700 MHz, 298K, a protein concentration of 850 μM, and [inhibitor]/[protein] ratio of 1.2. (d) The ratio of residue-wise relaxation rates ^{15}N-R_2/R_1, and (e) their products ^{15}N-$R_2 \times R_1$.

For the active site residues, the increase in the amplitude of dynamics on the pico- to nanosecond timescale is accompanied by a concomitant increase in dynamics on the slower micro- to millisecond times suggested by the loss of the HSQC peak intensities for the active site residues K73 and D131. This was further examined by initial CPMG relaxation dispersion (CPMG RD) experiments performed again at a single field of 16.4

T (700 MHz ^{1}H). The CPMG-RD exchange rate, R_{ex}(CPMG-RD), was calculated as the difference in apparent transverse relaxation rate (R_2') measured at effective fields of ν_{CPMG} = 25 Hz and 1000 Hz. Significant refocusing of magnetization under CPMG occurs when the CPMG rate is greater than twice the rate of the dynamics (Kleckner and Foster 2011), making R_{ex}(CPMG-RD) as defined here most sensitive to millisecond timescale motions. As seen in Fig 7.7, there is a stark increase in millisecond timescale dynamics within the active site upon binding of avibactam. Millisecond timescale motions are expected to play an important role in catalytic promiscuity.

Three additional class-A β lactamases have been characterized with MFA: TEM-1, PSE-4, and BlaC (Savard and Gagné 2006; Morin and Gagné 2009; Fisette et al. 2010; Elings et al. 2020). This class of enzymes shows similar dynamic features, including being well-ordered overall and having active sites that are mostly rigid on the pico- to nanosecond timescale, while displaying microsecond-to-millisecond motions for key catalytic residues such as S70 and S237 (or their flanking residues). It was originally proposed for TEM-1 that the high degree of order for the apo form could indicate that catalytic residues are appropriately positioned to play their role as soon as the substrate enters the active site (Savard and Gagné 2006). Such a high degree of ordering may also help mitigate any potentially unfavorable changes in conformational entropy upon substrate binding. Both the Toho-1 β-lactamase and BlaC have been characterized in complex with the inhibitor avibactam. Notably the active sites of both display larger amplitude dynamics on the microsecond-millisecond timescales upon inhibitor binding. These induced active site dynamics are thought to assist catalytic promiscuity by accommodating variously sized substrates and increasing the probability of sampling the optimal conformation for hydrolysis of the bound ligand (Elings et al. 2020). These

slower motions may also be relevant for catalysis by facilitating substrate entry and product release.

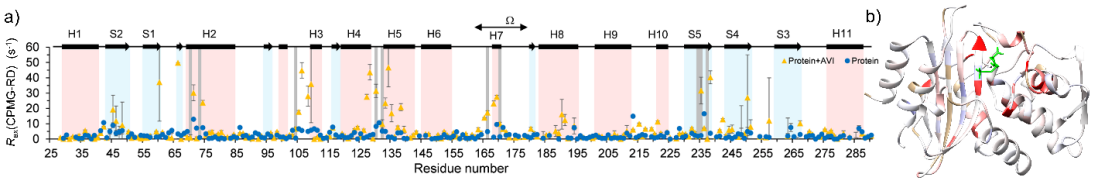


Fig 7.7. (a) CPMG relaxation dispersion exchange rates for residues of the Toho-1 β-lactamase with (orange triangles) and without (blue circles) avibactam at 700 MHz (^{1}H) and 298K from. The CPMG-RD exchange rate is calculated as the difference between R_2' at υ_{CPMG} = 25 s^{-1} and R_2' at υ_{CPMG} = 1,000 s^{-1} and reports primarily on millisecond timescale motions. Grey bars indicate active site residues that interact with avibactam. (b) Changes in CPMG exchange rates upon binding of avibactam mapped onto the Toho-1 β-lactamase crystal structure 2ZQ8. Red indicates residues that become more dynamic upon inhibitor addition, blue indicates residues that rigidify, and white indicates no change.

7.4 Conclusion

Sequential assignments of backbone C, N, and H^N chemical shifts for the Toho-1 β-lactamase R274N/R276N double mutant were accomplished using 2D and 3D TROSY-based NMR experiments. The backbone assignments permitted the prediction of secondary structure that compares well to the structure obtained from X-ray diffraction. The chemical shift assignments also enabled the site-specific characterization of inhibitor binding. Titration with the non β-lactam β-lactamase inhibitor avibactam was found to give perturbations in backbone chemical shifts indicative of slow exchange, 1:1 tight binding that are consistent with essentially irreversible covalent binding of the inhibitor on the NMR timescale. The extended lifetime of the inhibitor-enzyme complex allowed site-specific relaxation rates to be measured. Model-free analysis of the ^{15}N relaxation data revealed well-ordered structures for the ligand-free and avibactam-bound forms with generalized order parameters of 0.86 and 0.85, respectively. The high degree of ordering on the pico- to nanosecond timescale for the free enzyme is thought

to allow precise positioning of the sidechains for catalysis and help mitigate any potentially unfavorable changes in conformational entropy upon substrate binding. Complementary relaxation dispersion experiments show that upon substrate binding, there is actually a significant upturn in active site motions on the millisecond timescale. Promiscuous enzymes are expected to exhibit active-site flexibility to help accommodate variously sized substrates and increase the probability of sampling the optimal conformation for catalysis. The combination of high rigidity on short timescales and active site flexibility on longer timescales is consistent with the hypotheses of dynamical duality: the Toho-1 β-lactamase is rigid on the pico- to nanosecond timescale but exhibits flexibility on the millisecond timescale in and around the active site, as observed for other members of the class-A β-lactamases.

The assignment of backbone chemical shifts and the characterization of dynamics of the Toho-1 β-lactamase highlight a need for complementary data focusing on the protonation states of the catalytic residues' sidechains (Damblon et al. 1996; Pemberton et al. 2020). This motivates the ongoing investigation of the detailed chemical structure for Toho-1 β-lactamase using NMR and other techniques to better understand the catalytic cycle of the modern members of Class-A β-lactamases.

7.5 References

Abragam A (1961) The Principles of Nuclear Magnetism. Clarendon Press, Oxford

Ambler RP, Coulson AFW (1991) A standard numbering scheme for the Class A beta-lactamases. Biochem J 276:1990–1991

Bahrami A, Assadi AH, Markley JL, Eghbalnia HR (2009) Probabilistic Interaction Network of Evidence Algorithm and its Application to Complete Labeling of Peak Lists from Protein NMR Spectroscopy. PLoS Comput Biol 5:1–15 . https://doi.org/10.1371/journal.pcbi.1000307

Banerjee S, Pieper U, Kapadia G, Pannell LK, Herzbere O (1998) Role of the Ω-loop in the activity, substrate specificity, and structure of class A β-lactamase. Biochemistry 37:3286–3296 . https://doi.org/10.1021/bi972127f

Barbato G, Ikura M, Kay LE, Pastor RW, Bax A (1992) Backbone Dynamics of Calmodulin Studied by 15N Relaxation Using Inverse Detected Two-Dimensional NMR Spectrocopy: The Central Helix Is Flexible. Biochemistry 31:5269–5278

Bebrone C, Lassaux P, Vercheval L, Sohier JS, Jehaes A, Sauvage E, Galleni M (2010) Current challenges in antimicrobial chemotherapy: Focus on β-lactamase inhibition. Drugs 70:651–679 . https://doi.org/10.2165/11318430-000000000-00000

Best RB, Clarke J, Karplus M, Pasteur L (2005) What Contributions to Protein Side-chain Dynamics are Probed by NMR Experiments ? A Molecular Dynamics Simulation Analysis. J Mol Biol 349:185–203 . https://doi.org/10.1016/j.jmb.2005.03.001

Bloom M, Reeves LW, Wells EJ (1965) Spin Echoes and Chemical Exchange. J Chem Phys 42:1615–1624 . https://doi.org/10.1063/1.1696168

Bush K (2010) Bench-to-bedside review: The role of β-lactamases in antibiotic-resistant Gram-negative infections. Crit Care 14:224–232 . https://doi.org/10.1186/cc8892

Bush K (2013) The ABCD's of β-lactamase nomenclature. J Infect Chemother 19:549–559 . https://doi.org/10.1007/s10156-013-0640-7

Cavanagh J, Fairbrother W, Palmer III A, Rance M, Skelton N (2007) Protein NMR Spectroscopy: principles and practice, 2nd edn. Elsevier Inc., San Diego

Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM (1990) Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins. J Am Chem Soc 112:4989–4991 . https://doi.org/10.1021/ja00168a070

Cortina GA, Kasson PM (2016) Excess positional mutual information predicts both local and allosteric mutations affecting beta lactamase drug resistance. Bioinformatics

32:3420–3427 . https://doi.org/10.1093/bioinformatics/btw492

d'Auvergne EJ, Gooley PR (2003) The use of model selection in the model-free analysis of protein dynamics. J Biomol NMR 25:25–39 . https://doi.org/10.1023/A:1021902006114

Damblon C, Raquet X, Lian L, Lamotte-brasseur J, Fonze E, Charlier P, Roberts GCK, Frère J (1996) The Catalytic Mechanism of β -Lactamases : NMR Titration of an Active-Site Lysine Residue of the TEM-1 Enzyme. PNAS 93:1747–1752

Drawz SM, Bonomo RA (2010) Three decades of β-lactamase inhibitors. Clin Microbiol Rev 23:160–201 . https://doi.org/10.1128/CMR.00037-09

Ehmann DE, Jahić H, Ross PL, Gu RF, Hu J, Kern G, Walkup GK, Fisher SL (2012) Avibactam is a covalent, reversible, non-β-lactam β-lactamase inhibitor. Proc Natl Acad Sci U S A 109:11663–11668 . https://doi.org/10.1073/pnas.1205073109

Elings W, Gaur A, J. Blok A, Timmer M, van Ingen H, Ubbink M (2020) ℕ-Lactamase of Mycobacterium tuberculosis Shows Dynamics in the Active Site That Increase upon Inhibitor Binding Wouter. Antimicrob Agents Chemother 64:1–13

Esque J, Oguey C, Brevern AG De, Pontoise DC, Chauvin A (2011) Comparative Analysis of Threshold and Tessellation Methods for Determining Protein Contacts. J Chem Inf Model 51:493–507 . https://doi.org/10.1021/ci100195t

Feng C, Kovrigin EL, Post CB (2019) NmrLineGuru: Standalone and User-Friendly GUIs for Fast 1D NMR Lineshape Simulation and Analysis of Multi-State Equilibrium Binding Models. Sci Rep 9:1–14 . https://doi.org/10.1038/s41598-019-52451-8

Fernández C, Wider G (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. Curr Opin Struct Biol 13:570–580 . https://doi.org/10.1016/j.sbi.2003.09.009

Fielding L (2007) NMR methods for the determination of protein – ligand dissociation constants. Princ Nucl Magnaetic Reson Spectrosc 51:219–242 . https://doi.org/10.1016/j.pnmrs.2007.04.001

Fisette O, Gagné SM, Lagüe P (2012) Molecular dynamics of class A β-lactamases - Effects of substrate binding. Biophys J 103:1790–1801 . https://doi.org/10.1016/j.bpj.2012.09.009

Fisette O, Morin S, Savard PY, Lagüe P, Gagné SM (2010) TEM-1 backbone dynamics - Insights from combined molecular dynamics and nuclear magnetic resonance. Biophys J 98:637–645 . https://doi.org/10.1016/j.bpj.2009.08.061

Gu Y, Li D-W, Brüschweiler R (2014) NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. J Chem Theory Comput 10:2599–2607 . https://doi.org/10.1021/ct500181v

Hart KM, Ho CMW, Dutta S, Gross ML, Bowman GR (2016) Modelling proteins' hidden conformations to predict antibiotic resistance. Nat Commun 7:1–10 . https://doi.org/10.1038/ncomms12965

Helfand MS, Totir MA, Carey MP, Hujer AM, Bonomo RA, Carey PR (2003) Following the Reactions of Mechanism-Based Inhibitors with β-Lactamase by Raman Crystallography. Biochemistry 42:13386–13392 . https://doi.org/10.1021/bi035716w

Huang J, Mackerell AD (2013) CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. J Comput Chem 34:2135–2145 . https://doi.org/10.1002/jcc.23354

Humphrey W, Dalke A, Schulten K (1996) VMD: Visual Molecular Dynamics. J Mol Graph 14:33–38 . https://doi.org/10.1016/0263-7855(96)00018-5

Ishii Y, Ohno A, Taguchi H, Imajo S, Ishiguro M, Matsuzawa H (1995) Cloning and Sequence of the Gene Encoding a Cefotaxime-Hydrolyzing Class A β-lactamase Isolated from Escherichia coli. Antimicrob Agents Chemother 39:2269–2275 . https://doi.org/10.1128/AAC.39.10.2269

Ishima R, Torchia DA (2005) Error estimation and global fitting in transverse-relaxation dispersion experiments to determine chemical-exchange parameters. J Biomol NMR 32:41–54 . https://doi.org/10.1007/s10858-005-3593-z

Jacoby GA (2006) MINIREVIEW Beta-Lactamase Nomenclature. Antimicrob Agents Chemother 50:1123–1129 . https://doi.org/10.1128/AAC.50.4.1123

Kay LE (1998) Protein dynamics from NMR. Biochem Cell Biol 76:145–152 . https://doi.org/10.1038/755

Kay LE, Torchia DA, Bax A (1989) Backbone Dynamics of Proteins As Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. Biochemistry 28:8972–8979

King DT, King AM, Lal SM, Wright GD, Strynadka NCJ (2016) Molecular Mechanism of Avibactam-Mediated β-Lactamase Inhibition. ACS Infect Dis 1:175–184 . https://doi.org/10.1021/acsinfecdis.5b00007

Kleckner IR, Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. Biochim Biophys Acta - Proteins Proteomics 1814:942–968 . https://doi.org/10.1016/j.bbapap.2010.10.012

Kneller JM, Lu M, Bracken C (2002) An Effective Method for the Discrimination of Motional Anisotropy and Chemical Exchange. J Am Chem Soc 124:1852–1853 . https://doi.org/10.1021/ja017461k

Kong KF, Schneper L, Mathee K (2010) Beta-lactam antibiotics: From antibiosis to resistance and bacteriology. APMIS 118:1–36 . https://doi.org/10.1111/j.1600-

0463.2009.02563.x

Lagacé-Wiens P, Walkty A, Karlowsky JA (2014) Ceftazidime-avibactam: An evidence-based review of its pharmacology and potential use in the treatment of Gram-negative bacterial infections. Core Evid 9:13–25 . https://doi.org/10.2147/CE.S40698

Lahiri SD, Mangani S, Durand-Reville T, Benvenuti M, De Luca F, Sanyal G, Docquierd J-D (2013) Structural Insight into Potent Broad-Spectrum Inhibition with Reversible Recyclization Mechanism : Avibactam in Complex with CTX-M-15 and Pseudomonas aeruginosa AmpC β-Lactamases. Antimicrob Agents Chemother 57:2496–2505 . https://doi.org/10.1128/AAC.02247-12

Langan PS, Sullivan B, Weiss KL, Coates L (2020) Probing the role of the conserved residue Glu166 in a class A β-lactamase using neutron and X-ray protein crystallography. Acta Crystallogr Sect D Struct Biol 76:118–123 . https://doi.org/10.1107/S2059798319016334

Langan PS, Vandavasi VG, Cooper SJ, Weiss KL, Ginell SL, Parks JM, Coates L (2018) Substrate Binding Induces Conformational Changes in a Class A β-lactamase That Prime It for Catalysis. ACS Catal 8:2428–2437 . https://doi.org/10.1021/acscatal.7b04114

Langan PS, Vandavasi VG, Weiss KL, Cooper JB, Ginell SL, Coates L (2016) The structure of Toho1 β-lactamase in complex with penicillin reveals the role of Tyr105 in substrate recognition. FEBS Open Bio 6:1170–1177 . https://doi.org/10.1002/2211-5463.12132

Lee W, Bahrami A, Dashti HT, Eghbalnia HR, Tonelli M, Westler WM, Markley JL (2019) I-PINE web server: an integrative probabilistic NMR assignment system for proteins. J Biomol NMR 73:213–222 . https://doi.org/10.1007/s10858-019-00255-3

Lee W, Markley JL (2018) PINE-SPARKY.2 for automated NMR-based protein structure research. Bioinformatics 34:1586–1588 . https://doi.org/10.1093/bioinformatics/btx785

Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31:1325–1327 . https://doi.org/10.1093/bioinformatics/btu830

Lipari G, Szabo A (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory And Range Of Validity. J Am Chem Soc 104:4546–4559 . https://doi.org/10.1002/chin.198247060

Lizana I, Delgado EJ (2018) Theoretical insights on the inhibition mechanism of a class A Serine Hydrolase by avibactam. J Comput Chem 39:1943–1948 . https://doi.org/10.1002/jcc.25340

Lizana I, Delgado EJ (2019) Molecular Insights on the Release of Avibactam from the Acyl-Enzyme Complex. Biophys J 116:1650–1657 . https://doi.org/10.1016/j.bpj.2019.03.027

Lobkovsky E, Moews PC, Liu H, Zhao H, Frere JM, Knox JR (1993) Evolution of an enzyme activity: Crystallographic structure at 2-Å resolution of cephalosporinase from the ampC gene of Enterobacter cloacae P99 and comparison with a class a penicillinase. Proc Natl Acad Sci U S A 90:11257–11261 . https://doi.org/10.1073/pnas.90.23.11257

Mandel AM, Akke M, Palmer AG (1995) Backbone dynamics of Escherichia coli ribonuclease HI: Correlations with structure and function in an active enzyme. J Mol Biol 246:144–163 . https://doi.org/10.1006/jmbi.1994.0073

Mielke S, Krishnan V (2009) Characterization of protein secondary structure from NMR chemical shifts. Prog Nucl Magn Reson Spectrosc 54:141–165 . https://doi.org/10.1016/b978-1-4832-2943-0.50008-x

Morin S, Gagné SM (2009) NMR dynamics of PSE-4 β-lactamase: An interplay of ps-ns order and μs-ms motions in the active site. Biophys J 96:4681–4691 . https://doi.org/10.1016/j.bpj.2009.02.068

Ortega A, Amoros D, Garcia de la Torre J (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. 101:892–898 . https://doi.org/10.1016/j.bpj.2011.06.046

Pastor N, Amero C (2015) Information flow and protein dynamics: the interplay between nuclear magnetic resonance spectroscopy and molecular dynamics simulations. Front Plant Sci 6:1–13 . https://doi.org/10.3389/fpls.2015.00306

Patel MP, Hu L, Stojanoski V, Sankaran B, Prasad BVV, Palzkill T (2017) The Drug-Resistant Variant P167S Expands the Substrate Profile of CTX-M β-Lactamases for Oxyimino-Cephalosporin Antibiotics by Enlarging the Active Site upon Acylation. Biochemistry 56:3443–3453 . https://doi.org/10.1021/acs.biochem.7b00176

Pemberton OA, Noor RE, Vasantha Kumar M V., Sanishvili R, Trent Kemp M, Kearns FL, Lee Woodcock H, Gelis I, Chen Y (2020) Mechanism of proton transfer in class A β-lactamase catalysis and inhibition by avibactam. Proc Natl Acad Sci U S A 117:5818–5825 . https://doi.org/10.1073/pnas.1922203117

Petrosino JF, Palzkill T (1996) Systematic Mutagenesis of the Active Site Omega Loop of TEM-l β-Lactamase. J Bacteriol 178:1821–1828 . https://doi.org/10.1128/jb.178.7.1821-1828.1996

Petrovic D, Risso VA, Kamerlin SCL, Sanchez-Ruiz JM (2018) Conformational dynamics and enzyme evolution. J R Soc Interface 15: . https://doi.org/10.1098/rsif.2018.0330

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE

(2004) UCSF Chimera - A Visualization System for Exploratory Research and Analysis. J Comput Chem 25:1605–1612 . https://doi.org/10.1002/jcc.20084

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable Molecular Dynamics with NAMD. J Comput Chem 26:1781–1802 . https://doi.org/10.1002/jcc.20289

Riek R, Pervushin K, Wüthrich K (2000) TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. Trends Biochem Sci 25:462–468 . https://doi.org/10.1016/S0968-0004(00)01665-0

Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. J Comput Phys 23:327–341 . https://doi.org/10.1016/0021-9991(77)90098-5

Sakhrani, V. V., Ghosh, R. K., Hilario, E., Weiss, K. L., Coates, L., & Mueller, L. J. (2021). Toho-1 β-lactamase: backbone chemical shift assignments and changes in dynamics upon binding with avibactam. Journal of biomolecular NMR, 75, 303-318.

Savard PY, Gagné SM (2006) Backbone dynamics of TEM-1 determined by NMR: Evidence for a highly ordered protein. Biochemistry 45:11414–11424 . https://doi.org/10.1021/bi060414q

Schumann FH, Riepl H, Maurer T, Gronwald W, Neidig KP, Kalbitzer HR (2007) Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. J Biomol NMR 39:275–289 . https://doi.org/10.1007/s10858-007-9197-z

Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J Biomol NMR 56:227–241 . https://doi.org/10.1007/s10858-013-9741-y

Shimamura T, Ibuka A, Fushinobu S, Wakagi T, Ishiguro M, Ishii Y, Matsuzaw H (2002) Acyl-intermediate structures of the Extended-spectrum Class A β-lactamase, Toho-1, in Complex with Cefotaxime, Cephalothin, and Benzylpenicillin. J Biol Chem 277:46601–46608 . https://doi.org/10.1074/jbc.M207884200

Shimamura T, Nitanai Y, Uchiyama T, Matsuzawa H (2009) Improvement of crystal quality by surface mutations of beta-lactamase Toho-1 crystallization communications. Acta Crystallogr Sect F Struct Biol Cryst Commun F65:379–382 . https://doi.org/10.1107/S1744309109008240

Shimizu A, Ikeguchi M, Sugai S (1994) Appropriateness of DSS and TSP as internal references for 1H NMR studies of molten globule proteins in aqueous media. J Biomol NMR 4:859–862 . https://doi.org/10.1007/BF00398414

Shin J, Lee W, Weontae L (2006) Structural proteomics by NMR spectroscopy.

Proteomics 5:171–185 . https://doi.org/10.1002/0470007745.ch9

Shirley M (2018) Ceftazidime-Avibactam: A Review in the Treatment of Serious Gram-Negative Bacterial Infections. Drugs 78:675–692 . https://doi.org/10.1007/s40265-018-0902-x

Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW (2016) CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. J Biomol NMR 66:111–124 . https://doi.org/10.1007/s10858-016-0060-y

Spyracopoulos L (2006) A suite of Mathematica notebooks for the analysis of protein main chain 15N NMR relaxation data. J Biomol NMR 36:215–224 . https://doi.org/10.1007/s10858-006-9083-0

Stojanoski V, Chow DC, Hu L, Sankaran B, Gilbert HF, Prasad BVV, Palzkill T (2015) A Triple Mutant in the Ω-loop of TEM-1 β-Lactamase Changes the Substrate Profile via a Large Conformational Change and an Altered General Base for Catalysis. J Biol Chem 290:10382–10394 . https://doi.org/10.1074/jbc.M114.633438

Tjandra N, Feller SE, Pastor RW, Bax A (1995) Rotational Diffusion Anisotropy of Human Ubiquitin from 15N NMR Relaxation. J Am Chem Soc 117:12562–12566 . https://doi.org/10.1021/ja00155a020

Tollinger M, Skrynnikov NR, Mulder FAA, Forman-Kay JD, Kay LE (2001) Slow Dynamics in Folded and Unfolded States of an SH3 Domain. J Am Chem Soc 123:11341–11352 . https://doi.org/10.1021/ja011300z

Tomanicek SJ, Blakeley MP, Cooper J, Chen Y, Afonine P V., Coates L (2010) Neutron Diffraction Studies of a Class A β-lactamase Toho-1 E166A/R274N/R276N Triple Mutant. J Mol Biol 396:1070–1080 . https://doi.org/10.1016/j.jmb.2009.12.036

Tomanicek SJ, Standaert RF, Weiss KL, Ostermann A, Schrader TE, Ng JD, Coates L (2013) Neutron and X-ray Crystal Structures of a Perdeuterated Enzyme Inhibitor Complex Reveal the Catalytic Proton Network of the Toho-1 β-Lactamase for the Acylation Reaction. J Biol Chem 288:4715–4722 . https://doi.org/10.1074/jbc.M112.436238

Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VHA, Takebayashi Y, Spencer J (2019) β-Lactamases and β-Lactamase Inhibitors in the 21st Century. J Mol Biol 431:3472–3500 . https://doi.org/10.1016/j.jmb.2019.04.002

Vandavasi VG, Weiss KL, Cooper JB, Erskine PT, Tomanicek SJ, Ostermann A, Schrader TE, Ginell SL, Coates L (2016) Exploring the Mechanism of β-Lactam Ring Protonation in the Class A β-lactamase Acylation Mechanism Using Neutron and X-ray Crystallography. J Med Chem 59:474–479 . https://doi.org/10.1021/acs.jmedchem.5b01215

Verma D, Jacobs DJ, Livesay DR (2013) Variations within Class-A β-Lactamase

Physiochemical Properties Reflect Evolutionary and Environmental Patterns, but not Antibiotic Specificity. PLOS Comput Biol 9: . https://doi.org/10.1371/journal.pcbi.1003155

Wang L, Eghbalnia HR, Bahrami A, Markley JL (2005) Linear analysis of carbon-13 chemical shift differences and its application to the detection and correction of errors in referencing and spin system identifications. J Biomol NMR 32:13–22 . https://doi.org/10.1007/s10858-005-1717-0

Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc 73:1–16 . https://doi.org/10.1016/j.pnmrs.2013.02.001

Woessner DE (1962) Nuclear spin relaxation in ellipsoids undergoing rotational brownian motion. J Chem Phys 37:647–654 . https://doi.org/10.1063/1.1701390

Zou T, Risso VA, Gavira JA, Sanchez-Ruiz JM, Ozkan SB (2015) Evolution of Conformational Dynamics Determines the Conversion of a Promiscuous Generalist into a Specialist Enzyme. Mol Biol Evol 32:132–143 . https://doi.org/10.1093/molbev/msu281

Chapter 8

Insights into the Catalytic Mechanism of Tyrosine Phenol Lyase via DNP enhanced NMR Assisted Crystallography.

8.1 Introduction

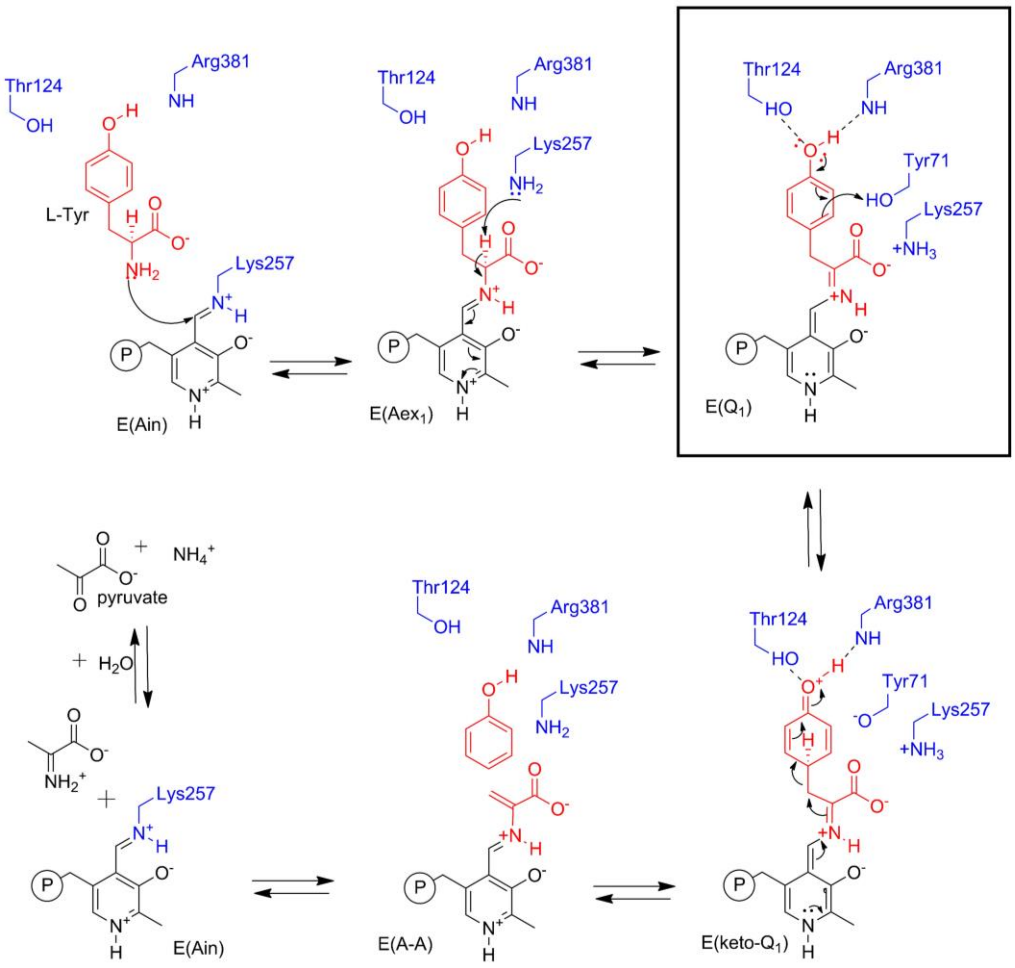
Pyridoxal-5'-phosphate (PLP) is an important biologically active phosphorylated derivative of vitamin B_6 (Di Salvo et al., 2011; Tran & Brown, 2022). The family of enzymes utilizing PLP as their workhorse has long been studied for their chemical diversity (Toney, 2014). The power of PLP as a cofactor in enzymology can be attributed to the fact that it is present in ~ 4% of all classified enzyme systems, participating in, but not limited to transamination, β- and γ-elimination, β-substitution, decarboxylation, phosphorylation, aldol cleavage, racemization, and replacement reactions (Riccardo Percudani & Alessio Peracchi, 2003; Steffen-Munsberg et al., 2015; Tran & Brown, 2022). A more elusive aspect of PLP chemistry is demonstrated in β-elimination reactions, facilitating the elimination of poor leaving groups from the β-carbon of amino acids (Phillips, 1987). Owing to their participation in synthesis of amino acid derivatives and metabolism, many PLP dependent enzymes have been identified as potential drug targets and model systems for directed evolution to produce non canonical amino acids (ncAA) (Andrew R Buller et al., 2015). Therefore, detailed understanding of PLP-facilitated enzyme catalysis has become necessary.

Common to all PLP dependent enzymes is the presence of active site lysine in an aldimine linkage with the C4' carbon of the cofactor and formation of substrate mediated carbanionic Schiff base structures (Metzler et al., 1988; Walsh, 1979). From the external aldimine and carbanion onwards, the diversification and reaction specificity arise. How PLP-mediated reaction specificity is achieved is an open-ended question and is

modulated by two primary chemistry concepts (Dajnowicz et al., 2017). First, Dunathan hypothesis (Dunathan, 1966), which postulates that bond cleavage in PLP-dependent enzymes is promoted by the arrangement of the scissile bond perpendicular to the conjugated system of the PLP. Second, the ability of the PLP to act as an electron sink to withdraw electrons and stabilize carbanion structures and subsequently assist in bond cleavage and rearrangements (Walsh, 1979). However, these basic concepts are not enough to describe the reaction diversity of PLP dependent enzyme activity. Selective protonation states of the cofactor modulated by active site electrostatics become important for explaining specific chemistries of PLP dependent enzymes. Hence, determination of protonation states remains a challenge to understand PLP-dependent reaction specificity and catalysis in biological systems (Bethany G. Caulkins et al., 2014; B. G. Caulkins et al., 2016a; Jinfeng Lai et al., 2011).

Recently, significant strides have been made in the development of structural biology tools to investigate protonation states in many enzyme intermediates (Gauto et al., 2019; Perilla et al., 2017). Methodologies such as X-ray crystallography, joint neutron/X-ray diffraction studies and Cryogenic electron microscopy (Cryo-EM) push the boundaries for resolution necessary for hydrogen atom detection and location (Guo & Rubinstein, 2018; Perilla et al., 2017). However, current state of the art for some of these methods, like, the requirements for large perdeuterated crystals and multiweek long acquisition times for neutron diffraction studies impedes the study of catalytically relevant and active enzyme intermediates (Kovalevsky et al., 2008; Michalczyk et al., 2015; Tomanicek et al., 2013). This is where NMR assisted crystallography – the joint application of solid-state NMR (SSNMR), X-ray crystallography, and first-principles computational chemistry will play a significant role in the determination of chemically rich, three-dimensional

structures of enzyme active sites (Jacob B. Holmes et al., 2022). Originally developed for molecular organic and inorganic solids, NMR crystallography played a critical role in delineating the conformations and molecular structures of these small molecules. Our group (B. G. Caulkins et al., 2016a; Jacob B. Holmes et al., 2022) and others (Gupta et al., 2015) have been working to extend this approach to mechanistic enzymology, to provide testable models of enzyme structure and function. Our approach is threefold. First, X-ray structure is used to provide a coarse structural framework for the active site. Second, chemically detailed models of the active site are built on this framework using computational chemistry, and various chemistries are explored. Third, these models are quantitatively distinguished by comparing their predicted or computed chemical shifts to the results of SSNMR experiments. Provided that enough chemical shifts are measured, NMR assisted crystallography can identify a unique, consistent structure. For the case of α-aminoacrylate intermediate in PLP dependent enzyme Tryptophan Synthase (TS), this approach identified a weak binding pocket for the water molecule released in the transformation. This integrated approach also highlighted the reaction coordinate for the loss of the substrate L-Ser hydroxyl and the protonation states of the active site Lys residue (βLys87) as one of the deciding factors to why the indole isostere benzimidiazole (BZI) does not react while the natural substrate indole does (Jacob B. Holmes et al., 2022).



Scheme 8.1 Proposed mechanism of TPL. Tyrosine Phenol Lyase splits L-Tyr into phenol and ammonium pyruvate. The proposed mechanism shows that Tyrosine coming in from the solution makes a nucleophilic attack on the C4' carbon of the PLP cofactor and passes through an external aldimine state to then give this quinonoid intermediate. At this stage of the reaction, the Cγ carbon abstracts the hydroxyl proton from Tyr71, which leads to bond cleavage between Cβ and Cγ with the release of phenol and formation of the so-called amino acrylate intermediate. This eventually releases ammonium ion and pyruvate regenerating the resting internal aldimine state.

Tyrosine phenol-lyase (TPL, [EC 4.1.99.2]) is a 206 kDa PLP-dependent homotetrameric enzyme (Kumagai et al., 1970) that catalyzes the reversible elimination of L-tyrosine to from phenol and ammonium pyruvate (Scheme 8.1) (Phillips et al., 2019; Phillips et al., 2016). This reaction pathway is directed in part by chemical and structural interactions

with acid-base groups such as Lys256, and residues' sidechains of Tyr71, Ser51, Arg404, Asn183, Asp214 and Phe448 and Phe449. Hence, the catalytic residues interacting with the cofactor establish the appropriate electrostatic environment and chemical dynamics in the active site, favoring a particular protonation state and reaction pathway. The proposed mechanism of TPL postulates that within the active subunit of TPL, catalytic cycle begins when the α-amino group of the bound substrate L-Tyr makes a nucleophilic attack on the C4' carbon of the internal aldimine form of the PLP cofactor, E(Ain) and in the subsequent steps is converted via gem diamine, E(GD), external aldimine, E(Aex), to the open form of the quinonoid intermediate, E(Q). In the next step, the open form of the E(Q) intermediate is interconverted to the closed complex of the E(Q). At this step of the catalytic pathway, the closed α-aminoacrylate Schiff's base E(A-A) is formed after the elimination of the β-leaving group (phenol, Scheme 8.1) (H. Chen et al., 1995; H. Y. Chen et al., 1995; Milic et al., 2008; Phillips et al., 2019; Phillips et al., 2016). Interconversion of the closed E(A-A) to the open E(A-A) form results in the release of the phenol from the active site of the TPL. Then, nucleophilic attack of the Lys257 at the PLP C4' carbon results in the formation of the α-aminoacrylate gem-diamine intermediate with the release of pyruvate and ammonium ion and finally regeneration of the E(Ain) (Scheme 8.1).

In addition to L-tyrosine, other amino acids such as S-alkyl and S-aryl-L-cysteines, L-serine, L-cysteine, β-Cl-L-alanine, L-alanine, L-phenylalanine are substrates(Phillips, 1987). In contrast to the reaction of L-tyrosine, these reactions are not readily reversible (Phillips et al., 2019). TPL has been shown to be a good enzyme for the production of non-canonical amino acids (ncAA) (Martínez-Rodríguez et al., 2020), however the catalytic mechanism of the enzyme is not yet fully understood with few studies directed

towards examining the protonation states of the PLP atoms in the TPL enzyme to modulate and control reaction specificity. This family of enzymes consist of aminotransferases, decarboxylases as well as enzymes that catalyze the α-, β-, or γ-eliminations. Most of these enzyme systems have common structural features in the active site, where there is an Asp residue positioned to make H-bonding interaction with the N1 of the PLP cofactor and an Asn residue making a H-bond to the PLP O3' atom (Barbolina et al., 2000; Dajnowicz et al., 2017; Phillips et al., 2016). Tautomers with the protonated Schiff base nitrogen are expected to be the catalytically active species when bond scission/formation at the C-α or the nucleophilic attack of the C4' atom of the PLP cofactor is required. In most of the fold type I PLP dependent enzymes, the PLP O3' oxygen is hydrogen bonded to the side chain of an Asn residue (N185 in TPL) (Barbolina et al., 2000), presumably acting as a hydrogen bond donor and biasing protonation of the Schiff base nitrogen. The carboxylate oxygen atoms of the PLP cofactor are oriented in the active site via the hydrogen bonding and salt bridge interactions with residues including guanidium group of Arg217 and the amide group of the Asn185 (Milic et al., 2008; Milic et al., 2011). These interactions dictate the importance of the Protonated Schiff Base (PSB) hypothesis in the TPL, which posits that a positively charged Schiff base tautomer activates catalysis via the formation of a significantly more reactive target for nucleophilic attack rather than the neutral imine.

Here, NMR-assisted crystallography is used to characterize the quinonoid intermediate in the *C.freundii* TPL active site, including first such measurements using SSNMR spectroscopy and characterization of the PLP cofactor pyridine ring nitrogen. In this work, we interrogate the active site complex formed by the reaction of L-Met with the PLP cofactor in TPL. This intermediate, termed as the L-Met-quinonoid ($E(Q)_{L\text{-}Met}$), is a

long-lived analogue of the transient quinonoids formed by the natural substrate L-Tyr. For the TPL L-Met quinonoid analogue, NMR assisted crystallography provides direct, atomic resolution support for the protonation states on the cofactor and delineate the protonated Schiff base and pyridine nitrogen tautomer as the dominant and catalytically significant species. This is consistent with the chemistry of a true canonical quinonoid intermediate and modulation of protonation states of the PLP cofactor in directing reaction specificity in TPL. The following research work will show how NMR assisted crystallography has made giant progress to linking both atomic resolution structure and chemical dynamics to mechanistic studies in a large enzyme complex, such as 206-*kDa* TPL.

8.2 Experimental Section

8.2.1 Protein Preparation. Tyrosine Phenol Lyase was prepared by overexpression of *C.freundii* in *E.coli* as previously described (Phillips et al., 2016). Samples were prepared with the following isotopic labelling schemes for the cofactor and protein components: (1) protein and cofactor unlabeled/natural abundance isotopomer concentration; (2) protein ^{15}N labeled at lysine ε-nitrogen sites; PLP unlabeled (ε-^{15}N-Lys TPL); (3) PLP cofactor selectively ^{15}N enriched at the pyridine ring nitrogen; protein unlabeled (^{15}N-PLP; TPL).

Microcrystalline protein samples for Solid-State NMR. Microcrystalline samples of TPL were prepared by diluting enzyme solution 1:1 with 50 mM TEA-HCl buffer, supplemented with 50 mM KCl, pH 8.0, containing 45% mPEG-5000 and 3 mM spermine. Microcrystals were collected and washed with 50 mM KCl, TEA-HCl buffer, pH 8.0, containing 40% mPEG-5000 and 2 mM spermine. The crystals were packed at 10 000 rpm into a Bruker 4 mm magic-angle spinning (MAS) rotor with an approximate

volume of 50 µL; each rotor contained 25-30 mg of protein. To form the L-Met quinonoid intermediate, L-Met was introduced by direct addition of ~4-5 µL of 0.2 M L-Met to the packed MAS rotor.

8.2.2 NMR Spectroscopy. ^{13}C and ^{15}N Solid-State NMR Spectroscopy.

^{13}C and `^{15}N cross-polarization (CP) magic angle spinning (MAS) experiments were performed at 9.4 T (400.37 MHz ^{1}H, 100.69 MHz ^{13}C, 40.57 MHz ^{15}N) on a Bruker AVIII spectrometer equipped with a double resonance, 4 mm MAS probe, spinning at standard MAS rates of 8 kHz; the bearing gas was cooled to -15 ^{0}C, giving an effective sample temperature of ~ -5 ^{0}C. Cross-polarization was accomplished at ^{1} (^{13}C) and 37 kHz (^{15}N) (ramped ±5 kHz); 85 kHz Spinal64 ^{1}H decoupling was used throughout. Standard ^{15}N spectra consist of the sum of ~500,000 transients acquired at a relaxation delay of 4 s, for a total acquisition time of 5 d 20 h. ^{15}N chemical shifts were referenced indirectly to liq-NH_3 (25 ^{0}C) via an external solid-state sample of $^{15}NH_4Cl$, in which the resonance frequency was set to 39.27 ppm. Standard ^{13}C spectra consist of the sum of ~500,000 transients acquired with a relaxation delay of 4 s, for a total time of 1 d 20 h. ^{13}C chemical shifts were referenced indirectly to neat TMS via an external solid-state sample of adamantane with the downfield-shifted peak set to 38.48 ppm.

The acquisition of solid-state NMR spectra was interleaved with single pulse, low-power decoupling experiments (64 scans ^{13}C, 256 scans ^{15}N) reporting predominantly on free ligand and reaction products (if any) in the mother liquor.

8.2.3 DNP Enhanced Solid-State NMR Spectroscopy.

MAS-DNP NMR spectra were collected at 600 MHz (14.1 T) using a Bruker AVIII DNP NMR system (Bruker Biospin) equipped with a Bruker 3.2 mm, $^{1}H/^{13}C/^{15}N$ MAS-DNP probe cooled to ~95 K. Microwaves were delivered using a gyrotron source operating at 395 GHz and output optimized for ~12 W at the probe base (Tran et al., 2020). Microwaves on/off spectra were collected with a MAS frequency of 10 kHz and a ^{13}C CP-echo pulse sequence with a ^{1}H CP ramp from 40 to 70 kHz, a ^{13}C RF field of 50 kHz, 85 to 100 kHz Spinal64 ^{1}H decoupling (Fung et al., 2000) and a recycle delay of 1.5 s. An AsymPolPOK (Salnikov et al., 2017; Wang et al., 2019) concentration of ~5 mM was used, resulting in enhancements >35 and buildup times on the order of 1-2 s. (*These experiments were performed in collaboration with Dr. Joanna Long's group at National MagLab, Tallahassee, Florida, USA*).

8.2.4 Evaluation of MAS-DNP samples by X-band EPR spectroscopy

CW EPR spectra were collected on a Bruker X-Band EMX Nano benchtop spectrometer (Bruker Biospin, Billerica, MA) to verify polarizing agent concentrations and PA dispersion upon freezing as described in Tran et al. (Tran et al., 2020). DNP sample rotors were placed in 5 mm quartz sample tubes and placed into the cavity with its associated adaptor. Spectra were collected at room temperature (298 K) and at 100 K by using cold nitrogen gas to cool the cavity. (*These experiments were performed in collaboration with Dr. Joanna Long's group at National MagLab, Tallahassee, Florida, USA*).

8.2.5 2D ^{15}N-^{13}C Correlation Spectroscopy using DNP Enhanced NMR Spectroscopy

2D $^{13}15$-^{13}C NC_aC_x were collected at 14.1 T (600.03 MHz ^{1}H, 150.9 MHz ^{13}C) on a standard bore Bruker NEO spectrometer equipped with a Bruker 3.2 mm, $^{1}H/^{13}C/^{15}N$ MAS-DNP probe. Experiments were performed at a MAS rate of 10.5 kHz. Cross Polarization was accomplished with a ^{13}C spin lock of 50 kHz and ^{1}H spin lock of 85-100 kHz (linearly ramped ±15%). 62.5 kHz ^{13}C and 100 kHz ^{1}H pulses were used throughout and 100 kHz Spinal64 ^{1}H decoupling was used during detection. The C-C DARR mixing time was 25 ms. Spectra were collected at room temperature (298 K) and at 100 K by using cold nitrogen gas to cool the cavity. (*These experiments were performed in collaboration with Dr. Joanna Long's group at National MagLab, Tallahassee, Florida, USA*).

8.2.6 ^{31}P Solid-State NMR Spectroscopy. ^{31}P CPMAS experiments were performed at 11.4 T (600 MHz ^{1}H, 242 MHz ^{31}P) on a Bruker Neo spectrometer equipped with an ^{1}H-X double resonance 3.2 mm MAS probe, spinning at a MAS rate of 10 kHz. The bearing gas was cooled to -27 ^{0}C, giving an effective sample temperature of -10 ^{0}C. Cross-polarization was accomplished with a ^{1}H spin-lock field of 45 kHz and a ^{31}P spin-lock of 47 kHz (ramped ±5 kHz); 58 kHz Spinal64 ^{1}H decoupling was used during detection. The ^{31}P spectra consist of the sum of 475,136 transients acquired with a relaxation delay of 3 s, for a total acquisition time of 20d 18 h. ^{31}P chemical shifts were indirectly referenced to 85% H_3PO_4 (MAS). For comparison to measurements in solution, δ[85% H3PO4 (capillary)] = δ[85% H3PO4 (sphere/ MAS)] + 0.36 ppm.

8.2.7 Chemical Shift Anisotropy (CSA) Tensor Analysis.

^{15}N CPMAS experiments on E(Q)$_{L\text{-}Met}$ were performed at 14.1 T (600.01 MHz ^{1}H, 60.82 MHz ^{15}N) on a Bruker a Bruker 3.2 mm, ^{1}H/^{13}C/^{15}N MAS-DNP probe cooled to ~95 K, spinning at a MAS rate of 10.5 kHz. Chemical shift tensor principal axis components were determined by a fit of the sideband intensities in the MAS spectra using Herzfeld-Berger analysis (Herzfeld & Berger, 1980), implemented within Bruker BioSpin's *Topspin 3.6* processing software.

8.2.8 First Principles Calculations.

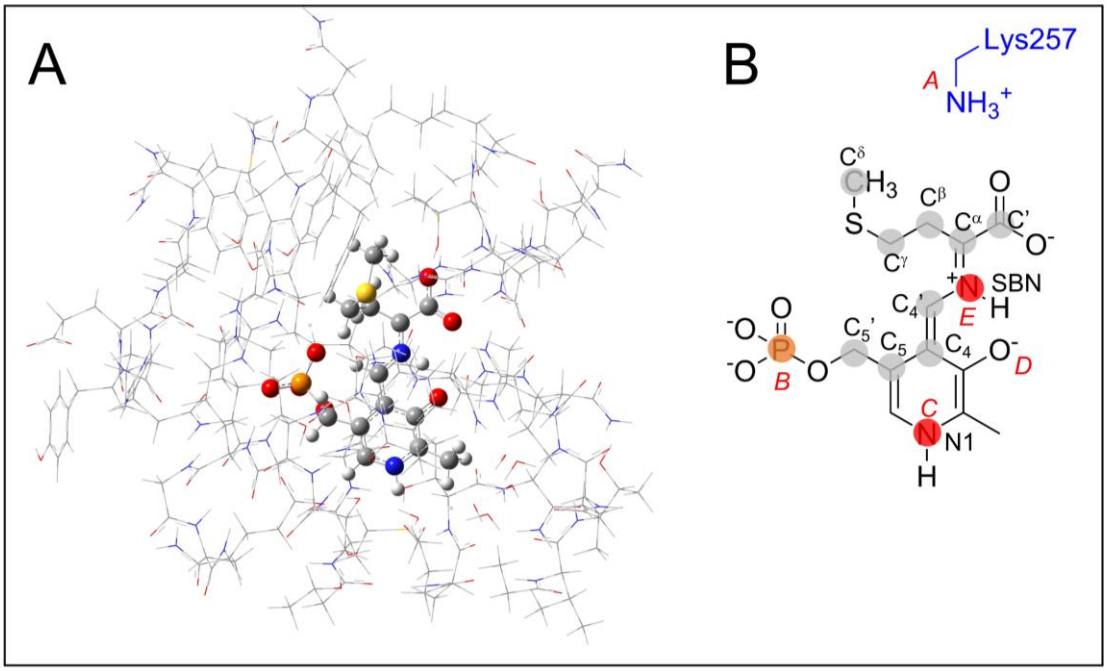


Fig 8.1 Cluster model of the E(Q)$_{L\text{-}Met}$ active site. (A) Cluster model of the active site for first-principles geometry optimization and chemical shift calculations; protein side chains displayed in wireframe and cofactor and substrate in ball-and-stick. (B) Protonation sites on and near the cofactor/substrate complex: A: the Lys257 side chain; B: the PLP phosphate group; C: the PLP pyridine ring nitrogen; D: the PLP phenolic oxygen; and E: the Schiff-base nitrogen. Shaded nuclei indicate sites for which experimental NMR chemical shifts are reported.

First principles calculations were performed using a cluster-based model of the active site as described previously (B. G. Caulkins et al., 2016a). The clusters for E(Q)$_{L\text{-}Met}$

were constructed from the corresponding crystal structures 2VLH by selecting all atoms within 8 Å of the cofactor-substrate complex (Fig 8.1). This selection was expanded to include complete residues and were modified as follows: (1) residues that were not part of continuous backbone segments and with only two atoms within the initial 8 Å cut were deleted; (2) residues with two backbone atoms and no side chain residues within 8 Å were converted to alanine; (3) residues with two side chain atoms and no backbone atoms within 8 Å were truncated by removing the backbone atoms and replacing C^{α} with a methyl group; (4) N-terminal nitrogen atoms were replaced with a hydrogen atom and C-terminal carbonyls were capped with an $–NH_2$ group and (6) hydrogen atoms were added. The final model for $E(Q)_{L\text{-}Met}$ included the amino acid residues following the above residues, the PLP-L-Met complex, and 8 water molecules and had a total of 848 atoms. The structures were further modified depending on the protonation states of the cofactor-substrate complex and Lys257.

Within this cluster, candidate structures were systematically generated by varying the protonation of the following seven ionizable sites on or near the PLP-ligand complex (Fig 8.1): the ε-amino group of βLys257, the pyridine nitrogen, the pyridoxal oxygen, the Schiff base nitrogen, and both carboxylate oxygen atoms. For these structural, the phosphate group was taken to be dianionic based on its chemical shift measurements and previous studies, and models that had more than a single proton placed at either the pyridoxal oxygen or the Schiff base nitrogen were not considered, nor were structures with a doubly protonated carboxylate. In sum, 11 variations of protonation state were constructed for $E(Q)_{L\text{-}Met}$. These models were labeled using a binary code to indicate whether a site was protonated (“1” – yes, “0” – no) in the order: (Nε of βLys257)-

(phosphate group)-(pyridine nitrogen)-(pyridoxyl phenolic oxygen)(Schiff base nitrogen); these sites are designated as AB-C-DE in Fig 8.1 and shown explicitly in Scheme 8.2.

Initial proton and sidechain orientations in the active site clusters were set using molecular dynamics (MD) simulations of the protein. These simulations were performed using the Amber package with GPU acceleration (Case et al., 2018; Salomon-Ferrer et al., 2013). Coordinates for *Cf*TPL were obtained from crystal structures 2VLH for $E(Q)_{L\text{-}Met}$ with crystal waters retained (D. Niks et al., 2013). Amber force field FF14SB was applied to protein atoms (Maier et al., 2015). Ligands were parametrized using General Amber Force Field (GAFF) and AM1-bcc charge model (Jakalian et al., 2002; Wang et al., 2004). Only hydrogens were minimized and the MD simulations were carried out using Generalized Born implicit solvent model (GB-Neck2) (Nguyen et al., 2013) at constant temperature of 298 K controlled by Langevin thermostat. Two MD trajectories for each system were collected over 50 ns at 1 ps interval with 2 fs time step. The trajectories were further processed with CPPTRAJ software to contain 5000 frames for analysis (Roe & Cheatham, 2013). The systems were visualized and analyzed using Visual Molecular Dynamics software (Humphrey et al., 1996).

All cluster models were subsequently geometry optimized in Gaussian09 (M. J. Frisch et al., 2009) at the DFT B3LYP level of theory using the Grimme D3 empirical dispersion correction (Grimme et al., 2010) and a two-tier, locally-dense basis set (Hartman et al., 2016; J. D. Hartman et al., 2015a; Moon & Case, 2006) with 6-31G(d,p) for the PLP/substrate complex and 6-31G for all other atoms. Except as noted below, all atoms on or within 4 Å of the PLP/substrate complex and all hydrogen atoms in the cluster

could adjust while the remaining atom coordinates were fixed at their crystallographic values.

For each optimized structure, NMR shieldings were calculated at the DFT B3LYP level of theory, employing a three-tier, locally dense basis set with the PLP/substrate, βLys257, and proximal sidechain atoms of the residues with H-bonding interactions to the PLP cofactor at 6-311+G(2d,p), atoms within 4Å of this subgroup at 6-311G(d,p), and all remaining atoms at 6-31G. NMR shielding calculations were performed under the explicit solvent (SCRF) protocol using a dielectric of 3.0 in *Gaussian 9.0*. NMR shielding values were converted to chemical shifts using the following linear rescaling relationships, derived at the same level of theory and bases (Hartman et al., 2016; J. D. Hartman et al., 2015a).

^{13}C: $\delta\left[\mathrm{TMS}(l)\right] = 173.70 - 0.9685\,\sigma_{\mathrm{calc}}$ (1)

^{15}N: $\delta\left[\mathrm{NH_3}(l)\right] = \delta\left[\mathrm{NH_4Cl}(s)\right] + 39.27 = 230.45 - 0.9996\,\sigma_{\mathrm{calc}}$ (2)

^{17}O: $\delta\left[\mathrm{H_2O}(l)\right] = 266.32 - 1.0551\,\sigma_{\mathrm{calc}}$ (3)

Previous benchmark studies across test sets of solid-state structures used to establish these relations, demonstrated root-mean-square errors (RMSE) for isotropic shifts to be 1.5 ppm for ^{13}C, 4.3 ppm for ^{15}N, and 7.5 ppm for ^{17}O (Hartman et al., 2016). RMSE for the chemical shift tensor components were 4.2 ppm and 13.7 ppm for ^{13}C and ^{15}N respectively (B. G. Caulkins et al., 2016a).

8.2.9 Incorporation of ^{13}C and ^{15}N PLP in TPL

The natural abundance PLP was extracted from the TPL using hydroxylamine and the apo enzyme was dialyzed (~30 kDa cutoff) against 0.1M Phosphate buffer, pH 8.0 overnight. Next day, the apo enzyme was fortified with ^{13}C and ^{15}N PLP prepared enzymatically (see Appendix 1) and incubated at 35ºc for 10 minutes. L-Met was added to the enzyme to a final concentration of ~ 2 mM and again incubated for 10 minutes at 35ºc and kept on ice for another 30 minutes. The reaction mixture was dialyzed again overnight to get rid of the excess PLP and the L-Met, with the enzyme returning to the E(Ain) resting state, confirmed via UV/Vis measurements (λ_{max} = 420 nm). Microcrystals were prepared using the protocol described in section 8.2.2.

8.3 Results and Discussion

8.3.1 X-ray Crystallography

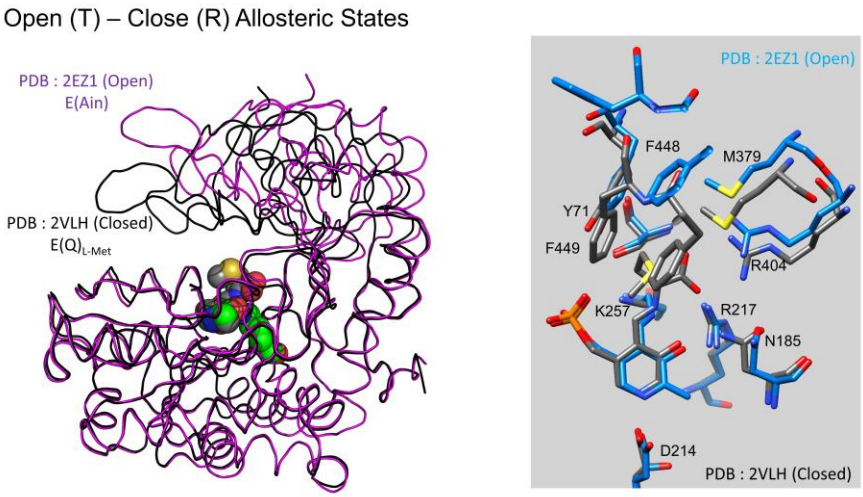


Fig 8.2. Open-close allosteric states in TPL. (*Left*) Overlay of the E(Ain) and $E(Q)_{L\text{-}Met}$ structures showing global structural changes occurring because of transition from the resting holoenzyme to the quinonoid intermediate. (*Right*) Key active site residues showing the structural changes in the conformations of the sidechains involved in interactions with the PLP-substrate complex. Coloring Scheme: (*Left*) – E(Ain) and $E(Q)_{L\text{-}Met}$ is presented in a ribbon structure with magenta coloring for the internal aldimine and black coloring for the quinonoid structure. The PLP and the PLP ligands are represented in CPK coloring scheme. (*Right*) – Carbons are represented in teal blue color and gray color for the E(Ain) and $E(Q)_{L\text{-}Met}$ respectively. O,S,P,N atoms are represented in general CPK coloring scheme. The closed conformation in quinonoid of TPL is created by the rotation of the small rigid region by about 16°around a hinge connecting it with the large rigid region. Local structural changes are shown on the *right*, where the Phe448, Ph449 rotate in and protect the bound intermediate from solvent exposure changing the electrostatics of the active site. In quinonoid, the Asn185 is rotated toward the phenolic oxygen, in the internal aldimine state, the Asn185 is rotated away.

The literature contains reports of X-ray crystal structures of $E(Q)_{L\text{-}Met}$ complexes formed via the reaction of PLP and L-Met in *Cf*TPL. The formation of the E(Q) intermediate involves the catalytically active subunit adopting a closed conformation. In the crystal structures of the $E(Q)_{L\text{-}Met}$ and E(Ain) intermediates, there are notable differences in the

active site residues surrounding the PLP moiety. While both subunits of the holo enzyme are in an open conformation, in the $E(Q)_{L\text{-}Met}$ intermediate, one of the subunits is in a closed conformation due to positional adjustments of small and large rigid regions. This information is illustrated in Fig 8.2. Analysis of the electron density map for the crystal structure of the TPL-L-Met complex reveals that the L-Met sidechain is rotated towards the phosphate group of the PLP moiety and away from Van der Waal's contact with the M379 sidechain. This raises an interesting question about the importance of Van der Waal's interactions with the original substrate L-Tyr sidechain in maintaining the electrostatics of the active site for the next step of the catalytic pathway, which involves the Cgamma carbon abstracting the hydroxyl proton from Tyr71, leading to bond cleavage between Cb and Cg with the release of phenol and formation of the amino acrylate intermediate (Scheme 8.1). Recently, Phillips et al. solved the crystal structure of the mutant M379A *Cf*TPL complexed with L-Met, which showed that the L-Met sidechain is positioned to make Van der Waal's contact with the M379 sidechain sulfhydryl group, unlike in the wild-type enzyme where it is in a different orientation. While crystal structures provide important information about the global structural states relevant to enzyme catalysis, they may not capture the dynamic nature of the active site residues or the protonation states of the complex and this is where NMR Assisted Crystallography can complement crystallographic studies by providing information about the dynamics of active site residues and the protonation states of the complex.

8.3.2 NMR Spectroscopy

ssNMR experiments were performed on microcrystalline protein samples prepared in close analogy to the single crystals used for X-ray crystallography. Microcrystals of TPL were prepared in analogous manner to the single crystal for diffraction studies to

maintain the same crystal habit as the larger crystals (Ahmed et al., 1987), giving high confidence that the NMR and X-ray data can be directly compared.

Both $E(Q)_{L\text{-}Met}$ and E(Ain) can be observed with ssNMR under conditions that favor the accumulation of the enzyme-bound species while the microcrystals remain catalytically active. ssNMR spectra of the reaction of L-Met with microcrystalline TPL to form $E(Q)_{L\text{-}Met}$. The signals in the cross-polarization magic-angle-spinning (CPMAS) spectra correspond to crystalline protein and bound substrate. The spectra were acquired with various combinations of ^{15}N and ^{13}C labels on the PLP cofactor, L-Met substrate, and protein Lys ε-amino groups. With unlabeled substrate and protein, the spectra show only unresolved background signals from the protein, including the backbone amide nitrogens (~120 ppm) and backbone carbonyl (~170 ppm), aliphatic (~10-70 ppm), and aromatic (~110-160 ppm) carbons.

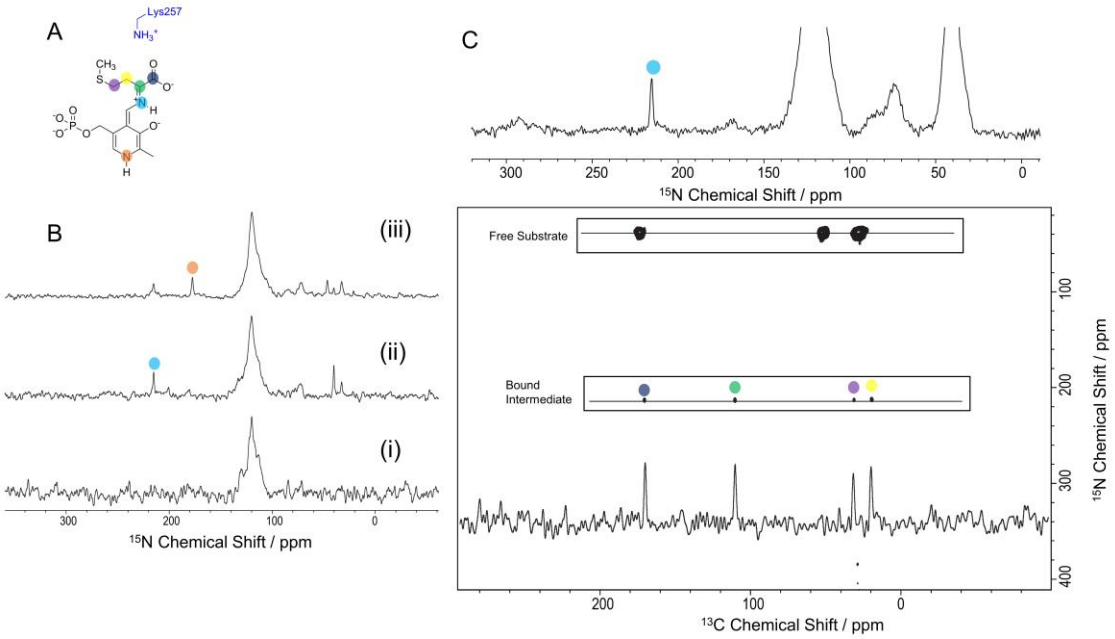


Fig 8.3 (A) Organic structure of $E(Q)_{L\text{-}Met}$ (B) ^{15}N CPMAS RT SSNMR spectra of microcrystalline TPL $E(Q)_{L\text{-}Met}$ prepared with the following isotopic labeling: (i) natural abundance TPL reacted with natural abundance L-Met; (ii) ^{15}N-enriched on L-Met reacted with natural abundance TPL; (iii) ^{15}N-enriched at PLP reacted with ^{15}N L-Met and natural abundance protein; (C) NCaCx correlation spectroscopy collected using DNP enhanced ssNMR spectroscopy; sample prepared with uniformly ^{13}C and ^{15}N labeled L-Met substrate show correlations in the ^{13}C dimension to the SB N (*215.1 ppm*). The large cross peaks (*top*) is from free L-Met. Acquisition details for the RT CPMAS and DNP enhanced CPMAS NMR spectroscopy are described in the Experimental section.

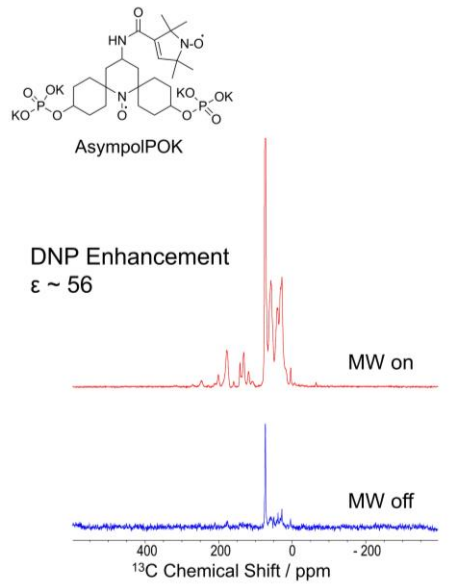


Fig 8.4 DNP enhancements achieved in TPL active site using the biradical AsympolPOK.

With the magnification of the isotopic enrichment, distinct resonances for the active site species become visible. 1D ^{15}N CPMAS NMR spectra collected at room temperature on the natural abundance protein sample prepared with uniformly ^{13}C-^{15}N enriched L-Met shows a unique resonance at 215.1 ppm which is tentatively assigned to the SB N linkage between the α- amino group of the incoming amino acid, L-Met and the C4' carbon of the PLP cofactor.

However, we observe no distinct ^{13}C resonances from the bound intermediate at room temperature. We take advantage of DNP enhanced CPMAS ssNMR spectroscopy, where we get enhancements of a factor of ~56 in the active site of the enzyme (Fig 8.4) allowing for the assignments of the carbon on the bound species.

Missing from the correlations are the methyl group (C^{δ}) assignments of the bound intermediate. To assign the methyl carbons, we performed RT CPMAS on TPL sample

reacted with ^{13}C-Methyl-L-Met (Fig 8.5A) showing the unique single resonance at 16.9 ppm distinct from the free L-Met. The Schiff base nitrogen of the cofactor also displayed no significant temperature dependence even at ~95 K. For this atom, the chemical shift tensor was measured using xCSA experiments (Fig 8.5B), and the principal axis components, [(δ_{11}, δ_{22}, δ_{33})] =[327.5, 281.5 and 35.9] ppm, were extracted by fitting the intensity of the spinning sideband manifold (Herzfeld & Berger, 1980).

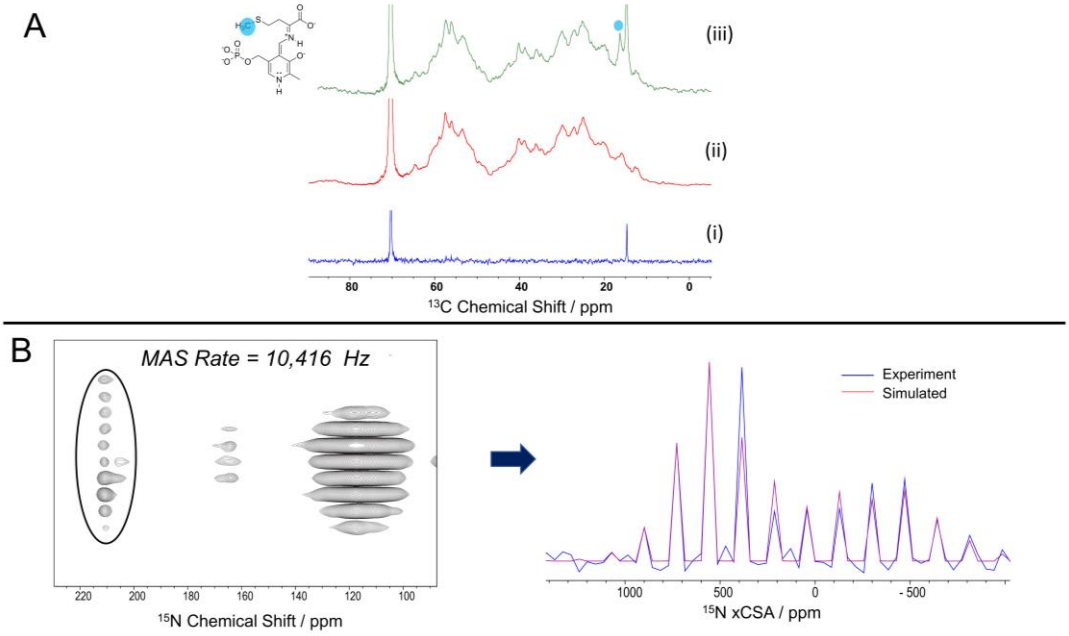


Fig 8.5 A. 1D ^{13}C NMR Spectroscopy. (i)^{13}C spectrum showing the components of the mother liquor (~15 ppm free methyl group of the L-Met substrate and ~70 ppm for PEG molecule) acquired using low ^{1}H decoupling pulses; TPL E(Q)$_{L\text{-}Met}$ prepared with (i) natural abundance TPL reacted with natural abundance L-Met; (iii) ^{13}C-enriched on the methyl carbon of L-Met reacted with natural abundance TPL. 8.5 B. ^{15}N xCSA measurements of the TPL quinonoid intermediate formed with U-^{15}N-L-Met in 90% D_2O buffer. The isotropic chemical shifts appear on the direct *F2* dimension and the anisotropies are extracted from the *F1* dimension and fitted using the SOLA program in Bruker Topspin 3.6.

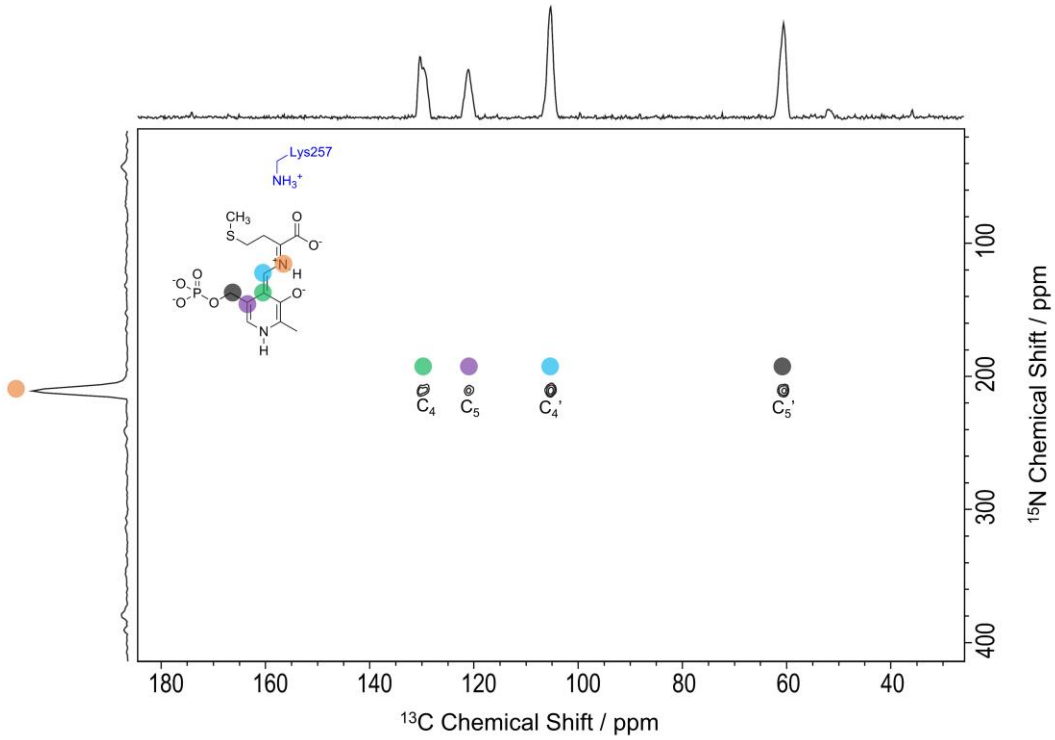


Fig 8.6 NCaCx correlation spectroscopy collected using DNP enhanced ssNMR spectroscopy; sample prepared containing ^{13}C labels on C4, C4', C5 and C5' carbons of the PLP and reacted with ^{15}N labeled L-Met show correlations in the ^{13}C dimension to the SB N (*215.1 ppm*). Acquisition details for DNP enhanced CPMAS NMR spectroscopy are described in the Experimental section.

At DNP regime, the spectra of the sample prepared with ^{15}N L-Met and ^{13}C on the PLP cofactor (Fig 8.6) show a nitrogen resonance at 215.1 ppm and four additional carbon resonances at 132, 107.5, 123.3 and 63 ppm that are assigned to the SB N and cofactor atoms C4, C4', C5 and C5', respectively achieved via NCaCx correlation spectroscopy (atom labeling given in Fig 8.2). The sample prepared using ^{13}C- and ^{15}N-L-Met as a substrate (Fig 8.3) displays similar nitrogen resonance at 215.1 ppm correlated to additional carbon resonances at 170.6 ppm, 110.3 ppm, 19.3 ppm and 31.1 that are assigned to the carbons that derive from the L-Met C′, C^{α},C^{β} and C^{γ} respectively.

These shifts help establish several key elements of the chemical structure for $E(Q)_{L\text{-}Met}$. At this point in the catalytic cycle, the L-Met has formed the quinonoid complex and there is a double bond between C^{α} and the SB N, which is confirmed by the C^{α} chemical shift at 110.3 ppm. Second, with a shift of 214.1 ppm, the Schiff base linkage is found to be protonated, which is further supported by the small span of its chemical shift tensor (Copie et al., 1988). This isotropic chemical shift, however, is higher than the limiting value for a fully protonated Schiff base, and the carbon dynamics at room temperature suggests altered conformations of the complex involving transient protonation of this site.

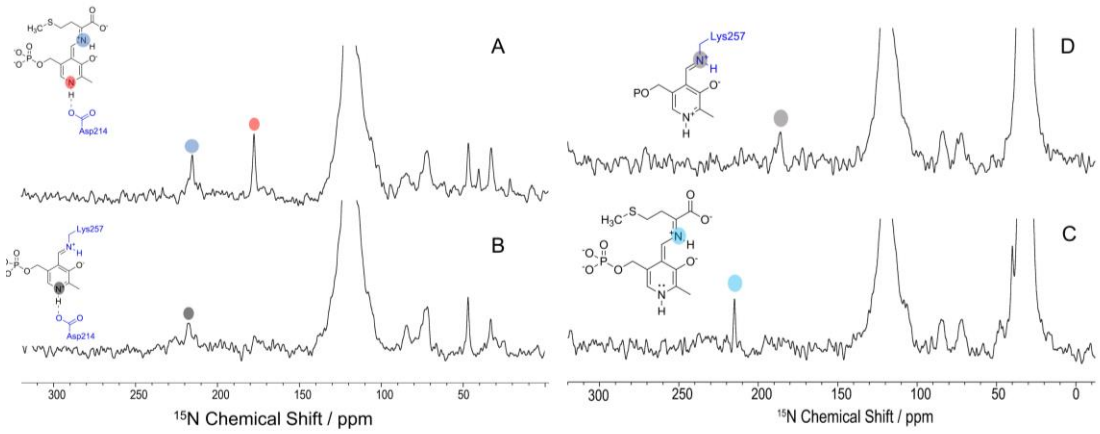


Fig 8.7 ^{15}N CPMAS RT SSNMR spectra of microcrystalline TPL E(Ain) and $E(Q)_{L\text{-}Met}$ prepared with the following isotopic labeling: (A) ^{15}N enriched at PLP in TPL and reacted with ^{15}N L-Met; (B) 15N-enriched at PLP inside TPL; (C) ε-^{15}N Lys labelled TPL reacted with ^{15}N L-Met (D) ε-^{15}N Lys labelled TPL.

The ^{15}N side chain chemical shift of βLys257 was measured on a TPL sample in which all lysine residues were ^{15}N-labeled on the ε-amino group (ε-$^{15}NH_3$-Lys TPL). In the initial resting state, E(Ain), this sample shows a broad peak at 186 ppm, corresponding to the protonated Schiff base linkage between βLys257 and the PLP cofactor. In the sample prepared with ^{15}N labelled PLP cofactor, a new resonance is observed at 217.4

ppm (Fig 8.7D), assigned to the N1 atom of the PLP in the holo enzyme. Upon addition of L-Met, this resonance is lost and two new resonances at 214.1 and 178.4 ppm are observed corresponding to the formation of the $E(Q)_{L\text{-}Met}$ complex (Fig 8.7(A,B)), indicating that the ε-^{15}N-Lys is now detached from the cofactor and is acting as acid/base catalyst in the reaction pathway.

The PLP nitrogen shift of 217.4 and 178.4 ppm for both the E(Ain) and $E(Q)_{L\text{-}Met}$ complex indicates a protonated pyridine ring N1 and confirms the formation of the canonical quinonoid complex unlike in Tryptophan Synthase.

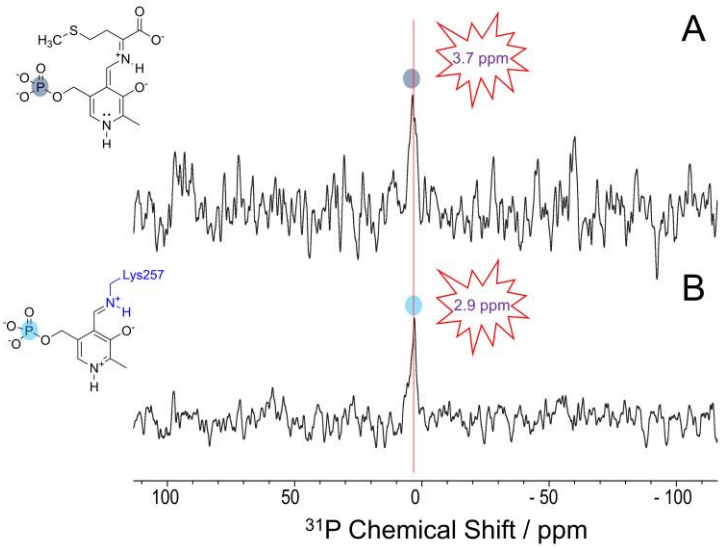


Fig 8.8 ^{31}P spectra of the (A) $E(Q)_{L\text{-}Met}$ and (B) E(Ain) intermediate in TPL showing phosphorous resonances at 3.7 ppm and 2.9 ppm respectively. A line broadening of 250 Hz is applied to each spectrum.

The ^{31}P isotropic chemical shift were also measured for the cofactor's phosphate group for both the resting state and the quinonoid complex with an isotropic shift of 2.9 and 3.7 ppm respectively (Fig 8.8) indicating that it is in a dianionic regime, in keeping with other PLP-dependent enzymes (Schnackerz et al., 2011).

8.3.3 First-Principles Calculations

The preliminary chemical structures mentioned above have provided valuable insights into the chemistry of the active site, but they are limited in two important ways. Firstly, they lack detailed three-dimensional structure, which is necessary for a complete understanding of the chemical interactions that occur within the active site. Secondly, the chemical shifts observed in experiments do not fully match the expected limiting shifts based on model compound studies. To fully interpret these experimental results and to place the chemistry of the active site in the proper structural context, it is necessary to use first-principles computational chemistry. By using computational methods that are based on fundamental physical principles, we can simulate the behavior of the molecules within the active site and obtain detailed information about their structures, energies, and chemical properties. This approach can help us to understand the complex interactions that occur within the active site and to make predictions about the behavior of the enzyme under different conditions. Therefore, the use of first-principles computational chemistry is a powerful tool for understanding enzyme catalysis and for guiding the design of new enzymatic systems.

The current state-of-the-art for first-principles chemical shift calculations is highly advanced, as demonstrated by previous studies (B. G. Caulkins et al., 2016a; Hartman et al., 2016; J. D. Hartman et al., 2015a; Jacob B. Holmes et al., 2022). These studies have shown that if the correct structure of the molecule is known, including the position of every atom, then NMR chemical shifts can be predicted with high accuracy. Specifically, the predicted chemical shifts can be within 1.5 ppm RMSD for carbon, and 4.3 ppm for nitrogen (J. D. Hartman et al., 2015a). This level of accuracy allows for the

establishment of screening protocols in which proposed structures can be evaluated and ranked based on their consistency with the experimental chemical shifts. The degree of agreement between the predicted and experimental chemical shifts can be quantified using the reduced-$\chi 2$ statistic, which calculates the squared deviation of the predicted and experimental chemical shifts weighted by the nuclide-dependent mean squared deviations (as described in Eq. 4 in the Experimental section). A solved structure is considered to be properly identified if it satisfies the 95% confidence limits of the reduced-$\chi 2$. Typically, the differentiation of proposed models requires the comparison of 10-12 chemical shifts throughout the active site.

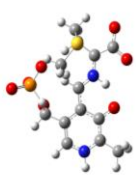
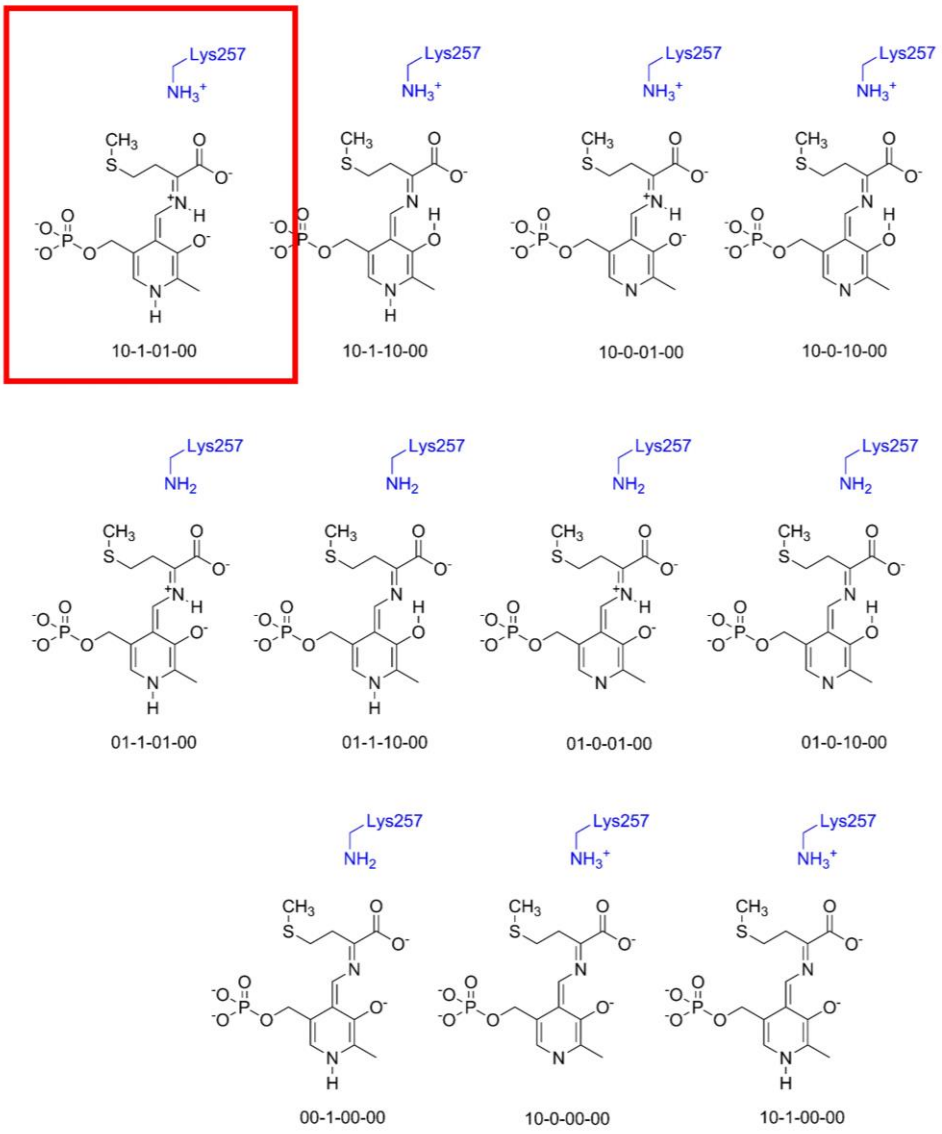
$E(Q)_{L\text{-}Met}$ ***(A)***	Atom	Model (X-Ray Refined)	Expt	***A***  ***B*** 	$E(Q)_{L\text{-}Met}$ ***(B)***	Atom	Model (Refined)	Expt
PLP	N1	188.5	177.4		PLP	N1	187.5	177.4
	C4	130.1	132.2			C4	130.1	132.2
	C4'	106.2	107.5			C4'	106.8	107.5
	C5	123.8	123.3			C5	123.8	123.3
	C5'	63.7	63			C5'	63.2	63
L-Met	SB N	215.9	215.1		L-Met	SB N	213.2	215.1
	Cα	110.09	110.3			Cα	108.0	110.3
	C'	171.1	170.6			C'	171.2	170.6
	Cβ	21.7	19.3			Cβ	20.1	19.3
	Cγ	37.3	31.1			Cγ	30.9	31.1
	Cδ	22.7	16.9			Cδ	23.0	16.9
	red-χ^2	2.95				red-χ^2	1.09	

Table 8.1. $E(Q)_{L\text{-}Met}$ experimental and first-principles chemical shifts (ppm) for the conformation of the $E(Q)_{L\text{-}Met}$ refined (A) from X-ray structure (PDB 2VLH) and remodeled and refined (B) $E(Q)_{L\text{-}Met}$ conformation.

To limit bias and broadly consider various active site chemistries, a pool of candidate structures was generated by systematically varying the protonation states of the cofactor, substrates, and Lys257 ε-amino group (Fig 8.1A, B). The candidate structures were constructed directly as three-dimensional cluster models of the active site, built on the framework of the X-ray crystal structure, and included all residues and crystallographic waters within 8 Å of the substrate and cofactor. The atoms at the exterior of the cluster were fixed at their crystallographic locations, and quantum-mechanical, DFT-based geometry optimization and chemical shift calculations were performed. The clusters contained 840-850 atoms (Fig 8.1A), a size for which the convergence and accuracy of the chemical shift calculations have been established (J. D. Hartman et al., 2015a).



Scheme 8.2 $E(Q)_{L\text{-}Met}$ candidate structures

11 models for $E(Q)_{L\text{-}Met}$ were generated (Scheme 8.2, Fig 8.9), geometry optimized, and their chemical shifts predicted (Table 8.1). The structures were ranked based on their agreement with the experimental chemical shifts (Fig 8.9). There is a clear

differentiation of the models, but initially all candidate structures were found to fall outside of the reduced-χ^2 95% confidence limits, indicating that none reproduced the experimental chemical shifts with the expected accuracy. One of the largest discrepancies for all models occurred for the $C^{\gamma}$and C^{δ}, where the closest predicted shifts were still much greater than standard errors from the experimental value.

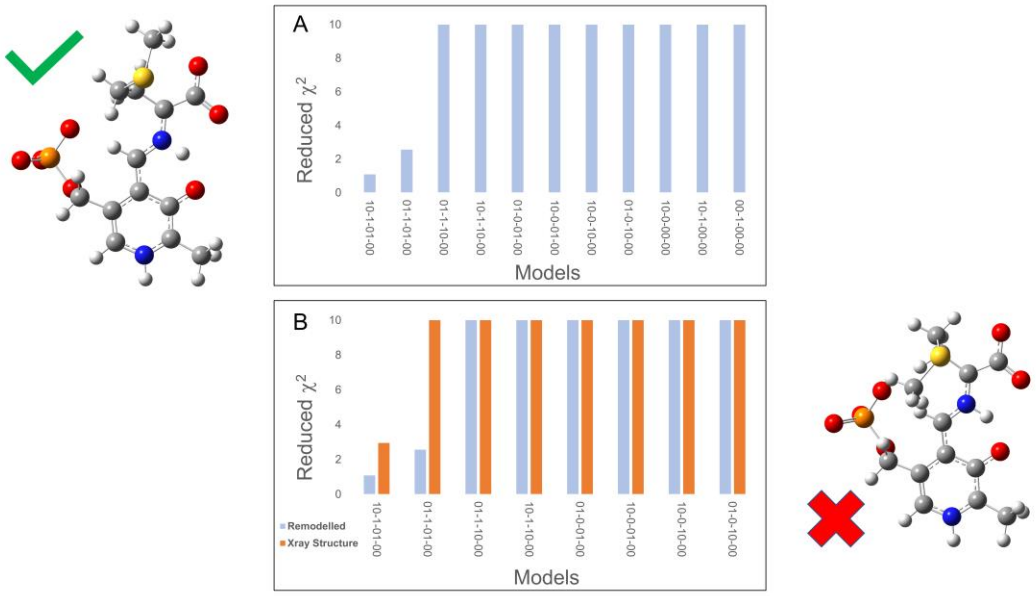


Fig 8.9 Ranking of structural models based on agreement between the experimental and first principles chemical shifts as quantified by the reduced-χ^2 statistic. (A) The 11 best geometry-optimized active site models for the $E(Q)_{L\text{-}Met}$ intermediate and (B) Rankings of the 8 best structural models comparing the experimental and first-principles isotropic chemical shifts of the remodeled PLP-L-Met conformation (blue) and the PLP-L-Met complex geometry optimized without any modification from X-ray crystal structure (orange).

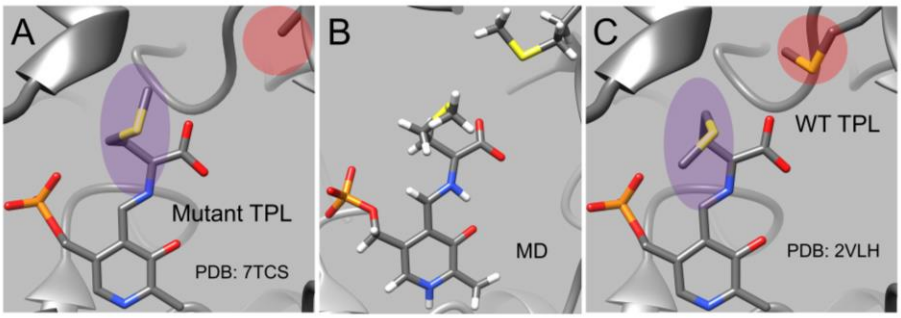


Fig 8.10 X-ray and Molecular Dynamics generated different conformations of the cofactor-substrate complex in TPL. (A) M379A TPL, (B) MD generated structure from a 100 ns run, (C) Wild-type (WT) TPL. The purple and the red ovals highlight the main differences in the structures reflected by different orientation of the methyl sidechain connected to the PLP moiety and the methyl sidechain of the M379/M379A residue.

Motivated by the recently solved crystal structure of the mutant M379A TPL $E(Q)_{L\text{-}Met}$, we performed MD simulations on the crystal structure, where it was observed that the methionine sidechain of the bound complex is relaxing in an orientation completely different from the static crystal structure in a 100 ns run (see Fig 8.10). A new conformation of the cofactor-substrate complex was built like the one observed in the mutant and was positioned in place of the X-ray crystallography determined complex and the active site was geometry optimized as previously described and NMR chemical shift calculations of all the tautomers were calculated and analyzed. Only a single tautomer of the $E(Q)_{L\text{-}Met}$ was found to satisfy the isotropic chemical shift restraints for *10* chemical shifts with a reduced-χ^2 of 1.08 falling between 95-99 % confidence intervals (Fig 8.9).

For both $E(Q)_{L\text{-}Met}$ structural models, the best-fit model was found to be the structural model with a protonated SB N and Pyridine nitrogen, which is contrasting to another PLP dependent enzyme TS where the winning model was found to be a tautomeric exchange

pair between SB N and PO for both the carbanionic and the amino-acrylate intermediate. This difference is quite unique and based on the structural organization of PLP fold type I enzymes, is expected, where there is a Aspartate (Asp214) H-bonding to the N1 atom of the pyridine ring and forming a canonical quinonoid intermediate, whereas in TS, there is a Ser (Ser377) hydroxyl group H-bonding to the N1 atom.

As a final assessment, the overall confidence in the identification of the experimental structure was quantified using Bayesian probability analysis following the approach of Engel et al. (Engel et al., 2019). Considering the isotropic chemical shift data, along with different conformations of the cofactor-substrate complex, the proposed winning structural model was found to be the most probable experimental structure with 93.4% confidence for $E(Q)_{L\text{-}Met}$. Dynamics of the L-Met methyl sidechain may explain why the overall agreement between theory and experiment is lower than expected confidence interval of 95%.

8.3.4 Chemically Rich Crystal Structures and Mechanistic Implications

Perhaps most striking in the structure of the $E(Q)_{L\text{-}Met}$ active site is the position of the methyl sidechain of the L-Met in the bound intermediate of the wild-type X-ray crystal structure. This sidechain is perfectly positioned for the Van der Waal's interactions with the M379 sulfhydryl sidechain. The interactions with of the Met side chain with Phe36, Phe448, Phe449, Met379, and Arg381 provide the required electrostatic environment The interaction of the water molecule with βLys257 and its proximity to the C^{β} carbon support the hypothesis that βLys257 is the acid catalyst that protonates the β-hydroxyl during C-O bond scission as E(Q) is converted to E(A-A) in TPL.

Close examination of the X-ray crystal structures shows that the electron density map could be fitted with a different conformation of the L-Met sidechain bound to the PLP. Initial model refinements performed using NMR Crystallography show that the refined structural model from this integrative method fits well into the electron density of the quinonoid in the active site of the enzyme (*in progress*). Critical inspection of the NMR data suggests that the methyl sidechain is too dynamic and unstable to be present in the conformation that has been revealed by the X-ray crystallography. We hypothesize that there are multiple conformations of the methyl sidechain that are present in the crystal structure (*Ensemble Refinement in progress*). The fact that we cannot get any methyl correlations in the NCaCx correlation spectroscopy suggests a disorder of the Methionine sidechain in the bound species in the active site.

In addition to conformational changes, NMR Crystallography also points out the unique protonation states of the SB N in the E(Ain) and E(Q) intermediate. The NMR chemical shifts show that the internal aldimine SB N is protonated at physiological pH. Therefore, it is activating the C4' carbon for the nucleophilic attack from the substrate's α-amino group. The bond between the PLP and the conserved lysine residue is cleaved during the formation of the external aldimine. N1 atom of the PLP ring is also protonated in the E(Ain) unlike in TS. This N1-PLP protonation influences the coplanarity of the SB N relative to the pyridinium ring, with the SB C=N torsion angle tilted by ~ 46º. In addition, N1-PLP protonation is important for the enhanced electron sink effect for stabilization of the carbanion intermediate. We can contrast this with the quinoid intermediate, where the PLP nitrogen is at 178 ppm and the PLP ring is reoriented as confirmed previously. This chemical shift of the N1 atom tells us that the PN is still protonated and has a very different electronic structure. The fact that it is hydrogen bonded to the Asp in both the

cases and the PLP is already charged in the internal aldimine, making it a good electron sink. However, for the quinonoid, there is an excess negative charge in the π electron system pulling the hydrogen closer away from the Asp that is hydrogen bonded to. The charge is compensated by the positive charge in the nitrogen.

8.4 Conclusions

NMR Crystallography in TPL provides a means to directly determine the protonation states in Fold type I PLP enzymes, which can shed light on how these enzymes regulate reaction specificity and catalytic pathways. When the cofactor and substrate bind to the enzyme, conformational and electrostatic changes occur in the active site, which can lead to changes in the required electronic configurations. By selectively protonating PLP sites and certain residues within the active site, nature has evolved a way to finely control biochemical reactions through stereo electronic control and electronic modulation in PLP-dependent enzymes.

Here the joint application of ssNMR spectroscopy, X-ray crystallography and first principles chemistry offers unprecedented, chemically rich views of the $E(Q)_{L\text{-}Met}$ and the E(Ain) active sites. Through the combined determination of heavy atom positions, identification of the protonation states, and orientation of the cofactor-substrate complex, a clear picture of the structure and dynamics emerges.

NMR crystallography takes advantage of one of the well-established strengths of NMR spectroscopy – remarkable sensitivity to chemical structure and chemical dynamics. This is, we would argue, where NMR will continue to interface most strongly with the other tools of structural biology, including X-ray and neutron crystallography and cryo-EM. When combined with first-principles computational chemistry, these complementary techniques can build consistent, testable models of structure and reactivity in enzyme

active sites. Importantly, this can be accomplished for samples near room temperature and under conditions of active catalysis.

8.5 References

Ahmed, S. A., Hyde, C. C., Thomas, G., & Miles, E. W. (1987). Microcrystals of tryptophan synthase $\alpha_2\beta_2$ complex from Salmonella typhimurium are catalytically active. *Biochemistry*, *26*(17), 5492-5498.

Barbolina, M., Phillips, R., Gollnick, P., Faleev, N., & Demidkina, T. (2000). Citrobacter freundii tyrosine phenol-lyase: the role of asparagine 185 in modulating enzyme function through stabilization of a quinonoid intermediate. *Protein engineering*, *13*(3), 207-215.

Buller, A. R., Brinkmann-Chen, S., Romney, D. K., Herger, M., Murciano-Calles, J., & Arnold, F. H. (2015). Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proceedings of the National Academy of Sciences*, *112*(47), 14599-14604.

Case, D. A., Ben-Shalom, Ido, Brozell, R., S., Cerutti, David, Cheatham, III, T. E., Cruzeiro, Wilian, V., Darden, T., Duke, R., E., Ghoreishi, Delaram, Gilson, M., Gohlke, Holger, . . . A., P. (2018). *AMBER 2018*. In University of California.

Caulkins, B. G., Bastin, B., Yang, C., Neubauer, T. J., Young, R. P., Hilario, E., Huang, Y.-m. M., Chang, C.-e. A., Fan, L., Dunn, M. F., Marsella, M. J., & Mueller, L. J. (2014). Protonation States of the Tryptophan Synthase Internal Aldimine Active Site from Solid-State NMR Spectroscopy: Direct Observation of the Protonated Schiff Base Linkage to Pyridoxal-5′-Phosphate. *Journal of the American Chemical Society*, *136*(37), 12824-12827. https://doi.org/10.1021/ja506267d

Caulkins, B. G., Young, R. P., Kudla, R. A., Yang, C., Bittbauer, T. J., Bastin, B., Hilario, E., Fan, L., Marsella, M. J., Dunn, M. F., & Mueller, L. J. (2016). NMR Crystallography of a Carbanionic Intermediate in Tryptophan Synthase: Chemical Structure, Tautomerization, and Reaction Specificity. *Journal of the American Chemical Society*, *138*(46), 15214-15226. https://doi.org/10.1021/jacs.6b08937

Chen, H., Gollnick, P., & Phillips, R. S. (1995). Site-Directed Mutagenesis of His343→ Ala in Citrobacter freundii Tyrosine Phenol-Lyase: Effects on the Kinetic Mechanism and Rate-Determining Step. *European Journal of Biochemistry*, *229*(2), 540-549.

Chen, H. Y., Demidkina, T. V., & Phillips, R. S. (1995). Site-directed mutagenesis of tyrosine-71 to phenylalanine in Citrobacter freundii tyrosine phenol-lyase: evidence for dual roles of tyrosine-71 as a general acid catalyst in the reaction mechanism and in cofactor binding. *Biochemistry*, *34*(38), 12276-12283.

Copie, V., Faraci, W. S., Walsh, C. T., & Griffin, R. G. (1988). Inhibition of alanine racemase by alanine phosphonate: detection of an imine linkage to pyridoxal 5'-phosphate in the enzyme-inhibitor complex by solid-state nitrogen-15 nuclear magnetic resonance. *Biochemistry*, *27*(14), 4966-4970.

Dajnowicz, S., Johnston, R. C., Parks, J. M., Blakeley, M. P., Keen, D. A., Weiss, K. L., Gerlits, O., Kovalevsky, A., & Mueser, T. C. (2017). Direct visualization of critical hydrogen atoms in a pyridoxal 5′-phosphate enzyme. *Nature communications*, *8*(1), 1-9.

Di Salvo, M. L., Contestabile, R., & Safo, M. K. (2011). Vitamin B6 salvage enzymes: mechanism, structure and regulation. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1814*(11), 1597-1608.

Dunathan, H. C. (1966). Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proceedings of the National Academy of Sciences*, *55*(4), 712-716.

Engel, E. A., Anelli, A., Hofstetter, A., Paruzzo, F., Emsley, L., & Ceriotti, M. (2019). A Bayesian approach to NMR crystal structure determination [10.1039/C9CP04489B]. *Physical Chemistry Chemical Physics*, *21*(42), 23385-23400. https://doi.org/10.1039/C9CP04489B

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., . . . Fox, D. J. (2009). *Gaussian 09*. In Gaussian, Inc.

Fung, B. M., Khitrin, A. K., & Ermolaev, K. (2000). An improved broadband decoupling sequence for liquid crystals and solids. *J Magn Reson*, *142*(1), 97-101. https://doi.org/10.1006/jmre.1999.1896

Gauto, D. F., Estrozi, L. F., Schwieters, C. D., Effantin, G., Macek, P., Sounier, R., Sivertsen, A. C., Schmidt, E., Kerfah, R., & Mas, G. (2019). Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nature communications*, *10*(1), 2697.

Grimme, S., Antony, J., Ehrlich, S., & Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics*, *132*(15). https://doi.org/Artn 154104

10.1063/1.3382344

Guo, H., & Rubinstein, J. L. (2018). Cryo-EM of ATP synthases. *Current Opinion in Structural Biology*, *52*, 71-79. https://doi.org/https://doi.org/10.1016/j.sbi.2018.08.005

Gupta, R., Hou, G., Renirie, R., Wever, R., & Polenova, T. (2015). 51V NMR crystallography of vanadium chloroperoxidase and its directed evolution P395D/L241V/T343A mutant: protonation environments of the active site. *Journal of the American Chemical Society*, *137*(16), 5618-5628.

Hartman, J. D., Kudla, R. A., Day, G. M., Mueller, L. J., & Beran, G. J. (2016). Benchmark fragment-based ^{1}H, ^{13}C, ^{15}N and ^{17}O chemical shift predictions in molecular crystals. *Physical Chemistry Chemical Physics*, *18*(31), 21686-21709.

Hartman, J. D., Neubauer, T. J., Caulkins, B. G., Mueller, L. J., & Beran, G. J. (2015). Converging nuclear magnetic shielding calculations with respect to basis and system size in protein systems [Research Support, N.I.H., Extramural

Research Support, U.S. Gov't, Non-P.H.S.]. *Journal of Biomolecular NMR*, *62*(3), 327-340. https://doi.org/10.1007/s10858-015-9947-2

Herzfeld, J., & Berger, A. E. (1980). Sideband intensities in NMR spectra of samples spinning at the magic angle. *Journal of Chemical Physics*, *73*, 6021-6030. https://doi.org/10.1063/1.440136

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., Ghosh, R. K., Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., Bogie, P. M., Paulino, J., Wang, X., Riviere, G., Bosken, Y. K., Struppe, J., Hassan, A., Guidoulianov, J., Perrone, B., Mentink-Vigier, F., . . . Mueller, L. J. (2022). Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *Proceedings of the National Academy of Sciences*, *119*(2), e2109235119. https://doi.org/doi:10.1073/pnas.2109235119

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33-38. https://doi.org/http://dx.doi.org/10.1016/0263-7855(96)00018-5

Jakalian, A., Jack, D. B., & Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry*, *23*(16), 1623-1641. https://doi.org/10.1002/jcc.10128

Kovalevsky, A. Y., Katz, A. K., Carrell, H. L., Hanson, L., Mustyakimov, M., Fisher, S. Z., Coates, L., Schoenborn, B. P., Bunick, G. J., Glusker, J. P., & Langan, P. (2008). Hydrogen Location in Stages of an Enzyme-Catalyzed Reaction: Time-of-Flight Neutron Structure of d-Xylose Isomerase with Bound d-Xylulose. *Biochemistry*, *47*(29), 7595-7597. https://doi.org/10.1021/bi8005434

Kumagai, H., Yamada, H., Matsui, H., Ohkishi, H., & Ogata, K. (1970). Tyrosine phenol lyase I. Purification, crystallization, and properties. *Journal of Biological Chemistry*, *245*(7), 1767-1772.

Lai, J., Niks, D., Wang, Y., Domratcheva, T., Barends, T. R. M., Schwarz, F., Olsen, R. A., Elliott, D. W., Fatmi, M. Q., Chang, C.-e. A., Schlichting, I., Dunn, M. F., & Mueller, L. J. (2011). X-ray and NMR Crystallography in an Enzyme Active Site: The Indoline Quinonoid Intermediate in Tryptophan Synthase. *Journal of the American Chemical Society*, *133*(1), 4-7. https://doi.org/10.1021/ja106555c

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB [research-article]. https://doi.org/10.1021/acs.jctc.5b00255

Martínez-Rodríguez, S., Torres, J. M., Sánchez, P., & Ortega, E. (2020). Overview on Multienzymatic Cascades for the Production of Non-canonical α-Amino Acids. *Frontiers in bioengineering and biotechnology*, *8*, 887.

Metzler, C. M., Harris, A. G., & Metzler, D. E. (1988). Spectroscopic studies of quinonoid species from pyridoxal 5'-phosphate. *Biochemistry*, *27*(13), 4923-4933.

Michalczyk, R., Unkefer, C. J., Bacik, J.-P., Schrader, T. E., Ostermann, A., Kovalevsky, A. Y., McKenna, R., & Fisher, S. Z. (2015). Joint neutron crystallographic and NMR solution studies of Tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proceedings of the National Academy of Sciences*, *112*(18), 5673-5678. https://doi.org/doi:10.1073/pnas.1502255112

Milic, D., Demidkina, T. V., Faleev, N. G., Matkovic-Cˇalogovic, D., & Antson, A. A. (2008). Insights into the Catalytic Mechanism of Tyrosine Phenol-lyase from X-ray Structures of Quinonoid Intermediates∗. *Journal of Biological Chemistry*, *283*(43), 29206-29214.

Milic, D., Demidkina, T. V., Faleev, N. G., Phillips, R. S., Matkovic-Calogovic, D., & Antson, A. A. (2011). Crystallographic snapshots of tyrosine phenol-lyase show that substrate strain plays a role in C–C bond cleavage. *Journal of the American Chemical Society*, *133*(41), 16468-16476.

Moon, S., & Case, D. A. (2006). A comparison of quantum chemical models for calculating NMR shielding parameters in peptides: Mixed basis set and ONIOM methods combined with a complete basis set extrapolation. *Journal of Computational Chemistry*, *27*(7), 825-836. https://doi.org/10.1002/jcc.20388

Nguyen, H., Roe, D. R., & Simmerling, C. (2013). Improved Generalized Born Solvent Model Parameters for Protein Simulations [Article]. *Journal of Chemical Theory and Computation*, *9*(4), 2020-2034. https://doi.org/10.1021/ct3010485

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T. J., Fan, L., Mueller, L. J., & Dunn, M. F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: evidence for two subunit conformations and four quaternary States. *Biochemistry*, *52*(37), 6396-6411. https://doi.org/10.1021/bi400795e

Percudani, R., & Peracchi, A. (2003). A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO reports*, *4*(9), 850-854.

Perilla, J. R., Zhao, G., Lu, M., Ning, J., Hou, G., Byeon, I.-J. L., Gronenborn, A. M., Polenova, T., & Zhang, P. (2017). CryoEM Structure Refinement by Integrating NMR Chemical Shifts with Molecular Dynamics Simulations. *The Journal of Physical Chemistry B*, *121*(15), 3853-3863. https://doi.org/10.1021/acs.jpcb.6b13105

Phillips, R. S. (1987). Reactions of O-acyl-L-serines with tryptophanase, tyrosine phenol-lyase, and tryptophan synthase. *Archives of biochemistry and biophysics*, *256*(1), 302-310.

Phillips, R. S., Craig, S., Kovalevsky, A., Gerlits, O., Weiss, K., Iorgu, A. I., Heyes, D. J., & Hay, S. (2019). Pressure and temperature effects on the formation of aminoacrylate intermediates of tyrosine phenol-lyase demonstrate reaction dynamics. *ACS Catalysis*, *10*(3), 1692-1703.

Phillips, R. S., Vita, A., Spivey, J. B., Rudloff, A. P., Driscoll, M. D., & Hay, S. (2016). Ground-state destabilization by Phe-448 and Phe-449 contributes to tyrosine phenol-lyase catalysis. *ACS Catalysis*, *6*(10), 6770-6779.

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, *9*(7), 3084-3095. https://doi.org/10.1021/ct400341p

Salnikov, E. S., Abel, S., Karthikeyan, G., Karoui, H., Aussenac, F., Tordo, P., Bechinger, B., & Ouari, O. (2017). Dynamic Nuclear Polarization/Solid-State NMR Spectroscopy of Membrane Polypeptides: Free-Radical Optimization for Matrix-Free Lipid Bilayer Samples. *ChemPhysChem*, *18*(15), 2103-2113.

Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald [research-article]. *J Chem Theory Comput*, *9*(9), 3878-3888. https://doi.org/10.1021/ct400314y

Schnackerz, K. D., Andi, B., & Cook, P. F. (2011). ^{31}P NMR spectroscopy senses the microenvironment of the 5'-phosphate group of enzyme-bound pyridoxal 5'-phosphate. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1814*(11), 1447-1458. https://doi.org/10.1016/j.bbapap.2011.02.001

Steffen-Munsberg, F., Vickers, C., Kohls, H., Land, H., Mallin, H., Nobili, A., Skalden, L., van den Bergh, T., Joosten, H.-J., & Berglund, P. (2015). Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. *Biotechnology advances*, *33*(5), 566-604.

Tomanicek, S. J., Standaert, R. F., Weiss, K. L., Ostermann, A., Schrader, T. E., Ng, J. D., & Coates, L. (2013). Neutron and X-ray Crystal Structures of a Perdeuterated Enzyme Inhibitor Complex Reveal the Catalytic Proton Network of the Toho-1 β-Lactamase for the Acylation Reaction *. *Journal of Biological Chemistry*, *288*(7), 4715-4722. https://doi.org/10.1074/jbc.M112.436238

Toney, M. D. (2014). Aspartate aminotransferase: an old dog teaches new tricks. *Archives of biochemistry and biophysics*, *544*, 119-127.

Tran, J. U., & Brown, B. L. (2022). Structural Basis for Allostery in PLP-dependent Enzymes [Review]. *Frontiers in Molecular Biosciences*, *9*. https://doi.org/10.3389/fmolb.2022.884281

Tran, N. T., Mentink-Vigier, F., & Long, J. R. (2020). Dynamic Nuclear Polarization of Biomembrane Assemblies. *Biomolecules*, *10*(9), 1246.

Walsh, C. (1979). *Enzymatic reaction mechanisms*. WH Freeman.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, *25*(9), 1157-1174.

Wang, X., Caulkins, B. G., Riviere, G., Mueller, L. J., Mentink-Vigier, F., & Long, J. R. (2019). Direct dynamic nuclear polarization of ^{15}N and ^{13}C spins at 14.1 T using a trityl radical and magic angle spinning. *Solid State Nuclear Magnetic Resonance*, *100*, 85-91.

Chapter 9

Positional Variance and Model Quality in NMR Crystallography: Anisotropic Displacement Parameters

9.1 Introduction

Enzyme chemistry relies heavily on understanding the relationship between a molecule's structure and function, and advancements in molecular science and enzymology in the last 50 years have been made possible by the ability to determine three-dimensional molecular structures (Orengo et al., 1999; Petsko, 2004). While single-crystal X-ray diffraction has been the traditional method used to characterize proteins and other molecules, it remains challenging to obtain accurate structural information from a single crystal (Harris & Cheung, 2004). This poses a problem for applications such as structural biology and pharmaceutical development, where the determination of correct structural information on protein-ligand complexes, enzyme-cofactor complexes, and drug polymorphs is crucial (Holzgrabe et al., 1999). However, solid-state nuclear magnetic resonance (NMR) can overcome this challenge by detecting local atomic environments regardless of long-range order (Dudenko et al., 2012). Significant progress has been made in this area with methods that can directly measure dipolar couplings(Partridge et al., 2015). The combination of solid-state NMR and computational methods has also advanced considerably over the last decade, with numerous examples of structure validation using chemical shift measurements in combination with density functional theory (DFT) calculations (Goward et al., 2003; Harris & Cheung, 2004; Harris, 2006; Ochsenfeld et al., 2001; Salager et al., 2010). Recent examples of de novo structure determination have also emerged, utilizing NMR shifts, DFT shift calculations, and crystal structure prediction (Baias et al., 2013; Baias et al., 2015; Salager et al., 2010).

However, unlike XRD methods, there is currently no established protocol for quantifying the positional errors of individual atoms in structures determined by chemical-shift-based NMR crystallography.

Here, we introduce a method, extended from Lyndon Emsley's group based on molecular dynamics (MD) simulations, and DFT calculations, to estimate the correlation between the root-mean-square deviation (RMSD) of the experimental and calculated shifts and the variances of atomic positions of individual atoms in structures determined by NMR crystallography, thereby making them directly comparable to structures determined by other methods (Baias et al., 2015; Hofstetter & Emsley, 2017). The approach is demonstrated on NMR Crystallography refined structures of intermediates in Tryptophan Synthase (TS) previously characterized by NMR crystallography (Jacob B Holmes et al., 2022). First, we generate an ensemble of slightly perturbed crystal structures with MD simulations at finite temperatures. By "slightly perturbed" we refer to structures that remain within the same local minima and do not undergo any significant conformational shifts. The temperature ranges used, and the methodology are detailed in the Experimental section.

Predicted ^{1}H and ^{13}C chemical shifts are then calculated for the members of the ensemble using GIAO DFT method (Kupka et al., 1999; Ośmiałowski et al., 2001). Given the estimated errors in the measured and predicted chemical shifts, we then correlate this directly with the atomic positions that are compatible with the measured chemical shifts to within the error, yielding a distribution of positions for each atom. The positional distributions are then converted into anisotropic displacement parameters (ADPs), which can be represented by ellipsoids on the determined structure. Here we use this

approach on the E(A-A) and E(A-A)(BZI) of TS and will be tested on other enzyme systems in the future.

9.2 Experimental Section

The positional uncertainties in the NMR-refined crystal structures were estimated following the method of Hofstetter and Emsley (Hofstetter & Emsley, 2017), adapted for NMR crystallography within an enzyme active site. The protocol is summarized in the following steps:

1. An ensemble of slightly perturbed structures was generated using low-temperature molecular dynamics

Low-T MD for E(A-A) and E(A-A)(BZI) were performed on the TS αβ dimer using the Amber package with GPU acceleration (Case et al., 2018; Salomon-Ferrer et al., 2013). Coordinates for *StTS* were obtained from crystal structures 4HN4 for E(A-A) and 4HPX for E(A-A)(BZI) with crystal waters retained (D. Niks et al., 2013) along with the refined active site and protonation states determined via NMR crystallography. Amber force field FF14SB was applied to protein atoms (Maier et al., 2015). Ligands were parametrized using the General Amber Force Field (GAFF) and AM1-bcc charge model (Jakalian et al., 2002; Wang et al., 2004).

Both the E(A-A) and E(A-A)(BZI) systems were minimized in a three-step process: hydrogens, side chains, and full protein, with the refined region within 6Å of the β subunit ligand “0JO” (the PLP cofactor + covalently attached ligands) restrained. Explicit TIP3P solvent was added to obtain a rectangular water box with edges at least 12 Å from any protein atom. Particle Mesh Ewald and SHAKE algorithms were applied for long range interactions and fixed heavy-atom-hydrogen bonds, respectively (Ryckaert et al., 1977;

Sagui et al., 2004). Positive ions (Na^+) were added to neutralize the overall charge of the proteins. The solvated systems were minimized again in two steps: water molecules only, and then the whole system, with the same region within 6 Å restrained.

Next, the systems were gradually heated at constant pressure from 5K to 150K at 5K intervals. 300 ps simulations were carried out at each temperature with a time step of 2 fs using the thermal fluctuation configuration ensemble. Trajectories were processed to contain 300 frames (1ps per frame) for each interval. The final ensemble of perturbed crystal structures was composed of 1200 structures, 300 structures each at 5, 10, 25 and 45 K.

2. First principles chemical shifts and reduced-χ^2 were calculated for the active site region of each structure

For each of the 1200 structures in the ensemble, the active site region consisting of the PLP-substrate covalent complex, the 3 adjacent waters or BZI, and residues βSer377(C^{α}), βHis86(C^{β}), βLys87(C^{γ}), βGlu109, βThr110, βGly111, βAla112, βGly232, βGly233, βGly234, βGly235, βAsn236, βGln114(bb), and βHis115(bb) were extracted for NMR chemical shift calculations. Only the side chain atoms of residues βSer377, βHis86, and βLys87 were included, truncated at the indicated atom and capped with hydrogen. Only the backbone atoms and side chains through C^{β} were included for βGln114 and βHis115. Terminal residues in contiguous chains were capped with hydrogen atoms. E(A-A) structures contained 192 atoms and E(A-A)(BZI) structures contained 199.

$$\chi_r^2(\text{model}) = \frac{1}{N-f}\sum_i^{\sum}\frac{(\delta_i^{\text{model}} - \delta_i^{exp}()^2)}{s_i^2} \quad (9.1)$$

For each structure, first principles chemical shifts were calculated at the DFT level of theory (B3LYP/6-31G). These shifts were used to calculate a reduced-χ^2 relative to a reference structure using Eq. 9.1; only atoms for which experimental chemical shifts were measured were included in the reduced-χ^2. The reference structure was chosen from the ensemble using the selection criterion that the sum of the reduced-χ^2 for all other structures was minimized. A separate reference structure was chosen at each temperature.

3. Anisotropic displacement parameters were defined at each atom by the correlation of positional deviations with reduced-χ^2

The positional deviations of each atom relative to its temperature-specific reference structure were used to define a covariance matrix of displacements and an atom-specific principal axis system (Hofstetter & Emsley, 2017). The correlation of positional deviations along a PAS axes (r_i) and the corresponding reduced-χ^2 was used to fit a 2D probability density at each atom (I)

$$G(r_{i,I}, \chi_r^2) = \frac{A_{i,I}\left(\chi_r^2\right)}{\sqrt{2\pi\sigma_{i,I}^2\chi_r^2}} \exp\left\{-\frac{r_{i,I}^2}{2\sigma_{i,I}^2\chi_r^2}\right\} \tag{9.2}$$

Here $A_{i,I}\left(\chi_r^2\right)$ is an arbitrary normalization that accounts for different numbers of samples at the various reduced-χ^2 values. Eq. (9.2) describes a Gaussian distribution with a width that is linearly dependent on $\sqrt{\chi_r^2}$; the dependence of width on χ_r^2 (or in their case RMSD) is a key result from Hofstetter and Emsley (Hofstetter & Emsley, 2017).

Fig 9.1 shows an example of the scatter plot of positional variance and reduced-χ^2. The NMR-derived anisotropic displacement parameters (ADP) were defined as the Gaussian widths in the principal axis system positional distributions drawn at the 95% confidence limit of the chemical shift reduced-χ^2

$$U_{ii,l}^{\mathrm{PAS}} = \sigma_{i,l}^2 \chi_{r,95\%}^2 \tag{9.3}$$

These were transformed back to the coordinate frame of the crystal structure using standard tensor methods (Mueller, 2011; Young et al., 2019), and were incorporated into an edited version of the protein crystal structure file along with the refined active site coordinates from NMR-assisted crystallography. The active sites with positional variances drawn at the 95% probability level are shown in Fig 9.2.

In the text, the average positional RMSD for a given atom type is reported as

$$\langle r_{\mathrm{av}} \rangle = \sqrt{3U_{\mathrm{eq}}} \tag{9.4}$$

where $U_{\mathrm{eq}} = \frac{1}{N}\sum_{l=1}^{N} U_{\mathrm{eq}}^{l}$ (9.5)

and $U_{eq}^{l} = \frac{1}{3}\left(U_{11,l}^{\mathrm{PAS}} + U_{22,l}^{\mathrm{PAS}} + U_{33,l}^{\mathrm{PAS}}\right)$ (9.6)

9.3 Results and Discussion

In this preliminary analysis, ADPs are reported only for atoms with measured chemical shifts and their covalently- or hydrogen-bonded neighbors. This restriction can likely be relaxed, a subject of ongoing work. The current restriction recognizes a fundamental difference in the application of NMR crystallography to an active site compared to its application in organic molecular crystals (Cordova et al., 2021; Hofstetter & Emsley, 2017): active sites are typically modeled using a cluster-based approach in which the measured chemical shifts are for atoms concentrated at the center of the model, while molecular crystals are typically modeled in a unit-cell based approach and have measured chemical shifts distributed throughout. Displacements of atoms at the periphery of cluster models are expected to have little influence on the measured chemical shifts at the center of the cluster, and therefore are expected to be unconstrained. To better quantify this, the sensitivity of the reduced-χ^2 to atomic displacements was assessed by forming its gradient with respect to single atom displacements using finite differences. It is known that single atom displacements lead to unphysical models and unrealistically large changes in chemical shifts within a covalent network (Hofstetter & Emsley, 2017), but this is expected to be a less severe problem for displacements outside of the shift region. These simulations confirm that atoms whose ADP have been reported show the largest gradients – more than 3 orders of magnitude larger than for atoms at the edge of the model.

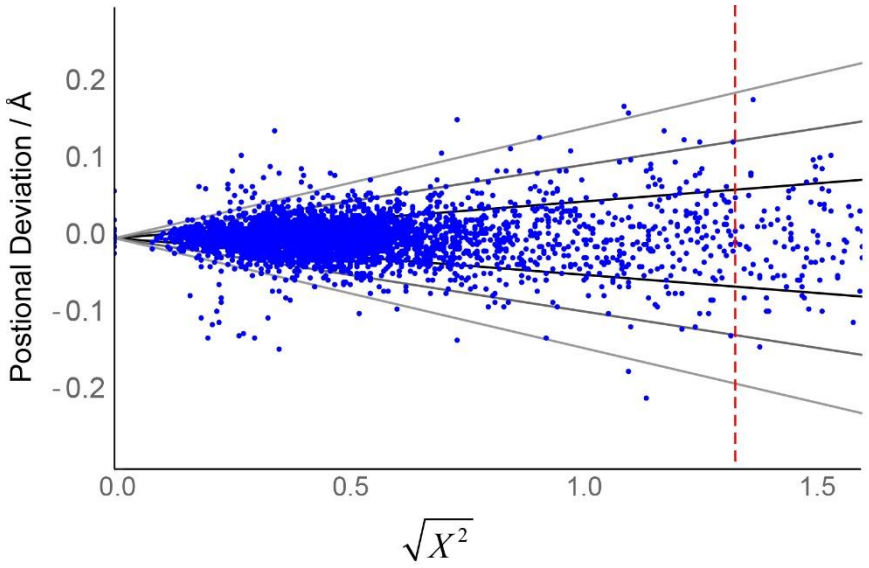


Fig 9.1: Scatter plot of $\sqrt{\chi_r^2}$ and positional deviation along one principal axis of the anisotropic displacement tensor for C^{α} of E(A-A)(BZI). The vertical line at $\sqrt{\chi_r^2} = \sqrt{1.79}$ shows the 95% confidence cutoff in the reduced-χ^2. The black and gray lines indicate ±1, 2, and 3 times the width of the Gaussian distribution at a given reduced-χ^2.

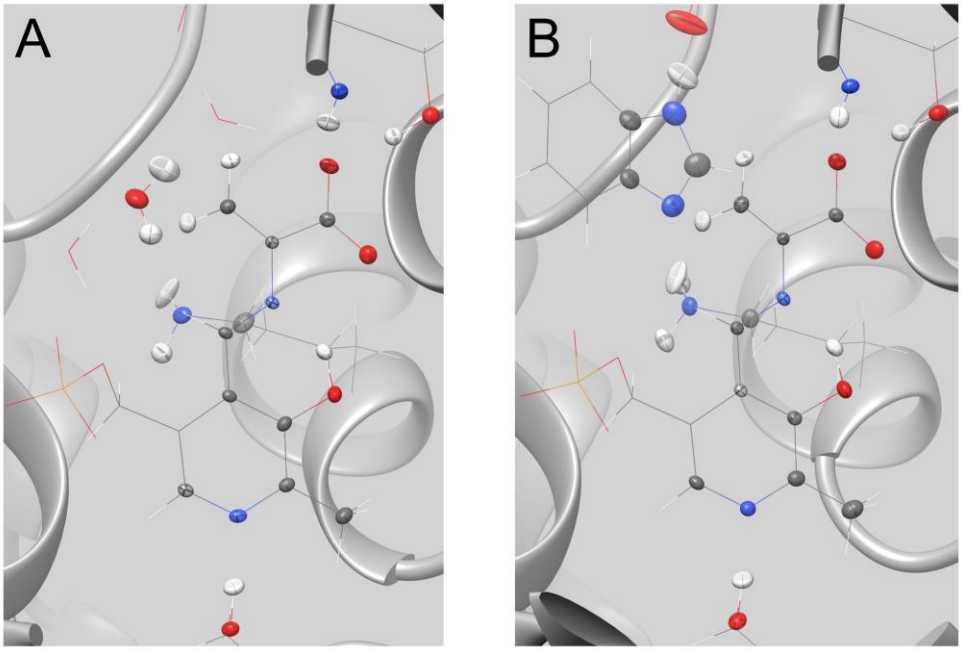


Fig 9.2: Anisotropic displacement parameters for (A) E(A-A) and (B) E(A-A)(BZI) displayed for atoms with experimental chemical shifts and their neighboring atoms. ADP ellipsoids are drawn at the 95% probability level (scale factor 2.8). The CPK scheme is used to designate atoms and their ellipsoids (C grey; H white; O red; P orange). Images rendered in UCSF Chimera (Petsko, 2004; Pettersen et al., 2004b).

9.4 Conclusions

To sum up, we have presented a novel approach for determining the positional uncertainties in crystal structures of enzyme active sites obtained from NMR chemical shifts. While our method was developed specifically for chemical-shift-based NMR crystallography, it can be applied to structures obtained from other methods, such as XRD. By generating an ensemble of structures around the experimentally determined structure, we can predict chemical shift deviations and compare them to positional deviations, which allows us to determine the average positional error for each atom in the crystal structure. Our results demonstrate that the average positional uncertainty in the five structures studied here is comparable to molecular solids, with an RMSD of 0.17 Å for 1H and 0.11 Å for heavy atoms. We also find that chemical-shift-based NMR

crystallography methods provide a significant improvement in positional accuracy compared to XRD, with a gain of approximately a factor of 6.5, primarily due to the lack of limitations posed by thermal vibrations.

9.5 References

Baias, M., Dumez, J.-N., Svensson, P. H., Schantz, S., Day, G. M., & Emsley, L. (2013). De novo determination of the crystal structure of a large drug molecule by crystal structure prediction-based powder NMR crystallography. *Journal of the American Chemical Society*, *135*(46), 17501-17507.

Baias, M., Lesage, A., Aguado, S., Canivet, J., Moizan-Basle, V., Audebrand, N., Farrusseng, D., & Emsley, L. (2015). Superstructure of a Substituted Zeolitic Imidazolate Metal–Organic Framework Determined by Combining Proton Solid-State NMR Spectroscopy and DFT Calculations. *Angewandte Chemie*, *127*(20), 6069-6074.

Case, D. A., Ben-Shalom, Ido, Brozell, R., S., Cerutti, David, Cheatham, III, T. E., Cruzeiro, Wilian, V., Darden, T., Duke, R., E., Ghoreishi, Delaram, Gilson, M., Gohlke, Holger, . . . A., P. (2018). *AMBER 2018*. In University of California.

Cordova, M., Balodis, M., Hofstetter, A., Paruzzo, F., Nilsson Lill, S. O., Eriksson, E. S. E., Berruyer, P., Simões de Almeida, B., Quayle, M. J., Norberg, S. T., Svensk Ankarberg, A., Schantz, S., & Emsley, L. (2021). Structure determination of an amorphous drug through large-scale NMR predictions. *Nature Communications*, *12*(1), 2964. https://doi.org/10.1038/s41467-021-23208-7

Dudenko, D., Kiersnowski, A., Shu, J., Pisula, W., Sebastiani, D., Spiess, H. W., & Hansen, M. R. (2012). A strategy for revealing the packing in semicrystalline π-conjugated polymers: crystal structure of bulk poly-3-hexyl-thiophene (P3HT). *Angewandte Chemie*, *124*(44), 11230-11234.

Goward, G. R., Sebastiani, D., Schnell, I., Spiess, H. W., Kim, H.-D., & Ishida, H. (2003). Benzoxazine oligomers: Evidence for a helical structure from solid-state NMR spectroscopy and DFT-based dynamics and chemical shift calculations. *Journal of the American Chemical Society*, *125*(19), 5792-5800.

Harris, K. D. M., & Cheung, E. Y. (2004). How to determine structures when single crystals cannot be grown: opportunities for structure determination of molecular materials using powder diffraction data [10.1039/B409059B]. *Chemical Society Reviews*, *33*(8), 526-538. https://doi.org/10.1039/B409059B

Harris, R. K. (2006). NMR studies of organic polymorphs & solvates. *Analyst*, *131*(3), 351-373.

Hofstetter, A., & Emsley, L. (2017). Positional Variance in NMR Crystallography. *Journal of the American Chemical Society*, *139*(7), 2573-2576. https://doi.org/10.1021/jacs.6b12705

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., Ghosh, R. K., Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., & Bogie, P. M. (2022). Imaging active site chemistry

and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *Proceedings of the National Academy of Sciences*, *119*(2), e2109235119.

Holzgrabe, U., Wawer, I., & Diehl, B. (1999). *NMR spectroscopy in drug development and analysis*. John Wiley & Sons.

Jakalian, A., Jack, D. B., & Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry*, *23*(16), 1623-1641. https://doi.org/10.1002/jcc.10128

Kupka, T., Pasterna, G., Lodowski, P., & Szeja, W. (1999). GIAO–DFT prediction of accurate NMR parameters in selected glucose derivatives. *Magnetic Resonance in Chemistry*, *37*(6), 421-426.

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB [research-article]. https://doi.org/10.1021/acs.jctc.5b00255

Mueller, L. J. (2011). Tensors and Rotations in NMR. *Concepts in Magnetic Resonance Part A*, *38A*(5), 221-235. https://doi.org/10.1002/cmr.a.20224

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T. J., Fan, L., Mueller, L. J., & Dunn, M. F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: evidence for two subunit conformations and four quaternary States. *Biochemistry*, *52*(37), 6396-6411. https://doi.org/10.1021/bi400795e

Ochsenfeld, C., Brown, S. P., Schnell, I., Gauss, J., & Spiess, H. W. (2001). Structure assignment in the solid state by the coupling of quantum chemical calculations with NMR experiments: a columnar hexabenzocoronene derivative. *Journal of the American Chemical Society*, *123*(11), 2597-2606.

Orengo, C. A., Todd, A. E., & Thornton, J. M. (1999). From protein structure to function. *Current Opinion in Structural Biology*, *9*(3), 374-382. https://doi.org/https://doi.org/10.1016/S0959-440X(99)80051-7

Ośmiałowski, B., Kolehmainen, E., & Gawinecki, R. (2001). GIAO/DFT calculated chemical shifts of tautomeric species. 2-Phenacylpyridines and (Z)-2-(2-hydroxy-2-phenylvinyl) pyridines. *Magnetic Resonance in Chemistry*, *39*(6), 334-340.

Partridge, B. E., Leowanawat, P., Aqad, E., Imam, M. R., Sun, H.-J., Peterca, M., Heiney, P. A., Graf, R., Spiess, H. W., & Zeng, X. (2015). Increasing 3D supramolecular order by decreasing molecular order. A comparative study of helical assemblies of dendronized nonchlorinated and tetrachlorinated perylene bisimides. *Journal of the American Chemical Society*, *137*(15), 5210-5224.

Petsko, G. A., & Ringe, D. . (2004). *Protein structure and function*. New Science Press.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605-1612. https://doi.org/10.1002/jcc.20084

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327-341. https://doi.org/https://doi.org/10.1016/0021-9991(77)90098-5

Sagui, C., Pedersen, L. G., & Darden, T. A. (2004). Towards an accurate representation of electrostatics in classical force fields: efficient implementation of multipolar interactions in biomolecular simulations [research-article]. *Journal of Chemical Physics*, *120*(1), 73-87. https://doi.org/10.1063/1.1630791

Salager, E., Day, G. M., Stein, R. S., Pickard, C. J., Elena, B., & Emsley, L. (2010). Powder crystallography by combined crystal structure prediction and high-resolution 1H solid-state NMR spectroscopy. *Journal of the American Chemical Society*, *132*(8), 2564-2566.

Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald [research-article]. *J Chem Theory Comput*, *9*(9), 3878-3888. https://doi.org/10.1021/ct400314y

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, *25*(9), 1157-1174.

Young, R. P., Lewis, C. R., Yang, C., Wang, L., Harper, J. K., & Mueller, L. J. (2019). TensorView: A software tool for displaying NMR tensors. *Magnetic Resonance in Chemistry*, *57*(5), 211-223. https://doi.org/10.1002/mrc.4793

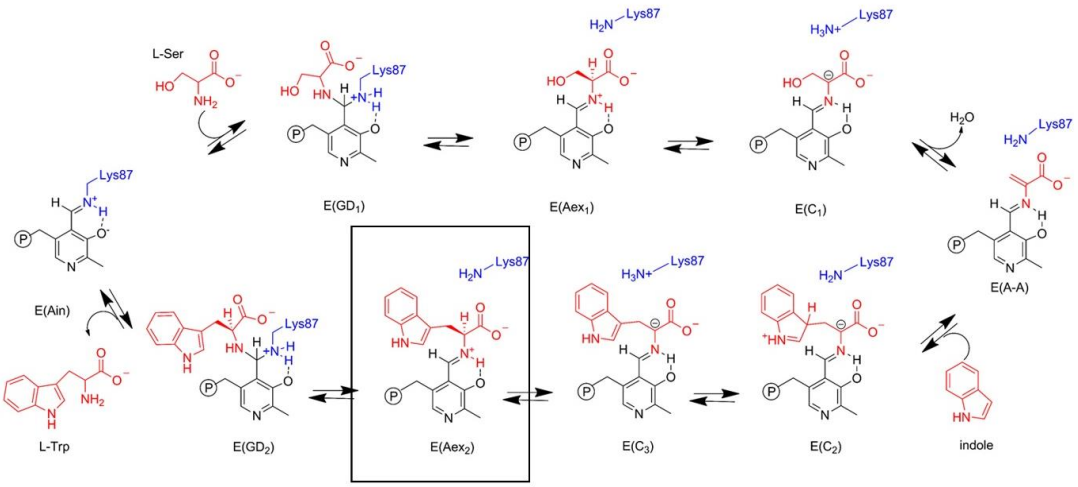
Chapter 10

Into the Future: Seeing the Bonds that X-ray Misses.

10.1 Introduction

The aim of this concluding chapter is to provide readers with a roadmap for the future of mechanistic enzymology, utilizing the innovative approach of "NMR Crystallography" in integrated structural biology. Within this chapter, we will explore two ongoing projects that will establish the basis for comprehending the relationships between structure and function in covalent adducts and intermediates that currently lack a crystal structure.

Our goal is to define the transformation of substrate to product with atomic resolution, determining the position of all atoms, including hydrogens (Mueller & Dunn, 2013). To accomplish this, our group is pioneering the development and application of NMR-assisted crystallography – the integrated application of solid-state NMR, X-ray crystallography, and computational chemistry – to enzyme active sites (Bryce & Taulelle, 2017; Facelli & Grant, 1993; Harris et al., 2009). Our traditional approach is three-fold. First, X-ray crystallography is used to provide a structural framework for the active site. Second, computational chemistry is used to build chemically-detailed models on this framework, and various active site chemistries are explored (Bethany G. Caulkins et al., 2014; B. G. Caulkins et al., 2016a; Joshua D. Hartman et al., 2015; Jacob B. Holmes et al., 2022; Jinfeng Lai et al., 2011). Third, these models are quantitatively distinguished by comparing their predicted NMR chemical shifts to the results of solid-state NMR experiments. Here we present our efforts to characterize intermediates that lack X-ray crystal structures, taking “NMR-assisted crystallography” to true “NMR crystallography.” Our target is the tryptophan external aldimine intermediate in tryptophan synthase.



Scheme 10.1 Proposed reaction mechanism in TS (the boxed intermediate is under investigation in this chapter).

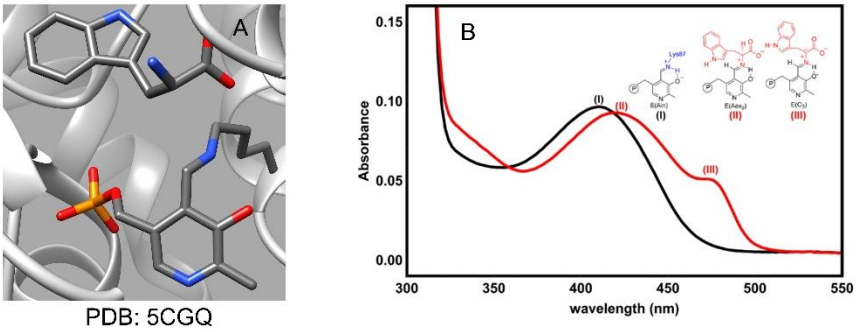


Fig 10.1 (A) Crystal structure of L-Trp in the active site of TS rotated away from the PLP cofactor. (B) Reaction of wild-type TS with L-Trp . A double-difference quartz cuvette was used to acquire the spectra, the solutions in the two compartments of the cuvette are designated as follows: (black line), the spectra before mixing (red line), spectra following mixing. The reactions were carried out in the presence of F9 (1 mM), with [TS], = 10 µM; when added, [L-Trp] = 50 mM.

External aldimines play a central role in catalytic transformations of amino acids performed by pyridoxal 5'- phosphate (PLP) enzymes. The X-ray crystal structure of the external aldimine formed by the reaction of TS with L-Trp (solved by our group and others) (Phillips & Harris, 2021) presents a curious problem: the nitrogen of the L-Trp is rotated away from the cofactor and no covalent bond is formed at 100 K. However, UV/Vis studies indicate that the bond to the cofactor forms at room temperature. We are

therefore seeing an intermediate that X-ray crystallography misses (Fig 10.1). To determine the structure via NMR crystallography, we present a new protocol that makes use of homology modelling rather than X-ray crystallography to define the initial structural framework of the active site. We then refine the coordinates of this close structural relative using first principles calculations and compare the predicted NMR chemical shifts to the experimental NMR chemical shifts measured throughout the active site. With enough shifts measured throughout the active site, we can determine both the chemical and full three-dimensional structure.

The resulting *de novo* NMR crystal structure offers several surprises. Most interestingly, the bound product is locked in a closed confirmation of the β subunit active and severely restricted by steric interactions within the active site. We posit that this is mechanistically significant, suggesting why product release is the rate limiting step in the overall catalytic transformation. The fact that X-ray crystallography is blind to the formation of the intermediate underscores the complementary nature of NMR crystallography to the other tools of structural biology.

10.2 Experimental Section

10.2.1 Protein Preparation. TS was prepared by overexpression of *S.typhimurium* in *E.coli* as previously described (Ghosh et al., 2021). Samples were prepared with the following isotopic labelling schemes for the cofactor and protein components: (1) protein ^{15}N enriched at lysine ε-nitrogen sites; PLP unlabeled (ε-^{15}N-Lys TS); (2) natural abundance protein and U-^{13}C-^{15}N L-Trp.

Microcrystalline protein samples for Solid-State NMR. Microcrystalline samples of TS were prepared by diluting enzyme solution 1:1 with 50 mM Cs^+ Bicine buffer, supplemented with 50 mM CsCl, pH 7.8, containing 20% PEG-8000 and 3 mM

spermine. Microcrystals were collected and washed with 50 mM Cs^+ Bicine, buffer, pH 7.8, containing 15% PEG-8000 and 2 mM spermine and 15 mM L-Trp and 1 mM high affinity α-site ligand F9 (Ngo, Harris, et al., 2007). The crystals were packed at 10 000 rpm into a Bruker 4 mm magic-angle spinning (MAS) rotor with an approximate volume of 50 µL; each rotor contained 25-30 mg of protein.

10.2.2 NMR Spectroscopy. ^{13}C and ^{15}N Solid-State NMR Spectroscopy.

^{13}C and `^{15}N cross-polarization (CP) magic angle spinning (MAS) experiments were performed at 9.4 T (400.37 MHz ^{1}H, 100.69 MHz ^{13}C, 40.57 MHz ^{15}N) on a Bruker AVIII spectrometer equipped with a double resonance, 4 mm MAS probe, spinning at standard MAS rates of 8 kHz; the bearing gas was cooled to -15 ^{0}C, giving an effective sample temperature of ~ -5 ^{0}C. Cross-polarization was accomplished at ^{1} (^{13}C) and 37 kHz (^{15}N) (ramped ±5 kHz); 85 kHz Spinal64 ^{1}H decoupling was used throughout. Standard ^{15}N spectra consist of the sum of ~500,000 transients acquired at a relaxation delay of 4 s, for a total acquisition time of 5 d 20 h. ^{15}N chemical shifts were referenced indirectly to liq-NH_3 (25 ^{0}C) via an external solid-state sample of $^{15}NH_4Cl$, in which the resonance frequency was set to 39.27 ppm. Standard ^{13}C spectra consist of the sum of ~500,000 transients acquired with a relaxation delay of 4 s, for a total time of 1 d 20 h. ^{13}C chemical shifts were referenced indirectly to neat TMS via an external solid-state sample of adamantane with the downfield-shifted peak set to 38.48 ppm.

The acquisition of solid-state NMR spectra was interleaved with single pulse, low-power decoupling experiments (64 scans ^{13}C, 256 scans ^{15}N) reporting predominantly on free ligand and reaction products (if any) in the mother liquor.

10.3 Results and Discussion

10.3.1 NMR Spectroscopy

ssNMR experiments were performed on microcrystalline protein samples prepared in close analogy to the single crystals used for X-ray crystallography. Microcrystals of TS were prepared in analogous manner to the single crystal for diffraction studies to maintain the same crystal habit as the larger crystals (Ahmed et al., 1987), giving high confidence that the NMR and X-ray data can be directly compared.

$E(Aex_2)$ can be observed with ssNMR under conditions that favor the accumulation of the enzyme-bound species while the microcrystals remain catalytically active. ssNMR spectra of the reaction of L-Trp with microcrystalline TS to form $E(Aex_2)$. The signals in the cross-polarization magic-angle-spinning (CPMAS) spectra correspond to crystalline protein and bound substrate. The spectra were acquired with various combinations of ^{15}N and ^{13}C labels on the L-Trp substrate, and protein Lys ε-amino groups. With unlabeled substrate and protein, the spectra show only unresolved background signals from the protein, including the backbone amide nitrogens (~120 ppm) and backbone carbonyl (~170 ppm), aliphatic (~10-70 ppm), and aromatic (~110-160 ppm) carbons.

1D ^{15}N and ^{13}C SSNMR CPMAS Spectroscopy

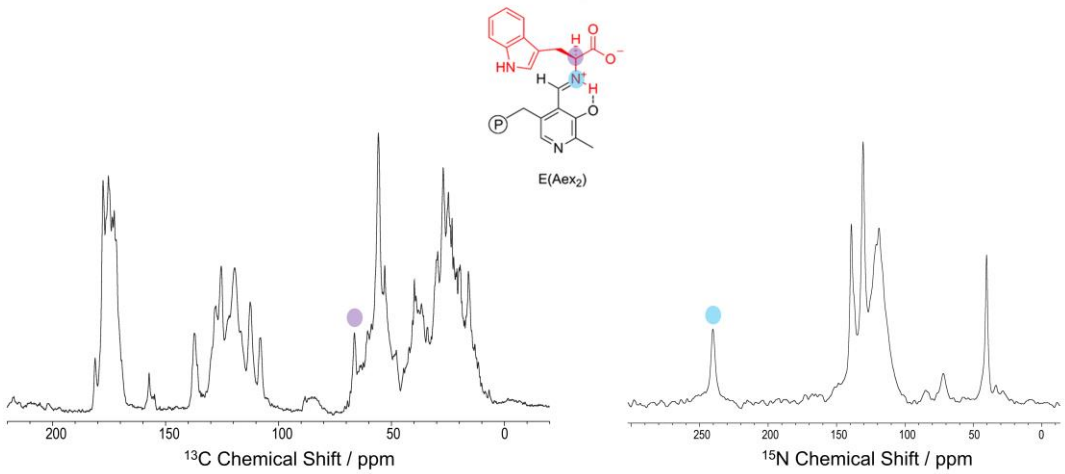


Fig 10.2 (*Left*) 1D ^{13}C CPMAS NMR spectra of natural abundance TS reacted with U-^{13}C-^{15}N-LTrp showing unique resonance at 66 ppm tentatively assigned to the Cα of the bound intermediate. (*Right*) 1D ^{15}N CPMAS NMR spectra of natural abundance TS reacted with U-^{13}C-^{15}N-LTrp showing unique resonance at 240 ppm assigned to the SB N linkage in E(Aex_2).

With the magnification of the isotopic enrichment, distinct resonances for the active site species become visible. 1D ^{15}N CPMAS NMR spectra collected at room temperature on the natural abundance protein sample prepared with uniformly ^{13}C-^{15}N enriched L-Trp shows a unique resonance at 240.1 ppm which is tentatively assigned to the SB N linkage between the α- amino group of the incoming amino acid, L-Trp and the C4' carbon of the PLP cofactor.

However, we observe distinct ^{13}C resonance at 66 ppm from the bound intermediate at room temperature. We take advantage of multi-dimension correlation spectroscopy to take advantage of direct bonded connections both in carbon and nitrogen dimensions (Fig 10.3) allowing for the assignments of the carbon and the nitrogen on the bound species.

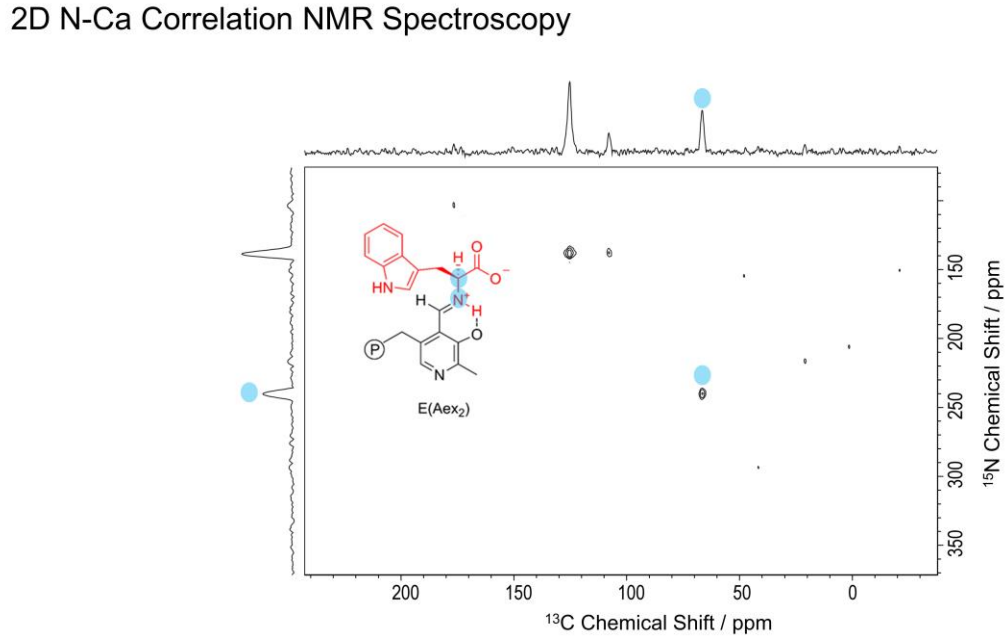


Fig 10.3 ^{15}N CPMAS RT SSNMR spectra of microcrystalline TS E(Aex$_2$) prepared with the following isotopic labeling: (i) natural abundance TS reacted with U-^{13}C-^{15}N L-Trp; NCa correlation spectroscopy collected on the same sample prepared with uniformly ^{13}C and ^{15}N enriched L-Trp substrate show correlations in the ^{13}C dimension to the SB N (*240.1 ppm*).

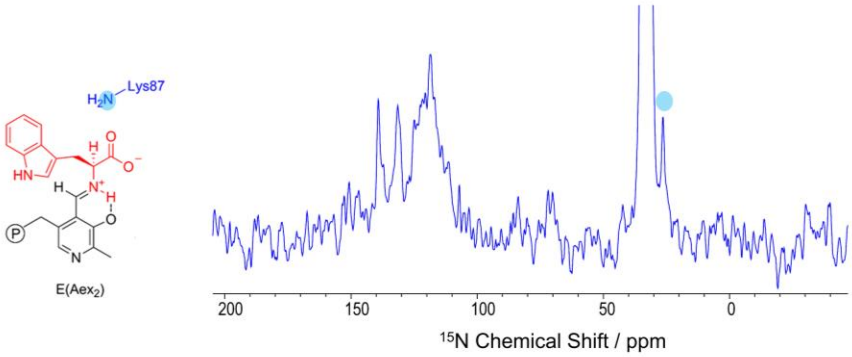


Fig 10.4 ^{15}N CPMAS RT SSNMR spectra of microcrystalline TS E(Aex$_2$) prepared with the following isotopic labeling: (i) ε-^{15}N-Lys TS reacted with U-^{13}C-^{15}N L-Trp; a unique single resonance shows up at 27.1 ppm assigned to the deprotonated ε-^{15}N group of the βLys87.

The preliminary ^{15}N and ^{13}C SSNMR data collected on this sample establishes that the bound intermediate is present in the micro-crystalline state of the enzyme, an

observation which precludes the diffraction data. Based on chemical intuition and previous studies on TS(B. G. Caulkins et al., 2016a; Jacob B. Holmes et al., 2022; Jinfeng Lai et al., 2011; Mueller & Dunn, 2013), it is expected that the SB N is a rapid tautomeric proton exchange with phenolic oxygen and that the βLys87 is deprotonated and ready to make the gem-diamine of the L-Trp to release the L-Trp from the sites.

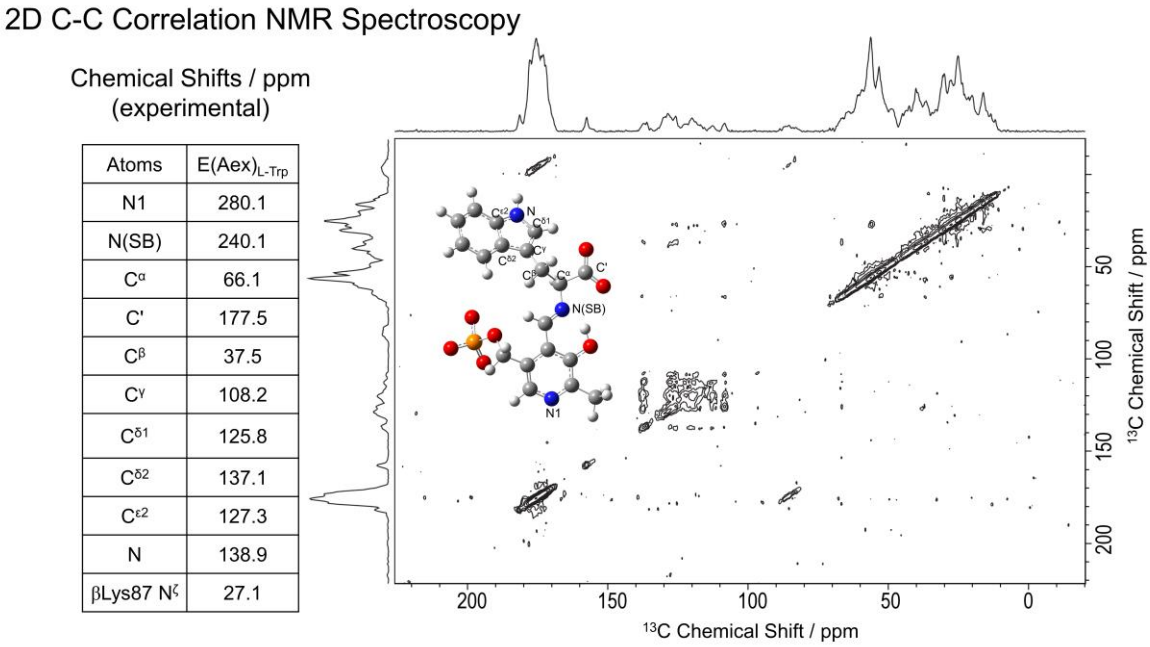


Fig 10.5 ^{13}C CPMAS RT SSNMR spectra of microcrystalline TS E(Aex_2) prepared with the following isotopic labeling: (i) natural abundance TS reacted with U-^{13}C-^{15}N L-Trp; C-C correlation spectroscopy (DARR mixing time = 50 msec) collected on the same sample prepared with uniformly ^{13}C and ^{15}N enriched L-Trp substrate show correlations for the carbons in the bound intermediate state.

^{13}C chemical shift at 66 ppm assigned via N-Ca (Fig 10.3) and C-C DARR (Fig 10.5) correlations establish the sp3 hybridization state of the bound E(Aex_2) intermediate, different from what is observed in E(A-A) or the E(C) complexes of TS or other PLP enzymes. This is expected based on the chemistry of the β-reaction where there is a proton attached to the Cα carbon.

^{31}P Spectroscopy

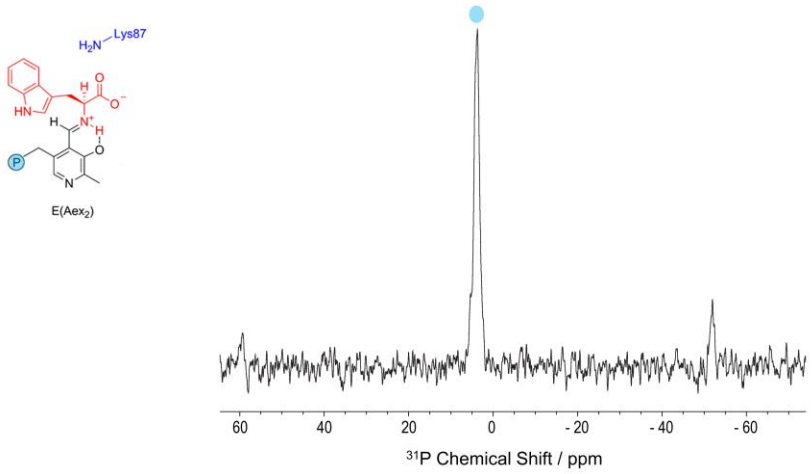


Fig 10.6 ^{31}P spectrum of the E(Aex)$_{L\text{-}Trp}$ intermediate in TS showing phosphorous resonances at 4.0 ppm.

The ^{31}P isotropic chemical shift were also measured for the cofactor's phosphate group for the (Aex$_2$) intermediate with an isotropic shift of 4.0 ppm respectively (Fig 10.6) indicating that it is in a dianionic regime, in keeping with other PLP-dependent enzymes (Schnackerz et al., 2011).

10.3.2 NMR First Principles Calculations

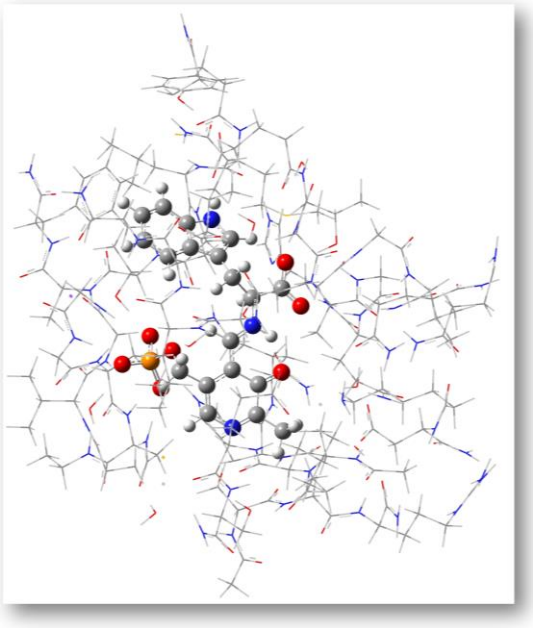


Fig 10.7 Active site model of the E(Aex_2) built from $E(C)_{indoline}$ crystal structure (Jinfeng Lai et al., 2011).

DFT First Principles calculations have been performed on the cofactor-substrate complex and the correlation of the experimental data and the calculated chemical shifts are underway.

10.4 Conclusions

To conclude, this is an ongoing work and we are trying new methodologies to calculate *denovo* structures of the active site in a large enzyme complex such as TS. This would alleviate some of the problems where there are no proper crystal structures for the enzymatic intermediates and initial guesses can be made using AlphaFold. This is where NMR Crystallography can play a central role in delineating structure function relationships in enzyme complexes.

10.5 References

Ahmed, S. A., Hyde, C. C., Thomas, G., & Miles, E. W. (1987). Microcrystals of tryptophan synthase $\alpha_2\beta_2$ complex from Salmonella typhimurium are catalytically active. *Biochemistry*, *26*(17), 5492-5498. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3314989

Bryce, D. L., & Taulelle, F. (2017). NMR crystallography. In (Vol. 73, pp. 126-127): International Union of Crystallography 5 Abbey Square, Chester, Cheshire CH1 ….

Caulkins, B. G., Bastin, B., Yang, C., Neubauer, T. J., Young, R. P., Hilario, E., Huang, Y.-m. M., Chang, C.-e. A., Fan, L., Dunn, M. F., Marsella, M. J., & Mueller, L. J. (2014). Protonation States of the Tryptophan Synthase Internal Aldimine Active Site from Solid-State NMR Spectroscopy: Direct Observation of the Protonated Schiff Base Linkage to Pyridoxal-5′-Phosphate. *Journal of the American Chemical Society*, *136*(37), 12824-12827. https://doi.org/10.1021/ja506267d

Caulkins, B. G., Young, R. P., Kudla, R. A., Yang, C., Bittbauer, T. J., Bastin, B., Hilario, E., Fan, L., Marsella, M. J., Dunn, M. F., & Mueller, L. J. (2016). NMR Crystallography of a Carbanionic Intermediate in Tryptophan Synthase: Chemical Structure, Tautomerization, and Reaction Specificity. *Journal of the American Chemical Society*, *138*(46), 15214-15226. https://doi.org/10.1021/jacs.6b08937

Facelli, J. C., & Grant, D. M. (1993). Determination of molecular symmetry in crystalline naphthalene using solid-state NMR. *Nature*, *365*(6444), 325-327. https://doi.org/10.1038/365325a0

Ghosh, R. K., Hilario, E., Liu, V., Wang, Y., Niks, D., Holmes, J. B., Sakhrani, V. V., Mueller, L. J., & Dunn, M. F. (2021). Mutation of βGln114 to Ala alters the stabilities of allosteric states in tryptophan synthase catalysis. *Biochemistry*, *60*(42), 3173-3186.

Harris, R. K., Wasylishen, R. E., & Duer, M. J. (2009). *NMR crystallography* (Vol. 4). John Wiley & Sons.

Hartman, J. D., Neubauer, T. J., Caulkins, B. G., Mueller, L. J., & Beran, G. J. O. (2015). Converging nuclear magnetic shielding calculations with respect to basis and system size in protein systems. *Journal of Biomolecular NMR*, *62*(3), 327-340. https://doi.org/10.1007/s10858-015-9947-2

Holmes, J. B., Liu, V., Caulkins, B. G., Hilario, E., Ghosh, R. K., Drago, V. N., Young, R. P., Romero, J. A., Gill, A. D., Bogie, P. M., Paulino, J., Wang, X., Riviere, G., Bosken, Y. K., Struppe, J., Hassan, A., Guidoulianov, J., Perrone, B., Mentink-

Vigier, F., . . . Mueller, L. J. (2022). Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *Proceedings of the National Academy of Sciences*, *119*(2), e2109235119. https://doi.org/doi:10.1073/pnas.2109235119

Lai, J., Niks, D., Wang, Y., Domratcheva, T., Barends, T. R. M., Schwarz, F., Olsen, R. A., Elliott, D. W., Fatmi, M. Q., Chang, C.-e. A., Schlichting, I., Dunn, M. F., & Mueller, L. J. (2011). X-ray and NMR Crystallography in an Enzyme Active Site: The Indoline Quinonoid Intermediate in Tryptophan Synthase. *Journal of the American Chemical Society*, *133*(1), 4-7. https://doi.org/10.1021/ja106555c

Mueller, L. J., & Dunn, M. F. (2013). NMR Crystallography of Enzyme Active Sites: Probing Chemically Detailed, Three-Dimensional Structure in Tryptophan Synthase. *Accounts of chemical research*, *46*(9), 2008-2017. https://doi.org/10.1021/ar3003333

Ngo, H., Harris, R., Kimmich, N., Casino, P., Niks, D., Blumenstein, L., Barends, T. R., Kulik, V., Weyand, M., & Schlichting, I. (2007). Synthesis and characterization of allosteric probes of substrate channeling in the tryptophan synthase bienzyme complex. *Biochemistry*, *46*(26), 7713-7727.

Phillips, R. S., & Harris, A. P. (2021). Structural basis of the stereochemistry of inhibition of tryptophan synthase by tryptophan and derivatives. *Biochemistry*, *60*(3), 231-244.

Schnackerz, K. D., Andi, B., & Cook, P. F. (2011). ^{31}P NMR spectroscopy senses the microenvironment of the 5'-phosphate group of enzyme-bound pyridoxal 5'-phosphate. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1814*(11), 1447-1458. https://doi.org/10.1016/j.bbapap.2011.02.001

Appendix 1
PLP Phosphorylation and Incorporation Protocol

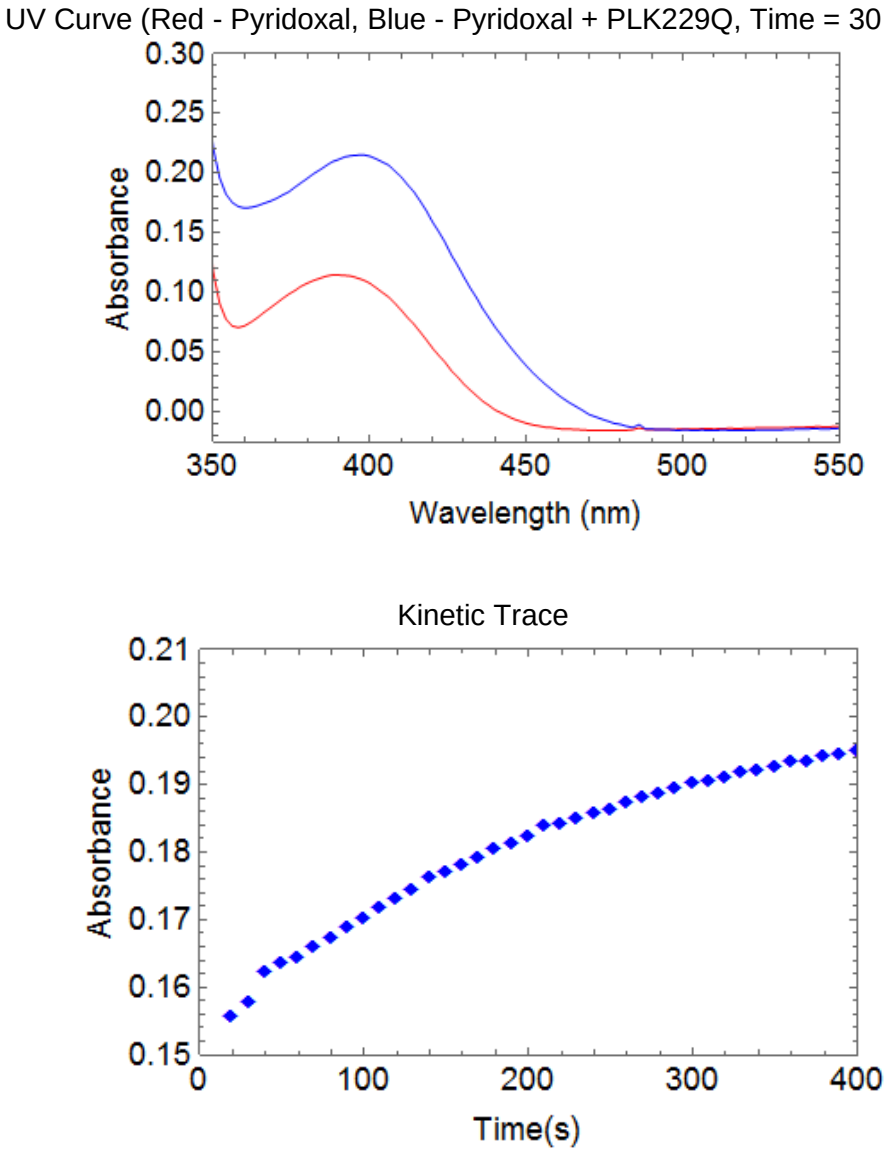


Fig A1. (*top*) Red curve showing the Pyridoxal (PL) absorption profile; blue curve showing the conversion taking place from PL to PLP (~1 minute time interval). *(bottom)* Kinetic trace of phosporylation event

Kevin Kou provided the 13C-PL (TFA salt) to us:
24.5 mgs (Mol wt. 321.24 g/mol)
Dissolved in water (3 mL)
Final stock concentration = 24.87 mM

PL Phosphorylation

- 200 uL 1M HEPES, pH 7.5
- 50 uL 1M KCl
- 320 uL 100 mM ATP
- 20 uL 20 mM MgCl2
- 160 uL 20 mM PL (when we have natural abundance PL)
- 100 uL PLKinase
- 110 uL dd water

After a few minutes, observe the color change from light yellow to very bright yellow.

Let it sit at room temperature for a day and the color stays and let the reaction complete (cover the tube).

PLP exchange into the enzyme:

- 1 mL (TS/SHMT/TPL) enzyme in eppendorf tube
- Add 10 uL of 1M $NH_2OH.HCl$ (pH 7-7.5) to the tube
- Should see the yellow color go away
- Mix it by moving up and down the tube
- Let it sit at room temperature for 5 minutes
- Start the dialysis in your buffer of choice (TS - Bicine; TPL - Phosphate; SHMT - Tris, all will have pH around 7.8 to 8.)
- o/n dialysis in 4c

Next Day

- Take the dialyzed sample out
- Add 100 uL PLP to the sample
- Mix it well
- Transfer it to water bath for 10 minutes (35c)
- Add substrate (15 uL 1M L-Ser for SHMT; TS, 15 uL 0.18 M L-Met for TPL)
- Mix well
- Transfer it to water bath for 10 minutes (35c)
- Transfer to ice bath for 30 minutes
- Dialyze o/n in buffer of choice (TS - Bicine; TPL - Phosphate; SHMT - Tris, all will have pH around 7.8 to 8.)
- o/n dialysis in 4c

Next Day

- Take the sample out from dialysis bag
- Run it down the centrifuge filter for 10 mins 4c, 3k rpm.
- If needed do the buffer exchange to crystallization buffer
- Crystallize the enzymes
- Store in -80c