Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate

Jacob B. Holmes[a,1], Viktoriia Liu[a,1], Bethany G. Caulkins[a,b,1], Eduardo Hilario[a], Rittik K. Ghosh[c], Victoria N. Drago[d], Robert P. Young[e], Jennifer A. Romero[a], Adam D. Gill[c], Paul M. Bogie[a], Joana Paulino[f], Xiaoling Wang[f], Gwladys Riviere[g], Yuliana K. Bosken[c], Jochem Struppe[h], Alia Hassan[i], Jevgeni Guidoulianov[i], Barbara Perrone[i], Frederic Mentink-Vigier[f], Chia-en A. Chang[a], Joanna R. Long[g], Richard J. Hooley[a,c], Timothy C. Mueser[d], Michael F. Dunn[c,2], and Leonard J. Mueller[a,2]

[a]Department of Chemistry, University of California, Riverside, CA 92521; [b]W.M. Keck Science Department, Claremont McKenna, Pitzer, and Scripps Colleges, Claremont, CA 91711; [c]Department of Biochemistry, University of California, Riverside, CA 92521; [d]Department of Chemistry and Biochemistry, University of Toledo, Toledo, OH 43606; [e]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354; [f]National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310; [g]Department of Biochemistry and Molecular Biology, McKnight Brain Institute, National High Magnetic Field Laboratory, University of Florida, Gainesville, FL 32610; [h]Bruker Biospin Corporation, Billerica, MA 01821; and [i]Bruker Switzerland AG 8117 Fällanden, Switzerland

NMR-assisted crystallography—the integrated application of solid-state NMR, X-ray crystallography, and first-principles computational chemistry—holds significant promise for mechanistic enzymology: by providing atomic-resolution characterization of stable intermediates in enzyme active sites, including hydrogen atom locations and tautomeric equilibria, NMR crystallography offers insight into both structure and chemical dynamics. Here, this integrated approach is used to characterize the tryptophan synthase α-aminoacrylate intermediate, a defining species for pyridoxal-5′-phosphate–dependent enzymes that catalyze β-elimination and replacement reactions. For this intermediate, NMR-assisted crystallography is able to identify the protonation states of the ionizable sites on the cofactor, substrate, and catalytic side chains as well as the location and orientation of crystallographic waters within the active site. Most notable is the water molecule immediately adjacent to the substrate β-carbon, which serves as a hydrogen bond donor to the ε-amino group of the acid–base catalytic residue βLys87. From this analysis, a detailed three-dimensional picture of structure and reactivity emerges, highlighting the fate of the L-serine hydroxyl leaving group and the reaction pathway back to the preceding transition state. Reaction of the α-aminoacrylate intermediate with benzimidazole, an isostere of the natural substrate indole, shows benzimidazole bound in the active site and poised for, but unable to initiate, the subsequent bond formation step. When modeled into the benzimidazole position, indole is positioned with C3 in contact with the α-aminoacrylate Cβ and aligned for nucleophilic attack. Here, the chemically detailed, three-dimensional structure from NMR-assisted crystallography is key to understanding why benzimidazole does not react, while indole does.

NMR-assisted crystallography | tryptophan synthase | pyridoxal-5′-phosphate | integrated structural biology | solid-state NMR

Pyridoxal-5′-phosphate (PLP; Fig. 1) participates in numerous enzyme-catalyzed reactions essential for amino acid metabolism, including transamination, decarboxylation, and α/β/γ-elimination and substitution (1–3). The power of PLP as a cofactor comes from its ability to act as an electron sink, allowing for the stabilization of carbanionic intermediates. A more subtle aspect of PLP chemistry demonstrated in β-elimination and replacement reactions is the ability of the cofactor to fine tune the polarity at the β-carbon of amino acids, facilitating the elimination of poor leaving groups and their replacement with weak nucleophiles.

Tryptophan synthase (TS) is the prototypical example of a PLP-dependent enzyme that catalyzes β-elimination and substitution reactions. The *Salmonella typhimurium* tryptophan synthase (*St*TS) studied here is a 143 kDa, αββα bienzyme complex (4). TS performs the final two steps in the biosynthesis of L-tryptophan: the α-subunit cleaves indole-3-glycerol 3′-phosphate to glyceraldehyde-3-phosphate and indole, while the β-subunit catalyzes the PLP-dependent β-elimination and replacement of the substrate L-Ser hydroxyl with indole to produce L-Trp (Fig. 1) (5, 6). The fidelity of the proton transfers in the TS catalytic cycle are critical for maintaining the β-elimination and substitution pathway. In addition to the stereoelectronic control of the initial bond-breaking step (7), it has been proposed that reaction specificity in PLP-dependent enzymes is modulated by the protonation states of the PLP cofactor–substrate complex (1, 3,

Significance

The determination of active site protonation states is critical for a full mechanistic understanding of enzymatic transformations. However, hydrogen atom positions are challenging to extract using the standard tools of structural biology. Here, we make use of a joint solid-state NMR, X-ray crystallography, and first-principles computational approach that enables the investigation of enzyme catalysis at this fine level of chemical detail. For tryptophan synthase, this allows us to peer along the reaction coordinates into and out of the α-aminoacrylate intermediate. Through this process, we are developing a high-resolution probe for structural biology that is keenly sensitive to hydrogen atom positions—complementing diffraction methods yet able to be applied under conditions of active catalysis in microcrystalline and non-crystalline materials.
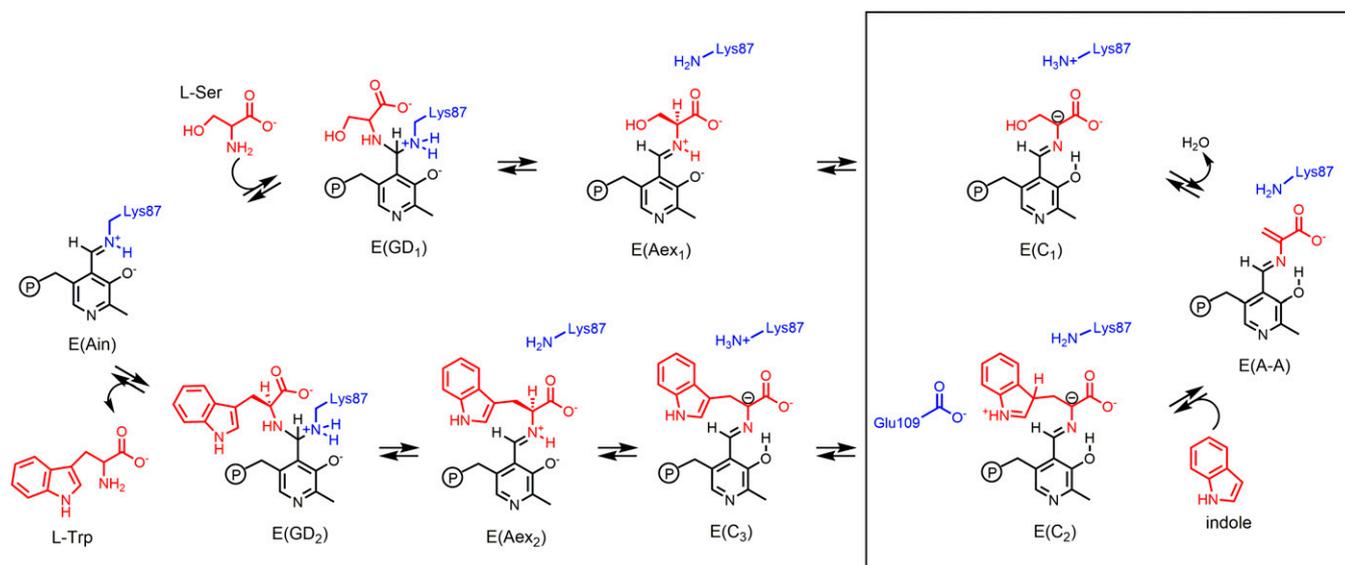
Fig. 1. TS β-site reaction, highlighting the α-aminoacrylate intermediate (41, 82). The PLP cofactor is drawn in black. In stage I of the reaction, L-Ser reacts with the internal aldimine E(Ain) to give the gem-diamine, E(GD$_1$), and external aldimine, E(Aex$_1$), intermediates. Subsequent proton abstraction via βLys87 leads to the first carbanionic intermediate, E(C$_1$), and elimination of the β-hydroxyl group as water gives the α-aminoacrylate intermediate, E(A-A). In stage II, indole makes a nucleophilic attack at C$^\beta$ of E(A-A) to yield a new C–C bond and the L-Trp carbanionic intermediate, E(C$_2$), and, upon deprotonation, E(C$_3$). In the final stages, E(C$_3$) is reprotonated, leading to the eventual release of L-Trp and the regeneration of E(Ain).

8), which are in turn directed by chemical interactions with acid–base groups such as βLys87 and βGlu109 in TS. Hence, the catalytic residues interacting with the cofactor establish the appropriate chemical and electrostatic environment to favor a particular protonation state and reaction pathway.

The determination of protonation states remains a significant challenge to the tools of structural biology. Even high-resolution X-ray crystal structures are challenged to place hydrogen atoms. Neutron crystallography can locate hydrogen atoms, offering remarkable insights into enzyme mechanism (9, 10). However, the requirements for neutron crystallography include large, preferably perdeuterated crystals and multiweek-long acquisition times at room temperature. The latter currently precludes analysis of reactive intermediates such as those studied here. Cryogenic electron microscopy (cryo-EM) is also pushing to the boundaries of resolution necessary for hydrogen atom detection (11), an extraordinary achievement yet currently far from routine. In combined application with NMR spectroscopy, however, cryo-EM and diffraction methods continue to work toward complete atomic-resolution descriptions of structure and function (12–24). For delineating the chemistry of the active site, NMR and diffraction become even more powerful when combined with first-principles computational chemistry (25–31).

We are developing NMR-assisted crystallography—the joint application of solid-state NMR (SSNMR), X-ray crystallography, and first-principles computational chemistry—to solve for the chemically rich, three-dimensional structures of enzyme active sites (26–28). By "chemically rich," we mean structures in which the location of all atoms, including hydrogens, are specified. NMR crystallography was originally developed within the context of molecular organic and inorganic solids (32–38). Our group (26–28) and others (25, 29–31) have been working to extend this approach to structural biology, in which it can provide consistent and testable models of enzyme structure and function. Our approach is threefold. First, X-ray crystallography is used to provide a structural framework for the active site. Second, chemically detailed models are built on this framework using computational chemistry, and various active site chemistries are explored. Third, these models are quantitatively

distinguished by comparing their predicted NMR chemical shifts to the results of SSNMR experiments. Provided that a sufficient number of chemical shifts are measured, NMR-assisted crystallography can identify a unique, consistent structure or, equally important, determine that none of the candidates is consistent with the experimental observations. For the case of the "quinonoid" intermediate in tryptophan synthase, this approach demonstrated that the intermediate is better described as a carbanionic species with a deprotonated pyridine ring nitrogen—a structure that is fundamental to understanding reaction specificity in TS (28).

Here, NMR-assisted crystallography is used to characterize the TS α-aminoacrylate intermediate [E(A-A); Fig. 1], a species that marks a divergent step in PLP catalysis, as only enzymes that perform β-elimination reactions generate this intermediate. For E(A-A), NMR-assisted crystallography is able to identify the protonation states of ionizable sites on the cofactor, substrates, and catalytic side chains as well as the location and orientation of active site waters. From this, a detailed, three-dimensional picture of structure and reactivity emerges, highlighting βLys87 as the acid–base catalytic residue and delineating the reaction coordinate for the elimination of the substrate β-hydroxyl. Subsequent characterization of the Michaelis complex formed with the indole isostere benzimidazole (BZI) shows BZI bound in the active site and poised for, but unable to initiate, the subsequent bond formation step. When modeled into the BZI position, indole is properly aligned for nucleophilic attack. The chemically rich structure from NMR-assisted crystallography is key to understanding why BZI does not react, while indole does.

Results and Discussion

X-Ray Crystallography. X-ray crystal structures for the *St*TS E(A-A) and E(A-A)(BZI) complexes have been reported by our group (39) and others (40). The formation of E(A-A) is characterized by both the α- and β-subunits adopting closed conformations (6, 39–43). The crystal structures for E(A-A) show three crystallographic waters adjacent to the serine substrate in the β-subunit active site, forming a hydrogen bonded

Holmes et al.
Imaging active site chemistry and protonation states: NMR crystallography
of the tryptophan synthase α-aminoacrylate intermediate

chain extending from the carboxylate of the catalytically essential βGlu109 (Fig. 2A). The position of the central water molecule is particularly striking, as it forms close contacts to both the substrate C^β (d_{CO} = 3.2 Å) and the βLys87 ε-amino group (d_{NO} = 3.0 Å). This places the water close to the site that the substrate β-hydroxyl is expected to occupy before it is eliminated. The binding of BZI (Fig. 2B) displaces the waters and causes a slight perturbation of the α-aminoacryloyl group. The binding also induces small movements of the βGlu109 carboxylate and the βLys87 ε-amino groups so that they are within hydrogen bonding distances of the BZI nitrogen atoms. These interactions and the close structural similarity of BZI to indole make the BZI complex a good mimic for the expected alignment of indole for nucleophilic attack at C^β. Absent in Fig. 2, however, are the hydrogen atoms, which are critical to identifying the hydrogen bond donors and acceptors and their mechanistic roles.

SSNMR. SSNMR experiments were performed on microcrystalline protein samples prepared in close analogy to the single crystals used for X-ray crystallography. The single crystals were grown slowly using sitting drop vapor diffusion, while the microcrystals were formed rapidly using a batch approach in which the protein and crystallization buffer solutions were mixed to give the same final reagent concentrations as the crystal plate reservoirs. Microcrystals of TS prepared in this manner have been shown to maintain the same crystal habit as the larger crystals (44), giving high confidence that the NMR and X-ray data can be directly compared.

Both E(A-A) and E(A-A)(BZI) can be observed with SSNMR under conditions that favor the accumulation of the enzyme-bound species while the microcrystals remain catalytically active (44–46). SSNMR spectra of E(A-A) formed by the reaction of L-Ser with microcrystalline TS are shown in Fig. 3. 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-Ser substrate, and protein Lys ε-amino groups. With unlabeled substrate and protein (Fig. 3D), the spectra show only unresolved background signals from the protein, including the backbone amide
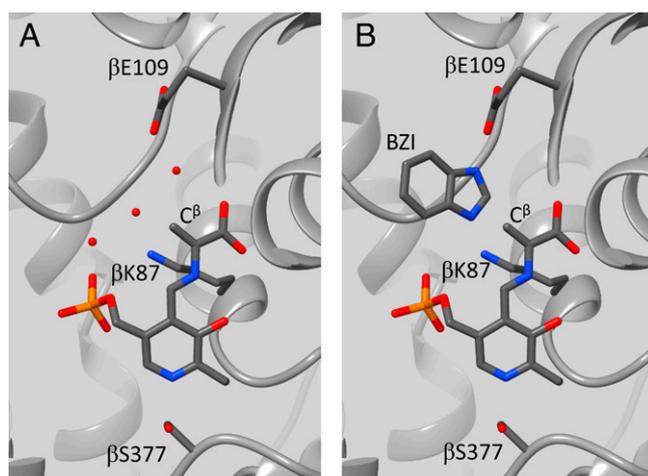

Fig. 2. The β-subunit active sites for E(A-A) and E(A-A)(BZI) from crystal structures 2J9X and 4HPX, respectively. (A) The E(A-A) intermediate shows three active site waters adjacent to the substrate with the central water forming close contacts to both the substrate C^β and the βLys87 ε-amino group. (B) The E(A-A)(BZI) complex shows BZI displacing the three waters but otherwise inducing only small changes in the active site structure. Images rendered in UCSF Chimera (83).

nitrogens (~120 ppm) and backbone carbonyl (~170 ppm), aliphatic (~10 to 70 ppm), and aromatic (~110 to 160 ppm) carbons.

With the magnification factor of the isotopic enrichment (32), distinct resonances for the active site species become visible. The spectra of the sample prepared with ^{15}N and ^{13}C on the PLP cofactor (Fig. 3 B and F) show a nitrogen resonance at 297.6 ppm and three additional carbon resonances at 17.5, 151.2, and 158.1 ppm that are assigned to the cofactor atoms N1, C2′, C2, and C3, respectively (atom labeling given in Fig. 3). The sample prepared using ^{13}C- and ^{15}N-L-Ser as a substrate (Fig. 3 A, E, F, and G) displays a nitrogen resonance at 286.7 ppm and additional carbon resonances at 170.9, 145.6, and 118.8 ppm that are assigned to the Schiff base nitrogen and carbons that derive from the serine C′, C^α, and C^β, respectively. The Schiff base nitrogen displays a significant temperature dependence of ~−0.14 ppm/°C, while the PLP ring nitrogen shows no discernable temperature dependence (*SI Appendix,* Fig. S6). The Schiff base nitrogen chemical shift tensor was measured under slow MAS conditions (*SI Appendix,* Fig. S7), and the principal axis components, $\{\delta_{11}, \delta_{22}, \delta_{33}\}$ = $\{525 \pm 11, 305 \pm 10, 30 \pm 14\}$ ppm, were extracted by fitting the intensity of the spinning sideband manifold (47).

These shifts help establish several key elements of the chemical structure for E(A-A). At this point in the catalytic cycle, the serine substrate has lost its β-hydroxyl, and there is a double bond between C^α and C^β, which is confirmed by their chemical shifts of 145.6 and 118.8 ppm, respectively. Second, with a shift of 286.7 ppm, the Schiff base nitrogen is found to be neutral, which is further supported by the large span of its chemical shift tensor (48). This isotropic chemical shift, however, is below the limiting value for a fully neutral Schiff base, and the temperature dependence suggests tautomeric exchange involving transient protonation of this site (8, 28). At the same time, the PLP C2 and C3 shifts of 151.2 and 158.1 ppm indicate that the PLP phenolic oxygen is neutral (28). The PLP nitrogen shift of 297.6 ppm also indicates a neutral pyridine ring N1.

The ^{15}N side-chain chemical shift of βLys87 was measured on a TS sample in which all lysine residues were ^{15}N-labeled on the ε-amino group as reported previously (45). In the initial internal aldimine state, E(Ain), this sample shows a peak at 202 ppm, corresponding to the protonated Schiff base linkage between βLys87 and the PLP cofactor (49). Upon addition of L-Ser, this resonance is lost, and a new resonance at 24.2 ppm is observed (Fig. 3C), which is assigned to a neutral ε-amino group on βLys87.

The ^{31}P isotropic chemical shift and chemical shift tensor were also measured for the cofactor's phosphate group (*SI Appendix,* Fig. S8). The phosphate chemical shift shows a characteristic response in going from the mono- to the dianionic charge state. The phosphate group's isotropic shift of 5.2 ppm and its shift tensor indicate that it is dianionic, in keeping with other TS intermediates (28, 49) and PLP-dependent enzymes (50).

Finally, the ^{17}O chemical shifts of the substrate carboxylate group were measured using ^{17}O quadrupole central transition (QCT) NMR in solution (52–54) (*SI Appendix,* Fig. S4). 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) (54). Here, an advantage is taken of additional measurements on the 35.2 T series connect hybrid magnet (55) at the National High Magnetic Field Laboratory to improve the accuracy of the extracted parameters (*SI Appendix*). The ^{17}O chemical shifts of 257 and 289 ppm are consistent with an ionized carboxylate group. The up-field shifted signal is
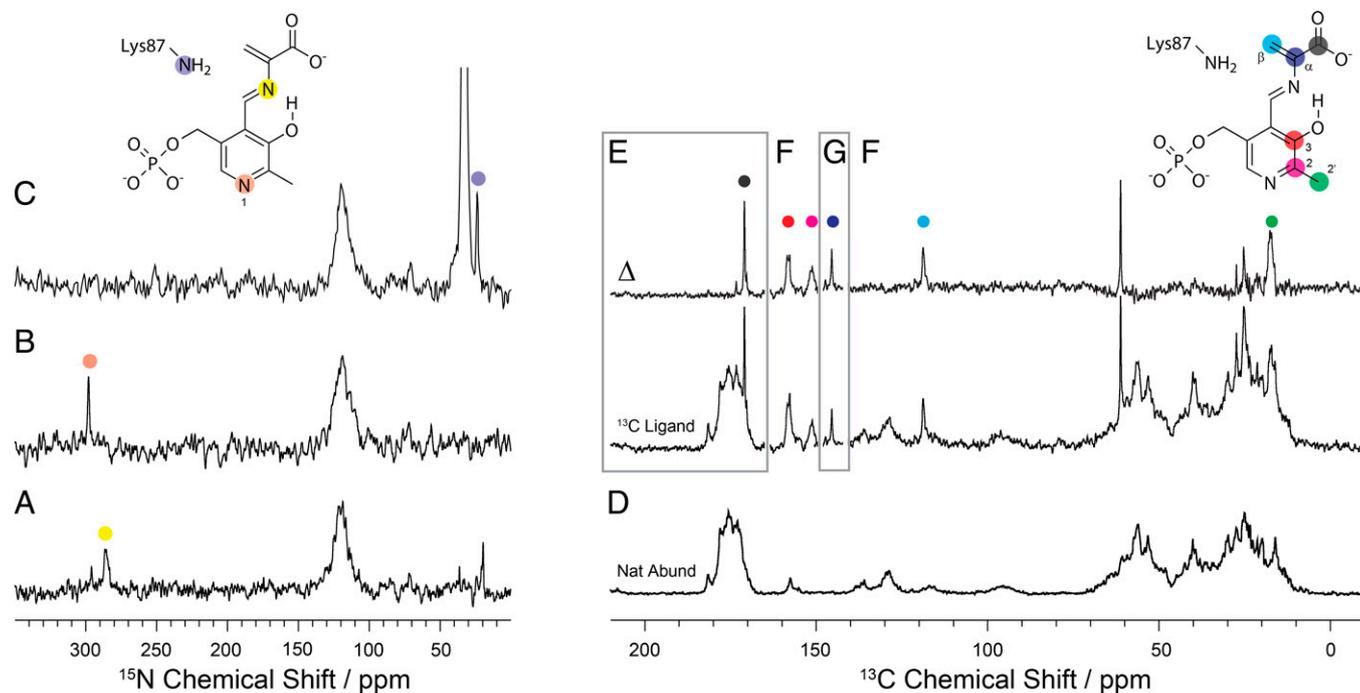
Fig. 3. ^{15}N and ^{13}C CPMAS SSNMR spectra of microcrystalline TS E(A-A) prepared with the following isotopic labeling: (*A*) ^{15}*N*-labeled on the substrate L-Ser; (*B*) ^{15}*N*-enriched on the PLP cofactor; (*C*) ^{15}*N*-enriched at protein lysine side-chain ε-amino groups; (*D*) natural abundance isotopomer concentration; (*E*) ^{13}C-labeled at C′ of the L-Ser substrate; (*F*) ^{13}C,^{15}*N*-enriched on the PLP cofactor and C$^\beta$ of the substrate L-Ser; and (*G*) ^{13}C-labeled at C$^\alpha$ of the substrate L-Ser. The top spectra in (*E–G*) are formed as the difference between the E(A-A) spectra with various cofactor/ligand isotopic labels and the same spectra acquired at natural abundance, emphasizing the resonances for the specific site labels. The large peak at 63.1 ppm is from free serine. Spectra acquired at 9.4 T, −10 °C, and 8 kHz MAS as described in *Materials and Methods*.

tentatively assigned to the oxygen hydrogen bonded to the side chain of βThr110, which forms the closest contact.

Chemical shifts for the E(A-A)(BZI) complex were measured analogously (*SI Appendix*, Figs. S4–S10) and are summarized along with the E(A-A) shifts in Table 1. The spectra for E(A-A)(BZI) show two additional nitrogen resonances at 165.5 and 227.8 ppm that are assigned to N1 and N3 of bound BZI, respectively (atom labeling in Table 1). The spectra are otherwise similar to those for E(A-A) with one notable exception: the loss of the neutral amino resonance at 24.1 ppm and the appearance of an additional charged amino group at 35.6 ppm. The latter was assigned to the active site βLys87 based on ^{15}N{^{31}P} rotational echo double-resonance (56) experiments that place it in close spatial proximity (<4 Å) to the phosphorus

atom of the cofactor (*SI Appendix*, Fig. S10). Thus, despite significant spectral overlap with the 26 other charged Lys side chains in TS, the catalytic active site residue can be uniquely identified. Similar strategies that make use of distinct chemical shifts, nuclides, and/or mixed labeling to assign and localize atoms within specific regions of proteins have been extensively demonstrated in SSNMR (57, 58).

First-Principles Calculations. While the preliminary chemical structures shed some light on the chemistry of the active site, they are incomplete in two important respects. First, they lack detailed three-dimensional structure. Second, the chemical shifts do not fully conform to the expected limiting shifts from model compound studies. To interpret the experimental

Table 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) computational models 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
	C$^\alpha$	144.4	133.1	143.2	145.6			C$^\alpha$	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
	C$^\beta$	123.3	121.3	123.0	118.8			C$^\beta$	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	N$^\zeta$	20.5	21.1	20.6	24.2		βK87	N$^\zeta$	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	

chemical shifts and place the chemistry of the active site in full structural context, we turn to first-principles computational chemistry.

The state-of-the-art for first-principles chemical shift calculations is advanced (25, 59–67), and extensive benchmarking and testing have shown that if the correct structure is known down to the position of every atom, then the NMR chemical shifts can be predicted better than 1.5 ppm RMSD for carbon, 4.3 ppm for nitrogen, and 7.5 ppm for oxygen (66, 67). This expected accuracy enables screening protocols to be established in which proposed structures can be evaluated and ranked for consistency with the experimental chemical shifts. Here, this agreement is quantified using the reduced χ^2 statistic, the squared deviation of the predicted and experimental chemical shifts weighted by the nuclide-dependent mean squared deviations (*SI Appendix*, Eq. **S4**). The benchmark for a solved structure is taken as the identification of a single structure (or a fast-exchange equilibrium between closely related tautomers) that satisfies the 95% confidence limits of the reduced χ^2. To fully distinguish proposed models typically requires 10 or more chemical shifts throughout the active site, supplemented by chemical shift tensor measurements and chemical shift temperature coefficients (28).

To limit bias and broadly consider potential active site chemistries, a pool of candidate structures was generated by systematically varying the protonation states of the cofactor, substrates, and βLys87 ε-amino group (Fig. 4 and *SI Appendix*). 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 7 Å of the substrate and cofactor. The atoms at the exterior of the cluster were fixed at their crystallographic locations, and quantum-mechanical geometry optimization and chemical shift calculations were performed using density functional theory. The clusters contained 703 to 706 atoms (Fig. 4*B*), a size for which the convergence and accuracy of the chemical shift calculations have been established (66).

Thirty-five initial models for E(A-A) and E(A-A)(BZI) were generated (*SI Appendix*, Schemes S1 and S2), their geometries were optimized, and their chemical shifts were predicted (*SI Appendix*, Tables S1 and S2). The structures were then ranked based on their agreement with the experimental chemical shifts as shown in Fig. 5 *A* and *C*. There is a clear differentiation of models for both E(A-A) and E(A-A)(BZI), yet the reduced χ^2 values for all candidate structures fall outside of the 95% confidence limits, indicating that none reproduce the experimental chemical shifts with the expected accuracy. One of the largest discrepancies for all models occurs for the Schiff base nitrogen, in which the closest predicted shifts are still greater than three standard errors from the experimental value.

Motivated by the temperature dependence of the Schiff base nitrogen, a two-site, fast-exchange equilibrium model was considered next in which the effective chemical shift is given as the population-weighted average of the individual shifts. To remain unbiased, all structures that differed by the position of a single hydrogen atom were paired, and their populations were optimized for best overall agreement with the experimental data. These models were again ranked and are summarized in Fig. 5 *B* and *D*. Multiple models now satisfy or come close to satisfying the 95% confidence limits. To differentiate these models, the predicted principal components for the Schiff base nitrogen chemical shift tensor were also compared to the experimental values. The chemical shift tensor is extremely sensitive to dynamics, as it averages in both magnitude and orientation. For both E(A-A) and E(A-A)(BZI), only a single tautomeric exchange model simultaneously satisfies the isotropic and tensor
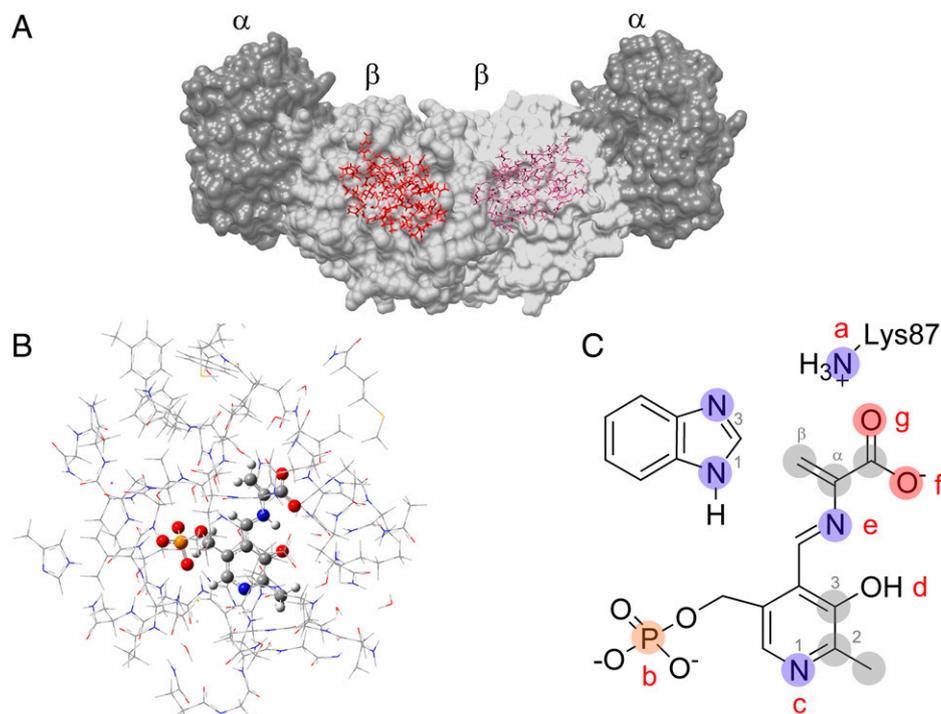
Fig. 4. Cluster model of the E(A-A) active site. (*A*) X-ray crystal structure of the TS $\alpha_2\beta_2$ heterodimer with the β-subunit active site highlighted in red. (*B*) Cluster model of the active site for first-principles geometry optimization and chemical shift calculations with protein side chains displayed in wireframe and cofactor and substrate in ball and stick. (*C*) Protonation sites on or 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.
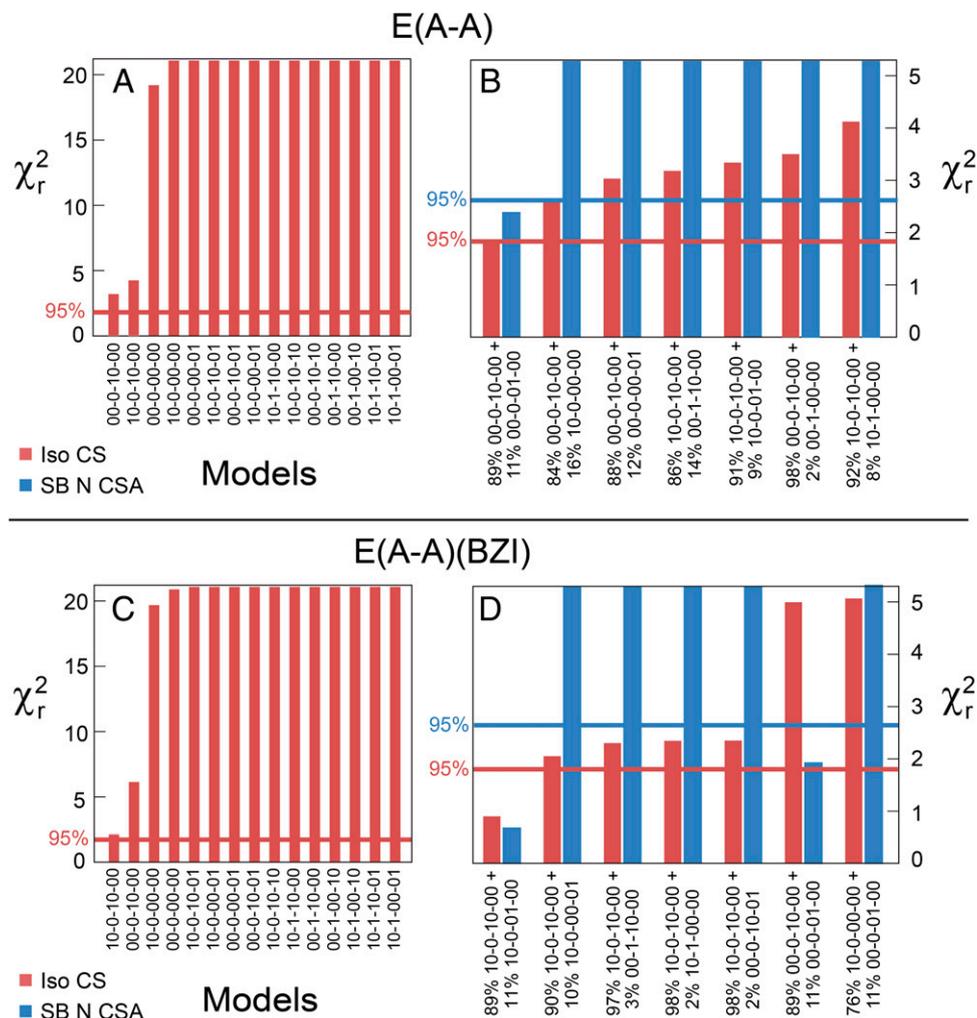
Holmes et al.
Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate

PNAS | 5 of 10
https://doi.org/10.1073/pnas.2109235119

Fig. 5. 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 15 best geometry-optimized active site models for the E(A-A) intermediate; structures and labeling are given in *SI Appendix*, Schemes S1 and S2. (*B*) Rankings of the seven best fast-exchange equilibrium models comparing the experimental and first-principles isotropic chemical shifts (red) and Schiff base nitrogen tensor components (blue). 95% confidence limits are shown as the correspondingly colored horizontal bars. (*C* and *D*) Model rankings for the E(A-A)(BZI) complex. For both E(A-A) and E(A-A)(BZI), only a single tautomeric exchange model simultaneously satisfies the isotropic and tensor chemical shift restraints.

chemical shift restraints, indicating that the benchmark for a solved structure has been reached.

For both E(A-A) and E(A-A)(BZI), the best-fit equilibrium is between the phenolic (enolimine; 89%) and protonated Schiff base (ketoenamine; 11%) tautomers, with proton exchange across the internal hydrogen bond (Table 1, center inset). The chemical shifts for the best-fit models are summarized in Table 1 along with the experimental shifts. An analogous exchange was also found for the TS carbanionic intermediate (28) and appears to be a common feature in many PLP-dependent enzymes (8). The only difference in protonation states between the E(A-A) and E(A-A)(BZI) structures is for the βLys87 ε-amino group: for E(A-A) it is neutral, while for E(A-A)(BZI) it is positively charged.

To further test the exchange model, the ^{15}N chemical shift of the Schiff base nitrogen was measured at 95 K under conditions of dynamic nuclear polarization (*SI Appendix*, Fig. S6). At this temperature, the tautomeric equilibrium is expected to shift predominantly to the phenolic form with less than 1% of the exchange partner present. This was confirmed experimentally, with chemical shifts of 301.4 ppm for E(A-A) and 302.1 ppm for E(A-A)(BZI) at 95 K compared to the predicted shifts for the individual structures of 302.9 and 310.2 ppm, respectively.

Thus, both the major tautomers and the dynamic equilibrium of these intermediates can be established.

As a final assessment, the overall confidence in the identification of the experimental structure was quantified using the Bayesian probability analysis following the approach of Engel et al. (68) as described in *SI Appendix*. Taking into account both the isotropic and anisotropic chemical shifts along with the observed temperature dependence of the Schiff base nitrogen, the proposed exchange model is found to be the most probable experimental structure with 88.5% confidence for E(A-A) and 99.6% confidence for E(A-A)(BZI). The decreased confidence for E(A-A) can be traced back to the lower overall agreement between the predicted and experimental shifts for the best-fit model as reflected in the reduced χ^2. Dynamics of the water molecules adjacent to the substrate may explain why the overall agreement between theory and experiment is lower in this case.

NMR-Assisted Crystallography. What emerges from the application of NMR-assisted crystallography to the α-aminoacrylate intermediate is a chemically detailed view of the enzyme active site in which the locations of all atoms, including hydrogens, have been specified (Fig. 6). The quality of these structures can be judged by several metrics. First, the refined heavy atom
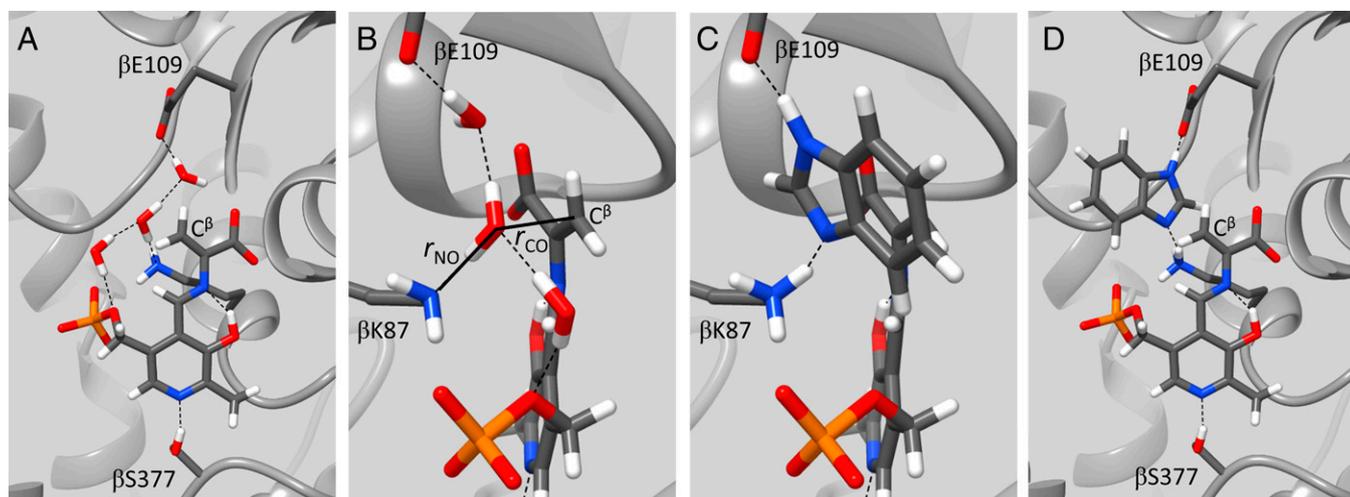
Holmes et al.
Imaging active site chemistry and protonation states: NMR crystallography
of the tryptophan synthase α-aminoacrylate intermediate

Fig. 6. Protonation states, hydrogen bonding interactions, and the placement and orientation of structural waters as revealed by NMR-assisted crystallography in the TS β-subunit active site. (*A* and *B*) The E(A-A) intermediate and (*C* and *D*) the E(A-A)(BZI) complex. For E(A-A), the position and orientation of the central water molecule highlights the reaction coordinate for the loss of the serine β-hydroxyl and confirms βLys87 as the active site acid–base catalytic residue. The E(A-A)(BZI) complex shows BZI bound in the active site with hydrogen bonding interactions to βGlu109 and the charged ε-amino group of βLys87. BZI is poised for, but unable to initiate, the subsequent bond formation step. Images rendered in UCSF Chimera (83).

positions are found to agree well with the original X-ray crystal coordinates, with an RMSD of 0.077 Å for E(A-A) and 0.091 Å for E(A-A)(BZI). Second, the positional uncertainties in the structures from NMR-assisted crystallography can be estimated following the method of Hofstetter and Emsley (*SI Appendix*) (69). This approach correlates changes in the predicted shifts with displacements in the structural coordinates, defining a positional uncertainty that is consistent with the experimental shifts. Average positional RMSD of 0.11 Å for heavy atoms and 0.17 Å for hydrogen atoms are found for the cofactor and substrates in both E(A-A) and E(A-A)(BZI). These positional uncertainties are similar in size to those found in organic molecular crystals (38, 69, 70) and are 6.5 times smaller than the heavy atom uncertainties reported in the corresponding protein X-ray crystal structures. While preliminary, these results suggest that positional uncertainties in NMR crystallography are independent of molecular size and that they will depend primarily on the nature and number of the measured NMR observables.

Chemically Rich Crystal Structures and Mechanistic Implications. Perhaps most striking in the structure of the E(A-A) active site is the position of the central water molecule immediately adjacent to the serine substrate C^β and oriented with one hydrogen pointing toward the ε-amino group of βLys87 (Fig. 6*B*). This water is located 2.9 Å from C^β and is perfectly aligned for the reverse nucleophilic attack to regenerate E(C₁). The interaction of this water with βLys87 and its proximity to the C^β carbon support the hypothesis that βLys87 is the acid catalyst that protonates the β-hydroxyl during C–O bond scission as E(C₁) is converted to E(A-A) and a water molecule. We postulate this site is the initial binding pocket for the eliminated water molecule.

Critical examination of the crystal structures for E(A-A), however, shows that this central water is not consistently present. The electron density maps for 2J9X and 4HN4 (both with Cs⁺ in the monovalent cation-binding site) show significantly lower electron density for this water and likely reflect occupancy less than 50% (*SI Appendix*, Fig. S11). To explore this variation, two additional E(A-A) crystal structures for the NH₄⁺ and Cs⁺ forms were solved at 1.40 and 1.50 Å, respectively (Protein Data Bank identification codes 7MT4

and 7MT5) (71, 72). The diffraction data (*SI Appendix*, Table S9) show well-defined electron density for the central water in the NH₄⁺ form but an absence of this density in the Cs⁺ structure (*SI Appendix*, Fig. S11). Close inspection of the NMR data for E(A-A) finds that each of the ¹³C resonances of the bound serine is accompanied by slightly shifted (±1 to 2 ppm) minor species of ∼20 to 30% intensity, while those for E(A-A)(BZI) show no such peaks. These resonances are most easily observed with single site ¹³C serine labels (Fig. 7*A*) and display the same multiplet pattern as the major resonances upon the incorporation of 2,3-¹³C₂-L-Ser. The minor and major resonances also show distinct sets of cross-peaks (minor to minor, major to major) in the two-dimensional (2D) dipolar-driven correlation experiment acquired with a CPMAS cryoprobe to improve sensitivity (Fig. 7*C*) (73, 74). We hypothesize that the major and minor resonances belong to two independent E(A-A) species that derive from varying occupancy of the central water position across the macroscopic crystals. To test this, geometry-optimized active site models for the primary phenolic and protonated Schiff base tautomers were constructed in which the water nearest C^β was removed. Their predicted chemical shifts track the experimental chemical shifts of the minor species, and these secondary peaks are therefore tentatively assigned to a minor population in which the central water is no longer bound. Taken together, the SSNMR, X-ray, and computational results are consistent with this site being a weak binding site for water.

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. BZI is an isostere of indole and a potent inhibitor of the β-reaction (39, 75–77). The view from NMR-assisted crystallography (Fig. 6 *C* and *D*) shows BZI bound in the active site, displacing the three waters, and poised for nucleophilic attack. Despite being more nucleophilic than indole, BZI does not react to form a covalent bond. A comparison to crystal structures for the carbanionic intermediate analogs formed with indoline and 2-aminophenol (28) makes it clear that BZI is bound to the same subsite. Thus, this intermediate analog appears to be poised close to the transition state between two reactive intermediates in the TS catalytic cycle.
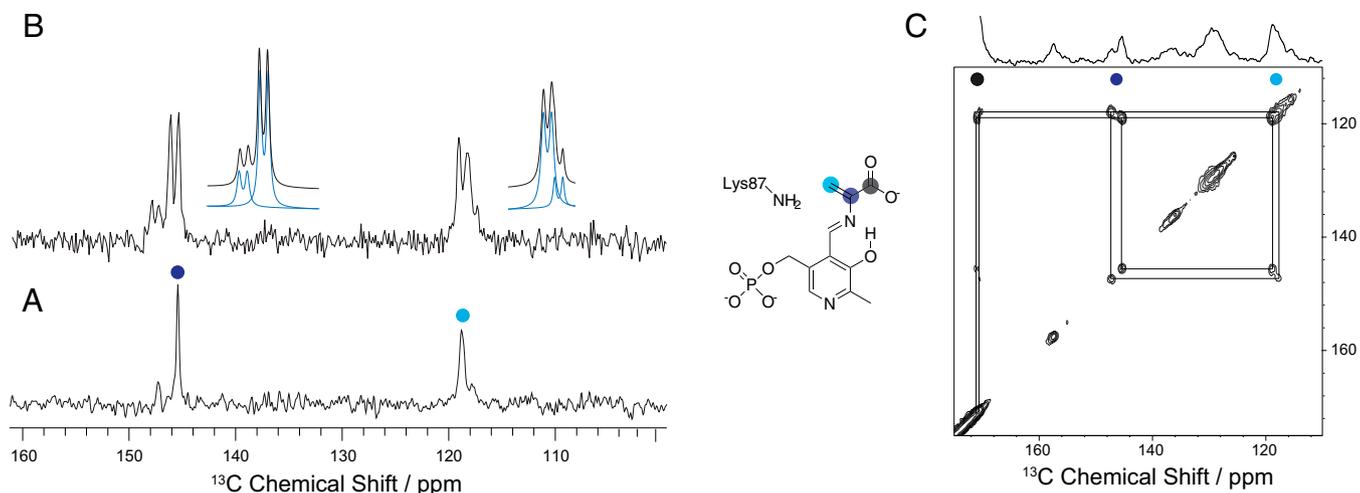
Fig. 7. ^{13}C SSNMR CPMAS spectra of microcrystalline TS E(A-A) prepared with (*A*) a 50:50 mixture of singly labeled 2-^{13}C-L-Ser and 3-^{13}C-L-Ser and (*B*) doubly labeled 2,3-^{13}C$_2$-L-Ser. For both, the natural abundance ^{13}C protein background has been subtracted as in Fig. 3. The C$^\alpha$ and C$^\beta$ resonances show a minor peak that, along with the major peak, splits into a doublet upon incorporation of doubly labeled L-Ser substrate (deconvolutions shown adjacent to the spectrum). (*C*) The 2D dipolar driven ^{13}C correlation spectrum of E(A-A) formed with U-^{13}C$_3$-L-Ser displays distinct cross-peaks for the major and minor resonances, indicating that they belong to two independent E(A-A) species. One-dimensional spectra acquired as in Fig. 3; 2D Spectra acquired using combined R2-driven (CORD) mixing (73) on a Bruker BioSolids CryoProbe (74) at 14.1 T, −10 °C (sample temp) and 8 kHz MAS as described in *Methods and Materials*.

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 (75). NMR-assisted crystallography identifies that it is the nucleophilic N3 that is adjacent to the substrate C$^\beta$ in the E(A-A)(BZI) complex. However, this structure shows that the BZI lone pair does not point toward the E(A-A) C$^\beta$ 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 is modeled in place of BZI, the C3 carbon of the five-membered ring is perfectly aligned to form the new C–C bond. The attacking p-orbital on indole points toward the electron-deficient C$^\beta$ 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$^\beta$. The critical role played by βGlu109 can be seen in the βGlu109Asp mutation, which reduces the β reaction rate 27-fold (78). 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.

Conclusions

The integrated application of SSNMR spectroscopy, X-ray crystallography, and first-principles computational chemistry offers unprecedented, chemically rich views of the TS E(A-A) and E(A-A)(BZI) active sites. Through the combined determination of heavy atom positions, the identification of protonation states, and the placement and orientation of active site waters, an exceptionally detailed picture of structure and reactivity

emerges. Through this lens, the E(A-A) intermediate shows a central water that is well positioned for the reverse nucleophilic attack on C$^\beta$, with an orientation that points back to the active site acid–base catalytic residue βLys87. This chemically detailed view highlights the reaction coordinate for the elimination of the substrate β-hydroxyl and the initial binding pocket for the eliminated water. Both X-ray crystallography and SSNMR indicate variable occupancy for this site. Upon reaction of E(A-A) with BZI, the three waters adjacent to the substrate are displaced, and NMR-assisted crystallography shows BZI occupying the indole binding pocket but unable to react. Here, the 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.

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.

Materials and Methods

First-principles calculations were performed using a cluster-based model of the active site as described previously (28). TS was prepared by overexpression of *St*TS in *Escherichia coli* (45, 49). Protein crystals of E(A-A) and E(A-A)(BZI) were prepared and structures solved following our earlier protocols (39, 79). SSNMR experiments followed our prior experimental design (28, 54), including DNP experiments at 95 K (78), and 2D correlation experiments performed on an MAS cryoprobe (74). ^{15}N-BZI was synthesized from ^{15}NH$_4$OH and 1-fluoro-2-nitrobenzene; 2,2′,3-^{13}C$_3$,^{15}N-PLP was prepared from U-^{13}C$_3$,^{15}N-Ala as previously detailed (49), while ^{15}N-PLP was prepared using a new synthetic strategy including enzymatic phosphorylation via the ePL Kinase K229Q mutant (81). Detailed materials and methods are included as part of *SI Appendix*.

1. C. Walsh, *Enzymatic Reaction Mechanisms* (W. H. Freeman and Company, San Francisco, 1979), pp. 978.
2. H. Hayashi, Pyridoxal enzymes: Mechanistic diversity and uniformity. *J. Biochem.* **118**, 463–473 (1995).
3. R. Percudani, A. Peracchi, A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep.* **4**, 850–854 (2003).
4. M. D. Toney, Controlling reaction specificity in pyridoxal phosphate enzymes. BBA - Proteins *Proteomics* **1814**, 1407–1418 (2011).
5. E. W. Miles, Tryptophan synthase: Structure, function, and subunit interaction. *Adv. Enzymol. Relat. Areas Mol. Biol.* **49**, 127–186 (1979).
6. C. Yanofsky, I. P. Crawford, "Tryptophan synthetase" in *The Enzymes*, P. D. Boyer, Ed. (Academic Press, Inc., New York, ed. 3, 1972), **vol. 7**, chap. 1, pp. 1–32.
7. M. F. Dunn, D. Niks, H. Ngo, T. R. Barends, I. Schlichting, Tryptophan synthase: The workings of a channeling nanomachine. *Trends Biochem. Sci.* **33**, 254–264 (2008).
8. H. C. Dunathan, Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **55**, 712–716 (1966).
9. M. Chan-Huot et al., 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. *J. Am. Chem. Soc.* **135**, 18160–18175 (2013).
10. P. S. Langan et al., Substrate binding induces conformational changes in a class A β-lactamase that prime it for catalysis. *ACS Catal.* **8**, 2428–2437 (2018).
11. S. Dajnowicz et al., Direct visualization of critical hydrogen atoms in a pyridoxal 5′-phosphate enzyme. *Nat. Commun.* **8**, 955 (2017).
12. K. M. Yip, N. Fischer, E. Paknia, A. Chari, H. Stark, Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).
13. A. T. Petkova et al., A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16742–16747 (2002).
14. C. Luchinat, G. Parigi, E. Ravera, M. Rinaldelli, Solid-state NMR crystallography through paramagnetic restraints. *J. Am. Chem. Soc.* **134**, 5006–5009 (2012).
15. J.-P. Demers et al., High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nat. Commun.* **5**, 4976 (2014).
16. L. Sborgi et al., Structure and assembly of the mouse ASC inflammasome by combined NMR spectroscopy and cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13237–13242 (2015).
17. R. Michalczyk et al., Joint neutron crystallographic and NMR solution studies of Tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5673–5678 (2015).
18. A. Carlon et al., Improved accuracy from joint X-ray and NMR refinement of a protein-RNA complex structure. *J. Am. Chem. Soc.* **138**, 1601–1610 (2016).
19. L. Gremer et al., Fibril structure of amyloid-β(1-42) by cryo-electron microscopy. *Science* **358**, 116–119 (2017).
20. J. R. Perilla et al., CryoEM structure refinement by integrating NMR chemical shifts with molecular dynamics simulations. *J. Phys. Chem. B* **121**, 3853–3863 (2017).
21. K. Zhang et al., Structure of the 30 kDa HIV-1 RNA dimerization signal by a hybrid Cryo-EM, NMR, and molecular dynamics approach. *Structure* **26**, 490–498.e3 (2018).
22. L. Cerofolini et al., Characterization of PEGylated asparaginase: New opportunities from NMR analysis of large PEGylated therapeutics. *Chemistry* **25**, 1984–1991 (2019).
23. D. F. Gauto et al., Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nat. Commun.* **10**, 2697 (2019).
24. J. Felix et al., Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors. *Sci. Adv.* **5**, eaaw3818 (2019).
25. S. Kosol et al., Structural basis for chain release from the enacyloxin polyketide synthase. *Nat. Chem.* **11**, 913–923 (2019).
26. J. A. Gascón, E. M. Sproviero, V. S. Batista, QM/MM study of the NMR spectroscopy of the retinyl chromophore in visual rhodopsin. *J. Chem. Theory Comput.* **1**, 674–685 (2005).
27. J. Lai et al., X-ray and NMR crystallography in an enzyme active site: The indoline quinonoid intermediate in tryptophan synthase. *J. Am. Chem. Soc.* **133**, 4–7 (2011).
28. L. J. Mueller, M. F. Dunn, NMR crystallography of enzyme active sites: Probing chemically detailed, three-dimensional structure in tryptophan synthase. *Acc. Chem. Res.* **46**, 2008–2017 (2013).
29. B. G. Caulkins et al., NMR crystallography of a carbanionic intermediate in tryptophan synthase: Chemical structure, tautomerization, and reaction specificity. *J. Am. Chem. Soc.* **138**, 15214–15226 (2016).
30. H. Zhang et al., HIV-1 capsid function is regulated by dynamics: Quantitative atomic-resolution insights by integrating magic-angle-spinning NMR, QM/MM, and MD. *J. Am. Chem. Soc.* **138**, 14066–14075 (2016).
31. H. Singh et al., Fast microsecond dynamics of the protein-water network in the active site of human carbonic anhydrase II studied by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **141**, 19276–19288 (2019).
32. A. Bertarello et al., Picometer resolution structure of the coordination sphere in the metal-binding site in a metalloprotein by NMR. *J. Am. Chem. Soc.* **142**, 16757–16765 (2020).
33. J. C. Facelli, D. M. Grant, Determination of molecular symmetry in crystalline naphthalene using solid-state NMR. *Nature* **365**, 325–327 (1993).
34. M. Baias et al., De novo determination of the crystal structure of a large drug molecule by crystal structure prediction-based powder NMR crystallography. *J. Am. Chem. Soc.* **135**, 17501–17507 (2013).
35. D. H. Brouwer et al., A general protocol for determining the structures of molecularly ordered but noncrystalline silicate frameworks. *J. Am. Chem. Soc.* **135**, 5641–5655 (2013).
36. C. Martineau, NMR crystallography: Applications to inorganic materials. *Solid State Nucl. Magn. Reson.* **63-64**, 1–12 (2014).
37. D. L. Bryce, NMR crystallography: Structure and properties of materials from solid-state nuclear magnetic resonance observables. *IUCrJ* **4**, 350–359 (2017).
38. P. Hodgkinson, NMR crystallography of molecular organics. *Prog. Nucl. Magn. Reson. Spectrosc.* **118-119**, 10–53 (2020).
39. M. Cordova et al., Structure determination of an amorphous drug through large-scale NMR predictions. *Nat. Commun.* **12**, 2964 (2021).
40. D. Niks et al., Allostery and substrate channeling in the tryptophan synthase bienzyme complex: Evidence for two subunit conformations and four quaternary states. *Biochemistry* **52**, 6396–6411 (2013).
41. H. Ngo et al., Allosteric regulation of substrate channeling in tryptophan synthase: Modulation of the L-serine reaction in stage I of the beta-reaction by alpha-site ligands. *Biochemistry* **46**, 7740–7753 (2007).
42. M. F. Dunn, Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Arch. Biochem. Biophys.* **519**, 154–166 (2012).
43. M. A. Maria-Solano, J. Iglesias-Fernández, S. Osuna, Deciphering the allosterically driven conformational ensemble in tryptophan synthase evolution. *J. Am. Chem. Soc.* **141**, 13049–13056 (2019).
44. A. R. Buller et al., Directed evolution mimics allosteric activation by stepwise tuning of the conformational ensemble. *J. Am. Chem. Soc.* **140**, 7256–7266 (2018).
45. S. A. Ahmed, C. C. Hyde, G. Thomas, E. W. Miles, Microcrystals of tryptophan synthase alpha.2.beta.2 complex from *Salmonella typhimurium* are catalytically active. *Biochemistry* **26**, 5492–5498 (1987).
46. B. G. Caulkins et al., Catalytic roles of βLys87 in tryptophan synthase: 15N solid state NMR studies. *BBA - Proteins Proteomics* **1854**, 1194–1199 (2015).
47. L. M. McDowell, M. S. Lee, J. Schaefer, K. S. Anderson, Observation of an aminoacrylate enzyme intermediate in the tryptophan synthase reaction by solid-state NMR. *J. Am. Chem. Soc.* **117**, 12352–12353 (1995).
48. J. Herzfeld, A. E. Berger, Sideband intensities in NMR spectra of samples spinning at the magic angle. *J. Chem. Phys.* **73**, 6021–6030 (1980).
49. V. Copié, W. S. Faraci, C. T. Walsh, R. G. Griffin, Inhibition of alanine racemase by alanine phosphonate: Detection of an imine linkage to pyridoxal 5′-phosphate in the enzyme-inhibitor complex by solid-state 15N nuclear magnetic resonance. *Biochemistry* **27**, 4966–4970 (1988).
50. B. G. Caulkins et al., 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**, 12824–12827 (2014).
51. K. D. Schnackerz, B. Andi, P. F. Cook, 31P NMR spectroscopy senses the microenvironment of the 5′-phosphate group of enzyme-bound pyridoxal 5′-phosphate. *BBA - Proteins Proteomics* **1814**, 1447–1458 (2011).
52. H. C. Lee, E. Oldfield, Oxygen-17 nuclear magnetic resonance-spectroscopic studies of carbonmonoxy hemoproteins. *J. Am. Chem. Soc.* **111**, 1584–1590 (1989).
53. J. Zhu, G. Wu, Quadrupole central transition 17O NMR spectroscopy of biological macromolecules in aqueous solution. *J. Am. Chem. Soc.* **133**, 920–932 (2011).
54. R. P. Young et al., Solution-state (17O) quadrupole central-transition NMR spectroscopy in the active site of tryptophan synthase. *Angew. Chem. Int. Ed. Engl.* **55**, 1350–1354 (2016).
55. Z. Gan et al., NMR spectroscopy up to 35.2T using a series-connected hybrid magnet. *J. Magn. Reson.* **284**, 125–136 (2017).
56. T. Gullion, J. Schaefer, NMR rotational-echo double-resonance. *J. Magn. Reson.* **81**, 196–200 (1989).

Holmes et al.
Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate

57. A. K. Schütz, Solid-state NMR approaches to investigate large enzymes in complex with substrates and inhibitors. *Biochem. Soc. Trans.* **49**, 131–144 (2021).

58. J. Kraus, S. Sarkar, C. M. Quinn, T. Polenova, "Solid-state NMR spectroscopy of microcrystalline proteins" in *Annual Reports on NMR Spectroscopy*, G. A. Webb, Ed. (Academic Press, **vol. 102**, 2021), pp. 81–151.

59. A. C. de Dios, J. G. Pearson, E. Oldfield, Secondary and tertiary structural effects on protein NMR chemical shifts: An ab initio approach. *Science* **260**, 1491–1496 (1993).

60. Q. Cui, M. Karplus, Molecular properties from combined QM/MM methods. 2. Chemical shifts in large molecules. *J. Phys. Chem. B* **104**, 3721–3743 (2000).

61. L. B. Casabianca, A. C. de Dios, Ab initio calculations of NMR chemical shifts. *J. Chem. Phys.* **128**, 052201 (2008).

62. J. C. Facelli, Chemical shift tensors: Theory and application to molecular structural problems. *Prog. Nucl. Magn. Reson. Spectrosc.* **58**, 176–201 (2011).

63. T. E. Exner, A. Frank, I. Onila, H. M. Möller, Toward the quantum chemical calculation of NMR chemical shifts of proteins. 3. Conformational sampling and explicit solvents model. *J. Chem. Theory Comput.* **8**, 4818–4827 (2012).

64. J. Swails, T. Zhu, X. He, D. A. Case, AFNMR: Automated fragmentation quantum mechanical calculation of NMR chemical shifts for biomolecules. *J. Biomol. NMR* **63**, 125–139 (2015).

65. S. K. Chandy, B. Thapa, K. Raghavachari, Accurate and cost-effective NMR chemical shift predictions for proteins using a molecules-in-molecules fragmentation-based method. *Phys. Chem. Chem. Phys.* **22**, 27781–27799 (2020).

66. J. D. Hartman, T. J. Neubauer, B. G. Caulkins, L. J. Mueller, G. J. Beran, Converging nuclear magnetic shielding calculations with respect to basis and system size in protein systems. *J. Biomol. NMR* **62**, 327–340 (2015).

67. J. D. Hartman, R. A. Kudla, G. M. Day, L. J. Mueller, G. J. Beran, Benchmark fragment-based ^{1}H, ^{13}C, ^{15}N and ^{17}O chemical shift predictions in molecular crystals. *Phys. Chem. Chem. Phys.* **18**, 21686–21709 (2016).

68. E. A. Engel et al., A Bayesian approach to NMR crystal structure determination. *Phys. Chem. Chem. Phys.* **21**, 23385–23400 (2019).

69. A. Hofstetter, L. Emsley, Positional variance in NMR crystallography. *J. Am. Chem. Soc.* **139**, 2573–2576 (2017).

70. L. Wang, F. J. Uribe-Romo, L. J. Mueller, J. K. Harper, Predicting anisotropic thermal displacements for hydrogens from solid-state NMR: A study on hydrogen bonding in polymorphs of palmitic acid. *Phys. Chem. Chem. Phys.* **20**, 8475–8487 (2018).

71. V. Drago, E. Hilario, M. F. Dunn, T. C. Mueser, L. J. Mueller, Crystal structure of tryptophan synthase in complex with F9, NH4+, pH7.8 - alpha aminoacrylate form - E(A-A). Protein Data Bank. https://www.rcsb.org/structure/7MT4. Deposited 12 May 2021.

72. V. Drago, E. Hilario, M. F. Dunn, T. C. Mueser, L. J. Mueller, Crystal structure of tryptophan synthase in complex with F9, Cs+, pH7.8 - alpha aminoacrylate form - E(A-A). Protein Data Bank. https://www.rcsb.org/structure/7MT5. Deposited 12 May 2021.

73. G. Hou, S. Yan, J. Trébosc, J. P. Amoureux, T. Polenova, Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J. Magn. Reson.* **232**, 18–30 (2013).

74. A. Hassan et al., Sensitivity boosts by the CPMAS CryoProbe for challenging biological assemblies. *J. Magn. Reson.* **311**, 106680 (2020).

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

76. M. F. Dunn et al., The tryptophan synthase bienzyme complex transfers indole between the alpha- and beta-sites via a 25-30 A long tunnel. *Biochemistry* **29**, 8598–8607 (1990).

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

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

79. E. Hilario et al., Visualizing the tunnel in tryptophan synthase with crystallography: Insights into a selective filter for accommodating indole and rejecting water. *BBA - Proteins Proteomics* **1864**, 268–279 (2016).

80. X. Wang et al., 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 Nucl. Magn. Reson.* **100**, 85–91 (2019).

81. M. S. Ghatge et al., Pyridoxal 5′-phosphate is a slow tight binding inhibitor of *E. coli* pyridoxal kinase. *PLoS One* **7**, e41680 (2012).

82. E. W. Miles, Stereochemistry and mechanism of a new single-turnover, half-transamination reaction catalyzed by the tryptophan synthase α2β2 complex. *Biochemistry* **26**, 597–603 (1987).

83. E. F. Pettersen et al., UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

84. V. Drago, E. Hilario, M. F. Dunn, T. C. Mueser, L. J. Mueller, Crystal structure of tryptophan synthase in complex with F9, Cs+, benzimidazole, pH7.8 - alpha aminoacrylate form - E(A-A)(BZI). Protein Data Bank. https://www.rcsb.org/structure/7MT6. Deposited 12 May 2021.

10 of 10 | PNAS
https://doi.org/10.1073/pnas.2109235119

Holmes et al.
Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate