

**Revealing protonation states and tracking substrate in serine hydroxymethyltransferase with room-temperature X-ray and neutron crystallography**

Victoria N. Drago,<sup>1</sup> Claudia Campos,<sup>2</sup> Mattea Hooper,<sup>2</sup> Aliyah Collins,<sup>2</sup> Oksana Gerlits,<sup>2</sup> Kevin L. Weiss,<sup>1</sup> Matthew P. Blakeley,<sup>3</sup> Robert S. Phillips,<sup>4,5</sup> and Andrey Kovalevsky<sup>1\*</sup>

*<sup>1</sup>Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA*

*<sup>2</sup>Department of Natural Sciences, Tennessee Wesleyan University, Athens, TN 37303, USA*

*<sup>3</sup>Large Scale Structures Group, Institut Laue–Langevin, 71 Avenue des Martyrs, 38000 Grenoble, France*

*<sup>4</sup>Department of Chemistry, University of Georgia, Athens, GA, 30602, USA*

*<sup>5</sup>Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, 30602, USA*

\* To whom correspondence should be addressed: Andrey Kovalevsky: [kovalevskyay@ornl.gov](mailto:kovalevskyay@ornl.gov)

## Abstract

Pyridoxal 5'-phosphate (PLP)-dependent enzymes utilize a vitamin B<sub>6</sub>-derived cofactor to perform a myriad of chemical transformations on amino acids and other small molecules. Some PLP-dependent enzymes, such as serine hydroxymethyltransferase (SHMT), are promising drug targets for the design of small-molecule antimicrobials and anticancer therapeutics, while others have been used to synthesize pharmaceutical building blocks. Understanding PLP-dependent catalysis and the reaction specificity is crucial to advance structure-assisted drug design and enzyme engineering. Here we report direct determination of the protonation states in the active site of *Thermus thermophilus* SHMT (*Tth*SHMT) in the internal aldimine state using room-temperature joint X-ray/neutron crystallography. Conserved active site architecture of the model enzyme *Tth*SHMT and of human mitochondrial SHMT (hSHMT2) were compared by obtaining a room-temperature X-ray structure of hSHMT2, suggesting identical protonation states in the human enzyme. Amino acid substrate serine pathway through the *Tth*SHMT active site cavity was tracked, revealing the peripheral and cationic binding sites that correspond to the *pre*-Michaelis and *pseudo*-Michaelis complexes, respectively. At the peripheral binding site, the substrate is bound in the zwitterionic form. By analyzing the observed protonation states, Glu53, but not His residues, is proposed as the general base catalyst, orchestrating the retro-aldol transformation of L-serine into glycine.

**Keywords:** serine hydroxymethyltransferase, enzyme mechanism, protonation state, L-serine, D-serine, neutron crystallography, room-temperature X-ray crystallography.

## INTRODUCTION

Found in all living organisms, pyridoxal 5'-phosphate (PLP)-dependent enzymes utilize a phosphorylated, biologically active form of vitamin B<sub>6</sub> (Figure 1a) to catalyze a myriad of chemical reactions including transamination,  $\beta$ - and  $\gamma$ -elimination,  $\alpha$ -decarboxylation, retro-aldol cleavages, and glycogen phosphorylation.<sup>1,2</sup> PLP-dependent enzymes have been organized into seven fold types based on their evolutionary lineages, however, within these groupings there are multiple catalytic activities.<sup>3-5</sup> The largest group of PLP-dependent enzymes is the aminotransferase superfamily (Fold I). Hence, PLP-dependent enzymes, owing to their versatility, are physiologically significant for the metabolism, interconversion, and synthesis of various amino acids. Unsurprisingly, many PLP-dependent enzymes are recognized as drug targets for the development of small-molecule therapeutics and have also been exploited for the synthesis of building blocks for pharmaceutical applications.<sup>6,7</sup>

PLP-dependent catalysis is enabled by the stereoelectronic control of the labile covalent linkage, called the Schiff base, formed between the PLP cofactor and the side-chain amino group of a catalytic lysine (Lys) residue to create the internal aldimine functional state (Figure 1a), and by the ability of the enzyme active sites to stabilize charged intermediates through the electron withdrawing properties of the pyridine ring.<sup>8-10</sup> These principles, however, do not fully explain the catalytic diversity of the PLP-dependent enzymes, but rather it appears that the manifold of PLP-dependent activities is governed by the surrounding protein environment and the electronic modulation of the cofactor through selective protonation. Consequently, PLP-dependent catalysis cannot be completely understood without knowing the protonation states of the cofactor and the surrounding residues that are dictated by the locations of hydrogen (H) atoms, which determine the electrostatic environment within the active site.

Serine hydroxymethyltransferase (SHMT), a ubiquitous PLP-dependent enzyme in the aminotransferase superfamily, catalyzes the reversible conversion of L-serine (L-Ser) to glycine (Gly), transferring a one-carbon unit to tetrahydrofolate (THF) to yield 5,10-methylenetetrahydrofolate (5,10-MTHF) (Figure 1b).<sup>5,11</sup> SHMT exhibits some catalytic promiscuity, similar to many PLP-dependent enzymes, and can catalyze THF-independent reactions such as the reversible cleavage of  $\beta$ -hydroxy amino acids<sup>12</sup>, decarboxylation of

aminomalonates<sup>13</sup>, and the racemization and transamination of D- and L-alanine.<sup>14</sup> As an enzyme in one-carbon metabolism, SHMT is essential to the synthesis of thymine nucleotides, purines, methionine and other essential biomolecules.<sup>15,16</sup> SHMT is found as a homodimer in prokaryotes, such as *Thermus thermophilus* (*Th*), and a homotetramer in eukaryotes, such as humans (Figure 1c and 1d).<sup>17,18</sup> One-carbon metabolism is compartmentalized in mammals, and hence requires two SHMT isoforms, SHMT1 and SHMT2, located in the cytosol and mitochondria, respectively.<sup>18,19</sup> SHMT1 and SHMT2 catalyze equivalent biochemical reactions and have a sequence identity of ~66%.<sup>20</sup> Interestingly, in many cancer cell lines these reactions proceed in opposite directions in SHMT1 and SHMT2, with SHMT2 providing the majority of one-carbon units in mitochondria from serine synthesized by SHMT1 in the cytosol.<sup>21</sup>

Human SHMT2 (hSHMT2, Figure 1c) has been found to be highly overexpressed in many types of cancers as a result of metabolic reprogramming of the one-carbon metabolism.<sup>22-27</sup> Additionally, high levels of hSHMT2 expression have been linked to poor cancer prognoses and to the resistance to 5-fluorouracil, one of the most commonly used drugs to treat cancer.<sup>28-31</sup> The significance of hSHMT2 in the proliferation and drug resistance of cancer cells has made this enzyme an attractive drug target for antimetabolite chemotherapies,<sup>16,32,33</sup> and several promising hSHMT2 inhibitors have been discovered.<sup>34-37</sup> To advance drug design<sup>38</sup> of hSHMT2-specific inhibitors it is critical to understand the SHMT catalysis at the atomic level of detail. However, there is a gap in our knowledge of the SHMT-catalyzed reaction mechanism, which is based on the acid-base chemistry, because the locations and movement of H atoms along the reaction pathway are not known.

Hydrogen (H) atoms comprise ~50% of all atoms in proteins and contribute to most noncovalent interactions, such as H-bonding, van der Waal's forces, and electrostatics. Thus, H atoms are key players in enzyme substrate and ligand binding, protein-protein interactions, and proton transfer reactions.<sup>39,40</sup> Moreover, the presence or absence of an H atom in a chemical group, such as imidazole of histidine (His) or carboxylate of aspartate/glutamate (Asp/Glu), determines the protonation state and thereby the electrical charge of that amino acid residue. Whereas X-ray diffraction continues to be the gold standard in protein structure determination, the resultant structures provide little to no information on the positions of functional H atoms even at ultrahigh resolutions.<sup>41</sup> This is a consequence of the X-ray scattering magnitude of H being significantly

smaller than that for the other atoms in biomacromolecular structures. Because H and its heavier isotope deuterium (D) atoms have comparable neutron scattering lengths to heavier atoms (C, N, O) within the protein, neutron diffraction can resolve the positions of H/D at moderate resolutions.<sup>42-45</sup> The ability to assign protonation states based on the location of H/D atoms by neutron diffraction is critical to understanding the mechanism of catalysis, molecular recognition, and in designing specific inhibitors through structure-based drug design.<sup>46-50</sup> As a result, a neutron structure can help resolve a biochemical issue for a specific protein<sup>51,52</sup> and provide an accurate model for drug design.<sup>49,50</sup>

The first step in the SHMT-catalyzed conversion of L-Ser to Gly is the formation of the external aldimine through a transaldimination reaction with the L-Ser substrate, a common feature for PLP-dependent enzymes. The amino group of L-Ser displaces the catalytic Lys, initially covalently linked to PLP within the internal aldimine state, to generate an external aldimine state followed by THF binding that ultimately leads to the THF-dependent conversion of L-Ser to Gly and 5,10-MTHF production. However, there is no accepted catalytic mechanism for this conversion, but the available literature contains two predominant hypotheses (Figure S1).<sup>53-56</sup> In a proposed two-step mechanism, after the external aldimine is formed, the  $\beta$ -hydroxyl of L-Ser is deprotonated by a general base, presumably Glu53 in *Tth*SHMT (Glu98 in hSHMT2, Figure 1d), in a retro-aldol reaction to promote the release of formaldehyde. The transient formaldehyde intermediate then undergoes a nucleophilic attack by the endocyclic N5 atom of THF, likely without dissociating from the active site. A histidine residue (His122 in *Tth*SHMT, His171 in hSHMT2) has also been suggested as a possible general base in this mechanism.<sup>53,57,58</sup> In the competing one-step mechanism, C $\beta$  of L-Ser within the external aldimine undergoes a direct nucleophilic attack by THF N5, facilitating  $\alpha$ -elimination. In this reaction mechanism, protonation of Glu53 is obligatory as it acts as a general acid to protonate the hydroxyl to dehydrate C $\beta$ .<sup>55,56,59,60</sup> Interestingly, within the one-step mechanism, the reaction has been proposed to proceed either by a nucleophilic displacement route through an intermediate having a covalent bond between L-Ser C $\beta$  and THF N5 that links the external aldimine to THF, or by a concerted mechanism in which the C $\beta$ -N5 bond formation and C $\beta$ -C $\alpha$  bond breakage occur simultaneously in the transition state to generate an intermediate product, 5-hydroxymethyl-THF. Furthermore, additional protonation/deprotonation events must occur in both proposed mechanisms that involve the THF

molecule and other active site residues for the reaction to reach the final products, Gly and 5,10-MTHF. It is also known that SHMT can convert some  $\beta$ -hydroxyamino acids, such as L-threonine and  $\beta$ -phenylserine, in the THF-independent fashion to afford corresponding aldehydes<sup>17,55,61</sup>, which would presumably require the same general base as in the THF-dependent reaction to deprotonate the hydroxyl of the  $\beta$ -substituted amino acid within the respective external aldimine. Taken together, the described THF-dependent SHMT catalytic mechanisms have been proposed each to require as many as six general acid-base groups, although some of them would likely be involved multiple times throughout the reaction pathway, possibly switching their roles, and thereby reducing the number to two or three.<sup>54,55</sup> Discriminating between these reaction mechanisms and pinpointing the general acids and bases involved will require knowledge of the protonation states of the enzyme active site residues, PLP, and THF at different reaction stages – details that can be divulged through neutron crystallography.

PLP-bound hSHMT2 forms hexagonal rod-shaped crystals, space group P6<sub>3</sub>22, with long cell edges (>150 Å) that diffract X-rays to resolutions of 2.3-2.5 Å using synchrotron radiation at cryogenic temperatures.<sup>18,37,62,63</sup> Such hSHMT2 crystals are not amenable to neutron crystallography on the currently available neutron crystallographic beamlines where neutron beam fluxes are significantly weaker than the X-ray beam fluxes at modern synchrotron facilities.<sup>64</sup> Like many metabolic enzymes, SHMT is evolutionarily highly conserved, especially in the active site, enabling us to utilize SHMT from *Thermus thermophilus* (*TthSHMT*, Figure 1c and S2) as a model of the mitochondrial human enzyme. *TthSHMT* and hSHMT2, both fold type I PLP-dependent enzymes, have completely conserved active sites (Figure 1d and S3), but the crystallographic unit cell volume of *TthSHMT* is considerably smaller than that for its human counterpart. *TthSHMT* crystals are well-diffracting, permitting us to use them for room-temperature X-ray and neutron diffraction experiments. Here, we present a 2.3 Å joint X-ray/neutron (XN) structure of the homodimeric *TthSHMT* in the open conformation. The XN structure depicts equivalent active sites in the two protomers with the PLP cofactor covalently bound to the catalytic Lys226 in the internal aldimine state and a sulfate ion filling the substrate binding site in each protomer. For direct comparison with the joint XN structure, we also determined a 2.5 Å room-temperature X-ray structure of hSHMT2 in the internal aldimine state. To track an amino acid substrate within the SHMT active site cavity we obtained a second joint XN structure by soaking the same *TthSHMT*

crystal with deuterated L-Ser (L-Ser-d<sub>7</sub>). This structure revealed L-Ser-d<sub>7</sub> bound in a solvent-exposed peripheral substrate binding site at the entrance to the active site cavity in an apparent *pre*-Michaelis complex, with the sulfate anion continuing to block the L-Ser-d<sub>7</sub> entry deeper into the active site cavity. We further tracked the substrate through the active site cavity by obtaining a room-temperature X-ray structure of a *pseudo*-Michaelis complex by soaking a *Tth*SHMT crystal with D-Ser, a non-reactive enantiomer of L-Ser, which displaced the sulfate anion in protomer A occupying the cationic substrate binding site. Positions of H atoms were revealed by the nuclear density maps in the XN structures, thereby allowing us to accurately assign the protonation states and electrical charges for all amino acid residues, L-Ser substrate and the PLP. By directly observing H atom locations and tracking the substrate positions, our study provides unique atomic-level understanding of the SHMT active site that sheds new light on the enzyme's catalytic mechanism by eliminating speculation on the roles of many active site residues. We believe this knowledge can be also employed to inform the design of anticancer drugs targeting hSHMT2.

## RESULTS

### Accurate map of protonation states in substrate-free *Tth*SHMT

SHMTs from prokaryotic sources function as biological homodimers, whereas eukaryotic SHMTs exist as homotetramers composed of a dimer of dimers.<sup>17,18</sup> *Tth*SHMT crystallizes as a dimer with virtually equivalent active sites in protomers A and B, as seen in the cryogenic X-ray structure (PDB ID 2DKJ, unpublished). Similar to the other fold type I PLP-dependent enzymes, each protomer of *Tth*SHMT can be divided into large and small domains.<sup>1,65</sup> The large domain (residues 33-284) is composed of a seven-strand mixed  $\beta$ -sheet surrounded by  $\alpha$ -helices in an  $\alpha/\beta/\alpha$  sandwich fold (Figure S4). The small domain (residues 7-32 and 285-407) is made up of residues from both the N- and C-termini and the structure is dictated by anti-parallel  $\beta$  strands and  $\alpha$ -helices in a two-fold  $\alpha/\beta$  sandwich. The small domain is primarily responsible for establishing inter-subunit contacts.<sup>54,66,67</sup> Each active site is located at the corresponding dimer interface and is constructed from residues of both protomers.

To visualize the active site of SHMT at the atomic level of detail, we obtained a 2.3 Å neutron crystallographic structure of substrate-free *Tth*SHMT in the internal aldimine state jointly

refined with 2.0 Å X-ray diffraction data (Table S1, Supplementary Data 1, PDB ID 8SUJ). We were unable to obtain neutron diffraction data from the P6<sub>5</sub>22 crystals of hSHMT2 because they diffracted X-rays at a synchrotron only to 2.5 Å (Table S2). Therefore, to directly compare hSHMT2 with the joint XN structure of *Tth*SHMT, we also determined a 2.5 Å room-temperature X-ray structure of hSHMT2 in the internal aldimine state (Supplementary Data 2, PDB ID 8SSJ), because all previous hSHMT2 X-ray structures were done at cryogenic temperatures.<sup>18,37,62,63</sup> We accurately mapped the H atom positions (observed as D atoms), revealing the active site protonation states, H-bonding networks, water molecule orientations, and consequently electrostatics. Both diffraction datasets were collected at room temperature and pH of 5.5 from the same H/D exchanged protein crystal (Table S1). In *Tth*SHMT, PLP is covalently bound to Lys226 (Lys280 in hSHMT2) through a Schiff base linkage with the side chain ζ-amino group forming the internal aldimine. Because the active site structures are identical within the experimental coordinate error of 0.3 Å in the protomers A and B of the *Tth*SHMT homodimer, we selected protomer B for our structural analyses below, unless stated otherwise. The pyridine nitrogen, N1, of the PLP cofactor is protonated. Thus, it is positively charged (Figure 2a). N1 participates in an H-bond with the negatively charged side chain carboxylate of Asp197 (Asp251 in hSHMT2) conserved in all fold type I PLP-dependent enzymes at an O...D distance of 1.7 Å. H-bond distances observed in the joint XN structures reported here are given henceforth as the distance between a D (or H) atom of an H-bond donor and a heavy atom of an H-bond acceptor, unless otherwise specified. The Asp197 side chain is correctly situated to interact with PLP's N1 by a network of H-bonds with Asn98 (Asn147 in hSHMT2), His125 (His174 in hSHMT2) and the backbone amide ND of Ala199 (Ala253 in hSHMT2). The pyridine N1 protonation has a significant role in the PLP-dependent catalysis, increasing the efficiency of electron delocalization to the pyridine ring and leading to the stabilization of negatively charged reaction intermediates, but the enzyme microenvironment near N1 regulates its protonation state.<sup>68,69</sup> Protonated N1 was also observed in the joint XN structure of aspartate aminotransferase (AAT), another fold type I PLP-dependent enzyme,<sup>44,70</sup> where it is similarly H-bonded to a conserved aspartate residue. Because the active sites of *Tth*SHMT and hSHMT2 are conserved we propose that N1 is also protonated in hSHMT2, participating in an H-bond with Asp251 having the N...O distance of 2.9 Å (Figure S5).

The Schiff base nitrogen,  $N_{SB}$ , of the PLP cofactor is found to be not protonated, thus neutral, and the C=N group is rotated by  $35^\circ$  and  $29^\circ$  above the plane of the pyridine ring on the *si* face of the active site in protomers A and B, respectively.  $N_{SB}$  makes no interactions with the active site residues. In hSHMT2, the Schiff base rotates to a much greater extent, with the corresponding dihedral angles in protomers A and B of  $67^\circ$  and  $73^\circ$  (Figure S6). The Schiff base out-of-plane rotation appears to be the attribute of the  $N_{SB}$  neutrality as was also observed in the joint XN structure of AAT.<sup>44</sup> Therefore, the above considerations point to the high probability of  $N_{SB}$  atom not being protonated in hSHMT2 either. Interestingly, in the high-resolution cryogenic X-ray diffraction structure of *Th*SHMT (PDB ID 2DKJ) this torsion angle is reduced to  $26^\circ$  and  $15^\circ$  in protomers A and B, respectively. Protonation of  $N_{SB}$  promotes the Schiff base to be more co-planar with the pyridine ring due to an H-bond formation with the phenolic oxygen O3', as demonstrated by previous DFT calculations.<sup>9,44,71</sup> It is thus reasonable to suggest that  $N_{SB}$  may be protonated at lower temperatures, pointing to possible temperature dependence of the Schiff base protonation state. To gain more insight into the energy profile of the Schiff base rotation around the pyridine ring, we performed DFT calculations on the PLP-Lys model (Figure S7). Indeed, when  $N_{SB}$  is not protonated the potential energy surface is shallow with two local energy minima at  $\sim 21^\circ$  and  $-21^\circ$  in agreement with the observed dihedral angles in the joint XN structures. The global minimum occurs when C=N bond is rotated away from O3' and is coplanar with the pyridine ring, a geometry that cannot be achieved in the protein due to geometric constraints of Lys being connected to other residues. In sharp contrast, the protonated Schiff base shows a very different potential energy profile, with the global minimum found for the C=N bond fully coplanar with the pyridine ring facing the deprotonated O3'.

The nuclear (neutron scattering length) density near the phenolic oxygen O3' indicates the absence of a D atom. Therefore O3' is observed as a deprotonated phenolate in the current joint XN structure. The resulting negative charge on O3' is delocalized through resonance with the protonated pyridine ring. Both  $2F_O - F_C$  and omit  $F_O - F_C$  nuclear density maps clearly show that O3' engages in an unconventional C-H...O hydrogen bond with C $\delta$ 2 of His200 imidazole (His254 in hSHMT2) with the distance of 2.0 Å (Figure 2b). This orients the protonated N $\epsilon$ 2 of His200 to donate a 1.7 Å H-bond with the sulfate ion ( $SO_4^{2-}$ ) located in the substrate binding pocket near Arg358 (Arg425 in hSHMT2) (Figure 2c). The sulfate is present in the crystallization solutions

and its presence in the structure is an artefact of the crystallization conditions. In contrast to the current joint XN structure, in some previously published X-ray crystal structures of SHMT from other species, the His residues equivalent to *Tth*SHMT His200 were incorrectly modeled to participate in H-bonds with O3' with their N $\delta$ 1 atoms.<sup>53,54</sup> In *Tth*SHMT, His200 is found to be singly protonated, thus neutral, its N $\delta$ 1 accepting a 1.9 Å H-bond with N $\epsilon$ 2 of His312 (His380 in hSHMT2). Furthermore, His312 N $\epsilon$ 2 faces its own main chain amide ND to form a 2.2 Å H-bond and is also singly protonated and neutral. Thus, a network of unconventional C-H...O and conventional N-H...N H-bonds (Figure 2b) that connects the negatively charged PLP O3' to the His312 main chain ensures the neutral charge of both His200 and His312, which would undoubtedly remain neutral at the physiological pH of 7.4. The H-bond network connecting PLP O3' to the main chain NH of His380 is identical in hSHMT2, suggesting the same protonation states of O3', His254, and His380 in the human enzyme as observed in *Tth*SHMT.

Found on the *re* face of PLP, the sulfate ion mimics substrate binding, occupying the presumed substrate binding pocket. The sulfate position is stabilized by an H-bonding network consisting of Ser31 (Ser76 in hSHMT2), His200, Arg358, Tyr61\* (Tyr106\* in hSHMT2, the asterisk denotes that the residue belongs to the other protomer), and two water molecules (Figure 2c). The sulfate participates in a salt bridge with Arg358, with H-bond distances of 1.9 Å and 2.0 Å. Ser31 and Tyr61\* donate their hydroxyl D atoms in 2.0 and 1.9 Å H-bonds with the sulfate, respectively. The two water molecules coordinate the sulfate with H-bond distances of 1.8 and 1.9 Å and link it to Glu53\* (Glu98 in hSHMT2). Glu53, a conserved active site residue that is proposed to take part in the THF-dependent cleavage of L-Ser as a general acid/base catalyst,<sup>55,56,58,72</sup> is deprotonated and positioned on the *re* face of PLP near the substrate binding pocket (Figure 2c). We believe Glu98 in hSHMT2 is also deprotonated because the electrostatic environment is the same in the substrate binding pocket of the human enzyme. In our room-temperature structure of hSHMT2 the substrate binding pocket is occupied by a chloride anion, Cl<sup>-</sup>. The presence of Cl<sup>-</sup> can be physiologically relevant as intracellular Cl<sup>-</sup> concentration can be ~ 20-100mM.<sup>73,74</sup> The affinity of this pocket for anionic ligands is probably dictated by the nearby positively charged arginine residue, Arg358 in *Tth*SHMT or Arg425 in hSHMT2. In contrast, a glycerol molecule is found in the substrate binding pocket in the X-ray structure of hSHMT2 obtained under cryogenic conditions with glycerol added as the cryoprotectant.<sup>18</sup> Glycerol is normally used at concentrations

of  $\sim 2\text{M}$  for cryoprotection of protein crystals and can undoubtedly outcompete an anion for binding in this pocket.

The PLP pyridine ring is sandwiched between Ala199 (Ala253 in hSHMT2) on the *si* face and His122 (His171 in hSHMT2) on the *re* face (Figure 2d). The side chain of Ala199 makes C-H $\cdots\pi$  interactions with the  $\pi$ -conjugated system of the pyridine ring, with the heavy atom distances in the range of 3.5-3.9 Å. This residue is conserved in fold type I PLP-dependent enzymes. Site-directed mutagenesis studies suggest the residue at this position may help modulate the  $\text{pK}_a$  of the internal aldimine  $\text{N}_{\text{SB}}$ .<sup>75</sup> The His122 imidazole ring sits parallel to the pyridine ring in a  $\pi$ - $\pi$  stacking interaction with the distances of 3.5-3.7 Å between the heavy atoms. His122 is singly protonated on  $\text{N}_{\epsilon 2}$  (Figure 2c and 2d). The neutral charge of His122 is the result of the imidazole  $\text{N}_{\delta 1}$  accepting a bifurcated H-bond with the side chain hydroxyl and main chain amide ND of Thr124 (Thr173 in hSHMT2) with the distances of 2.0 and 2.2 Å, respectively. His122 also donates a 1.8 Å H-bond with a water molecule that mediates the imidazole interactions with the sulfate and PLP phosphate group (Figure 2c). Again, in hSHMT2 His171 and Ala253 make very similar interactions with PLP and H-bonds with adjacent residues, allowing us to postulate that the His171 protonation state is the same as that of His122 in *Tth*SHMT (Figure S5).

In addition to making a covalent bond with the catalytic Lys226 and an H-bond between pyridine  $\text{N}_1$  and Asp197, the PLP position in the active site is further stabilized through an intricate network of H-bonds and water-mediated interactions made by the phosphate group with the phosphate binding pocket. In *Tth*SHMT, the phosphate binding pocket is composed of the side chains of Ser95 (Ser144 in hSHMT2) and Tyr51\* (Tyr96 in hSHMT2), the main chain amide NDs of Gly94 (Gly143 in hSHMT2), Ser95, and Gly258\* (Gly326 in hSHMT2), and three water molecules. The phosphate group makes 1.8 and 1.9 Å hydrogen bonds with Ser95 and Tyr51\* hydroxyls, but somewhat longer 2.0-2.3 Å H-bonds with the main chain amides (Figure 2e). The water molecules mediate the phosphate group interactions with the main chain carbonyl of Gly258\*, and with the side chain of His122 mentioned above.

### **L-Ser binds as a zwitterion in the *pre*-Michaelis complex**

To obtain a complex with bound substrate, we repurposed the *Tth*SHMT crystal from the first neutron diffraction experiment by soaking it with the deuterated serine substrate, L-Ser- $\text{d}_7$  at

500 mM concentration. We then collected a 2.3 Å room-temperature neutron diffraction dataset on this crystal. Joint XN refinement was performed using 2.0 Å room-temperature X-ray diffraction data collected on the same crystal immediately after the neutron experiment. Although it was expected that L-Ser-d<sub>7</sub> would displace Lys226 in the Schiff base linkage to PLP to form the external aldimine, the intact internal aldimine was observed in this structure with the sulfate ions still occupying the substrate binding sites in both protomers. However, we observed L-Ser-d<sub>7</sub> bound in a pocket on the periphery of the active site in protomer A, with the substrate positioned in front of the sulfate ion (Figure 3a and 3b; Supplementary Data 3, PDB ID 8SUI). We therefore call this complex, *Tth*SHMT/L-Ser, a *pre*-Michaelis complex, in which both protomers display the same open conformation observed in the substrate-free internal aldimine joint XN structure. The enzyme active site entrance resembles a funnel, with a wide opening that narrows down going deeper towards PLP. The peripheral binding site of L-Ser-d<sub>7</sub> is where various ligands, including THF analogs and non-THF high-affinity inhibitors such as SHIN-1, are located in the structures of SHMT from different sources;<sup>37,62,76,77</sup> however, amino acid substrates have not been captured previously in this site.

In protomer A of *Tth*SHMT/L-Ser, L-Ser-d<sub>7</sub> is observed in the zwitterionic state with the protonated  $\alpha$ -amino group facing a hydrophobic patch made up of Leu117 (Leu166 in hSHMT2), Leu123 (Leu172 in hSHMT2), and Phe252\* (Phe320\* in hSHMT2), whereas the deprotonated carboxyl group is directed into the bulk solvent (Figure 3b). The side chain hydroxyl is positioned towards the substrate binding site next to the sulfate ion. The hydroxyl group of L-Ser-d<sub>7</sub> donates its D atom in a weak bifurcated H-bond with a water molecule (2.20 Å), mediating its contact with Glu53\*, and the sulfate ion (2.6 Å) (Figure 3c). The hydroxyl group of Tyr61\* (Tyr106\* in hSHMT2) rotates away from the sulfate ion to lose its H-bond as found in the substrate-free *Tth*SHMT structure, and now makes a 2.2 Å H-bond with the L-Ser hydroxyl oxygen. It is important to note that in the *Tth*SHMT/L-Ser joint XN structure we found the same protonation states in the active sites of both protomers that were observed in the substrate-free structure. The N<sub>SB</sub> atoms of the Schiff bases are neutral, as are all the His residues, whereas the pyridine N1 is protonated. Thus, the non-covalent binding of the substrate in the *pre*-Michaelis complex does not disturb the locations of H atoms and hence the electrostatics of the active site cavity.

### D-Ser occupies the substrate binding site in a *pseudo*-Michaelis complex

In the *Tth*SHMT joint XN structures described above, the position occupied by the sulfate ion (or Cl<sup>-</sup> in hSHMT2) in the active sites is presumably the binding site for amino acid substrate en route to generate the external aldimine intermediate. To capture a substrate bound at this site, we soaked a *Tth*SHMT crystal with D-Ser and obtained a 1.8 Å room-temperature X-ray structure of the complex (Supplementary Data 4, PDB ID 8SSY). The crystal soaking experiment produced a complex with D-Ser in protomer A replacing the sulfate, whereas protomer B retained the sulfate ion in the substrate binding site. Both protomers remained in the open conformation. D-Ser, the enantiomer of the physiological substrate L-Ser, did not proceed to react with C4' of the internal aldimine thus generating a *pseudo*-Michaelis complex, *Tth*SHMT/D-Ser. It was not however possible to produce crystals of this complex of sufficient volume for neutron diffraction experiments.

H-bond interactions made by the D-Ser carboxylate and side chain hydroxyl groups orient the substrate enantiomer in the active site in a specific fashion (Figure 4a). Arg358 acts as the primary anchoring point, forming a salt bridge with the D-Ser carboxylate by making two H-bonds with the N...O distances of 2.7 Å and 2.8 Å. Tyr61\* that interacts with the D-Ser hydroxyl adopts two alternate conformations (Figure 4b). In one conformation, at 57% occupancy, the Tyr61\* phenolic hydroxyl is shifted towards the hydroxyl group of D-Ser with an O...O distance of 2.6 Å. In the other conformation, at 43% occupancy, the Tyr61\* phenol is rotated slightly away to form H-bonds with the D-Ser carboxylate and Ser31 side chain with both distances of 2.8 Å, resulting in a longer interaction of 3.1 Å with the D-Ser hydroxyl group. Tyr61 occupies a single conformation in each of the *Tth*SHMT joint XN structures, therefore its positional disorder in *Tth*SHMT/D-Ser is due to D-Ser binding. In the *pseudo*-Michaelis complex, Glu53\* was observed within an H-bond distance of 2.9 Å to the D-Ser hydroxyl. All these H-bonds would probably be retained when L-Ser binds in the Michaelis complex, before its nucleophilic  $\alpha$ -amino group attacks C4' of the internal aldimine for the reaction to proceed to the external aldimine. Importantly, the H-bonds with the residues corresponding to Arg358, Glu53\* and Tyr61\* are also observed in the X-ray structures of the L-Ser external aldimine state for SHMTs from different species (PDB IDs: 4PFN, 6CDI),<sup>77-79</sup> indicating minimal structural rearrangements upon the external aldimine formation.

In the *Tth*SHMT/D-Ser complex, the D-Ser  $\alpha$ -amino group is 3.0 Å away from N<sub>SB</sub> but is angled away from C4' at a distance of 3.5 Å, possibly making the nucleophilic attack unfavorable and producing the stable *pseudo*-Michaelis complex. In contrast, D-Ser readily reacted with the internal aldimine of *Plasmodium vivax* SHMT to generate the corresponding external aldimine.<sup>78</sup> Because H atoms cannot be observed in the *Tth*SHMT/D-Ser X-ray structure, the protonation states of the cofactor and D-Ser cannot be determined. Therefore, we do not know whether the substrate would still be zwitterionic in the *pseudo*-Michaelis complex as observed in our joint XN structure of the *pre*-Michaelis complex. Previously, based on the XN structures of AAT in the internal and external aldimine states<sup>44</sup>, we suggested that the protonated  $\alpha$ -amino group of an amino acid substrate might transfer a proton to N<sub>SB</sub>, generating the reactive nucleophilic amine, in the Michaelis complex. In *Tth*SHMT/D-Ser, the D-Ser  $\alpha$ -amino group makes a 3.0 Å H-bond with His200 Nε2 that is protonated, indicating the amine may already be in the reactive state. Moreover, the Schiff base out-of-plane dihedral relative to the pyridine ring is flattened to 20° indicative of N<sub>SB</sub> being protonated and positively charged. The corresponding Schiff base dihedral in protomer B is 26°, similar to the value of 29° found in protomer B of the substrate-free *Tth*SHMT XN structure.

### Protonation of Glu residues

Outside of the active site, we detected three protonated glutamate residues in the *Tth*SHMT homodimer (Figure S8). The buried Glu261 is protonated in both protomers A and B on Oε2 atoms forming H-bonds of 1.7 Å and 1.6 Å, respectively, with Glu71 from the neighboring  $\alpha$ -helix. The O...O distances of these H-bonds are 2.7 (protomer A) and 2.5 Å (protomer B). We therefore observed fully localized D atoms participating in short conventional H-bonds in the Glu-Glu pairs. Similar observations were made in high-resolution X-ray structures of protein DJ-1 (Lin et al. 2017) and in the neutron structure of HIV-1 protease gemdiol intermediate.<sup>52</sup> Conversely, in concanavalin A and HIV-1 protease/drug complexes<sup>80,81</sup> D atoms within Glu-Asp or Asp-Asp pairs, respectively, were found participating in the apparent low-barrier hydrogen bonds manifested by the D atoms located midway between the oxygen atoms. It is possible that the H-bonding and electrostatic environment around the carboxylic groups modulates the nature of the formed H-bond. For example, in *Tth*SHMT, positively charged Arg59 is positioned near Glu71 donating its D atoms in two H-bonds with Oε1. Therefore, Glu71 Oε1 forms a bifurcated H-bond

with Arg59 and protonated Glu261, possibly driving the localization of D on Glu261 Oε2. In hSHMT2, residue 329 corresponding to *Tth*SHMT Glu261 is substituted to a smaller asparagine, which makes a water-mediated contact with Glu116 (Glu71 in *Tth*SHMT) unable to form a direct H-bond.

Another protonated glutamic acid residue, Glu65 in protomer A of *Tth*SHMT, participates in a superficial long-range, electrostatic interaction with Asp333\* from the small domain of protomer B (O...O distance of ~4 Å). In protomer B, Glu65\* is not protonated, curling in to make an H-bond with its own main chain amide ND, and is > 5 Å away from the Asp333 carboxylate of protomer A. Asp333 is not conserved in hSHMT2, where it is Glu401. Although glutamate has a larger side chain compared to aspartate, Glu401 is > 4 Å away from Glu110\* for all pairs of these residues in hSHMT2 homotetramer.

## DISCUSSION

Many biochemical reactions proceeding inside active sites of enzymes are governed by the general acid-base catalysis. The catalytic reaction is typically facilitated by proton transfer events that occur along the reaction pathway. Moreover, ligand (substrate or inhibitor) binding in the enzyme active site may be accompanied by protonation state changes that alter the electrostatic landscape of the binding site relative to its initial unbound state. As a result, to fully understand enzyme catalysis and to inform structure-based drug design it is essential to capture these protonation and deprotonation events by observing the locations and movement of H atoms before and after a ligand binds and throughout the catalytic reaction. The H atom positions determine the protonation states and electrical charges of amino acid residues and ligands, and hence electrostatics, within the enzyme active sites. In this work, we utilized the power of neutrons to directly detect and visualize positions of H atoms in a protein structure to study the catalytic mechanism and substrate binding for *Tth*SHMT enzyme. We extended the structural and mechanistic insights gleaned from the joint XN structures to the homologous human enzyme hSHMT2 whose active site structure and amino acid composition is conserved relative to that of *Tth*SHMT.

The serine substrate enters the *Tth*SHMT active site through a funnel-shaped entrance first being captured at the peripheral substrate binding site (*pre*-Michaelis state) and then tracked deeper

to the cationic substrate binding site (Michaelis state) (Figure 5a). We observed in the room-temperature joint XN structure of *Tth*SHMT/L-Ser complex that L-Ser is captured in the zwitterionic state ( $^+\text{NH}_3\text{-C}\alpha\text{-COO}^-$ ) at the peripheral substrate binding site that is mainly hydrophobic and exposed to the bulk solvent. The L-Ser position is stabilized by the substrate's hydroxyl making a direct H-bond with Tyr61\* and a water-mediated contact with Glu53\*. The Tyr61\* phenolic hydroxyl rotates to donate its D in an H-bond with L-Ser from its position in the substrate-free structure, where it formed an H-bond with the sulfate anion. Previous structural and mutagenesis data suggested that Tyr61 plays a role in the interconversion of the open and closed conformational states of the active site.<sup>82</sup> We further tracked the substrate deeper into the active site. In the *Tth*SHMT/D-Ser room-temperature X-ray structure, D-Ser occupies the substrate binding site that evidently has some affinity to anions, as sulfate is found in *Tth*SHMT, and chloride is observed in hSHMT2 room-temperature structures. Therefore, this binding site, likely occupied in the Michaelis complex, can be called a cationic substrate binding site. The D-Ser  $\alpha$ -amino group is 3.5 Å away from PLP C4' and is expected to be positioned closer in the actual Michaelis complex with L-Ser. At the cationic substrate binding site, D-Ser is anchored by a salt bridge with Arg358 and by direct H-bonds with Glu53\* and Tyr61\*, the H-bonds anticipated to be conserved in the Michaelis complex to orient the L-Ser substrate correctly to carry out a nucleophilic attack on the PLP Schiff base C4' atom. Interestingly, SHIN-1, a high-affinity inhibitor of hSHMT2, binds at the peripheral substrate binding site in the X-ray structure of *E. faecium* SHMT displacing the Tyr61\* side chain by a bulky isopropyl group.<sup>77</sup> Based on our structures, we propose that the future inhibitor design could take advantage of the existence of the two substrate binding sites that provide opportunities for additional H-bonds and electrostatic interactions to be made by designed inhibitors.

In the internal aldimine state, when the SHMT active site is substrate-free or L-Ser has not yet reached the cationic substrate binding site to react with the Schiff base, the PLP N<sub>SB</sub> atom is not protonated whereas the pyridine N1 is protonated. Identical protonation states for PLP in the internal aldimine state have been observed in the joint XN structure of AAT.<sup>44</sup> In disagreement with our experimental results, however, several published theoretical calculations of the SHMT catalytic mechanism used protonated, positively charged, N<sub>SB</sub> in the internal aldimine state.<sup>60,83,84</sup> In order to proceed through the transaldimination reaction, the incoming amino acid substrate needs to be deprotonated. We believe that N<sub>SB</sub> receives the proton from the  $\alpha$ -amino group of the

zwitterionic amino acid substrate just before or in concert with its attack on the PLP C4' (Figure 5b). Although the anionic  $\alpha$ -amino group of L-Ser ( $-\text{NH}-\text{C}\alpha-\text{COO}^-$ ) was previously suggested as the true substrate<sup>85</sup>, our joint XN structures are not in favor of this scenario because the only nearby residue, His200, capable of accepting an H atom from the substrate is protonated on N $\epsilon$ 2 that would face the  $\alpha$ -amino group.

According to the retro-aldol catalytic mechanism of SHMT catalysis, a general base should abstract a proton from the  $\beta$ -hydroxyl of L-Ser when it becomes part of the external aldimine. The previously proposed candidates for this role are Glu53, His122, His125 and His200.<sup>53,66,86,87</sup> Analysis of our joint XN structures reveals that for the THF-dependent catalysis the best candidate for the general base is Glu53 that is found in the deprotonated, carboxylate, form with an apparently low  $\text{pK}_a$ . We believe that Glu53 would also act as the general base in the THF-independent reaction converting  $\beta$ -hydroxyamino acids to the corresponding aldehydes and Gly. Conversely, none of the three histidine residues, although neutral, can act as a general base because they are locked in H-bonds that would preclude their protonation, making them non-titratable residues at physiological pH. Specifically, His122,  $\pi$ - $\pi$  stacked against the PLP pyridine ring, participates in H-bonds with Thr124 and its own main chain amide NH using N $\delta$ 1, whereas N $\epsilon$ 2-H is directed towards the substrate. His125 donates N $\epsilon$ 2-H in an H-bond with Asp197 and N $\delta$ 1 makes an H-bond with the main chain amide NH of His122. Moreover, His125 is  $> 9 \text{ \AA}$  away from the cationic substrate binding site, too far to participate in the reaction. His200 also cannot be considered as a possible catalytic base. It is part of the H-bond network that links O3' to the His312 main chain. His200 N $\epsilon$ 2-H faces the substrate, as mentioned above, and its N $\delta$ 1 is in an H-bond with His312 N $\epsilon$ 2-H. Consequently, none of these histidine residues can be protonated during the SHMT catalysis. Moreover, they remained neutral even at the acidic pH of 5.5 used to grow *Tth*SHMT crystals. Our results thus indicate that Glu53 plays the general base role to deprotonate  $\beta$ -hydroxyl of L-Ser (Figure 5b), in agreement with the recent QM/MM studies which concluded that the retro-aldol mechanism is most probable for SHMT.<sup>60</sup>

In conclusion, neutron crystallography at near-physiological (room) temperature was used to obtain all-atom structures of *Tth*SHMT where positions of H and D atoms were accurately located revealing protonation states of key residues and, as a result, mapping electrostatics in the enzyme active site. We observed protonated (positively charged) pyridine N1, deprotonated

(negatively charged) phenolic O3', and non-protonated (neutral) Schiff base N<sub>SB</sub> atoms within the PLP cofactor. Interestingly, all active site His residues were found in the monoprotonated (neutral) state, participating in H-bonding networks that would prevent their protonation at physiological pH. Based on the observed protonation states in the active site, Glu53 is proposed as the best candidate for the general base catalyst to orchestrate the retro-aldol transformation of L-Ser into Gly. The pathway of the substrate amino acid serine entering the active site cavity was tracked revealing a peripheral binding site, where the amino acid exists in the zwitterionic state, and the cationic binding site, where the substrate's  $\alpha$ -amino group is poised for the nucleophilic attack on Schiff base C4' atom. Understanding the next stages in the SHMT catalysis will require mapping the H atom movements in the external aldimine with bound L-Ser and in complexes with THF or its analogs.

## **MATERIALS and METHODS**

### **General information**

Columns for protein purification were purchased from Cytiva (Piscataway, New Jersey, USA). His-tagged Tobacco Etch Virus (TEV) protease was produced in-house. Crystallization reagents and supplies were purchased from Hampton Research (Aliso Viejo, California, USA). Crystallographic supplies for crystal mounting and X-ray and neutron diffraction data collection at room temperature were purchased from MiTeGen (Ithaca, New York, USA) and Vitrocom (Mountain Lakes, New Jersey, USA). Deuterated L-Ser and hydrogenous D-Ser were purchased from Millipore Sigma (St. Louis, Montana, USA).

### ***Tth*SHMT expression and purification**

The *glyA* gene (Figure S9) encoding SHMT enzyme from bacterium *Thermus thermophilus* was codon optimized, synthesized, and cloned into kanamycin-resistance plasmid, pJ411 (ATUM, Newark, CA), in addition to a DNA sequence encoding for an N-terminal polyhistidine-(His<sub>6</sub>)-tag with a 34 amino acid long linker. A TEV protease tag ENLYFQ was introduced at the *Tth*SHMT N-terminus sequence so that after cleavage the enzyme sequence started from Ser3 (Figure S9). This was done because residues Met1 and Val2 are not visible in the electron density maps in the

previously deposited *Tth*SHMT structure (PDB ID 2DKJ). The plasmid was transformed into BL21(DE3) competent *E. coli* cells for expression. Transformed cells were grown in Luria-Bertani (LB) media supplemented with 50 µg/mL kanamycin antibiotic at 37 °C to an optical density of 0.8-1.0 and induced overnight with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 22 °C (approximately 16-18 h). Induced cells were harvested by centrifugation at 5,660 rpm at 4° C, producing typical yields of 10 g of cells per 1 L of cell culture. A buffer containing 50 mM sodium phosphate pH 7.5, 500 mM NaCl, and 10 mM imidazole was used to resuspend the cell pellet, utilizing 5 mL of lysis buffer per gram of wet cell paste. The cells were stirred on ice for 30 minutes prior to mechanical sonication. The cell lysate was clarified by centrifugation at 17,000 rpm (~30,000 g) for 30 min. The supernatant was loaded onto a 5 mL HisTrap FF nickel column equilibrated with 20 mM HEPES pH 7.5, 100 mM NaCl, and 10 mM imidazole and washed with 5 column volumes (CV) of 20 mM HEPES pH 7.5, 100 mM NaCl, and 20 mM imidazole. The pure, tagged protein was eluted with 20 mM HEPES pH 7.5, 100 mM NaCl, and 500 mM imidazole in a linear gradient at relatively low imidazole concentrations (~50-60 mM). SHMT-containing fraction were pooled, and TEV protease was added to cleave the poly-histidine tag (1 mg TEV/100 mg of tagged protein). After room temperature overnight dialysis against 20 mM HEPES pH 7.5, 100 mM NaCl, and 1 mM EDTA, the TEV-treated fractions were loaded onto the 5 mL HisTrap FF nickel column and eluted in the flow-through. Pure *Tth*SHMT, verified by SDS-PAGE, was then dialyzed overnight against 40 mM NaOAc pH 5.4 and 1 mM PLP at 4 °C, concentrated to 19 mg/mL, and stored at -30 °C in the presence of 20% (v/v) glycerol.

### **hSHMT2 expression and purification**

The *SHMT2* gene encoding the SHMT enzyme from human mitochondrion (residues 37-504, Uniprot ID P34897, Figure S10) with mitochondrial leader sequence deleted was codon optimized, synthesized, and cloned into kanamycin-resistance plasmid, pJ411 (ATUM, Newark, CA). The DNA sequence encoded for a 35 amino acid long linking sequence containing an N-terminal His<sub>6</sub>-tag and the TEV protease cleavage tag so that after cleavage the enzyme sequence starts with Gly37 (Figure S10). The plasmid was transformed into BL21(DE3) competent *E. coli* cells for expression and grown in LB media supplemented with 50 µg/mL kanamycin antibiotic at 37 °C. The culture was grown to an optical density of 0.8-1.0 and induced overnight with 1 mM IPTG at 22 °C (approximately 16-18 h). Induced cells were harvested by centrifugation at 5,660 rpm at 4°

C. Using 5 mL of lysis buffer per gram of wet cell paste, the cell pellet was resuspended in a buffer containing 20 mM HEPES pH 7.5, 500 mM NaCl, and 10 mM imidazole. Lysozyme was added at 0.1 mg/mL as the cells were stirred on ice for 30 min then subsequently mechanically sonicated. The lysates were clarified by centrifugation at 30,000 g for 30 min and then loaded onto a 5 mL HisTrap FF nickel column equilibrated with 20 mM HEPES pH 7.5, 500 mM NaCl, and 10 mM imidazole. The column was washed with 20 mM HEPES pH 7.5, 500 mM NaCl, and 20 mM imidazole and tagged hSHMT2 was eluted with a linear gradient of 20 mM HEPES pH 7.5, 500 mM NaCl, and 500 mM imidazole. TEV protease was added to the pooled hSHMT2-containing fractions in order to cleave the poly-histidine tag (1 mg TEV/100 mg of tagged protein). The sample was dialyzed overnight at room-temperature against 20 mM HEPES pH 7.5, 250 mM NaCl, and 1 mM EDTA. The TEV-treated fractions were loaded onto a 5 mL HisTrap FF nickel column and eluted in the flow-through. Pure hSHMT2, verified by SDS-PAGE, was then dialyzed overnight against 20 mM HEPES pH 7.5, 300 mM NaCl, and 1 mM PLP at 4 °C and concentrated to 18 mg/mL for crystallization setups. hSHMT2 was crystallized in 50 mM Tricine pH 8.2-8.4 and 11-13% PEG 3350 in sitting drop vapor diffusion experiments.

### **Crystallization and H/D-exchange**

Aliquots of pure *Tth*SHMT were thawed and dialyzed against 40 mM NaOAc pH 5.4 and 1 mM PLP to remove the glycerol. *Tth*SHMT (19 mg/mL) was crystallized in 40 mM NaOAc pH 5.5, 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 0.5 M Li<sub>2</sub>SO<sub>4</sub> in sitting drop vapor diffusion experiments, producing showers of crystals and some crystal aggregates. Several crystal aggregates were crushed in their crystallization drops to create microcrystals for microseeding experiments. Large, single crystals for neutron diffraction were grown using a streak seeding method in 9-well glass plates and sandwich box setups. Specifically, after large crystallization drops were set up, a seeding tool from Hampton Research was dipped into the crushed crystals and then dipped quickly into the new crystal drops to transfer microcrystals. In this approach, only a few, but larger, crystals per drop of *Tth*SHMT grow. A crystal of ~2 mm<sup>3</sup> in volume suitable for neutron diffraction was mounted in a 2 mm-inner diameter quartz capillary with a liquid plug of 40 mM NaOAc pH 5.5, 1.0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 0.5 M Li<sub>2</sub>SO<sub>4</sub> prepared in 100% D<sub>2</sub>O to perform H/D-vapor exchange. The same crystal was used for both neutron crystallographic experiments, in the absence and presence of substrate L-Ser-d<sub>7</sub>. To prepare the crystal for soaking with L-Ser, the quartz capillary was unsealed,

and the H/D-exchange liquid plug was removed. The capillary was then filled with 500 mM deuterated L-Ser, 40 mM NaOAc pH 5.5, 1.0M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 0.5 Li<sub>2</sub>SO<sub>4</sub> left to soak overnight. The following day, the soaking solution was removed and replaced with a liquid plug of 40 mM NaOAc pH 5.5, 1.0M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 0.5 Li<sub>2</sub>SO<sub>4</sub> in 100% D<sub>2</sub>O to perform H/D-vapor exchange. *Tth*SHMT crystals for D-Ser soaking experiments were first transferred to a fresh drop containing 0.1 M NaOAc pH 5.5 and 15% PEG 4000, to remove excess sulfate, then moved to a drop containing 0.5 M D-Ser in 0.1 M NaOAc pH 5.5 and 15% PEG 4000. hSHMT2 at 18 mg/mL was crystallized in 50 mM Tricine pH 8.2-8.4 and 11-13% PEG 3350 in sitting drop vapor diffusion experiments producing hexagonal rod-shaped crystals.

### Neutron Diffraction Data Collection

Neutron diffraction was tested at room temperature and a preliminary dataset was obtained on the LADI-DALI beamline<sup>88</sup> at the Institut Laue-Langevin (ILL) in Grenoble. A complete room-temperature neutron diffraction dataset for the *Tth*SHMT internal aldimine were collected on the IMAGINE<sup>89-92</sup> single-crystal diffractometer at the HFIR at ORNL using a neutron wavelength range of 2.8-4.5 Å. Each neutron image was composed of a 20 h exposure of the crystal held in a stationary position. The crystal was rotated along the vertical axis ( $\Delta\phi=8^\circ$ ) before collecting each successive image. The crystal orientation was changed three times by tilting the capillary with respect to the incident neutron beam to improve data completeness. In total, 44 neutron diffraction images were collected. Neutron diffraction data processing was performed with a version of LAUEGEN<sup>93,94</sup> modified to account for the geometry of the cylindrical image plate detector. The wavelength-normalization curve was determined using the intensities of symmetry-equivalent reflections at different wavelengths in LSCALE.<sup>95</sup> No explicit absorption corrections were applied. The data were scaled and merged in SCALA (Weiss 2001).

Room-temperature neutron diffraction data for the *Tth*SHMT-L-Ser pre-Michaelis complex were collected on the instrument MaNDi<sup>96,97</sup> at the Spallation Neutron Source (SNS) at ORNL. The crystal was held stationary for 20 h exposures and rotated 10° around the  $\phi$ -axis before collecting the next image. All neutrons between 2 and 4.16 Å were used to collect the frames, with a total of 25 images collected. Neutron diffraction data collection on MaNDi were processed and integrated with 3D time-of-flight profile fitting in Mantid.<sup>98,99</sup> Wavelength normalization of the

data was performed with LAUENORM<sup>94,100</sup> and the data were scaled and merged in SCALA.<sup>101</sup> Neutron data collection statistics for both datasets are shown in Table S1.

### **Room-temperature X-ray diffraction data collection and structure refinement**

Room temperature X-ray diffraction data collection for *Tth*SHMT crystals was carried out on a Rigaku HighFlux HomeLab instrument equipped with a MicroMax-007 HF X-ray generator, Osmic VariMax optics, and a DECTRIS Eiger R 4M detector at Oak Ridge National Laboratory. The data were indexed and integrated using the CrysAlis Pro software (Rigaku, The Woodlands, TX), then reduced and scaled with the AIMLESS program in the CCP4 software suite.<sup>102,103</sup> The *Tth*SHMT room-temperature X-ray structures were solved by molecular replacement in PHASER<sup>104</sup> using phases from PDB code 2DKJ. Room temperature X-ray diffraction data collection for hSHMT2 crystals was performed from a single crystal on the ID19 beamline at SBC-CAT using a Pilatus3 X 6M detector at the Advanced Photon Source (APS). X-ray diffraction data were integrated and scaled using the HKL3000 software suite.<sup>105</sup> To minimize radiation damage to the hSHMT2 crystal, the X-ray beam intensity was attenuated 40 times and the data were collected with 0.2sec/frame rate. The radiation damage to the crystal was estimated by the HKL3000 software to be less than 5%. The hSHMT2 structure was solved by molecular replacement using PHASER.<sup>104</sup> The cryo-temperature X-ray structure of hSHMT2 (PDB ID 4PVF)<sup>18</sup> was used as a starting model. All the structures were subsequently refined against the room temperature data with Phenix.refine from the PHENIX suite<sup>106,107</sup> and COOT.<sup>108-110</sup> Geometry validation was aided by Molprobity.<sup>111</sup> All ligand restraints were generated with eLBOW<sup>112</sup> using geometry optimized by quantum mechanical calculations in Gaussian 16<sup>113</sup> at B3LYP/6-31g(d,p) level of theory. Final data collection and refinement statistics can be found in Table S2.

### ***Joint XN refinement***

Joint XN-refinement of the *Tth*SHMT internal aldimine structure and *Tth*SHMT pre-Michaelis complex were performed using the nCNS<sup>114,115</sup> patch of the Crystallography & NMR Systems (CNS)<sup>115,116</sup> software suite for macromolecular structure determination. The refinement procedure began with a single rigid body refinement followed by a series of atomic position, atomic displacement parameters, and D atom occupancy refinements. The structures were

visualized in the graphics program COOT<sup>108-110</sup>, in between rounds of refinements to inspect side chain modeling and correctly rotate side chain hydroxyl, thiol, and ammonium groups, as well as rotate water molecules to make appropriate H-bonding networks based on both  $2F_o - F_c$  and  $F_o - F_c$  nuclear scattering length density maps. All water molecules in the model are assumed to be and refined as D<sub>2</sub>O as a consequence of the H/D-vapor exchange. Because hydrogenous protein was used in this experiment, the protein was modelled with H atoms at non-exchangeable positions, that is in C-H bonds. All labile, thus exchangeable, H positions in the structure were initially modeled as D until their occupancies were refined. An individual occupancy of 0.56 is associated with the presence of pure H at that position, whereas occupancy of 1.00 is indicative of pure D, because the neutron scattering length of H is -0.56 times that of D. Before depositing the neutron structures to the PDB, coordinates of each D atom were split into two records corresponding to an H and a D partially occupying the same site, both with positive partial occupancies that add up to unity. The percent D at a specific site is calculated according to the following formula: % D = (occupancy(D) + 0.56)/1.56. Neutron refinement statistics can be found in Table S1.

### Quantum chemical calculations

The 2D relaxed potential energy profiles of the C4'-N<sub>SB</sub> bond rotation around the PLP pyridine ring were calculated with a simplified model of the internal aldimine (Figure S7). The model was truncated at C $\beta$  of the lysine-portion of the internal aldimine and the phosphate group was removed, making C5' a methyl group. The scans were completed at B3PW91/Def2-TZVPP level of theory<sup>117,118</sup> and reproduced the geometry observed in the experimental methods. Starting from the torsion angle observed in the *Tth*SHMT joint XN substrate-free internal aldimine structure, the full rotation of C3-C4- C4'-N<sub>SB</sub> was scanned. The scan was also performed on the model with a protonated N<sub>SB</sub>. All calculations were performed with Gaussian16.<sup>113</sup>

### Data Availability

The structures and corresponding structure factors have been deposited into the protein data bank with the PDB accession codes 8SUJ for *Tth*SHMT (Supplementary Data 1), 8SSJ for hSHMT2 (Supplementary Data 2) 8SUI for *Tth*SHMT/L-Ser (Supplementary Data 3), and 8SSY for *Tth*SHMT/D-Ser (Supplementary Data 4). Supporting information is available online.

## **Acknowledgements**

This research at ORNL's High Flux Isotope Reactor (IMAGINE beamline) and at ORNL's Spallation Neutron Source (MaNDi beamline) was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The Office of Biological and Environmental Research supported research at ORNL's Center for Structural Molecular Biology (CSMB), a DOE Office of Science User Facility. ORNL is managed by UT-Battelle LLC for DOE's Office of Science, the single largest supporter of basic research in the physical sciences in the United States. X-ray crystallographic data were in part collected at Argonne National Laboratory using Structural Biology Center (SBC) beamline ID19 at the Advanced Photon Source. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357. The authors thank the Institut Laue Langevin (beamline LADI-DALI) for awarded additional neutron beamtime. This research was supported by a grant from NIH-GMS (R01GM137008) to R.S.P. and A.K.

## **Author contributions**

R.S.P. and A.K. designed the study. V.N.D., C.C., M.H., A.C., O.G., K.L.W. and R.S.P. expressed and purified the proteins. V.N.D., C.C., O.G., and R.S.P. crystallized the proteins. V.N.D., M.H. and A.K. collected X-ray diffraction data. V.N.D. and A.K. reduced X-ray data and refined the structures. V.N.D., M.P.B., and A.K. collected and reduced neutron diffraction data. V.N.D. and A.K. refined joint XN structures. V.N.D. carried out quantum chemical calculations. V.N.D., R.S.P., and A.K. wrote the paper with help from all co-authors.

## **Competing interests**

The authors declare no competing interests.

## References

- 1 Eliot, A. C. & Kirsch, J. F. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu Rev Biochem* **73**, 383-415 (2004). <https://doi.org/10.1146/annurev.biochem.73.011303.074021>
- 2 Liang, J., Han, Q., Tan, Y., Ding, H. & Li, J. Current Advances on Structure-Function Relationships of Pyridoxal 5'-Phosphate-Dependent Enzymes. *Front Mol Biosci* **6**, 4 (2019). <https://doi.org/10.3389/fmolb.2019.00004>
- 3 Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. Modeling of the Spatial Structure of Eukaryotic Ornithine Decarboxylases. *Protein Sci* **4**, 1291-1304 (1995). <https://doi.org/10.1002/pro.5560040705>
- 4 Jansonius, J. N. Structure, evolution and action of vitamin B6-dependent enzymes. *Curr Opin Struct Biol* **8**, 759-769 (1998). [https://doi.org/10.1016/s0959-440x\(98\)80096-1](https://doi.org/10.1016/s0959-440x(98)80096-1)
- 5 Percudani, R. & Peracchi, A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep* **4**, 850-854 (2003). <https://doi.org/10.1038/sj.embor.embor914>
- 6 Amadasi, A. *et al.* Pyridoxal 5'-phosphate enzymes as targets for therapeutic agents. *Curr Med Chem* **14**, 1291-1324 (2007). <https://doi.org/10.2174/092986707780597899>
- 7 Steffen-Munsberg, F. *et al.* Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. *Biotechnol Adv* **33**, 566-604 (2015). <https://doi.org/10.1016/j.biotechadv.2014.12.012>
- 8 Dunathan, H. C. Conformation and Reaction Specificity in Pyridoxal Phosphate Enzymes. *P Natl Acad Sci USA* **55**, 712-& (1966). <https://doi.org/10.1073/pnas.55.4.712>
- 9 Richard, J. P., Amyes, T. L., Crujeiras, J. & Rios, A. Pyridoxal 5'-phosphate: electrophilic catalyst extraordinaire. *Current Opinion in Chemical Biology* **13**, 475-483 (2009). <https://doi.org/10.1016/j.cbpa.2009.06.023>
- 10 Toney, M. D. Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochim Biophys Acta* **1814**, 1407-1418 (2011). <https://doi.org/10.1016/j.bbapap.2011.05.019>
- 11 Schirch, L. Serine Hydroxymethyltransferase. *Adv Enzymol Ramb* **53**, 83-112 (1982).
- 12 Ulevitch, R. J. & Kallen, R. G. Studies of Reactions of Lamb Liver Serine Hydroxymethylase with L-Phenylalanine - Kinetic Isotope Effects Upon Quinonoid Intermediate Formation. *Biochemistry* **16**, 5350-5354 (1977). <https://doi.org/10.1021/bi00643a028>
- 13 Thomas, N. R., Schirch, V. & Gani, D. Synthesis of (2r)-[1-C-13]-2-Amino-2-Methylmalonic and (2s)-[1-C-13]-2-Amino-2-Methylmalonic Acid, Probes for the Serine Hydroxymethyltransferase Reaction - Stereospecific Decarboxylation of the 2-Pro-R Carboxy Group with the Retention of Configuration. *J Chem Soc Chem Comm*, 400-402 (1990). <https://doi.org/10.1039/c39900000400>
- 14 Shostak, K. & Schirch, V. Serine hydroxymethyltransferase: mechanism of the racemization and transamination of D- and L-alanine. *Biochemistry* **27**, 8007-8014 (1988). <https://doi.org/10.1021/bi00421a006>
- 15 Herbig, K. *et al.* Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. *Journal of Biological Chemistry* **277**, 38381-38389 (2002). <https://doi.org/10.1074/jbc.M205000200>

- 16 Xie, M. & Pei, D. S. Serine hydroxymethyltransferase 2: a novel target for human cancer therapy. *Invest New Drug* **39**, 1671-1681 (2021). <https://doi.org:10.1007/s10637-021-01144-z>
- 17 Florio, R., di Salvo, M. L., Vivoli, M. & Contestabile, R. Serine hydroxymethyltransferase: a model enzyme for mechanistic, structural, and evolutionary studies. *Biochim Biophys Acta* **1814**, 1489-1496 (2011). <https://doi.org:10.1016/j.bbapap.2010.10.010>
- 18 Giardina, G. *et al.* How pyridoxal 5'-phosphate differentially regulates human cytosolic and mitochondrial serine hydroxymethyltransferase oligomeric state. *FEBS J* **282**, 1225-1241 (2015). <https://doi.org:10.1111/febs.13211>
- 19 Tramonti, A. *et al.* Human Cytosolic and Mitochondrial Serine Hydroxymethyltransferase Isoforms in Comparison: Full Kinetic Characterization and Substrate Inhibition Properties. *Biochemistry* **57**, 6984-6996 (2018). <https://doi.org:10.1021/acs.biochem.8b01074>
- 20 Garrow, T. A. *et al.* Cloning of Human Cdnas Encoding Mitochondrial and Cytosolic Serine Hydroxymethyltransferases and Chromosomal Localization. *Journal of Biological Chemistry* **268**, 11910-11916 (1993).
- 21 Ducker, G. S. & Rabinowitz, J. D. One-Carbon Metabolism in Health and Disease. *Cell Metab* **25**, 27-42 (2017). <https://doi.org:10.1016/j.cmet.2016.08.009>
- 22 Lee, G. Y. *et al.* Comparative oncogenomics identifies PSMB4 and SHMT2 as potential cancer driver genes. *Cancer Res* **74**, 3114-3126 (2014). <https://doi.org:10.1158/0008-5472.CAN-13-2683>
- 23 Giardina, G. *et al.* The catalytic activity of serine hydroxymethyltransferase is essential for de novo nuclear dTMP synthesis in lung cancer cells. *FEBS J* **285**, 3238-3253 (2018). <https://doi.org:10.1111/febs.14610>
- 24 Li, A. M. *et al.* Metabolic Profiling Reveals a Dependency of Human Metastatic Breast Cancer on Mitochondrial Serine and One-Carbon Unit Metabolism. *Mol Cancer Res* **18**, 599-611 (2020). <https://doi.org:10.1158/1541-7786.Mcr-19-0606>
- 25 Wu, Z. Z. *et al.* Increased Expression of SHMT2 Is Associated With Poor Prognosis and Advanced Pathological Grade in Oral Squamous Cell Carcinoma. *Front Oncol* **10** (2020). <https://doi.org:ARTN58853010.3389/fonc.2020.588530>
- 26 Yang, C. C. *et al.* Folate-mediated one-carbon metabolism: a targeting strategy in cancer therapy. *Drug Discov Today* **26**, 817-825 (2021). <https://doi.org:10.1016/j.drudis.2020.12.006>
- 27 Cui, X. M. *et al.* SHMT2 Drives the Progression of Colorectal Cancer by Regulating UHRF1 Expression. *Can J Gastroenterol* **2022** (2022). <https://doi.org:Artn375869710.1155/2022/3758697>
- 28 Zeng, Y. *et al.* Roles of Mitochondrial Serine Hydroxymethyltransferase 2 (SHMT2) in Human Carcinogenesis. *J Cancer* **12**, 5888-5894 (2021). <https://doi.org:10.7150/jca.60170>
- 29 Du, J. *et al.* Serine hydroxymethyltransferase 2 predicts unfavorable outcomes in multiple cancer: a systematic review and meta-analysis. *Transl Cancer Res* **11**, 444-455 (2022). <https://doi.org:10.21037/tcr-21-2485>
- 30 Pranzini, E. *et al.* SHMT2-mediated mitochondrial serine metabolism drives 5-FU resistance by fueling nucleotide biosynthesis. *Cell Rep* **40**, 111233 (2022). <https://doi.org:10.1016/j.celrep.2022.111233>
- 31 Clark, R. A., Qiao, J., Jacobson, J. C. & Chung, D. H. Induction of serine hydroxymethyltransferase 2 promotes tumorigenesis and metastasis in neuroblastoma. *Oncotarget* **13**, 32-45 (2022). <https://doi.org:10.18632/oncotarget.28168>

- 32 Cuthbertson, C. R., Arabzada, Z., Bankhead, A., Kyani, A. & Neamati, N. A Review of Small-Molecule Inhibitors of One-Carbon Enzymes: SHMT2 and MTHFD2 in the Spotlight. *Acs Pharmacol Transl* **4**, 624-646 (2021). <https://doi.org/10.1021/acspsci.0c00223>
- 33 Stine, Z. E., Schug, Z. T., Salvino, J. M. & Dang, C. V. Targeting cancer metabolism in the era of precision oncology. *Nat Rev Drug Discov* **21**, 141-162 (2022). <https://doi.org/10.1038/s41573-021-00339-6>
- 34 Han, Y. *et al.* Identification of three new compounds that directly target human serine hydroxymethyltransferase 2. *Chem Biol Drug Des* **97**, 221-230 (2021). <https://doi.org/10.1111/cbdd.13774>
- 35 Garcia-Canaveras, J. C. *et al.* SHMT inhibition is effective and synergizes with methotrexate in T-cell acute lymphoblastic leukemia. *Leukemia* **35**, 377-388 (2021). <https://doi.org/10.1038/s41375-020-0845-6>
- 36 Pikman, Y. *et al.* Targeting serine hydroxymethyltransferases 1 and 2 for T-cell acute lymphoblastic leukemia therapy. *Leukemia* **36**, 348-360 (2022). <https://doi.org/10.1038/s41375-021-01361-8>
- 37 Ducker, G. S. *et al.* Human SHMT inhibitors reveal defective glycine import as a targetable metabolic vulnerability of diffuse large B-cell lymphoma. *P Natl Acad Sci USA* **114**, 11404-11409 (2017). <https://doi.org/10.1073/pnas.1706617114>
- 38 McConnell, D. B. Biotin's Lessons in Drug Design. *J Med Chem* **64**, 16319-16327 (2021). <https://doi.org/10.1021/acs.jmedchem.1c00975>
- 39 Engler, N., Ostermann, A., Niimura, N. & Parak, F. G. Hydrogen atoms in proteins: positions and dynamics. *Proc Natl Acad Sci U S A* **100**, 10243-10248 (2003). <https://doi.org/10.1073/pnas.1834279100>
- 40 Bax, B., Chung, C. & Edge, C. Getting the chemistry right: protonation, tautomers and the importance of H atoms in biological chemistry. *Acta Crystallographica Section D* **73**, 131-140 (2017). <https://doi.org/doi:10.1107/S2059798316020283>
- 41 Gardberg, A. S. *et al.* Unambiguous determination of H-atom positions: comparing results from neutron and high-resolution X-ray crystallography. *Acta Crystallographica Section D-Structural Biology* **66**, 558-567 (2010). <https://doi.org/10.1107/S0907444910005494>
- 42 Niimura, N., Podjarny, A. & Crystallography, I. U. o. *Neutron Protein Crystallography: Hydrogen, Protons, and Hydration in Bio-macromolecules*. (OUP Oxford, 2011).
- 43 Gerlits, O. *et al.* Long-Range Electrostatics-Induced Two-Proton Transfer Captured by Neutron Crystallography in an Enzyme Catalytic Site. *Angew Chem Int Ed Engl* **55**, 4924-4927 (2016). <https://doi.org/10.1002/anie.201509989>
- 44 Dajnowicz, S. *et al.* Direct visualization of critical hydrogen atoms in a pyridoxal 5'-phosphate enzyme. *Nat Commun* **8**, 955 (2017). <https://doi.org/10.1038/s41467-017-01060-y>
- 45 Drago, V. N. *et al.* An N...H...N low-barrier hydrogen bond preorganizes the catalytic site of aspartate aminotransferase to facilitate the second half-reaction. *Chem Sci* **13**, 10057-10065 (2022). <https://doi.org/10.1039/d2sc02285k>
- 46 Fisher, S. Z., Aggarwal, M., Kovalevsky, A. Y., Silverman, D. N. & McKenna, R. Neutron diffraction of acetazolamide-bound human carbonic anhydrase II reveals atomic details of drug binding. *J Am Chem Soc* **134**, 14726-14729 (2012). <https://doi.org/10.1021/ja3068098>

- 47 Weber, I. T. *et al.* Joint X-ray/neutron crystallographic study of HIV-1 protease with clinical inhibitor amprenavir: insights for drug design. *J Med Chem* **56**, 5631-5635 (2013). <https://doi.org:10.1021/jm400684f>
- 48 Manzoni, F. *et al.* Elucidation of Hydrogen Bonding Patterns in Ligand-Free, Lactose- and Glycerol-Bound Galectin-3C by Neutron Crystallography to Guide Drug Design. *J Med Chem* **61**, 4412-4420 (2018). <https://doi.org:10.1021/acs.jmedchem.8b00081>
- 49 Kneller, D. W. *et al.* Structural, Electronic, and Electrostatic Determinants for Inhibitor Binding to Subsites S1 and S2 in SARS-CoV-2 Main Protease. *J Med Chem* **64**, 17366-17383 (2021). <https://doi.org:10.1021/acs.jmedchem.1c01475>
- 50 Kneller, D. *et al.* Covalent nartlaprevir- and boceprevir-derived hybrid inhibitors of SARS-CoV-2 main protease: room-temperature X-ray and neutron crystallography, binding thermodynamics, and antiviral activity. *Res Sq* (2022). <https://doi.org:10.21203/rs.3.rs-1318037/v1>
- 51 Gerlits, O. *et al.* Zooming in on protons: Neutron structure of protein kinase A trapped in a product complex. *Sci Adv* **5**, eaav0482 (2019). <https://doi.org:10.1126/sciadv.aav0482>
- 52 Kumar, M. *et al.* Visualizing Tetrahedral Oxyanion Bound in HIV-1 Protease Using Neutrons: Implications for the Catalytic Mechanism and Drug Design. *ACS Omega* **5**, 11605-11617 (2020). <https://doi.org:10.1021/acsomega.0c00835>
- 53 Renwick, S. B., Snell, K. & Baumann, U. The crystal structure of human cytosolic serine hydroxymethyltransferase: a target for cancer chemotherapy. *Structure* **6**, 1105-1116 (1998). [https://doi.org:10.1016/S0969-2126\(98\)00112-9](https://doi.org:10.1016/S0969-2126(98)00112-9)
- 54 Trivedi, V. *et al.* Crystal structure of binary and ternary complexes of serine hydroxymethyltransferase from *Bacillus stearothermophilus* - Insights into the catalytic mechanism. *Journal of Biological Chemistry* **277**, 17161-17169 (2002). <https://doi.org:10.1074/jbc.M111976200>
- 55 Szebenyi, D. M. E., Musayev, F. N., di Salvo, M. L., Safo, M. K. & Schirch, V. Serine hydroxymethyltransferase: Role of Glu75 and evidence that serine is cleaved by a retroaldol mechanism. *Biochemistry* **43**, 6865-6876 (2004). <https://doi.org:10.1021/bi049791y>
- 56 Schirch, V. & Szebenyi, D. M. Serine hydroxymethyltransferase revisited. *Curr Opin Chem Biol* **9**, 482-487 (2005). <https://doi.org:10.1016/j.cbpa.2005.08.017>
- 57 Jagath, J. R. *et al.* Importance of the amino terminus in maintenance of oligomeric structure of sheep liver cytosolic serine hydroxymethyltransferase. *Eur J Biochem* **247**, 372-379 (1997). <https://doi.org:10.1111/j.1432-1033.1997.00372.x>
- 58 Rao, N. A., Ambili, M., Jala, V. R., Subramanya, H. S. & Savithri, H. S. Structure-function relationship in serine hydroxymethyltransferase. *Bba-Proteins Proteom* **1647**, 24-29 (2003). [https://doi.org:10.1016/S1570-9639\(03\)00043-8](https://doi.org:10.1016/S1570-9639(03)00043-8)
- 59 Matthews, R. G. & Drummond, J. T. Providing One-Carbon Units for Biological Methylations - Mechanistic Studies on Serine Hydroxymethyltransferase, Methylenetetrahydrofolate Reductase, and Methyltetrahydrofolate-Homocysteine Methyltransferase. *Chem Rev* **90**, 1275-1290 (1990). <https://doi.org:10.1021/cr00105a010>
- 60 Fernandes, H. S., Ramos, M. J. & Cerqueira, N. M. F. S. A. Catalytic Mechanism of the Serine Hydroxymethyltransferase: A Computational ONIOM QM/MM Study. *ACS Catal* **8**, 10096-10110 (2018). <https://doi.org:10.1021/acscatal.8b02321>

- 61 Nonaka, H. *et al.* Design strategy for serine hydroxymethyltransferase probes based on retro-aldol-type reaction. *Nat Commun* **10**, 876 (2019). <https://doi.org/10.1038/s41467-019-08833-7>
- 62 Scaletti, E., Jemth, A. S., Helleday, T. & Stenmark, P. Structural basis of inhibition of the human serine hydroxymethyltransferase SHMT2 by antifolate drugs. *Febs Lett* **593**, 1863-1873 (2019). <https://doi.org/10.1002/1873-3468.13455>
- 63 Ota, T. *et al.* Structural basis for selective inhibition of human serine hydroxymethyltransferase by secondary bile acid conjugate. *iScience* **24**, 102036 (2021). <https://doi.org/10.1016/j.isci.2021.102036>
- 64 Blakeley, M. P. & Podjarny, A. D. Neutron macromolecular crystallography. *Emerg Top Life Sci* **2**, 39-55 (2018). <https://doi.org/10.1042/Etls20170083>
- 65 Schneider, G., Kack, H. & Lindqvist, Y. The manifold of vitamin B6 dependent enzymes. *Structure* **8**, R1-6 (2000). [https://doi.org/10.1016/s0969-2126\(00\)00085-x](https://doi.org/10.1016/s0969-2126(00)00085-x)
- 66 Scarsdale, J. N., Radaev, S., Kazanina, G., Schirch, V. & Wright, H. T. Crystal structure at 2.4 Å resolution of E. coli serine hydroxymethyltransferase in complex with glycine substrate and 5-formyl tetrahydrofolate<sup>11</sup>Edited by I. A. Wilson. *Journal of Molecular Biology* **296**, 155-168 (2000). [https://doi.org:https://doi.org/10.1006/jmbi.1999.3453](https://doi.org/https://doi.org/10.1006/jmbi.1999.3453)
- 67 Szebenyi, D. M. E., Liu, X. W., Kriksunov, I. A., Stover, P. J. & Thiel, D. J. Structure of a murine cytoplasmic serine hydroxymethyltransferase quinonoid ternary complex: Evidence for asymmetric obligate dimers. *Biochemistry* **39**, 13313-13323 (2000). <https://doi.org/10.1021/bi000635a>
- 68 Casasnovas, R., Salvà, A., Frau, J., Donoso, J. & Muñoz, F. Theoretical study on the distribution of atomic charges in the Schiff bases of 3-hydroxypyridine-4-aldehyde and alanine. The effect of the protonation state of the pyridine and imine nitrogen atoms. *Chemical Physics* **355**, 149-156 (2009).
- 69 Griswold, W. R. & Toney, M. D. Role of the pyridine nitrogen in pyridoxal 5'-phosphate catalysis: activity of three classes of PLP enzymes reconstituted with deazapyridoxal 5'-phosphate. *J Am Chem Soc* **133**, 14823-14830 (2011). <https://doi.org/10.1021/ja2061006>
- 70 Mueser, T. C., Drago, V., Kovalevsky, A. & Dajnowicz, S. Pyridoxal 5'-phosphate dependent reactions: Analyzing the mechanism of aspartate aminotransferase. *Methods Enzymol* **634**, 333-359 (2020). <https://doi.org/10.1016/bs.mie.2020.01.009>
- 71 Casasnovas, R. *et al.* C-H Activation in Pyridoxal-5'-phosphate Schiff Bases: The Role of the Imine Nitrogen. A Combined Experimental and Computational Study. *J Phys Chem B* **116**, 10665-10675 (2012). <https://doi.org/10.1021/jp303678n>
- 72 Rajaram, V. *et al.* Structure determination and biochemical studies on Bacillus stearothermophilus E53Q serine hydroxymethyltransferase and its complexes provide insights on function and enzyme memory. *FEBS J* **274**, 4148-4160 (2007). <https://doi.org/10.1111/j.1742-4658.2007.05943.x>
- 73 Shcheynikov, N. *et al.* Intracellular Cl<sup>-</sup> as a signaling ion that potently regulates Na<sup>+</sup>/HCO<sub>3</sub><sup>-</sup> transporters. *P Natl Acad Sci USA* **112**, E329-E337 (2015). <https://doi.org/10.1073/pnas.1415673112>
- 74 Jahn, S. C., Rowland-Faux, L., Stacpoole, P. W. & James, M. O. Chloride concentrations in human hepatic cytosol and mitochondria are a function of age. *Biochem Biophys Res Commun* **459**, 463-468 (2015). <https://doi.org/10.1016/j.bbrc.2015.02.128>

- 75 Eliot, A. C. & Kirsch, J. F. Modulation of the internal aldimine pK(a)'s of 1-aminocyclopropane-1-carboxylate synthase and aspartate aminotransferase by specific active site residues. *Biochemistry* **41**, 3836-3842 (2002). <https://doi.org:10.1021/bi016084l>
- 76 Ruzskowski, M. *et al.* Structural basis of methotrexate and pemetrexed action on serine hydroxymethyltransferases revealed using plant models. *Sci Rep-Uk* **9** (2019). <https://doi.org:ARTN1961410.1038/s41598-019-56043-4>
- 77 Makino, Y. *et al.* Serine hydroxymethyltransferase as a potential target of antibacterial agents acting synergistically with one-carbon metabolism-related inhibitors. *Commun Biol* **5**, 619 (2022). <https://doi.org:10.1038/s42003-022-03555-x>
- 78 Chitnumsub, P. *et al.* Structures of Plasmodium vivax serine hydroxymethyltransferase: implications for ligand-binding specificity and functional control. *Acta Crystallogr D* **70**, 3177-3186 (2014). <https://doi.org:10.1107/S1399004714023128>
- 79 Ruzskowski, M., Sekula, B., Ruzskowska, A. & Dauter, Z. Chloroplastic Serine Hydroxymethyltransferase From Medicago truncatula: A Structural Characterization. *Front Plant Sci* **9** (2018). <https://doi.org:ARTN58410.3389/fpls.2018.00584>
- 80 Gerlits, O. O., Coates, L., Woods, R. J. & Kovalevsky, A. Mannobiose Binding Induces Changes in Hydrogen Bonding and Protonation States of Acidic Residues in Concanavalin A As Revealed by Neutron Crystallography. *Biochemistry* **56**, 4747-4750 (2017). <https://doi.org:10.1021/acs.biochem.7b00654>
- 81 Gerlits, O. *et al.* Room Temperature Neutron Crystallography of Drug Resistant HIV-1 Protease Uncovers Limitations of X-ray Structural Analysis at 100 K. *J Med Chem* **60**, 2018-2025 (2017). <https://doi.org:10.1021/acs.jmedchem.6b01767>
- 82 Contestabile, R. *et al.* Role of tyrosine 65 in the mechanism of serine hydroxymethyltransferase. *Biochemistry* **39**, 7492-7500 (2000). <https://doi.org:10.1021/bi000032z>
- 83 Chiba, Y. *et al.* Mechanism for folate-independent aldolase reaction catalyzed by serine hydroxymethyltransferase. *Febs Journal* **279**, 504-514 (2012). <https://doi.org:10.1111/j.1742-4658.2011.08443.x>
- 84 Soniya, K. & Chandra, A. Free Energy Landscape and Proton Transfer Pathways of the Transamination Reaction at the Active site of the Serine Hydroxymethyltransferase Enzyme in Aqueous Medium. *J Phys Chem B* **125**, 11848-11856 (2021). <https://doi.org:10.1021/acs.jpccb.1c05864>
- 85 Di Salvo, M. L. *et al.* Structure-based mechanism for early PLP-mediated steps of rabbit cytosolic serine hydroxymethyltransferase reaction. *Biomed Res Int* **2013**, 458571 (2013). <https://doi.org:10.1155/2013/458571>
- 86 Talwar, R., Jagath, J. R., Rao, N. A. & Savithri, H. S. His230 of serine hydroxymethyltransferase facilitates the proton abstraction step in catalysis. *Eur J Biochem* **267**, 1441-1446 (2000). <https://doi.org:10.1046/j.1432-1327.2000.01142.x>
- 87 Rao, J. V. K., Prakash, V., Rao, N. A. & Savithri, H. S. The role of Glu74 and Tyr82 in the reaction catalyzed by sheep liver cytosolic serine hydroxymethyltransferase. *Eur J Biochem* **267**, 5967-5976 (2000). <https://doi.org:10.1046/j.1432-1327.2000.01667.x>
- 88 Blakeley, M. P. *et al.* Neutron macromolecular crystallography with LADI-III. *Acta Crystallogr D Biol Crystallogr* **66**, 1198-1205 (2010). <https://doi.org:10.1107/S0907444910019797>

- 89 Meilleur, F. *et al.* The IMAGINE instrument: first neutron protein structure and new capabilities for neutron macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **69**, 2157-2160 (2013). <https://doi.org/10.1107/S0907444913019604>
- 90 Meilleur, F., Coates, L., Cuneo, M. J., Kovalevsky, A. & Myles, D. A. A. The Neutron Macromolecular Crystallography Instruments at Oak Ridge National Laboratory: Advances, Challenges, and Opportunities. *Crystals* **8** (2018). <https://doi.org/ARTN38810.3390/cryst8100388>
- 91 Schroder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: neutrons reveal enzyme chemistry. *Acta Crystallogr D Struct Biol* **74**, 778-786 (2018). <https://doi.org/10.1107/S2059798318001626>
- 92 Meilleur, F., Kovalevsky, A. & Myles, D. A. A. IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. *Methods Enzymol* **634**, 69-85 (2020). <https://doi.org/10.1016/bs.mie.2019.11.016>
- 93 Campbell, J. W. Lauegen, an X-Windows-Based Program for the Processing of Laue X-Ray-Diffraction Data. *J Appl Crystallogr* **28**, 228-236 (1995). <https://doi.org/10.1107/S002188989400991x>
- 94 Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. LAUEGEN version 6.0 and INTLDM. *J Appl Crystallogr* **31**, 496-502 (1998). <https://doi.org/10.1107/S0021889897016683>
- 95 Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. LSCALE - the new normalization, scaling and absorption correction program in the Daresbury Laue software suite. *J Appl Crystallogr* **32**, 554-562 (1999). <https://doi.org/10.1107/S0021889898015350>
- 96 Coates, L. *et al.* The Macromolecular Neutron Diffractometer MaNDi at the Spallation Neutron Source. *J Appl Crystallogr* **48**, 1302-1306 (2015). <https://doi.org/10.1107/S1600576715011243>
- 97 Coates, L. & Sullivan, B. The macromolecular neutron diffractometer at the spallation neutron source. *Method Enzymol* **634**, 87-99 (2020). <https://doi.org/10.1016/bs.mie.2019.11.020>
- 98 Arnold, O. *et al.* Mantid—Data analysis and visualization package for neutron scattering and  $\mu$  SR experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **764**, 156-166 (2014). <https://doi.org/10.1016/j.nima.2014.07.029>
- 99 Sullivan, B. *et al.* Improving the accuracy and resolution of neutron crystallographic data by three-dimensional profile fitting of Bragg peaks in reciprocal space. *Acta Crystallographica Section D-Structural Biology* **74**, 1085-1095 (2018). <https://doi.org/10.1107/S2059798318013347>
- 100 Helliwell, J. R. *et al.* The Recording and Analysis of Synchrotron X-Radiation Laue Diffraction Photographs. *J Appl Crystallogr* **22**, 483-497 (1989). <https://doi.org/10.1107/S0021889889006564>
- 101 Weiss, M. S. Global indicators of X-ray data quality. *J Appl Crystallogr* **34**, 130-135 (2001). <https://doi.org/10.1107/S0021889800018227>
- 102 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214 (2013). <https://doi.org/10.1107/s0907444913000061>

- 103 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Structural Biology* **67**, 235-242 (2011). <https://doi.org:10.1107/S0907444910045749>
- 104 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007). <https://doi.org:10.1107/S0021889807021206>
- 105 Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* **62**, 859-866 (2006). <https://doi.org:10.1107/S0907444906019949>
- 106 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). <https://doi.org:10.1107/S0907444909052925>
- 107 Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D-Structural Biology* **75**, 861-877 (2019). <https://doi.org:10.1107/S2059798319011471>
- 108 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Structural Biology* **60**, 2126-2132 (2004). <https://doi.org:10.1107/S0907444904019158>
- 109 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010). <https://doi.org:10.1107/S0907444910007493>
- 110 Casanal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci* **29**, 1069-1078 (2020). <https://doi.org:10.1002/pro.3791>
- 111 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010). <https://doi.org:10.1107/S0907444909042073>
- 112 Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. Electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074-1080 (2009). <https://doi.org:10.1107/S0907444909029436>
- 113 Gaussian 16 Rev. A.03 (Gaussian Inc., Wallingford, CT, 2016).
- 114 Mustyakimov, M. & Langan, P. nCNS: an open source distribution patch for CNS for macromolecular structure refinement. *Los Alamos National Security, Los Alamos, NM, USA* (2007).
- 115 Adams, P. D., Mustyakimov, M., Afonine, P. V. & Langan, P. Generalized X-ray and neutron crystallographic analysis: more accurate and complete structures for biological macromolecules. *Acta Crystallogr D Biol Crystallogr* **65**, 567-573 (2009). <https://doi.org:10.1107/S0907444909011548>
- 116 Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**, 905-921 (1998). <https://doi.org:10.1107/S0907444998003254>
- 117 Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* **7**, 3297-3305 (2005). <https://doi.org:10.1039/b508541a>

118 Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys Chem Chem Phys* **8**, 1057-1065 (2006). <https://doi.org/10.1039/b515623h>

## Figure legends.

**Figure 1. SHMT catalyzed reaction, enzyme 3D structure and sequence alignment of *Tth*SHMT and hSHMT2.** **a)** Chemical structures of vitamin B<sub>6</sub>, PLP, and the internal aldimine. Atom labels are given on the internal aldimine structure. **b)** THF-dependent conversion of L-Ser to Gly catalyzed by SHMT. **c)** Overview of the overall fold of SHMT proteins from human mitochondria (hSHMT2) and *Thermus thermophilus* (*Tth*SHMT). Individual protomers are depicted in different color schemes and PLP cofactors are shown with CPK representation. hSHMT2 exists as a homotetramer made up of obligate dimers, whereas *Tth*SHMT is a homodimer. **d)** *Tth*SHMT and hSHMT2 sequence alignment. Active sites are conserved between the two enzymes, highlighted in red. The sequence identity is 41%.

**Figure 2. Neutron scattering length density maps for *Tth*SHMT active site residues for the joint XN internal aldimine structure.** The  $2F_O-F_C$  neutron scattering length density contoured at  $1\sigma$  level is depicted by wheat mesh. The omit  $F_O-F_C$  difference neutron scattering length density map contoured at  $2\sigma$  level is represented by magenta mesh. D atoms are shown in green, and H atoms are white. **a)** N1 of the PLP cofactor is protonated (positively charged), but Schiff base N<sub>SB</sub> is not (neutral). An H-bonding network with Asn98, His125, and Ala199 fixes the position of deprotonated Asp197 (negatively charged) that forms an H-bond with N1. Above PLP, the side chain of Ala199 further stabilizes PLP position through C-H $\cdots\pi$  interactions. **b)** O3' of PLP is deprotonated with a partial negative charge stabilized through an unconventional C-H $\cdots$ O bond with His200. A network of H-bonds made by neutral His200 and His312 connect O3' to the main chain amide ND of His312. **c)** A sulfate ion accommodates the amino acid substrate binding site, and its position is stabilized through several H-bonds. **d)** The pyridine ring of PLP is sandwiched between His122 and Ala199 making  $\pi$ - $\pi$  stacking and C-H $\cdots\pi$  interactions. His122, positioned on the *re* face of the cofactor, is neutral and its imidazole group accepts D atoms from the side chain and amide backbone of Thr124 in a bifurcated H-bond. **e)** Phosphate group of PLP binds in the phosphate binding pocket made up of the side chains of Tyr51\* and Ser95, the amide backbone of Gly94, Ser95, and Gly258, and three crystallographic waters.

**Figure 3. Binding of L-Ser-d<sub>7</sub> to the peripheral binding site within the funnel-shaped active site cavity as observed in the joint XN structure of *Tth*SHMT/L-Ser *pre*-Michaelis complex.** **a)** L-Ser-d<sub>7</sub> is found at the entrance to the active site cavity in front of the sulfate ion. **b)**  $\alpha$ -Amino group of L-Ser-d<sub>7</sub> faces a hydrophobic patch (Leu117 and Leu123) and its carboxyl group is directed towards the bulk solvent. The side chain hydroxyl of L-Ser-d<sub>7</sub> makes an H-bond with Tyr61\*. **c)** The  $2F_O-F_C$  neutron scattering length density in the L-Ser-d<sub>7</sub> binding site contoured at  $1\sigma$  level is depicted by wheat mesh. In addition to an H-bond with Tyr61\*, the substrate has a water-mediated interaction with Glu53\*.

**Figure 4. Binding of D-Ser to the substrate binding site as observed in the room-temperature X-ray structure of *Tth*SHMT/D-Ser *pseudo*-Michaelis complex.** The  $2F_O-F_C$  electron density contoured at  $1\sigma$  level is depicted by gray mesh. **a)** D-Ser interacts makes numerous H-bonds with the active site residues. Tyr61\* adopts two alternate conformations which make different substrate interactions depicted in panel **b**. The D-Ser  $\alpha$ -amino group is oriented away from C4', presumably preventing the formation of the external aldimine. **b)** Zoomed-in view of Tyr61\* interactions with D-Ser.

**Figure 5. Proposed L-Ser substrate pathway through the *Tth*SHMT active site cavity and the mechanism of the external aldimine formation.** **a)** Overview of substrate tracking into the active site of *Tth*SHMT revealed by X-ray and joint X-ray/neutron crystallography. The composite image was made from superimposing the *Tth*SHMT/L-Ser (RMS = 0.081) neutron structures and the *Tth*SHMT/D-Ser (RMSD= 0.122) room-temperature X-ray structure with the substrate-free *Tth*SHMT. **b)** Proposed mechanism for the generation of the external aldimine based on the current joint X-ray/neutron structures. Glu53 in *Tth*SHMT is proposed to act as the general base to abstract the  $\beta$ -hydroxyl of L-Ser to subsequently form formaldehyde.