

Role of a Conserved Active Site Cation– π Interaction in *Escherichia coli* Serine Hydroxymethyltransferase[†]

Mirella Vivoli,^{†,||} Francesco Angelucci,^{†,||} Andrea Ilari,[§] Veronica Morea,[§] Sebastiana Angelaccio,[‡]
Martino Luigi di Salvo,^{*,‡} and Roberto Contestabile^{*,‡}

[‡]Dipartimento di Scienze Biochimiche and Istituto Pasteur, Fondazione Cenci Bolognetti, Sapienza Università di Roma, Piazzale Aldo Moro, 5-00185 Roma, Italy and [§]Istituto di Biologia e Patologia Molecolari, Consiglio Nazionale delle Ricerche, Piazzale Aldo Moro, 5-00185 Roma, Italy. ^{||}These authors contributed equally to this work.

Received September 8, 2009; Revised Manuscript Received November 2, 2009

ABSTRACT: Serine hydroxymethyltransferase is a pyridoxal 5'-phosphate-dependent enzyme that catalyzes the interconversion of serine and glycine using tetrahydropteroylglutamate as the one-carbon carrier. In all pyridoxal phosphate-dependent enzymes, amino acid substrates are bound and released through a transaldimination process, in which an internal aldimine and an external aldimine are interconverted via gem-diamine intermediates. Bioinformatic analyses of serine hydroxymethyltransferase sequences and structures showed the presence of two highly conserved residues, a tyrosine and an arginine, engaged in a cation– π interaction. In *Escherichia coli* serine hydroxymethyltransferase, the hydroxyl group of this conserved tyrosine (Tyr55) is located in a position compatible with a role as hydrogen exchanger in the transaldimination reaction. Because of the location of Tyr55 at the active site, the enhancement of its acidic properties caused by the cation– π interaction with Arg235, and the hydrogen bonds established by its hydroxyl group, a role of this residue as acid–base catalyst in the transaldimination process was envisaged. The role played by this cation– π interaction in the *E. coli* serine hydroxymethyltransferase was investigated by crystallography and site-directed mutagenesis using Y55F and three R235 mutant forms. The crystal structure of the Y55F mutant suggests that the presence of Tyr55 is indispensable for a correct positioning of the cofactor and for the maintenance of the structure of several loops involved in substrate and cofactor binding. The kinetic properties of all mutant enzymes are profoundly altered. Substrate binding and rapid kinetic experiments showed that both Y55 and R235 are required for a correct progress of the transaldimination reaction.

Serine hydroxymethyltransferase (SHMT;¹ EC 2.1.2.1) is a pyridoxal 5'-phosphate (PLP) dependent enzyme that catalyzes the reversible transfer of the C β of serine to tetrahydropteroylglutamate (H₄PteGlu), with formation of glycine and 5,10-methylene-H₄PteGlu (*I*). SHMT is a ubiquitous enzyme because of its essential function as a major source of one-carbon groups, which are required for purine, thymidylate, and methionine biosynthesis, as well as for many other methylation reactions.

Increased activity of this enzyme coupled with enhanced DNA synthesis in neoplastic tissues has suggested SHMT to be a suitable target for cancer chemotherapy (2). The reaction catalyzed by SHMT is of great interest since it represents a link between amino acid and folate metabolisms.

Several hundred nonidentical SHMT primary structures currently appear in sequence data banks, and crystal structures of the enzyme from various sources have been solved (2–6). In SHMTs, monomers are organized into obligate homodimers (in eukaryotes the enzyme occurs as a dimer of dimers), and the active site is located at the interface region. The extensive similarity of amino acid sequences and conservation of functional residues indicate that all SHMTs are closely related both in structure and in catalytic mechanism (7, 8). In the past years, many active site residues have been changed by site-directed mutagenesis, and their role in binding and catalysis has been discussed (9, 10). Several reaction mechanisms have been proposed to explain the complex operation mode of SHMT. The currently accepted mechanism consists in a modified folate-dependent retroaldol cleavage via direct nucleophilic attack of N⁵ of H₄PteGlu to C β of serine, which results in the elimination of the quinonoid intermediate (11). On the other hand, little is known on the earliest steps of the reaction: the transaldimination process. This is a common passage in all reactions catalyzed by PLP-dependent enzymes, essential for substrate binding and

[†]This work was supported by grants of the Italian Ministero dell'Università e della Ricerca, Istituto Pasteur, Fondazione Cenci Bolognetti, and European Community, Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme (through the "Integrated Infrastructure Initiative" Integrating Activity on Synchrotron and Free Electron Laser Science, Contract R II 3-CT-2004-506008). F.A. was supported by FIRB/Proteomica 2007-protRBRN07BMCT.

*To whom correspondence should be addressed. Tel: +39 06 49917569. Fax: +39 06 49917566. E-mail: roberto.contestabile@uniroma1.it; martino.disalvo@uniroma1.it.

¹Abbreviations: SHMT, serine hydroxymethyltransferase; eSHMT, *Escherichia coli* SHMT; bsSHMT, *Bacillus stearothermophilus* SHMT; PLP, pyridoxal 5'-phosphate; H₄PteGlu, tetrahydropteroylglutamate; $K_m(\text{Ser})$, K_m for L-serine; $K_m(\text{Gly})$, K_m for glycine; $K_m(\text{THF})$, K_m for H₄PteGlu; $K_d(\text{PLP})$, dissociation constant for PLP binding equilibrium; $K_d(\text{Ser})$, dissociation constant for L-serine binding equilibrium; $K_d(\text{Gly})$, dissociation constant for glycine binding equilibrium; $k_{\text{obs}}(\text{Ser})$, observed rate constant for L-serine binding; $k_{\text{obs}}(\text{Gly})$, observed rate constant for glycine binding.

product release. In the transaldimination reaction, the incoming amino acid substrate displaces the active site lysine residue from the cofactor Schiff base linkage via the formation of the *gem*-diamine I intermediate (Scheme 1). The reaction then implies the transfer of a proton from the substrate amino group to the amino group of the leaving lysine, with the formation of *gem*-diamine II.

In the crystal structure of *Escherichia coli* SHMT (*e*SHMT) in complex with glycine and 5-formyl- H_4 PteGlu (PDB ID 1DFO (3)), Tyr55' (the prime indicates that the residue is contributed by the symmetry-related monomer, with respect to the PLP-binding subunit) points toward the protein region where the transaldimination reaction takes place, and its phenolic oxygen is located within hydrogen-bonding distance from the nitrogen atom of glycine (Figure 1). This led us to hypothesize that the Tyr55' hydroxyl group might play a role as general acid–base catalyst in the proton transfers occurring during the transaldimination reaction (Scheme 1). An interesting feature of Tyr55' resides in its engagement in a conserved cation– π interaction with the positively charged Arg235. Cation– π interaction, the electrostatic attraction between a cation and the negative electrostatic potential associated with a π system, is recognized to play an important role in the structure and function of biological systems (12). In proteins, cation– π interactions occur between the cationic side chain of lysine or arginine and the aromatic side chain of phenylalanine, tyrosine, or tryptophan; the energies of such interactions are between -2 and -4 kcal/mol (13), similar to those of hydrogen bonds. Beside their structural

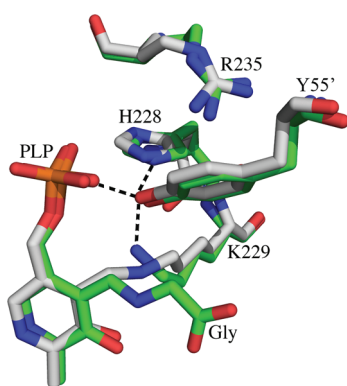
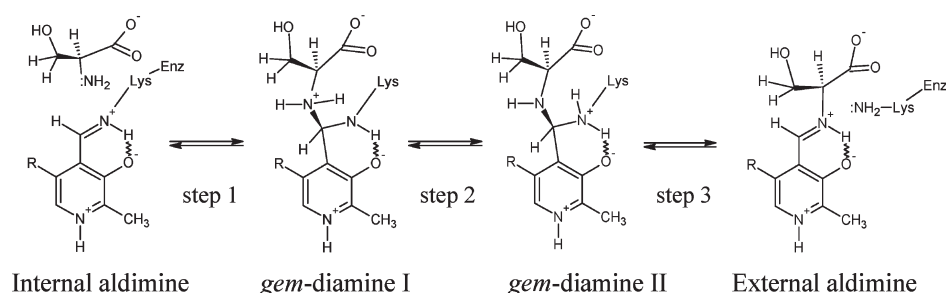


FIGURE 1: Superimposition of the active site residues of wild-type *e*SHMT (green) in complex with glycine and 5-formyl- H_4 PteGlu (PDB ID 1DFO; external aldimine form) and unliganded *Bacillus stearothermophilus* SHMT (*bs*SHMT; gray) (PDB ID 1KKJ; internal aldimine form). Active site residues, PLP, and the PLP–Gly complex are shown as sticks and colored according to atom type: N, blue; O, red; P, orange; C, green for *e*SHMT and gray for *bs*SHMT. Hydrogen bonds of *e*SHMT are indicated by dashed lines. The picture was generated using PyMol (<http://www.pymol.org>).

Scheme 1: Transaldimination Reaction



function, cation– π interactions in proteins are important in the recognition of ligands and in the catalytic mechanism of several enzymes (14, 15). It has been shown that Tyr residues involved in cation– π interactions may have their pK_a lowered by 1 to 3 pH units (the normal pK_a of tyrosine residue is about 9.5) and thus efficiently act as general base catalysts in enzyme reactions (16, 17).

In the present work, a bioinformatic study was carried out, showing that Tyr55' and Arg235 are strongly conserved within the SHMT family and are involved in a cation– π interaction in the large majority of the available SHMT structures. The role of Tyr55' and Arg235 in the catalytic and structural properties of *e*SHMT was then investigated by site-directed mutagenesis and protein crystallography. Y55F, R235Q, R235L, and R235K *e*SHMT mutants were produced and thoroughly characterized with respect to their catalytic properties, substrate and cofactor binding properties, and thermal denaturation properties. The three-dimensional structure of the Y55F mutant was solved by X-ray crystallography.

EXPERIMENTAL PROCEDURES

Materials. Ingredients for bacterial growth and chemicals for the purification of the enzymes were from Sigma-Aldrich; DEAE-Sepharose and phenyl-Sepharose were from GE Healthcare. Wild-type and mutant forms of *e*SHMT and methylenetetrahydrofolate dehydrogenase were purified as previously described (18, 19). The subunit concentration of the holoenzyme was calculated according to a molar absorptivity value of $\epsilon_{280} = 44884 \text{ cm}^{-1} \text{ M}^{-1}$ (20). All enzymes were stored and all experiments performed in 50 mM sodium Hepes (*N*-(2-hydroxyethyl)piperazine-*N'*-2-ethanesulfonic acid), pH 7.2, containing 0.2 mM DTT and 0.1 mM EDTA. PLP was added to protein samples during the purification procedure, but it was left out in the final dialysis step. (6*S*)- H_4 PteGlu, was a gift from Eprova AG, Schaffhausen, Switzerland. PLP was from Sigma-Aldrich (98% pure). All other reagents were from Sigma-Aldrich.

Preparation of Apoenzyme Samples. Apo-*e*SHMT was prepared using L-cysteine as previously described (21). The apoenzyme, whose subunit concentration was calculated according to a molar absorptivity value of $\epsilon_{280} = 42790 \text{ cm}^{-1} \text{ M}^{-1}$ (20), was stored in 10% glycerol at -20°C for no more than 3 days before use. A small, residual fraction (less than 5%) of holoenzyme, estimated by activity assays, was present in the apoenzyme samples. This observation was made with either wild-type or mutant forms of the enzyme.

Site-Directed Mutagenesis. Site-directed mutagenesis of the *E. coli glyA* (SHMT encoding gene) coding region was performed with the QuickChange kit from Stratagene (La Jolla, CA) using the pBS::*glyA* plasmid as template (18) and two

complementary oligonucleotide primers containing the mutations (synthesized by MWG-Biotech AG, Anzinger, Germany). The Y55F, R235K, R235Q, and R235L genes were produced using the primers 5'-GCTGACCAACAAATTTGCTGAAGGTTATC-3', 5'-CTGGCGGGTCCGCAGGGCGGCCTGATCC-3', 5'-CTGGCGGGTCCGAAGGGCGGCCTGATCC-3', and 5'-CTGGCGGGTCCGCTGGGCGGCCTGATCC-3', respectively, and their complementary oligonucleotides (the mutated bases are underlined). *E. coli* DH5 α cells were used to amplify the mutated plasmids. Both strands of the coding region of the mutated genes were sequenced. The only differences with respect to the wild-type nucleotide sequence were those intended. Enzyme expression was performed using the GS1993 *recA*⁻ strain of *E. coli* (18).

Spectroscopic Measurements. Fluorescence emission measurements were carried out at 20 °C with a LS50B spectrofluorometer (PerkinElmer Life Sciences) using a 1 cm path length quartz cuvette. Fluorescence emission spectra were recorded from 300 to 450 nm (1 nm sampling interval), with the excitation wavelength set at 280 nm. UV-visible CD spectra were recorded with a Jasco 725 spectropolarimeter in the circular dichroism mode. Far- (190–250 nm) and near-UV (250–310 nm) and visible (310–500 nm) CD spectra were measured using 0.2 and 1 cm path length quartz cuvettes. Kinetic measurements in the activity assays were performed on a Hewlett-Packard 8453 diode-array spectrophotometer. All spectroscopic measurements were carried out at 20 °C in 50 mM sodium Hepes, pH 7.2, containing 0.2 mM DTT and 0.1 mM EDTA.

Measurement of the K_d of PLP Binding Equilibrium ($K_{d(PLP)}$). PLP binding equilibria were analyzed taking advantage of the protein intrinsic fluorescence quenching observed upon the binding event (21). Dissociation constants of binding equilibria were calculated from saturation curves obtained by measuring the protein fluorescence emission intensity as a function of increasing PLP concentrations, as previously described (22). The cofactor (from 1 to 400 nM) was added to apoenzyme samples (25 nM) at 20 °C in 50 mM sodium Hepes, pH 7.2, containing 0.2 mM DTT and 0.1 mM EDTA. Preliminary experiments demonstrated that the binding equilibrium of all enzyme forms was established within the mixing time. Fluorescence emission spectra (300–450 nm, 5 nm emission slit) were recorded immediately after mixing PLP and apoenzyme with a Horiba-Jobin Yvon FluoroMax-3 spectrofluorometer, with excitation wavelength set at 280 nm (1 nm excitation slit), at the same temperature and with a 1 cm path length quartz cell.

Measurement of the K_d of Substrate Binding Equilibrium ($K_{d(Ser)}$, $K_{d(Gly)}$). Affinity of amino acid ligands for eSHMT was determined from the decrease in ellipticity of the PLP chromophore at 425 nm during titration of eSHMT (33 μ M) with successive increasing concentrations of the amino acid ligand. L-Serine concentrations were from 0.1 to 500 mM. Glycine concentrations were varied between 0.2 and 500 mM. Spectra were recorded between 500 and 350 nm at 30 °C. Data were analyzed according to eq 11.

Rapid Kinetics. Stopped-flow absorbance experiments were performed with an Applied Photophysics SX18 apparatus (Leatherhead, U.K.) equipped with a 1 cm optical path observation chamber. Transient spectroscopy experiments were performed with the same apparatus reconfigured with the photodiode-array accessory. All experiments were carried out in 50 mM sodium Hepes buffer, containing 0.2 mM DTT and

0.1 mM EDTA, pH 7.2, at 30 °C. Each study was an average of four to six traces.

Activity Assays. All assays were carried out at 30 °C in 50 mM sodium Hepes, pH 7.2, containing 0.2 mM DTT and 0.1 mM EDTA. The serine hydroxymethyltransferase activity was measured with 0.05 μ M enzyme samples with L-serine and H₄PteGlu as substrates, using a coupled assay as previously described (23). In order to determine the K_m for L-serine ($K_{m(Ser)}$), H₄PteGlu was maintained at 0.8 mM, and the L-serine concentration was varied between 0.06 and 25 mM. For K_m determinations of H₄PteGlu ($K_{m(THF)}$), L-serine concentrations were held constant at 25 mM, and H₄PteGlu concentrations varied from 3 to 500 μ M. The rate of L-*allo*-threonine cleavage (3 μ M enzyme samples) was followed by determining the rate of reduction of the product acetaldehyde by alcohol dehydrogenase and NADH at 340 nm (24).

Data Analysis. All data analyses were carried out using the software Prism (GraphPad Software Inc., San Diego, CA). Rapid kinetics were analyzed according to the equation describing a single exponential process (25). Steady-state kinetic parameters were obtained by nonlinear least-squares fitting of initial velocity data to the Michaelis–Menten equation.

Fluorescence data obtained in PLP-binding equilibrium experiments were analyzed as previously described (22).

The values of K_d for each amino acid ligand were obtained by computer fitting to eq 11, where $\Delta\Theta$ is the ellipticity change upon substrate titration, $\Delta\Theta_{max}$ is the maximum change of ellipticity, and $[S]$ is the concentration of the amino acid ligand.

$$\Delta\Theta = \Delta\Theta_{max} \frac{[S]}{[S] + K_d} \quad (1)$$

Thermal Denaturation Experiments. The protein samples (2.3 μ M) in 50 mM sodium Hepes buffer, pH 7.2, containing 0.2 μ M DTT and 0.1 μ M EDTA were heated from 30 to 95 °C with a heating rate of 1 deg min⁻¹ controlled by a Jasco programmable Peltier element. The dichroic activity at 220 nm was monitored continuously every 0.5 °C. All thermal scans were corrected for solvent contribution at the different temperatures. Melting temperature (T_m) values were calculated by taking the first derivative of the ellipticity at 220 nm with respect to temperature (26). All denaturation experiments were performed in triplicate.

Protein Crystallization. Crystallization conditions were initially screened by the Phoenix nanoliter dispensing robot (ArtRobbins) and then refined manually. All attempts to crystallize the Arg235 mutants failed, whereas crystallization of the Y55F mutant succeeded only after many trials. Crystallization was achieved at 293 K by the hanging drop vapor diffusion technique. A 2 μ L volume of the protein sample concentrated to 20 mg/mL, equilibrated against 10 mM Tris-HCl, pH 7.5, was mixed with an equal amount of the reservoir solution containing 0.1 M sodium Hepes at pH 7.5 and sodium citrate at a concentration of 1.4 M. Only a few crystals were obtained, and it was not possible to crystallize the protein in the presence of glycine or L-serine. The crystals grew in 3–4 weeks to 0.4 \times 0.3 \times 0.2 mm³.

Data Collection and Processing. Diffraction data from different crystals were collected at different synchrotron radiation sources (ESRF (Grenoble, France), BESSY (Berlin, Germany), and DESY (Hamburg, Germany)). The best data set was collected as 1.0° oscillation frames using the MAR CCD

Table 1: Crystal Parameters, Data Collection, and Refinement Statistics of Y55F eSHMT^a

data reduction and crystal parameters	
space group	P41212
<i>a</i> (Å)	112.6
<i>b</i> (Å)	112.6
<i>c</i> (Å)	60.1
no. of molecules in ASU	1
solvent content (%)	48.9
Matthews coeff (Å ³ /Da)	2.4
resolution range (Å)	50–3.3 (3.42–3.3)
unique reflections	6600
completeness (%)	94.9 (97.3)
<i>R</i> _{merge} ^b	0.13 (0.48)
χ^2	1.3 (0.86)
$\langle I/\sigma(I) \rangle$	9.7 (2.5)
refinement	
resolution range (Å)	50–3.3
reflections used for refinement	6356
<i>R</i> _{crys} (%)	25.5
<i>R</i> _{free} (%)	27.5
correlation coeff <i>F</i> _o – <i>F</i> _c	0.896
correlation coeff <i>F</i> _o – <i>F</i> _c free	0.897
Ramachandran plot	
residues in most favored region (%)	90.7
residues in additionally allowed region (%)	7.3
residues in generously allowed region (%)	2.0

^aValues in parentheses are for the highest resolution shell. ^b*R*_{merge} = $\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where *I*_i(*hkl*) is the *i*th observation of the reflection (*hkl*) and $\langle I(hkl) \rangle$ is the mean intensity of the (*hkl*) reflection.

detector on the BL14-1 beamline at BESSY (Berlin, Germany) at a wavelength of 0.97 Å. Data analysis performed with DENZO (27) indicated that the crystals are tetragonal (P41212) with the following unit cell dimensions: *a* = 112.6 Å; *b* = 112.6 Å; *c* = 68.9 Å. Data were scaled using SCALEPACK (27) and are 94.0% complete at 3.3 Å resolution, with an *R*_{merge} of 13% and a χ^2 of 1.3. The crystal contains two monomers per asymmetric unit, with a *V*_M of 2.41 Å³ Da^{−1} and a solvent content of 48.9% (Table 1).

Structure Solution and Refinement. The structure was solved by molecular replacement using a polyalanine-truncated model built from the wild-type protein (PDB ID 1DFO) as search probe for the rotational and translational searches. The rotational and translational searches, performed with PHASER (28) in the resolution range 10–3.0 Å, produced a clear solution corresponding to a monomer in the asymmetric unit. Refinement of the atomic coordinates and displacement parameters was carried out by Refmac5 (29–31). The final *R* factor and *R* free at 3.3 Å resolution are 25.5% and 27.5%, respectively, with rmsd values of 0.005 Å on the bond lengths and 0.776° on the bond angles. All refinement statistics are presented in Table 1. Model building was performed using the program package COOT (32). The final model (a monomer) includes 417 residues and one PLP molecule. The quality of the model was assessed using PROCHECK (33); 100% of the residues were within the allowed or generously allowed regions of the Ramachandran plot.

Analysis of SHMT Structures and Sequences. The coordinates of the 29 structures of SHMTs available from the Protein Data Bank (34) were downloaded from the PDB Web site (<http://www.rcsb.org/pdb/>). For the 17 structures containing coordinates only for one monomer we generated the symmetry-related subunits using the program COOT (32). To identify cation– π interactions, we adopted the geometric criteria described in ref 35, namely, (i) distance between the midpoint

of the aromatic ring of the Tyr residue and the nitrogen atom of the Arg residue closest to the Tyr ring ≤ 4.0 Å, and (ii) angle between the line connecting the two aforementioned atoms and the normal to the plane where the aromatic ring of the Tyr lies $\leq 30^\circ$. Both parameters were calculated using in-house built programs. Structure visualization and analysis (e.g., atomic coordinate superpositions) were performed using the software InsightII (Accelrys Inc.) (36).

The multiple sequence alignment of the 1271 proteins assigned to the SHMT family by the Pfam database (37) was retrieved from the Pfam Web site. Sequences incomplete in the regions of interest for our analysis (e.g., those around eSHMT Tyr55' and Arg235) were eliminated, reducing the data set to 1172 sequences. Calculations of amino acid frequencies were performed by in-house built programs.

Accession Numbers. Atomic coordinates and structural factors have been deposited in the Protein Data Bank with accession number 3G8M.

RESULTS

The Cation– π Interaction at the Active Site of eSHMT.

In the crystal structure of the eSHMT ternary complex with glycine and 5-formyl-H₄PteGlu (all structures available for the *E. coli* enzyme are in the external aldimine form) Tyr55' and Arg235 are placed at the dimer interface and form a T-geometry (12) cation– π interaction (Figure 1). In eSHMT, the positive charge of Arg235 is not counterbalanced by any negative residue in its proximity, and therefore the cation– π interaction is expected to be quite strong, 3–10-fold stronger than usual (38). The distance between Arg235 NH₂ and the centroid of the Tyr55' aromatic ring is 3.4 Å. The side chains of Tyr55' and Arg235 are involved in additional interactions with both protein and ligands: the Tyr55' hydroxyl group is hydrogen bonded to the side chain of His228 (OH–NE2 = 3.0 Å), to the side chain of Lys229 (OH–NZ = 3.5 Å), and to the phosphate group of PLP (OH–OP = 2.5 Å); Arg235 is hydrogen bonded to the main chain carbonyl oxygen of residue Thr52' (NH1–O = 3.0 Å). In the unliganded enzyme from *Bacillus stearothermophilus* (bsSHMT; this is the unliganded structure most similar to eSHMT available in the PDB), PLP is covalently bound to the protein as an internal aldimine through a Schiff base linkage between its C4' and the NZ of Lys229. The distance between the Tyr55' hydroxyl oxygen and the internal aldimine nitrogen is 3.4 Å (Figure 1). When amino substrates bind to the enzyme, an external aldimine is formed (Scheme 1). The crystal structure of eSHMT in complex with glycine shows a 3.6 Å distance between the Tyr55' hydroxyl oxygen and the external aldimine nitrogen. The mechanism of transaldimination (i.e., the transition between internal and external aldimines) implies the presence of two *gem*-diamine intermediates (Scheme 1), which differ for the position of a proton. The superposition of the internal and external aldimine SHMT structures (Figure 1) shows that the Tyr55' hydroxyl group is ideally positioned to function as the general acid–base catalyst in this proton transfer and in proton transfers that might take place in the previous or following steps of the transaldimination reaction. Moreover, as said in the introduction, the involvement of Tyr55' in a cation– π interaction is expected to enhance the acidic properties of this residue. The O–H bond of Tyr55' is further polarized by the hydrogen bonds established with His228 and the phosphate group of PLP. The constraints imposed by the cation– π and hydrogen bond interactions help keeping the hydroxyl group in a suitable orientation for acting as the acid–base catalyst in the transaldimination process.

Analysis of SHMT Structures and Sequences. Analysis of the 29 SHMT structures available from the PDB (34) showed that the interactions established by the side chain of Tyr55' in *e*SHMT, namely, the cation- π interaction with Arg235 and the hydrogen bonds with the side chain nitrogen of His228 and the phosphate moiety of PLP, are conserved in 41 of the 52 nonidentical active sites (see Supporting Information). In particular, this set of interactions is present in all structures of wild-type SHMTs in the holoenzyme form and in the large majority of the mutant forms (17 and 24 active sites, respectively), whereas it is altered in the structure of the enzymes in the apo form and in some of the structures containing mutations of the glutamate residue homologous to Glu57' in *e*SHMT, which is essential for catalysis (11) (5 and 6 nonidentical active sites, respectively). In all of the 41 active sites where the tyrosine interactions are conserved, the hydroxyl group of this residue points toward the region where the transaldimination reaction takes place, as indicated by its distance from the C4' atom of PLP, the side chain nitrogen atom of the lysine residue homologous to Lys229, and the nitrogen atom of the amino acid substrate (Table A1). In the 6 active sites of mutant SHMTs where the set of interactions is altered (E53Q *bs*SHMT and E75L rabbit cytosolic SHMT) the cation- π interaction and the hydrogen bond of the tyrosine hydroxyl group with the histidine side chain are not present (Table A1). However, also in these cases, the hydrogen bond interaction with PLP is present, and the tyrosine hydroxyl group points toward the active site region where the transaldimination reaction takes place. The main chain conformation of these mutants is the same as that of the wild-type enzymes (see Supporting Information), suggesting that small changes in the side chain conformation would allow all interactions to occur. Conversely, in the active site structure of the apoenzyme forms, the tyrosine side chain assumes a completely different orientation and does not establish any of the typical interactions. Evidently, the presence of PLP and the formation of a hydrogen bond between the tyrosine residue and the cofactor phosphate are required for these interactions to occur. For an expanded analysis of SHMT structures, see Supporting Information.

In addition to the structural conservation of the cation- π interaction, Tyr55' and Arg235 are almost completely conserved in SHMT sequences. Tyrosine and arginine residues are present in the corresponding positions in 98% and 96%, respectively, of the proteins assigned to the SHMT family by the Pfam database, the most frequent replacement for arginine being glutamine (3% of the sequences, which interestingly correspond to SHMTs from Archaea).

Interestingly, the cation- π interaction is not present in other enzymes closely related to SHMTs, such as those assigned to the GABA-aminotransferase-like family by the SCOP database (see Table 2 in Supporting Information), or in more distantly related PLP-dependent enzymes, such as chicken heart mitochondrial and cytosolic aspartate aminotransferase (PDB IDs 7AAT and 2CST, respectively), *Citrobacter freundii* tyrosine phenol-lyase (PDB ID 1TPL), porcine DOPA decarboxylase (PDB ID 1JS3), and *E. coli* glutamate decarboxylase (A and B form) (PDB IDs 1XEY and 1PMM, respectively). With the only exception of aspartate aminotransferases, none of these enzymes contain tyrosine or arginine residues in a position similar to that observed in SHMT structures (i.e., within a radius of 6.0 Å from the NZ atom of the catalytic lysine and the phosphate and C4A atoms of PLP). In the two isoforms of aspartate aminotransferase the

hydroxyl group of Tyr70' is hydrogen bonded to the phosphate moiety of PLP, but the side chain of Arg266, equivalent to *e*SHMT Arg235, is salt-bridged to the PLP phosphate and does not form a cation- π interaction with Tyr70'.

Y55F *e*SHMT Crystal Structure. The structure of Y55F *e*SHMT in the absence of substrates was solved by molecular replacement at 3.3 Å resolution. The crystallographic data are summarized in Table 1, and details of structure refinement are described in Experimental Procedures. Y55F *e*SHMT displays the classic SHMT fold. It is a homodimer whose single subunits comprise three different domains: the N-terminus (residues 1–33), containing one α -helix and one short β -strand; a large N-terminal domain (residues 34–281), which is an $\alpha\beta\alpha$ sandwich, made of ten α -helices wrapped around a seven-stranded β -sheet; and a small C-terminal domain (residues 282–417), which is an $\alpha\beta$ sandwich formed by a four-stranded β -sheet and five α -helices, plus one small β -strand pairing with that of the N-terminus. The active site is located at the interface between the large and small domains and is made of amino acid residues contributed by both subunits of the catalytic dimer.

The Y55F mutant shows significant structural differences with respect to wild-type *e*SHMT, concerning Phe55', the PLP cofactor, and the interacting protein residues. With respect to wild-type Tyr55', the aromatic ring of Phe55' is shifted and rotated about 90°, so that it is too far away from Arg235 for a cation- π interaction (Figure 2, panels A and C). In the mutant structure, Arg235 forms a salt bridge with Glu36 (NH1-OE1 = 3.8 Å), which in the wild-type *e*SHMT interacts with the amide nitrogen of Tyr55', balancing the loss of the cation- π interaction. In the wild-type *e*SHMT, Tyr65' and Ser35 interact with the carboxylate group of substrates, and the loop regions made of residues 34–37 and 64'–67' protrude toward PLP. In the Y55F mutant, both loop regions have moved away from the PLP binding site, and the hydroxyl group of Ser35 is oriented in the opposite direction with respect to the wild-type structure.

The mode of binding of PLP is substantially different with respect to that observed in wild-type *e*SHMT (Figure 2) and in all other structures of the SHMT family, in which a number of conserved interactions can be observed (Figure 2, panel B): (i) a stacking interaction between the pyridine ring and His126; (ii) a salt bridge between the pyridinium nitrogen and Asp200 (N1-OD2 distance = 2.6 Å); (iii) a hydrogen bond of the C3 hydroxyl group with the side chain hydroxyl group of Ser175 (O3-OG distance = 2.8 Å); (iv) hydrogen bonds of phosphate oxygens with the side chain hydroxyl groups of Tyr55' (O2P-OH distance = 2.8 Å) and Ser99 (O1P-OG distance = 3.5 Å). Other conserved interactions (not shown in Figure 2) include the hydrogen bonds between the C3 hydroxyl group and the nitrogen of His203 (O3-ND1 distance = 3.1 Å) and between the phosphate oxygens and the main chain nitrogens of Gly98 (O1P-N = 3.4 Å), Ser97 (O1P-N = 3.0 Å), and Gly263' (O2P-N = 3.0 Å). In the Y55F mutant, the PLP cofactor is found in two different positions, marked PLP-A and PLP-B in Figure 2 (panel A), having an occupancy of 0.5 each and higher mobility than the rest of the structure (the averaged *B* factor of PLP is 13 Å² higher than that of the rest of the protein). In both conformations, the cofactor, whose C4' is covalently linked to Lys229 NZ to form an internal aldimine, establishes peculiar interactions with the protein moiety. The pyridine ring is rotated about 90° with respect to the wild-type structure and has moved away from His126 (in the direction of Phe55'), losing the stacking interaction with this residue (Figure 1, panels A and C). The salt-bridge

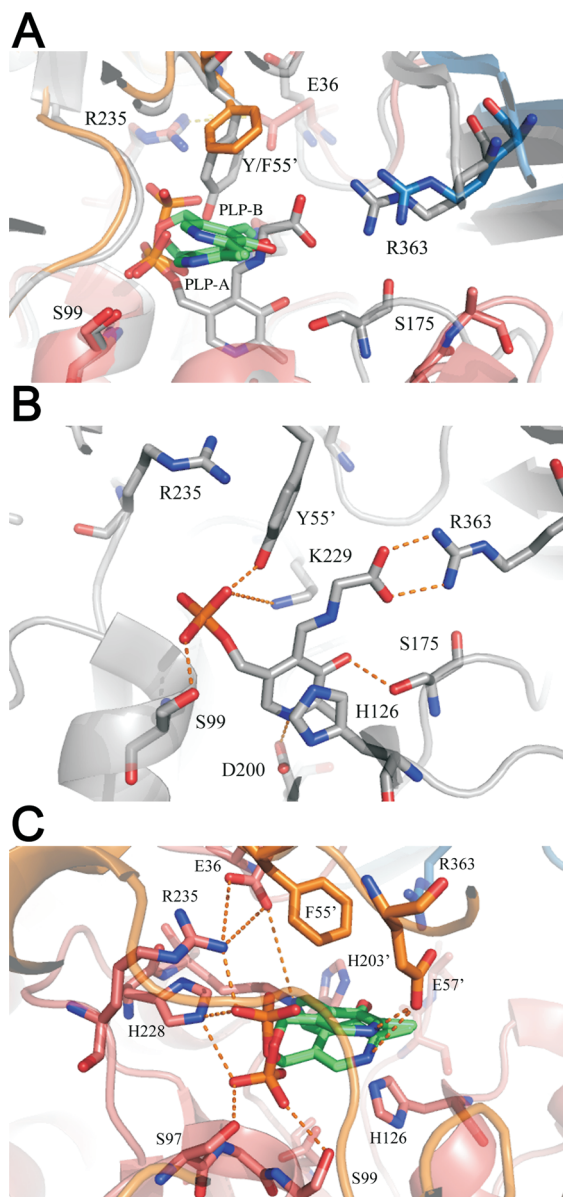


FIGURE 2: Comparison between the active sites of wild-type and Y55F *e*SHMTs. The PLP cofactor and the residues lining the catalytic pocket are indicated as sticks. N, O, and P atoms are colored as in Figure 1. Carbon atoms and secondary structure elements are colored in gray for wild-type *e*SHMT, and orange, red, and blue for the B subunit and the N- and C-terminal domains of the A subunit of Y55F, respectively. Carbon atoms of PLP in Y55F are in green. Hydrogen bonds are shown as dotted lines. (A) Superimposition of the active site structures of wild-type *e*SHMT (in the ternary complex with glycine and 5-formyl- H_4 PteGlu; PDB ID 1DFO) and of the Y55F mutant (unliganded form; PDB ID 3G8M). (B) Close-up of the active site of wild-type *e*SHMT. (C) Close-up of the active site of Y55F *e*SHMT. The pictures were generated by PyMol.

interaction between the pyridinium nitrogen and Asp200 (Figure 2, panel B), which is conserved in all enzymes belonging to the SHMT family and is indispensable for catalysis (39), is replaced by a salt bridge with Glu57' (N1-OE2 = 2.8 and 3.7 Å in PLP-B and PLP-A, respectively; Figure 2, panel C). In wild-type *e*SHMT, Glu57' binds to the hydroxyl group of the L-serine substrate and is essential for catalysis (11). In Y55F, the hydrogen bond between the C3' hydroxyl group of PLP and Ser175 is lost, since the loop of residues 171–182 is shifted apart from the PLP binding site and the side chain hydroxyl group of Ser175 is oriented in the opposite direction (Figure 2, panel A). In PLP-A,

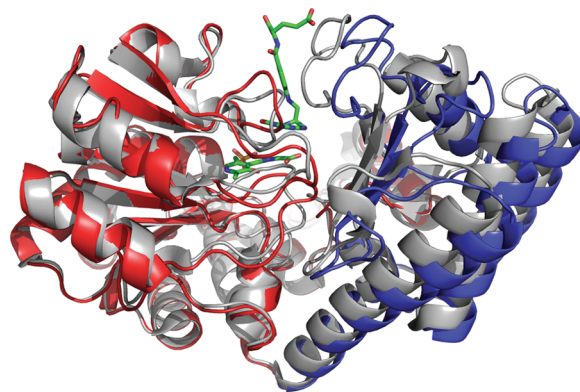


FIGURE 3: Monomeric crystal structures of wild-type *e*SHMT in complex with glycine and 5-formyl- H_4 PteGlu (PDB ID 1DFO) and Y55F *e*SHMT (PDB ID 3G8M) after structural superimposition of the large N-terminal domains. Wild-type *e*SHMT is colored in gray, and the large N-terminal and small C-terminal domains of the Y55F mutant are red and blue, respectively. The PLP-glycine complex and 5-formyl- H_4 PteGlu in wild-type *e*SHMT are indicated as sticks and colored by atom type: N, blue; O, red; P, orange; C, green. The picture was generated by PyMol.

the phosphate group points to the opposite direction with respect to wild-type *e*SHMT. Therefore, it establishes novel interactions with Arg235 (O3P-NH2 = 2.8 Å; O1P-NH2 = 3.3 Å) and His228 (O3P-NE2 = 3.3 Å) (Figure 2, panel C), which in wild-type *e*SHMT are respectively involved in a cation- π interaction and a hydrogen bond with Tyr55'. In PLP-B, the phosphate group is in a similar position with respect to wild-type *e*SHMT, and accordingly, it establishes hydrogen bonds with Ser97, Gly98, and Ser99 (Ser 97 OG-O3P = 2.7 Å; Ser 97 OG-O2P = 2.9 Å; Gly98 N-O2P = 2.9 Å; O3P-Ser99 OG = 3.0 Å; O3P-Ser99 N = 3.2 Å) and with His228 (NE2-O2P = 3.5 Å) (Figure 2, panel C). Finally, the C2 and C2A atoms of both PLP conformations point toward a cavity which, in the wild-type ternary complex, is occupied by 5-formyl- H_4 PteGlu. The side chain of residues Tyr65' and Phe357 also overlaps to the wild-type 5-formyl- H_4 PteGlu binding site. To allow for 5-formyl- H_4 PteGlu binding, PLP, Tyr65', and Phe357 must then be displaced and assume a conformation similar to that of the wild-type enzyme.

Conformational differences between wild-type and Y55F *e*SHMT can be also observed with respect to the relative orientation of the large N-terminal and small C-terminal domains, which enclose the active site (Figure 3). Separate superposition of domain main chain atoms yielded rmsd values of 0.9 Å for the large domain and 0.7 Å for the small domain, whereas superposition of both domains yielded the somewhat higher rmsd value of 1.4 Å. By comparison, in *bs*SHMT, which is a close homologue of *e*SHMT (58% sequence identity), substrate-induced conformational changes are much smaller. Structures of wild-type *bs*SHMT are available both in the absence of substrates (PDB ID 1KKJ) and in the presence of glycine and 5-formyl- H_4 PteGlu (PDB ID 1KL2). Separate superposition of the large and small domains provided rmsd values of 0.4 and 0.3 Å, respectively, while the superposition of both domains gave an rmsd value of 0.6 Å. As shown in Figure 3, where the superposition of wild-type and Y55F *e*SHMT is limited to the large domain, in the mutant enzyme the small domain is further apart from the large domain, assuming a so-called open conformation.

Spectroscopic Studies. Wild-type *e*SHMT exhibits a single absorption band in the visible region of the spectrum, with

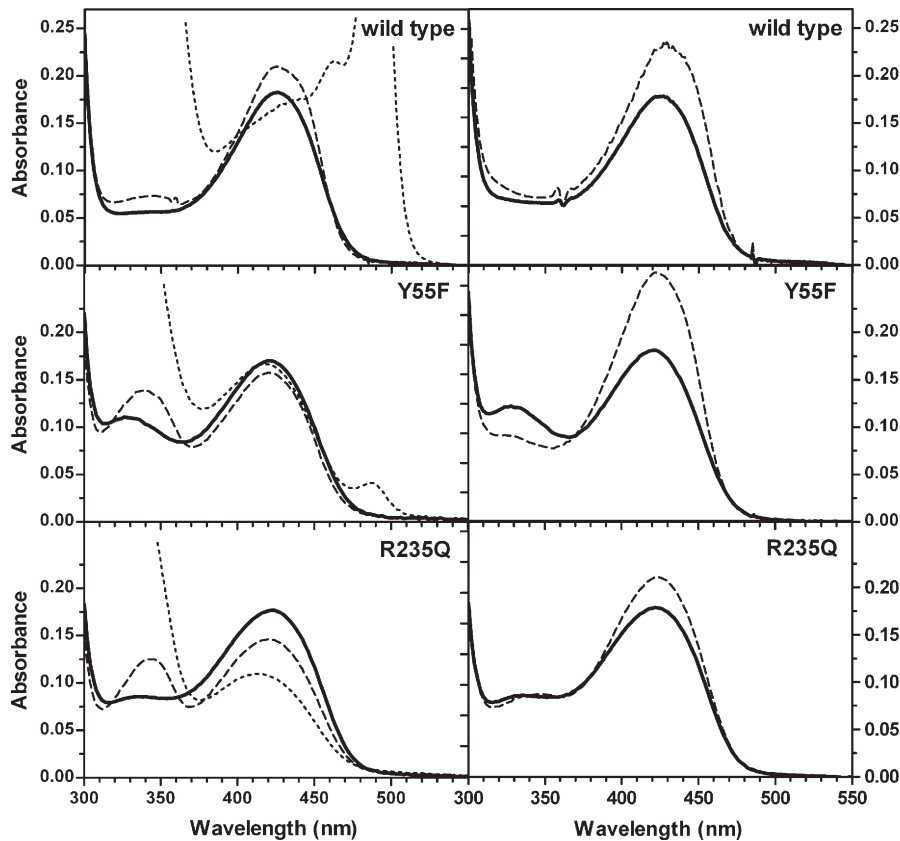


FIGURE 4: Absorption spectra of wild-type and mutant *eSHMT* forms in the absence and presence of substrates. Left panels show the absorption spectra of 30 μM wild-type, Y55F, and R235Q *eSHMT*s at 30 $^{\circ}\text{C}$ before (continuous, thick lines) and after the addition of 90% saturating glycine (dashed lines). Dotted lines are spectra taken after the addition of 100 μM H_4PteGlu to the samples containing glycine. Right panels show the absorption spectra of the same enzymes before (continuous, thick lines) and after the addition of 90% saturating L-serine (dashed lines).

maximum intensity at 422 nm, typical of the protonated internal aldimine of PLP. All absorption spectra are shown in Figure 4. In the spectrum of the Y55F mutant, the corresponding band has slightly lesser intensity and is shifted toward lower wavelengths; moreover, a second band centered at about 326 nm is present. The R235K, R235Q, and R235L mutants show very similar absorbance spectra, with a major band at 422 nm and a small band at 334 nm (for simplicity, Figure 4 shows only the spectrum of the R235Q mutant).

The addition of 90% saturating glycine (60 mM) to wild-type *eSHMT* determined an increase of the 422 nm absorbing band, corresponding to the formation of the external aldimine, and the appearance of a band with a maximum at 343 nm, interpreted as the *gem*-diamine intermediates (Scheme 1) (1). When 100 μM H_4PteGlu was added to the enzyme solution containing glycine, another band with maximum absorbance at 492 nm appeared, to the detriment of the external aldimine. It is known from literature that the 492 nm absorbing band corresponds to the accumulation of the quinonoid intermediate resulting from deprotonation of the C_α of glycine in the $\text{SHMT}\cdot\text{Gly}\cdot\text{H}_4\text{PteGlu}$ ternary complex (1). The effect of H_4PteGlu addition on the *gem*-diamine intermediate concentration is obscured by the intense absorption of this compound in the UV region. The absorbance changes observed upon binding of 90% saturating glycine to the mutant enzymes (calculated on the basis of the relevant dissociation constants; Table 2) are characterized by the appearance of a much more prominent 343 nm absorbing band. The bigger accumulation of the *gem*-diamine intermediate corresponds to a decreased

Table 2: Dissociation Constants (K_d) of PLP and Substrate Binding Equilibria and Kinetic Constants (k_{obs}) for Substrate Binding

enzyme form	$K_d(\text{PLP})^a$ (nM)	$K_d(\text{Gly})$ (mM)	$K_d(\text{Ser})$ (mM)	$k_{\text{obs}}(\text{Gly})^a$ (s^{-1})	$k_{\text{obs}}(\text{Ser})$ (s^{-1})
wild type	5.0	6.7	1.5	300	23
Y55F	19.5	70	23	0.32	0.55
R235K	4.2	204	103	0.42	0.66
R235Q	2.3	154	54	0.36	0.84
R235L	3.1	118	35	0.44	0.60

^aThe values of K_d and k_{obs} are the average of three determinations, which varied within a range of $\pm 5\%$.

intensity of the external aldimine band. The formation of the quinonoid intermediate upon addition of 100 μM H_4PteGlu was observed only with the Y55F mutant, although to a much lesser extent with respect to wild-type *eSHMT*.

The addition of 90% saturating L-serine (13.5 mM) to wild-type *eSHMT* determined an increase of the 422 nm band that also shifted to higher wavelengths. The presence of a *gem*-diamine intermediate is not evident from the absorption spectrum. Similarly, L-serine binding to the mutant enzymes had the effect to increase the 422 nm band. However, a small band around 330 nm, which may correspond to the *gem*-diamine intermediate, is visible in the spectrum of Y55F and R235Q mutant enzymes. Again, spectral changes observed with R235L and R235K mutants are not shown, being very similar to those observed with the R235Q mutant.

The far-UV CD spectra of wild-type and mutant enzymes at 2.5 μM subunit concentration were virtually identical

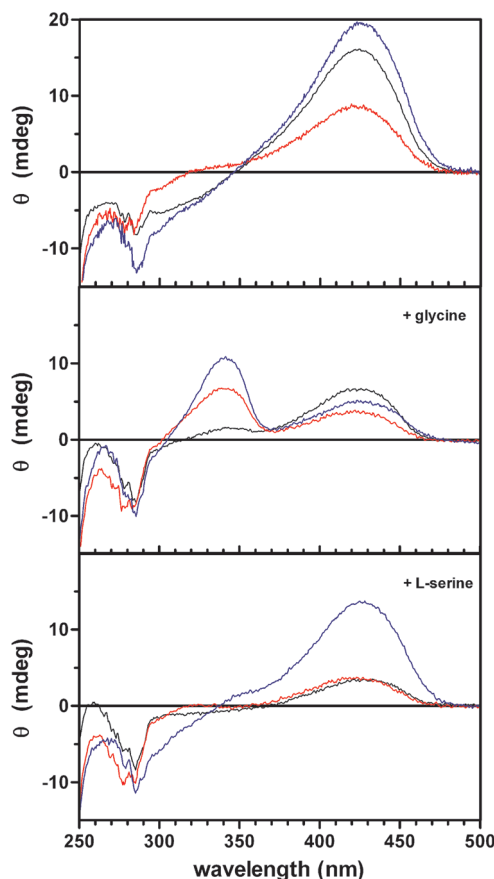


FIGURE 5: Circular dichroism spectra of wild-type and mutant *e*SHMT forms in the absence and presence of substrates. All CD spectra were recorded on 30 μ M enzyme samples at 30 $^{\circ}$ C. The spectra of wild type (black line), Y55F (red line), and R235Q (blue line) are shown before (upper panel) and after the addition of 90% saturating glycine (middle panel) or L-serine (lower panel).

(data not shown), indicating that the mutations did not alter the secondary structure of the enzyme.

The visible CD spectrum of wild-type *e*SHMT is characterized by a positive cofactor band centered around 422 nm. All CD spectra are shown in Figure 5. In the near-UV region, where the ellipticity becomes negative, a broad band between 320 and 340 nm, also attributable to the cofactor, is visible. Below 310 nm, the spectrum shows the characteristic CD bands of aromatic residues (40). The cofactor contribution to the CD spectrum of the Y55F mutant consists of two positive bands: one at 422 nm, with lower ellipticity with respect to wild-type *e*SHMT, and the other one in the 320–340 nm region. This latter band may correspond to the 326 nm band visible in the absorption spectrum of the Y55F mutant (Figure 4). The aromatic CD bands of Y55F are substantially different from those of the wild-type enzyme. As far as the Arg235 mutants are concerned, the CD spectrum in the whole near-UV/visible region is quite similar in shape to that of wild-type *e*SHMT, although it is characterized by higher ellipticity values.

Changes of visible CD spectra upon binding of 90% saturating glycine to wild-type and mutant *e*SHMTs reflect what is seen in the absorbance spectra. The accumulation of a *gem*-diamine intermediate, corresponding to a positive 343 nm CD band, which can be barely seen with the wild type, is particularly evident with the Y55F and Arg235 mutants (Figure 5, middle panel). Interestingly, the aromatic CD bands of the wild-type and mutant enzymes, so different from each other when the enzymes

are in the unliganded form, become more alike when glycine is bound. On the other hand, a substantial difference in the CD spectra of wild-type and mutant enzymes is observed upon binding of 90% saturating L-serine (Figure 5, lower panel). In particular, the CD spectrum of the serine-bound Arg235 mutants is profoundly different with respect to that of the wild-type and Y55F forms, in both the visible and aromatic regions; similarly to the unliganded form, this spectrum is characterized by a prominent positive cofactor band at 422 nm and a strong negative contribution between 300 and 340 nm.

Rapid Kinetics of Substrate Binding ($k_{\text{obs(Ser)}}$, $k_{\text{obs(Gly)}}$). The spectral changes observed upon rapid mixing of 60 μ M wild-type *e*SHMT with a pseudosaturating concentration of glycine (25 mM), in a diode array stopped-flow spectrophotometer, were very small and fast. Two absorption bands with maxima at 343 and 426 nm were present in the first spectrum acquired after 2 ms from the stop of flow (data not shown). From comparison of this spectrum to that of the enzyme without ligands, it is clear that the absorption band at 343 nm has formed in the dead time of the instrument and that at the same time the internal aldimine 422 nm band has turned into a less intense band with a maximum at 426 nm. In the following course of the reaction, the amplitude of the 343 nm band decreased, while the 426 nm band increased. Single wavelength kinetics at 426 and 343 nm are symmetric and correspond to the same exponential process with a rate constant of approximately 300 s^{-1} (data not shown). Stopped-flow and temperature-jump experiments with glycine had previously shown that a 343 nm band corresponding to the *gem*-diamine intermediates is formed very rapidly in a bimolecular step and then transformed into a 425 nm absorbing external aldimine (41). It follows that the spectral changes we observed correspond to the conversion of the *gem*-diamine intermediates into the external aldimine (Scheme 1, steps 2 and 3).

The same experiment performed with 60 μ M Y55F and R235Q mutant enzymes and 400 mM glycine gave similar spectral changes but much slower kinetics (Figure 6). With each mutant, single wavelength absorbance changes at 343 and 426 nm were globally fitted to an exponential equation with shared rate constant. Table 2 shows that the observed rate constants ($k_{\text{obs(Gly)}}$) of external aldimine formation for the mutant enzymes are about 3 orders of magnitude smaller than that of the wild-type enzyme.

Rapid kinetic experiments with L-serine gave very similar spectral changes to those observed with glycine. When wild-type *e*SHMT (60 μ M) was mixed with a pseudosaturating concentration of L-serine (25 mM), symmetrical absorbance changes at 330 and 426 nm were obtained (data not shown). These results are consistent with an exponential process with a rate constant of 23 s^{-1} . Experiments carried out with the mutant enzymes, with 400 mM L-serine, gave 30–40-fold slower rate constants (Table 2, Figure 7).

Cofactor and Substrate Binding Equilibria. The affinity of wild-type and mutant forms for the cofactor was measured in order to evaluate the impact of the mutations on the structure of the active site. Since PLP binding to apo-*e*SHMT is known to quench the intrinsic fluorescence emission of the enzyme (21), the dissociation constant of the binding equilibrium was calculated from saturation curves obtained by measuring the fluorescence emission of apoenzyme (25 nM/subunit) at increasing PLP concentrations (22). Table 2 shows that the calculated apparent $K_{\text{d(PLP)}}$ value is similar to wild type for all Arg235 mutants and about 4-fold higher with Y55F.

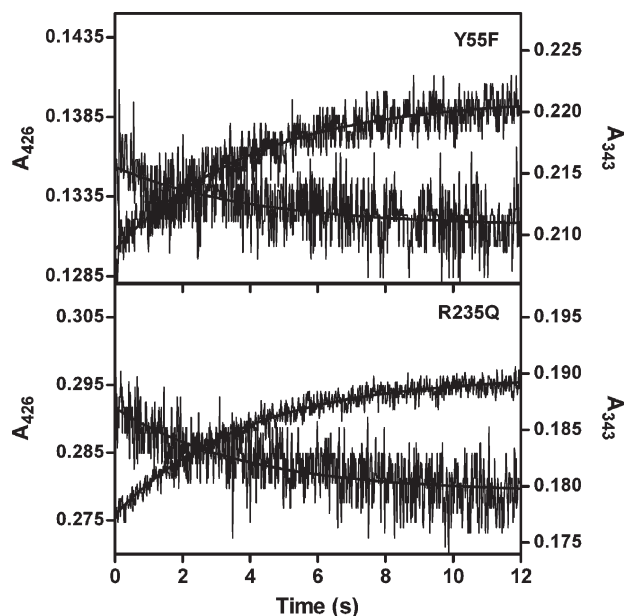


FIGURE 6: Rapid kinetics of glycine binding to mutant *e*SHMTs. Equal volumes of enzyme and glycine solutions in 50 mM sodium Hepes buffer, pH 7.2, were stopped-flow mixed at 30 °C and absorbance changes at 343 and 426 nm recorded. Enzyme concentration after mixing was 60 μ M; glycine concentration was 400 mM with all mutant enzymes. Experimental traces shown in the figure are the average of four separate experiments and correspond to an increase of absorbance at 426 nm and a decrease at 343 nm. Continuous lines through the experimental data were obtained from global fitting of data from both wavelengths to an exponential process with rate constants listed in Table 2.

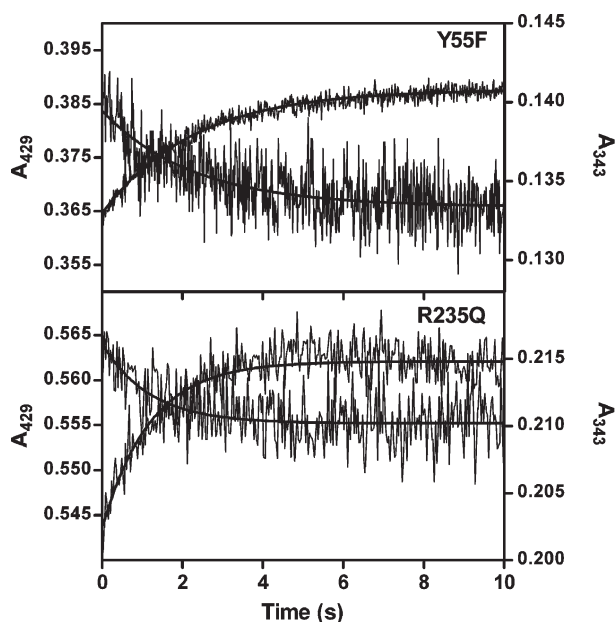


FIGURE 7: Rapid kinetics of L-serine binding to mutant *e*SHMTs. Stopped-flow experiments with L-serine were carried out in the same conditions described in Figure 6 and at the same enzyme concentration. L-Serine concentration after mixing was 400 mM with all mutant enzymes. Kinetic traces were acquired at 330 and 426 nm. Experimental data from both wavelengths were globally fitted to an exponential process with rate constants listed in Table 2.

Binding of either L-serine or glycine to *e*SHMT results in an extensive change of the circular dichroism spectrum of the enzyme in the region where the cofactor absorbs light (42).

The change of ellipticity at 422 nm was measured as a function of substrate concentration in order to determine, according to eq 11, the dissociation constant of substrate binding equilibria, which are reported in Table 2. With respect to the wild-type enzyme, calculated dissociation constants for glycine and L-serine are about 10–15-fold higher for Y55F and about 25–35-fold higher for the Arg235 mutants.

Catalytic Properties of Mutant Enzymes. The Y55F mutation resulted in a 4-fold reduction of k_{cat} for the serine hydroxymethyltransferase reaction; 15-, 30-, and 60-fold decreases of the same parameter were observed with R235K, R235Q, and R235L, respectively (Table 3). All mutations had the effect to increase the $K_{\text{m(Ser)}}$, which was 50-fold bigger with Y55F and 1500-, 900-, and 450-fold bigger with R235K, R235Q, and R235L, respectively. The $K_{\text{m(THF)}}$ was also increased by all mutations, although to a minor extent (less than 20-fold). In the retroaldol cleavage of L-*allo*-threonine, the Y55F mutation resulted in 25-fold reduction of k_{cat} ; 200- and 150-fold decrease of the same parameter was observed for R235K and R235Q, respectively. Also in this case, the mutations had the effect to increase the $K_{\text{m(}l\text{-}l\text{-}Thr)}$, which was 40-fold bigger with Y55F and 60-fold bigger with R235K and R235Q. Parameters for L-*allo*-threonine cleavage could not be measured for R235L mutant with reasonable accuracy, due to very low activity. Catalytic efficiencies with respect to amino acid substrates were strongly reduced by all mutations. In particular, Y55F showed a 3 order of magnitude decrease for both serine hydroxymethyltransferase and L-*allo*-threonine cleavage reactions; Arg235 mutants showed a 5 order of magnitude decrease for serine hydroxymethyltransferase reaction and a 4 order of magnitude decrease for the L-*allo*-threonine cleavage reaction.

Thermal Denaturation Experiments. The thermal stability of wild-type and mutant *e*SHMTs was investigated by monitoring the change of ellipticity at 220 nm in the far-UV CD spectrum, when the temperature was increased from 30 to 95 °C. The observed sigmoid transitions were irreversible, and the CD spectra measured at the end of the cooling phase differed from those of the protein at 30 °C. The parameter chosen to compare the transition curves was the apparent melting temperature (T_{m}), defined as the midpoint of the sigmoid denaturation process and calculated by plotting the first derivative of the molar ellipticity values as a function of temperature. The experiments were performed on either holo or apo forms of the enzymes. The presence of the cofactor stabilizes both wild-type and mutant enzymes by 10–15 °C. However, Y55F and Arg235 mutations have opposite effects on enzyme stability. In particular, while the Y55F mutation destabilizes apo- and holo-*e*SHMT by about 7 and 4 °C, respectively, the Arg235 mutations stabilize apo and holo forms by about 5–7 °C (Table 4). Binding of 65% saturating L-serine to wild type further increased T_{m} by 3.6 °C. Because of the elevated optical activity of L-serine, which interferes with the measurement of the CD signal at 220 nm, this amino acid had to be added to no more than 65% saturation. Interestingly, L-serine binding to the mutant enzymes did not have any stabilizing effect.

DISCUSSION

Although PLP-dependent enzymes have been intensively studied for more than 40 years, little is known about the catalysis of one of their common mechanistic features: the transaldimination reaction, responsible for substrate binding and product

Table 3: Kinetic Parameters of Serine Hydroxymethyltransferase and L-*allo*-Threonine Retroaldol Cleavage Reactions Catalyzed by Wild-Type and Mutant *e*SHMT Forms^a

enzyme form	serine hydroxymethyltransferase reaction				retroaldol cleavage reaction		
	k_{cat} (min ⁻¹)	$K_{\text{m(Ser)}}$ ^b (mM)	$K_{\text{m(THF)}}$ ^c (mM)	$k_{\text{cat}}/K_{\text{m(Ser)}}$ ^d	k_{cat} (min ⁻¹)	$K_{\text{m(}l\text{-}l\text{-}Thr)}$ (mM)	$k_{\text{cat}}/K_{\text{m(}l\text{-}l\text{-}Thr)}$ ^d
wild type	690	0.14	0.007	1	30	1.5	1
Y55F	180	7	0.11	5.2×10^{-3}	1.2	57	1×10^{-3}
R235K	48	218	0.14	4.4×10^{-5}	0.15	96	0.7×10^{-4}
R235Q	20	127	0.12	3.2×10^{-5}	0.2	95	1×10^{-4}
R235L	12	64	0.08	3.6×10^{-5}	nd ^e	nd	nd

^aKinetic constants are the average of three determinations. The range of values was always less than $\pm 5\%$. ^bApparent K_{m} at saturating [H_4PteGlu]. ^cApparent K_{m} at saturating [L-serine]. ^dCatalytic efficiency with respect to either L-serine or L-*allo*-threonine, expressed as the ratio between mutant and wild-type values. ^eNot determined.

Table 4: Apparent Melting Temperatures (T_{m}) of Wild-Type and Mutant *e*SHMTs

	T_{m} ^a of apoenzymes (°C)	T_{m} of holoenzymes (°C)	T_{m} with L-serine (65% saturation) (°C)
wild type	58.8 \pm 0.3	69.5 \pm 1.3	73.1 \pm 0.9
Y55F	51.6 \pm 0.1	65.6 \pm 0.4	65.7 \pm 0.3
R235Q	64.0 \pm 1.1	75.4 \pm 0.5	72.8 \pm 0.2
R235K	66.3 \pm 0.8	77.3 \pm 0.3	75.2 \pm 0.3
R235L	63.2 \pm 0.9	74.9 \pm 0.4	74.4 \pm 0.5

^a T_{m} values are expressed as average \pm standard deviation calculated on three independent determinations.

release. The hypothesis at the basis of this study was that residue Tyr55' in *e*SHMT, involved in a cation- π interaction with Arg235, may participate to this step of the catalytic cycle as a proton exchanger. The results obtained with the Y55F, R235Q, R235K, and R235L mutants clearly show that the residues under study are both necessary for the correct proceeding of the transaldimination reaction and play a role in the structure and stability of *e*SHMT.

The crystal structure of the Y55F mutant shows a number of differences with respect to wild-type *e*SHMT. In the active site structure of the Y55F mutant, the Phe55' ring has moved from its former position and has tilted by 90°, while Arg235 has not changed its orientation. As is clear from Figure 2, Phe55' and Arg235 do not establish a cation- π interaction. This suggests that the formation of the interaction is subordinated to the presence of a Tyr residue in position 55'. In particular, the analysis of known SHMT structures suggests that the formation of the hydrogen bond between the Tyr hydroxyl group and the PLP phosphate group is required for the cation- π interaction to occur, whereas the hydrogen bond with His228 is not. In agreement with this finding, the cation- π interaction between *bs*SHMT Tyr51' and Arg232, homologous to *e*SHMT Tyr55' and Arg235, respectively, is also lost in the recently reported structure of the Y51F mutant (43). In the *e*SHMT Y55F mutant, PLP is found in two equally populated positions, both with the pyridine ring in a completely different orientation with respect to wild-type (Figure 2) and to all other known SHMT structures. In *e*SHMT, the interaction between Tyr55' and the phosphate group of PLP seems to be of fundamental importance for a correct positioning of the cofactor at the enzyme active site. The atypical mode of PLP binding probably also takes place in the enzyme in solution, as indicated by the anomalous absorption and CD spectra of Y55F (Figures 4 and 5). Nevertheless, in the Y55F mutant, the cofactor establishes a series of alternative

interactions that are roughly equivalent to those found in the wild-type enzyme, and this justifies the relatively low $K_{\text{d(PLP)}}$, that is only 4-fold larger than wild type (Table 2).

The overall crystal structure of this mutant shows the enzyme to be in a more open conformation, the large and small domains being further apart than in the wild-type enzyme (Figure 3). The rearrangement of the 64'-67' loop, which as a consequence of the Y55F mutation has moved away from the PLP binding site, may force the movement of the small C-terminal domain and be responsible for this conformational change (Figure 3). In principle, another factor that might contribute to determine the conformational differences observed between Y55F and wild-type *e*SHMT is the absence of substrates in the first structure and their presence in the second. However, the structural variations are somewhat larger than those observed in other SHMT structures as a consequence of substrate binding, suggesting that they are likely to result mostly from the mutation.

The mutant enzyme in solution shows a lower apparent melting temperature (Table 4) and a lower stability to urea-induced denaturation with respect to wild-type *e*SHMT (data not shown). The Y55F mutation has therefore the general effect to destabilize the enzyme. The decrease in stability of the Y55F mutant may be ascribed to the loss of interactions present in the wild-type enzyme, as shown by the crystal structure, and in particular to the loss of the cation- π interaction with Arg235 (Figure 2, panel A), which is located at the dimer interface.

On the other hand, the R235 mutations do not seem to alter the active site structure to the extent observed with the Y55F mutant. Although the crystal structure of these mutants could not be obtained, their structural similarity to the wild-type *e*SHMT can be inferred by the absorption and CD spectra, as well as by the $K_{\text{d(PLP)}}$ values (Table 2). In particular, the aromatic CD bands of the R235 mutants, which reflect the tertiary structure of the enzyme, are similar in shape with respect to wild type, although showing a difference in the ellipticity values, while the aromatic CD spectrum of Y55F is profoundly different (Figure 5). The absorption spectra of the R235 mutants in the cofactor region (Figure 4) also indicate that PLP is surrounded by a chemical environment similar to that of the wild type and is therefore correctly positioned. Nevertheless, the catalytic efficiency of R235 mutants is 10⁵-fold lower than wild type, whereas, quite surprisingly, the Y55F mutant retains much more activity (Table 3). In all mutants, the catalytic impairment results partially from a decrease of k_{cat} (4-fold in Y55F and 15–60-fold in R235 mutants) but is mainly due to the increase of $K_{\text{m(Ser)}}$ (50-fold in Y55F and 450–1550-fold in R235 mutants). All mutants also show lower affinity for the amino acid substrates

(Table 2). These observations may be accounted for by an effect of the mutations on the transaldimination process. When glycine is bound, an evident accumulation of *gem*-diamine is observed, suggesting that the external aldimine is a relatively less stable intermediate than in the wild-type enzyme. The impairment of the transaldimination process appears to result from the difficulty of the *gem*-diamine to be converted into the external aldimine. In fact, while with all mutants the rapid spectral changes taking place in the dead time of the stopped-flow instrument correspond to the conversion of the internal aldimine into the *gem*-diamine (Scheme 1, step 1), the next step, the conversion of the *gem*-diamine into the external aldimine (Scheme 1, steps 2 and 3), is drastically slowed down by the mutations (Figures 6 and 7, Table 2). This may be thought to be a consequence of the absence of a proton exchanger that catalyzes the interconversion of the *gem*-diamine intermediates and that may also protonate the leaving lysine residue as the external aldimine is formed. This can be obviously due to the lack of the OH group in the Y55F mutant and, in the R235 mutants, to the increase of the pK_a of Tyr55' caused by the loss of the cation- π interaction. However, the k_{cat} of the serine hydroxymethyltransferase reaction catalyzed by the Y55F mutant is more than 5-fold higher than the $k_{obs(Ser)}$ measured upon L-serine binding to the same enzyme, suggesting that some structural rearrangement might take place upon H₄PteGlu binding and partially compensate for the slowing down of the transaldimination reaction. In agreement with the higher efficiency of the Y55F mutant in catalyzing the serine hydroxymethyltransferase reaction with respect to the Arg235 mutants, absorption spectra indicate that after addition of H₄PteGlu (Figure 4) the Y55F mutant is capable of producing the quinonoid intermediate resulting from deprotonation of the C α atom of glycine, although to a far lesser extent than wild-type *e*SHMT, whereas the Arg235 mutants do not form it. It seems worth reminding that, while Tyr55' is directly involved in PLP binding, Arg235 belongs to the second sphere residues, which are too far from PLP to play a direct functional role in catalysis. This observation reinforces the hypothesis that the presence of Arg235 influences catalysis through its interaction with Tyr55'. Although Lys, like Arg, can also establish cation- π interactions with Tyr residues, Arg and Lys do not appear to be interchangeable in this role (44, 45). Indeed, the mutant forms containing a hydrophobic leucine, a polar glutamine, or a positively charged lysine show a similar reduction in k_{obs} values toward both substrates (Table 2) with respect to wild-type *e*SHMT, indicating that none of those residues compensates efficiently for the loss of Arg235. As a support to the hypothesized role of Tyr55', it was recently shown that the Y51F mutant of *hs*SHMT (corresponding to *e*SHMT Y55F) has no detectable catalytic activity and accumulates the *gem*-diamine, although PLP is bound at the active site in the same position as in the wild-type enzyme (43). In order to explain the residual catalytic activity of *e*SHMT Y55F, it seems reasonable to presume that in the absence of the Tyr55' hydroxyl group another residue or a water molecule may act as proton exchanger in the transaldimination reaction. The observed catalytic activity of *e*SHMT Y55F is likely to be subordinated to a rearrangement of PLP upon substrate binding, which places it back into its canonical position. This statement is supported by the spectral analyses of the enzyme-substrate complexes: while the CD aromatic region of unliganded Y55F reveals a difference of tertiary structure with respect to wild type, these differences attenuate upon binding of either glycine or L-serine (Figure 5).

The transition between internal and external aldimine may also be analyzed by thermal denaturation experiments, since it is known that in SHMT the transaldimination process taking place upon L-serine binding is associated to a conformational change (from an open to a closed form) that determines an increase of T_m ; glycine and substrate analogues that bind but do not trigger the conformational change do not increase T_m (46–48). Thermal denaturation experiments show that T_m does not increase upon L-serine binding to mutant *e*SHMTs (Table 4), suggesting that their conformational change is reduced or absent. The conformational change induced by L-serine binding to wild-type SHMT is also indicated by a large decrease of ellipticity of the cofactor CD signal (47). It is clear that, upon binding of L-serine to the mutant enzymes, a much smaller decrease of ellipticity at 420 nm is observed (Figure 5). This hampering in conformational change could contribute, together with the impairment of the transaldimination reaction, to the lower catalytic efficiency of the mutants, in terms of impediment in binding and releasing of substrates and products. For sake of completion, it has to be reported that also in sheep cytosolic SHMT the mutation of the arginine residue homologous to Arg235 into an alanine led to a 30-fold reduction of catalytic activity and to the accumulation of the *gem*-diamine intermediate (10).

The fact that the substitution of Arg235 has led to mutant forms with higher stability is quite unusual, since in general cation- π interactions have a stabilizing effect on protein structure (13), although this finding is not unprecedented (15). In its structural context, Arg235 may represent a destabilizing residue for wild-type *e*SHMT because it does not interact with any negatively charged residue, its positive charges being only counterbalanced by the π -electron cloud of Tyr55' (Figure 2, panel A). The increased stability of the enzyme upon mutation of Arg235 might therefore be contributed by the loss of its only partially balanced positive charge. There are several examples in which the mutation of residues involved in catalysis results in more stable but less active enzymes (49). Interactions that conflict with protein stability (the so-called frustrated interactions) but often play a functionally relevant role have been shown to be evolutionarily conserved in proteins (50) and frequently belong to protein portions where charged amino acids are located in electrostatically unfavorable environments (49). These features are shared by Arg235, which might therefore represent a frustrated residue.

As shown in the Results section, Tyr55' and Arg235 are highly conserved among the SHMT family sequences. The known SHMT structures present an almost complete conservation of the cation- π interaction and of the other interactions that enhance the acidity of the Tyr55' hydroxyl group and maintain it in a suitable position to act as a proton exchanger in the transaldimination reaction. This suggests that the role of the *e*SHMT residues analyzed in this work might well have a more general relevance. The most relevant exception is represented by SHMTs from Archaea, where Arg is replaced by Gln. Based on our finding that the *e*SHMT R235Q mutation increases enzyme stability, it is tempting to speculate that the homologous substitution in SHMT from Archaea may represent one of the structural determinants accounting for adaptation at high temperatures. Given the fact that in the SHMT family the PLP binding site is at the dimer interface, the cation- π interaction between Tyr55' and Arg235 may then represent a compromise between the stability of the enzyme and its catalytic role.

ACKNOWLEDGMENT

We gratefully acknowledge the beamline scientists of the Bessy Synchrotron Radiation Source (Berlin), where the X-ray data were collected. We are also grateful to Andrea Bellelli for useful discussions.

SUPPORTING INFORMATION AVAILABLE

Details concerning the analysis of SHMT structures and sequences. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Schirch, V. (1998) Mechanism of folate-requiring enzymes in one-carbon metabolism, in *Comprehensive biological catalysis: a mechanistic reference* (Sinnott, M., Ed.) pp 211–252, Academic Press, San Diego, CA.
- Renwick, S. B., Snell, K., and Baumann, U. (1998) The crystal structure of human cytosolic serine hydroxymethyltransferase: a target for cancer chemotherapy. *Structure* 6, 1105–1116.
- Scarsdale, J. N., Radaev, S., Kazanina, G., Schirch, V., and Wright, H. T. (2000) Crystal structure at 2.4 Å resolution of *E. coli* serine hydroxymethyltransferase in complex with glycine substrate and 5-formyl tetrahydrofolate. *J. Mol. Biol.* 296, 155–168.
- Scarsdale, J. N., Kazanina, G., Radaev, S., Schirch, V., and Wright, H. T. (1999) Crystal structure of rabbit cytosolic serine hydroxymethyltransferase at 2.8 Å resolution: mechanistic implications. *Biochemistry* 38, 8347–8358.
- Trivedi, V., Gupta, A., Jala, V. R., Saravanan, P., Rao, G. S., Rao, N. A., Savithri, H. S., and Subramanya, H. S. (2002) Crystal structure of binary and ternary complexes of serine hydroxymethyltransferase from *Bacillus stearothermophilus*: insights into the catalytic mechanism. *J. Biol. Chem.* 277, 17161–17169.
- Szebenyi, D. M., Liu, X., Kriksunov, I. A., Stover, P. J., and Thiel, D. J. (2000) Structure of a murine cytoplasmic serine hydroxymethyltransferase quinonoid ternary complex: evidence for asymmetric obligate dimers. *Biochemistry* 39, 13313–13323.
- Paiardini, A., Gianese, G., Bossa, F., and Pascarella, S. (2003) Structural plasticity of thermophilic serine hydroxymethyltransferases. *Proteins* 50, 122–134.
- Angelaccio, S., Chiaraluce, R., Consalvi, V., Buchenau, B., Giangiacomo, L., Bossa, F., and Contestabile, R. (2003) Catalytic and thermodynamic properties of tetrahydromethanopterin-dependent serine hydroxymethyltransferase from *Methanococcus jannaschii*. *J. Biol. Chem.* 278, 41789–41797.
- Schirch, V., and Szebenyi, D. M. (2005) Serine hydroxymethyltransferase revisited. *Curr. Opin. Chem. Biol.* 9, 482–487.
- Appaji Rao, N., Ambili, M., Jala, V. R., Subramanya, H. S., and Savithri, H. S. (2003) Structure-function relationship in serine hydroxymethyltransferase. *Biochim. Biophys. Acta* 1647, 24–29.
- Szebenyi, D. M., Musayev, F. N., di Salvo, M. L., Safo, M. K., and Schirch, V. (2004) Serine hydroxymethyltransferase: role of Glu75 and evidence that serine is cleaved by a retroaldol mechanism. *Biochemistry* 43, 6865–6876.
- Ma, J. C., and Dougherty, D. A. (1997) The cation– π interaction. *Chem. Rev.* 97, 1303–1324.
- Gallivan, J. P., and Dougherty, D. A. (1999) Cation– π interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9459–9464.
- Zacharias, N., and Dougherty, D. A. (2002) Cation– π interactions in ligand recognition and catalysis. *Trends Pharmacol. Sci.* 23, 281–287.
- Prajapati, R. S., Sirajuddin, M., Durani, V., Sreeramulu, S., and Varadarajan, R. (2006) Contribution of cation– π interactions to protein stability. *Biochemistry* 45, 15000–15010.
- Angelucci, F., Baiocco, P., Brunori, M., Gourlay, L., Morea, V., and Bellelli, A. (2005) Insights into the catalytic mechanism of glutathione S-transferase: the lesson from *Schistosoma haematobium*. *Structure* 13, 1241–1246.
- Baiocco, P., Gourlay, L. J., Angelucci, F., Fontaine, J., Herve, M., Miele, A. E., Trottein, F., Brunori, M., and Bellelli, A. (2006) Probing the mechanism of GSH activation in *Schistosoma haematobium* glutathione-S-transferase by site-directed mutagenesis and X-ray crystallography. *J. Mol. Biol.* 360, 678–689.
- Iurescia, S., Condo, I., Angelaccio, S., Delle Fratte, S., and Bossa, F. (1996) Site-directed mutagenesis techniques in the study of *Escherichia coli* serine hydroxymethyltransferase. *Protein Expression Purif.* 7, 323–328.
- Schirch, V. (1997) Purification of folate-dependent enzymes from rabbit liver. *Methods Enzymol.* 281, 146–161.
- Malerba, F., Bellelli, A., Giorgi, A., Bossa, F., and Contestabile, R. (2007) The mechanism of addition of pyridoxal 5'-phosphate to *Escherichia coli* apo-serine hydroxymethyltransferase. *Biochem. J.* 404, 477–485.
- Cai, K., Schirch, D., and Schirch, V. (1995) The affinity of pyridoxal 5'-phosphate for folding intermediates of *Escherichia coli* serine hydroxymethyltransferase. *J. Biol. Chem.* 270, 19294–19299.
- Florio, R., Chiaraluce, R., Consalvi, V., Paiardini, A., Catacchio, B., Bossa, F., and Contestabile, R. (2009) The role of evolutionarily conserved hydrophobic contacts in the quaternary structure stability of *Escherichia coli* serine hydroxymethyltransferase. *FEBS J.* 276, 132–143.
- Schirch, V., Hopkins, S., Villar, E., and Angelaccio, S. (1985) Serine hydroxymethyltransferase from *Escherichia coli*: purification and properties. *J. Bacteriol.* 163, 1–7.
- Schirch, L., and Peterson, D. (1980) Purification and properties of mitochondrial serine hydroxymethyltransferase. *J. Biol. Chem.* 255, 7801–7806.
- Cornish-Bowden, A. (1995) Fundamentals of enzyme kinetics, revised ed., Portland, London.
- Bellapadrona, G., Chiaraluce, R., Consalvi, V., Ilari, A., Stefanini, S., and Chiancone, E. (2007) The mutations Lys 114 \rightarrow Gln and Asp 126 \rightarrow Asn disrupt an intersubunit salt bridge and convert *Listeria innocua* Dps into its natural mutant *Listeria monocytogenes* Dps. Effects on protein stability at low pH. *Proteins* 66, 975–983.
- Otwinowski, Z., and Minor, W. (1997) Macromolecular crystallography. *Methods Enzymol.* 276, 307–326.
- Read, R. J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 57, 1373–1382.
- Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 57, 122–133.
- Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 55, 247–255.
- The CCP4 suite (1994) Programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 50, 760–763.
- Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 60, 2126–2132.
- Laskowsky, R. A., MacArthur, M. W., Moss, D. W., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Crystallogr.* 26, 283–281.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242.
- Babu, M. M. (2003) NCI: a server to identify non-canonical interactions in protein structures. *Nucleic Acids Res.* 31, 3345–3348.
- Dayringer, H. E., Tramontano, A., Sprang, S. R., and Fletterick, R. J. (1986) INSIGHT: an interactive molecular graphics package. *J. Mol. Graphics* 4, 82–87.
- Finn, R. D., Tate, J., Mistry, J., Coghill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2008) The Pfam protein families database. *Nucleic Acids Res.* 36, D281–288.
- Anderson, M. A., Ogbay, B., Arimoto, R., Sha, W., Kisselev, O. G., Cistola, D. P., and Marshall, G. R. (2006) Relative strength of cation– π vs salt-bridge interactions: the Gtalpha(340–350) peptide/rhodopsin system. *J. Am. Chem. Soc.* 128, 7531–7541.
- Mehta, P. K., and Christen, P. (2000) The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Adv. Enzymol. Relat. Areas Mol. Biol.* 74, 129–184.
- Strickland, E. H. (1974) Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit. Rev. Biochem.* 2, 113–175.
- Schirch, L. (1975) Serine transhydroxymethylase. Relaxation and transient kinetic study of the formation and interconversion of the enzyme-glycine complexes. *J. Biol. Chem.* 250, 1939–1945.
- Contestabile, R., Angelaccio, S., Bossa, F., Wright, H. T., Scarsdale, N., Kazanina, G., and Schirch, V. (2000) Role of tyrosine 65 in the mechanism of serine hydroxymethyltransferase. *Biochemistry* 39, 7492–7500.

43. Bhavani, B. S., Rajaram, V., Bisht, S., Kaul, P., Prakash, V., Murthy, M. R., Appaji Rao, N., and Savithri, H. S. (2008) Importance of tyrosine residues of *Bacillus stearothermophilus* serine hydroxymethyltransferase in cofactor binding and L-allo-Thr cleavage. *FEBS J.* 275, 4606–4619.
44. Paddock, M. L., Weber, K. H., Chang, C., and Okamura, M. Y. (2005) Interactions between cytochrome c_2 and the photosynthetic reaction center from *Rhodobacter sphaeroides*: the cation– π interaction. *Biochemistry* 44, 9619–9625.
45. Luque, L. E., Grape, K. P., and Junker, M. (2002) A highly conserved arginine is critical for the functional folding of inhibitor of apoptosis (IAP) BIR domains. *Biochemistry* 41, 13663–13671.
46. Stover, P., Zamora, M., Shostak, K., Gautam-Basak, M., and Schirch, V. (1992) *Escherichia coli* serine hydroxymethyltransferase. The role of histidine 228 in determining reaction specificity. *J. Biol. Chem.* 267, 17679–17687.
47. Schirch, V., Shostak, K., Zamora, M., and Gautam-Basak, M. (1991) The origin of reaction specificity in serine hydroxymethyltransferase. *J. Biol. Chem.* 266, 759–764.
48. Rao, J. V., Prakash, V., Rao, N. A., and Savithri, H. S. (2000) The role of Glu74 and Tyr82 in the reaction catalyzed by sheep liver cytosolic serine hydroxymethyltransferase. *Eur. J. Biochem.* 267, 5967–5976.
49. Elcock, A. H. (2001) Prediction of functionally important residues based solely on the computed energetics of protein structure. *J. Mol. Biol.* 312, 885–896.
50. Ferreiro, D. U., Hegler, J. A., Komives, E. A., and Wolynes, P. G. (2007) Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19819–19824.