Location of the Pteroylpolyglutamate-binding Site on Rabbit Cytosolic Serine Hydroxymethyltransferase*

Tzu-Fun Fuû, J. Neel Scarsdale, Galina Kazanina, Verne Schirch, and H. Tonie Wright§

From the Department of Biochemistry and the Institute for Structural Biology and Drug Discovery, Virginia Commonwealth University, Richmond, Virginia 23219-1570

Serine hydroxymethyltransferase (SHMT; EC 2.1.2.1) catalyzes the reversible interconversion of serine and glycine with transfer of the serine side chain one-carbon group to tetrahydropteroylglutamate (H₄PteGlu), and also the conversion of 5,10-methenyl-H₄PteGlu to 5-formyl-H₂PteGlu. In the cell, H₂PteGlu carries a polyγ-glutamyl tail of at least 3 glutamate residues that is required for physiological activity. This study combines solution binding and mutagenesis studies with crystallographic structure determination to identify the extended binding site for tetrahydropteroylglutamate on rabbit cytosolic SHMT. Equilibrium binding and kinetic measurements of H₄PteGlu₃ and H₄PteGlu₅ with wild-type and Lys → Gln or Glu site mutant homotetrameric rabbit cytosolic SHMTs identified lysine residues that contribute to the binding of the polyglutamate tail. The crystal structure of the enzyme in complex with 5-formyl-H₂PteGlu₃ confirms the solution data and indicates that the conformation of the pteridine ring and its interactions with the enzyme differ slightly from those observed in complexes of the monoglutamate cofactor. The polyglutamate chain, which does not contribute to catalysis, exists in multiple conformations in each of the two occupied binding sites and appears to be bound by the electrostatic field created by the cationic residues, with only limited interactions with specific individual residues.

The polyglutamate forms of tetrahydropteroylglutamate (H₄PteGlu) serve as carriers of one-carbon groups that are required in the biosynthesis of purines, thymidylate, choline, methionine, and other important metabolites (1, 2). The circulating form of the coenzyme in blood is the monoglutamate derivative (n = 1), which is readily transported into cells. Once in the cell, additional glutamate residues are added to form H₄PteGlu₃-₇ in a reaction catalyzed by folypolyglutamate synthetase utilizing both ATP and glutamate as substrates (3). Only polyglutamate forms are physiologically active in the cell, although all but one enzyme involved in one-carbon metabolism utilize the monoglutamate derivative in vitro. Polyglutamylation blocks cellular export by multidrug resistance proteins of H₂PteGlu₃ and chemotherapeutic anti-folate analogs such as methotrexate and also increases H₂PteGlu₅ affinity for most enzymes in one-carbon metabolism (4).

The polyglutamate chain may also act as a tether in “channeling” the H₂PteGlu₅ coenzyme between enzymes involved in one-carbon metabolism (5, 6). The strongest evidence for channeling in folate enzymes is between the dehydrogenase and cyclohydrolase activities of the trifunctional enzyme C₅-tetrahydrofolate synthase (7, 8) and between the dihydrofolate reductase and thymidylate synthase sites on a bifunctional enzyme from *Plasmodium falciparum* (9–13). Kinetic evidence suggests that channeling of H₂PteGlu₅ may also be important for SHMT (14, 15).

There are now ~16 crystal structures of folate-requiring enzymes, but few of these address the binding determinants of the polyglutamate chain (16–21). Most of these structures have been determined without any bound folate or folate analog or with folate analogs lacking a polyglutamate chain or having only a single glutamyl residue. This deficiency is due to several technical constraints and may also be inherent in the nature of polyglutamylated H₂PteGlu₅ binding. Even in cases where polyglutamylated folates might be used, the enzymes are often crystallized in high salt, and the anions of these salts compete for binding with the anionic polyglutamate part of the coenzyme, which may cause weak or disordered binding of the polyglutamate chain. It is also possible that the polyglutamate chain occurs as multiple conformers whose occupancies are too low to be interpreted from an electron density map. The only two published crystallographic structures with a bound polyglutamyl chain are those for tetraglutamylfolate analogs bound to thymidylate synthase (16, 17). In the first of these structures (folate analog CB3717-Glu₄), there was unambiguous electron density for triglutamate in one active site and diglutamate in the other, but the group temperature factors for Glu₂ and Glu₃ in the first site were high. The failure to see the complete ligand in this structure may have been a consequence of the high ionic strength (I = 3.5 M) of the crystallization medium (2.5 M ammonium sulfate), which limited occupancy of the binding site. This inference is supported by the second thymidylate synthase-folate analog crystal structure (folate analog LY231514-Glu₄), in which Glu₃-Glu₄ are defined, but the elec-
tron density for Glu$_4$ is poorer. Crystals of this complex were obtained at lower ionic strength ($I \approx 2.0$ M).

SHMT catalyzes the conversion of serine and H$_4$PteGlu$_{n}$ to glycine and 5,10-methylene-H$_4$PteGlu$_{n}$, which is the major entry point of one-carbon groups into one-carbon metabolism (1).

An increase in glutamate chain length from 1 to 6 increases the affinity of H$_4$PteGlu for several mammalian SHMTs by orders of magnitude, but results in only a 2-fold increase in affinity for Escherichia coli SHMT (22, 23). Crystal structures have been determined for the cytosolic SHMTs of human (24), rabbit (25), and mouse (26) and for E. coli (27) and Bacillus steavorthermophilus (28) SHMTs. The human and rabbit SHMT structures have no bound folate ligand, whereas the E. coli and mouse SHMT structures are ternary complexes of the monoglutamate of 5-CHO-H$_4$PteGlu with glycine substrate. Crystal structures for bsSHMT include the enzyme with no bound ligands, with either serine or glycine bound, and the ternary complex of glycine and 5-CHO-H$_4$PteGlu. However, none of these structures defines the polyglutamate-binding site.

In addition to the conversion of serine to glycine and the concomitant transfer of the one-carbon group to H$_4$PteGlu, SHMT also catalyzes the physiologically important conversion of 5,10-methenyl-H$_4$PteGlu to 5-CHO-H$_4$PteGlu (29–31). The polyglutamate forms of the product 5-CHO-H$_4$PteGlu are slow tightly binding inhibitors of the physiological reaction catalyzed by SHMT, but the monoglutamate derivative binds rapidly (23). The polyglutamate forms of H$_4$PteGlu and 5-methyl-H$_4$PteGlu have affinity for the active site of rcSHMT comparable to 5-CHO-H$_4$PteGlu$_{n}$, but are not slow binding. Of all the reduced folate derivatives, only 5-CHO-H$_4$PteGlu$_{n}$ is stable to degradative oxidation, and this has permitted extensive spectroscopic binding and isothermal titration calorimetry studies of this compound with rcSHMT (23, 32).

In the work described here, we have measured solution equilibrium binding and kinetic constants for wild-type and structure-based site mutant tetrameric rcSHMTs with 5-CHO-H$_4$PteGlu$_{n}$ and H$_4$PteGlu$_{n}$ to define the binding site of the polyglutamate tail. We have also determined the crystal structure at 2.7 Å resolution of a complex of rcSHMT with active site-bound 5-CHO-H$_4$PteGlu$_{n}$ obtained by soaking unliganded crystals of rcSHMT in 5-CHO-H$_4$PteGlu$_{n}$, which is the major entry point of one-carbon groups into one-carbon metabolism (1). Site-directed mutagenesis was performed using a PCR-based method as previously described (34). Mutant enzymes were expressed from a recA strain of E. coli and purified as described for wild-type rcSHMT (35). $K_m$ and $V_{max}$ values with 1-serine as substrate were determined at pH 7.3 and 30 °C by coupling to the NADP-dependent oxidation of the product 5,10-methylene-H$_4$PteGlu (36). $K_m$ values for L-serine and glycine were determined by a spectrophotometric method and analyzed by the method of Scatchard as described (32). $T_m$ values were determined in a Microlab differential scanning calorimeter.

**EXPERIMENTAL PROCEDURES**

Preparation of Site Mutants, Purification of Expressed Enzymes, and Determination of Kinetic Constants—H$_4$PteGlu$_{n}$ were purchased from Schircks Laboratory (Jona, Switzerland). These compounds were reduced and converted to (6S)-H$_4$PteGlu, as previously described (33). (6S)-5-CHO-H$_4$PteGlu$_{n}$ was also made by this method. Site-directed mutagenesis was performed using a PCR-based method as previously described (34). Mutant enzymes were expressed from a recA strain of E. coli and purified as described for wild-type rcSHMT (35). $K_m$ and $V_{max}$ values with 1-serine as substrate were determined at pH 7.3 and 30 °C by coupling to the NADP-dependent oxidation of the product 5,10-methylene-H$_4$PteGlu (36). $K_m$ values for L-serine and glycine were determined by a spectrophotometric method and analyzed by the method of Scatchard as described (32). $T_m$ values were determined in a Microlab differential scanning calorimeter.

Crystallization and Diffraction Data—Crystals of rcSHMT were grown in 0.5 M l-Eppendorf tubes at room temperature from 10 ul drops of 10% PEG 1000, 0.1 M HEPES, pH 7.5, 10% 50% potassium MES (pH 7.0), 0.7 M, 0.4 mM potassium MES (pH 7.0) and flash-frozen in liquid N$_2$ for data collection. Data for a 10° sector were collected on an RAXIS II with Osmics confocal optics at 60 kV and 150 mA. Oscillation frames were integrated with Denzo and merged with Scalepack (38). Merged intensity data were converted to structure factor amplitudes using Truncate (39). Data collection statistics are summarized in Table I. Although the unliganded rcSHMT structure was previously assigned to space group P4$_2$2$_1$ with one dimer/asymmetric unit, in this work, we indexed these crystals in the lower symmetry P4$_1$, space group with one tetramer/asymmetric unit.

**Model Refinement**—Initial phases were based on a 2.1 Å structure of unliganded rcSHMT, which was nearly isomorphous with the rcSHMT-5-CHO-H$_4$PteGlu$_{n}$ structure. Initially, the space group was assigned as P4$_2$2$_1$, but after 20 cycles of rigid body refinement in CSS Version 1.0 (40), R$_{work}$ and R$_{free}$ were 35.8 and 35.6%, respectively. The resulting phases were used to calculate a SIGMAA-weighted 2mFo – $dF_c$ map that was visualized using the program O (41). In this initial map, broken density corresponding to 5-CHO-H$_4$PteGlu was observed in monomer B of the dimer in the asymmetric unit, but density was not observed for either of the two remaining glutamyl groups. The model was iteratively refined with high temperature torsion angle dynamics and a maximum likelihood target function (42) using CSS Version 1.0, with each refinement cycle followed by manual rebuilding in program O using SIGMAA-weighted 2mFo – $dF_c$ cross-validated maps or composite omit maps (43). Non-crystallographic symmetry was enforced via positional NCS restraints between monomers comprising the dimer. During iterative rebuilding, residue geometries and interactions were monitored with the programs OOPS (44), WHATCHET (45), and PROBES (46). We noted that for the 2.1 Å unliganded rcSHMT structure, for which diffraction data had been acquired under similar conditions, the value of R$_{work}$ near 24% in P4$_2$2$_1$, even though visual inspection of both SIGMAA-weighted and composite omit maps did not indicate significant errors in the model. By re-merging the integrated intensity data in the lower P4$_1$, symmetry space group and expanding the asymmetric unit to a pair of dimers with 222 symmetry, both R$_{work}$ and R$_{free}$ dropped significantly, which led us to conclude that the correct space group is P4$_1$. In the lower symmetry space group, additional NCS restraints were added between monomers in the two tight dimers comprising the asymmetric unit.

After several rounds of rebuilding and refinement, during which the weight of the NCS restraints was reduced from 300 to 37.5 kcal/mol and 390 water molecules and 3 glycerol molecules were added to the model, R$_{work}$ and R$_{free}$ improved to 29.4 and 22.0%, respectively. In the latter stages of the refinement, positional restraints were removed for regions of the structure that differed significantly between NCS-related monomers, and model adjustments were made in SIGMAA-weighted 2mFo – $dF_c$ maps. In the final cycles of simulated annealing via high temperature torsion angle dynamics, the positions of the 5-CHO-H$_4$PteGlu residues and side chain amino acids in partially disordered loops that interact with 5-CHO-H$_4$PteGlu, were refined via harmonic restraint errors in the refined structure were estimated using the formalism of Luazetti (47) as implemented in CNS. Refinement statistics are summarized in Table I. Of the 470 residues in the rcSHMT sequence, there is no interpretable electron density for the amino-terminal 13 residues. Also, as noted below, there is variable missing density in the different subunits for residues in the insert just after residue 244 and broken or missing localized density around residues 134, 355, and 399.

**Calculation of Surface Charge Distribution**—Electrostatic field calculations were carried out by numerical solution of the Poisson-Boltzmann equation using the program DELPHI as implemented in the INSIGHT II, San Diego, CA. The calculation, the protein formal charge set, which corresponds to unit charges on Lys, Arg, Glu, and Asp residues, was used. For the triglutamate chain of 5-CHO-H$_4$PteGlu$_{n}$, unit negative charges were assigned to the a-carboxyl groups of the glutamate residues and to the side chain carboxyl group for the terminal glutamate residue. In the rcSHMT structure, surface charge distributions for which density was not observed, as well as the surfaces of residues modeled as alanine due to missing side chain density, were added to the structure with the aid of a side chain rotamer data base (49) in O Version 8.0. The calculation was carried out on a cubic grid with 165 points along each axis, corresponding to a grid spacing of 3 Å. To investigate the effect of binding of the triglutamate chain on the surface charge distribution of the rcSHMT, the bound folate residues were excluded from a set of calculations.

$^2$ J. N. Scarsdale, G. Kazanina, V. Schirch, and H. T. Wright, unpublished data.
**RESULTS**

Kinetic Properties and Ligand Affinity of rcSHMT and Its Mutants—The surface of rcSHMT has a net positive charge at neutral pH (25), and we previously described a possible cationic binding locus for the extended polyglutamate tail of H₄PteGlu₅ emanating from the folate-binding site of each of the subunits (27). Both Lys → Gln or Glu and Arg → Gln or Glu point mutants in and near this channel were made, and the mutant proteins were purified and assayed with serine and H₄PteGlu₅ to determine specific activity, stoichiometry of H₄PteGlu₅ binding, and affinity with serine for both H₄PteGlu₅ (Table II). However, none of the residues tested plays a role in the putative open-to-closed transformation.

Crystal Structure of rcSHMT Crystals Soaked in 5-CHO-H₄PteGlu₅—Crystals of unliganded rcSHMT in space group P4₁, with a single homotetramer of 54-kDa monomers/asymmetric unit were used in the soaking experiment. The asymmetric unit of these crystals consists of two tight dimers, A and B, with a large anion that has been shown not to inhibit substrate binding. Numerous screenings of conditions did not yield any crystals of the preformed mono- or polyglutamyl rcSHMT-5-CHO-H₄PteGlu₅ complex.

We previously described an inferred binding locus in rcSHMT for the polyglutamate chain of H₄PteGlu₅ based on the structures of the eSHMT ternary complex and unliganded rcSHMT (25, 27). The position of the 5-CHO-H₄PteGlu₅ inhibitor in the rcSHMT structure, subsequently confirmed in mocSHMT and bsSHMT crystal structures (26, 28), defines the binding site of the pteridine ring, the p-aminobenzoic acid (PABA) group, and the first glutamate of the polyglutamate tail. The computed surface charge distribution of the rcSHMT tetramer shows a positively charged channel emanating from the folate-binding site of each of the subunits (27), and with this as a guide, we constructed a hypothetical model for the binding of the extended polyglutamate tail of H₄PteGlu₅. This hypothetical model was the basis for identifying residues that could interact with the extended polyglutamate chain.

The SIGMAA-weighted 3mFo – 2dFc electron density map of the soaked rcSHMT-5-CHO-H₄PteGlu₅ complex showed density with several breaks in this putative H₄PteGlu₅ site at the 0.95σ level in two (subunits B and D) of the four crystallographically independent monomers of the rcSHMT tetramer.
that constitute the tetrameric asymmetric unit. This density in one subunit of each tight dimer coincided with the site at which 5-CHO-H4PteGlu binds in the eSHMT:5-CHO-H4PteGlu-glycine and mocSHMT:5-CHO-H4PteGlu-glycine ternary complex structures and extends into the cationic channel hypothesized to be the binding site of the polyglutamate tail (Fig. 1). The electron density in the putative polyglutamate-binding site could be fit with a triglutamyl chain bound in two distinct conformations in each of subunits B and D (see below). In subunit B, there are alternate conformations for Glu3 of 5-CHO-H4PteGlu3, and in subunit D, there is bifurcation from Glu1 to give two alternate conformations of Glu2–Glu3 (Fig. 2).

Subunits A and C (but not B and D) had electron density in the amino acid substrate-binding site contiguous with the PLP ring and consistent with the presence of a glycine bound as a gem-diamine. No exogenous glycine was added either to the enzyme solution prior to crystallization or to the crystals soaked in 5-CHO-H4PteGlu3, and we surmise that adventitious glycine was present in some component of the sample. In folate-bound subunits B and D, there was a single peak of electron density in the amino acid-binding site, but it was not connected to C-4' of the PLP ring and was assigned as a solvent molecule.

Interactions of Bound 5-CHO-H4PteGlu with rcSHMT—The binding of 5-CHO-H4PteGlu3 did not cause any large-scale changes in the structure of the enzyme. The molecular interactions between the pteridine and PABA portions of 5-CHO-H4PteGlu in the soaked rcSHMT crystals were similar (but not identical) to those observed in the mocSHMT and eSHMT structures that were grown in the presence of 5-CHO-H4PteGlu. Electron density for the pteridine ring in subunit B is broken, with C-7 missing; in subunit D, the tetrahydrodpyr-azine N-5, C-6, and C-7 positions have weak or missing density, and the edges of the PABA ring are shaved. The C-6–C-9 bond is axial to the formyltetrahydroprazine of the pteridine so that the PABA ring is nearly perpendicular to the pteridine ring. The formyltetrahydroprazine ring of the pteridine is in the half-chair form, as is observed in the ternary complexes of eSHMT and mocSHMT with 5-CHO-H4PteGlu; but the broken density in this ring indicates conformational flexibility not seen in the structures of the other complexes. The formyl group in the complex described here is oriented roughly perpendicular to the pseudoplanar of the formyltetrahydroprazine ring, resembling that of the mocSHMT ternary complex and differing from that of the eSHMT ternary complex, in which the formyl oxygen is oriented diatomically away from O-4 of the pteridine ring. The formyl oxygen of 5-CHO-H4PteGlu3 in subunits B and D makes hydrogen bonds with the solvent in the amino acid-binding site. In subunit B, the carboxylate of Glu3 also makes a long hydrogen bond with this formyl oxygen, but this distance is too long in subunit D for a hydrogen bond.

The side chain amide of Asn347 is not visible in subunit B and is broken in subunit D, but is well defined in subunits A and C. The hydrogen bond contacts of Asn347 with the pteridine ring vary in three of the four complexes for which structures are available. In the eSHMT and bsSHMT ternary complexes, the Asn347 side chain makes a double hydrogen bond with N-1 and N-8 of the pteridine, whereas in the mocSHMT ternary complex, this side chain makes two different sets of contacts, one with N-1 and the exocyclic amino group of the pteridine and the other with the exocyclic amino group alone. The disorder in the buried Asn347 residue in subunits B and D of the rcSHMT:5-CHO-H4PteGlu3 complex, but not in unliganded subunits A and C, may be transmitted from the adjacent flexible loop around position 355, which, in subunits B and D, interacts with the polyglutamate chain.

Other hydrogen bonds between the pteridine and the protein observed in the ternary complex structures of mouse and E. coli

![Fig. 1. Surface charge distribution of the tetrameric rcSHMT complex showing 5-CHO-H4PteGlu3 (green space filling) extending into the cationic region centered on intersecting 2-fold rotation axes at the center of the rcSHMT homotetramer. Blue is + and red is − charge field.](http://www.jbc.org/Downloaded from)

**TABLE II**

| rcSHMT | H4PteGlu | H4PteGlu2 | H4PteGlu3 | H4PteGlu5 |
|--------|---------|----------|----------|----------|
| Wild-type | 10 (1.0) | 6.0 (1.0) | 0.7 (1.0) | 0.4 (1.0) |
| K1341Q | 26 (2.0) | 8.0 (1.3) | 10 (14) | 0.8 (2.0) |
| K1342Q | 100 (10) | 12 (2.0) | 7.0 (10) | 0.9 (2.0) |
| K346Q | 50 (5.0) | 10 (1.7) | 3.0 (4.3) | 4.0 (10) |
| K354Q | 21 (2.0) | 4.0 (0.7) | 0.8 (1.0) | 0.3 (0.8) |
| K399Q | 11 (1.1) | 7.0 (1.2) | 6.0 (8.5) | 12 (30) |
| K399Q | 13 (1.3) | 5.0 (0.8) | 2.0 (3.0) | 5.4 (13) |

| rcSHMT mutations | Kd (μM) | Kcat (μM) | Km (μM) |
|------------------|---------|---------|---------|
| K335Q, R3963Q, K3994E, K3994Q, K401E, K401Q, K405E, K405Q, H406D, H406N, R407E, R407Q, R412E, R412N, and Loop deletion (residues 244–249 deleted) | 26 (2.6) | 10 (1.7) | 3.0 (4.3) |

* Ratios of the Km values of the mutant enzymes compared with the wild-type enzyme are shown in light-face.
* Ratios of Kd values that exceed a factor of 2 are shown in boldface.
structures are similar in the soaked rcSHMT complex, although there are variations in angle and bond length. The PABA ring is approximately parallel with Tyr64 of the opposing subunit of the tight dimer (subunits A and C), as observed in the structures of the eSHMT and mocSHMT ternary complexes. Although the location of 5-CHO-H4PteGlu3 in this soaked crystal structure is the same as that in the eSHMT, bsSHMT, and mocSHMT ternary complexes with glycine, the small differences noted above and the incomplete electron density for the pteridine indicate that the cofactor is not optimally bound and is present in more than one conformation.

Structure of the Polyglutamate Chain in the rcSHMT Complex—Omit map electron density (3mFo - 2Fc) contoured at the 0.95σ level for the 5-CHO-H4PteGlu3 ligand in subunit B with several proximal binding site residues.

In subunit B, the α-carboxylate of Glu2 is oriented toward the side chains of LysB1341 and LysB1342, the interaction with LysB1341 being the tighter one. Neither carboxylate of Glu3 in both conformers in subunit B is close enough to any groups in the enzyme to make strong interactions. In subunit D, both the α- and γ-carboxylates of Glu3 of one conformer can interact with LysD1341 and LysD1342. In the other conformer, either the α- or γ-carboxylate can interact with LysD1346. These differences in interactions of the polyglutamyl chain in subunits B and D are the result of differences in the conformation of the triglutamyl group and also of small differences in the loop conformations around residues 134 and 356 in these two subunits. The electron density in both of these loops in subunits B and D is broken and apparently less ordered than their counterparts in subunits A and C. In the unpublished 2.1-Å resolution electron density map of unliganded rcSHMT, these loops are closely similar and well ordered in all four subunits, in contrast to the four loops in the 5-CHO-H4PteGlu3 complex structure, two of which (subunits A and C) are well ordered and two of which (subunits B and D) are disordered.

The half-site binding of 5-CHO-H4PteGlu3 in the crystal structure of the soaked complex is consistent with the equilibrium binding studies showing half-site occupancy of H4PteGlu3-bound rcSHMT for all mutants and all congeners of the folate. This strengthens the inference that the form of the enzyme in the crystal is similar (if not identical) to that on
which the solution measurements were made. The lattice packing presents no obvious obstruction of access to ligands diffusing into the active sites on all four subunits. When both glycine and 5-CHO-H₄PteGlu₃ were included in the soak, the crystals cracked, most likely due to the enzyme going from an open-to-closed structure.

Comparison of the SHMT-5-CHO-H₄PteGlu₃ Complex with Related Structures—In the mocSHMT ternary complex structure with 5-CHO-H₄PteGlu and glycine, all subunits have a glycine conjugated to PLP. One tight dimer binds 5-CHO-H₄PteGlu in both subunits (A and B), whereas the other dimer binds 5-CHO-H₄PteGlu in only one subunit and in a slightly different orientation and at a lower occupancy than the other tight dimer (26). Although the resolution of this structure may not be sufficient to draw detailed conclusions about local stereochemistry, the Protein Data Bank entry shows the linkage of the glycine to the PLP group in each subunit to be variable, suggesting the coexistence of different intermediates in the tetramer. There also appears to be some variation between the 2 inferred glycine residues bound to PLP in each subunit to be variable, supporting the coexistence of different intermediates in the tetramer. There also appears to be some variation between the 2 inferred glycine residues bound to PLP in each subunit to be variable, supporting the coexistence of different intermediates in the tetramer. There also appears to be some variation between the 2 inferred glycine residues bound to PLP in each subunit to be variable, supporting the coexistence of different intermediates in the tetramer. There also appears to be some variation between the 2 inferred glycine residues bound to PLP in each subunit to be variable, supporting the coexistence of different intermediates in the tetramer.

There is breakdown of the 222 symmetry of rcSHMT in crystals soaked in 5-CHO-H₄PteGlu₃, although it differs from that observed in the mocSHMT ternary complex. The soaked rcSHMT-5-CHO-H₄PteGlu₃ complex shows degraded pseudo-symmetry about two of the three local dyads due to the differential ligand binding to subunits B and D versus subunits A and C. The third local dyad relating the two tight dimers retains the symmetry of ligand binding, but because the ligands and the loops around residues 134 and 356 have different conformations, this operation also deviates from the near-perfect symmetry of the unliganded structure.

DISCUSSION

Correlation of Crystal Structure and Solution Studies of rcSHMT Binding of 5-CHO-H₄PteGlu₃—In this work, we used the formyl derivative of H₄PteGlu₃ because of its high stability and comparable affinity to H₄PteGlu₃ in binding SHMT. In solution, the pteridine ring of 5-CHO-H₄PteGlu is reported to be in a half-chair conformation with the C-6–C-9 bond axial to the formyltetrahydropyrazine ring (52, 53), whereas H₄PteGlu exists as a roughly equal mixture of two half-chair conformations with C-6–C-9 either axial or equatorial (54). However, the data reported in two NMR studies (52, 53) on the conformation of 5-CHO-H₄PteGlu in solution are not in complete agreement.
and are consistent with more than one conformation of the pteridine ring. Such polymorphism is also suggested by the pattern of broken density in the formyltetrahydropterin ring in both subunits B and D of the soaked rcSHMT-5-CHO-H₄PteGlu₃ complex structure described here. As found in other enzymes that bind reduced folates, the C-6–C-9 bond is axial in eSHMT, bsSHMT, rcSHMT, and mocSHMT.

The binding of the pteridine and PABA portions of 5-CHO-H₄PteGlu₃ to preformed rcSHMT crystals occurs at the same site as that observed in the crystal structures of the eSHMT, bsSHMT, and mocSHMT ternary complexes with glycine (26, 27). However, the enzyme ligands to NH₂, C-2, N-3, O-4, and the N-5 formyl group of the pteridine ring are similar (but not identical) in the three structures. The electron density of the complex described here shows suboptimal hydrogen bonding for some of the pteridine donors and acceptors and weak electron density for N-3 and the tetrahydropterin ring of the pteridine and for the side chains of AsnB347 and AsnD347 in the folate-occupied subunits. The mocSHMT ternary complex also shows variable and, in some cases, suboptimal hydrogen bonding to the pteridine ring in the three liganded subunits. In the eSHMT and bsSHMT complexes with 5-CHO-H₄PteGlu, the side chain amide of AsnB347 makes two hydrogen bonds with N-1 and N-8 of the pteridine ring, and all hydrogen-bonding groups have good stereochemistry. This suggests that, in the current structure, 5-CHO-H₄PteGlu₃ is not optimally aligned at the active site of rcSHMT and that the disorder in the side chain of AsnB347 in subunits B and D is due to the presence of 5-CHO-H₄PteGlu₃, since these side chains have clear density in unliganded subunits A and C.

Solution studies suggest that the slow step in binding 5-CHO-H₄PteGlu₃ with glycine to SHMT requires a conformational change that occurs late in the binding process (23). Because the crystals of the soaked complex of 5-CHO-H₄PteGlu₃ with rcSHMT are isomorphous with the unliganded native crystals, lattice forces or the absence of glycine is likely to have blocked this putative conformational change in the soaked complex described here, and some of the observed poorly formed enzyme-ligand interactions of this complex may become optimal upon transition through this conformational change. This conformational change may account for the crystals cracking when glycine was included in the soak. It should be noted that, in the mocSHMT ternary complex, Asn347 makes variable and generally poorer hydrogen bonds with the pteridine ring than is observed in the “closed” eSHMT and bsSHMT ternary complexes, suggesting that, in the mocSHMT structure, 5-CHO-H₄PteGlu₃ may also not be optimally bound and may not, as thought, be in the closed conformation (26). The observations in the rcSHMT-5-CHO-H₄PteGlu₃ complex that the formyl group of 5-CHO-H₄PteGlu₃ is oriented away from O-4 of the pteridine and that the PABA ring is stacked and parallel to Tyr⁶⁴⁹, as found in the other structures, support the current structure as being on the pathway to the final structure characterized in our solution studies. Substitution of Asn347 with Ala in eSHMT essentially abolished the affinity for H₄PteGlu₃, demonstrating that the interactions of this side chain with the pteridine are essential for the optimal folate binding assumed in the productive reaction. Ligand-induced flexibility in loop 355 in an early complex of the SHMT reaction may be coupled to limited disorder of the pteridine-binding groups of the active site and to multiple conformations of the tetrahydropterin ring that are necessary to realize productive binding of a single conformer in a subsequent reaction step. A pre-binding folate site has also been proposed for a folate-soaked complex of thymidylate synthase (55).

Structural and Site Mutant Studies of the Binding of the Polyglutamate Chain—Two studies have previously used chemical modification to locate the folate-binding site on SHMT. In one study (32), 5-CHO-H₄PteGlu₃ and 5-CHO-H₄PteGlu₅ were activated by a carbodiimide, and both were found to cross-link uniquely to Lys⁸⁹⁹⁴ and to block further binding of folates. In the second study (56), modification of the guanidinium groups of 2 Arg residues in sheep liver SHMT, equivalent to Lys⁸⁹⁹⁴ and Arg⁸⁹¹² in rcSHMT, blocked H₄PteGlu₅ binding. These 3 residues are near the polyglutamate-binding site, but only Lys⁸⁹⁹⁴, in the structure described here could possibly interact with the polyglutamate tail of folate and only with Glu₃ of H₄PteGlu₃. Solution studies of site mutants also did not implicate these residues in folate binding (Table II). Cross-linking in solution of Lys⁸⁹⁹⁴ to the activated polyglutamate chain most likely arises from collisions of the entering, nonspecifically bound folate ligand with Lys⁸⁹⁹⁴, whose main chain and side chain are flexible.

The site mutants that affect the affinity of H₄PteGlu₃ are closely consistent with the crystal structure. LysB1341 and LysB1342 could interact with the γ-carboxylate of Glu₃ in subunit B and with the Glu₃ carboxylates of one conformer in subunit D. These lysines could also hydrogen-bond to the carbonyl oxygen of the Glu₃ peptide bond in polyglutamate chain lengths greater than 1. H₄PteGlu₃ and H₄PteGlu₅ showed higher Kᵣ values when either of these residues was mutated. LysD346 makes a close approach to both carboxylates of Glu₅ in subunit D and could approach Glu₃ in subunit D if LysD346 adopted a different side chain conformation. Mutation of LysB1345 similarly increased the Kᵣ values for H₄PteGlu₃ and H₄PteGlu₅. LysB1341 and LysB1342 approach Glu₅ in subunit B, but mutation of any of the lysines only weakly affected the Kᵣ for the binding of 5-CHO-H₄PteGlu₃ and H₄PteGlu₃. This insensitivity to the loss of the positive charges at LysB1341, LysB1342, or LysD346 for the diglutamate form of the folate may arise from alternative interaction(s) between the γ-carboxylate group of Glu₃ and these residues that are absent in the mono- and triglutamate forms of 5-CHO-H₄PteGlu. Extrapolation of the 5-CHO-H₄PteGlu₃ structure to Glu₃ and Glu₅ would lead in both sites to interactions with Lys⁸⁹⁹⁴ and Lys⁸⁹⁹⁴, mutation of which increased the Kᵣ for H₄PteGlu₃ and H₄PteGlu₅. A longer polyglutamate chain might also interact with the region in which LysB1341, LysB1346, and LysD346 and LysD358 also the more distal Arg²⁴⁶³ lie. Isotothermal titration calorimetry data show that the binding of 5-CHO-H₄PteGlu₃ is entropy-driven, most likely by displacement of ordered water from the pteroyl-binding site (32). In the 2.1-Å structure of unliganded rcSHMT, there are ordered solvent molecules in this site. The isothermal calorimetry titration experiments further suggest that the binding of the polyglutamate chain occurs with a negative enthalpy and little change in entropy. This suggests that the polyglutamate chain does not make specific interaction with groups on the enzyme, but binds electrostatically in multiple conformations.

This inference is consistent with the crystal structure of 5-CHO-H₄PteGlu₃ complexed with rcSHMT, where only a small number of the potential interactions between cationic groups on the enzyme and anionic polyglutamyl carboxylates are actually made. Other stabilizing electrostatic interactions could be made with relatively minor changes in the orientation of lysine and arginine side chains on the enzyme. These observations parallel those made on the binding of tetraglutamylfolate analogs to thymidylate synthase and the conclusions drawn from those structures (16, 17). In those two structures, the binding mode of the polyglutamate chain is very similar.

---

4 S. Angelaccio and R. Contestabile, personal communication.
but few (if any) direct hydrogen bonds are made between the tetraglutamyl chain and groups on the thymidylate synthase. The failure to form specific interactions between the polyglutamyl carboxylates and accessible groups on the enzymes in these two complexes is consistent with the proposal that this binding is driven by electrostatic attraction of the polyanion for a polycationic site that has been identified at the active sites of both enzymes. In the 5-CHO-H4PteGlu3 complex with SHMT, we observed multiple conformations, which were not observed in the thymidylate synthase complex. The thymidylate synthase crystals were in a high ionic strength solution, which may have left only the most stable of multiple complex conformations in the crystal structure.

The occupancy by 5-CHO-H4PteGlu3 of subunits B and D (but not A and C) is consistent with the calorimetric observation that the stoichiometry of folate binding is 0.5/subunit. The binding of 5-CHO-H4PteGlu3 to subunits B and D disperses the two poly glutamyl tails on opposite faces of the tetramer (Fig. 2), thereby maximizing the distance between extended polyglutamyl chains. This arrangement retains more of the cationic character of each of the two tetramer surfaces centered at the pseudo-222 intersection than if the ligands were bound pairwise to subunits A and D or B and C, where the polyglutamate chains would lie on the same cationic face. It is noteworthy that the charge surface calculated with the 5-CHO-H4PteGlu3 ligand bound (data not shown) is significantly altered, having diminished cationic character at some distance from the ligand itself.

Comparison of rcSHMT with SHMT Proteins in Other Studies—The stoichiometry and subunit occupancy for folate binding are different for each of the four SHMT structures that have been determined with bound folate. In eSHMT, both monomers of the homodimer contain bound 5-CHO-H4PteGlu, whereas in bsSHMT, one monomer binding site of the homodimer is fully occupied by 5-CHO-H4PteGlu, but the other is only partially occupied. Tetrameric rcSHMT has 2 molecules of bound 5-CHO-H4PteGlu/tetramer, one in each tight dimer, whereas moSHMT has 5-CHO-H4PteGlu bound to both subunits of one tight dimer and only one subunit of the other tight dimer. No obvious reason for these differences is evident from the structures.

SHMT belongs to the fold type I family of PLP enzymes, which have a core amino acid sequence of ~45 kDa that is conserved in the >50 SHMT sequences available (57, 58). The amino acid- and PLP-binding sites are highly conserved in all SHMTs, and the canonical SHMT structure supports the inference that the folate interaction site evolved from a fold type I PLP precursor enzyme through sequence insertions and not by domain swapping from other folate-requiring enzymes (27). This may be the reason why anti-folate compounds developed as chemotherapeutic agents are ineffective as inhibitors of SHMT and suggests that an effective anti-folate inhibitor of SHMT might not inhibit other folate enzymes.

Previous studies have shown that increasing the chain length of H4PteGlu from 1 to 5 lowers the $K_p$ by ~2 orders of magnitude for mammalian rcSHMTs. In E. coli, only the first 2 glutamate residues are linked through the $\gamma$-carboxylate group, additional glutamatines being linked through the $\alpha$-carboxylate (59). The affinities of cognate monoglumylfolates versus pentaglutamylfolates (with the $\alpha$-linkage for the last 2 glutamate residues) for eSHMT differ by only a factor of 2, indicating that eSHMT does not have a well defined polyglutamate-binding site.5,6 There are several insertions of amino acids in the mammalian enzymes compared with eSHMT, three of which are located in the polyglutamate-binding site. The insertion at position 134 introduces a KKK sequence that makes interactions with Glu5 of SHMT, suggesting that an effective anti-folate inhibitor of SHMT and suggests that a role in either channeling of the reduced folate from other folate-binding enzymes or guiding the folate substrate into the active site, without directly affecting the binding. In contrast to PLP- amino acid- and pteridine-binding sites of SHMTs, which are highly conserved, the residues identified in this work that interact with the polyglutamate tail of H4PteGlu3 are less conserved in many SHMTs. This suggests that other domains may interact with the polyglutamate chain or have different binding loci in the enzymes from different organisms.

REFERENCES

1. Schirch, V. (1984) in Folates and Pterins: Chemistry and Biochemistry of Folates (Blakely, R. L., and Benkovic, S. J., eds) Vol. 1, pp. 399–432, John Wiley and Sons, Inc., New York
2. Schirch, V., and Strong, W. B. (1991) J. Biol. Chem. 266, 371–380
3. Sanghani, P. C., and Moran, R. G. (2000) Protein Expression Purif. 18, 36–45
4. Zeng, H., Chen, Z. S., Belinsky, M. G., Rea, P. A., and Kruh, G. D. (2001) Cancer Res. 61, 7225–7232
5. Orandi, J., and Stree, P. A. (2000) Int. Rev. Cytol. 192, 255–280
6. Huang, X., Holden, H. M., and Rauschel, F. M. (2001) Annu. Rev. Biochem. 70, 149–160
7. Pawelek, P. D., Aliaire, M., Cygler, M., and MacKenzie, R. E. (2000) Biochim. Biophys. Acta 1470, 59–68
8. Kohls, D., Sulea, T., Purisma, E. O., MacKenzie, R. E., and Vrieilink, A. (2000) Struct. Fold. Des. 8, 35–46
9. Stroud, R. M. (1994) Nat. Struct. Biol. 1, 131–134
10. Knighton, D. R., Kan, C.-C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M., and Matthews, D. A. (1994) Nat. Struct. Biol. 1, 186–194
11. Trujillo, M., Duncan, R., and Santi, D. V. (1997) Protein Eng. 10, 567–573
12. Trujillo, M., Donald, R. G. K., Roos, D. S., Greene, P. J., and Santi, D. V. (1996) Biochemistry 35, 6366–6374
13. Leguay, P. H., and Anderson, S. F. (1998) Biochemistry 37, 12195–12205
14. Strong, W. B., and Schirch, V. (1989) Biochemistry 28, 9430–9439
15. Kim, D. W., Huang, T., Schirch, D., and Schirch, V. (1996) Biochemistry 35, 15772–15783
16. Kamb, A., Finer-Moore, J., Calvert, A. H., and Stroud, R. M. (1992) Biochemistry 31, 9883–9890
17. Fiser, J., Maley, G. F., Maley, F., Montfort, W. R., and Stroud, R. M. (1994) Biochemistry 33, 15459–15468
18. Sayre, P. H., Fiser, J. S., Fritz, T. A., Biermann, D., Gates, S. B., MaxKellar, W. C., Patel, V. F., and Stroud, R. M. (2001) J. Mol. Biol. 313, 813–829
19. Klein, C., Chen, P., Arevalo, J. H., Stura, E. A., Marlewiiski, A., Warren, M. S., Benkovic, S. J., and Wilson, I. A. (1995) J. Mol. Biol. 249, 153–175
20. Shen, B. W., Dyer, D. H., Hung, J.-Y., D’Ari, L., Rabinowitz, J., and Stodard, B. L. (1999) Protein Sci. 8, 1342–1349
21. Phan, J., Seli, M., Minor, W., Dunlap, R. B., Berger, S. H., and Lebioda, L. (2001) Biochemistry 40, 1897–1902
22. Matthews, R. G., Ross, J., Baugh, C. M., Cook, J. D., and Davis, L. (1982) J. Mol. Biol. 81, 829–829
23. Stover, P., and Schirch, V. (1991) J. Biol. Chem. 266, 1543–1550
24. Renwick, S. B., Snell, K., and Baumann, U. (1998) Structure 6, 1105–1116
25. Scarsdale, J. N., Karazina, G., Radev, S., Schirch, V., and Wright, H. T. (1999) Biochemistry 38, 8347–8358
26. Scarsdale, J. N., Karazina, G., Radev, S., Schirch, V., and Wright, H. T. (2000) J. Mol. Biol. 296, 125–138
27. Trivedi, V., Gupta, A., Jala, V. R., Saravanam, P., Rao, G. S. J., Rao, N. A., Savithri, H. S., and Subramanya, H. S. (2002) J. Biol. Chem. 277, 17161–17169
28. Stover, P., and Schirch, V. (1990) J. Biol. Chem. 265, 14227–14233
29. Girgis, S., Suh, J. R., Jolivet, J., and Stover, P. J. (1997) J. Biol. Chem. 272, 4729–4734
30. Holmes, W. B., and Appling, D. R. (2002) J. Biol. Chem. 277, 20205–20213
31. Huang, T., Wang, C., Maras, B., Barra, D., and Schirch, V. (1998) Biochemistry 37, 15358–15362
32. Stover, P., and Schirch, V. (1992) Anal. Biochem. 202, 82–88
33. Iresecia, S., Condo, I., Angelacchio, D., Delle Fratte, S., and Bossa, F. (1996) Protein Expression Purif. 7, 323–328
35. di Salvo, M. L., Delle Fratte, S., De Biase, D., Bossa, F., and Schirch, V. (1998) *Protein Expression Purif.* **13**, 177–183
36. Schirch, V., Shostak, K., Zamora, M., and Gautam-Basak, M. (1991) *J. Biol. Chem.* **266**, 759–764
37. Kazanina, G., Radaev, S., Wright, H. T., and Schirch, V. (1998) *J. Struct. Biol.* **123**, 169–174
38. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
39. Collaborative Computing Project No. 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
40. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., and Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
41. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* **277**, 173–208
42. Pannu, N. S., and Read, R. J. (1996) *Acta Crystallogr. Sect. A* **52**, 659–668
43. Bhat, T. N. (1988) *J. Appl. Crystallogr.* **21**, 279–281
44. Kleywegt, G., and Jones, T. A. (1996) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **52**, 829–832
45. Hoofi, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature* **381**, 272
46. Werd, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S., and Richardson, D. S. (1999) *J. Mol. Biol.* **285**, 1711–1733
47. Luzatti, P. V. (1952) *Acta Crystallogr.* **5**, 802–810
48. Vaguine, A. A., Richelle, J., and Wlodak, S. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 191–205
49. Ponder, J. W., and Richards, F. M. (1987) *J. Mol. Biol.* **193**, 775–779
50. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng.* **8**, 127–134
51. Schirch, V. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., ed.) Vol. I, pp. 211–252, Academic Press Ltd., London
52. Poe, M., and Benkovic, S. J. (1980) *Biochemistry* **19**, 4576–4582
53. Feeney, J., Albrand, J. P., and Boicelli, C. A. (1980) *J. Chem. Soc. Perkin Trans. 1* **2**, 176–180
54. Poe, M., and Hugsteen, K. (1978) *J. Biol. Chem.* **253**, 543–546
55. Birdsall, D. L., Finer-Moore, J., and Stroud, R. M. (1996) *J. Mol. Biol.* **255**, 522–535
56. Usha, R., Savithri, H. S., and Rao, N. A. (1992) *J. Biol. Chem.* **267**, 9289–9293
57. Pascarella, S., Schirch, V., and Bossa, F. (1993) *FEBS Lett.* **331**, 145–149
58. Grishin, N. V., Phillips, M. A., and Goldsmith, E. J. (1995) *Protein Sci.* **4**, 1291–1304
59. Ferone, R., Handin, M. H., Singer, S. C., and Hunt, D. F. (1986) *J. Biol. Chem.* **261**, 16356–16362
