ELSEVIER

Biophysical Chemistry 115 (2005) 1–10

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

A three-dimensional structure of *Plasmodium falciparum* serine hydroxymethyltransferase in complex with glycine and 5-formyl-tetrahydrofolate. Homology modeling and molecular dynamics

Tanos C.C. França[a], Pedro G. Pascutti[b], Teodorico C. Ramalho[a], José D. Figueroa-Villar[a,*]

[a]*Departamento de Química, Instituto Militar de Engenharia, Praça General Tibúrcio 80-Urca, 22290-270 Rio de Janeiro, Brazil*
[b]*Instituto de Biofísica Carlos Chagas Filho, CCS. bloco G, sala 29, Av Brigadeiro Trompowski s/n°-Ilha do Fundão-UFRJ, 21949-900 Rio de Janeiro, Brazil*

Received 3 August 2004; received in revised form 29 November 2004; accepted 2 December 2004
Available online 15 December 2004

Abstract

Cytosolic *Plasmodium falciparum* serine hydroxymethyltransferase (pfSHMT) is a potential target for antimalarial chemotherapy. Contrasting with the other enzymes involved in the parasite folate cycle, little information is available about this enzyme, and its crystallographic structure is unknown yet.

In this paper, we propose a theoretical low-resolution 3D model for pfSHMT in complex with glycine and 5-formyl tetrahydrofolate (5-FTHF) based on homology modeling by multiple alignment followed by intensive optimization, validation and dynamics simulations in water. Comparison between the active sites of our model and that of crystallographic Human SHMT (hSHMT) revealed key differences that could be useful for the design of new selective inhibitors of pfSHMT.

Keywords: Malaria; *Plasmodium falciparum*; pfSHMT; Homology modeling; Molecular dynamics

1. Introduction

The so-called parasitic diseases affect today a great part of the world population, causing many deaths and having a great limiting influence in the life quality and development of several countries, especially in the tropics. These diseases are usually caused by complex organisms, mainly worms and protozoa.

Protozoa are responsible for most parasitic infections affecting human beings. Protozoonoses are caused by about 10.000 known species of protozoa and are most common at the less developed regions of the planet, being usually endemic. The most lethal protozoa disease in humans is malaria [1], which affects about 300 and 500 million people worldwide, causing between 1.0 and 2.5 million deaths annually, mostly among children. Today, the threat is more significant for nations in Africa, Asia, Latin America and some regions of the South Pacific. It is estimated that approximately 40% of the world population is at risk of infection with malaria [2]. This disease is caused by protozoa of the genus *Plasmodium*. In humans, four species are responsible for Malaria: *Plasmodium falciparum* (the most dangerous), *P. vivax*, *P ovale* and *P. malariae.*

Two aspects have currently stimulated new efforts regarding the development of chemotherapy, vaccines and sanitary studies about malaria: the rapid emergence of *P. falciparum* strains which are resistant to currently available antimalarial drugs [3–6] and the inefficacy of antimalarial vaccines [7]. Since a vaccine for malaria seems to be far in the future, a more immediate solution would be the development of new antimalarial chemotherapy. Still, we have to assume that the ability of *Plasmodium* to adapt to new chemotherapy under the pressure of the drugs would eventually render any new antimalarial drugs less efficient, thus making all research on aspects of malaria even more

* Corresponding author. Tel.: +55 21 25467057; fax: +55 21 25467059.
E-mail address: figueroa@ime.eb.br (J.D. Figueroa-Villar).

0301-4622/$ - see front matter © 2005 Elsevier B.V. All rights reseved.
doi:10.1016/j.bpc.2004.12.002

important. Many strategies have been adopted to develop new drugs for malaria. Among them, we have been interested on the design of new inhibitors for three enzymes involved in the parasite methylenetetrahydrofolate cycle using molecular modeling and dynamics studies. It is well known that inhibitors of folate metabolism are quite important drugs, not only in the chemotherapy of malaria, but also of bacterial infections and cancer [8]. The effectiveness of antifolates is based on the perturbations they cause in the folate pathways, which rapidly lead to nucleotide imbalances and cell death [9,10]. This turns the enzymes involved in this cycle into good targets to antimalarial chemotherapy. These three enzymes are thymidylate synthase (TS), dihydrofolate reductase (DHFR) and serine hydroxymethyltransferase (SHMT). Among those enzymes, DHFR have been the main target for antimalarial chemotherapy mainly because this enzyme shows important active site differences for different species. For example, the similarity between the human and the *P. falciparum* DHFR active sites is only 45%, a fact that allows for the development of antifolates, which are very selective towards the parasite enzyme. We believe that this variability on the primary sequence of DHFR for different species is a clear indication of the potential of this enzyme to suffer important mutations in its active site without loss of its activity. In fact, those mutations are responsible for the widespread resistance of *P. falciparum* to antifolates.

The fact that all protozoa, including *Plasmodium*, have the enzymes DHFR and TS as parts of a bifunctional homodimeric enzyme, in contrast with mammalians that have these enzymes as monofunctional separate proteins, also opens a new perspective that could be used in the development of new and more selective antimalarials: targeting the mechanism of substrate transport from one active site to the other. We have already considered this possibility in a previous work [11].

The other enzyme of the folate cycle, SHMT, is a member of the α-class of the pyridoxal-5′-phosphate-dependent enzymes; it reversibly catalyses the conversion of serine into glycine, while the hydroxymethyl group is transferred to 5,6,7,8-tetrahydrofolate, transforming it in 5,10 methylenetetrahydrofolate, the sole precursor of purine biosynthesis and a key intermediate in the biosynthesis of thymidine, choline and methionine [12,13]. This enzyme also catalyses THF-independent aldolytic cleavege, decarboxylation, racemization and transamination reactions [14], being ubiquitous in nature.

The indispensable role of SHMT in DNA biosynthesis, allied to the fact that a high level of this enzyme activity was observed in rapidly proliferating cells during the S phase of the cell cycle and in a wide variety of tumors cells [15–17], points to SHMT as a potential target for the development of anticancer and antimicrobial agents [15,18,19]. However, SHMT shows great sequence similarity between species. In fact, the similarity between the active site of the *P. falciparum* SHMT and the human enzyme is 86.95%. This fact makes the discovery of selective inhibitors for this enzyme quite difficult. On the other hand, if those inhibitors are somehow developed, it is likely that the natural selection for resistant strains would be much slower. Still, this enzyme has not been validated yet as a new target for antimalarial chemotherapy. While DHFRTS have been well studied in protozoan parasites [11,20–24], little is known about parasitic SHMT [25–28]. Its crystallographic structure is yet unknown, and only recently its gene organization was revealed [29].

In this paper, we propose a 3D model for *P. falciparum* serine hydroxymethyltransferase (pfSHMT) in complex with *N*-glycine-[3-hydroxy-2-methyl-5-phosphonooxymethyl-pyridin-4-yl-methane] (PLG) and 5-formyl tetrahydrofolate (5-FTHF) or 5-formyl-6-hydrofolic acid (FFO). This model was built based on homology modeling by multiple alignment using as templates the crystallographic structures of SHMT available in the Protein Data Bank (PDB) [30,31].

2. Experimental

In the search for good templates, all the amino acid sequences of SHMT available in the PDB were used. The SHMTs from *E. coli* (eSHMT; PDB entry 1DF0, resolution=2,40 Ă, *R*-value=0,174) [32], Human (hSHMT; PDB entry 1BJ4, resolution=2,65 Ă, *R*-value=0210) [33], Murine (mSHMT; PDB entry 1EJI, resolution=2,90 Ă, *R*-value=0,271) [34], Rabbit (rSHMT; PDB entry 1CJ0, resolution=2,80 Ă, *R*-value=0,214) [35] and *Bacillus stearothermophyllus* (bsSHMT; PDB entry 1KKJ, resolution=1,93 *R*-value=0,178) [36] had their sequences aligned with the sequence of pfSHMT using the BLAST server [37,38], presenting 49.22%, 48.98%, 45.87%, 45.52% and 40.50% sequence homology, respectively. The alignments from BLAST were manually refined by comparison with the alignment of pfSHMT to sequences of other species proposed by Alfadhli and Rathod [29] and the alignment between eSHMT, rSHMT, mSHMT and hSHMT proposed by Scarsdale et al. [35]. The final model was built based on this refined multiple alignment, and the residue numbering followed was the same proposed by Scarsdale et al. [35].

The refined alignment was then submitted to the SWISS Model server [39–41] in the optimize mode to generate the initial models of each protomer. To prepare the templates for the quaternary structure of the model, the protomers were visually inspected and joined using Swiss-PDBViewer [39] by superposition of the backbones to the dimeric structures of the respective templates. These templates were then aligned with the sequence of the dimeric structure of pfSHMT and resubmitted to the SWISS Model server [39–41], again in the optimize mode, to generate the final model of the apoenzyme.

After building the structure of the pfSHMT apoenzyme model, the next step was to build the corresponding

holoenzyme by docking PLG and FFO into the respective active sites. This task was accomplished by superposition of the backbone of our model on that of the crystallographic structure of eSHMT complexed with PLG and FFO (PDB entry 1DFO) and copying the coordinates of the ligands into our model. The atomic partial charges of PLG and FFO were previously calculated at the Hartree Fock level with the 6-31G* basis set using the CHELPG approach of the Gaussian98 package [42]. The holoenzyme obtained by this procedure was set into a 1.700 nm^3 cubic box containing about 129.000 water molecules and minimized with the GROMOS force field [43] implemented in the GROMACS 3.1.4 package [44,45]. This minimization was carried out using the steepest descent algorithm until reaching an energy gradient of 100 kcal/mol·nm, which was found to be the most appropriate energy gradient to relax the models and afford good Ramachandran plots [46].

To check the quality of our model, it was submitted to validation analysis using PROCHECK [47] and WHATIF [48], available at Biotech Validation Suite for Protein Structures [49]. Additionally the software PROSA 2003 [50] was used to calculate the *Z*-scores [51] of our model and hSHMT.

The molecular dynamics steps were carried out according to the following procedure: First, 50 ps of molecular dynamics at 300 K in the water molecules inside the box to allow for the equilibration of the solvent around the protein residues. In this dynamics, all protein atoms had their positions restrained. Then, a full molecular dynamics simulation of 1000 ps at 300 K with no restrictions using 1 fs of integration time and a cut-off of 14 Angstroms (Å) for long-range interactions was carried out. As a whole, 100 conformations were stored during this simulation. In this step, the pairlists were updated every 500 time steps, all the Lys and Arg residues were positively charged, and the Glu and Asp residues were negatively charged.

Furthermore, the crystallographic structure of dimeric hSHMT, with PLG and FFO docked into its active site, was submitted to several rounds of energy minimization and dynamic simulations following the same procedure described to pfSHMT. After the molecular dynamic simulations, the backbone of the active site residues of our model was superimposed to that of hSHMT using Swiss-PDBViewer [39].

The hardware resources used in this work were a PC AMD Atlhon 2000 MHz and a cluster of computers composed by a PC Pentium IV 2.4 and two PC Athlon 2200 MHz.

3. Results and discussion

3.1. Homology modeling

Fig. 1 shows our final alignment, which was submitted to the SWISS Model server to build the initial model of pfSHMT. This alignment was obtained after manual adjustments of the initial alignment from the BLAST server [37,38]. The adjustments were performed after inspection of the alignment of the pfSHMT sequence with the sequences of other species proposed by Alfadhli and Rathod [29] and the alignment between the sequences of eSHMT, rSHMT, mSHMT and hSHMT proposed by Scarsdale et al. [35]. Visual inspection showed that the BLAST alignment did not match well important regions of the pfSHMT sequence, for example, the insertion observed in residue Asn250 (Fig. 1). The alignments in these regions were corrected accordingly to the mentioned alignments. The sequence of bsSHMT was aligned manually first to the templates and then to pfSHMT following this same procedure.

As it is the case for the templates [32–36], our model has a quaternary structure formed by at least two identical monomers that fold into three domains. The Ramachandran plot [46] (data not shown) shows more than 96% of the amino acid residues in the favorable regions of the plot for the whole enzyme. The main structural elements of the optimized pfSHMT homology model are shown in Fig. 2. The model proposed here is composed of two identical monomeric chains packed into a tight dimer, with two functional active sites shared between its monomeric subunits, similar to the experimentally observed structure of eSHMT in water [32] or the crystal of bsSHMT [36]. In contrast, the mammalian SHMT templates present a tetrameric structure in solution, a dimer of dimers with 222 (D2) symmetry [33]. However, each dimer, in the same way as eSHMT and bsSHMT, presents two functional active sites shared between its monomeric subunits, and none of the active site residues is involved in the interactions responsible for the tetramer formation.

Some structural features are fundamental to the maintenance of the tetrameric structure of all the mammalian templates in solution. One of them is the residue His135. For these enzymes, His135 from the different monomers hydrogen bond to each other in the dimers, thus assembling the symmetric tetrameric quaternary structure, in which very center they are located [33]. In eSHMT and pfSHMT sequences, there are gaps in the positions corresponding to these histidines, and the nearest residues are, respectively, a proline and a glycine for eSHMT and a valine and a lysine for pfSHMT (see Fig. 1). None of those residues is able to form the type of hydrogen bond that is necessary for the stabilization of the tetrameric structure in solution, as proposed by Renwick et al. [33]. Other important structural features to the tetramer-stabilizing interactions are two insertions present in the mammalian enzymes. The first insertion is KKKISATSI at residue Lys157. This insertion is changed to KKKVSITSD in pfSHMT. The second insertion is RKGVKSVDPKTGKEILY, which starts at residue Arg270. This insertion is incomplete in pfSHMT corresponding only to KKRNPG (see alignment in Fig. 1). Scarsdale et al. [32] propose that eSHMT could not form the same tetrameric quaternary

```
pfSHMT    1                   M-F NNDPLQKYDK ELFDLLEKEK NRQIETINLI ASENLTNTAV RECLGDRISN KYSEGYPHKR YYGGNDYVDK IEELCYKRAL
eSHMT    -1                 MRMLK REMNIADYDA ELWQAMEQEK VRQEEHIELI ASENYTSPRV MQAQGSQLTN KYAEGYPGKR YYGGCEYVDI VEQLAIDRAK
hSHMT    11       DAD LWSSHDA--MLA QPLK--DSDV EVYNIIKKES NRQRVGLELI ASENFASRAV LEALGSCLNN KYSEGYPGQR YYGGTEFIDE LETLCQKRAL
mSHMT     7    MADRDAT LWASHEK--MLS QPLK--DSDA EVYSIIKKES NRQRVGLELI ASENFASRAV LEALGSSLNN KYSEGYPGQR YYGGTEFIDE LEXLCQKRAL
rSHMT    14             WSSHEQ--MLA QPLK--DSDA EVYDIIKKES NRQRVGLELI ASENFASRAV LEALGSCLNN KYSEGYPGQR YYGGTEHIDE LETLCQKRAL
bsSHMT    1                 MKYLP QQDP------ QVFAAIEQER KRQHAKIELI ASENFVSRAV MEAQGSVLTN KYAEGYPGRR YYGGCEYVDI VEELARERAK
                                               ..   .  *   **   ..** **** .  *  .  *  . * **.****  * **** . .*  .* *   **

pfSHMT   83   EAFNVSEEEW GVNVQPLSGS AANVQALYAL VGV-KGKIMG MHLCSGGHLT HGFFDEKKKV SITSDLFESK LYKCNSE-GY VDMESVRNLA LSFQPKVIIC
eSHMT    84   ELFGAD---- YANVQPHSGS QANFAVYTAL LEP-GDTVLG MNLAHGGHLT HGSPVN---F SGK-LY-NIV PYGIDA-TGH IDYADLEKQA KEHKPKMIIG
hSHMT   102   QAYKLDPQCW GVNVQPYSGS PANFAVYTAL VEPHG-RIMG LDLPDGGHLT HGFMTDKKKI SATSIFFESM PYKVNPDTGY INYDQLEENA RLFHPKLIIA
mSHMT   102   QAYHLDPQCW GVNVQPYSGS PANFAVYTAL VEPHG-RIMG LDLPDGGHLT HGFXTDKKKI SATSIFFESX PYKVYPETGY INYDQLEENA SLFHPKLIIA
rSHMT   102   QAYGLDPQCW GVNVQPYSGS PANFAVYTAL VEPHG-RIMG LDLPDGGHLT HGFMTDKKKI SATSIFFESM AYKVNPDTGY IDYDRLEENA RLFHPKLIIA
bsSHMT   80   QLFGAE---- HANVQPHSGA QANMAVYFTV LE-HGDTVLG MNLSHGGHLT HGSPVN---F SGV-QY-NFV AYGVDPETHV IDYDDVREKA RLHRPKLIVA
              . .          **** **.  **     .. .      . * ..*  ***** **   .     *       .    *   .      ..   .   *     **.*.

pfSHMT  181   GYTSYPRDID YKGFREICDE VNAYLFADIS HISSFVACNL LNNPFTYADV VTTTTHKILR GPRSALIFFN ---KKRNPG- -------IDQ KINSSVFPSF
eSHMT   173   GFSAYSGVVD WAKMREIADS IGAYLFVDMA HVAGLVAAGV YPNPVPHAHV VTTTTHKTLA GPRGGLILAK -G-------- GSEELYK--K L-NSAVFPGG
hSHMT   201   GTSCYSRNLE YARLRKIADE NGAYLMADMA HISGLVAAGV VPSPFEHCHV VTTTTHKTLR GCRAGMIFYR KGVKSVDPAT GKEILYNLES LINSAVFPGL
mSHMT   201   GTSCYSRNLD YARLRKIADD NGAYLXADXA HISGLVAAGV VPSPFEHCHV VTTTTHKTLR GCRAGXIFYR KGVRSVDPKT GKETYYELES LINSAVFPGL
rSHMT   201   GTSCYSRNLD YGRLRKIADE NGAYLMADMA HISGLVVAGV VPSPFEHCHV VTTTTHKTLR GCRAGMIFYR RGVRSVDPKT GKEILYNLES LINSAVFPGL
bsSHMT  170   AASAYPRIID FAKFREIADE VGAYLMVDMA HIAGLVAAGL HPNPVPYAHF VTTTTHKTLR GPRGGMIL-- ---------- CQEQFA---K QIDKAIFPGI
              . . *.  ..     * * *    ***  * . *....*   .   .*       ******* *  * *.. *.                          . ..**.

pfSHMT  270   QGGPHNNKIA AVACQLKEVN TPFFKEYTKQ VLLNSKALAE CLLKRNLDLV TNGTDNHLIV VDLRKYNITG SKLQETCNAI NIALNKNTIP SD-VDCVSPS
eSHMT   261   QGGPLMHVIA GKAVALKEAM EPEFKTYQQQ VAKNAKAMVE VFLERGYKVV SGGTDNHLFL VDLVDKNLTG KEADAALGRA NITVNKNSVP NDPKSPFVTS
hSHMT   301   QGGPHNHAIA GVAVALKQAM TLEFKVYQHQ VVANCRALSE ALTELGYKIV TGGSDNHLIL VDLRSKGTDG GRAEKVLEAC SIACNKNTCP GD-RSALRPS
mSHMT   301   QGGPHNHAIA GVAVALKQAX TTEFKIYQLQ VLANCRALSD ALTELGYKIV TGGSDNHLIL XDLRSKGTDG GRAEKVLEAC SIACNKNTCP GD-KSALRPS
rSHMT   301   QGGPHNHAIA GVAVALKQAM TPEFKEYQRQ VVANCRALSA ALVELGYKIV TGGSDNHLIL VDLRSKGTDG GRAEKVLEAC SIACNKNTCP GD-KSALRPS
bsSHMT  255   QGGPLMHVIA AKAVAFGEAL QDDFKAYAKR VVDNAKRLAS ALQNEGFTLV SGGTDNHLLL VDLRPQQLTG KTAEKVLDEV GITVNKNTIP YDPESPFVTS
              ****  . ** . *  . .      ** *  . *  * . .    .      .* . *.****..  **     *    .         *. ***. *  *       *

pfSHMT  369   GIRIGTPALT TRGCKEKDME FIADMLLKAI LLTDELQQKY G--KKLVDFK KGLVNNPK-- --IDELKKEV VQWAKNL--- --PFA
eSHMT   361   GIRVGTPAIT RRGFKE--AE AKELAGWMCD VLD-SIN--- DE---AVI-- --ERIKG--- KV-LDICA-R YPV--YAH
hSHMT   400   GLRLGTPALT SRGLLEKDFQ -K-VAHFIHR GIELTLQIQS DTGVAATLKE FKERLAGD-- KYQAAVQALR EEVESFASLF PLPGL
mSHMT   400   GLRLGTPALT SRGLLEEDFQ -K-VAHFIHR GIELTLQIQS HTATKATLKE FKEKLAGD-E KIQSAVATLR EEVENFASNF SLPGLPDF
rSHMT   400   GLRLGTPALT SRGLLEKDFQ -K-VAHFIHR GIELTVQIQD DTGPRATLKE FKEKLAGD-E KHQRAVRALR QEVESFAALF PLPGLPGF
bsSHMT  355   GIRIGTAAVT TRGF--GLEE MDEIAAII-- GLVL-KNVGS E---QA-LEE ARQRVAALTD
              *.*.**.*.*  **      .         .    .
```

Fig. 1. Alignment of pfSHMT with the sequences of the crystallographic SHMTs available in PDB. The sequences numbering is the same adopted by Scarsdale et al. [35].

structure observed for mammalian SHMT because, besides not having a residue equivalent to His135, it lacks the tetramer-stabilizing interactions. Accordingly, there are relatively few specific interactions between the dimers in the eSHMT tetramers, and their persistence in solution seems unlikely [32]. It is believed that bsSHMT is also a

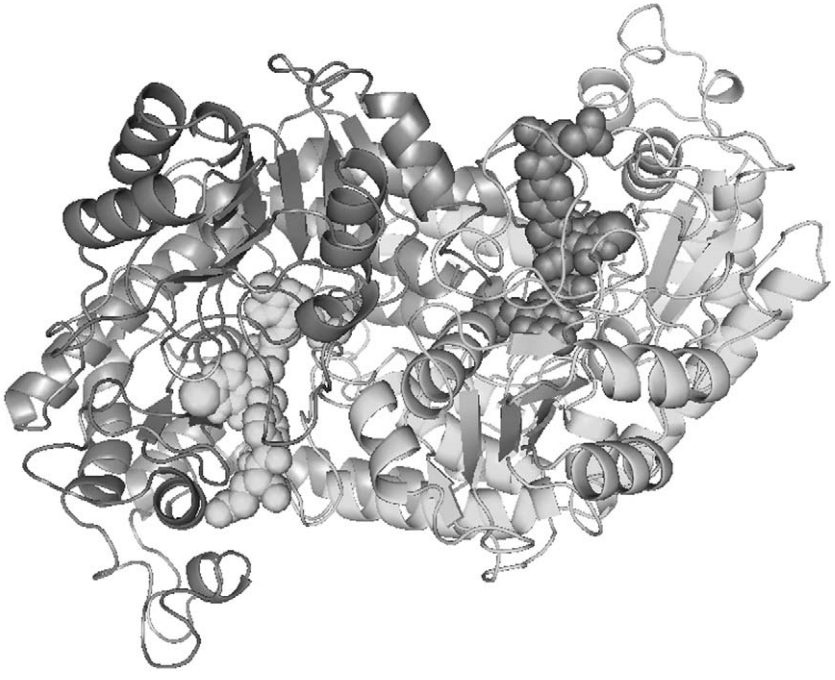

Fig. 2. Proposed homology model for pfSHMT based on the multiple alignment; *N*-glycine-[3-hydroxy-2-methyl-5-phosphonooxymethyl-pyridin-4-yl-methane] (PLG) and 5-formyl tetrahydrofolate (5-CHO-H_4PteGlu) (FFO) are shown in CPK notation.

dimer since, according to Trivedi et al. [36], no tetrameric contacts are visible in its crystal structure, and this is consistent with gel filtration experiments, which show that the protein elutes at a position corresponding to a dimer. His108 in bsSHMT is the equivalent to His135 in the mammalian enzymes, but this residue has been shown to be involved in stacking interactions, which are important for the formation of a dimeric quaternary structure instead of a tetrameric one [36]. It can be noticed that pfSHMT possesses the first tetramer stabilizing insertion, with only two different residues, but lacks part of the second one. This enzyme also lacks the very important central histidines. Since there is no experimental evidence available, it is difficult to affirm whether or not pfSHMT would form a tetrameric quaternary structure in solution. Instead of this and considering that all SHMT are active as dimers [28], we decided to restrict our study to a dimeric structure. This decision was also important to reduce the computation time necessary for the molecular dynamics simulations.

As it is the case for all the templates [32–36], each monomer in our model is composed by three domains, the N-terminal domain, the second N-terminal domain and the small domain. The N-terminal domain (residues 1–34) mediates intersubunit contacts and folds into two α-helices and one β-strand. The second N-terminal domain or the "large" domain (residues 34 a 290) binds PLP, has most of the active site residues and folds into an α–β–α sandwich containing nine α-helices wrapped around a seven-stranded mixed β-sheet. Finally, the C-terminal domain or small domain (residues 291–442) folds into an α–β sandwich which intermediates the contacts with the small domain of the other pfSHMT monomers in the dimer.

The backbones of the modeled monomers of pfSHMT and the templates used in the multiple alignments were superimposed to calculate the RMSD values and to check their structural compatibility with the templates. The RMSD were of 1.34 Å for eSHMT, mSHMT and bsSHMT, 1.37 Å for rSHMT and 1.44 Å for hSHMT, thus confirming that the models are satisfactory regarding the utilization of the chosen templates for the homology modeling process.

3.2. Validation of the models

Regarding the main chain properties of the modeled enzymes, no considerable bad contacts nor C_α tetrahedron distortion nor hydrogen bond energy problems were found. Moreover, the average G factor, the measure of the normality degree of the protein properties, was of −1.04, which is inside the permitted values for homology models. Furthermore, there were no distortions of the side chain torsion angles found. The *Z*-scores calculated using the software PROSA 2003 [50] showed that hSHMT and pfSHMT are well inside the range of a typical native structure [51].

3.3. Active site determination

The active site determination in our model was accomplished based on its alignment to the templates. There were found degrees of identity above 82% for the active site in all the cases, as shown in Table 1. An inspection of Table 1 shows that the active site of pfSHMT presents three different residues when compared to the active site of hSHMT: Asp(A)136, Glu(A)137 and Thr(A)183 in pfSHMT instead of Thr(A)145, Asp(A)146 and Ser(A)193 in hSHMT. At first hand, it seems that it would be very difficult to design selective inhibitors for pfSHMT, however, if this is possible, we expect that those compounds would be more durable as efficient antimalarial drugs, once the close similarity between the SHMT active sites of the different species is an indication of the low potential for mutation of this enzyme. Furthermore, Fu et al. [52] propose that the canonical SHMT structure supports the inference that the folate interactions site evolved from a type I PLP precursor enzyme through sequence insertions and not by domain swapping from other folate-requiring enzymes. This may be the reason why antifolate compounds developed as chemotherapeutic agents are ineffective as inhibitors of SHMT and suggests that an effective antifolate inhibitor of SHMT may not inhibit other folate enzymes. Thus, the challenge of developing selective inhibitors for SHMT would bring the additional advantage of achieving selectivity to the other enzymes of the folate cycle.

Table 1
Comparison between the active site residues of site one of pfSHMT and the templates

bsSHMT	rSHMT	mSHMT	hSHMT	eSHMT	pfSHMT
Ser(A)31	Ser(A)35	Ser(A)53	Ser(A)43	Ser(A)35	Ser(A)34
Tyr(B)51	Tyr(B)55	Tyr(B)73	Tyr(B)63	Tyr(B)55	Tyr(B)54
Glu(B)53	Glu(B)57	Glu(B)75	Glu(B)65	Glu(B)57	Glu(B)56
Tyr(B)60	Tyr(B)64	Tyr(B)82	Tyr(B)72	Tyr(B)64	Tyr(B)63
Tyr(B)61	Tyr(B)65	Tyr(B)83	Tyr(B)73	Tyr(B)65	Tyr(B)64
Gly(A)95	Gly(A)102	Gly(A)120	Gly(A)110	Gly(A)98	Gly(A)101
Ala(A)96	Ser(A)103	Ser(A)121	Ser(A)111	Ser(A)99	Ser(A)102
Leu(A)107	**Ser(A)125**	Leu(A)143	Leu(A)133	Leu(A)121	Leu(A)124
Gly(A)111	Gly(A)129	Gly(A)147	Gly(A)137	Gly(A)125	Gly(A)128
His(A)112	His(A)130	His(A)148	His(A)138	His(A)126	His(A)129
Leu(A)113	Leu(A)131	Leu(A)149	Leu(A)139	Leu(A)127	Leu(A)130
Val(A)119	**Thr(A)133**	**Thr(A)145**	**Thr(A)145**	**Val(A)137**	Asp(A)136
Asn(A)120	**Asp(A)134**	**Asp(A)146**	**Asp(A)146**	**Asn(A)138**	Glu(A)137
–	Lys(A)136	Lys(A)148	Lys(A)148	–	Lys(A)139
Ser(A)173	**Ser(A)175**	**Ser(A)203**	**Ser(A)193**	**Ser(A)175**	Thr(A)183
Asp(A)198	Asp(A)200	Asp(A)228	Asp(A)218	Asp(A)200	Asp(A)208
His(A)200	His(A)203	His(A)231	His(A)221	His(A)203	His(A)211
Thr(A)223	Thr(A)226	Thr(A)254	Thr(A)244	Thr(A)226	Thr(A)234
His(A)225	His(A)228	His(A)256	His(A)246	His(A)228	His(A)236
Lys(A)226	Lys(A)229	Lys(A)257	Lys(A)247	Lys(A)229	Lys(A)237
Gly(B)256	Gly(B)263	Gly(B)303	Gly(B)293	Gly(B)263	Gly(B)272
Asn(A)341	Asn(A)347	Asn(A)387	Asn(A)377	Asn(A)347	Asn(A)356
Arg(A)357	Arg(A)363	Arg(A)402	Arg(A)392	Arg(A)363	Arg(A)371

Residues that does not match in our model are shown in bold.

In literature, there are only four crystal structures of SHMT complexes: eSHMT, bsSHMT, mSHMT and rSHMT. While the structures of eSHMT, mSHMT and bsSHMT are ternary complexes of the monoglutamate of 5-CHO-H_4PteGlu (FFO) with glycine [32,34,36], rSHMT includes the complexes with 5-CHO-H_4PteGlu$_3$ and H_4PteGlu. Fu et al. [52] have discussed the pteroylpolyglutamate-binding site in rSHMT and observed that the anchorage of PLP, as well as of the pteridine ring and the PABA portions of FFO, occurs at the same site and in a very similar way for the four crystals. The polyglutamate tail, on the other hand, binds differently. There are several insertions of amino acids in the mammalian enzymes compared with eSHMT and bsSHMT, three of which are located at the polyglutamate-binding site. In eSHMT (or bsSHMT), the insertion at positons 134 (130) introduces a KKK sequence that leads to interactions with glu_2 and glu_3. Furthermore, the insertions of Ile399 (Leu393 for bsSHMT) and Gly244 (Cys238 for bsSHMT) have been shown to be important for the interactions and/or channeling of glu_5 of an extended polyglutamate tail. This indicates that eSHMT and bsSHMT do not have a well-defined polyglutamate-binding site. As for the mammalian enzymes, our model of pfSHMT presents, relatively conserved, all the insertions considered important for polyglutamate binding, thus leading to the expectation that it would anchor well not only glu but also glu_3 and glu_5 tails. In this work, however, we built and investigated only the interactions of a model docked with a monoglutamate tail.

As mentioned before, the dimeric structure of the templates and our model present two active sites composed by residues from two monomers. Site one is composed by residues from monomers A and B and site two by residues from monomers B and A. Comparing the residues of site one from our model to the pictures of site one from eSHMT and mSHMT, reported by Scarsdale et al. [32] and Szebenyi et al. [34], it can be initially inferred, based on the alignment in Fig. 1, that, in pfSHMT, residues Tyr(B)54, Tyr(B)69, Gly(A)101, Thr(A)183, Asp(A)208, His(A)211, Thr(A)234, His(A)236, Gly(A)272 and Arg(A)371 would be the ones that stabilize the binding of PLG, while Lys(A)237 would be the residue that forms the internal aldimine with PLG in the first stage of the enzymatic reaction mechanism [32]. The anchorage of FFO as a monoglutamate form inside the active site would be accomplished by residues Glu(A)57, Tyr(B)64, Leu(A)121, Gly(A)125, His(A)126, Leu(A)127, Asp(A)136, Glu(A)137, Thr(A)183 and Asn(A)347.

3.4. Molecular dynamics simulations

After the energy minimization, the 3D pfSHMT model and the crystallographic hSHMT holoenzymes were sub-

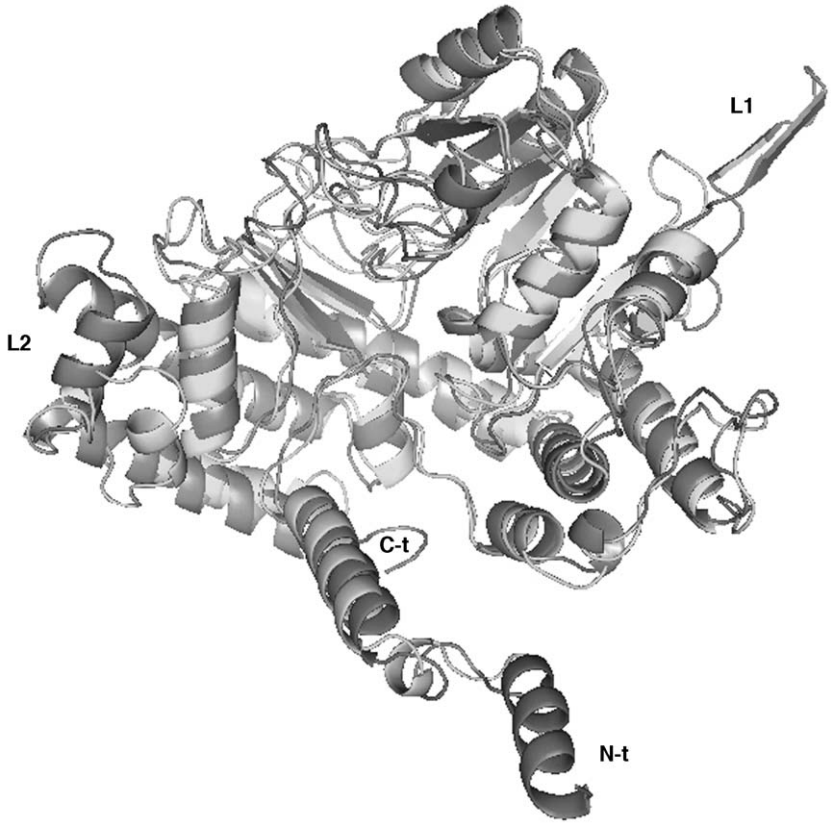

Fig. 3. Overall structure comparison of the backbones of our model (in light gray) and hSHMT (in dark gray). The loops with major structural differences between the two enzymes and their C- and N-terminal extremities are labeled as L1, L2, C-t and N-t.

mitted to molecular dynamics simulations using the GROMOS 96 force field. This procedure had the goal of putting both enzymes in the same physiologic conditions for further superposition and also of providing a way for additional validation and refinement of the structure of pfSHMT. In fact, an overall average structure conserved in time during a molecular dynamics simulation is essential to consider a model acceptable.

After the dynamics simulations, the RMSD deviation between the final model and the hSHMT was calculated again. The considerably low RMSD value presented, 1.44, confirmed that the pfSHMT model was very similar to the crystallographic structure. The overall folding of our pfSHMT model remained practically unchanged during the dynamics simulations, thus ensuring its good quality.

3.5. Overall structural comparison of pfSHMT and hSHMT

The proposed pfSHMT dimer contains 884 amino acid residues, while dimeric hSHMT has 940 residues. The backbone superposition of the monomer of our model with hSHMT (Fig. 3) shows that the general fold of pfSHMT is essentially the same as that of the human protein. The major structural differences between both structures occur at the external loops L1 and L2 and at the C- and N-terminal extremities, as shown in Fig. 3. Loops L1 and L2 in pfSHMT are almost half the size as in hSHMT, and the C- and N-terminal extremities of hSHMT are longer than in pfSHMT. In fact, hSHMT has 28 additional amino acid residues which are located at those regions; seven in loop L1 (residues 278–284), six in loop L2 (residues 294–299), nine at the C-terminal region and six at the N-terminal. The structural differences at the SHMT active site between the *P. falciparum* and the human enzyme are of particular interest for the design of compounds that would potentially have greater affinity for the parasite enzyme active site. Despite the great sequence identity between both active sites (86.96%), some structural aspects, which were observed after the molecular dynamics simulations, appear to differentiate significantly both enzymes.

3.6. Differences in the binding sites regions between pfSHMT and hSHMT

As shown in Table 1, pfSHMT presents three different residues at its active site in comparison with hSHMT: Thr(A)145, Asp(A)146 and Ser(A)193 in hSHMT instead of Asp(A)136, Glu(A)137 and Thr(A)183 in pfSHMT. Analysis of the behavior of those residues along the MD simulation (Figs. 4 and 5) showed that the side chain of Glu(A)137 in pfSHMT is positioned about at 4 to 6 Å from the FFO monoglutamate tail but that it does not interact with them due to the repulsion between the negative charges of both regions. On the other hand, the corresponding amino acid

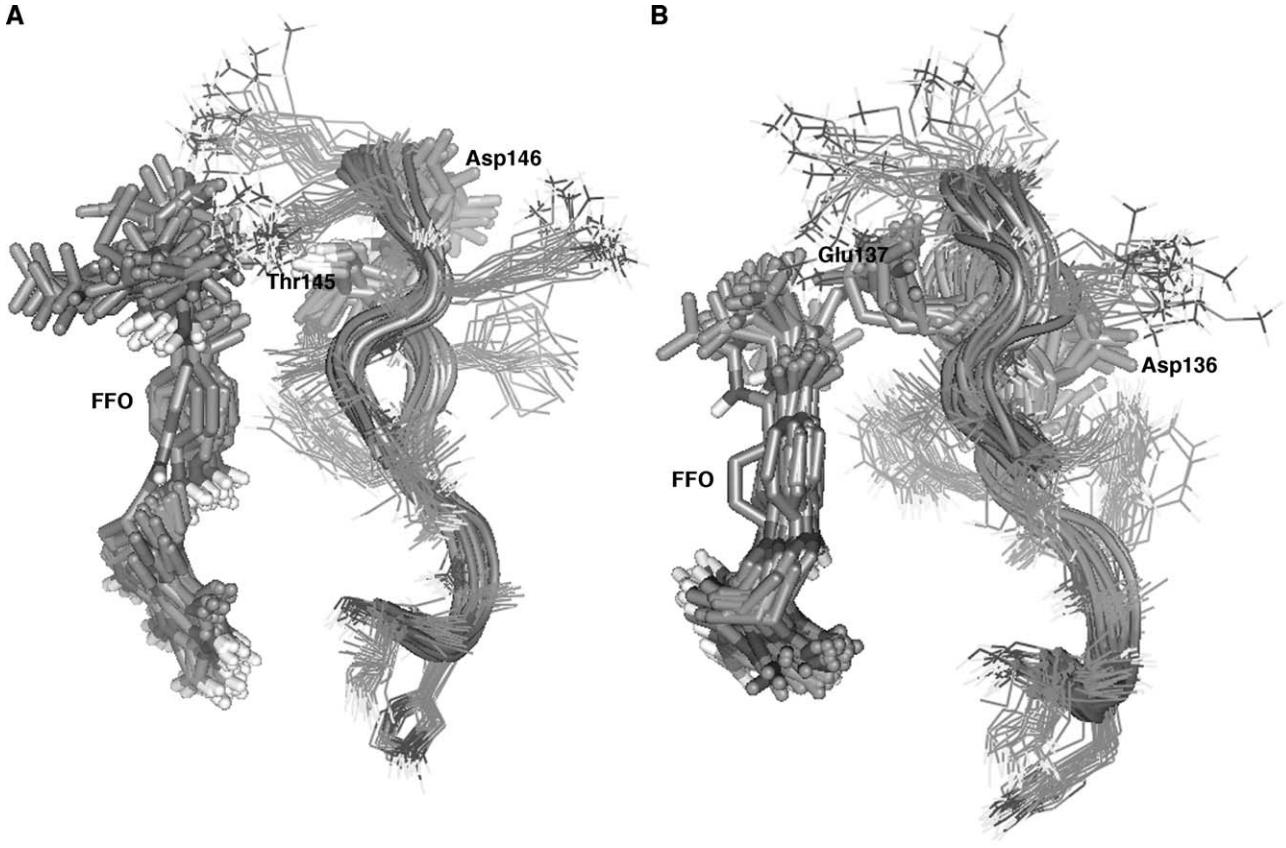

Fig. 4. Behavior of some residues along the molecular dynamics simulations. (A) Thr145 and Asp146 in hSHMT; (B) Asp136 and Glu137 in pfSHMT.

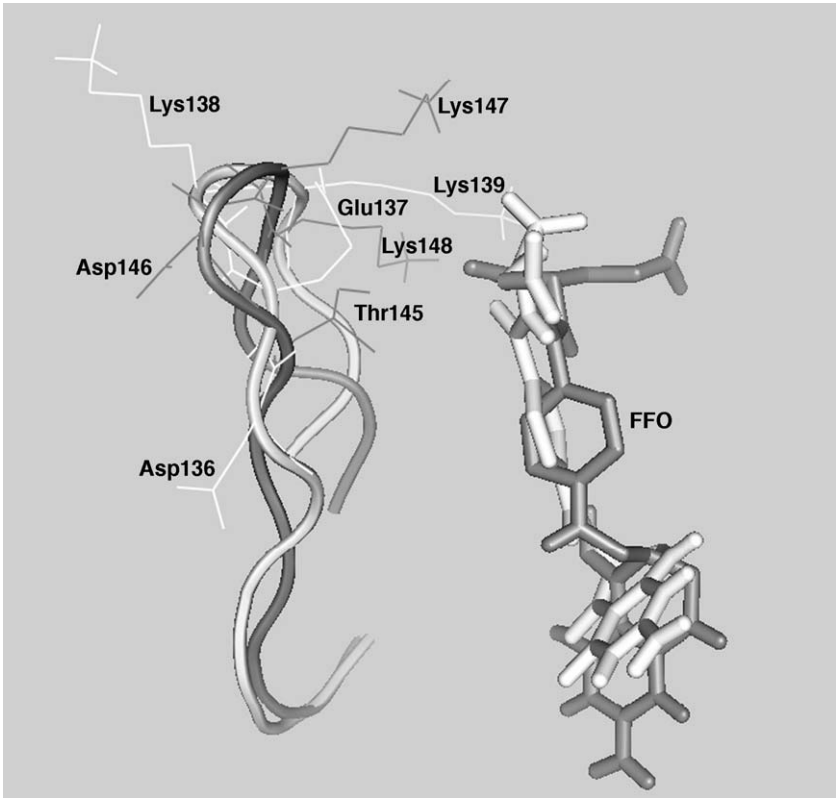

Fig. 5. Superposition of residues of active site one around the formyl tetrahydrofolate (5-CHO-H_4PteGlu), FFO, tail after molecular dynamics simulations. Light gray—pfSHMT; Dark gray—hSHMT. PLG—[3-hydroxy-2-methyl-5-phosphonooxymethyl-pyridin-4-yl-methane].

in the human enzyme, Asp(A)146, is forced by the protein backbone to orient its carboxylate group in opposition to the monoglutamate tail of FFO (about 10 Å), thus avoiding any possibility of interaction even if it were a positively charged side chain. Another significant difference between both enzymes is that Thr(A)145 in hSHMT sets its side chain at about 5 to 6 Å from FFO, with its OH group pointing towards FFO, while Asp(A)136 in pfSHMT have its carboxylate group in an opposite position, i.e., about 7 to 8 Å from FFO. Finally, Ser(A)193 in hSHMT and Thr(A)183 in pfSHMT set their side chains between PLG and FFO, and, in both cases, they establish hydrogen bonds with FFO, thus behaving practically in the same way.

The differences between Glu(A)137 and Asp(A)146 and between Asp(A)136 and Thr(A)145 in both enzymes seem to afford the only good opportunities to be exploited in the design of compounds that would bind selectively to pfSHMT. For example, an FFO analogue having a positively charged group in the position corresponding to the first carboxylate group at the monoglutamate tail of FFO could interact strongly with Glu137(A) in pfSHMT and be repelled by Lys(A)147 and Lys(A)148 in hSHMT (see Fig. 5). Such inhibitor could bind selectively to pfSHMT. The shorter distance of Thr(A)145 from FFO in hSHMT also could be exploited by a derivative with a bulkier group that would fill the extra 2 Å observed in pfSHMT. Such a derivative would be too large to be well accommodated in the active site of hSHMT and would bind more tightly to pfSHMT.

4. Conclusions

In this work, we have used homology modeling and molecular dynamics studies to propose the first 3D structure for pfSHMT. The comparison between the known crystallographic structures for several SHMT and our pfSHMT model indicates that the *Plasmodium* enzyme could be a dimer in solution. We have also used molecular dynamics and docking studies to explore the opportunities opened by the differences found for the interactions of a monoglutamate tailed substrate with the active site of the pfSHMT model and that of the crystallographic structure of hSHMT. The differences between the two enzymes point out to compounds analogous to FFO, with either a positively charged tail instead of a glutamate tail or with bulkier substituents at the tail as potential selective inhibitors. With the availability of the X-ray diffraction structure of rSHMT in complex with glu3, which could be used as template for docking studies [52], there is much room for further theoretical investigations of the opportunities that a polyglutamate tail docked in the active site could open for the design of selective inhibitors for the parasite enzyme. This work is already in progress in our laboratory, and its results, associated with the results presented here, will be the

starting point for further molecular dynamics simulations that could lead to the proposition of the first potential selective inhibitors of pfSHMT.

Acknowledgments

The authors wish to thank the Brazilian financial agencies CNPq, FAPERJ and CAPES for financial support and Dr. Alan Wilter Souza da Silva for his help with the GROMACS Package.

References

[1] A.O. Santos Filho, Molecular modeling of dihydrofolate reductase of *Plasmodium falciparum* and study of its inhibitors by QSAR-3D, PhD thesis—(Chemical Engineering Department, Instituto Militar de Engenharia—Rio de Janeiro, Brazil 2000) Chapter 1.
[2] The World Health Organization Report, WHO Publications, Geneva, 1997.
[3] W. Peters, Drug resistance in malaria parasites of animals in man, Adv. Parasitol. 41 (1998) 1–62.
[4] S.J. Foote, A.F. Cowman, The mode of action and the mechanism of resistance to antimalarial drugs, Acta Trop. 56 (1994) 157–171.
[5] P. Newton, N. White, Malaria: new developments in treatment and prevention, Annu. Rev. Med. 50 (1999) 179–192.
[6] N.J. White, P.L. Olliaro, Strategies for the prevention of antimalarial drug resistance: rationale for combination chemotherapy of malaria, Parasitol. Today 12 (1996) 399–401.
[7] I.S. Soares, M.M. Rodrigues, Malaria vaccine: road-blocks and possible solutions, Braz. J. Med. Biol. Res. 31 (1998) 317–332.
[8] F.M. Huennekens, T.H. Duffy, K.S. Vitols, Folic acid metabolism and its disruption by pharmacological agents, NCI Monogr. 5 (1987) 1–8.
[9] H.A. Ingraham, L. Dickey, M. Goulian, DNA fragmentation and cytotoxicity from increased cellular deoxyuridylate, Biochemistry 25 (1986) 3225–3230.
[10] A. Yoshida, S. Tanaka, O. Hiraoka, Y. Koyama, Y. Hirota, D. Ayusawa, T. Seno, C. Garrett, Y. Wataya, Deoxyribonuceloside triphosphate imbalance. 5-fluorodeoxyuridine-induced DNA double strand breaks in mouse FM3A cells and the mechanism of cell death, J. Biol. Chem. 262 (1987) 8235–8241.
[11] T.C.C. França, A.L. Medeiros, O.A. Santos Filho, E.C.P. Santos, J.D. Figueroa-Villar, A complete model of the *Plasmodium falciparum* bifunctional enzyme dihydrofolate reductase–thymidylate synthase. A model to design new antimalarials, J. Braz. Chem. Soc. 15 (3) (2004) 450–454.
[12] R.L. Blakely, The interconversion of serine and glycine. Participation of pyridoxal-phosphate, Biochem. J. 61 (1995) 315–323.
[13] D.R. Appling, Compartmentation of folate-mediated one-carbon metabolism in eukharyotes, FASEB J. 5 (1991) 2645–2651.
[14] L. Schirch, in: R.L. Blakely, S.J. Benkovic (Eds.), Folates and Proteins: Chemistry and Biochemistry of Folates, Wiley Interscience, New York, 1984, pp. 399–412.
[15] H.G. Eichler, R. Hubbard, K. Snell, The role of serine hydroxymethyltransferase in cell proliferation: DNA synthesis from serine following mitogenic stimulation of lymphocytes, Biosci. Rep. 1 (1981) 101–106.
[16] K. Snell, Y. Natsumeda, J.N. Eble, J.L. Glover, G. Weber, Enzymatic imbalance in serine metabolism in human colon carcinoma and rat sarcoma, Br. J. Cancer 57 (1988) 87–90.
[17] K. Snell, Y. Natsumeda, G. Weber, The modulation of serine metabolism in hepatoma 3924A during different phases of cellular proliferation in culture, Biochem. J. 245 (1987) 609–612.
[18] J. Thorndike, T.T. Pelliniemi, W.S. Beck, Serine hydroxymethyltransferase activity and serine incorporation in leukocytes, Cancer Res. 39 (1979), 3435–3440.
[19] N. Appaji Rao, in: T. Ozawa (Ed.), New Trends in Biological Chemistry, Japan Scientific Press, Tokyo, 1991, pp. 333–340.
[20] G.X. Chen, C. Mueller, M. Wendlinger, J.W. Zolg, Kinetic and molecular properties of the dihydrofolate reductase from pyrimethamine-sensitive and pyrimethamine-resistant clones of the human malaria parasite *Plasmodium falciparum*, Mol. Pharmacol. 31 (1987) 430–437.
[21] W. Sirawaraporn, T. Sathitkul, R. Sirawaraporn, Y. Yuthavong, D.V. Santi, Antifolate-resistant mutants of *Plasmodium falciparum* dihydrofolate reductase, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 1124–1129.
[22] M. Hekmat-Nejad, P.K. Rathod, Kinetics of *Plasmodium falciparum* thymidylate synthase: interactions with high-affinity metabolites of 5-fluoroorotate and D1694, Chemotherapy 40 (1996) 1628–1632.
[23] R.T. Delfino, O.A. Santos Filho, J.D. Figueroa-Villar, Molecular modeling of wild-type and antifolate resistant mutant *Plasmodium falciparum* DHFR, Biophys. Chemist. 98 (2002) 287–300.
[24] A.L. Medeiros, Molecular modeling by homology of the bifunctional enzyme dihydrofolate reductase–thymidylate synthase of *Plasmodium falciparum,* MSc thesis, (Chemistry Department, Military Institute of de Engineering—Rio de Janeiro, Brazil 2002).
[25] P. Ruenwongsa, M. Luanvararat, W. O'Sullivan, Serine hydroxymethyltransferase from pyrimethamine-sensitive and resistant strains of *Plasmodium chobaudi*, J. Mol. Biochem. Parasitol. 33 (1989) 265–271.
[26] E.G. Platzer, H.C.J. Campuzano, The serine hydroxymethyltransferase of *Plasmodium lophurae*, Protozoologia 23 (1976) 282–286.
[27] E.G. Platzer, Subcellular distribution of serine hydroxymethyltransferase in *Plasmodium lophurae*, Life Sci. 20 (1977) 1417–1424.
[28] D.C. Kaslw, S. Hill, Cloning metabolic pathway genes by complementation in *Escherichia coli*. Isolation and expression of *Plasmodium falciparum* glucose phosphate isomerase, J. Biol. Chem. 165 (1990) 12337–12341.
[29] S. Alfadhli, P.K. Rathod, Gene organization of a *Plasmodium falciparum* serine hydroxymethyltransferase and its functional expression in *Escherichia coli*, Mol. Biochem. Parasitol. 110 (2000) 283–291.
[30] F.C. Bernstein, T.F. Koetzle, G.J. Williams, E.E. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, The protein data bank. A computer-based archival file for macromolecular structures, J. Mol. Biol. 112 (1977) 535–542.
[31] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[32] J.N. Scarsdale, S. Radaev, G. Kazanina, V. Schirch, T. Wright, 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 (2000) 155–168.
[33] S.B. Renwick, K. Snell, U. Baumann, The crystal structure of human cytosolic serine hydroxymethyltransferase: a target for cancer chemotherapy, Structure 6 (9) (1998) 1105–1116.
[34] D.M.E. Szebenyi, X. Liu, A. Kriksunov, P.J. Stover, D.J. Thiel, Structure of a murine cytoplasmic serine hydroxymethyltransferase quinonoid ternary complex: evidence for asymmetric obligate dimers, Biochemistry 39 (2000) 13313–13323.
[35] J.N. Scarsdale, G. Kazanina, S. Radaev, V. Schirch, T. Wright, Crystal structure of rabbit cytosolic serine hydroxymethyltransferase at 2.8 Å resolution: mechanistic implications, Biochemistry 38 (1999) 8347–8358.
[36] V. Trivedi, A. Gupta, V.R. Jal, P. Saravanan, G.S.J. Rao, N.A. Rao, H.S. Savithri, H.S. Subramanya, Crystal structure of binary and ternary complexes of serine hydroxymethyltransferase from *Bacillus stearothermophilus*, J. Biol. Chem. 10 (2002) 17161–17169.

[37] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[38] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Mille, D. Lipman, Gapped BLAST and PSIBLAST: a new generation of protein database search programs, J. Nucleic Acids Res. 25 (1997) 3389–3402.
[39] N. Guex, M.C. Peitsch, SWISSMODEL and the SwissPdb viewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[40] M.C. Peitsch, Protein modeling by mail, Biotechnology 13 (1995) 658–660.
[41] M.C. Peitsch, Biochem Promod and SWISS-MODEL: internet-based tools for automated comparative protein modeling, Soc. Trans. 24 (1996) 274–279.
[42] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrewski, J.A. Montgomery, R.E. Stratman, J.C. Vurant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennuci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuch, K. Raghavachari, J.B. Foresman, J. Cioslowshi, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Foz, T.M. AlLaham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Jonson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Reploge, J.A. Pople, Gaussian 98, revison A.6, Gaussian, Pittsburg, PA, 1998.
[43] W.L. Jorgenson, J. Tyrado-Reeves, The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin, J. Am. Chem. Soc. 110 (1988) 1657–1666.
[44] E. Lindahl, B. Hess, D. Van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, J. Mol. Model. 7 (2001) 306–317.
[45] H.J.C. Berendsen, D. Van der Spoel, R. Van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1995) 43–56.
[46] G.N. Ramachandran, V. Sasisekharan, Conformations of polypeptides and proteins, Adv. Protein Chem. 23 (1968) 283–437.
[47] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[48] G.J. Vriend, WHAT IF: a molecular modeling and drug design program, Mol. Graph. 8 (1990) 52–56.
[49] (http://www.biotech.embl-heidelberg.de:8400/) accessed in February 2004.
[50] M.J. Sippl, Recognition of errors in three-dimensional structures in proteins, Proteins 17 (1993) 355–362.
[51] M. Wiederstein, P. Lackner, F. Kienberger, Direct in silico mutagenesis, in: S. Brakmann, A. Schwienhorst (Eds.), Evolutionary Methods in Biotechnology, Wiley-VCH, 2004.
[52] T.F. Fu, J.N. Scarsdale, G. Kazanina, V. Schirch, H.T. Wright, Location of the pteroylpolyglutamate-binding site on rabbit cytosolic serine hydroxymethyltransferase, J. Biol. Chem. 24 (2003) 2645–2653.