Structure, evolution, and inhibitor interaction of S-adenosyl-L-homocysteine hydrolase from Plasmodium falciparum

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · SEPTEMBER 2003

Impact Factor: 2.63 · DOI: 10.1002/prot.10446 · Source: PubMed

PROTEINS: Structure, Function, and Genetics 00:000–000 (2003)

Structure, Evolution, and Inhibitor Interaction of *S*-Adenosyl-L-Homocysteine Hydrolase from *Plasmodium falciparum*

Janusz M. Bujnicki,[1]* Sean T. Prigge,[2] Diana Caridha,[3] and Peter K. Chiang[3,4]
[1]*Bioinformatics Laboratory, International Institute of Molecular and Cell Biology, Warsaw, Poland*
[2]*Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland*
[3]*Walter Reed Army Institute of Research, Silver Spring, Maryland*
[4]*National Institute of Child Health and Human Development, National Institute of Health, Bethesda, Maryland*

ABSTRACT **S-adenosylhomocysteine hydrolase (SAHH) is a key regulator of *S*-adenosylmethionine–dependent methylation reactions and an interesting pharmacologic target. We cloned the SAHH gene from *Plasmodium falciparum* (PfSAHH), with an amino acid sequence agreeing with that of the PlasmoDB genomic database. Even though the expressed recombinant enzyme, PfSAHH, could use 3-deaza-adenosine (DZA) as an alternative substrate in contrast to the human SAHH, it has a unique inability to substitute 3-deaza-(±)aristeromycin (DZAri) for adenosine. Among the analogs of DZA, including neplanocin A, DZAri was the most potent inhibitor of the PfSAHH enzyme activity, with a K_i of about 150 n*M*, whether Ado or DZA was used as a substrate. When the same DZA analogs were tested for their antimalarial activity, they also inhibited the in vitro growth of *P. falciparum* parasites potently. Homology-modeling analysis revealed that a single substitution (Thr60-Cys59) between the human and malarial PfSAHH, in an otherwise similar SAH-binding pocket, might account for the differential interactions with the nucleoside analogs. This subtle difference in the active site may be exploited in the development of novel drugs that selectively inhibit PfSAHH. We performed a comprehensive phylogenetic analysis of the SAHH superfamily and inferred that SAHH evolved in the common ancestor of Archaea and Eukaryota, and was subsequently horizontally transferred to Bacteria. Additionally, an analysis of the unusual and uncharacterized AHCYL1 family of the SAHH paralogs extant only in animals reveals striking divergence of its SAH-binding pocket and the loss of key conserved residues, thus suggesting an evolution of novel function(s). Proteins 2003;00:000–000.**

Key words: malaria; protein–ligand interactions; homology modeling; molecular phylogenetics

INTRODUCTION

Biochemically, *S*-adenosylmethionine (SAM) is the most widely used methyl donor.[1] Numerous methyltransferases (MTases) transfer the methyl group from SAM to their respective biologic acceptors, forming *S*-adenosylhomocysteine (SAH), which in turn is hydrolyzed by *S*-adenosylhomocysteine hydrolase (SAHH, EC 3.3.1.1) to adenosine (Ado) and L-homocysteine (Hcy).[2] The reaction catalyzed by SAHH is reversible, with an equilibrium constant (K_{eq}) of about 1 μ<I>M, favoring the synthesis of SAH. Under physiologic conditions, however, the removal of both Ado and Hcy is sufficiently rapid, so that the net reaction proceeds in the direction of hydrolysis.[3]

SAH is a potent inhibitor of SAM-dependent MTases. In mammals, SAHH is the only enzyme involved in SAH metabolism, which makes it is an endogenous regulator of methylation reactions and thereby an attractive pharmacologic target.[4] Indeed, deletion of the SAHH gene in mice is associated with embryo lethality,[5] and inhibitors of this enzyme affect the methylation status of nucleic acids, phospholipids, proteins, and small molecules, and have impressive antiviral, antiparasitic, antiarthritic, and immunosuppressive properties.[6]

SAHH has broad substrate specificity and can substitute Ado with a number of other nucleoside analogs (e.g., 3-deaza-adenosine, DZA), toward the formation of *S*-nucleosidylhomocysteine (e.g., DZA + Hcy ↔ DZAHcy). The enzyme can react with a number of mechanistic inhibitors, some of which covalently modify the active-site residues and are time-dependent inactivators, whereas others inhibit SAHH by steady-state kinetics.[6] Four crystal structures of rat and human SAHH have been determined, providing snapshots of conformational changes at different stages of the reaction. The substrate-free enzyme[7] [1b3r in the Protein Data Bank (PDB)] exhibits a large, open cleft between the NAD-binding domain and the catalytic domain, whereas the structures containing the

The nucleotide sequence data reported have been submitted to the GenBank® under the accession number AF525293.

Grant sponsor: J.M.B. is supported by the Young Investigator grant from EMBO and Howard Hughes Medical Institute.

*Correspondence to: Janusz M. Bujnicki, Bioinformatics Laboratory, International Institute of Molecular and Cell Biology, 02-109 Warsaw, Poland. E-mail: iamb@genesilico.pl

26 November 2002; 24 January 2003; 27 January 2003

Published online 00 Month 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.10446

inhibitor[8,9] (ADC, 1a7a; or DEA, 1k0u) or the reaction intermediate[10] (3-keto-adenosine complexed with the D244E mutant enzyme, 1d4f) have a closed conformation. Analysis of the crystal structures reveals the interactions between the protein and the ligands, and demonstrates how the adenine ring in various compounds is recognized by the active site of the enzyme, and also how the binding strength of Ado analogs is influenced by different hydrogen-bonding capabilities of their sugar moieties. It has been hypothesized that potent reversible inhibitors can fit into the active site of SAHH in the open conformation, binding tightly via all possible hydrogen bonds with the protein, and stabilize the enzyme in the closed conformation. Analysis of the protein–ligand interactions has also led to designs of novel inhibitors that bear substitutions in the adenine ring and may be even more selective.[9]

The DZA analogs possess remarkable anti-infective efficacies in vitro and in vivo.[6,11] A very potent inhibitor of SAHH, neplanocin A (NepA), has been shown to inhibit the growth of *Plasmodium falciparum* in infected red blood cells.[12] The latter result, combined with the availability of structural information, prompted us to study the conservation of protein–ligand interactions in the phylogenetic context. Our primary aim in this study was to compare the activity of selected inhibitors against the human and the cloned malarial SAHH, and to identify differences that may guide the design of novel antimalarial drugs. Secondarily, our aim was to provide a comprehensive phylogenetic analysis of the SAHH family, including possible identification of yet uncharacterized members that may be interesting from a biomedical point of view.

MATERIALS AND METHODS

Cloning of the PfSAHH Gene

The SAHH gene from *P. falciparum* is located on chromosome 5 and codes for a 479 amino acid protein with a predicted mass of 53,839 Da. We amplified the SAHH gene from gDNA of the 3D7 strain of *P. falciparum*, using PfuTurbo polymerase (Stratagene) and two primers (5′-GGTGGTGAATTCATGGTTGAAAATAAGAGCAAGG-3′ and 5′-GGTGGTGTCGACTTAATATCTGTATTCGTTACTC-3′). The resulting amplicon was digested with *EcoRI* and *SalI* endonucleases and gel-purified with the QIAquick Gel Extraction Kit from Qiagen. The digested amplicon was ligated with T4 DNA ligase (Boehringer Mannheim) into the pMAL-c2X vector (New England Biolabs), which had previously been digested with the same endonucleases and treated with Shrimp Alkaline Phosphatase (USB). This amplicon was also ligated into a similar vector (pMAL-cHT) in which the linker region was modified to contain nucleotides encoding a TEV (Tobacco Etch Virus) protease-cut site (ENLYFQG) followed by a six-histidine tag (HHHHHH). Thus, the fusion protein produced by the pMAL-cHT vector will have an amino-terminal histidine tag after cleavage by the TEV protease.

Overexpression of PfSAHH in Bacterial Culture

The pMAL-cHT vector containing the SAHH construct was transformed into BL21-CodonPlus (DE3) cells (Stratagene). These cells were then augmented with an additional plasmid encoding the TEV protease (pRK586).[13] Bacterial cultures were grown in shaker flasks at 37° to an OD_{600} of 0.6, and then induced with IPTG (Sigma) to a final concentration of 0.4 m*M*. Induced cultures were transferred to a 20° shaker and incubated for an additional 12 h. After this period, the cells were harvested by centrifugation at 5000 × G for 15 min, and the cell pellet was frozen at −20°.

Purification of Recombinant PfSAHH

Cell lysis buffer [20 m*M* Na/K phosphate pH 7.5, 1 mg/mL lysozyme (Sigma), 2.5 μg/mL DNAse I (Sigma), 200 m*M* NaCl] was added to the frozen cell pellets (20 mL per liter of original culture) and gently vortexed. Resuspended cells were incubated on ice for 10 min, followed by 30 s of sonication. Cell lysate was clarified by centrifugation at 20,000 × G for 15 min at 4° and applied to a 5 mL HiTrap Chelating HP column (Pharmacia) equilibrated in 20 m*M* Na/K phosphate (pH 7.5), and 200 m*M* NaCl. The column was washed with 10 column volumes of 10 m*M* imidazole pH 7.5, 200 m*M* NaCl, and eluted with a linear gradient to 500 m*M* imidazole (pH 7.5). Fractions containing the cut SAHH were loaded on a 5-mL HiTrap Q Fast Flow column equilibrated in 20 m*M* Bis-Tris pH 6.0. The column was washed with 10 column volumes of 20 m*M* Bis-Tris (pH 6.0) and eluted with a linear gradient to 500 m*M* NaCl.

Kinetics Characterization of Recombinant PfSAHH and Human Placental SAHH, and In Vitro Antimalarial Test

The assay conditions for PfSAHH enzyme activity were as described previously[14] with [C^{14}Ado] or [C^{14}DZA] and Hcy, and the radioactive SAH produced was separated by thin-layer chromatography (TLC; cellulose with fluorescent indicator: 2-propanol/concentrated ammonia/water, 7:1:2). After we cut the radiolabeled SAH spots from the TLC plates and placed them in scintillation vials, we next determined the radioactivity with a scintillation counter. All the kinetic constants were determined by the software Enfitter (Biosfotft, Cambridge, UK), and each mean was based on the average of at least three independent assays. The human placental SAHH (kindly provided by Michael Hershfield of Duke University). The in vitro antimalarial assay was carried out as previously described.[15] The kinetic constants were determined with the use of EnzFitter (Biosoft, Cambridge, UK) based on three independent experiments.

Sequence Database Searches

The nonredundant (nr) protein sequence database and putative translations of the EST (expressed sequence tag), STS (sequence-tagged site), HTGS (high-throughput genomic sequence), and GSS (genome survey sequence) divisions of the GenBank database and partially sequenced genomes were searched online using the position-specific iterated basic local alignment search tool (PSI-BLAST) algorithm[16] (National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov) and the Gen-

AQ: 3 eSilico server http://www.genesilico.pl). We generated multiple-sequence alignment using CLUSTALX[17] with default parameters. Manual adjustments were guided by inspection of pairwise alignments reported by PSI-BLAST and the SAHH crystal structures and homology models.

Homology Modeling

Secondary and tertiary structure prediction of PfSAHH was carried out via the MetaServer interface,[18] providing a set of target–template alignments. Homology modeling was carried out with the SWISS-MODEL/PROMOD II server,[19] with the superimposed structures of the human and rat SAHH (1a7a and 1k0u) as the templates. The homology model was energy-minimized in vacuo with use of the GROMOS forcefield[20] to relieve steric clashes. The structure was validated by VERIFY3D.[21]

Phylogenetic Analysis

Phylogenetic inference was carried out with use of the conserved blocks of the multiple-sequence alignment. The matrix of pairwise distances was calculated from sequences according to the JTT model,[22] with gaps treated as missing data. We inferred the neighbor-joining (NJ) tree using the method of Saitou and Nei.[23] The sampling variance of the distance values was estimated from 1000 bootstrap resamplings of the alignment columns (random seed value 111). All nodes with statistical support <50% were regarded as unresolved and were collapsed.

RESULTS AND DISCUSSION

Molecular Characterization of PfSAHH

The gene coding for *P. falciparum* SAHH (PfSAHH) was first cloned and shown to possess the SAHH activity by Creedon et al.[24] However, the nucleotide sequence of the PfSAHH gene published by them differs from the sequence found in the PlasmoDB genomic database. To resolve this difference, we cloned the gene from the 3D7 strain of *P. falciparum* and confirmed its 100% identity with the PlasmoDB sequence, with the Leu^{234} and Pro^{244} instead of Tyr^{243} and Thr^{244} reported previously by Creedon et al.[24] The corrected sequence has been deposited in GenBank under the accession number AF525293. PfSAHH encodes a 479 aa protein with a calculated mass of 53,838 Da and a theoretical pI of 5.64. The purified recombinant PfSAHH has an apparent mass of about 50 kDa by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) F1 (Fig. 1).

The kinetics properties of recombinant PfSAHH were T1-T2 compared with those of human placental SAHH (HsSAHH) F2-F3 (Tables I and II, Figs. 2 and 3). Although the malarial PfSAHH could replace Ado with DZA as a substrate at comparable rates (Figs. 2 and 3), it was unable to use 3-deaza-(±)aristeromycin (DZAri). In contrast, the human SAHH used Ado at a higher rate than either DZA (three times higher) or DZAri (four times higher). The K_m for PfSAHH was for 4.7 μ*M* Ado and 5.4 μ*M* for DZA (Figs. 2 and 3, Table I). For the human SAHH, the K_m values were 9.3 μ*M* for Ado, 15.8 μ*M* for DZA, and 9.1 μ*M* for DZAri (Table I). Interestingly, these results point to the distinct inability of PfSAHH to substitute Ado with DZAri, in contrast to the human placental SAHH. Both the malarial and human SAHH were unable to use NepA or 3-deazaneplanocin (DZNep) as substrates.

When tested as inhibitors of malarial PfSAHH with Ado as a substrate (Table II), the K_i was 159 n*M* for DZAri, 275 n*M* for DZNep, and 720 n*M* for NepA. In comparison, with

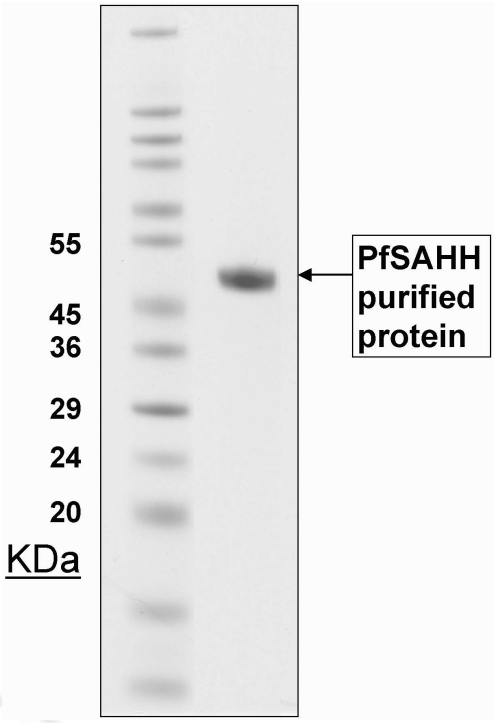


Fig. 1. Gel electrophoresis analysis of PfSAHH. Protein molecular mass standards (Sigma wide range) are shown in lane 1, and purified PfSAHH in lane 2. Samples were resolved by SDS-PAGE on a 4–12% BisTris Novex NuPAGE gel (Invitrogen) and visualized using SimplyBlue Coomassie stain (Invitrogen).

TABLE I. Kinetics Properties of PfSAHH and Human SAHH

Substrate	PfSAHH K_m (μ*M*)	PfSAHH V_{max} (μ*M*/min/mg)	Human K_m (μ*M*)
Ado	4.7 ± 0.5	0.45 ± 0.014	9.3 ± 0.9
DZA	5.4 ± 1.4	0.55 ± 0.088	15.8 ± 1.4
DZAri	Not a substrate	Not a substrate	9.1 ± 1.1

AQ: T1

TABLE II. K_i Values for Three PfSAHH Inhibitors, With DZA or Ado as a Substrate

Substrate	Inhibitor	K_i in n*M*
DZA	DZAri	141 ± 28
	Nep	373 ± 46
	DZNep	165 ± 22
Ado	DZAri	159 ± 16
	Nep	720 ± 50
	DZNep	275 ± 35

tapraid5/7e-protein/7e-protein/7e1103/7e1362d03g | royerl | S=5 | 5/16/03 | 14:20 | Art: 04192002

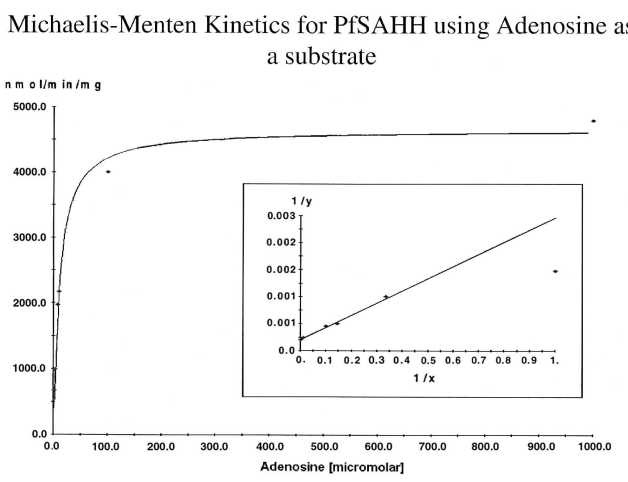


Fig. 2. Michaelis–Menten kinetics for PfSAHH with the use of Ado as a substrate.

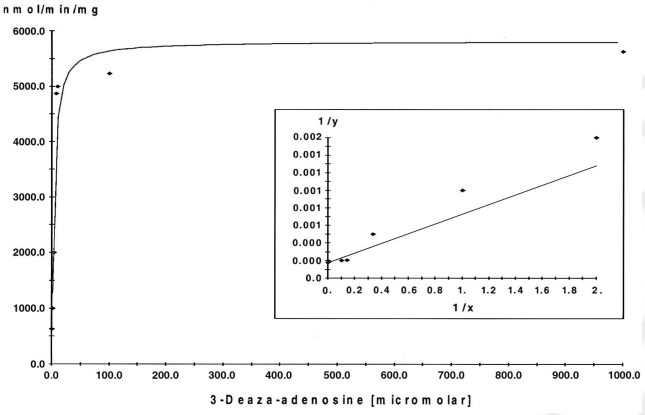


Fig. 3. Michaelis–Menten kinetics for PfSAHH with the use of DZA as a substrate.

DZA as a substrate for PfSAHH, the same analogs displayed comparable K_i values: 141 n*M* for DZAri, 165 n*M* for DZNep, and 373 n*M* for NepA (Table II). DZAri was the most potent inhibitor of PfSAHH regardless of whether we used Ado or DZA as a substrate.

Both DZA and NepA have been shown to interfere with the proliferation of *Plasmodium* parasites in vitro.[12,25] When we tested DZAri, DZNep, and NepA in a similar antimalarial in vitro assay, these compounds all inhibited the growth of the chloroquine-resistant W2 strain and the chloroquine-susceptible D6 strain of *P. falciparum*. DZNep was the most potent, with an IC_{50} value of 27 n*M* for D6 versus 191 n*M* for W2. Both NepA and DZAri showed equal potencies, with an IC_{50} value of about 400 n*M* for both the D6 and W2 strains.

Sequence Comparisons and Evolutionary Analysis

The amino acid sequence of the human HsSAHH was used to search sequence databases with PSI-BLAST to identify homologous proteins. Because we were interested in the identification of closely related enzymes with similar functions and not other enzymes [e.g., formate/glycerate dehydrogenases—closest relatives of SAHH according to structural classification of proteins (SCOP): http://www.scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.ce.bc.html], a relatively stringent cutoff of expectation (e) value = 10^{-10} was used to avoid expansion of the profile onto the whole superfamily. The search performed against AQ: 3 the nr database converged in the second iteration, yielding genuine and putative SAHH sequences. Additional searches allowed us to retrieve SAHH homologs from unfinished genomic databases, as well as from the HTGS, EST, and GSS divisions of GenBank (see Materials and Methods section). Among the reported sequences, an uncharacterized family of SAHH paralogs in animals (51% identity to the genuine SAHH) was uncovered. Multiple sequence alignment (Fig. 4) revealed several motifs conserved among all SAHH homologs, as well as sequence segments present only in certain subfamilies. Based on the F4 alignment, two classes of SAHH homologs can be identified: those that possess a characteristic insert in the middle of the sequence, and those that do not. Strikingly, members of the insert-containing class occur exclusively among Eukaryota and Bacteria, but not among Archaea. Among the insert-less SAHH homologs, another subclass can be identified, which groups together all archaeal and some bacterial members that lack the C-terminal sequence segment (Fig. 4).

To resolve the evolutionary relationships in the SAHH family, phylogenetic trees were constructed using Minimum Evolution, Maximum Parsimony, and Maximum Likelihood methods. The consensus evolutionary tree of the SAHH family (Fig. 5), revealed two primary branches (100% bootstrap support in all analyses). Although the F5 whole tree itself was unrooted, the two primary branches served as outgroups to root one another. This division separated the eukaryotic SAHH homologs from the archaeal sequences; hence, for the sake of simplicity, we hereafter refer to both sequence subfamilies (including the bacterial members present in both branches) as "eukaryotic" and "archaeal."

The inferred topology of the evolutionary tree of the SAHH family did not fully agree with the established topology of the Tree of Life, suggesting horizontal gene transfer and/or nonorthologous gene replacement. Within the archaeal branch, the primary division between Euryarchaeota and Crenarchaeota is preserved; the crenarchaeal branch received 96% bootstrap support, whereas the cluster of remaining sequences, including all euryarchaeal members, received 52% bootstrap support. Several bacterial sequences (*Thermotoga maritima*, *Desulfitobacterium hafniense*, and *Carboxydothermus hydrogeniformans*) are observed within the euryarchaeal branch, whereas Cyanobacteria form a separate branch (bootstrap support 100%) together with *Aquifex aeolicus*, *Dehalococcoides ethenogenes*, and *C. aurantiacus*. In the case of thermophilic bacteria *A. aeolicus*[26] and *T. maritima*,[27] massive horizontal transfer of genes from Archaea has been reported, and it is tempting here to speculate that all bacterial SAHH homologs within

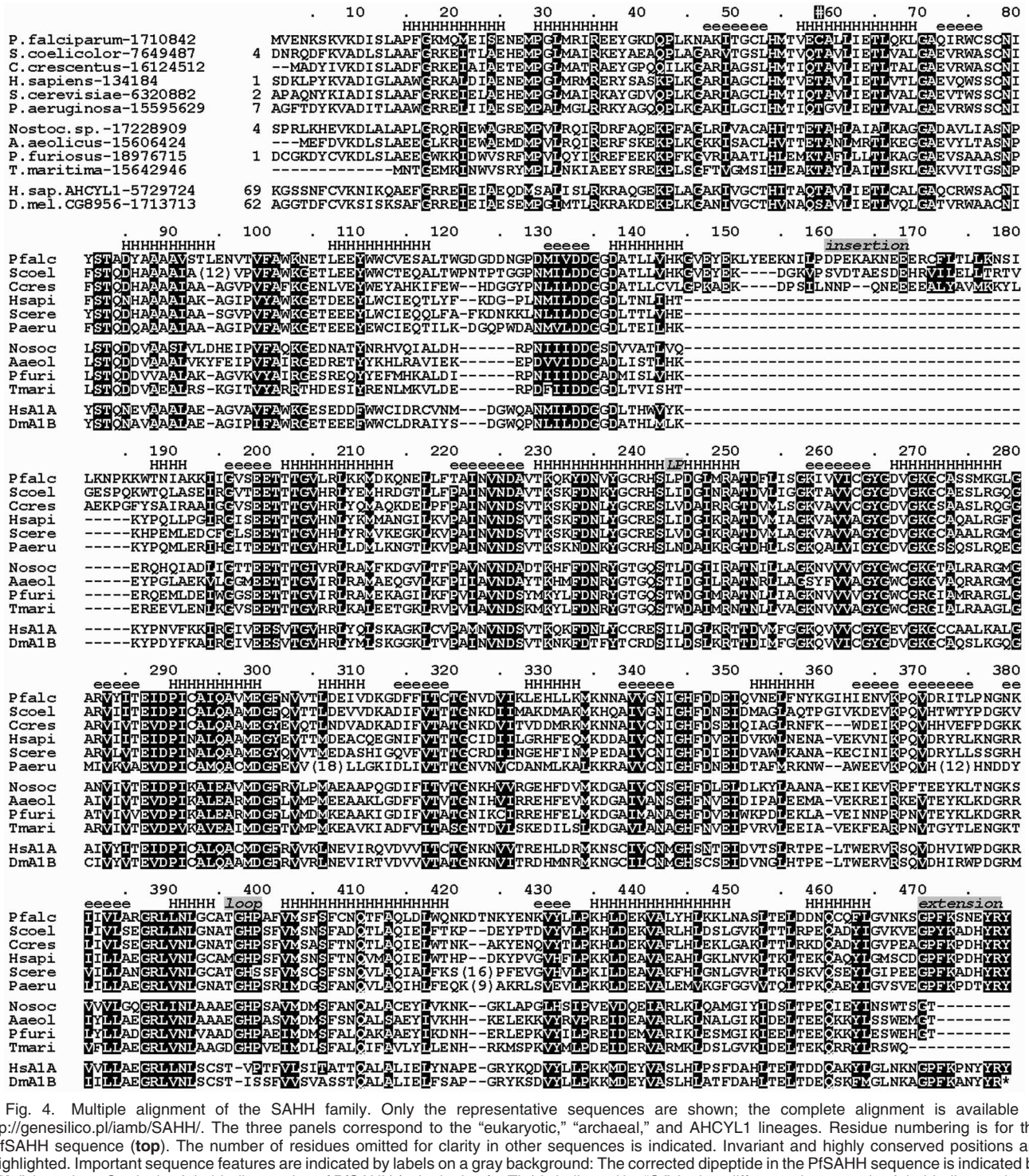


Fig. 4. Multiple alignment of the SAHH family. Only the representative sequences are shown; the complete alignment is available at ftp://genesilico.pl/iamb/SAHH/. The three panels correspond to the "eukaryotic," "archaeal," and AHCYL1 lineages. Residue numbering is for the PfSAHH sequence (**top**). The number of residues omitted for clarity in other sequences is indicated. Invariant and highly conserved positions are highlighted. Important sequence features are indicated by labels on a gray background: The corrected dipeptide in the PfSAHH sequence is indicated by "**PL**," the unique Cys in the Ado-binding pocket of PfSAHH (substituting for Thr) is indicated by "**C**," the key difference between the Ado-binding pocket of all orthodox SAHH and the AHCYL1 family is indicated by "**loop**," the insertion and the C-terminal extension are also labeled accordingly.

AQ: 3

the "archaeal" branch radiated from an ancient archaeal enzyme. Significantly, all members of this branch lack the characteristic insert present in some members of the eukaryotic branch (Fig. 4).

The phylogeny of SAHH homologs from the eukaryotic branch is even more complicated. Three main clusters include the insert-containing SAHH from *Plasmodium* (bootstrap support 100%), the intron-less AHCYL1 family of uncharacterized metazoan SAHH paralogs (bootstrap support 100%), and the main group including all other sequences (bootstrap support 97%). Within the main group, there are several well-resolved lineages, including insert-

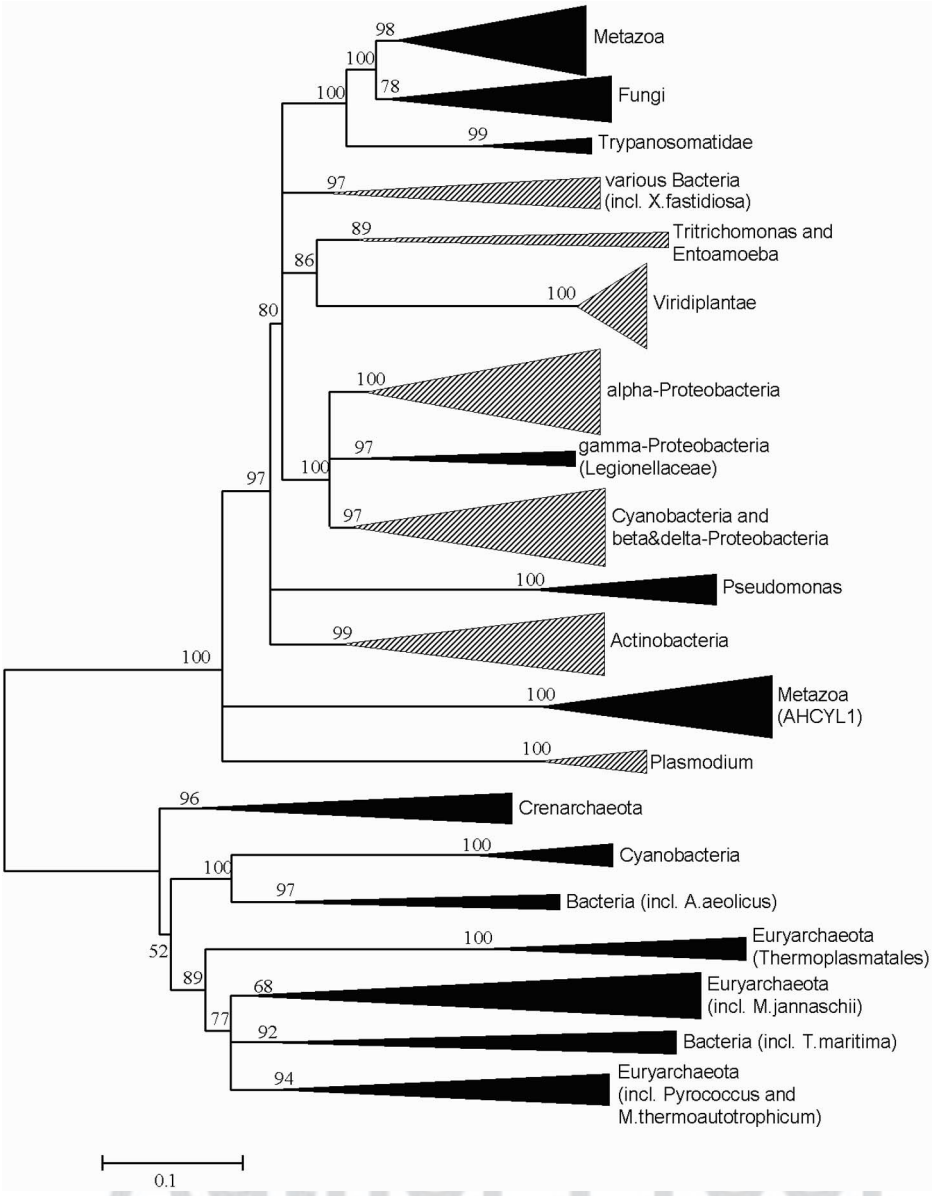


Fig. 5. The phylogenetic tree of the SAHH family. Branches corresponding to well-defined, monophyletic lineages have been collapsed and shown as triangles. Presence of the unusual insert is indicated by hatched shading, and families comprising insert-less sequences are represented by black triangles. The numbers at the nodes indicate the statistical support of the branching order by the bootstrap criterion. The nodes with bootstrap support <50% are shown as unresolved. The bar at the bottom of the phylogram indicates the evolutionary distance, to which the branch lengths are scaled based on the estimated divergence.

containing Actinobacteria, insert-less Pseudomonas, and a cluster of remaining sequences, of which some do contain the insert (green plants, several proteobacterial lineages), and some do not (animals, fungi, Trypanosomatidae, Legionellaceae). It is interesting that cyanobacterial members are present also in this branch, but the distribution of Cyanobacteria between the archaeal and eukaryotic branches is complementary (i.e., *Nostoc* sp. PCC 7120, *N. punctiforme*, and *Synechocystis* PCC6803 vs *Synechococcus* sp. PCC 7120 and *Prochlorococcus marinus*). If the AHCYL1 and bacterial sequences were excluded, the branching topology of eukaryotic SAHH homologs is largely consistent with the phylogeny of the species. In contrast, the distribution of SAHH homologs among bacteria is sporadic and inconsistent with the established microbial phylogeny.

The most parsimonious explanation for the observed pattern of distribution and phylogenetic relationships of SAHH homologs among all three Domains of Life is that SAHH evolved in the common ancestor of Archaea and Eukarya, and that after the radiation of these two domains, it was independently transferred to several bacterial lineages. Moreover, sporadic distribution of the semi-conserved insert in some lineages of the eukaryotic branch and its absence from the archaeal branch suggest that the insert appeared early in the evolution of the eukaryotic

branch and was independently lost in several lineages. Significant similarity of all insert sequences regardless of the subfamily (Figs. 4 and 5) argues against their independent acquisition; however, the possibility of a convergent evolution cannot be completely ruled out. The function of this insert remains to be demonstrated empirically, for instance, by deletion mutagenesis.

The above-mentioned scenario is also supported by the apparent synapomorphy of all members of the archaeal branch, namely, the lack of eight C-terminal amino acids typical for the other SAHH homologs. The octapeptide (F/Y)-K-(P/A/S)-(D/E/H/N)-(H/Y)-Y-R-Y is relatively well-conserved in all members of the eukaryotic lineage. In the crystal structures of rat and human SAHH, it forms a lid of the NAD-binding site, with the invariant Lys and Tyr residues making hydrogen bonds to the cofactor.[7–10] Preliminary homology modeling of the three-dimensional (3D) structure of members of the archaeal lineage suggests that deletion of this region opens the central channel in the SAHH dimer and makes the cofactor-binding sites more accessible (a representative model of *Methanococcus jannaschii* SAHH is available online at ftp://genesilico.pl/iamb/SAHH/). It will be interesting to test whether the C-terminal octapeptide is essential for the activity of the eukaryotic SAHH and how its absence does influence the biochemical properties of the archaeal members of the family.

AQ: 3

The phylogenetic position of the AHCYL1 family within the eukaryotic branch and its putative function remains unclear. It is possible that it represents a product of duplication within the metazoan lineage, which underwent accelerated evolution and, hence, its current position at the bottom of the tree is due to the long-branch-attraction artifact. The presence of two AHCYL1 gene copies both in *Homo sapiens* (localized on chromosomes 1 and 7, respectively) and *D. melanogaster*, which are more diverged in respect to one another than the genuine SAHH from human and mouse seems to support this scenario. Nevertheless, it cannot be ruled out that the AHCYL1 family represents an ancient duplication that was retained only in certain Metazoa. Strikingly, in all AHCYL1 family members, a counterpart of the otherwise invariant His residue, specifically, H^{353} in HsSAHH, is missing. The backbone of H^{353} makes key hydrogen bonds to the N7 and N6 atoms of the Ado moiety, and its sidechain delineates the bottom of the catalytic pocket.[8–10] In the corresponding loop, the AHCYL1 proteins have one residue less than the orthodox SAHH family members (Fig. 2). Because the AHCYL1 sequences are strongly divergent in this region, it is not clear what residue from the orthodox SAHH (A^{350}, M^{351}, G^{352}, or H^{353} in HsSAHH) should be deleted to produce the loop of length equal to that observed in AHCYL1. Depending on how the sequences in the corresponding region are aligned, the invariant GH dipeptide localized in the heart of the SAHH-binding site is replaced by either single aliphatic residue (V or I), or by a dipeptide TV or TI (Fig. 4). In both cases, the key His residue is substituted by an aliphatic residue, which, together with modification of the main chain (deletion of one residue), results in a significant modification of the binding pocket. Preliminary modeling of the AHCYL1 structure, based on several alternative target–template alignments, always produced structures in which the original shape of the SAHH-binding pocket was significantly altered (a representative model of HsAHCYL1 is available online at ftp://genesilico.pl/iamb/SAHH/). Hence, we hypothesize that the AHCYL1 family members may display different substrate specificity than SAHH or could represent an inactivated version of the enzyme, which could have evolved novel and unknown biologic functions. It is noteworthy that the C-terminal octapeptide and other elements of the NAD-binding site are retained in the AHCYL1 family, suggesting that these proteins retain the ability to bind the cofactor. Only in the *D. melanogaster* CG8956 protein the C-terminal Tyr is replaced by a unique extension "LVTLLSLSILHSS."

AQ: 4

AQ: 3

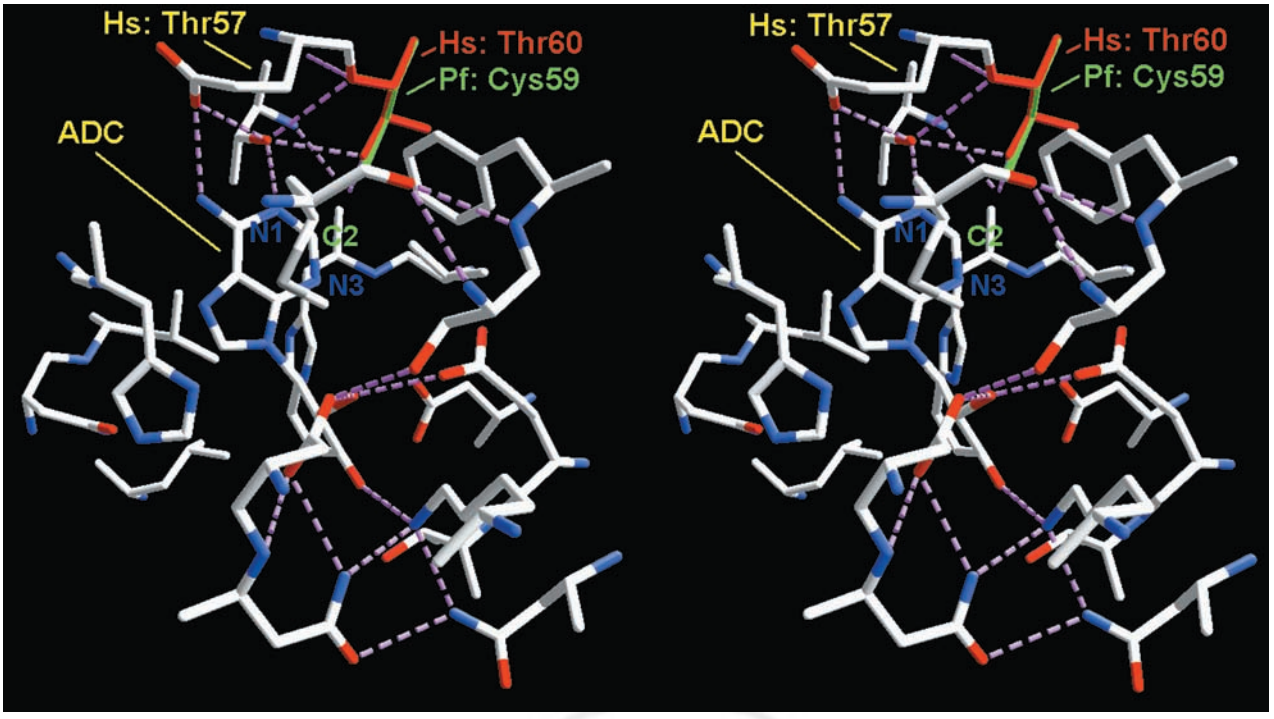
Homology Modeling

The structure of PfSAHH was homology-modeled based on the superimposed closed forms of human and rat SAHH structures (1a7a and 1k0u), according to the protocol described in the Materials and Methods section. The insert (PfSAHH residues 145–185) could not be modeled with confidence and was therefore omitted from the model, because the mammalian SAHH structures lack the corresponding segment (see above), and prediction of such large polypeptide fragments ab initio is currently not plausible. The ligand molecules (NAD and Ado) were copied to the PfSAHH model from the coordinates of the template. The final model of the PfSAHH dimer was energy-minimized to alleviate steric clashes; it passed the quality tests implemented in Verify3D (average score 0.442 and no dips below 0). The model is available online at ftp://genesilico.pl/iamb/SAHH/.

AQ: 3

There are only three differences between the binding pockets of PfSAHH and HsSAHH. Two of them map to the NAD-binding region, and both the PfSAHH- and HsSAHH-like amino acids in these positions [Asn (Pf) vs Cys (Hs) in one case, and Cys vs Asn in the other] do occur in various organisms (Fig. 4). In other words, SAHH can apparently have C or N, or even other amino acids in these semiconserved positions. However, one of the amino acids in the vicinity of the substrate-binding site, a nearly invariant Thr residue (T60 in HsSAHH), is substituted by Cys in *P. falciparum* (C59) and *P. knowlesi* (Fig. 4). In HsSAHH, the conserved Thr does not interact directly with the substrate but participates in a hydrogen-bond network that stabilizes Ado. It has been noted that the N3 atom of Ado is not involved in any hydrogen bonding, and there is a relatively large space in front of N3, which could be utilized by an additional functional group of a novel SAHH inhibitor[9] (Fig. 6). Our model does not provide a straightforward explanation for different interactions of human and malarial SAHH with the inhibitors, because the difference between DZAri and DZA is located in the ribose ring, positioned away from T60 in HsSAHH.[9] Nonetheless, the sulfhydryl group unique to the SAHH from *Plasmodium* may represent a target for specific inhibitors. For instance,

F6



COLOR

Fig. 6. Crossed stereoview of the ligand-binding site of HsSAHH (1a7a). Only the residues within the distance of 5Å from the inhibitor are shown. The pair of nonidentical residues is shown in color (PfSAHH, green; HsSAHH, red) and labeled. Other residues (identical between HsSAHH and PfSAHH) and the inhibitor are colored according to their atomic structure (carbon atoms, white; oxygen, red; nitrogen, blue; sulfur, yellow). The N1, C2, and N3 atoms of the inhibitor are labeled.

a reagent able to form a covalent bond with C59 may become a potent inactivator highly specific to malarial SAHH.

CONCLUSIONS

SAHH is an attractive pharmacologic target, because its inhibition can affect the methylation status of nucleic acids, phospholipids, proteins, and small molecules. We cloned the gene for *P. falciparum* SAHH with the correct sequence and expressed this enzyme in bacterial culture. The purified recombinant PfSAHH can use Ado or DZA as substrate, but unlike the human SAHH, the malarial ortholog was unable to use DZAri as a substrate. The compounds DZAri, NepA, and DZNep all inhibited PfSAHH with K_i values in the range of 141–720 n*M*. These compounds also blocked the growth of malaria parasites in cultures with IC_{50} values of 27–400 n*M*. The unique inability of PfSAHH to use DZAri as a substrate, combined with the observation that DZAri was the most potent in vitro inhibitor of the growth of *P. falciparum*, points to the potential exploitation of DZAri for the designing of the next generation of PSAHH inhibitors.

After we performed a comprehensive phylogenetic analysis of the SAHH superfamily, we deduced three key findings: (1) SAHH evolved in the common ancestor of Archaea and Eukaryota, and was subsequently horizontally transferred to Bacteria on multiple occasions and independently from both these lineages; (2) the AHCYL1 family represents unusual metazoan SAHH paralogs, in which the Ado-binding loop has strongly diverged, suggesting the evolution of a novel specificity; (3) members of the archaeal branch lack the C-terminal segment, which in eukaryotic homologs forms a "lid" of the NAD-binding site.

Remarkably, homology-modeling analysis revealed that a nearly universally conserved Thr residue in the binding pocket of SAHH is substituted by Cys in the malarial enzyme. Whether this sole substitution is responsible for the differential reactivity of PfSAHH and HsSAHH toward some Ado analogs remains to be investigated. More importantly, the discovery of a unique sulfhydryl group in the active site of malarial SAHH may be exploited in developing novel drugs that selectively inhibit PfSAHH.

REFERENCES

1. Cheng X, Blumenthal RM. *S*-adenosylmethionine-dependent methyltransferases: Structures and functions. Singapore: World Scientific Inc.; 1999.
2. Chiang PK, Gordon RK, Tal J, Zeng GC, Doctor BP, Pardhasaradhi K, McCann PP. *S*-adenosylmethionine and methylation. FASEB J 1996;10:471–480.
3. Richards HH, Chiang PK, Cantoni GL. Adenosylhomocysteine hydrolase: Crystallization of the purified enzyme and its properties. J Biol Chem 1978;253:4476–4480.
4. de la Haba G, Cantoni G. The enzymatic synthesis of *S*-adenosyl-L-homocysteine from adenosine and homocysteine. J Biol Chem 1959;234:603–608.
5. Miller MW, Duhl DM, Winkes BM, Arredondo-Vega F, Saxon PJ, Wolff GL, Epstein CJ, Hershfield MS, Barsh GS. The mouse lethal nonagouti [a(x)] mutation deletes the *S*-adenosylhomocysteine hydrolase (Ahcy) gene. EMBO J 1994;13:1806–1816.
6. Chiang PK. Biological effects of inhibitors of *S*-adenosylhomocysteine hydrolase. Pharmacol Ther 1998;77:115–134.
7. Hu Y, Komoto J, Huang Y, Gomi T, Ogawa H, Takata Y, Fujioka

M, Takusagawa F. Crystal structure of *S*-adenosylhomocysteine hydrolase from rat liver. Biochemistry 1999;38:8323–8333.
8. Turner MA, Yuan CS, Borchardt RT, Hershfield MS, Smith GD, Howell PL. Structure determination of selenomethionyl *S*-adenosylhomocysteine hydrolase using data at a single wavelength. Nat Struct Biol 1998;5:369–376.
9. Huang Y, Komoto J, Takata Y, Powell DR, Gomi T, Ogawa H, Fujioka M, Takusagawa F. Inhibition of *S*-adenosylhomocysteine hydrolase by acyclic sugar adenosine analogue *D*-eritadenine: Crystal structure of *S*-adenosylhomocysteine hydrolase complexed with *D*-eritadenine. J Biol Chem 2002;277:7477–7482.
10. Komoto J, Huang Y, Gomi T, Ogawa H, Takata Y, Fujioka M, Takusagawa F. Effects of site-directed mutagenesis on structure and function of recombinant rat liver *S*-adenosylhomocysteine hydrolase: Crystal structure of D244E mutant enzyme. J Biol Chem 2000;275:32147–32156.
11. Prigge ST, Chiang PK. *S*-Adenosylhomocysteine hydrolase. In: Carmel R, Jacobsen D, editors. Homocysteine in health and disease. Cambridge, UK: Cambridge University Press; 2001. p 79–91.
12. Whaun JM, Miura GA, Brown ND, Gordon RK, Chiang PK. Antimalarial activity of neplanocin A with perturbations in the metabolism of purines, polyamines and *S*-adenosylmethionine. J Pharmacol Exp Ther 1986;236:277–283.
13. Kapust RB, Waugh DS. Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr Purif 2000;19:312–318.
14. Chiang PK. *S*-adenosylhomocysteine hydrolase: Measurement of activity and use of inhibitors. In: Paton OM, editor. Methods in pharmacology: Vol. 6. Methods used in adenosine research. New York: Plenum Press; 1985. p 127–145.
15. Zhang P, Nicholson DE, Bujnicki JM, Su X, Brendle JJ, Ferdig M, Kyle DE, Milhous WK, Chiang PK. Angiogenesis inhibitors specific for methionine aminopeptidase 2 as drugs for malaria and leishmaniasis. J Biomed Sci 2002;9:34–40.
16. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.
17. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876–4882.
18. Bujnicki JM, Elofsson A, Fischer D, Rychlewski L. Structure prediction Meta Server. Bioinformatics 2001;17:750–751.
19. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 1997;18:2714–2723.
20. Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF. The GROMOS biomolecular simulation program package. J Phys Chem 1999;103:3596–3607.
21. Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature 1992;356:83–85.
22. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 1992;8:275–282.
23. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–425.
24. Creedon KA, Rathod PK, Wellems TE. *Plasmodium falciparum S*-adenosylhomocysteine hydrolase: cDNA identification, predicted protein sequence, and expression in *Escherichia coli*. J Biol Chem 1994;269:16364–16370.
25. Trager W, Tershakovec M, Chiang PK, Cantoni GL. *Plasmodium falciparum*: Antimalarial activity in culture of sinefungin and other methylation inhibitors. Exp Parasitol 1980;50:83–89.
26. Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham DE, Overbeek R, Snead MA, Keller M, Aujay M, et al. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature 1998;392:353–358. **AQ: 5**
27. Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Nelson WC, Ketchum KA, et al. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature 1999;399:323–329. **AQ: 6**