

# Molecular evolution-directed approach for designing archaeal formyltetrahydrofolate ligase

[Moleküler evrim güdümlü yaklaşım ile arkeal formiltetrahidrofolat ligaz dizaynı]

Paulchamy Chellapandi<sup>1</sup>,  
Jayachandrabal Balachandramohan<sup>1</sup>

<sup>1</sup>Department of Bioinformatics, Centre for Excellence in School of Life Sciences, Bharathidasan University, Tiruchirappalli-620024, Tamil Nadu, India

Yazışma Adresi  
[Correspondence Address]

Paulchamy Chellapandi

Assistant Professor

Department of Bioinformatics, Centre for Excellence in School of Life Sciences, Bharathidasan University, Tiruchirappalli-620024, Tamil Nadu, India

Tel: +91-431-2407071

Fax: +91-431-2407045

Email: pchellapandi@gmail.com

Registered: 20 May 1010; Accepted: 28 March 2011

[Kayıt Tarihi : 20 Mayıs 2010; Kabul Tarihi : 28 Mart 2011]

## ABSTRACT

**Objective:** The objective was to design a biocatalyst formyltetrahydrofolate ligase from the sequences of archaea bacteria based on the evolutionary conservation at metal- and substrate-binding regions.

**Methods:** Complete formyltetrahydrofolate ligase sequences of archaea were retrieved from GenPept of National Center for Biotechnology Information (NCBI). The best structure identity and the shortest metal- and substrate-binding regions with templates were concerned for generating homology models and then selected for docking with respective substrates using AutoDock. Every homology model of this study was energy minimized to generate the lowest energetic conformers. The best enzyme-substrate complex models were chosen based on the mode of catalysis, types of molecular interactions, and binding affinity.

**Results:** Two stable biocatalyst variants were developed from the sequences of *Haloquadratum walsbyi* DSM 16790 and *Methanocorpusculum labreanum* Z. The binding energy of variant 1-substrate complex and variant 2-substrate complex were ranged from -5.74 to -3.27 kcal/mol and from -5.23 to -2.78 kcal/mol, respectively. The most important interactions contributing to the high binding affinity occur between 5, 10-methenyltetrahydrofolate and the side chains of Glu113, Ser115 and Asp147 in variant 1 and Glu100, Ser102 and Asp134 in variant 2. Both variants have cesium and substrate-binding sites within 30 amino acids length wherein enzyme active sites residues were noted and thus evolutionary conservation in sequence as well as structure would make contribution in enzyme catalysis.

**Conclusion:** This computer-aided protein designing approach is merely based on evolutionary hypothesis. It is also constructive to develop a stable biocatalyst variant with more catalytic efficiency because original conserved amino acids residues are preserved as it is. Both variants have catalytic functions as similar to 5-hydrolyase, ammonia forming and amino lyase on 5, 10-methenyltetrahydrofolate.

**Keywords:** Molecular docking; Molecular dynamics; Formyltetrahydrofolate ligase; Archaea; Molecular evolution; Enzyme design

**Conflict of Interest:** Authors have no conflict of interest.

## ÖZET

**Amaç:** Metal ve substrat bağlanma bölgelerindeki evrimsel konservasyon temel alınarak arkeal bakterilerin sekanslarından biyokatalizör formiltetrahidrofolat ligazın dizayn edilmesi amaçlanmıştır.

**Yöntemler:** Arkeal formiltetrahidrofolat ligazın sekanslarının tamamı Ulusal Biyoteknoloji Enformasyon Merkezi (*National Center for Biotechnology Information* – NCBI)'nde bulunan GenPept'ten elde edildi. Şablonlardaki en iyi yapısal birim ve en kısa metal ve substrat bağlayıcı bölgeler homolog modellerin oluşturulması ve daha sonra AutoDock kullanılarak ilgili substrat ile bağdaştırılması için kullanıldı. Çalışmadaki her bir homolog modelinin enerjisi en düşük enerjili konformerlerin oluşturulması için en aza indirildi. En iyi enzim-substrat kompleks modelleri katalizleme moduna, moleküler etkileşimlerin tipine ve bağlanma eğilimine göre seçildi.

**Bulgular:** *Haloquadratum walsbyi* DSM 16790 ve *Methanocorpusculum labreanum* Z sekanslarından iki adet stabil biyokatalizör varyantı geliştirildi. Varyant 1-substrat kompleksi ve varyant 2-substrat kompleksinin bağlanma enerjileri sırasıyla -5.74 ile -3.27 kcal/mol ve -5.23 ile -2.78 kcal/mol arasındaydı. Yüksek bağlanma eğilimine katkıda bulunan en önemli etkileşimler 5,10 meteniltetrahidrofolat ve varyant 1'deki Glu113, Ser115 ve Asp147 yan zincirleri ile varyant 2'deki Glu100, Ser102 ve Asp134 yan zincirleri arasında gözlemlendi. Sekansın evrimsel konservasyonu kadar yapısının da enzimin katalizlemesine katkıda bulunmasını sağlayacak şekilde her iki varyantta da sezyum ve substrat bağlanma bölgelerinin 30 amino asit uzunluğundaki bölge içerisinde yer aldığı belirlenmiş, aynı bölgede enzimin aktif bölgelerinin rezidüleri de saptanmıştır.

**Sonuç:** Sunulan bilgisayar aracılığıyla protein dizayn edilmesi yaklaşımı tamamıyla evrim hipotezine dayanmaktadır. Bu yöntem, orijinal amino asit parçaları korunduğundan, daha fazla katalitik etkinliğe sahip stabil biyokatalizör varyantının geliştirilmesi açısından da faydalıdır. Her iki varyantın katalitik fonksiyonu 5,10 meteniltetrahidrofolattan amonyak oluşturan ve amino grubunu ayıran 5-hidrolyaza benzerdir.

**Anahtar Kelimeler:** Moleküler bağdaştırma, moleküler dinamik, formiltetrahidrofolat ligaz, arkeal, moleküler evrim, enzim dizaynı

**Çıkar Çatışması:** Yazarların çıkar çatışması bulunmamaktadır.

## Introduction

A wide range of archaeon enzymes' application and the use of the organisms themselves in biotechnology are restricted due to complicated purification strategies and lack of expression systems [1]. However, the increasing interest in enzyme applications in industrial processes has spurred the search for biocatalysts with new or improved properties [2,3]. The use of biotransformation in industry will increase and it has been claimed that the number of industrially established biocatalytic processes will double every decade [4-6]. Unfortunately, naturally available enzymes are usually not optimally suited for industrial applications due to the stability of the enzymes under process conditions. With the advances in protein engineering technologies, a variety of enzyme properties can be altered simultaneously, if the appropriate screening parameters are employed. Although it sometimes is beneficial to adapt industrial processes to mild and environmentally benign conditions favored by the enzyme, the use of more extreme conditions is often desirable [7]. The development and use of computational methods for searching the sequence space associated with a particular target structure has led to successful designs of small (less than 75 residues) monomeric proteins [8, 9], protein oligomers [10], and the redesign of natural proteins to confer novel functionalities [11, 12].

Formyltetrahydrofolate ligase (FTHFL) (EC 6.3.4.3) belongs to the family of ligases, specifically those forming generic carbon-nitrogen bonds. The systematic name of this enzyme class is formate: tetrahydrofolate and other commonly used names include formyltetrahydrofolate synthetase, 10-formyltetrahydrofolate synthetase, tetrahydrofolic formylase, and tetrahydrofolate formylase. FTHFL catalyzes the ATP-dependent activation of formate ion via its addition to the N10 position of tetrahydrofolate. It is a highly expressed key enzyme in both the Wood-Ljungdahl pathway of autotrophic CO<sub>2</sub> fixation (acetogenesis) and the glycine synthase/reductase pathways of purinolysis. FTHFL has a key physiological role in acetogens to catalyze the formylation of tetrahydrofolate, an initial step in the reduction of carbon dioxide and other one-carbon precursors to acetate. Moreover, the enzymatic reaction can be reversed in purinolytic organisms on liberating formate from 10-formyltetrahydrofolate with concurrent production of ATP [13]. The monovalent cations potassium, ammonium and cesium were reported for maximal thermostability and improved catalytic activity of this enzyme [14, 15]. X-ray crystallography studies have revealed the 3D structures of FTHFL with cesium, potassium and ammonium ions [16]. The side chain of Glu98 is conserved in all known bacterial FTHFL sequences that can be participated in metal ion binding. Other ligands in the Cs<sup>+</sup> binding site are four oxygen atoms of main chain carbonyls and water molecules. When FTHFL complexes with ammonium and cesium ions, minor structural differences were observed, even these ions bind to different sites [17].

The maturation of enzyme technology is shown by the development of the theory concerning structure-function relationship of enzymes and how this is related to their primary structure through the formation and configuration of 3D structure. The binding of a substrate close to functional groups in the enzyme causes catalysis by so-called proximity effects. Active biocatalysts generated from dramatically reduced amino acid alphabets has provided a strong support for the idea that primordial enzyme can be made from only a handful of building blocks [18]. It was stated that the ability to bind substrate with catalytic functional groups can be promising when similar catalysts obtained from small molecule mimics of enzyme active sites combined in a small molecule. The present study was aimed to identify conserved domain architecture enabling to catalyze metal-dependent biochemical reaction, generate 3D model structure of enzyme variant from the sequence carrying catalytic domain, and calculate binding energy of enzyme variant-substrate complexes. Thus, evolutionary conservation of FTHFL sequences at metal- and substrate-binding regions has to be considered for designing enzyme variants with wide substrate specificity.

## Materials and Methods

### *Evolutionary conservation analysis*

Complete archaeal FTHFL sequences were retrieved from GenPept of National Center for Biotechnology Information-NIH (NCBI). Conserved domains architecture of these sequences were searched from NCBI-Conserved Domain Database (NCBI-CDD) [19] using conserved domain (CD) search tool [20] with expected value threshold 10 and a low complexity filter. Metal- and substrate-binding sites from CDs were compared with available Protein Data Bank (PDB) structures. The parameter was 10 expected threshold, 3 word size, 11 gap existence cost, 1 gap extension cost, Blossum 62 scoring matrix, conditional compositional score alignment, and automatically adjust the parameter for short input sequences. Query sequences were compared to a position-specific score matrix prepared from the underlying conserved domain alignment. The selected sequences were clustered together with complete deletion of gaps and multiple substitutions were corrected using ClustalX 2.0 software [21]. The aligned sequences were iterated at each alignment step and manually inspected to delete the low scoring sequences. Conserved domain regions of selected sequences were aligned with sequences of similarity structures and related sequences. Neighbor joining (NJ) algorithm was used to search homogeneous patterns among all lineages and then an unrooted phylogenetic tree was built by using MEGA 4.0 software [22] with 1000 bootstraps values [23], JTT model along 0.25 gamma distributions, at uniform rates among sites.

## ***Molecular modeling and enzyme designing***

ModWeb is an automatic comparative protein modeling server where the same sequences were uploaded to build 3D structure based on a template [24]. The structural templates used to build models in ModPipe consist of a set of non-redundant chains extracted from structures in the PDB. Sequence-structure matches were established using multiple variations of sequence-sequence, profile-sequence, sequence-profile and profile-profile alignment methods. A suitable template (homolog) for 3D structure modeling from these sequences was searched against PDB using PSI-BLAST tool [25] with a default parameter. Protein models were built for each one of the sequence-structure matches using comparative modeling by satisfaction of spatial restraints as implemented in Modeller [26]. Finally, the resulting models were evaluated using Structural Analysis and Verification Server (SAVS) (<http://nihserver.mbi.ucla.edu/SAVES>), and the best scoring models selected to superimpose on the corresponding PDB templates with Dali pairwise comparison tool in DALITE server ([http://ekhidna.biocenter.helsinki.fi/dali\\_lite/start](http://ekhidna.biocenter.helsinki.fi/dali_lite/start)). The function and active sites (nests) of these models were predicted by ProFunc [27]. ProFunc server helps to identify the likely biochemical function of a protein from its 3D structure. It uses a series of methods including fold matching, residue conservation, surface cleft analysis, and functional 3D templates to identify both the protein's likely active site and possible homologues in the PDB. These models were compared with crystallographic protein structures whose catalytic domains are similar to metal- and substrate-binding sites. Atomic coordinates of amino acid residues in the models which are not covering the position of metal- and substrate-binding sites and also active sites have been removed.

## ***Molecular dynamics simulation of enzyme constructs***

Standard dynamic simulation cascade module in Discovery Studio software was used to obtain structural conformers of these models (enzyme variants) by using CHARMM force field and steepest descent as well as adopted basis Newton-Raphson (NR) algorithms. Distance constraint was between N-terminal to C-terminal and dihedral restraint was started from C to C $\alpha$  ( $\Phi$ ) of first amino acid residue and C $\alpha$  to N ( $\psi$ ) of second amino acid residue until the last amino acid residue in a molecular dynamic ensemble. The top-five lower energy conformers were selected for computing molecular interaction energies. The parameter set for minimization 1 was 500 steps, 0.1 RMS (root mean square) gradient, distance and dihedral restraints as constraints in steepest descent algorithm, while parameter set for minimization 2 was 500 steps, 0.0001 RMS gradient and distance and dihedral restraints as constraints in adapted basis NR algorithm. In heating and equilibrium processes, 2000 steps, 0.001 time steps, 50 initial temperature,

300 target temperature and 50 adjust velocity were used. NVT ensemble and 50 temperature coupling decay time were used for production steps in MD simulation. Generalized Born with a simple switching implicit solvent model was used with spherical cut-off (electrostatic), 80 implicit dielectric constant, 1 dielectric constant, 0 salt concentration and non-polar surface area for conformational analysis.

## ***Molecular docking studies***

Genetic algorithm [28] and AMBER force field implemented in Autodock software 4.0 was used for finding inhibition constants, free energy of binding, intermolecular and internal energies of the selected lower energy conformers with substrates. The related structures to substrate were searched from KEGG database using SIMCOM software (<http://www.genome.jp/tools/simcomp>) in MOL2 files and then converted to PDB format. An enzyme variant was fixed as rigid while substrate was set as flexible in docking process. Binding site (cavity) of each variant was selected in an energy grid and a flexible substrate preferred to dock into it. The molecular mechanics-based and empirical terms were multiplied by coefficients that are determined by linear regression analysis of complexes with known 3D structures and known binding free energies. The desolvation free energy term, which was based on new atomic solvation parameters, depends on the absolute partial charge on the atom. Genetic algorithm (GA) parameter was 10 GA runs, 150 population size, 250000 energy evaluations, 27000 generations, 0.02 rate of gene mutation, 0.8 rate of cross over, and 1.0 variance of Cauchy distribution for gene mutation. Minimization of docked models was performed with smart energy minimization algorithm to refine the orientation of the substrate in the receptor site. After docking, this model was solvated with explicit periodic boundary solvation model. RMSD (root mean square deviation) is a frequently used measure of the differences between values computed by a model or an estimator, and the values actually observed from the thing being modeled or estimated. Thus, RMSD calculation was performed to evaluate the docking models.

## **Results**

### ***Evolutionary conservation analysis of FTHFL domains***

The FTHFL sequences were examined in different genera of archaeal domain. There was a little knowledge about crystallographic structures information of this enzyme from archaea so that its 3D structures has predicted from the sequences. The structural identity of FTHFL was ranged from 49 to 64% with PDB templates. NCBI accession Q18JB5 (enzyme variant 1) and A2SS72 (enzyme variant 2) were selected for further study as they showed maximum sequence identity, modeling score, and have the shortest length in metal binding position

comprising active sites. Other homology models were neglected due to the low structural quality and wide positions in the selected CDs (beyond the metal-binding position) (Tables 1 & 2). Variant 1 and variant 2 have metal-binding domain at the positions 110-140 and 100-130, respectively. Both variants have one functional FTHFL domain showing the similarity to P-loop NTPase and PRK13507 super families, which can necessary to carry out its catalytic function on substrate 5,10 formyltetrahydrofolate (Figure 1).

### **Phylogenetic analysis of FTHFL sequences**

As shown in Figure 2, FTHFL was extensively dispersed in thermophilic (*Thermoplasma volcanium*, *T. acidophilum*, *Picrophilus torridus*), methanogenic (*Methanococcus labreanum*) and halophilic (*Haloquadratum walsbyi*, *H. marismortui*, *H. lacusprofundi*) archaea. FTHFL sequences from different species were separately clustered in the phylogenetic tree with a considerable bootstrap confidence. Both variants were evolutionarily related only within the same species of organisms. It revealed that its enzyme function can be conserved within the particular sequence position and or a domain. Phylogenetic analysis of this enzyme suggested a high likelihood of using evolution-based enzyme designing approach as FTHFL sequences of archaea have more conservation at substrate- and metal-binding domains in CDs.

### **Molecular dynamic studies of enzyme variants**

Thirty energy conformers from each variant were generated by standard molecular dynamic simulation from which the top-five lower energy conformers were selected (Table 3). Torsion energy to every conformers of variants have ranged from 76 to 95 kcal/mol. Electrostatic energy of each conformer was higher than van der Waals energy of them. Among overall conformers, the total energy of variant showed the best energy score (-847 kcal/mol), which was somewhat better than variant 2 (-735 kcal/mol).

### **Structural quality and accuracy of enzyme variants**

Homology models of each variant was evaluated and validated for the structural quality and accuracy using Verify3D, ERRAT and Prove algorithms (Figure 3 & supplementary). Verify 3D results showed that 97.87% of the residues in variant 1 and 93.65% in variant 2 had an averaged 3D-1D score >0.2. As shown in Figure 4, Ramachandran plot indicated that core, allowed, general and disallowed residues were 89.8%, 8.6%, 1.3% and 0.4%, respectively, in the variant 1; whereas, 89.7%, 8.4%, 1.3% and 0.6%, respectively, in the variant 2. Overall quality factor computed by ERRAT2 program was 75.91% for variant 1 and 79.36% for variant 2. It is expressed as percentage of the protein for which the calculate error value falls below the 95% rejection limit. Good high resolution structures generally produce values around

and 95% or higher. For lower resolutions (2.5 to 3 Å) the average overall quality factor is around 91%. When entire structures were analyzed by Prove program, Z-score mean was reported as 1.902 and 2.409 for variant 1 and 2, respectively.

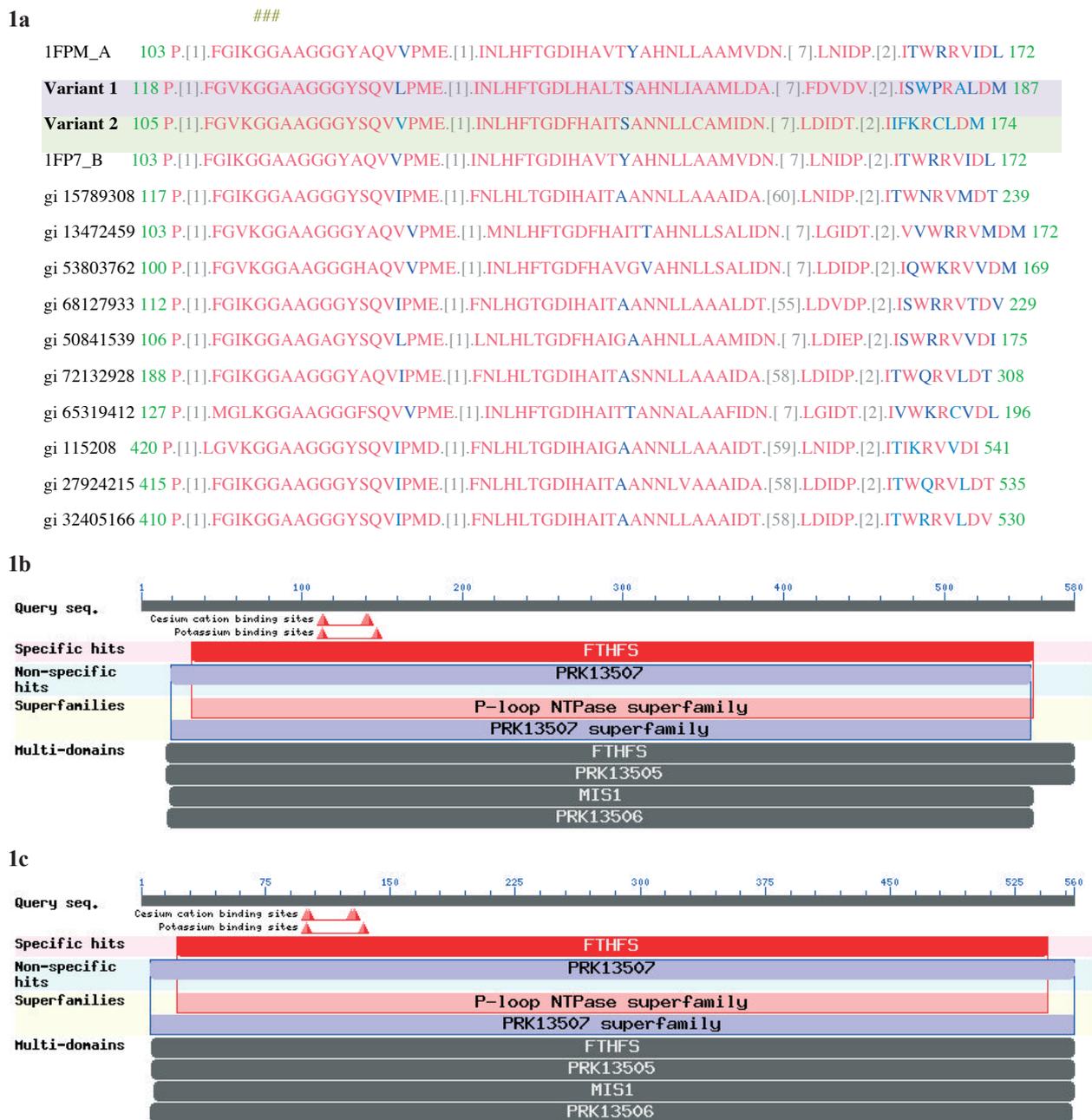
The pairwise structural similarity of each variant was analyzed by structurally superimposing it with a corresponding PBD template, and the resulted sequence and structural alignment are depicted in Figures 5 & 6. Z-score, RMSD and structural identity of variant 1 with PDB template were 66.6, 0.6 and 53%, respectively; whereas, that of variant 2 were 65.7, 0.2 and 64%, respectively.

Variant 1 has 17 H-bonds donors, 1 helices, 8 turns and 2 strand in its 3D structure while in variant 2, 22 H-bonds donors, 3 helix, 6 turns, and 0 strands were observed (Figure 7). These findings supported that more number of hydrogen donors of both variants could contribute on hydrogen bonding with their substrates.

### **Molecular docking studies of enzyme variant-substrate complexes**

The 10-formyltetrahydrofolate, 10-formyldihydrofolate, 10-formyltetrahydrofolyl L-glutamate, 5-methyltetrahydrofolate and 5,10-methenyltetrahydrofolate was substrates chosen to dock into enzyme variants (Table 4). The binding energy of variant 1-substrate complexes and variant 2-substrate complexes ranged from -5.74 to -3.27 kcal/mol and -5.23 to -2.78 kcal/mol, respectively. Variant 1 has showed as the best docking model, which has 60.44 Å RMSD, 1.09 mmol inhibition constant, -5.92 kcal/mol intermolecular energy and -5.00 kcal/mol internal energy when it interacted with 5,10-methenyltetrahydrofolate. Variant 2 was also favorably interacted with 5-methyltetrahydrofolate (-5.23 kcal/mol) and 5,10 methenyltetrahydrofolate (-4.07 kcal/mol). As a result, variant 1 was assumed to catalyze formyltetrahydrofolyl L-glutamate (-5.74 kcal/mol) and 10-formyltetrahydrofolate (-4.04 kcal/mol).

A suitable enzyme variant-substrate complex was chosen based on the mode of catalysis, types of molecular interactions, and binding affinity (Figure 8A). A strong binding affinity was measured between carbonyl group of 5,10-methenyltetrahydrofolate and the side chains of Glu113, Ser115 and Asp147 in variant 1, and its interaction sites were; H-Glu113 (3.04 Å), O-Ser115 (3.09 Å) O-Ser115 (3.01 Å) and O-Asp147 (2.78Å). Moreover, Asp147 residue formed three H-bonds with substrate and also was reported as a primary covalent attachment site (active site residue) to 10-formyltetrahydrofolate in variant 1. As represented in Figure 8B, five molecular interactions were formed between 5,10-methenyltetrahydrofolate and variant 2 at atomic positions of N<sub>3</sub>-Glu100 (3.22 Å), O-Glu100 (3.05 Å), O-Ser102 (2.95Å), N<sub>2</sub>-Asp134 (2.53 Å) and N<sub>2</sub>- Asp134 (2.67 Å). As similar to variant 1, Asp134 was reported as active site residue of this enzyme so that variant 2-substrate complex was



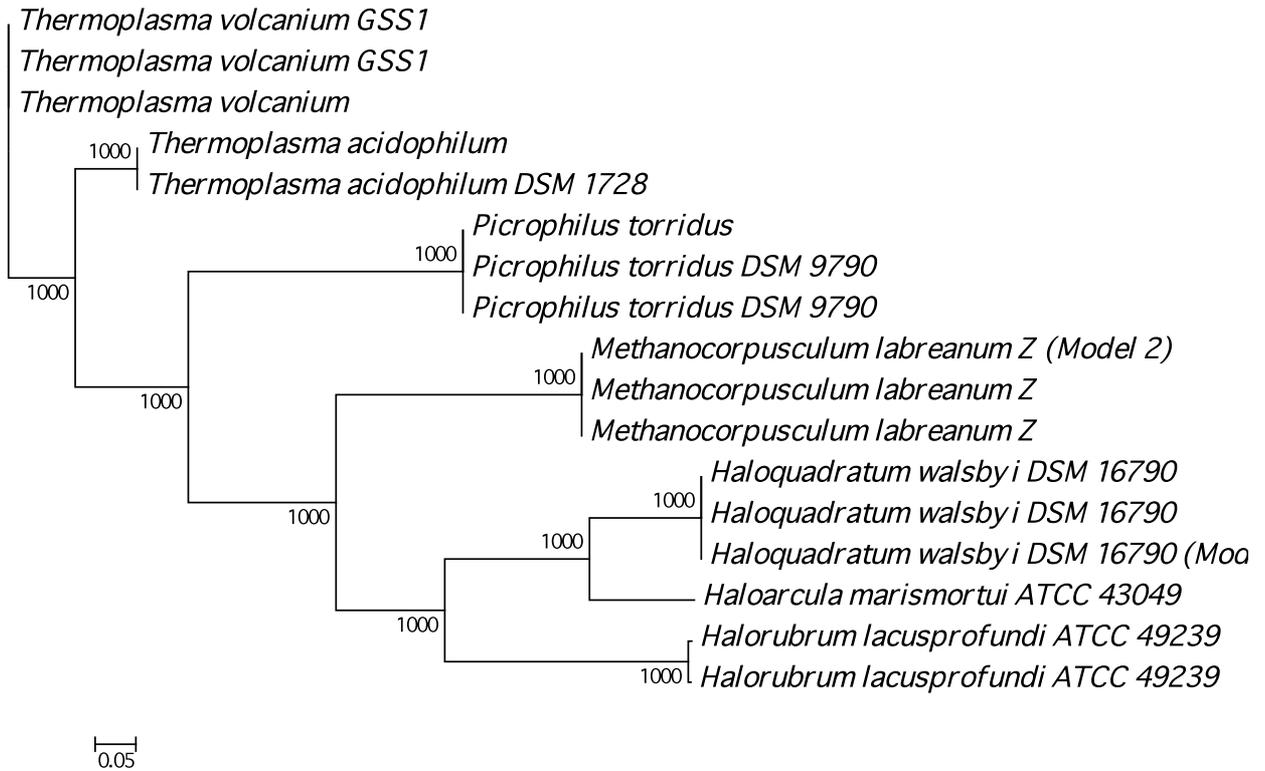
**Figure 1a.** Multiple sequence alignment of variants with related sequences at metal – and substrate-binding regions  
**1b.** Conserved domains of variant 1(Q18JB5)  
**1c.** Conserved domains of variant 2 (A2SS72)

also stable. Thus, it suggested that as both enzyme variants have ability to interact with 5,10-methenyltetrahydrofolate they may consider to be effective catalysts for biotransformation reactions.

## Discussion

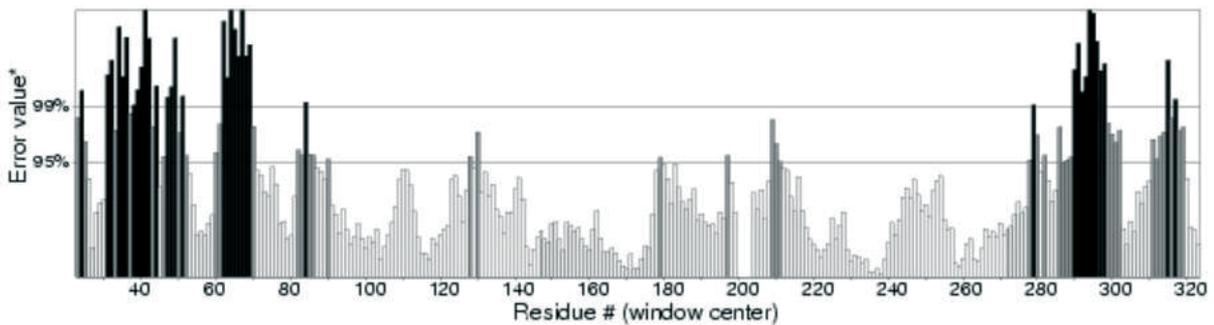
Metals are tightly bound in the active sites of metallo-enzymes to bring the chemical activity, which is often limited to a space extending only a few (20-100) Å around the active site. Metalloenzymes are potentially very

good models for designing chemical catalysts since one could theoretically “cut” out the metal ion together with its chelating amino acids and put the resulting metallo-complex in a flask and expect it to “work”. The first step towards designing a working catalyst is therefore knowledge of the structure of the enzyme and its active site [29]. The functional amino acids or peptides as characteristic molecular moieties and its conservations have led to a significant expansion of the field of artificial enzymes or enzyme mimics. Thus, when enzymes are in action, molecular evolution of enzymes is the main

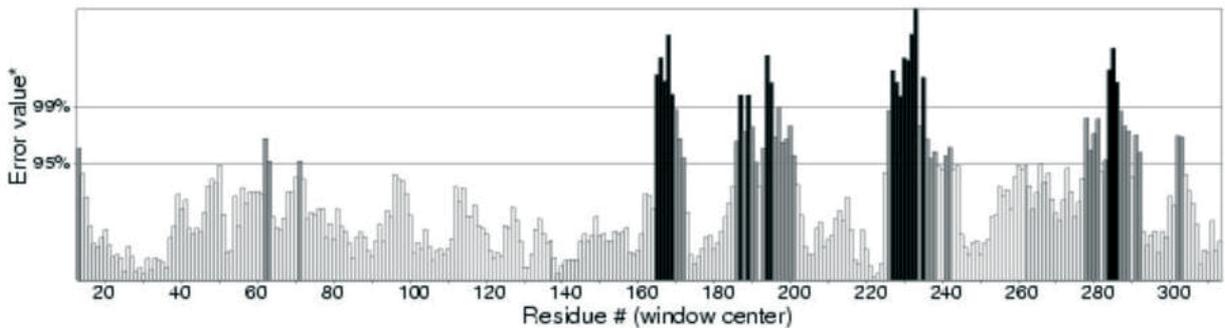


**Figure 2.** NJ Phylogenetic tree of FTHFL sequences obtained from archaea.

Variant 1

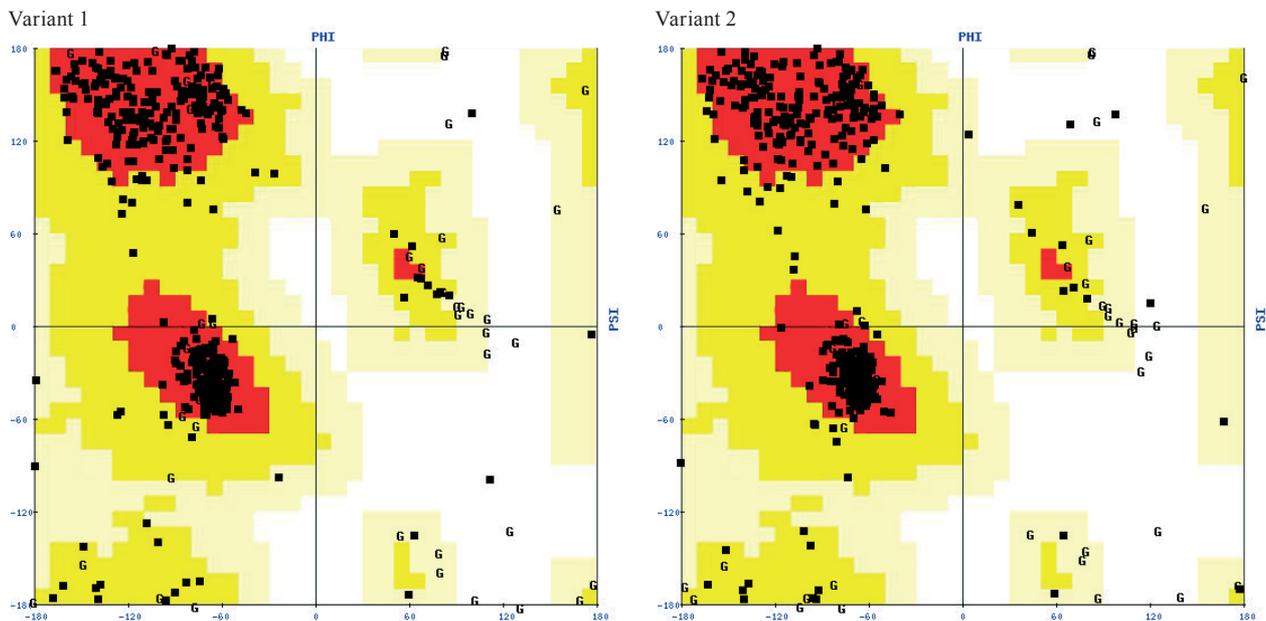


Variant 2

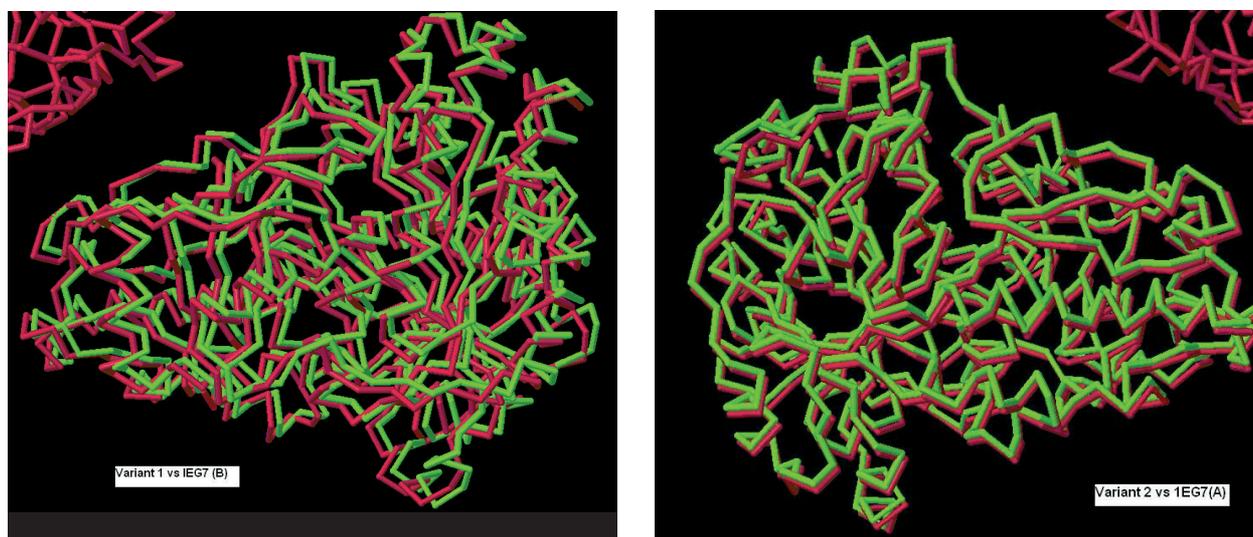


**Figure 3.** The quality of protein models (enzyme variants) validation by ERRAT2 program

\*On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value.



**Figure 4.** The quality of protein models (enzyme variants) validation by Ramachandran Plot



**Figure 5.** Graphical representation of structural superimposition by DaliLite server

determinant in biological systems.

The FTHFL sequence was reported in *Haloquadratum walsbyi* DSM 16790 which is well known as “Walsby’s Square Bacterium” of the family of *Halobacteriaceae* archaeon [30, 31]. It was also reported in *Methanococcus labreanum* Z isolated from surface sediments of Tar Pit Lake at the LaBrea tar pits in Los Angeles, belongs to the order *Methanomicrobiales* of archaeon [32]. Since both organisms were physiologically and ecologically diverged, FTHFL would be expected to be more conserved at sequence and structure levels. Therefore, phylogenetic analysis of this study revealed that FTHFL

sequences of these organisms are evolutionarily related within the species and limited the functional fidelity during evolutionary process.

ERROT program analyzes the statistics of non-bonded interactions between different atom types and plots the value of the error function versus position of a 9-residue sliding window, calculated by a comparison with statistics from highly refined structures. Verify 3D program determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar, etc.) and comparing



T: ALKAGMVNLEAHIQNLQTFGVPVVVAINRFSTDDTEFLAVLKEFCTAQGAEFASEVFAKGGEGGVELAKKVVASCE  
 KPQKFQCLYELNTPKIEKINALATRIYGADGVVYSPAA  
 V2: ALREGFANLEKHIENIGKFGVPAVVAINAFPTDTAEALNLLYELCAKAGAVALS--  
 WAKGGEGGLELARKVLQTLERPSNFHVLNLDLSIKDKIAKIATEIYGADGVNYTAEA  
 T: DAAIKDIDALGRSNLPICTMAKTQYSLDDPNKLRPKNFVINAATVRLCNGAGFIVETGDMITLPLGPAVPAACSIDVNNDGYISG  
 V2: DKAIQRYESLGYGNLPPVMAKTQYFSDDMTKLRPRNFITTVREVRLSAGRLIVPTGAIMTMPGLPKRPAACNIDIDADGVITG

T:  
 LLLHHHLLLLLLLLHHHHHHHHHLLHHHEEELLLLFEELLHHHHHHHLLLLLEEEEEELLLLLLLLLHHHHHHHHHHHHHHHLLLEEEEL  
 LLLHHHHHLLLLLEELLEEEELH  
 V2: LLLLLLLLLLLLLHHHHHHHLLLLHHHEEELLLLFEELLHHHHHHHLLLLLEEEEEELLLLLLLLLHHHHHHHHHHHHHHHLLLEEEEL  
 LLLLHHHHLLLLLEELLEEEELH

T:  
 HHHHHLLLLLLLLHH  
 LHHHHHHHHHLLLEEEELLLLFEEL  
 V2: HHHHLLLLLLLLHH  
 LLLHHHHHHHHHLLLEEEELLLLFEEL

T:  
 LLLLLHHHHHHHHHLLLLLEEEELLLLFEELLLLLLLLLLLLLLLLLHHHHHHHHHHHLLLEEEELLLLLLLLLHHHHHHHHHHHHHHHLLLEEEEL  
 LHHHHHHHLLLLLLLLLH  
 V2: HHHHLLHHHHHHHLLLLLEEEELLLLFEELLLLLLLLLLLLLLLLLHHHHHHHHHHHLLLEEEELLLLLLLLLHHHHHHHHHHHHHHHLLLEEE  
 EEEELHHHHHHLLLLHHHHLLLLLH

T: HHH  
 LLLLLLLLLLLLLHHHHHHHHHHHLLLLLEEEELHHH  
 V2: HHH  
 LLLHH  
 LLLHH

T: HHHHHHHHHHHHLLLLLEEEELLLLLLLLLLLLLLLLLLEEELEEEELLLLLEEEELLLLLLLLLLLLLHHHHHLEELLLLEEL  
 V2: HHHHHHHHHHLLLLLEEEELLLLLLLLLLLLLLLLLLEEEELLEEEELLLLLEEEELLLLLLLLLLLLLHHHHHLEELLLLEEL

Figure 6. Representation of pairwise structural alignment by DaliLite server (T: template, V1: variant 1, V2: variant 2, H: helix, E: strand, L: coil)

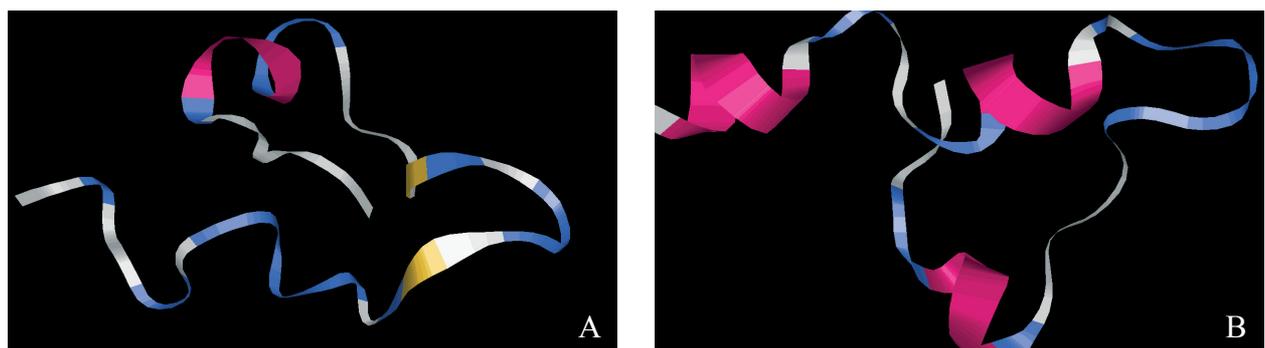


Figure 7. Molecular graphical representation of 3D structure (Rasmol view) of FTHFL (A) variant 1 and (B) variant 2

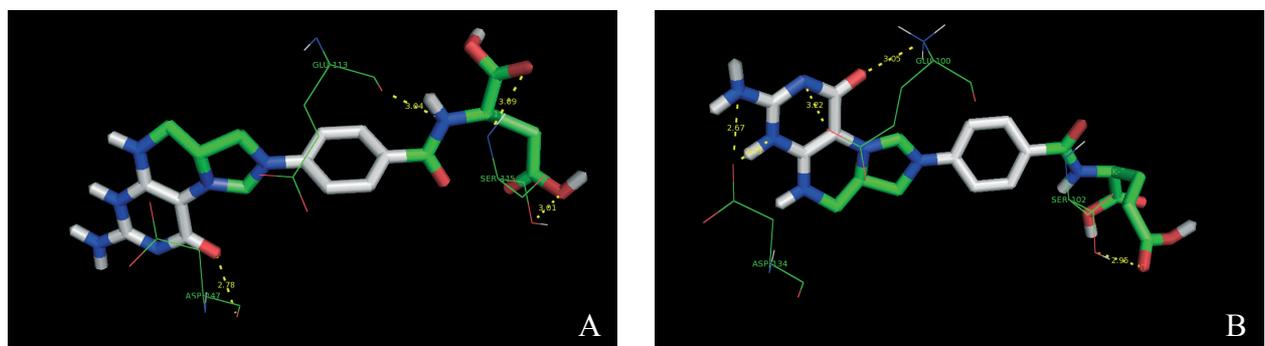


Figure 8. Molecular interaction view of 5, 10-methenyltetrahydrofolate into FTHFL (A) variant 1 and (B) variant 2. After docking, this model was solvated with explicit periodic boundary solvation model. The yellow dot lines denote the hydrogen bonds. All the amino acid residues which involved in molecular interaction are shown in line drawing and colored by residue types in which hydrogen is colored white, carbon green, oxygen red, nitrogen blue, and sulfur orange. Ligands are shown in stick in which carbon is colored tints, hydrogen gray, nitrogen blue, and sulfur orange. All the interaction distances are represented as RMSD and expressed as Angstrom (Å). The binding energy as well as intermolecular forces acting on this docking model is represented in Table 4.

**Table 1.** Homology modeling for predicting 3D structure of FTHFL

Accession	AA	Template	Identity	Position	MPQS	Z-Dope
Q9HI67	536	1eg7	53.00	1-534	1.68	-0.76
NP_394929	536	1eg7	53.00	1-534	1.67	-0.73
Q97CL3	536	1eg7	53.00	1-534	1.65	-0.54
BAB59230	536	1eg7	53.00	1-534	1.65	-0.57
Q6KZM3	535	3do6	49.00	1-534	1.68	-1.02
YP_024022	535	3do6	49.00	4-541	1.68	-1.02
AAT43829	535	3do6	49.00	1-534	1.68	-1.02
Q5V5Y2	553	3do6	53.00	1-552	1.68	-0.79
Q18JB5	580	1eg7	53.00	19-578	1.63	-0.63
A2SS72	560	1eg7	64.00	9-558	1.77	-0.56
YP_001030445	560	1eg7	64.00	9-558	1.77	-0.56
ABN07178	560	1eg7	64.00	9-558	1.77	-0.56
ZP_02016481	574	1eg7	54.00	20-572	1.64	-0.59
EDN49020	574	1eg7	54.00	20-572	1.64	-0.59
NP_110608	536	1eg7	53.00	1-534	1.65	-0.54
YP_134776	643	1eg7	53.00	82-641	1.52	-0.48
AAV45070	643	1eg7	53.00	82-641	1.52	-0.48
CAJ51896	580	1eg7	53.00	19-578	1.63	-0.63
YP_657533	580	1eg7	53.00	19-578	1.63	-0.63

Note: Modeler score and E-value for all the sequences were 1.00 and 0.00 respectively. Each template has matched with A chain only.

**Table 2.** Data mining for searching metal binding and active site similarity regions of FTHFL variants (PSSM-ID: 73210).

Accession	Template	Metal binding site	Active site	Score
Q9HI67	1eg7A	80-110	Thr113, Gly114, Asp115, Phe116	2.40
NP_394929	1eg7A	80-110	Leu71,Gly72,Lys73	1.93
Q97CL3	1eg7A	80-110	Thr113,Gly114,Asp115,Phe116	2.41
BAB59230	1eg7A	80-110	Thr113,Gly114,Asp115,Phe116	2.41
Q6KZM3	3do6A	80-110	Thr112, Gly113, Asp114, Phe115	1.90
YP_024022	3do6A	80-110	Thr112, Gly113, Asp114, Phe115	1.90
AAT43829	3do6A	80-110	Thr112, Gly113, Asp114, Phe115	1.90
Q5V5Y2	3do6A	80-110	Thr118,Gly119,Asp120Leu121	1.90
Q18JB5	1eg7A	110-140	Thr145, Gly146, Asp147, Leu148	2.88
A2SS72	1eg7A	100-130	Thr132,Gly133,Asp134,Phe135	2.90
YP_001030445	1eg7A	100-130	Thr132,Gly133,Asp134,Phe135	2.90
ABN07178	1eg7A	100-130	Thr132,Gly133,Asp134,Phe135	2.90
NP_110608	1eg7A	80-110	Thr113,Gly114,Asp115,Phe116	2.38
YP_134776	1eg7A	175-205	Trp101, Gly102, Leu103	2.30
AAV45070	1eg7A	175-205	Trp101, Gly102, Leu103	2.30
CAJ51896	1eg7A	115-145	Glu87, Gly88, Lys89, Thr90	5.47
YP_657533	1eg7A	115-145	Glu87, Gly88, Lys89, Thr90	5.47

**Table 3.** Molecular dynamic simulation for searching top-five lower energy conformers of FTHFL variants. All of the molecular energies are expressed as kcal/mole.

Accession	Total Energy	Vander Waals energy	Electrostatic energy	Torsion energy	Temperature (K)
Q18JB5	-847.36	-143.49	-875.67	95.03	301.32
	-847.13	-151.54	-873.40	91.85	302.32
	-847.13	-143.84	-875.48	92.22	297.31
	-847.11	-135.84	-841.49	88.17	305.84
	-847.03	-143.51	-880.65	87.66	302.11
A2SS72	-735.03	-147.59	-635.42	78.90	298.05
	-735.03	-141.20	-719.18	93.05	302.95
	-734.68	-138.28	-656.03	81.43	302.17
	-734.66	-137.97	-655.27	81.00	302.02
	-734.65	-140.39	-657.89	76.08	297.52
YP_001030445	-735.03	-141.20	-719.18	93.05	302.95
	-735.03	-147.59	-635.42	78.90	298.05
	-734.68	-138.28	-656.03	81.43	302.17
	-734.66	-137.97	-655.27	81.00	302.02
	-734.65	-140.39	-657.89	76.08	297.52
ABN07178	-735.03	-141.20	-719.18	93.05	302.95
	-735.03	-147.59	-635.42	78.90	298.05
	-734.68	-138.28	-656.03	81.43	302.17
	-734.66	-137.97	-655.27	81.00	302.02
	-734.65	-140.39	-657.89	76.08	297.52

**Table 4.** Molecular docking for predicting binding energy of FTHFL variant and THF derivative complexes

Substrate	RMSD (Å)	Binding energy (kcal/mol)	Inhibition constant (Ki mM)	Inter molecular energy (kcal/mol)	Internal energy (kcal/mol)
Variant 1					
10-Formyltetrahydrofolate	57.60	-3.41	3.18	-4.07	-4.13
10-Formyldihydrofolate	60.28	-3.27	4.02	-6.51	-4.92
10-Formyltetrahydrofolyl L-glutamate	59.03	-5.74	0.62	-10.64	-8.27
5-Methyltetrahydrofolate	63.87	-3.59	2.35	-6.18	-5.39
5,10-Methenyltetrahydrofolate	60.44	-4.04	1.09	-5.92	-5.00
Variant 2					
10-Formyltetrahydrofolate	59.04	-3.87	1.45	-6.30	-3.94
10-Formyldihydrofolate	59.01	-2.78	9.24	-6.36	-4.27
10-Formyltetrahydrofolyl L-glutamate	57.59	-4.46	0.54	-7.60	-5.11
5-Methyltetrahydrofolate	59.33	-5.23	0.15	-7.98	-6.26
5,10-Methenyltetrahydrofolate	60.75	-4.00	1.17	-5.86	-4.03

the results to good structures. The volumes of atoms in macromolecules using Prove algorithm which treats the atoms like hard spheres and calculates a statistical Z-score deviation for the model from highly resolved (2.0 Å or better) and refined (R-factor of 0.2 or better) PDB-deposited structures. All of these algorithms have been used to evaluate the structural quality of homology models of variants and reported as good structures with

high accuracy for further work.

One-carbon metabolism mediated by folate co-enzymes plays an essential role in several major cellular processes. Three folate-dependent enzymes [10-formyltetrahydrofolate synthetase, 5,10-methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9) and 5,10-methylenetetrahydrofolate dehydrogenase (EC 1.5.1.5)] are generally

existed as monofunctional or bifunctional proteins in prokaryotes. A reversible conversion of 10N-formyltetrahydrofolate to formate is also catalyzed by FTHFL and it is part of the tetrahydrofolate pathway of formaldehyde oxidation. [33] A molecular complex of FTHFL with tetrahydrofolate (THF) derivatives revealed that how FTHFL plays an essential role in different biochemical reactions in folate metabolism of *H. walsbyi* DSM 16790 and *M. labreanum* Z.

The main mechanism for increasing activity with monovalent cations is the enhancement of formate-binding. The high electron density of cesium is allowed unambiguous location of its binding site, but it differed from that of the physiological cation binding site. At high formate concentrations (40 mM), FTHFL activation by monovalent cations becomes insignificant [34]. Furthermore, monovalent cations have only weak polarizing abilities and usually have a role in FTHFL regulation in respect to thermal stability [14, 15]. Cation-binding site was buried and far from the putative active site [16, 17]. Studies on thermophilic homoacetogen, *Moorella thermoacetica* showed that monovalent cations such as cesium and potassium ions are required for both optimal activity and stabilization of tetrameric structure of FTHFL at higher temperatures [16]. Similarly, cesium-binding domain in variant 1 and variant 2 was buried at the amino acid positions 110-140 and 100-130, respectively. Cesium- and folate-binding sites in FTHFL are more important than other sites for its catalytic action. Glu98 has reported as a crucial residue in the interaction of monovalent cations with FTHFL [16]. Similarly, the side chains of Glu113, Ser115 and Asp147 were identified in variant 1, while Glu113, Ser115 and Asp147 were side chains in variant 2. This suggests that cesium-binding sites and side chains of both enzyme variants can be contributed for catalytic enhancement in chemical processes of designed enzyme constructs. The contribution of side chains in these enzyme variants were varied in some extent as compared to the side chains contribution of crystallographic structures (IFPM and IFP7) to monovalent cations (Supplementary). The side chains of amino acids may contain a variety of reactive residues and metal ions play a role in electrostatic stabilization of intermediates and transition states occur. It may be attributed to functional divergence and substrate binding affinity of the enzyme variants. THF and its derivatives are the biologically active forms of folic acid (a four-electron oxidized form) that are specialized co-substrates for a variety of enzymes involved in one-carbon metabolism. Catalytic activity and interactions of FTHFL variants with THF derivatives were examined in this study. More number of favorable substrate-binding sites and docking pose orientations were observed in FTHFL variants. It suggested that these variants have to show strong interactions with THF derivatives and thus provide diverse catalytic activity. Both enzyme variants have a potential for hydrogen bonding with 5,10 methenyltetrahydrofolate and this sug-

gests that they can enable to carry out different chemical activities. Besides, variant 1 was better than variant 2 in interacting with 5,10-methenyltetrahydrofolate as it has high binding energy, H-bonding with its active site, and a stable energy conformation. Both variants have three possible catalytic activities, apart from carbon-nitrogen forming activity, when they interacted with 5,10-methenyltetrahydrofolate. The possible catalytic activities could be; (i) as 5,10-methenyltetrahydrofolate 5-hydrolyase (decyclase) (EC 3.5.4.9) to form 10-formyltetrahydrofolate [35], (ii) as tetrahydrofolate aminomethyl transferase (EC 2.1.2.10) to form 5-formyltetrahydrofolate [36], and (iii) as 5-formiminotetrahydrofolate aminolyase (cyclizing) (EC 4.3.1.4) to produce 5-formiminotetrahydrofolate in presence of ammonia [13].

The FTHFL requires monovalent cations to prevent active tetramer from dissociating into inactive monomeric subunits [37]. Monomers associate to form a tight dimer complex through a series. The protein residues forming the cesium ion-binding site belong to two coils; the first coil connects helices 5 and 6 in domain 1, while the other coil links  $\beta$ -strand 5 and  $\alpha$ -helix 4 in domain 2 [16-17]. Enzyme variants of this study have only  $\alpha$ -helices in a stable monomeric structure so that they may form a domain 1 as similar to earlier reports [16-17].

The removal of monovalent cations in the presence of substrates results in an altered tetramer FTHFL, which has an increased  $K_m$  for formate, but still has the same maximal catalytic activity [16, 37-42]. This suggests that the binding of formate was assisted by monovalent cations, but that they are not an essential part of the reaction mechanism. The catalytic inactivation may result from a conformational change which accompanies dissociation. Even though the binding experiments showed that there are four sites for each substrate per mole of tetramer, the binding sites may not be intrinsic to each monomer [38-43]. There are two possible reasons; one or more of the substrate sites may be formed as a result of conformational changes which are caused by the protein-protein interactions responsible for the association of monomers, or one or more of the substrate sites may possibly be located at the site of interaction between subunits and that the substrate site was composed of segments of polypeptide chains from two distinct subunits. Thus, the subunit interactions resulting from the association may cause conformational changes sufficient to produce the new binding sites. Alternatively, the folate site may be composed of segments of polypeptide chains from the two distinct subunits and is created at the site of interaction between subunits.

Structure-based computational design techniques have been used to predict mutations for the construction of catalytically active sites in proteins of known structure [12, 43, 44]. There has been an increasing effort to combine rational design features into Darwinian evolutionary protocols [45-47]. Fischer *et al.* [48] developed TransCent, a computational enzyme design method, suppor-

ting the transfer of active sites from one enzyme to an alternative scaffold. They reported that protein stability, ligand binding, pKa values of active site residues and structural features of the active site are constraints needed to balance in the optimization process. Molecular evolution-directed approach has already been reported for designing of  $\beta$ -methylaspartate mutase variant from the sequences of Haloarchaea [49]. The present approach is also entirely based on evolutionary conservation of sequences as well as structures of FTHFL as it is in the native form so that there is no need on further optimizing process to evaluate these enzyme variants to work.

## Acknowledgements

P. Chellapandi is grateful to the University Grants Commission, New Delhi, India, for financial assistance (UGC Sanction No. 32-559/2006) to carry out the work.

## Conflict of Interest

Authors have no conflict of interest.

## References

- Schiraldi C, Giuliano M, De Rosa M. (2002) Perspectives on biotechnological applications of archaea. *Archaea* 1:75–86.
- Kirk O, Borchert TV, Fuglsang CC. (2002) Industrial enzyme applications. *Curr Opin Biotechnol* 13:345–51.
- Turner NJ. (2003) Directed evolution of enzymes for applied biocatalysis. *Trends Biotechnol* 21:474–8.
- Straathof AJ, Panke S, Schmid A. (2002) The production of fine chemicals by biotransformations. *Curr Opin Biotechnol* 13:548–56.
- Schmid A, Hollmann F, Park JB, Buhler B. (2002) The use of enzymes in the chemical industry in Europe. *Curr Opin Biotechnol* 13:59–366.
- Panke S, Held M, Wubbolts M. (2004) Trends and innovations in industrial biocatalysis for the production of fine chemicals. *Curr Opin Biotechnol* 15:272–8.
- Luetz S, Giver L, Lalonde J. (2008) Engineered enzymes for chemical production. *Biotechnol Bioeng* 101:647–53.
- Walsh ST, Cheng H, Bryson JW, Roder H, DeGrado WF. (1999) Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc Natl Acad Sci USA* 96:5486–91.
- Marshall SA, Mayo SL. (2001) Achieving stability and conformational specificity in designed proteins via binary patterning. *J Mol Bio* 305:619–31.
- Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS. (1998) High-resolution protein design with backbone freedom. *Science* 282:1462–7.
- Benson DE, Wisz MS, Hellinga HW. (2000) Rational design of nascent metalloenzymes. *Proc Natl Acad Sci USA* 97:6292–7.
- Looger LL, Dwyer MA, Smith JJ, Hellinga HW. (2003) Computational design of receptor and sensor proteins with novel functions. *Nature* 423:185–90.
- Rabinowitz JC, Pricer WE Jr. (1962) Formyltetrahydrofolate synthetase. I. Isolation and crystallization of the enzyme. *J Biol Chem* 237:2898–902.
- Hase T, Matsurbara H, Koike H, Katoh S. (1983) Cation binding and thermostability of FTHFL monovalent cation binding sites and thermostability of N10-formyltetrahydrofolate synthetase from *Moorella thermoacetica*. *Biochem Biophys Acta* 744:46–52.
- Dionisi HM, Alvarez CV, Viale AM. (1999) Alkali metal ions protect mitochondrial rhodanese against thermal inactivation. *Arch Biochem Biophys* 361:202–6.
- Radfar R, Shin R, Sheldrick GM, Minor W, Lovell CR, *et al.* (2000) The crystal structure of N(10)-formyltetrahydrofolate synthetase from *Moorella thermoacetica*. *Biochem* 39:3920–6.
- Radfar R, Leaphart A, Brewer JM, Minor W, Odom JD, *et al.* (2000b) Cation binding and thermostability of FTHFS monovalent cation binding sites and thermostability of N10-formyltetrahydrofolate synthetase from *Moorella thermoacetica*. *Biochem* 39:14481–6.
- Hibbert EG, Dalby PA. (2005) Directed evolution strategies for improved enzymatic performance. *Microb Cell Fact* 4:29.
- Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, *et al.* (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucl Acids Res* 37:205–10.
- Marchler-Bauer A, Bryant SH. (2004) CD-Search: protein domain annotations on the fly. *Nucl Acids Res* 32: 327–31.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl Acids Res* 24:4876–82.
- Tamura K, Dudley J, Nei M, Kumar S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–9.
- Efron B, Halloran E, Holmes S. (1996) Bootstrap confidence levels for phylogenetic trees. *Proc Natl Acad Sci USA* 93:13429–34.
- Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, *et al.* (2003) Tools for comparative protein structure modeling and analysis. *Nucl Acids Res* 31:3375–80.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* 25:3389–402.
- Sali A, Blundell TL. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815.
- Laskowski RA, Watson JD, Thornton JM. (2005) ProFunc: a server for predicting protein function from 3D structure. *Nucl Acids Res* 33:W89–W93.
- Willett P. (1995) Genetic algorithms in molecular recognition and design. *Trends in Biotechnol* 13:516–21.
- Schenk S, Weston S, Anders E. (2003) Computational studies on the mode of action of metalloenzymes quantum chemistry connects molecular biology with chemistry. *Berichte des IZWR Band* 2:1–18.
- Bolhuis H, Poele EM, Rodríguez-Valera F. (2004) Isolation and cultivation of Walsby's square archaeon. *Environ Microbiol* 6:1287–91.
- Bolhuis H, Palm P, Wende A, Falb M, Rampp M, *et al.* (2006) The genome of the square archaeon *Haloquadratum walsbyi*: life at the limits of water activity. *BMC Genomics* 7:169.
- Zhao Y, Boone DR, Mah RA, Xun L. (1989) Isolation and characterization of *Methanococcus labreanus* sp. nov. from the LaBrea Tar pits. *Int J Syst Bacteriol* 39:10–3.
- Christopher JM, Markus L, Julia AV, Mary EL. (2003) Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. *J Bacteriol* 185:7169–75.
- de Renobales M, Welch W Jr. (1982) The monovalent cation induced reassociation of formyltetrahydrofolate synthetase monitored by Rayleigh light scattering and enzymic activity. *Bioc*

hem 21:3530–7.

- [35] Tabor H, Wyngarden L. (1959) The enzymatic formation of formiminotetrahydrofolic acid, 5,10-methenyltetrahydrofolic acid, and 10-formyltetrahydrofolic acid in the metabolism of formiminoglutamic acid. *J Biol Chem* 234:1830-46.
- [36] Nesbitt NM, Baleanu-Gogonea C, Cicchillo RM, Goodson K, Iwig DF, *et al.* (2005) Expression, purification, and physical characterization of *Escherichia coli* lipoyl(octanoyl)transferase. *Protein Expr Purif* 39:269-82.
- [37] Harmony JAK, Himes RH, Schowen RL. (1975) Monovalent cation-induced association of formyltetrahydrofolate synthetase subunits. Solvent isotope effect. *Biochem* 14:5379-86.
- [38] Welch WH, Irwin C, Himes RH. (1968) *Biochem Biophys Res Commun* 30:255.
- [39] Curthoys NP, Straus LD, Rabinowitz JC. (1972) Formyltetrahydrofolate synthetase- substrate binding to monomeric subunits. *Biochem* 11:345-9.
- [40] Anguera MC, Suh JR, Ghandour H, Nasrallah IM, Selhub J, *et al.* (2003) Methenyltetrahydrofolate synthetase regulates folate turnover and accumulation. *J Biol* 278:29856-62.
- [41] Chen S, Yakunin AF, Proudfoot M, Kim R, Kim SH. (2005) Structural and functional characterization of a 5,10-methenyltetrahydrofolate synthetase from *Mycoplasma pneumoniae* (GI: 13508087). *Proteins* 61:433-43.
- [42] Field MS, Szebenyi DM, Perry CA, Stover PJ. (2007) Inhibition of 5,10-methenyltetrahydrofolate synthetase. *Arch Biochem Biophys* 458:194-201.
- [43] Hancock AN, Coleman RS, Johnson RT, Sarisky CA, Johann TW. (2008) Investigations of the roles of arginine 115 and lysine 120 in the active site of 5,10-methenyltetrahydrofolate synthetase from *Mycoplasma pneumoniae*. *Protein J* 27:303-8.
- [44] Leaphart AB, Trent Spencer H, Lovell CR. (2002) Site-directed mutagenesis of a potential catalytic and formyl phosphate binding site and substrate inhibition of N10-formyltetrahydrofolate synthetase. *Arch Biochem Biophys* 408:137-43.
- [45] Voigt CA, Gordon DB, Mayo SL. (2000) Trading accuracy for speed: a quantitative comparison of search algorithms in protein sequence design. *J Mol Biol* 299:789–803.
- [46] Reetz MT. (2000) Application of directed evolution in the development of enantioselective enzymes. *Pure Appl Chem* 72:1615–22.
- [47] Reetz MT, Carballeira JD. (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Prot* 2:891-903.
- [48] Fischer A, Enkler N, Neudert G, Bocola M, Sterner R, *et al.* (2009) TransCent: Computational enzyme design by transferring active sites and considering constraints relevant for catalysis. *BMC Bioinformatics* 10:54.
- [49] Chellapandi P, Balachandramohan J. (2011) Molecular evolution-directed approach for designing of  $\beta$ -methylaspartate mutase from the sequences of Haloarchaea. *International J Chemical Mod* 3(In Press). Abstract available at: [https://www.novapublishers.com/catalog/product\\_info.php?products\\_id=24749&osCsid=5e9a553a1b1129fed10cc0ad3598a587](https://www.novapublishers.com/catalog/product_info.php?products_id=24749&osCsid=5e9a553a1b1129fed10cc0ad3598a587); last accessed: 12.05.2011