J Mol Model (2011) 17:621–631
DOI 10.1007/s00894-010-0764-y

ORIGINAL PAPER

In Silico sequence analysis and molecular modeling of the three-dimensional structure of DAHP synthase from *Pseudomonas fragi*

Satya Tapas · Girijesh Kumar Patel · Sonali Dhindwal · Shailly Tomar

Received: 20 February 2010 / Accepted: 18 May 2010 / Published online: 3 June 2010

Abstract The shikimate pathway is involved in production of aromatic amino acids in microorganisms and plants. The enzymes of this biosynthetic pathway are a potential target for the design of antimicrobial compounds and herbicides. 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase (DAHPS) catalyzes the first step of the pathway. The gene encoding DAHPS was cloned and sequenced from *Pseudomonas fragi,* the bacterium responsible for spoilage of milk, dairy products and meat. Amino acid sequence deduced from the nucleotide sequence revealed that *P. fragi* DAHPS (*Pf*-DAHPS) consists of 448 amino acids with calculated molecular weight of ~50 kDa and isoelectric point of 5.81. Primary sequence analysis of *Pf*-DAHPS shows that it has more than 84% identity with DAHPS of other *Pseudomonas* species, 46% identity with *Mycobacterium tuberculosis* DAHPS (*Mt*-DAHPS), the type II DAHPS and less than 11% sequence identity with the type I DAHPS. The three-dimensional structure of *Pf*-DAHPS was predicted by homology modeling based on the crystal structure of *Mt*-DAHPS. *Pf*-DAHPS model contains a $(\beta/\alpha)_8$ TIM barrel structure. Sequence alignment, phylogenetic analysis and 3D structure model classifies *Pf*-DAHPS as a type II DAHPS. Sequence analysis revealed the presence of DAHPS signature motif DxxHxN in *Pf*-DAHPS. Highly conserved sequence motif RxxxxxxKPRT(S/T) and xGxR present in type II DAHPS were also identified in *Pf*-DAHPS sequence. High sequence homology of DAHPS within *Pseudomonas* species points to the option of designing a broad spectrum drug for the genus. *Pf*-DAHPS 3D model provides molecular insights that may be beneficial in rationale inhibitor design for developing effective food preservative against *P. fragi*.

Keywords 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase · Homology modeling · *Pseudomonas fragi*

Introduction

Pseudomonas fragi, a psychrophilic species of the genus is the major spoilage microorganism in milk, cottage cheese and dairy products that produces volatile aroma compounds with fruity fragrances [1]. Some strains of *P. fragi* are responsible for aerobic spoilage of meat [2]. The spoilage of food cannot be prevented even if milk, milk products and meat are stored in refrigerators because the bacterium has the capability of growing at psychrophilic temperatures. Water used for cleaning utensils, containers and equipments in food processing industries is a common cause for food contamination. As a result, these microorganisms are able to get into pasteurized milk and milk products and cause spoilage in few days even at low refrigeration temperatures [3]. Various food additives have been used for studying the effect of these on *P. fragi* growth at different temperature and pH [4]. No such food additives were found to be an effective inhibitor of *P. fragi* at different pH ranges and temperatures. Thus, there is a necessity to find novel antimicrobial preservatives that are safe for human consumption but kill or prevent the growth of *P. fragi*.

The essential biosynthetic pathways present only in microbes and absent in humans are potential antimicrobial targets. One such pathway is the shikimate pathway, which is found in all plants and microorganisms, and is absent in humans [5]. The pathway is responsible for synthesis of

Satya Tapas and Girijesh Kumar Patel have contributed equally.

S. Tapas · G. Kumar Patel · S. Dhindwal · S. Tomar (✉)
Department of Biotechnology, Indian Institute of Technology, Roorkee 247667 Uttarakhand, India
e-mail: shailfbt@iitr.ernet.in

chorismate which is a precursor for aromatic amino acids and secondary metabolites [6, 7]. This pathway is being exploited as a rational target for the development of antimicrobial drugs against many pathogens [8–10]. 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (DAHPS; EC 2.5.1.54) enzyme catalyzes the first committed step of the pathway. The reaction is the condensation of phosphoenol-pyruvate (PEP) and D-erythrose 4-phosphate (E4P) to produce 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) and inorganic phosphate. DAHPS enzymes have been classified into two distinct homology families, type I and type II [11, 12]. Less than 12% sequence identity is observed between the two families. Type I family consists of microbial DAHPS including *Escherichia coli*, *Saccharomyces cerevisiae, Thermotoga maritima*, *Pyrococcus furiosus etc.* The type II family DAHPS were first identified in plants and later in some microbes including *M. tuberculosis, Neurospora crassa* [13]. In microbes, DAHPS is regulated at the protein level through allosteric feedback inhibition by the endproducts of this pathway (*i.e*., tyrosine, tryptophan and phenylalanine), whereas in plants the feedback inhibition has not been reported. Instead, DAHPS enzymes in plants are regulated by redox regulation [14–16].

The present study describes cloning, sequence analysis and homology modeling of the deduced amino acid sequence of DAHPS enzyme from *P. fragi*. This is the first report on molecular cloning and structure prediction of an enzyme of the shikimate pathway from the genus *Pseudomonas*. DAHPS is not present in humans, thus it represents an important potential antimicrobial target and the predicted 3D structure of *Pseudomonas* DAHPS may provide a basis for rationale drug designing. Homology between *Pf*-DAHPS and DAHPS present in other species of *Pseudomonas* was also predicted, which indicates that the same inhibitor may be used to inhibit the growth of *P. fragi* and also be used as a drug for the various pathogenic *Pseudomonas* species.

Materials and methods

Bacterial strains and plasmid

Pseudomonas fragi NCIM 2391 was procured from National Collection of Industrial Microorganisms, National Chemical Laboratory (Pune, India). Luria-Bertani (LB) medium was used for the growth of *P. fragi* and *E. coli* strains. *P. fragi* was grown at 30 °C. pGEM-T vector system was obtained from Promega Corporation, Madison, USA and was used according to the manufacturer's instructions. *E. coli* DH5α competent cells were obtained from Invitrogen, Carlsbad, CA, USA.

Genomic DNA isolation

The lyophilized *P. fragi* cells were revived using LB broth. A loopful of the suspension was streaked onto an LB agar plate, which was incubated overnight at 30 °C. A single colony was used to inoculate 5 ml LB broth followed by its incubation overnight at 30 °C. Genomic DNA was isolated from 5 ml overnight culture of *P. fragi* using standard protocol [17].

PCR amplification of *Pf*-DAHPS

Oligonucleotide primers used for amplification of *Pf*-DAHPS were designed based on the putative DAHPS sequence from *P. aeruginosa* strain UCBPP-PA14 deposited in Pseudomonas genome database (NCBI accession no. YP_790330.1) [18]. The sequence of the forward primer was 5′-TGAGCCAGTCCTGGAGCCCCGAGAGCTGG-3′ and the reverse primer was 5′-GCGGCGTACCTGCTT CAGGGTTTCGGCGATC-3′. The reaction mixture used for gene amplification contained 2 µl of 10X *Taq* Buffer supplied with the enzyme, 200 µM of each dNTPs, 5 pmol of each primer, 150 ng of DNA template, and 2 units of *Taq* DNA polymerase (Bangalore *Genei*, Bangalore, India) and water to a final volume of 20 µl. After optimizing the conditions for polymerase chain reaction (PCR), the gene was successfully amplified using the following PCR conditions: 94 °C for 4 min followed by 30 cycles of denaturation at 94 °C for 1 min, annealing at 59 °C for 1 min, extension at 72 °C for 2 min and the final extension was carried out at 72 °C for 10 min on a PTC-100 Thermocycler (M.J. Research, Watertown, MA). The PCR product was analyzed by electrophoresis on a 1 % agarose gel and showed a DNA band of the expected size which was purified using a gel extraction kit (Qiagen, Inc. Valencia, CA).

Cloning and DNA sequencing

The PCR product was subcloned into plasmid DNA using the pGEM-T vector system (Promega, UK). 2 µl of purified PCR product was mixed with 0.5 µl linearized pGEM-T cloning vector in presence of 0.5 µl T4 DNA ligase (Bangalore *Genei*, Bangalore, India) and incubated overnight at 15 °C. Then the ligation mixture was directly used for the transformation of $CaCl_2$ competent DH5α cells by heat shock method [19]. Transformed cells were screened by ampicillin resistant, blue-white selection with X-gal and IPTG. Individual colonies were picked, grown overnight in 5 ml LB broth at 37 °C and plasmids were isolated using commercial mini-prep kit (Qiagen, Inc. Valencia, CA). Restriction digestion screening of the isolated plasmids was done to select the construct contain-

ing the correct size insert and selected constructs were sequenced. Sequencing was performed in both the directions using vector specific SP6 and T7 promoter primers by dye termination method using the ABI Prism automated DNA sequencer.

Sequence and phylogenetic analysis

Sequence identity was verified by doing homology searches using the basic local alignment search tool (BLAST) algorithm [20]. Sequence analysis tools of the ExPASy Server were used for processing nucleotide sequence of *Pf*-DAHPS to deduce the amino acid sequence [21]. Primary structure analysis was done using the ProtParam (http://www.expasy.ch/tools/protparam.html). The ClustalW multiple sequence alignment program was used to align the DAHPS sequences [22]. The ESPript server was used for generating secondary structure elements and to produce a representation of the sequence alignment [23]. The phylogenetic tree was inferred using Genebee's Multiline program [24].

Comparative molecular modeling

Homology modeling for DAHP synthase of *Pseudomonas fragi* was performed in the following sequential steps: template selection from Protein Data Bank (PDB), sequence-template alignment, model building, model refinement and validation [25]. Template search for *Pf*-DAHPS was done using NCBI BLAST search tool against PDB database. Blastp program was run with BLOSUM62 as a scoring matrix, word size of 3, gap penalty of 11 and gap extension penalty of 1. Crystal structure of *Mt*-DAHPS in complex with PEP, sulfate and Manganese ion (PDB ID: 2B7O) having 46% sequence identity with *Pf*-DAHPS was obtained as the best hit. *Mt*-DAHPS crystal structure was used as template to generate a comparative 3D model of *Pf*-DAHPS by MODELLER 9v6 [26].

ClustalW program was used for multiple sequence alignment of query sequence with template sequence [22]. Some manual corrections were done in the alignment file for missing residues in the template sequence. This was then used to build alignment file in PIR/PAR format as an input for MODELLER. Based on sequence alignment analysis, it was assumed that ligand binding modes of *Pf*-DAHPS are similar to *Mt*-DAHPS. Therefore PEP, Mn^{2+} ion and sulfate ion from the template structure were incorporated into the modeled structure of *Pf*-DAHPS. Using MODELLER, several preliminary models were generated which were ranked based on their DOPE scores. Five sets of models having lowest DOPE scores were selected and stereo-chemical quality of each was assessed by PROCHECK [27]. The model with the least number of residues in the disallowed region was further refined for relieving steric clashes and improper contacts. Energy minimization of the selected model was performed using Swiss-Pdb Viewer 4.01 (http://www.expasy.org/spdbv/). SPDBV implements GROMOS96 force field to compute energy and to execute energy minimization. PROCHECK was again used to evaluate the stereo-chemical quality of the model. Loop refinement tool of MODELLER was used in an iterative fashion to refine the loop conformation of the model. Structural validation after each loop refinement step was done using ERRAT plot which gives a measure of the structural error at each residue in the protein. This process was repeated iteratively until most of the amino acid residues were below 95% cut-off value in ERRAT plot [28]. The refined model was further validated by VERIFY-3D of SAVES server (http://nihserver.mbi.ucla.edu/SAVES/). ProSA 2003 was used to evaluate the generated 3D structure model of protein for potential errors [29].

Molecular dynamics simulations

Model assessment was done by molecular dynamics (MD) and simulation studies to determine the stability of the predicted 3D model of *Pf*-DAHPS. GROMACS simulation suite version v. 4.0.7 was used for molecular dynamics work [30]. PEP, sulfate and metal ion (Mn^{2+}) ligands were taken into consideration and included in the system. After solvating, the system was made electroneutral by adding the positive ions as counter ions and was energy minimized to remove high-energy interatomic contacts. Energy minimization was done using 2000 steps of the steepest descent method. Molecular dynamics simulations were performed in the isothermal isobaric ensemble (NPT). A constant pressure of 1 bar was applied independently in all the directions with Berendsen temperature coupling of 0.5 ps. Protein, solvent and ligands were coupled separately to the thermal bath at 300 K using a coupling constant of 0.1 ps. Finally, simulation was done which consisted of two phases: a short 100 ps canonical ensemble MD simulation allowing randomization of water molecules surrounding the protein molecule and a 1 ns isobaric-isothermal ensemble simulation. MD simulations were performed with GROMACS program installed in Red Hat Enterprise Linux 5 operation system (Red Hat Inc. Raleigh, NC) on a Dell Precision T5400 workstation. The protein stability during MD simulation time was assessed by calculating the root mean square deviation (RMSD) between the structures generated from the simulations and the starting structure. The generated model was visualized, inspected and analyzed using COOT [31] and PyMOL [32].

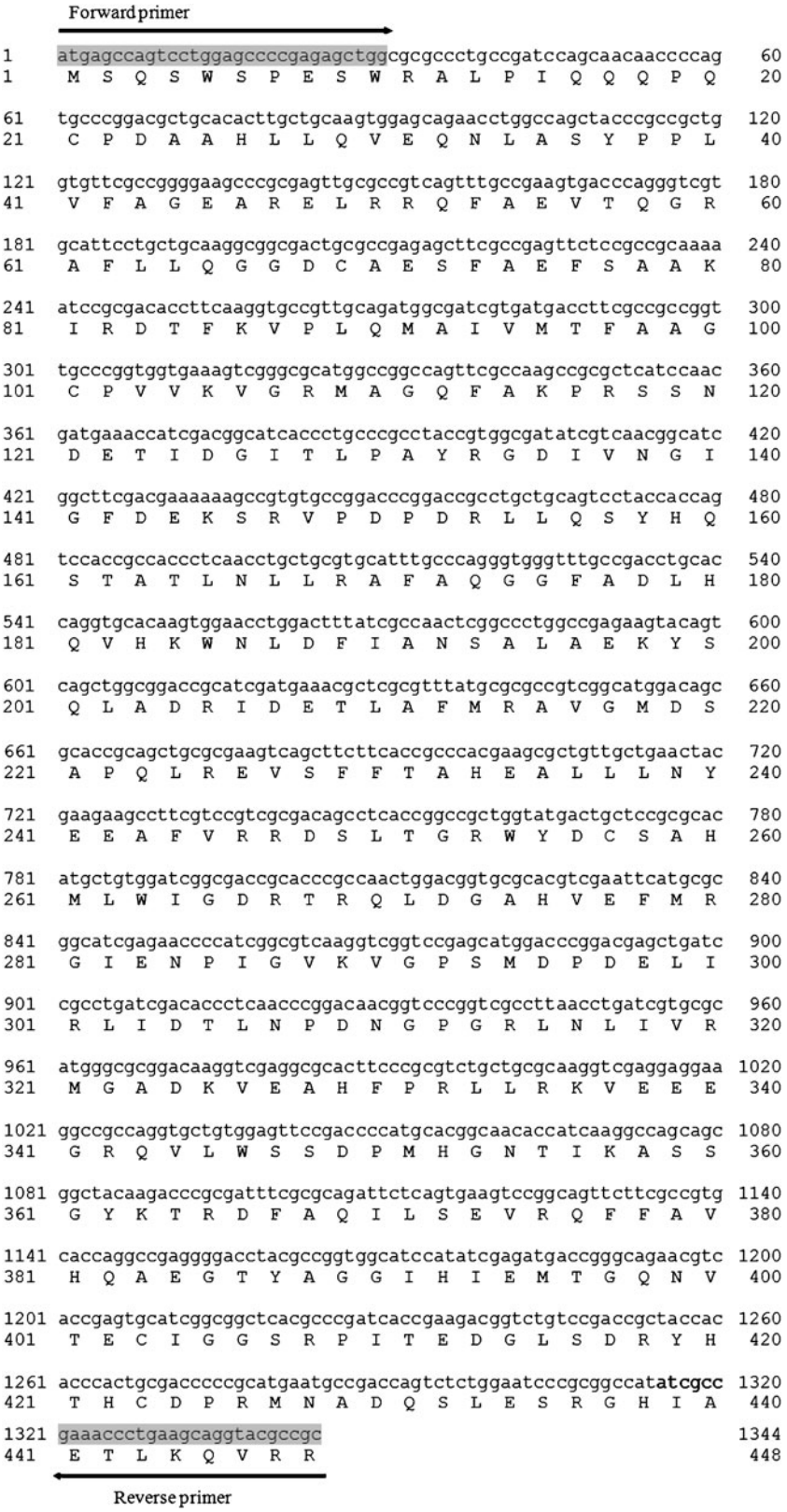

Fig. 1 Nucleotide sequence of *Pseudomonas fragi* DAHP Synthase. The deduced amino acid sequence is given in the one-letter code below the corresponding nucleotide sequence. Forward and reverse primers are highlighted in gray. *Pf*-DAHPS nucleotide sequence has been deposited in GenBank under accession number GQ904237

Nucleotide sequence and protein structure accession code

The sequence of *Pf*-DAHPS was deposited in GenBank database under accession number GQ904237 and homology model was submitted to PMDB database (http://mi.caspur.it/PMDB/; PMDB identifier no. PM0076256).

Results and discussion

Cloning and primary structure analysis

The gene search of Pseudomonas Genome database sequences for DAHPS reveals a single copy of gene in

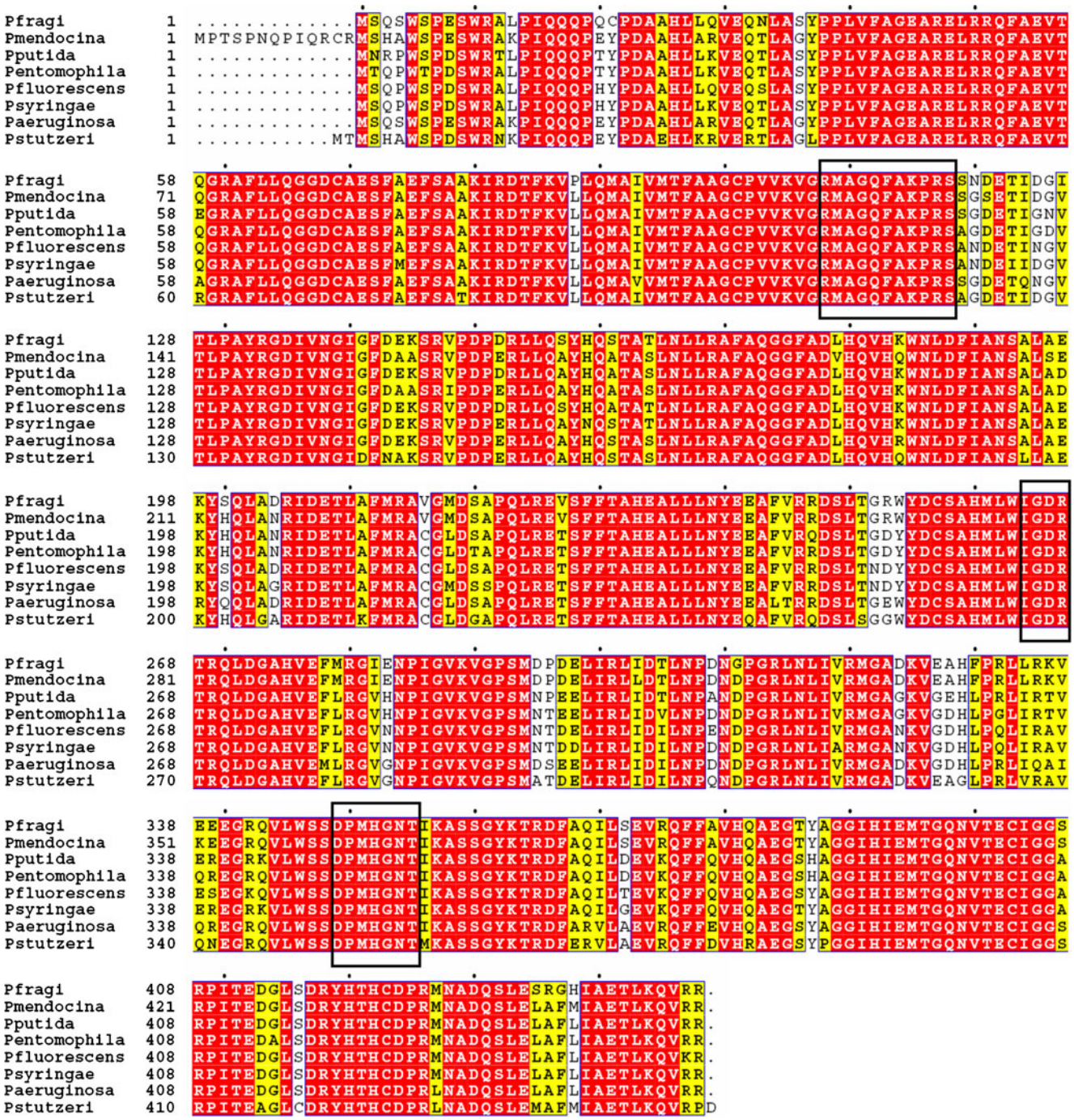

Fig. 2 Multiple sequence alignment of deduced amino acid sequence of *P. fragi* DAHP synthase with DAHP synthase of *P. mendocina* (YP_001187632.1), *P. putida* (YP_001748349.1), *P. entomophila* (YP_607242.1), *P. fluorescens* (YP_261346.1), *P. syringae* (YP_275779.1), *P. aeruginosa* (YP_790330.1) and *P. stutzeri* (YP_001173004.1). Identical and similar amino acids are highlighted in different background colors. Conserved RxxxxxxKPRS, xGxR and DxxHxN motifs are highlighted in boxes. Figure prepared using ESPript [23]

Pseudomonas species. To clone DAHPS gene from *Pseudomonas fragi*, oligonucleotide primers were designed based on the gene sequence of DAHPS from *P. aeruginosa* (*Pa*-DAHPS). The sense primer encodes for N-terminal amino acid sequence MSQSWSPESW of *Pa*-DAHPS exhibiting more than 80 % identity with DAHPS of other *Pseudomonas* species. The antisense primer encodes for C-terminal amino acid sequence IAETLKQVRR of *Pa*-DAHPS, which has more than 90 % homology with the DAHPS of other *Pseudomonas* species. PCR was carried out and a ~1344 bp fragment was amplified from *P. fragi* chromosomal DNA. The PCR product was purified by gel extraction method, cloned into pGEM-T vector and sequenced. *Pf*-DAHPS sequence has been deposited in NCBI GeneBank. BLASTP search showed that *Pf*-DAHPS shared over 84 % sequence identity with other DAHPS genes from the genus *Pseudomonas*. In particular it had highest sequence identity with *P. mendocina* DAHPS (93%).

Nucleotide sequence analysis of cloned *Pf*-DAHPS encodes for 448 amino acids with a predicted molecular mass of~50 kDa (Fig. 1). The predicted isoelectric point (pI) was determined to be 5.8, comparable to *Pa*-DAHPS (YP_790330.1, pI=5.76) and *Mt*-DAHPS (GI:82408027, pI = 5.47). Theoretical hydropathicity of *Pf*-DAHPS is −0.380. When the deduced primary structure of *Pf*-DAHPS was compared with DAHPS of other *Pseudomonas* species and with those of several type I and type II DAHPS, *Pf*-DAHPS had highest identity with that from genus *Pseudomonas*, such as *P. aeruginosa* (88 %) *P. fluorescens* (90 %) and *P. mendocina* (93 %). 337 amino acids out of 448 deduced amino acid residues of *Pf*-DAHPS are identical among DAHPS of *Pseudomonas* species (Fig. 2). *Pf*-DAHPS showed 46 % sequence identity with *Mt*-DAHPS, which belongs to type II and over 50 % sequence identity with other type II DAHPS [*Agrobacterium tumefaciens* (63 %) and *Helicobacter cinaedi* (59 %)]. While DAHPS of type I family showed very low sequence identity with *Pf*-DAHPS. High sequence homology with type II DAHPS and very low sequence identity with type I DAHPS indicates that *Pf*-DAHPS belongs to type II DAHPS family. In general, type II DAHPS (~ 50 kDa) are bigger in size than DAHPS from type I family (~ 40 kDa) and *Pf*-DAHPS is predicted to have~50 kDa molecular mass [11].

Further sequence analysis of *Pf*-DAHPS revealed the presence of $_{108}$RMAGQFAKPRS$_{118}$ sequence, which contains the RxxxxxxKPRT(S/T) phosphate binding motif

Fig. 3 Phylogenetic tree analysis of DAHP synthases. The type II DAHP synthases are *Psuedomonas fragi*, *P. fluorescens* (YP_261346.1), *P. syringae* (YP_275779.1), *P. putida* (YP_001748349.1), *P. entomophila* (YP_607242.1), *P. aeruginosa* (YP_790330.1), *P. mendocina* (YP_001187632.1), *P. stutzeri* (YP_001173004.1), *Arcobacter butzleri* (YP_001489169.1), *Helicobacter canadensis* (ZP_04870584.1), *Pseudovibrio* sp. (ZP_05086764.1), *Bradyrhizobium* sp. (YP_001239492.1), *Parvibaculum lavamentivorans* (YP_001412374.1), *Magnetospirillum magneticum* (YP_420582.1), *Rhodospirillum centenum* (YP_002297191.1), *Arabidopsis thaliana* (NP_195708.1) and *Mycobacterium tuberculosis* (2B7O_A). The type I DAHP synthases are *Thermotoga maritime* (1VR6_A), *Pyrococcus furiosus* (1ZCO_A), *Shigella dysenteriae* (ZP_03067276.1), *Escherichia coli* (1QR7_A), *Citrobacter koseri* (YP_001453944.1), *Salmonella enterica* (YP_001571195.1), *Enterobacter cancerogenus* (ZP_05969196.1), *Vibrio parahaemolyticus* (NP_800695.1) and *Saccharomyces cerevisiae* (1OFP_A)

highly conserved in type II DAHPS signifying that *Pf*-DAHPS belong to type II family (Fig. 2) [33]. The same motif $_{126}$RIAGQYAKPRS$_{136}$ is also present in *Mt*-DAHPS and crystal structure studies has revealed that R$_{126}$ of the motif interacts with phosphate group of PEP and S$_{136}$ interacts with the sulfate anion that occupies the phosphate group of E4P. Another significant characteristic of type II DAHPS is the presence of conserved xGxR sequence motif instead of (I/L)GAR found in DAHPS of type I, which is involved in interactions with PEP [33, 34]. The same $_{264}$IGDR$_{267}$ motif is completely conserved in DAHPS of the *Pseudomonas* genus (Fig. 2). However, the two types of DAHPS contain a well conserved DxxHxN motif ($_{349}$DPMHGN$_{354}$ in *Pf*-DAHPS) which is involved in co-ordination with both PEP and metal ions (Fig. 2) [35]. The phylogenetic tree analysis was performed for *Pf*-DAHPS, DAHPS from other *Pseudomonas* species, type II and type I DAHPS from different microbial and plant sources (Fig. 3). Due to high sequence identity between all the *Pseudomonas* DAHPS, they follow a single branch and cluster together at the top of the tree. The nearest homologues to *Pseudomonas* DAHPS were found to be

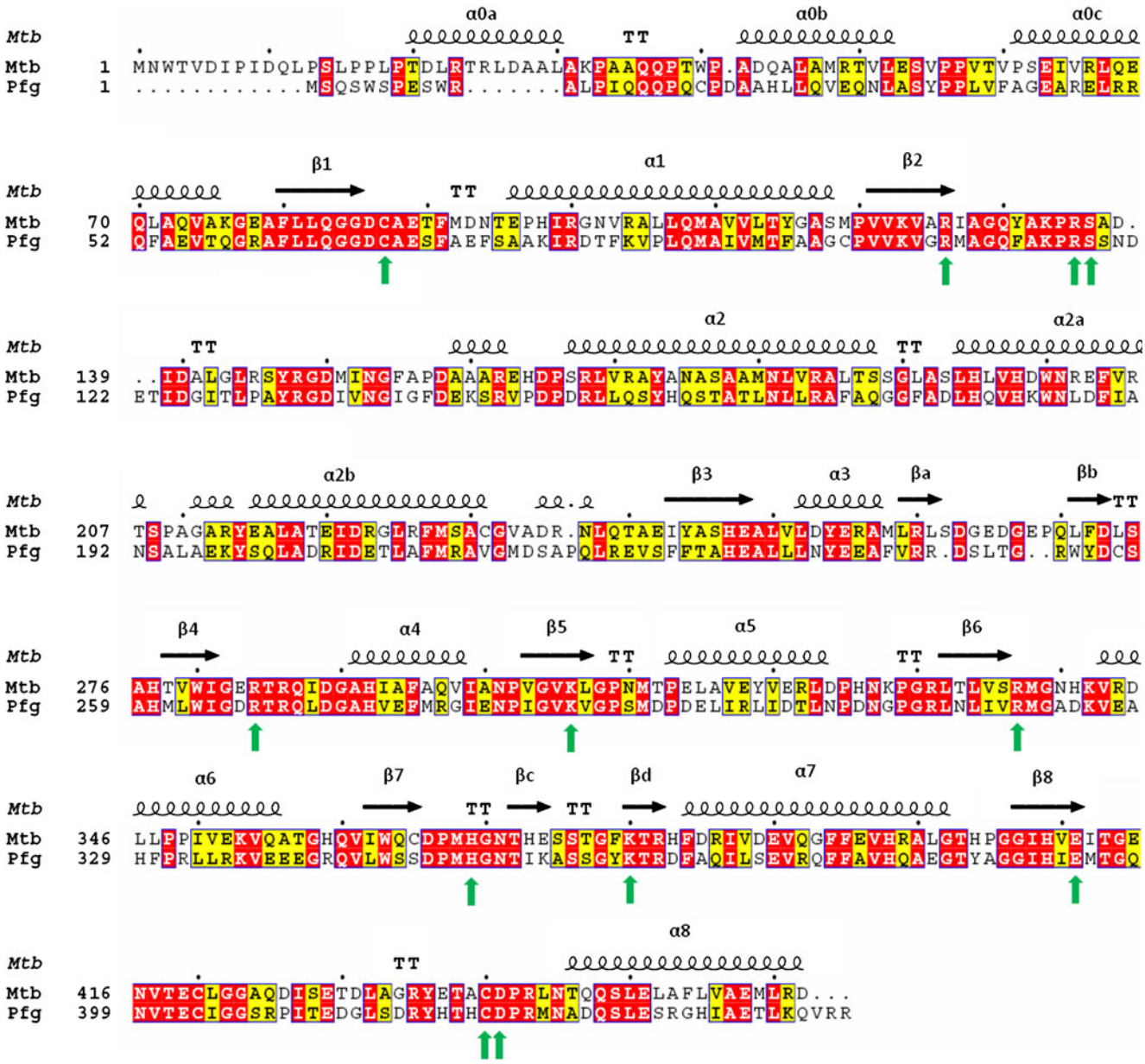

Fig. 4 Amino acid sequence alignment of *P. fragi* DAHP synthase and *M. tuberculosis* DAHP synthase. Identical and similar amino acids are highlighted in different background colors. The secondary structural elements predicted in the sequence of *P. fragi* are shown above the aligned sequences. The conserved amino acid residues at active site have been indicated by arrows at 69(C), 108(R), 117(R), 118(S), 267(R), 289(K), 320(R), 352(H), 363(K), 394(E), 423(C), 424 (D). Figure prepared using ESPript [23]

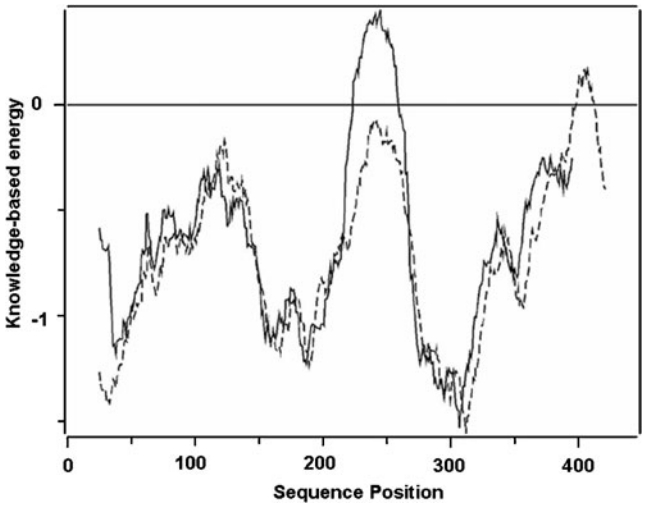

Fig. 5 ProSA energy profile for *P. fragi* DAHP synthase (dotted line) and *M. tuberculosis* DAHP synthase (PDB ID: 2B7O_A) (smooth line) [29]

type II DAHPS from *Arcobacter butzleri* and *Helicobacter Canadensis*. Whereas, type I DAHPS have diversified into a distinct group and form a separate cluster.

Three dimensional structure analysis

The sequence homology between *Pf*-DAHPS and the template was 46 % identity and 63 % similarity. In sequence alignment, the predicted secondary structural elements of *Pf*-DAHPS are roughly identical to the secondary structures present in the crystal structure of *Mt*-DAHPS (Fig. 4). *Pf*-DAHPS ligand-supported homology model was constructed based on the crystallographic 3D structure of *Mt*-DAHPS (PDB ID: 2B7O). To date, *Mt*-DAHPS is the only type II DAHPS that has been structurally characterized. The generated model was subjected to refinement, loop modeling and energy minimization. PROCHECK, Verify-3D and ERRAT plot were used for determining the stereo-chemical parameters of the energy minimized model of *Pf*-DAHPS. Ramachandran plot of the 3D model generated by PROCHECK shows 90.5% residues are present in the core region, 9.0% in allowed region, 0.3% in generously allowed region and 0.3% in disallowed region which includes only one residue. This residue in disallowed region could be ignored as it is not present near the active site nor is involved in ligand binding. Verify_3D shows that 98.43% of the residues have an averaged 3D–1D score greater than 0.2 and ERRAT plot gives an overall quality factor of 94.966 to the modeled structure. ProSA 2003 analysis showed that protein folding energy of our modeled structure is in good agreement with that of the template (Fig. 5).

Molecular dynamics simulations were carried out using the predicted 3D structure of *Pf*-DAHPS protein to determine the stability of the model in equilibration with solvent molecules, *i.e.*, in the physiological state. The obtained MD trajectories during the simulation run of 1 ns were monitored and found to be stable. Figure 6 shows the RMSD of *Pf*-DAHPS as time dependent functions of MD simulation. The graph revealed that RMSD values rise to 0.2 nm in the first 200 ps and then protein remained in the plateau state till the end of equilibration. An overall RMSD of 0.23 nm was obtained which indicates that the 3D modeled structure of *Pf*-DAHPS is good and has a stable configuration. The above results confirm that the 3D model constructed for DAHP synthase of *P. fragi* by comparative modeling is reliable for detail structural analysis.

The RMSD of Cα trace between *Mt*-DAHPS and *Pf*-DAHPS structures was 0.295 Å. The *Pf*-DAHPS model consists of a $(\beta/\alpha)_8$ TIM barrel at the core region of the molecule (Fig. 7). TIM barrel has eight parallel β-strands (residues 62–67, 103–108, 230–234, 261–264, 286–290, 315–319, 345–349, 390–394) that are each followed by α-helices. The active site located at the C-terminal end of the TIM barrel is defined by the residues that are mainly contributed by the loops that follow the β-strands. In addition to the core $(\beta/\alpha)_8$ TIM barrel, predicted *Pf*-DAHPS structure possesses three helices α0a, α0b and α0c (residues 9–12, 23–36 and 43–58) at the N-terminus of the barrel and two helices α2a and α2b (residues 179–192 and 196–216) in the α2-β3 loop region. Structural comparison of *Pf*-DAHPS with *Mt*-DAHPS revealed that the conserved active site residues that were found in the same order and relative spacing in the primary structure of both the enzymes occupy the same spatial position in the three-dimensional structure (Fig. 8). *Pf*-DAHPS residues which form the PEP binding site are Lys 289, Arg 320, Arg 267, Asp 266, Arg 108 which correspond to Lys 306, Arg

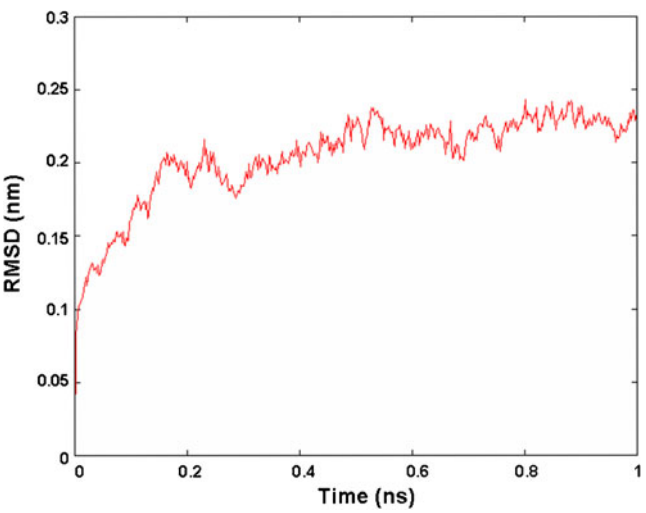

Fig. 6 The root mean square deviation (RMSD) of MD simulation *P. fragi* DAHP synthase relative to the structure of pre-MD simulation

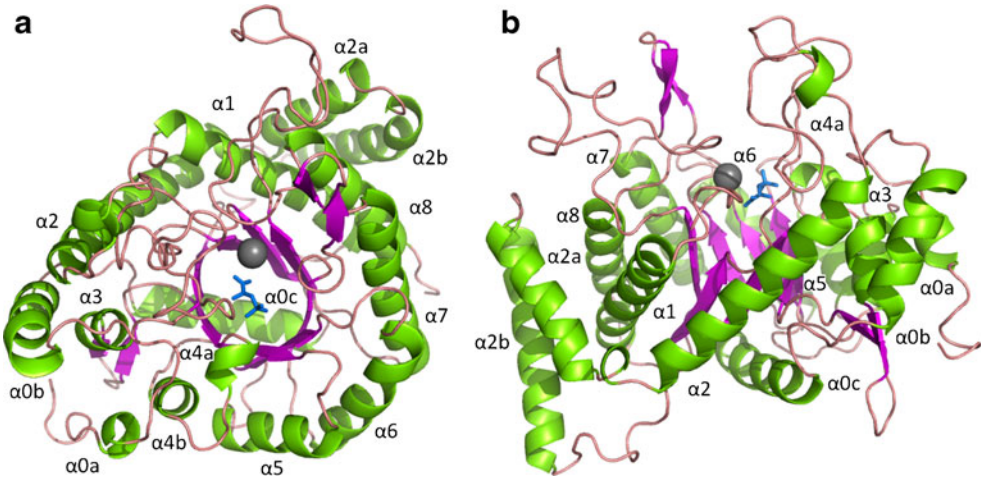

Fig. 7 (**a**) Top view of molecular model of *P. fragi* DAHP synthase developed by Modeler 9v6. (**b**) Side on view of the *P. fragi* DAHP synthase. Helices are shown by cylindrical spiral ribbons and β-strands by arrows. Substrate PEP is shown in the centre of the barrel and manganese ion in a sphere. Figure prepared using PyMOL [32]. 3D model of *Pf*-DAHPS has been submitted to PMDB database with identifier no. PM0076256

337, Arg 284, Glu 283, Arg 126 positions of *Mt*-DAHPS. In addition to these residues, *Pf*-DAHPS has Trp 263 also present in the crystal structure of *Mt*-DAHPS (Trp 280), which is conserved in type II DAHPS. The role of conserved Trp 263 in the catalytic mechanism of type II DAHPS is to be investigated. The divalent metal binding ligands are contributed by four conserved residues His 352, Glu 394, Cys 69 and Asp 424 in *Pf*-DAHPS structure. Similar to *Mt*-DAHPS, Cys 423 is present close to the metal binding Cys 69, which can form a disulfide bridge. Thus, it is expected that Cys 423 of *Pf*-DAHPS plays a role in the regulation of the enzyme activity under reducing conditions. Arg 267, Arg 117, Ser 118 of *Pf*-DAHPS model are expected to interact with the phosphate group of E4P substrate as these residues correspond to the Arg 135, Arg 284 and Ser 136 residues that interact with sulfate ion in the crystal structure of *Mt*-DAHPS. Arg 267 in *Pf*-DAHPS structure is forming a bridge between the phosphate groups of PEP and E4P.

Type I DAHPS enzymes have been reported to form active homodimers that further associate to produce homotetrameric oligomers, whereas *Mt*-DAHPS, a type II enzyme forms only dimeric structure. The N-terminal region that extends the core catalytic barrel of these enzymes plays a crucial role in DAHPS dimerization [11, 13, 33, 36, 37]. It is anticipated that *Pf*-DAHPS also exists as a dimer, based on the observation that all DAHPS enzymes exist in dimeric form. Superimposition of *Pf*-DAHPS onto the dimeric structure of *Mt*-DAHPS indicates that the interface of *Pf*-DAHPS is highly hydrophobic (~ 60 %). However, the main difference in the dimeric interface of the two structures is the absence of non-core β0 strand

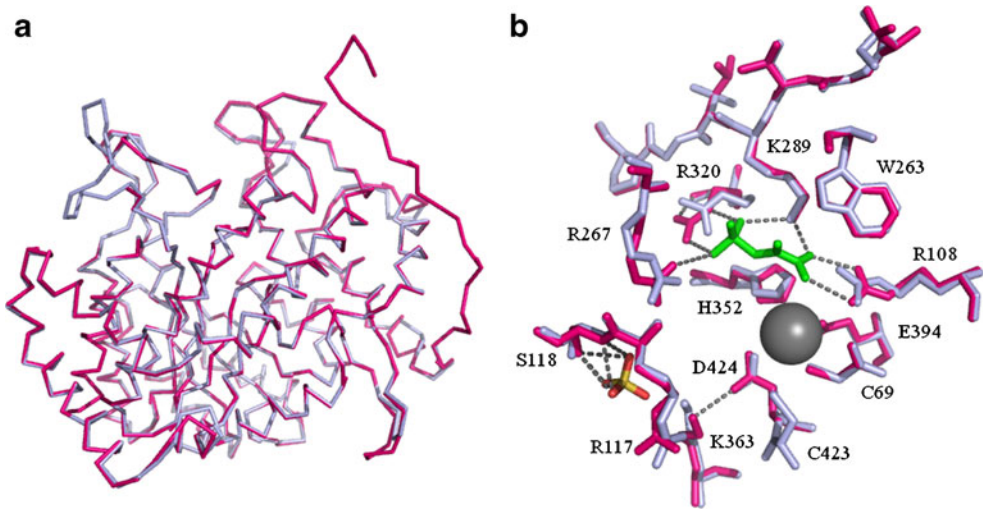

Fig. 8 (**a**) Superimposition of *P. fragi* DAHP synthase (light color) with *M. tuberculosis* DAHP synthase (Dark color) developed in PyMOL. (**b**) Superimposition of active site residues of *P. fragi* DAHP synthase (light color) and *M. tuberculosis* DAHP synthase (dark color). *Pf*-DAHPS active site residues interacting with substrate PEP, manganese ion (sphere) and sulfate are labeled. Polar interactions are shown with dotted lines. Figure prepared using PyMOL [32]

(Trp 3 - Ile 9) and the salt bridge between Asp 10_A and Arg 171_B of *Mt*-DAHPS in the structure of *Pf*-DAHPS. The role of non-core secondary structural elements and residues in the oligomerization of *Pf*-DAHPS can further be elucidated by site-directed mutagenesis.

The stretch of~41 residues in *Pf*-DAHPS (Tyr 164 to Arg 205) is characteristically present in DAHPS that are involved primarily in aromatic amino acid synthesis and is absent in DAHPS involved in secondary metabolite synthesis such as phenazine pigment production or antibiotic synthesis. This region in DAHPS has been implicated to be involved in the allosteric feedback regulation of the aromatic amino acid biosynthetic pathway [13, 35, 38]. Therefore, the internal non-core α2a and α2b helices of *Pf*-DAHPS structure which are part of this region are suspected to bind allosteric inhibitors and regulate the enzymatic activity.

Conclusions

In this study, the molecular model of DAHP synthase from *Pseudomnas fragi* was developed from the deduced amino acid sequence on the basis of homology modeling using the crystal structure of *Mycobacterium tuberculosis* as a template. The highly conserved DAHP synthase sequence motifs the $_{108}$RxxxxxxKPRT(S/T)$_{118}$, $_{264}$xGxR$_{267}$ and $_{349}$DxxHxN$_{354}$ were identified in *Pf*-DAHPS based on primary structure analysis. The 3D model of *Pf*-DAHPS and its comparison with *Mt*-DAHPS also indicates that these conserved motifs are involved in binding E4P, PEP and metal ion. Site directed mutagenesis and biochemical studies can further confirm the proposed role of the residues in these conserved sequence motifs. However, the major difference in the two structures is the absence of non-core N-terminal β0-strand in *Pf*-DAHPS 3D model which is participating in the dimer formation in crystal structure of *Mt*-DAHPS structure. Sequence analysis reveals that DAHPS of all the *Pseudomonas* species share high sequence identity and this points to the possibility of designing and developing a common DAHPS enzyme inhibitor for all the species of *Pseudomonas* genus. The molecular model of *Pf*-DAHPS may be helpful in the future for understanding the enzyme mechanism and structure-based molecular designing of inhibitors.

Acknowledgments We thank Dr. Pravindra Kumar, Dr. Ashwani K. Sharma and Ruchika Bajaj for helpful discussions and critical reading of the manuscript. We also thank Macromolecular Crystallographic Unit (MCU) for providing cloning and computational facilities. Satya thanks NDF, Girijesh and Sonali thank Ministry of Human Resource Development (MHRD) and Shailly thanks MHRD, FIG (Scheme B) for the financial support.

References

1. Pereira JN, Morgan ME (1957) Nutrition and physiology of Pseudomonas fragi. J Bacteriol 74:710–713
2. Lebert I, Begot C, Lebert A (1998) Growth of Pseudomonas fluorescens and Pseudomonas fragi in a meat medium as affected by pH (5.8–7.0), water activity (0.97–1.00) and temperature (7–25 °C). Int J Food Microbiol 39:53–60
3. Collins EB (1966) Use of Preservatives in Milk and Dairy Products. J Diary Sci 50:599–603
4. Moustafa HH, Collins EB (1969) Effects of selected food additives on growth of Pseudomonas fragi. J Diary Sci 52:335–340
5. Herrmann KM (1995) The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. Plant Cell 7:907–919
6. Bentley R (1990) The shikimate pathway: a metabolic tree with many branches. Crit Rev Biochem Mol Biol 25:307–384
7. Knaggs AR (2003) The biosynthesis of shikimate metabolites. Nat Prod Rep 20:119–136
8. McConkey GA (1999) Targeting the shikimate pathway in the malaria parasite Plasmodium falciparum. Antimicrob Agents Chemother 43:175–177
9. Coggins JR, Abell C, Evans LB, Frederickson M, Robinson DA, Roszak AW, Lapthorn AP (2003) Experiences with the shikimate-pathway enzymes as targets for rational drug design. Biochem Soc Trans 31:548–552
10. Kapnick SM, Zhang Y (2008) New tuberculosis drug development: targeting the shikimate pathway. Expert Opin Drug Discovery 3:565–577
11. Walker GE, Dunbar B, Hunter IS, Nimmo HG, Coggins JR (1996) Evidence for a novel class of microbial 3-Deoxy-D-arabino-Heptulosonate-7-Phosphate Synthase in Streptomyces coelicolor A3(2), Streptomyces rimosus and Neurospora crassa. Microbiology 142:1973–1982
12. Birck MR, Woodard RW (2001) Aquifex aeolicus 3-deoxy-D-manno-2-octulosonic acid 8-phosphate synthase: A new class of KDO 8-P synthase? J Mol Evol 52:205–214
13. Webby CJ, Baker HM, Lott JS, Baker EN, Parker EJ (2005) The structure of 3-Deoxy-D-arabino-Heptulosonate 7-Phosphate Synthase from Mycobacterium tuberculosis reveals a common catalytic scaffold and ancestry for type I and type II enzymes. J Mol Biol 354:927–939
14. Salcher O, Lingens F (1980) Regulation of phospho-2-keto-3-deoxy-heptonate aldolase (DAHP Synthase) and anthranilate synthase of Pseudomonas aureofaciens. J Gen Microbiol 121:473–476
15. Entus R, Poling M, Herrman KM (2002) Redox regulation of Arabidopsis 3-Deoxy-D-arabino-Heptulosonate 7-Phosphate Synthase. Plant Physiol 129:1866–1871
16. Hartmann M, Schneider TR, Pfeil A, Heinrich G, Lipscomb WN, Braus GH (2003) Evolution of feedback-inhibited beta/alpha barrel isoenzymes by gene duplication and a single mutation. Proc Natl Acad Sci 100:862–867
17. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Lee DG, Urbach JM, Wu G et al (2006) Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial. Genome Biol 7:R90
19. Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of Escherichia coli with plasmids. Gene 96:23–28
20. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389–3402

21. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucl Acids Res 31:3784–3788
22. Larkin MA, Blackshields G, Brown NP et al (2007) ClustalW and ClustalX version 2. Bioinformatics 23:2947–2948
23. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15:305–308
24. Brodsky LI, Ivanov VV, Kalai dzidis YL, Leontovich AM, Nikolaev VK, Feranchuk SI, Dracher VA (1995) GeneBee-NET: Internet-based server for analyzing biopolymers structure. Biochemistry 60:923–928
25. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 29:291–325
26. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815
27. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
28. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2:1511–1519
29. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucl Acids Res 35:W407–W410
30. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comput 4:435–447
31. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132
32. DeLano WL (2002) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
33. Schofield LR, Anderson BF, Patchett ML, Norris GE, Jameson GB, Parker EJ (2005) Substrate ambiguity and crystal structure of Pyrococcus furiosus 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase: an ancestral 3-deoxyald-2-ulosonate-phosphate synthase? Biochemistry 44:11950–11962
34. Silakowski B, Kunze B, Müller R (2000) Stigmatella aurantiaca Sg a15 carries genes encoding type I and type II 3-deoxy-D-arabino-heptulosonate-7-phosphate synthases: involvement of a type II synthase in aurachin biosynthesis. Arch Microbiol 173:403–411
35. Gosset G, Bonner CA, Jensen RA (2001) Microbial origin of plant-type 2-keto-3-deoxy-D-arabino-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from Xanthomonas campestris. J Bacteriol 183:4061–4070
36. Shumilin IA, Kretsinger RH, Bauerle RH (1999) Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from Escherichia coli. Structure 7:865–875
37. Shumilin IA, Bauerle R, Wu J, Woodard RW, Kretsinger RH (2004) Crystal structure of the reaction complex of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from Thermotoga maritima refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. J Mol Biol 341:455–466
38. Webby CJ, Patchett ML, Parker EJ (2005) Characterization of a recombinant type II 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from Helicobacter pylori. Biochem J 390:223–230