International Journal of Biological Macromolecules 49 (2011) 767–777

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Structural analysis of chorismate synthase from *Plasmodium falciparum*: A novel target for antimalaria drug discovery

Satya Tapas [1], Abhinav Kumar [1], Sonali Dhindwal, Preeti, Pravindra Kumar*

Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee 247667, Uttarakhand, India

ARTICLE INFO

Article history:
Received 10 May 2011
Received in revised form 12 July 2011
Accepted 13 July 2011
Available online 22 July 2011

Keywords:
Chorismate synthase
Shikimate pathway
Plasmodium falciparum
Homology modeling

ABSTRACT

The shikimate pathway in *Plasmodium falciparum* provides several targets for designing novel antiparasitic agents for the treatment of malaria. Chorismate synthase (CS) is a key enzyme in the shikimate pathway which catalyzes the seventh and final step of the pathway. *P. falciparum* chorismate synthase (*Pf*CS) is unique in terms of enzymatic behavior, cellular localization and in having two additional amino acid inserts compared to any other CS. The structure of *Pf*CS along with cofactor FMN was predicted by homology modeling using crystal structure of *Helicobacter pylori* chorismate synthase (*Hp*CS). The quality of the model was validated using structure analysis servers and molecular dynamics. Dimeric form of *Pf*CS was generated and the FMN binding mechanism involving movement of loop near active site has been proposed. Active site pocket has been identified and substrate 5-enolpyruvylshikimate 3-phosphate (EPSP) along with screened potent inhibitors has been docked. The study resulted in identification of putative inhibitors of *Pf*CS with binding efficiency in nanomolar range. The selected putative inhibitors could lead to the development of anti-malarial drugs.

1. Introduction

The phylum apicomplexa comprises of nearly 5000 species, most of which are parasites and are the causative agents of various pernicious diseases [1]. The most prominent apicomplexa is the *Plasmodium* species, which is the causative agent of malaria, a disease that is affecting nearly 300–500 million people every year worldwide. With the resurrection of drug resistant *Plasmodium falciparum*, the most fatal human malaria parasite, identification of new possible medicinal targets is an extreme priority. In such scenario, *P. falciparum* genome has facilitated the description of several metabolic pathways, particularly those differing from humans; thus providing new targets for drug development [2,3].

The shikimate pathway, in apicomplexan parasites provides a promising and exciting opportunity for reaching such objectives, as the pathway is only found in algae, plants, bacteria, protozoa and fungi, but is absent from mammals. The pathway is responsible for the formation of key aromatic amino acids involved in primary metabolism. In addition, the pathway in apicomplexa is important for the supply of folates, for which animals rely exclusively on an exogenous source [4]. Moreover, some of the enzymes of the pathway catalyze biochemically unique reactions in nature, making them excellent targets for new antiparasite drugs. Enzymes of this metabolic pathway have been studied extensively by various authors [5–7]. Chorismate synthase (EC 4.2.3.5) is the last enzyme of the pathway which catalyzes the conversion of the 5-enolpyruvylshikimate-3-phosphate (EPSP) to chorismate. The CS reaction comprises an *anti*-1,4-elimination of the 3-phosphate group and the C-6 *pro* R hydrogen with requirement of reduced FMN as a cofactor [8,9]. The catalysis action of CS does not involve any overall change in redox state, hence it is considered to be unique reaction in nature [9]. According to Bornemann et al., the reduced FMN donates an electron to EPSP to facilitate the loss of the phosphate and receive it back after the reaction [8].

According to the functionality, CS has been divided into two classes, monofunctional and bifunctional. Chorismate synthases from plants and eubacteria possess only *trans* elimination activity of substrate and are called monofunctional. However, CS from *Neurospora crassa* and *Saccharomyces cerevisiae* have an additional NADPH:FMN oxidoreductase activity, so-called bifunctional. Despite of resemblance with fungal CS, *Pf*CS is reported to be monofunctional [10].

In the present study, we performed *in silico* molecular modeling of three-dimensional structure of chorismate synthase enzyme from *P. falciparum*. We also performed an *in silico* structure-based inhibitors study using various substrate analogs and *in vitro* characterized inhibitors series. In absence of the crystal structure, we

* Corresponding author. Tel.: +91 1332 286312; fax: +91 1332 273560.
E-mail addresses: satyadbt@iitr.ernet.in (S. Tapas), abhi2pbt@iitr.ernet.in (A. Kumar), sonaldbt@iitr.ernet.in (S. Dhindwal), prityDBS@iitr.ernet.in (Preeti), pravshai@gmail.com, kumarfbs@iitr.ernet.in (P. Kumar).
[1] These authors contributed equally to this work.

0141-8130/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijbiomac.2011.07.011

hope that the proposed 3D model will be helpful for providing novel target for structure based drug design against malaria.

2. Methodology

The computational analysis was done on Intel Core i3-2.13 GHz Processor running on Windows 2007 Home Basic. The AUTODOCK version 4 and Molecular Dynamics (MD) simulations (GROMACS) were performed on Red Hat Enterprise Linux 5 operation system (Red Hat Inc., Raleigh, NC) installed on a Dell Precision T5400 workstation [11,12]. The Molecular Docking was performed on GLIDE 5.5 program (Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009) running on Windows 2003 on a HP xw8400 Workstation. All the graphical analysis and image production was done using PyMOL [13] and WinCoot software [14].

2.1. Multiple sequence alignment and phylogenetic tree construction

The CS sequences of *Plasmodium* species along with representative of bacteria, algae, fungi and plant were retrieved from NCBI database. A multiple sequence alignment for these sequences was generated using ClustalW with default parameters [15]. The graphical enhancement of the aligned sequences was performed using ESPript 2.2 server [16]. The phylogenetic tree was inferred using Phylogeny FR server [17].

2.2. Comparative molecular modeling, model optimization, evaluation and dimerization

Suitable template for modeling of *Pf*CS was searched using a PSI-BLAST against PDB database with default parameters [18]. From the best hits, crystal structure of chorismate synthase from *Helicobacter pylori* in complex with FMN (PDB 1UM0) was selected for model building. ClustalW program was used for multiple sequence alignment of query sequence with template sequence. Some manual corrections were done (as discussed in Section 3.2) in the alignment file for additional residues in *Pf*CS amino acid sequence. The sequence identity of modified *Pf*CS with the template was nearly 34%. Based on sequence alignment analysis, it was assumed that cofactor binding mode of *Pf*CS is similar to that of *Hp*CS. Therefore, FMN from the template was also incorporated in the modeled structure of *Pf*CS during model generation.

Program Modeller9v8 was employed to generate ten models of *Pf*CS [19]. Three sets of model having lowest DOPE scores were selected and stereo-chemical quality of each was evaluated by PROCHECK [20]. The model with least number of residues in the disallowed region was further refined for relieving steric clashes and improper contacts by energy minimization using Swiss PDB viewer 4.0.1 (http://spdbv.vital-it.ch/). Swiss PDB viewer implements GROMOS96 force field to compute energy and execute energy minimization. PROCHECK and ERRAT plot were further used to evaluate stereo-chemical quality of the model [21]. Loop refinement tool of the 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 that gives a measure of structural error at each residue in the protein. This process was repeated iteratively until most of the amino acid residues were below 95% cutoff value in ERRAT plot. The refined model was further validated by VERIFY 3D of SAVES server (http://nihserver.mbi.ucla.edu/SAVES/) [22]. ProSA-Web server (https://prosa.services.came.sbg.ac.at/prosa.php) was also used to evaluate the generated 3D model of protein for potential error [23]. The dimer of *Pf*CS was generated by Modeller9v8 using the dimeric state of *Hp*CS as template. Furthermore, the ClusPro 2.0 Protein-Protein docking server (http://cluspro.bu.edu/login.php?redir=/) was used to generate *Pf*CS dimer to compare with dimeric model generated by Modeller9v8 [24].

GROMACS simulation suite version v.4.0.7 was used to perform Molecular Dynamics of predicted *Pf*CS model [12]. The energy minimization was performed by giving 1000 steps of Steepest Descent to eliminate bad atomic contacts, which converged in less than 800 steps only. Position restrained molecular dynamics was executed for 50 ps by partially restraining the atomic positions of the protein while letting the water molecules move into it during simulation. The equilibrated system was then subjected to non-restrained molecular dynamics simulation of 1000 ps (1 ns). The protein stability was assessed by determining the root mean square deviation (RMSD) between the structures generated before and after MD simulation.

2.3. Docking studies of PfCS

The docking studies of substrate and inhibitors were performed using GLIDE (*Grid-based Ligand Docking with Energetics*) version 5.5 software by Schrödinger running on Windows 2003 on a HP xw8400 workstation. GLIDE approximates a complete systematic search of the conformational and positional space of the docked ligand [25]. For docking, the protein was prepared by adding hydrogens, assigning bond orders and minimizing overall structure to RMSD of 0.30 Å using OPLS2001 force-field in Maestro's protein preparation wizard. All the substrates were prepared using Maestro's Ligprep module (Ligprep, Schrödinger Inc., New York, NY). Receptor grid for docking was generated by the centroids of the selected amino acid residues of the active site. GLIDE was further used for the docking of the substrates into active site using Extra Precision (XP). Best pose was selected on the basis of Glidescore, rank and by visually inspecting the molecule in PyMOL. The best

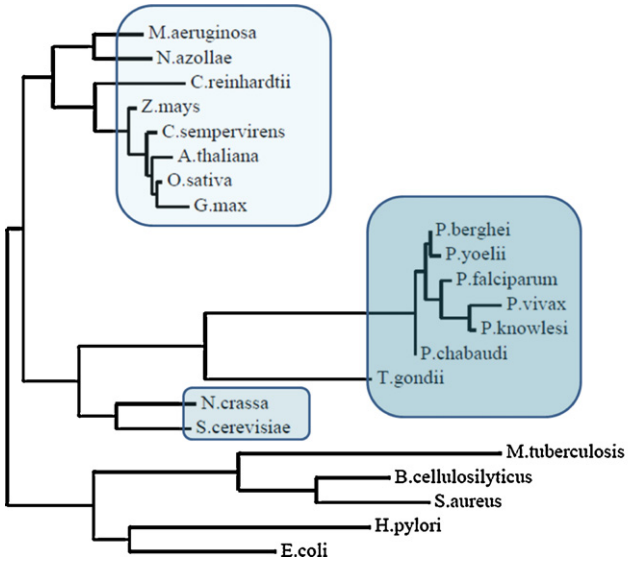

Fig. 1. Phylogeny of chorismate synthases. Phylogenetic reconstruction of chorismate synthase amino acid sequences from higher plant *Z. mays* (NP_001148583.1), *C. sempervirens* (CAA43034.1), *A. thaliana* (NP_001031158.1), *O. sativa* (BAD14928.1), *G. max* (ABA90483.1); apicomplexa *P. berghei* (XP_678920.1), *P. yoelii* (XP_724321.1), *P. falciparum* (AAB63293.1), *P. vivax* (AAL56611.1), *P. knowlesi* (XP_002261447), *P. chabaudi* (XP_743671.1), *T. gondii* (AAB52422.1); fungi *N. crassa* (AAC49056.1), *S. cerevisiae* (CAA42745.1); bacteria *M. aeruginosa* (YP_001656924.1), *N. azollae* (ADI63255.1), *M. tuberculosis* (NP_217056.1), *B. cellulosilyticus* (YP_004094866), *S. aureus* (NP_374580.1), *H. pylori* (AAB63293.1), *E. coli* (CAA68707.1); algae *C. reinhardtii* (XP_001695611.1). The apicomplexa CS follow a single branch and cluster together due to high sequence identity; however, they have nearest homologues to fungi CS. Though, apicomplexa CS physiologically similar to plant CS but make distant clusters in phylogenetic tree. The discussed separate clusters are highlighted in boxes. The phylogenetic tree was generated using Phylogeny FR server.