JPP JOURNAL OF PROTEINS AND PROTEOMICS
4(1), June 2013, pp. 21-25

Research Article

HOMOLOGY MODELING AND FUNCTIONAL CHARACTERIZATION OF THREE-DIMENSIONAL STRUCTURE OF DAHP SYNTHASE FROM *BRACHYPODIUM DISTACHYON*

Aditya Dev, Nandita Bodra, Pramod Kumar, Shivendra Pratap, Pravindra Kumar*

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

Abstract: The Shikimate pathway is an attractive target for herbicides and antimicrobial agents because it is essential in microbes and plants but absent in animals. The 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS) is the first enzyme of this pathway, which is involved in the condensation of phosphoenolpyruvate (PEP) and D-erythrose 4-phosphate (E4P) to produce 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP). DAHPS enzymes have been divided into two types, class I and class II, based on their primary amino acid sequence and three dimensional structures. The plant DAHPS belongs to class II and is regulated differently than DAHPS from microorganisms. To understand the structural basis of such differences in DAHPS from plants and its catalytic mechanism, we have used sequence analysis, homology modeling and docking approach to generate the three dimensional models of DAHP synthase from *Brachypodium distachyon (Bd*-DAHPS) complexed with substrate PEP for the first time. The three dimensional models of *Bd*-DAHPS provides a detailed knowledge of the active site and the important secondary structural regions that play significant roles in the regulatory mechanism and further may be helpful for design of specific inhibitors towards herbicide development.

Keywords: DAHP Synthase; Shikimate pathway; Molecular modeling; Three-dimensional structure; Multiple sequence alignment; Docking.

Introduction

The Shikimate pathway is an important and one of the common pathway for the biosynthesis of various important compounds including folic acid, vitamin K, ubiquinone and the aromatic compounds in microorganisms as well as in plants. The seventh enzyme of this pathway is responsible for the synthesis of chorismate which serve as precursors of the aromatic amino acids (Phe, Tyr, and Trp), catechols, *p*-aminobenzoic acid and number of other secondary metabolites (Pittard, 1987; Braus, 1991). Shikimate pathway is essential in bacteria, fungi algae and higher plants, while it is absent in animals, thus the pathway has been an attractive target for the development of herbicides and antimicrobial agents against number of diseases (Bentley, 1990; Arcuri *et al.*, 2004).

First step of this pathway consist of condensation of the glycolytic intermediate phosphoenolpyruvate (PEP) and the pentose phosphate pathway intermediate D-erythrose 4-phosphate (E4P) to give a seven-carbon, six-member heterocyclic compound, 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP). This reaction is catalyzed by DAHP synthase that displaces the high energy phosphoryl group of PEP, allowing the enolpyruvate methylene to attack the aldehyde of E4P (Arcuri *et al.*, 2004).

Corresponding Author: **Pravindra Kumar**
E-mail: kumarfbs@iitr.ernet.in, pravshai@gmail.com

Received: January 30, 2013
Accepted: June 24, 2013
Published: July 30, 2013

Further steps involve six other enzymes which serve to the subsequent formation of an aromatic compound, chorismate that is the final product of the main trunk of the Shikimate pathway. Chorismate then acts as a substrate for several anabolic metabolites leading to primary and secondary compounds (Bentley, 1990; Herrmann, 1995).

DAHPS is metal-activated enzyme and is ubiquitous in archaebacteria, fungi and plants. In micro-organisms, DAHPS regulates the pathway by negative feedback inhibition caused by three aromatic amino acids, Phenylalanine, Tyrosine or Tryptophan. This is in contrast from plants, where none of the DAHP synthase is inhibited by feedback mechanism. However, plant DAHP synthase is known to undergo metabolic regulation (Herrmann, 1995; Shumilin *et al.*, 2003; Konig *et al.*, 2004). On the basis of primary structure homology and molecular mass, two unrelated type (Class I & Class II) of DAHP synthase are classified (Bentley, 1990). Class I is majorly found in prokaryotic and archae organisms, although some eukaryotic examples have also been identified (*S. cerevisiae, N. crassa*). Class II enzymes were originally identified in plants but are now known to have some microbial proteins too, as in *M. tuberculosis*. Some key features are shared between Class I and Class II DAHP syntheses. The sequences of these classes of DAHPS are reasonably similar. Despite their presence in diverse organisms and apparent differences in architecture, they display common motifs and conserved active site residues. These results show a monophyletic origin of the DAHP synthases and thus giving the basis for molecular modeling studies to elucidate the structure function relationships among them (Webby *et al.*, 2005 "a"; Webby *et al.*, 2005 "b").

There are crystal structures available for microbial DAHPS like *E.coli*, *S. cerevisiae, N. crassa, M. tuberculosis* but plant DAHP synthase structure still remains elusive. In the present work the DAHP synthase from *Brachypodium distachyon,* the first member of Pooideae subfamily to be sequenced has been studied. This plant is a grass species native to northern Africa, southern Europe and southwestern Asia east to India, has zero commercial or agricultural importance. As a weed it grows easily without specialized growing conditions. The genome of *Brachypodium distachyon* was sequenced recently in 2010, which helped to study the enzyme from this plant using bioinformatics. *Brachypodium distachyon*, an important model system for developing new energy and food crops because of having the high-quality genome sequence, small size and rapid life cycle.

The three dimensional structural knowledge of this enzyme in plants is of utmost importance since its information can be used as target for designing selective and specific inhibitors for herbicide development. This study will enhance and act as an aid in the design of specific structure based inhibitor that may be used as herbicides against plant weeds.

Material and Methods

The amino acid sequence of *Bd*-DAHPS (Accession No. XP_003562680.1) was retrieved from the NCBI (*http://www.ncbi.nlm.nih.gov/*) and taken as target sequence. BLAST was used to search the homologous crystal structure available in Protein Data Bank (Altschul *et al.*, 1997; Berman *et al.*, 2000).

A potential template structure (PDB-ID:2B7O), representing the crystal structure of DAHPS from *Mycobacterium tuberculosis,* consisting of a metal ion Mn and a substrate PEP was used for building *Bd*-DAHP synthase structure. MODELLER 9v7 (Sali and Blundell, 1993) was used for the comparative modeling. Initially models for each of the isoenzyme were generated and then minimized using Swiss-PdbViewer 4.01 (Guex and Peitsch, 1997). Based on stereochemical properties, best model was taken for further refinement using MODELLER's loop refinement tool (Fiser *et al.*, 2000). Validation of the models was performed using Ramachandran plot, ERRAT PLOT and ProSA analysis (Luthy *et al.*, 1992; Sippl, 1993; Laskowski *et al.*, 1993). The generated model was visualized, inspected and analyzed using PyMOL (DeLano, 2002). Maestro suite Glide was used for the molecular docking (Halgren *et al.*, 2004).

Results and Discussion

Sequence Alignment

The multiple sequence alignment of DAHP synthase was done to locate the motif (RxxxxKPRS/T) in *Bd*-DAHPS as shown in Figure 1 with yellow background. This is a phosphate binding motif known to be involved in binding of E4P in DAHPS (Dev *et al.*, 2012). It was found that this motif has got two insertions in *Bd*-DAHPS and is present as Rxxxx**xx**KPRS.

E4P, the second substrate of DAHPS undergoes condensation with the other substrate PEP, to produce DAHP. There are a few crystal structures of DAHPS available with PEP but till now there is only one structure, (PDB ID:1RZM) which is available with E4P at the active site, reported from *Thermotoga maritime.* The available binding site information of E4P from *Thermotoga maritime* was used and compared in with that of *Bd*-DAHPS by multiple sequence alignment**.** The multiple sequence alignment of 1RZM with *Bd*-DAHP synthase and template sequence 2B7O was performed with ClustalW as shown in Figure 1 (Thompson *et al.*, 1994; Shumilin *et al.*, 2002).

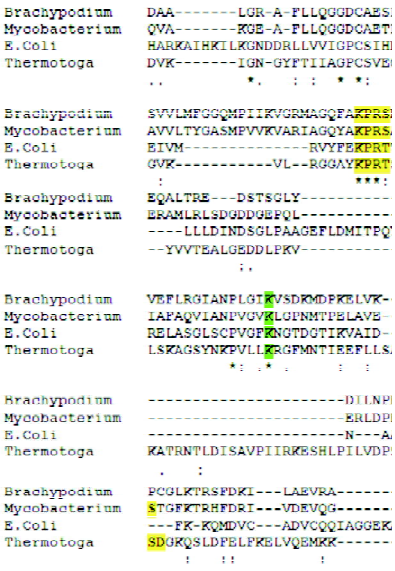

Figure 1: **Multiple sequence alignment of DAHP *of Brachypodium distachyon, Mycobacterium tuberculosis, E.Coli* and *Thermotoga maritima* (1RZM). Yellow and green background regions show the conserved residues involved in E4P and PEP binding respectively**

Homology modeling

The DAHP synthase from Mycobacterium tuberculosis (PDB ID: 2B7O) was used as template for homology modeling of DAHP synthase from *Brachypodium distachyon*. The 3D generated structure was subjected to energy minimization. The minimized structure was further validated using Ramachandran plot analysis by PROCHECK server. The amino acid environment was evaluated using ERRAT server. The *Mt*-DAHPS was also evaluated and compared with the generated model (Table 1).

The above result of stereochemical data of molecular backbone and side chain environment suggests a relatively accurate model. The structure prediction reveals that the *Bd*-DAHPS belongs to other DAHPS characterized till date, and consist of a core $(\beta/\alpha)_8$ TIM barrel. The superimposition of the model with template showed low RMSD of 0.207. This also suggests high similarity of the model to the template (Figure 2). Further the energy function of the model was evaluated using ProSA server. The

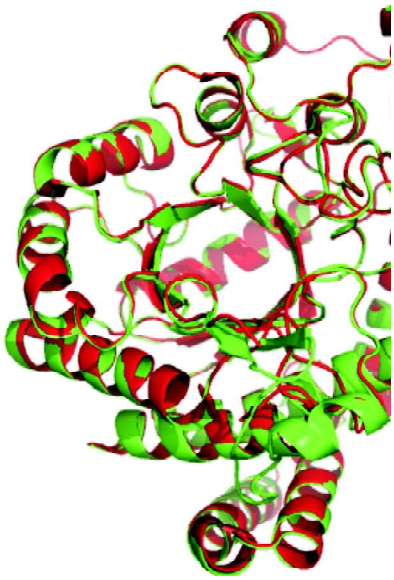

Figure 2: **Superimpostion between *Bd*-DAHPS model and template *Mt*-DAHPS (PDB ID: 2B7O) in green and red color respectively**

ProSA tool graph shows all over model quality of the structure and the location of the *z*-score for the structure. The overall Z-score evaluated for the model came to be -9.82. The negative Z-score of energy profile confirmed good overall quality of the model. This also represented the point of the structure that was within a range which was solved by X-ray and NMR (Figure 3).

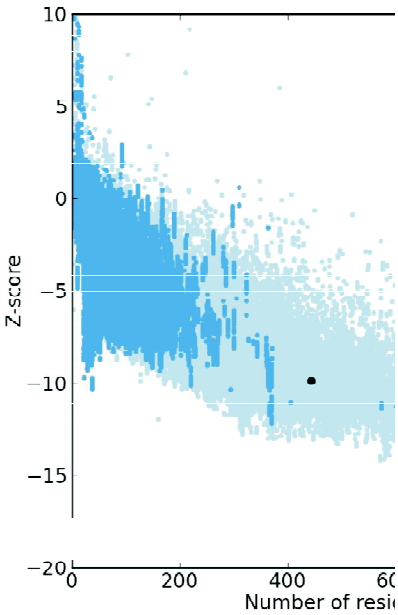

Figure 3: **Structural validation of *Bd*-DAHP synthase using ProSA tool. The overall Z-score of -9.82 has been shown with a black dot**

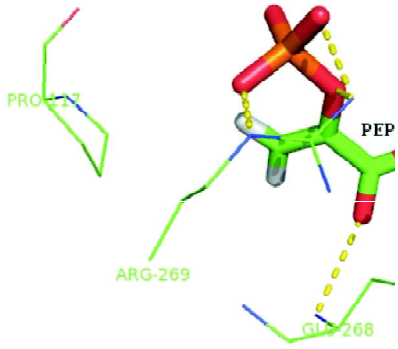

Figure 4: **The binding site of PEP in *Bd*-DAHP synthase has been shown with hydrogen bonds in yellow dashes**

Table 1
Comparison of energy minimized M1 model with the *Mycobacterium tuberculosis* template

Protein	*Ramachandran plot*	*Errat*
M1 Model	99.5%	86.9%
Mt-DAHPS template	100%	94.9%

Binding of E4P and PEP at the active site

The active site of *Bd*-DAHP synthase is located at the C-terminal end of barrel and connecting loop between β-strands & α-helices. To study the interaction of substrates with the residues of *Bd*-DAHPS, PEP was docked in the energy minimized model by Glide (Maestro Suite). The interacting residues for the binding of PEP and E4P were compared using the known structural information from *Mt*-DAHP synthase and *Tm*-DAHP synthase (Figure 1 & 4). The key residues for the binding of PEP in *Bd*-DAHPS include Pro117, Glu268, Arg269, Lys291, Arg322 and His354. Out of them Pro117, Arg269 and Lys291 were found to be highly conserved in all the species compared in Figure 1. The PEP is bound in the active site of the *Bd*-DAHPS by multiple hydrogen bonds involving both its phosphate and carboxylate groups. The PEP and E4P undergoes aldol-like condensation forming 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) with the release of the phosphate of PEP.

Conclusion

In the present work the homology modeling of DAHP synthase from *Brachypodium distachyon* has been done. The *Bd*-DAHPS was found to contain the same $(alpha\text{-}beta)_8$ TIM barrel fold which is a characteristic of all other known DAHP synthase. The primary sequence analysis has shown large variation in DAHP synthase from plants and microbes. The sequences of plant and microbial enzymes showed high variations which could be responsible for their differential modes of regulation. Most of the residues at the active site for the binding of PEP were found to be conserved through sequence alignment. The docking study also confirmed the same with few substitutions. This led to the speculation that inhibitor development for PEP binding could facilitate in the development of herbicides. The specific herbicides could be developed by targeting those specific residues or by developing resistant transgenic crop.

Acknowledgements

AD, NB and SP thanks Department of Biotechnology, Science & Technology and PK thanks Council of Scientific and Industrial Research, New Delhi for the research fellowships provided.

Abbreviations

DAHP, 3-deoxy-D-arabino-heptulosonate 7-phosphate; DAHPS, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; E4P, D-erythrose 4-phosphate; PEP, phosphoenolpyruvate.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Arcuri, H. A., Canduri, F., Pereira, J. H., Da Silveira, N. J. F., Camera Jr, J. C., de Oliveira, J. S., Basso, L. A., Palma, M. S., and Santos, D. S. (2004). Molecular models for shikimate pathway enzymes of Xylella fastidiosa. Biochem. Biophys. Res. Commun. 320, 979-991.

Bentley, R. (1990). The shikimate pathway: A metabolic tree with many branches, Crit. ReV. Biochem. Mol. Biol. 25, 307-384.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I.N., and Bourne, P. E. (2000). The protein data bank. Nucleic Acids Res. 28, 235–242.

Braus, G. H. (1991). Aromatic amino acid biosynthesis in the yeast Saccharomyces cerevisiae: a model system for the regulation of a eukaryotic biosynthetic pathway. Microbiol Mol Biol Rev. 5, 349.

DeLano, W. L. (2002). The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA.

Dev, A., Tapas, S., Pratap, S., and Kumar, P. (2012). Structure and function of enzymes of shikimate pathway, Current Bioinformatics. 7,00-04. 374-391.

Fiser, A., Do, R. K., and Sali, A. (2000). Modeling of loops in protein structures. Protein Sci. 9, 1753-1773.

Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 18, 2714-2723.

Halgren, T. A, Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. J. Med. Chem. 47, 1750–1759.

Herrmann, K. M. (1995). The shikimate pathway as an entry to aromatic secondary metabolism. Plant Physiol. 107, 7–12.

Konig, V., Pfeil, A., Braus, G. H., and Schneider, T. R. (2004). Substrate and metal complexes of 3-deoxy- D-arabino-heptulosonate-7-phosphate synthase from Saccharomyces cerevisiae provide new insights into the catalytic mechanism, J. Mol. Biol. 337, 675–690.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Luthy, R., Bowie, J. U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature. 356, 83-85.

Pittard, A. J. (1987). Biosynthesis of the aromatic amino acids. *Escherichia coli* and *Salmonella typhimurium*. Am. Soc. Microbiol. 1, 368-394.

Sali, A., and Blundell T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779-815.

Shumilin, I. A., Bauerle, R., and Kiesinger, R. H. (2003). The high-resolution structure of 3-deoxy-D-arabinoheptulosonate- 7-phosphate synthase reveals a twist in the plane of bound phosphoenolpyruvate, Biochemistry. 42, 3766–3776.

Shumilin, I. A., Bauerle, R., Woodard, R. W., and Kretsinger, R. H. (2002). Crystal structure of the reaction complex of 3-deoxy-D-arabino-heptulosonate- 7-phosphate synthase from *Thermotoga maritime* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. J. Mol. Biol. 341, 455–466.

Sippl, MJ. (1993). Recognition of errors in three-dimensional structures of proteins. Proteins. 17, 355–362.

Thompson, J. D., Higgins, D. G., and Gibson, T. J., (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Webby, C. J., Baker, H. M., Shaun J. L., Baker, E.N., and Emily, J. (2005 “a”). 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.

Webby, C. J., Patchett, M. L., and Parker, E. J. (2005 “b”). Characterization of a recombinant type II 3-deoxy-Darabino- heptulosonate-7-phosphate synthase from Helicobacter pylori. Biochem. J. 390, 223–230.