Nature & Environment
Vol. 20 (2), 2015: 64-72
Website: www.natureandenvironment.com

ISSN (Print) : 2321-810X
ISSN (Online) : 2321-8738

RESEARCH ARTICLE

Combatting the Whooping Cough Bacteria: *in-silico* Approaches

Sushmita Bhattacharya
Department of Biotechnology, Indian Institute of Technology, Roorkee, India
Email: sushmita.b8@gmail.com

Received: 11th April 2015, Revised: 8th May 2015, Accepted: 16th May 2015

ABSTRACT

Bordetella pertussis causes pertussis or the whooping cough disease which claims an alarmingly number of lives each year owing possibly to the emergence of resistance to the antimicrobials of choice for the treatment of Bordetella pertussis infections. All the shikimate pathway enzymes form attractive targets for herbicides and antimicrobial agent development because the pathway is essential in algae, higher plants, bacteria, and fungi, but absent from mammals. In the present work molecular model of the sixth enzyme of the pathway, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase in Bordetella pertussis has been generated. The architecture of the active site has been described based on the prediction of active site residues. This should provide a structural framework to guide enzymological studies and design of specific inhibitors that could serve as new and effective antibiotics against the Bordetella pertussis bacterium.

Key words: EPSP synthase, Bioinformatics, Molecular modeling, Bordetella pertussis

INTRODUCTION

Lower respiratory tract infections in recent times have figured as one of the 10 biggest killer diseases in the developing world, ranked by annual death tolls [1]. They claim a lot many lives each year, estimated to be more than 4 million. These are mostly diseases like pneumonia and other diseases of the lungs, windpipe or bronchial tubes, including Legionnaire's disease, tuberculosis and whooping cough. Whooping cough alone claims 200,000 to 300,000 lives each year; the infection rate is 20 million to 40 million cases every year.

Pertussis (or Whooping Cough) is a highly communicable, though vaccine-preventable disease. *Bordetella pertussis,* the microorganism that causes pertussis is a Gram-negative, aerobic coccobacillus of the genus *Bordetella* [2]. In children, the disease is manifested with paroxysmal spasms of severe coughing, whooping, and posttussive vomiting. The paroxysmal cough precedes a crowing inspiratory sound, this is characteristic of pertussis. Adults show milder symptoms, like prolonged coughing without the "whoop". A coughing spell may cause cyanosis, apnoea and seizures. Other major complications are hypoxia, pneumonia, encephalopathy, and malnutrition. A prolonged cough could be irritating and at times a disabling cough may go undiagnosed in adults for many months. This disease results in high morbidity and mortality in many countries every year [1, 2].

The first clear description of pertussis (whooping cough) was provided by Baillou in 1640 [3]. The incidence of pertussis decreased dramatically in the late 1940s, when the whole-cell pertussis vaccines were introduced. However, even in countries with good vaccination rates, there have been recent reports of increases in cases in The Netherlands [4] and in the United States [5]. In the United States, it has shown that the incidence of pertussis is relatively stable in infants and young children, but it has increased in adolescents and young adults [5]. Although erythromycin or macrolides are still the antimicrobials of choice for the treatment of *Bordetella pertussis* infections,

there have been reports of emergence of resistance to these agents in clinical isolates from the United States [6, 7].

The shikimate pathway connects the metabolism of carbohydrates to the biosynthesis of ring-containing compounds [8]. In plants and microorganisms, all the key aromatic compounds involved in primary metabolism, are produced via shikimate pathway. The shikimate pathway is essential in algae, higher plants, bacteria, and fungi, but absent from mammals that depend on these compounds for their diet. Due to this, there has been interest, extending back more than 25 years, in the shikimate-pathway enzymes as potential targets against non-toxic herbicides and anti-microbial compounds [9]. The pathway is composed of seven steps which take the exogenous substrates, phosphoenol pyruvate and 4-erythrose phosphate, to yield chorismate. It is now confirmed that inhibitors of the aromatic amino acid biosynthetic pathway have herbicidal activity and also glyphosate (*N*-phosphonomethylglycine), which is the active constituent of herbicide RoundUp targets the shikimate pathway effectively. Also, glyphosate has been shown to inhibit the growth of several apicomplexan parasites, including those that cause malaria and toxoplasmosis. The precise enzyme target for glyphosate is 5-enolpyruvylshikimate- 3-phosphate synthase (EPSP synthase), the sixth enzyme of the pathway, encoded by aroA gene, that catalyzes the transfer of the enolpyruvyl moiety from phosphoenolpyruvate (PEP) and inorganic phosphate [10].

Recent studies have confirmed that the pathway is present in many pathogens and there is scope for developing broad-spectrum antibiotics. It is possible that a single compound that inhibits the shikimate pathway or combinations of synergistic compounds that inhibit various enzymes in this pathway could be of use in treating patients with multiple infections [11]. We exploit this information in devising strategies to combat the whooping cough disease by elucidating the structure of *Bordetella pertussis* EPSP synthase (BpEPSPS).

The present paper describes the molecular model of BpEPSPS, in complex with 3-phosphoshikimate (S3P) and glyphosate. The homology modeling was performed using the crystallographic structure of EPSPS from Escherichia coli, solved to resolution 1.5 Å, as template. The results presented here should provide a three-dimensional model of BpEPSPS to both guide enzymological studies and aid in the design of specific inhibitors against *Bordetella pertussis*, the causative of the dreaded disease whooping cough.

METHODS

Databases

The amino acid sequences of the experimentally characterized, 5-enol-pyruvyl-shikimate-3-phosphate (EPSP) synthase enzyme from different organisms were retrieved from the protein sequence database at NCBI http://www.ncbi.nlm.nih.gov. The 3-D structures of proteins were obtained from the protein data bank [12].

Servers

Protein sequence databases were searched using BLAST [13] or PSI-BLAST [14] servers at NCBI. FFAS03 [15], FUGUE [16], PHYRE (successor of 3D-PSSM, [17]), SAMT06 [18] and Gene Silico Metaserver [19] were used for fold-recognition. Transmembrane helices were predicted using the TMHMM server v. 2.0 [20]. Secondary structures were predicted using the SAM-T06 [21], JPRED [22], NNPREDICT [23], PSIPRED [24], SOPMA [25] and SSPRO [26] servers. All the servers were used with default values for the various parameters, except where mentioned otherwise.

Secondary structure prediction

The secondary structure of the BpEPSPS was predicted separately using six prediction servers mentioned earlier. The secondary structures were predicted as three states, helix (H), strand (E) and coil (C). A consensus secondary structure was obtained by comparing the predictions of the eight servers. If different secondary structure states are predicted for a residue by the servers, the

state that has been predicted by at least four of six servers was taken as the consensus state; in other cases, it was marked as U (uncertain).

Multiple sequence alignment

Multiple sequence alignment was performed using the TCoffee server [27, 28]. The sequence of BpEPSPS was aligned with the sequences of the same enzyme from Escherichia coli (EcEPSPS) (PDB Id: 1G6T), *Streptococcus* (SpEPSPS) (PDB Id: 1RF4), *Mycobacterium tuberculosis* (MtEPSPS) (PDB Id: 2BJB) and *Agrobacterium* (CP4 EPSPS) (PDB Id: 2GG4) (obtained from the protein data bank (http://www.rcsb.org).

Template search for 3D modelling

The amino acid sequence of BpEPSPS (accession no. gi|34978354|sp|P12421.2, entry name: AROA_BORPE) was obtained from Swiss-Prot database [29]. The PDB BLAST through NCBI was used to identify homologous structures by searching the structural database of protein sequences in the protein data bank (PDB) using default parameters. The sequence of BpEPSPS was provided as input to the Geno3D server (http://pbil.ibcp.fr/htm/.) Also, the most common and highest scoring template suggested by different fold-recognition servers was taken in consideration when choosing the template.

Sequence/template alignment

We first performed a BLAST search for proteins with similar sequence and known 3Dstructure using the 442 residue long BpEPSPS sequence (SWISS-PROT: P12421.2). Significant similarities were found with several EPSPS enzyme family suggesting that BpEPSPS enzyme sequence bears high similarity to the sequence of the same enzyme from different organisms and belongs to the same family. Additionally, a conserved domain search [30] of BpEPSPS sequence also supported the above idea. A PROSITE analysis of BpEPSPS sequence also indicated the presence of EPSPS domain signature and profile. The amino acid sequence of BpEPSPS was aligned with sequence of EPSPS from *E.coli* (EcEPSPS) extracted from its crystal structure, using EsPript [31].

MODEL GENERATION

The homology model of Bordetella pertussis EPSPS was constructed using the MODELLER version 9v1 (default parameters), based upon the sequence/template alignment. MODELLER is a well-known comparative modeling methodology, which generates a refined 3D homology model of a protein sequence automatically and rapidly, based on a given sequence alignment to a known 3D protein structure. This computational technique is based on the assumption that tertiary structures of two proteins will be similar if their sequences are related, and it is the approach most likely to give accurate results [32]. There are two main approaches to homology modeling: (1) fragment-based comparative modeling [33, 34] and (2) restrained-based modeling [35]. For modeling of the BpEPSPS we used the second approach. The geometry of loop regions was corrected using MODELLER/Refine Loop command.

MODELLER is an implementation of an automated approach to comparative modeling by satisfaction of spatial restraints [36-38]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and sidechain non-hydrogen atoms. Since there are no three-dimensional structures available for EPSP synthase enzyme of the BpEPSPS (target) the models were built based on the atomic coordinates of EcEPSP enzyme (template), which was solved by crystallographic methods.

The percentage of primary sequence identity between BpEPSPS enzyme and template EcEPSPS indicates that the crystallographic structure of enzyme (template) is a good model to be used as template for the BpEPSPS enzyme (target). Further, the ligands (glyphosate and S3P) in the active

site of the template were copied onto the model of the target to infer the active site residues in the target structure.

MODEL VALIDATION

The overall stereochemical quality of the final model for the complex BpEPSPS: Glyphosate: S3P was assessed by the program RAMPAGE [39]. The root mean square deviation (RMSD) between the main chain atoms of the model and template was calculated for the reliability of the model.

RESULTS AND DISCUSSION

Primary Structure comparison

The sequence alignment of EcEPSPS (1G6S) as template and BpEPSPS (target) is shown in Fig. 1. The secondary structural elements are indicated in the figure. The sequence of BpEPSPS shows 55% identity with the sequence of EcEPSPS.

Fig. 1: The sequence alignment of EcEPSPS (1G6S) and BpEPSPS indicating the secondary structural elements. The sequence BpEPSPS shows 55% of identity with the sequence of EcEPSPS. The alignment was performed with the program EsPript [31].

Fig. 2: Ramachandran diagram φ-ψ plot for the EPSPS structure (A) Template EcEPSPS (B) Model BpEPSPS

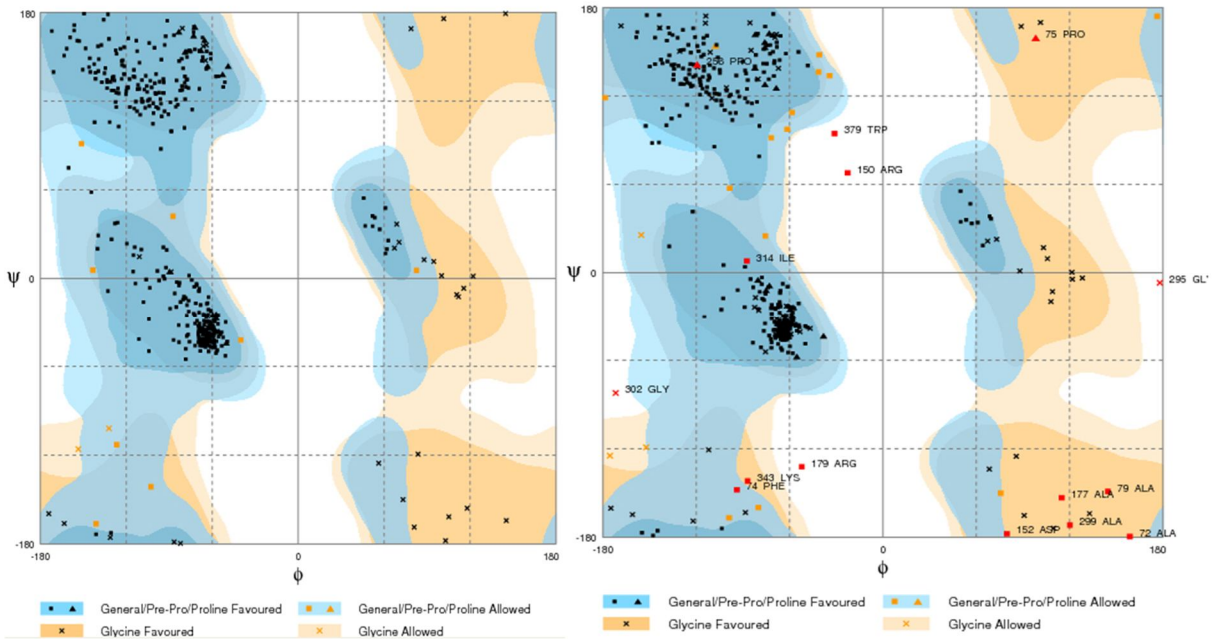


Fig. 3: (A) Predicted structure of BpEPSPS model (Glyphosate is represented bound as blue stick model and S3P as brown stick model). (B) Superposition of cartoon representations of BpEPSPS (green); EcEPSPS (1G6S) (cyan); Streptococcus (1RF6) (magenta), MtEPSPS (2BJB) (yellow); CP4 EPSPS (2GGA) (teal). (C) Residues in active site pocket of EPSPS.

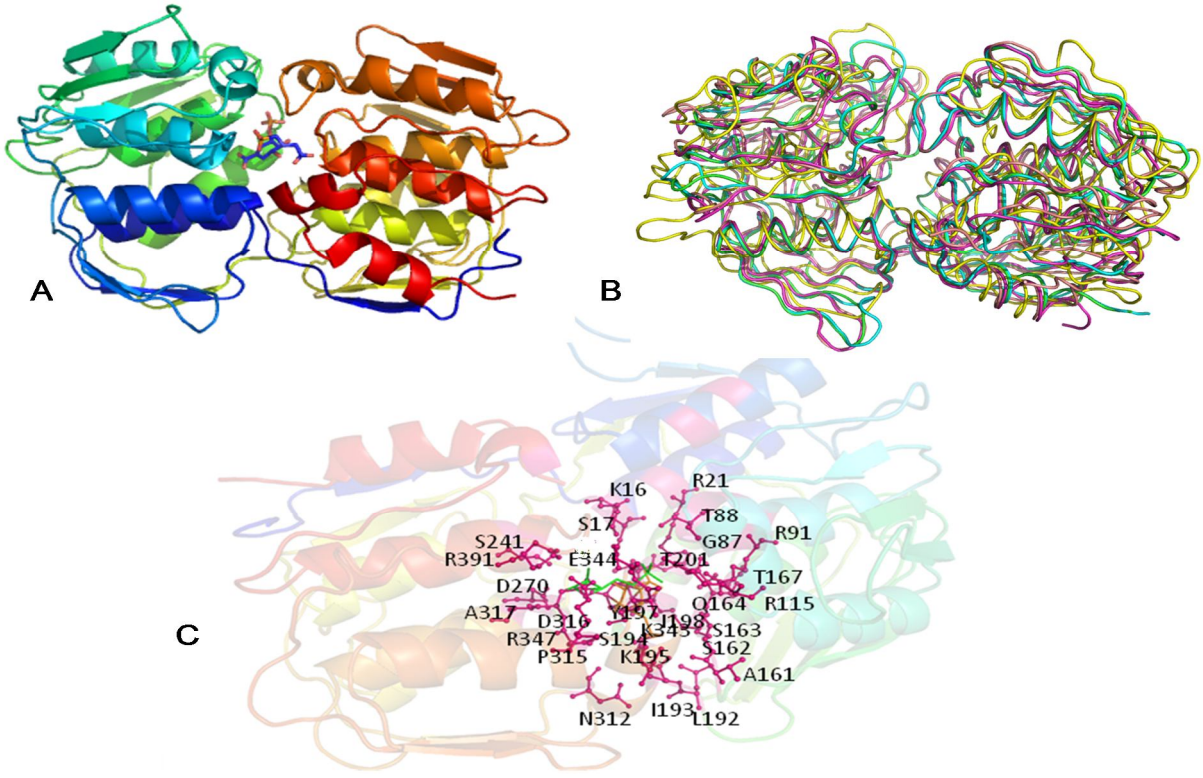
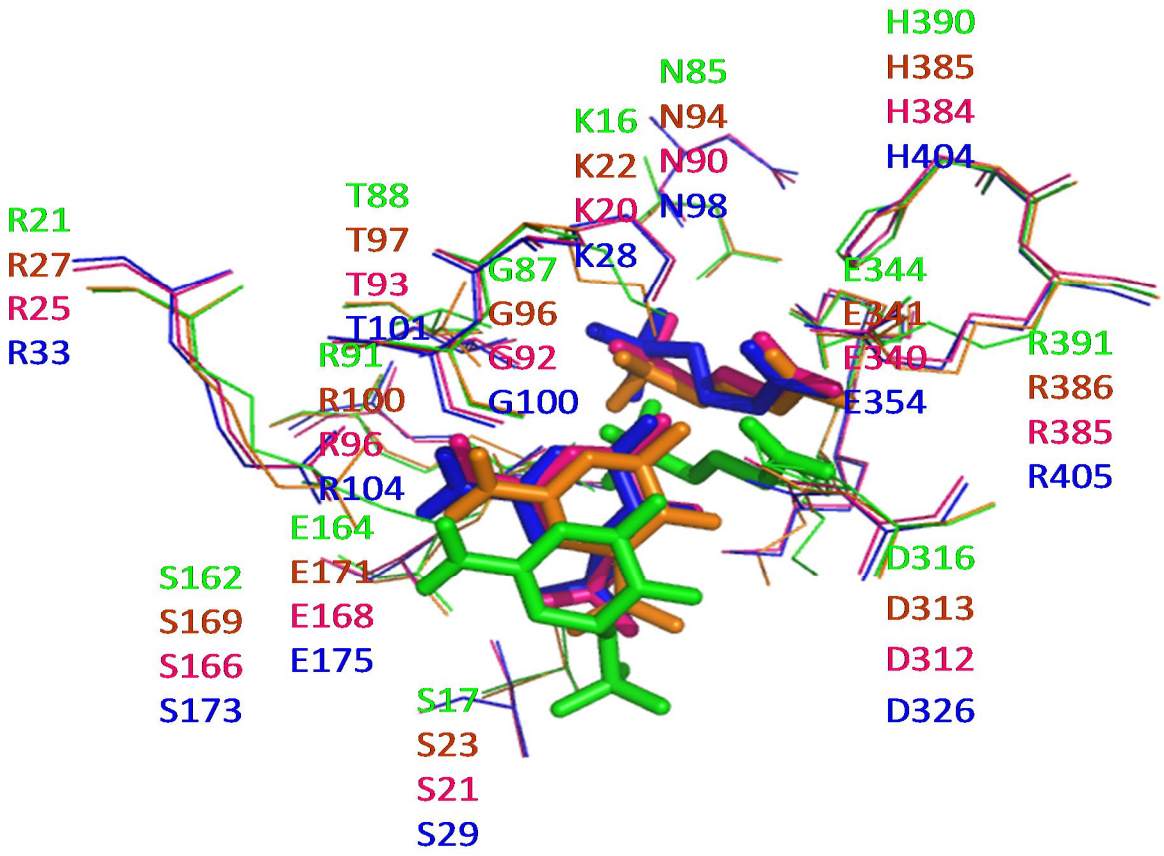


Fig. 4: Conserved residues in the Glyphosate and S3P binding cavity (Colour codes for residues in BpEPSPS: green; EcEPSPS (1G6S): brown; Streptococcus (1RF6): magenta and CP4 (2GGA): blue). The same colour code is used for the Glyphosate and S3P in the centre of the figure.



Fold recognition

The Geno 3D server returned the crystal structure of *Escherichia coli* EPSPS (PDB code: 1G6S, Chain A) as the possible template with the best score. Fold-recognition servers HHPred and FUGUE also reported EcEPSPS (PDB entry: 1G6S) as the best template with highly significant score. On the basis of these results, we can conclude that BpEPSPS has a fold similar to this protein. Therefore, we used this experimental three-dimensional structure as template to predict the 3D structure of BpEPSPS by using comparative modeling strategy.

Secondary structure prediction

A consensus secondary structure was derived for BpEPSPS based on the results from six secondary structure prediction servers (data not shown). Overall, about 33% of residues are in helices and about 19% residues in strands. The conservation of the nature and order of occurrence of secondary structural elements is strongly suggestive of the conservation of the overall fold in this EPSP synthase. It can be inferred from the predicted secondary structures that EPSP synthases belong to the α/β class, as defined in the SCOP database [40].

Multiple sequence alignment

The multiple sequence alignment revealed the conserved residues, the similar residues (data not shown) and these were further corroborated by the superposition of the modeled BpEPSPS with the previously solved crystallographic structures of Escherichia coli (EcEPSPS) (PDB Id: 1G6T),

Streptococcus (SpEPSPS) (PDB Id: 1RF4), *Mycobacterium tuberculosis* (MtEPSPS) (PDB Id: 2BJB) and *Agrobacterium* (CP4 EPSPS) (PDB Id: 2GG4) (obtained from the protein data bank.

Quality of the model

The Ramachandran plot for the EcEPSPS (1G6S) structure was generated in order to compare the overall stereochemical quality of BpEPSPS model against EPSP structures solved by biocrystallography. Analysis of the Ramachandran plot of the BpEPSPS model shows that 92.5% of the residues lie in the most favorable regions, 4% in the allowed region and the remaining 3.5% in the outlier region. The same analysis for the crystallographic EcEPSPS structure presents 97.6% of residues in the most favorable, 2.4% additional allowed regions, and 0.0% in the outlier region (Fig. 2). The overall rating for the BpEPSPS model is slightly poorer than the one obtained for the EcEPSP structure. However, it has over 90% of the residues in the most favorable regions.

Overall description

Fig. 3A shows the secondary structure of the BpEPSPS model. EPSP synthase is a α/β protein. The structure folds into two domains of nearly equal size, each with a radius of about 24.5 Å. Each domain contains β sheets and six parallel α helices aligned in a way that their macrodipoles create a significant electropositive attraction for the anionic ligands at the interface between the two domains [41]. The four-stranded β-sheet structures contain both parallel and antiparallel strands and the helices are parallel.

The amino and carboxyl termini of the protein lie in the same domain and domains are connected via two crossover chain segments. The glyphosate and S3P molecules are bound in the interdomain cleft and could promote the closure of the structure as seen in the case of the previously solved crystallographic structures of EcEPSPS and SpEPSPS [42-43].

When superposed, the EPSPS structures display overall similarity in folds but they differ in several regions (Fig. 3B). Superposition of Cα atoms of the constituent domains of EcEPSPS and BpEPSPS structure results in RMSD of 0.37 Å. The RMSD value for equivalent Cα atoms between SpEPSPS with BpEPSPS was evaluated to be 1.34 Å, between CP4 EPSPS and BpEPSPS was 1.65 Å and that between MtEPSPS and BpEPSPS was 1.53 Å. The low RMSD values, observed in the superpositions, indicate that the alternative conformation of the two-domain structures is due to relative motions of structurally conserved domains.

Active site architecture: 3-Phosphoshikimate and glyphosate binding

The glyphosate and the S3P bind in the cleft in the hinge region and are lined with several conserved and non-conserved amino acid residues (Fig. 3C). EPSPS structures solved till date like SpEPSPS and the CP4 EPSPS exist in an open, unliganded state and a closed, S3P liganded state, suggesting an induced-fit mechanism with binding of S3P as a prerequisite for the enzyme's interaction with PEP [42, 43]. The same may be proposed to hold true for BpEPSPS.

Residues lining the glyphosate binding cavity in BpEPSPS consist of Lys16, Asn85, His390, Arg391, Glu344, Lys343, Asp316, Gln164, Arg91, Gly87, Thr88, Ser241, Arg115, Pro315, Ser162, Ala317, Asp270, Tyr197. The former ten are conserved residues. Residues surrounding the S3P binding site consists of Ser17, Arg21, Val161, Ser162, Ser163, Gln164, Thr167, Leu192, Ile193, Ser194, Lys195, Tyr197, Ile198, Asn312, Pro315, Asp316, Ser339, Lys343. Ser17, Arg21, Ser162, Gln164, Asp316 and Lys343 are conserved residues. The conserved residues are shown in Fig. 4.

In the binary complex with S3P, the active site architecture is highly conserved as in CP4 EPSPS and EcEPSPS enzymes. Strictly conserved residues from both globular domains of the enzyme constitute the PEP-binding site. In the unliganded state, in analogy with the CP4 EPSPS, BpEPSPS could be proposed to contain highly flexible regions, particularly around the strictly conserved residue Glu 344. In the enzyme's binary complex with S3P, this loop becomes structured and interacts with the N-terminal domain to constitute the active site [42]. Also, interdomain interaction in the closed conformation may be stabilized through direct or water-mediated

hydrogen bonds and by salt bridges. Direct hydrogen bonds and salt bridges are observed in the hinge region and the cleft between domains. The majority of these interdomain interactions could be formed by the simple domain approach without significant motion of the participating residues as observed in SpEPSPS [43].

CONCLUSION

Shikimate pathway enzymes undoubtedly offer important targets for the development of drugs against bacterial, fungal and apicomplexan parasitic diseases. Each of the enzymatic steps is of potential interest and the availability of structures for the pathway enzymes will enable rapid strides in the drug development scenario. It is hoped that the molecular structure of BpEPSPS predicted in this work should guide biochemical experiments as well as provide insights on designing antimicrobial agents against the whooping cough disease. The description of the active site can be used for virtual screening or even *de novo* ligand design. While designing novel antimicrobial agents, it could be useful to construct molecules that block the binding of the substrates important for the function of the enzyme.

ACKNOWLEDGEMENT

The work was supported by funding from the All India Council for Technical Education National Doctoral Fellowship (AICTE-NDF). The author is grateful to the facilities in the Bioinformatics Lab, Indian Institute of Technology, Roorkee, India. The work has no bearing with Novo Nordisk.

REFERENCES

1. K.J. Ryan, C.G. Ray, Sherris Medical Microbiology, fourth ed., McGraw Hill, 2004.
2. http://en.wikipedia.org/w/index.php?title=Bordetella_pertussis&oldid=188342151
3. R. H. Major**,** A history of medicine, Charles C. Thomas, Springfield, Ill, 1954. vol. 1, p. 423.
4. De Melker, H. E., J. F. P. Schellekens, S. E. Neppelenbroek, F. R. Mooi, H. C. Rumke, M. A. E. Conynvan Spaendonck, Reemergence of pertussis in the highly vaccinated population of The Netherlands: observations on surveillance data, Emerg. Infect. Dis. 6 (2000) 348–357.
5. D.Gu¨ris, P. M. Strebel, B. Bardenheier, M. Brennan, R. Tachdjian, E. Finch, M. Wharton, J. R. Livengood, Changing epidemiology of pertussis in the United States: increasing reported incidence among adolescents and adults, 1990–1996. Clin. Infect. Dis. 28 (1999) 1230–1237.
6. E. K. Korgenski, J. A. Daly, Surveillance and detection of erythromycin resistance in *Bordetella pertussis* isolates recovered from a pediatric population in the Intermountain West Region of the United States, J. Clin. Microbiol. 35 (1997) 2989–2991.
7. K.Lewis, M. A. Saubolle, F. C. Tenover, M. F. Rudinsky, S. D. Barbour, J. D. Cherry**,** Pertussis caused by an erythromycin-resistant strain of *Bordetella pertussis*. Pediatr. Infect. Dis. J. 14 (1995) 388–391.
8. K.A. Brown, E.P. Carpenter, K.A. Watson, J.R. Coggins, A.R. Hawkins, M.H.J. Koch, D.I. Svergun, Enzyme Mechanism – A Structural Perspective Twists and turns: a tale of two shikimate-pathway enzymes, Biochem. Soc. Trans. 31 (2003) 3
9. J.R. Coggins, C. Abell, L.B. Evans, M. Frederickson, D.A. Robinson, A.W. Roszak, A.P. Lapthorn, Experiences with the shikimate-pathway enzymes as targets for rational drug design, Biochem. Soc. Trans. 31 (2003) 548–552.
10. A.J. Pittard, in: Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology (Niedhardt, F.C., ed.), pp. 1, American Society of Microbiology, Washington DC, 1996, pp. 458–484
11. F. Roberts, C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J. Ferguson, Ferguson, D. Chakrabarti, R. McLeod, Evidence for the shikimate pathway in apicomplexan parasites, Nature (London) 395 (6699) (1998) 801–805.
12. H.M. Berman, J.Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H.Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucl. Acids Res. 28 (2000) 235–242.
13. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J Mol Biol 215 (1990) 403-410.
14. S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucl. Acids. Res. 25 (1997) 3389-3402.
15. L. Jaroszewski, L. Rychlewski, Z. Li, W. Li, A. Godzik, FFAS03: a server for profile-profile sequence alignments, Nucl. Acids. Res. 33 (2005) W284-W288.
16. J. Shi, T.L .Blundell, K. Mizuguchi, FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure- dependent gap penalties, J Mol Biol, 310 (2001) 243-257.

17. L.A. Kelley, R.M. MacCallum, M.J.E. Sternberg, Enhanced Genome Annotation using Structural Profiles in the Program 3DPSSM, J. Mol. Biol. 299 (2000) 499-520.

18. K Karplus, R Karchin, J Draper, J Casper, Y Mandel-Gutfreund, M Diekhans, R Hughey, Combining local-structure, fold-recognition, and new-fold methods for protein structure prediction, Proteins, 53 (2003) 491-496.

19. M.A. Kurowski, J.M. Bujnicki, GeneSilico protein structure prediction meta-server. Nucl. Acids. Res. 31 (2003) 3305-3307.

20. A. Krogh, B. Larsson, G. von Heijne, E.L.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567-580.

21. K. Karplus, S. Katzman, G. Shackleford, M. Koeva, J. Draper, B. Barnes, M. Soriano, R Hughey, SAM-T04: what's new in protein-structure prediction for CASP6, Prosfb 61 (2005) 135-142.

22. J.A. Cuff, G.J. Barton, Application of enhanced multiple sequence alignment profiles to improve protein secondary structure prediction, Proteins 40 (2000) 502-511.

23. D.G. Kneller, F.E. Cohen, R. Langridge, Improvements in protein secondary structure prediction by an enhanced neural network, J. Mol. Biol. 214 (1990) 171-182.

24. L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (2000) 404-445.

25. C. Geourjon, G. Deleage, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Appl.Biosci. 11 (1995) 681-684.

26. G. Pollastri, D. Przybylski, B. Rost, P. Baldi, Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles, Proteins 47 (2002) 228-235.

27. C. Notredame, D. Higgins, J. Heringa, T-Coffee: a novel method for multiple sequence alignments. J. Mol. Biol., 302 (2000) 205-217.

28. O. Poirot, K. Suhre, C. Abergel, E. O'Toole, C. Notredame, 3DCoffee: a web server for mixing sequences and structures into multiple sequence alignments, Nucl. Acids. Res. 32 (2004) W37-40.

29. A. Bairoch, R. Apweiler, The SWISS-PROT protein sequence database: its relevance to human molecular medical research, J. Mol. Med. 75 (1997) 312–316.

30. A. Marchler-Bauer, S.H. Bryant, CD-Search: protein domain annotations on the fly, Nucl. Acids. Res. 32 (2004) W327-W331.

31. P.Gouet, E.Courcelle, D.I. Stuart, F.Metoz, ESPript: multiple sequence alignments in PostScript, Bioinformatics. 15 (1999) 305-308.

32. R.T. Kroemer, S.W. Doughty, A.J. Robinson, W.G. Richards, Prediction of the three-dimensional structure of human interleukin- 7 by homology modeling, Protein Eng. 9 (6) (1996) 493-498.

33. T.L. Blundell, B.L. Sibanda, M.J. Sternberg, J.M. Thornton, Knowledge-based prediction of protein structures and the design of novel molecules, Nature 326 (6111) (1987) 347–352.

34. T.L. Blundell, D. Carney, S. Gardner, F. Hayes, B. Howlin, T. Hubbard, J. Overington, D.A. Singh, B.L. Sibanda, M. SutCliffe, 18th Krebs, Hans lecture knowledge-based protein modeling and design, Eur. J. Biochem. 172 (3) (1988) 513–520.

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

36. A. Sali, J.P. Overington, Derivation of rules for comparative protein modeling from a database of protein structure alignments, Protein Sci. 3 (9) (1994) 1582–1596.

37. A. Sali, L. Potterton, F. Yuan, H. van Vlijmen, M. Karplus, Evaluation of comparative protein modeling by MODELLER, Proteins 23 (3) (1995) 318–326.

38. A. Sali, Modeling mutations and homologous proteins, Curr. Opin. Biotechnol. 6 (4) (1995) 437–451.

39. S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson and D.C. Richardson Structure validation by Calpha geometry: phi,psi and Cbeta deviation, Proteins: Structure, Function & Genetics 50 (2002) 437-450.

40. Murzin AG, Brenner SE, Hubbard T, Chothia C: SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol, 247 (1995) 536-540.

41. W.C. Stallings, S.S. Abdel-Meguid, L.W. Lim, H.-S. Shieh, H.E. Dayringer, N.K. Leimgruber, R.A. Stegeman, K.S. Anderson, J.A. Sikorski, S.R. Padgette, G.M. Kishore, Structure and topological symmetry of the glyphosate target 5-enol-pyruvylshikimate-3-phosphate synthase: a distinctive protein fold, Proc. Natl. Acad. Sci. USA 88 (1991) 5046–5050.

42. T. Funke, H. Han, L. Martha, Healy-Fried, M. Fischer, and E. Scho¨ nbrunn Molecular basis for the herbicide resistance of Roundup Ready crops, PNAS. 103 (35) (2006) 13010–13015.

43. H. Park, J. L. Hilsenbeck, H. J. Kim, W. A.Shuttleworth, Y. H. Park, J. N. Evans, C. Kang, Structural studies of *Streptococcus pneumoniae* EPSP synthase in unliganded state, tetrahedral intermediate-bound state and S3P-GLP-bound state, Mol. Microbiol. 51 (2004) 963–971.