Research & Reviews

ISSN: 2320-3528

Research and Reviews: Journal of Microbiology and Biotechnology

Homology Modeling of Chorismate Synthase from *Brucella melitensis*: A Novel Target Molecule.

Maskar AU[1]* and Meshram RJ[2]

[1]Institute of Life Sciences, University of Hyderabad Campus, Gachibowli, Hyderabad – 500046, Andhra Pradesh India.
[2]School of Life Sciences, Swami Ramanand Teerth Marathwada University, Nanded, Maharashtra, India.

Research Article

Received: 02/06/2013
Revised: 15/06/2013
Accepted: 26/06/2013

***For Correspondence**

Institute of Life Sciences, University of Hyderabad Campus, Gachibowli, Hyderabad – 500046, Andhra Pradesh India.

Keywords: homology, chorismate synthase, *brucella melitensis*

ABSTRACT

Brucellosis is one of the most common zoonotic diseases. Human Brucellosis is primarily caused by *B. melitensis*transmitted by goats and sheep along with other potential routes for the spread of infection. Till date no definite therapy is available for the permanent treatment. Also vaccines, that are available, are not effective or non-specific to stop the spread of the infection. So there is a sure call for finding newer specific target molecules to design the newer specific therapeutic agent/s. After comparative metabolomic investigation in *B. melitensis*, we found enzyme Chorismate Synthase as the probable target molecule in Human Brucellosis. The present work includes generation of homology model of Chorismate Synthase for *B. melitensis*using template from *H. pylori*. The model was developed using satisfaction of spatial restraints approach implemented in Modeller 9v3 software. The model generated is a tetramer and consists of β-α-β Sandwich Fold in each monomer which is a signature fold of this enzyme family. After thorough structural evaluations, the model is found to be quiet satisfactory and can surely be useful for further study to find out the potential therapeutic agent/s for Human Brucellosis. The model is available at Protein Model Data Base (PMDB ID PM0078183).

INTRODUCTION

Human Brucellosis

Brucellosis (Mediterranean/Malta/Undulant Fever) is one of the most important and most common zoonosis in the world affecting livestock and humans characterized by wave-like variations in the body temperature of afflicted victims. Its modern name bears tribute to Sir David Bruce, the military physician who discovered the etiologic agent [1-3]. A direct link with economic status has been a hallmark of the disease [4].

Causative Agent and Mode of Transmission

Brucellosis is caused by members of the bacterial genus Brucella. These are facultative intracellular Gram-negative pathogens. *Brucella melitensis* is present in and transmitted by goats and sheep and related animals and is most virulent for man. *B. abortus* is the dominant species in cattle and *B. suis* is mainly confined to pigs [5]. These are categorized as biological agents and were designated as Select Agents of Category B by the Centre for Disease Control in Atlanta, USA [1, 2]. The transmission of Brucella infection and its prevalence in a region depends upon several factors like food habits, methods of processing milk and milk products, social customs, husbandry practices, climatic conditions, socioeconomic status and environment hygiene. Brucellosis is almost invariably transmitted to man from infected domestic animals. Human brucellosis is found to have significant presence in rural/nomadic communities where people live in close association with animals [2].

ISSN: 2320-3528

Indian Scenario

Worldwide, reported incidence of human brucellosis in endemic disease areas varies widely, from <0.01 to >200 per 100,000 population. The true incidence of human brucellosis however, is unknown for most countries including India. In India 80% of the population live in villages and have close contact with animals owing to their occupation and hence there is a greater risk of acquiring brucellosis. The disease has an added importance in countries like India, where conditions are conducive for wide-spread human infection on account of unhygienic conditions and poverty. Several published reports including recent ones indicate that human brucellosis is quiet common disease in India [2]. Brucellosis in human occurs in all age groups and both males and females are affected equally. Brucellosis in children can be very common in particular in areas with *B. melitensis*. Infection also may occur through cuts and abrasions of the skin, via the conjunctiva and by inhalation [5].

Clinical Manifestations

Brucellosis in animals causes tremendous economic losses due to abortion, premature birth, decreased milk production and reduced reproduction rate [1]. Human brucellosis is known for protean manifestations. However, the most common presenting symptom is fever. The most common signs and symptoms are listed in Table 1. Epididymoorchitis is the most frequent genitourinary complication in men. Brucellosis during pregnancy poses a substantial risk of spontaneous abortion or intrauterine transmission of infection to the infant [2]. Thus brucellosis is a multi-systemic disease that shows wide clinical polymorphism (Table 2) [6, 7]. Osteoarticular, gastrointestinal, hepatobiliary, respiratory, genitourinary complications are observed. Bone and joint complications are the most frequent complications. Meningitis and meningo-encephalitis are the most common complications seen in neurobrucellosis. Complications of the cardiovascular system are rare but important as they have a high degree of mortality. Other examples of rare complications are those of skin and eyes [5].

Table 1: Common Signs and Symptoms of Human Brucellosis

Anorexia	Malaise
Back Pain	Myalgia
Cephalgia	Sweats
Fatigue	Weight Loss
Fever	Prostration
Chills	Headache
Arthritis	Loss of Appetite

Table 2: Clinical Manifestations of Human Brucellosis

Anemia	Nephritis
Deep Vein Thrombosis	Optic Neuritis
Endocarditis	Pancytopenia
Hepatomegaly	Papilledema
Leukocytoclastic Vasculitis	Splenic Abscess
Leukopenia	Splenomegaly
Liver Abscess	Spondylitis
Lymphadenopathy	Thrombocytopenia
Meningitis	Uveitis

Available Treatment

Currently, antibiotic treatment of animals is systematically discouraged and highly regulated. The limited number of effective antibiotics and the potential for accidental or malicious introduction of antibiotic resistances into the organism emphasizes the need for alternative solutions [4]. Although vaccination is probably the most economic control measure, administration of currently available vaccines alone is not sufficient for elimination of brucellosis in any host species [8]. A limited number of antibiotics, in humans, are effective against these organisms. The side-effects of drug combination schemes and the high incidence of relapses and therapeutic failures, have led to the investigation of new drugs to treat the disease [6]. The most effective, least toxic chemotherapy for human brucellosis is still undecided [9]. No vaccine is available for the prevention of brucellosis in human [5].

Hence, our present work includes finding out the novel specific target molecule in *B. melitensis* which can effectively treat human brucellosis in human.

ISSN: 2320-3528

MATERIALS AND METHODS

Comparative Metabolomic Study

Rahnuma is a tool for prediction and analysis of metabolic pathways and comparison of metabolic networks [10]. The Comparative Analysis of Type Standard Comparison for Organism Mode on Full Network Comparison was run for human (*Homo sapiens*, hsa) and brucellae (*B. melitensis*, bmi). The output tabular text file contained many reactions which can be classified as reactions unique to human, reactions present in brucellae alone and reactions common to both organisms. As our major objective was to find target specific to brucellae, reactions unique to brucellae were selected for further study.

Target Identification

The metabolic pathway analysis in KEGG Pathway Database [11] was carried out for two reasons, first being to make sure that chorismate biosynthesis is significantly different or absent in human and second is to confirm that there is no alternate pathway for chorismate biosynthesis in *B. melitensis*. In study, we observed the reaction which is catalyzed by enzyme chorismate synthase leading to formation of chorismic acid (KEGG Reaction ID R01714). This reaction is uniquely present in brucellae and not in human, making it an attractive therapeutic target molecule. Further Choke Point Analysis was carried out by using MetExplore Database [12].

Modeling of Chorismate Synthase

The amino acid sequence of chorismate synthase from *B. melitensis* was retrieved from UniProtKB/Swiss-Prot [13] with Accession Number P63607. This was subjected to BLAST-P [14, 15] using PDB Database [16] utilizing BLAST Server of NCBI. Further steps in modeling protocol were carried out using Modeller 9v3 which is a computer program that models three-dimensional structures of proteins and their assemblies by satisfaction of spatial restraints. 3D models were obtained by optimization of a molecular probability density function (pdf) [17].

Model Evaluation

The model produced by Modeller 9v3 was firstly evaluated using Structural Superimposition [18] followed by running standard evaluation tests like QMEAN Score [19], ProCheck [20], ProSA [21] and Verify3D [22]. The resulting model was visualized in PyMol v0.99 (W. L. DeLano *et al*) and was submitted in Protein Model Data Base [23] (PMDB ID PM0078183).

RESULTS AND DISCUSSION

Recent progress in genomics and proteomics has led to generation of enormous proteomic as well as biochemical data. The metabolic information specific to organism is now available in KEGG Database. Comparison of entire metaboloms is now possible using tools like Rahnuma. Comparative Metabolomic Approach [24] can be used to identify newer drug targets. There must be three main properties that any therapeutic target should have. Firstly, it must be unique to the pathogen so that its inactivation will only hamper the physiology of pathogen, not the host. Secondly, it must play an important role in the pathogen metabolism so that its inactivation will lead to the hazardous effects in the pathogen only. Thirdly, close homolog of target protein must not be present in the host, as the tendency of homologs is to share the common structural architecture. The chemical agent designed to modulate target's activity might interfere with functioning of its homologous target molecule in the host resulting sometimes in the severe adverse effects.

Comparative Metabolomic Analysis

The output file from Rahnuma from our analysis showed total of 1932 reactions out of which 1114 (58%) reactions were unique to human, 387 (20%) reactions were present in brucellae only and remaining 431 (22%) reactions were common to both human and brucellae (Figure 1).

Target Molecule

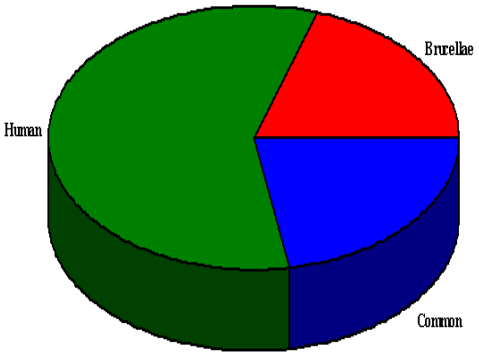
The Shikimate Pathway was discovered as the biosynthetic route to the aromatic amino acids phenylalanine, tyrosine and tryptophan. This pathway has been proved to be found only in microorganisms and plants. The seventh and last step of the Shikimate Pathway is the concerted lP-Trans Elimination of Phosphate from 5-enolpyruvylshikimate-3-phosphate (EPSP) to yield chorismate. Chorismate synthase with reduced flavin catalyzes this reaction (Figure 2) [25]. In this way our study indicates that, this reaction is uniquely present in the pathogen. In addition, the chorismate pool in the cell is necessary for the synthesis of other aromatic compounds such as

ISSN: 2320-3528

vitamin K, ubiquinone or p-aminobenzoate [26]. These facts suggest that chorismate synthase has the vital role in the pathogen physiology.

The Human BLAST result showed no significant sequence identity/similarity towards the target molecule. The hits obtained possessed very poor query coverage and had very high e-value. This clearly validate that there is no close homolog of chorismate synthase exists in human. Also the choke point analysis confirmed that chorismate synthase catalyzes a choke point reaction in *B. melitensis*. A choke point reaction is defined as a reaction that either uniquely consumes a specific substrate or uniquely produces a specific product. The inactivation of such reactions results in either the toxicity due to accumulation or the starvation owing to depletion of a particular metabolite in pathogen, thus increasing its importance in the drug designing.

Figure 1: Pie Chart Summarizing Comparative Metabolomic Reactions



(Human: Green, Brucellae: Red and Both Human and Brucellae: Blue)

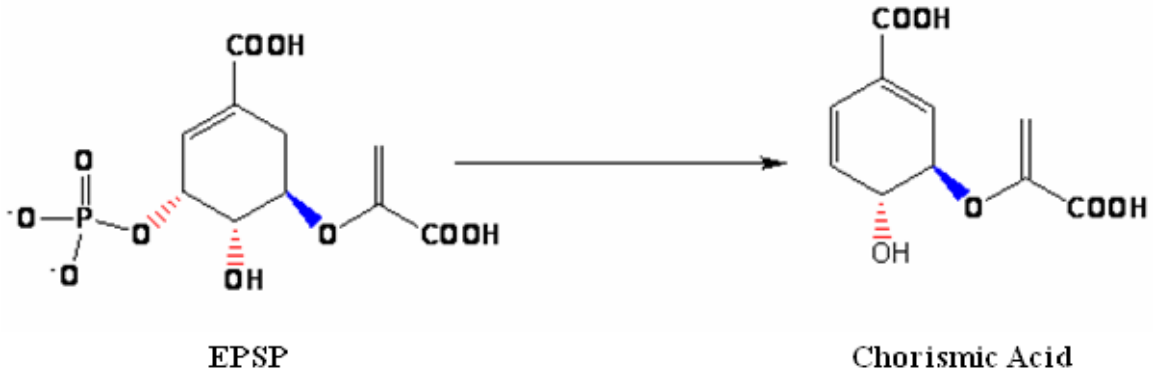


Figure 2: Reaction Catalyzed by Chorismate Synthase in Shikimate Pathway

Template Identification and Sequence Analysis

BLAST-P search performed aiming to identify suitable template for chorismate synthase from *B. melitensis* and yielded seven hits with comparable e-value and bit score. The most significant was the structure of chorismate synthase complexed with FMN from *Helicobacter pylori* (PDB ID 1UM0) having 96% query coverage and sharing 41% sequence identity with the query sequence. The template identification was also confirmed by Geno3D Server [27, 28]. The sequence alignment of chorismate synthase from *B. melitensis* and *H. pylori* indicate the presence of two inserts ^{75}EDGBTM^{80} and ^{273}DGKPI^{277} which are inductive of acquisition of new motifs in evolutionary path of chorismate synthase from *B. melitensis.* Focusing on FMN binding-site, it seems to be highly conserved across both the sequences (Figure 3). Three signature motifs are essentially present in all members of the family of chorismate synthase [29] and also found to be fairly conserved in chorismate synthase of *B. melitensis*, however structural and functional significance of these motifs is yet to be discovered. These motifs are found to be present conserved across all the six sequences of chorismate synthase in *H. pylori* (Figure 4). These three motifs are present in the form of ^{16}GESHGLALGCVVDGCP^{31}, ^{129}GRSSARETAARVAAGAI^{145} and ^{329}RHDPCVGIRAVPIGEAM^{345} in chorismate synthase of *B. melitensis*.

ISSN: 2320-3528

```
                          20                  40                  60                  80
Template:  --MNTLGRFLRLTTFGESHGDVIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFITPRKEDDKVEITSGVFED------
Query    : MSHNSFGHLFRVTTWGESHGLALGCVVDGCPPGITFTEAEIQSFLDKRKPGQSKYTTQRREPDQVRVLSGVLLGEDGVTM

                         100                 120                 140                 160
Template:  FSTGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLREIGIVCES
Query    : TTTGTPISMMIENTDQRSKDYGEIARQYRPGHADYAYDVKYGIRDYRGGGRSSARETAARVAAGAIARKVVPGLEVRGAL

                         180                 200                 220                 240
Template:  GIIEIGGIKAKNYDFNHALKSEIFALDEEQEEAQKTAIQNAIKNHDSIGGVALIRARSIKTNQKLPIGLGQGLYAKLDAK
Query    : VSIGAHDIDRSRWNWAEVDNNPFFTPDAGSVEVFADYLDGIRKNGSSVGAIIEIVA------EGVPAGIGAPIYGKLDQD

                         260                 280                 300                 320
Template:  IAEAMMGLNGVKAVEIGKGVESSLLKGSEYNDLMDQKG-----FLSNRSGGVLGGMSNGEEIIVRVHFKPTPSIFQPQRT
Query    : IASYLMSINAVKGVEIGNGFEAARLTGEENADEMRMGNDGKPIFLSNHAGGVLGGIATGAPVVARFAVKPTSSILTPRRS

                         340                 360                 380
Template:  IDINGNECECLLKGRHDPCIAIRGSVVCESLLALVLADMVLLN--LTSKIEYLKTIYNEN
Query    : IDKDGNEVDVMTRGRHDPCVGIRAVPIGEAMVACAIADHYLRHRGQTGRV----------
```

Figure 3: Pair-Wise Sequence Alignment of Active Site in Chorismate Synthase of *H. pylori* (Template) and *B. melitensis* (Query) with Conserved Residues in Chain A (Red) and D (Magenta)

Figure 4: Three Signature Motifs (Highlighted in Yellow) in Chorismate Synthase Conserved Across All Six Species of *H. pylori* (Hp) and *B. melitensis* (Bm) (Species of *H. pylori* shown are 700392, Shi470, HPAG1, P12, G27 and J99)

```
                          20                  40                  60                  80
Hp_700392:  --MNTLGRFLRLTTFGESHGDVIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFITPRKEDDKVEITSGVFED------
Hp_Shi470:  --MNTLGRFLRLTTFGESHGDVIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFITPRKEDDKVEITSGVFEG------
Hp_HPAG1 :  --MNTLGCFLRLTTFGESHGDMIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFITPRKEDDKVEITSGVFED------
Hp_P12   :  --MNTLGRFLRLTTFGESHGDVIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFITPRKEDDKVEITSGVFED------
Hp_G27   :  --MNTLGRFLRLTTFGESHGDVIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFTTPRKEDDKVEITSGVFED------
Hp_J99   :  --MNTLGRFLRLTTFGESHGDMIGGVLDGMPSGIKIDYDLLENEMKRRQGGRNIFTTPRKEDDKVEITSGVFEG------
Bm       : MSHNSFGHLFRVTTWGESHGLALGCVVDGCPPGITFTEAEIQSFLDKRKPGQSKYTTQRREPDQVRVLSGVLLGEDGVTM

                         100                 120                 140                 160
Hp_700392:  FSTGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLREIGIVCES
Hp_Shi470:  FSTGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLREIGIVCES
Hp_HPAG1 :  FSTGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLREIGIVCES
Hp_P12   :  FSTGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLREIGIVCES
Hp_G27   :  FSTGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLKEIGIVCES
Hp_J99   :  FSAGTPIGFLIHNQRARSKDYDNIKNLFRPSHADFTYFHKYGIRDFRGGGRSSARESAIRVAAGAFAKMLLREIGVVCES
Bm       :  TTTGTPISMMIENTDQRSKDYGEIARQYRPGHADYAYDVKYGIRDYRGGGRSSARETAARVAAGAIARKVVPGLEVRGAL

                         180                 200                 220                 240
Hp_700392:  GIIEIGGIKAKNYDFNHALKSEIFALDEEQEEAQKTAIQNAIKNHDSIGGVALIRARSIKTNQKLPIGLGQGLYAKLDAK
Hp_Shi470:  GIIEIGGIKAKNYDFNHALKSEIFALDEEQEEVQKTAIQNAIKNHDSIGGVALIRARSAKTNQKLPIGLGQGLYAKLDAK
HpHPAG1  :  GIIEIGGIEAKNYDFNHALKSEIFALDKEQEEAQKTAIQNAIKNHDSIGGVALIRARSVKRNQKLPIGLGQGLYAKLDAK
Hp_P12   :  GIIKIGGIEAKNYDFNHALKSEIFALDKEQEEAQKTAIQNAIKNHDSIGGVALIRARSAKTNQKLPIGLGQGLYAKLDAK
Hp_G27   :  GIIKIGGIEAKNYDFNHALKSEIFALDEEQEEAQKTAIQNAIKNHDSIGGVALIRARSIKTNQKLPIGLGQGLYAKLDAK
Hp_J99   :  GIIKIGGIEAKNYDFNHALKSEIFALDKEQEEAQKTAIQNAIKNHDSIGGVALIRARGAKTNQKLPIGLGQGLYAKLDAK
Bm       :  VSIGAHDIDRSRWNWAEVDNNPFFTPDAGSVEVFADYLDGIRKNGSSVGAIIEIVAEGV------PAGIGAPIYGKLDQD

                         260                 280                 300                 320
Hp_700392:  IAEAMMGLNGVKAVEIGKGVESSLLKGSEYNDLMDQKG-----FLSNRSGGVLGGMSNGEEIIVRVHFKPTPSIFQPQRT
Hp_Shi470:  IAEAMIGLNGVKAVEIGKGVESSLLKGSEYNDLMDQKG-----FLSNRSGGVLGGMSNGEEIIIRAHFKPTPSIFQPQQT
Hp_HPAG1 :  IAEAMMGLNGVKAVEIGKGVESSLLKGSEYNDLMDQKG-----FLSNRSGGVLGGMSNGEEIIVRVHFKPTPSIFQPQQT
Hp_P12   :  IAEAMMGFNGVKAVEIGKGVESSLLKGSEYNDLMNQKG-----FLSNRSGGVLGGMSNGEEIIVRVHFKPTPSIFQPQQT
Hp_G27   :  IAEAMMGLNGVKAVEIGKGVESSLLKGSEYNDLMNQKG-----FLSNHSGGVLGGMSNGEEIIVKAHFKPTPSIFQPQQT
Hp_J99   :  IAEAMMGLNGVKAVEIGKGVESSLLKGSEYNDLMSQKG-----FLSNHSGGVLGGMSNGEEIIVKVHFKPTPSIFQPQRT
Bm       :  IASYLMSINAVKGVEIGNGFEAARLTGEENADEMRMGNDGKPIFLSNHAGGVLGGIATGAPVVARFAVKPTSSILTPRRS

                         340                 360                 380
Hp_700392:  IDINGNECECLLKGRHDPCIAIRGSVVCESLLALVLADMVLLN--LTSKIEYLKTIYNEN
Hp_Shi470:  IDINNQECECLLKGRHDPCIAIRGSVVCESLLSLVLADMVLLN--LTSKIEYLKTIYNEN
Hp_HPAG1 :  IDINNHECECLLKGRHDPCIAIRGSVVCESLLSLVLADMVLLN--LTSKIEYLKTIYNEN
Hp_P12   :  IDINGNECECLLKGRHDPCIAIRGSVVCESLLSLVLADMVLLN--LTSKIEYLKTIYNEN
Hp_G27   :  IDINNNECECLLKGRHDPCIAIRGSVVCESLLSLVLADMVLLN--LTSKIEYLKTIYNEN
Hp_J99   :  IDIKGNECECLLKGRHDPCIAIRGSVVCESLLSLVLADMVLLN--LTSKIEYLKTIYNEN
Bm       :  IDKDGNEVDVMTRGRHDPCVGIRAVPIGEAMVACAIADHYLRHRGQTGRV----------
```

We have also observed sequentially conserved loops in our enzyme, of which four are present in all the subunits, while remaining two are confined to B and C subunits of chorismate synthase [29]. These regions correspond to ^{15}WGESHGLAL^{23}, ^{50}RKPGQSKYTTQRREPDQVR^{68}, ^{91}IENTDQRSKDYGEIAR^{106}, ^{122}GIRDYRGGGRSSARE^{136} ^{262}EENADEMRMGNDGKPI^{277} and ^{308}ILTPRRSIDKDGNEVDVMTRGRHD^{331} further will be referred as loops L1-L6. Except L5 loop, all other five loops are present around the FMN binding site (Figure 5). All these loops contain highly conserved amino acid residues around the active site.

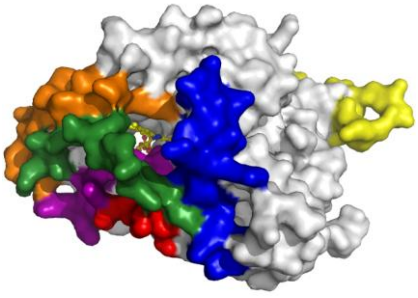


Figure 5: Loops Present in Particular Monomer of Chorismate Synthase from *B. melitensis* Around FMN (L1: Red, L2: Blue, L3: Green, L4: Magenta, L5: Yellow and L6: Orange)

ISSN: 2320-3528

Template Validation

The reliability of a homology model is always dependent on the quality of structural models used. Therefore, it is very essential to make sure that the template used in the study is of the good quality. B-factor (Atomic Displacement/Temperature Factors) gives information on the mobility of each of the atoms in the structure. This B-factor reflects the degree of thermal motion and static disorder of an atom in a protein crystal structure [30]. The good model should have mean B-factor value below 60 Å^2. In the template, both Chains A and D have considerably lower average B-factors, as 36.9 and 36.8 Å^2 respectively, than the other two Chains B and C, as 49.3 and 53.6 Å^2 respectively [29]. The crystallographic R-factor is the most widely used parameter to measure the disagreement between the amplitudes of the experimentally measured reflections and the protein model. It estimates the average coordinate error in a protein crystal structure. For protein structures solved and refined at better than 2 Å resolution, the average coordinate error is in the range 0.2-0.3 Å [31]. In case of 1UM0, value of both crystallographic terms, mean B-factor (44.70 Å) and R-Value (0.24 Å), falls well inside the threshold.

Considering the above mentioned facts, we decided to select chorismate synthase from *H. pylori* as the template in our study.

Homology Model Generation

Modeller 9v3 was allowed to generate five loop models from each of the four monomer templates Chain A, B, C and D (Table 3). We selected the best out of five generated loop models from each monomer template based on its molecular probability density function (pdf) values. Generally, models having lowest molecular pdf values are preferred. Hence, Model_2, Model_5, Model_3 and Model_4 were selected as apo forms from templates Chain A, B, C and D, respectively.

Later on hetero-atoms like FMN and water molecules were merged into the individual monomer models by using DeepView / SPdbV 3.7 [32] to generate holo forms of individual monomers. The quaternary structure (tetramer) was further generated by again merging these four holo forms of individual monomers by using same method described above (Figure 6).

Table 3: Summary of Successfully Produced Loop Models

Template	Model	PDF	Template	Model	PDF
Chain A	1	1123	Chain C	1	108
	2	127 [a]		2	77
	3	2376		3	31 [a]
	4	919		4	58
	5	2878		5	96
Chain B	1	2544	Chain D	1	2557
	2	1615		2	141
	3	1432		3	1154
	4	4222		4	98 [a]
	5	54 [a]		5	122

[a]Lowest PDF Values Among Five Models for Each Monomer

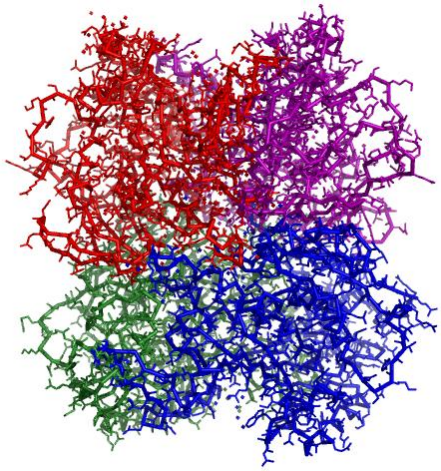


Figure 6: Tetramer Model of Chorismate Synthase from *B. melitensis* (Chain A: Red, B: Blue, C: Green and D: Magenta)

Model Evaluation

Structural Superimposition

In the same way that sequence comparisons can provide tremendous insight into the origins, function, location, interactions and activity of a given protein, so too can structure comparison. In fact, because structure is actually much more conserved than sequence, structure comparisons allow us to look even further back into biological prehistory. The most common method for three-dimensional (3D) structure comparison is called structure superposition. The superposition or superimposition is simply the process of rotating or orienting an object until it can be directly overlaid on top of a similar object. The root mean square deviation (RMSD) value denotes the degree of superimposition. Generally lower the value more is the degree of superimposition and closer are the structures. The generated model shows RMSD value of 0.5 Å for backbone atoms than the template using DeepView / Swiss-PDB Viewer 3.7 application. These figures illustrate that model generated in this study are very close to the experimentally verified template structure.

QMEAN Score

The QMEAN (Qualitative Model Energy ANalysis) is a composite scoring function that estimates the global quality of the models on the basis of a linear combination of six structural descriptors out of which four are statistical potentials of mean force. The score reflects the predicted global model reliability ranging from 0 to 1. The QMEAN Score of the model was predicted as 0.691 which is comparable to that of the template score as 0.762.

ProCheck

The ProCheck assess the stereo-chemical quality of a given protein structure. The analysis of the model shows the presence of 98.8% residues in the favorable region and total quality of G-factors is found to be -0.2 indicating its good stereo-chemical quality (Table 4) (acceptable G-factor values are between 0 and -0.5).

Table 4: Comparison of ProCheck Analysis of Template and Model

Residues in	Ramachandran Plot Region (%)			
	Most Favored	Additional Allowed	Generously Allowed	Disallowed
Template	83.0	14.0	1.6	1.4
Model	86.7	10.5	1.7	1.2
	Goodness (G) Factor			
	Dihedral Angle	Covalent Force	Overall Average	
Template	0.14	0.58	0.32	
Model	-0.15	-0.31	-0.20	

Verify 3D

Verify3D Structure Evaluation Server analyzes the compatibility of an atomic model (3D) with its own amino acid sequence (1D). The score ranges from -1 (bad score) to +1 (good score). After analyzing the Verify 3D data, 348 (95.6%) of 364 residues in the model were found to have the score above 0.2 (Figure 7). The active site residues have the score range from 0.3 to 0.6 indicating that the active site residues are attuned with the model 3D structure. The residues having score below 0.2 are ^{259}GFEAARLTGEENADEM^{274} which do not belong even near to the active site.

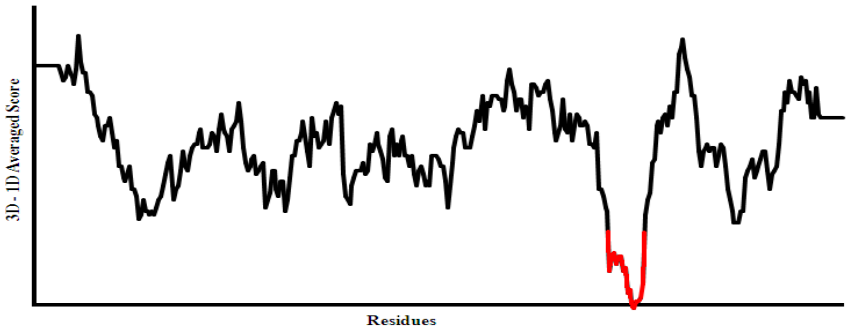


Figure 7: Verify 3D Plot for Residues with Score Above (Black) and Below (Red) Cut-Off Value in A Particular Monomer of Model

ProSA is a tool widely used to check 3D models of protein structures for potential errors. The ProSA Energy Plot for a monomer of model (Figure 8) showed negative values for all residues indicating the absence of any error in a model.

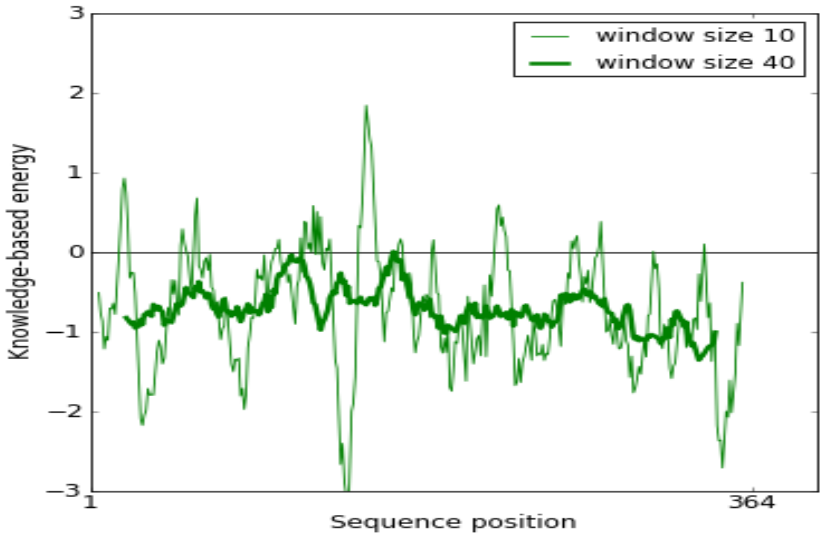


Figure 8: ProSA Energy Plot of Model for Chain A

Secondary Structure Analysis

The secondary structural analysis and interface interactions in monomers were done by using PDBsum Web Server [33]. Each monomer of model consists of nine α-helices (except monomers B and D, eight and ten respectively) and thirteen β-strands (except monomers C and D, fifteen and twelve respectively). The detailed structural information of secondary and super-secondary elements obtained from the server is described in Table 5.

Table 5: Secondary Structure Information (ProMotif) of Chains A, B, C and D of Model

Particular	Chain A	Chain B	Chain C	Chain D
Sheets	4	4	5	3
Beta Hairpins	5	5	5	6
Beta Bulges	3	2	2	3
Strands	13	13	15	12
Helices	9	8	9	10
Helix-Helix Interactions	8	6	8	9
Beta Turns	28	33	33	29
Gamma Turns	3	5	5	5

The core of each monomer consists of β-α-β sandwich fold which is found to be present in the experimentally verified chorismate synthase indicating that protein has acquired the desired fold (Figure 9). The β-α-β sandwich fold is composed of four α-helices sandwiched between two anti-parallel β-sheets containing four β-strands on each side.

The pair-wise structural alignment of all the four monomers showed the RMSD values in the range of 0.69 to 0.83 Å for backbone atoms only indicating that these asymmetric units have highly similar structures. These values are better than experimentally solved structure of chorismate synthase from *H. pylori*.

Inter-Chain Interactions in Monomers of Model

Monomers interact with each other in variety of interactions like hydrogen bonds, salt bridge and non-bonded interactions as described in Table 6. Figure 10 demonstrates the hydrogen bonds and salt bridge present in between Chains A and D of model.

ISSN: 2320-3528

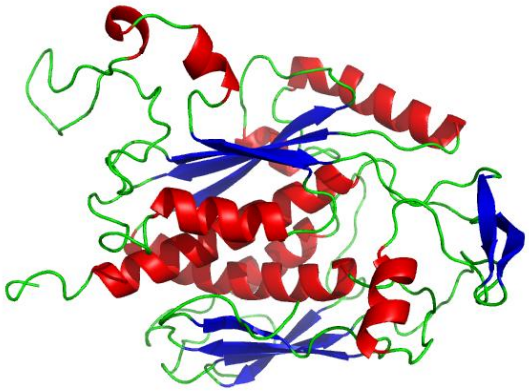


Figure 9: β-α-β Sandwich Fold in Monomer (α-Helix: Red and β-Sheet: Blue)

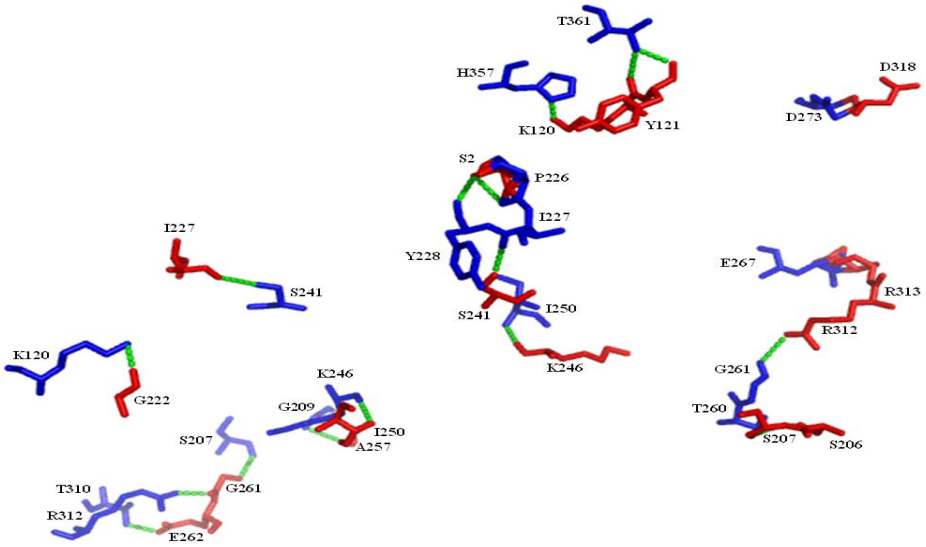


Figure 10: Pair-Wise Inter-Chain Interactions (Chain A: Red, Chain D: Blue, Hydrogen Bond: Green and Salt Bridge: Black Arrowed)

Table 6: Number and Type of Interactions in Paired Monomers in Model

Chains			Number of Interactions			
			Salt Bridges	Disulphide Bonds	Hydrogen Bonds	Non-Bonded Contacts
A	:	B	-	-	-	108
C	:	D	-	-	2	155
A	:	C	-	-	5	235
A	:	D	1	-	19	1298
B	:	C	4	-	23	1347
B	:	D	-	-	9	539

FMN Binding Site in Chorismate Synthase

The accuracy of the active site of chorismate synthase from *B. melitensis* was checked by performing its structural alignment with experimentally solved structure from *H. pylori*. The RMSD value for residues in active site was found to be 0.61 for backbone atoms only showing very high accuracy of the active site in chorismate synthase of *B. melitensis*. FMN binding site is highly conserved in the model and template [29]. The binding site in chorismate synthase model is composed of following strictly conserved residues ^{112}His, ^{131}Arg, ^{133}Ser, ^{140}Arg, ^{243}Asn, ^{246}Lys, ^{303}Lys, ^{305}Thr, ^{307}Ser, ^{331}Asp and ^{337}Arg from Chain A and ^{288}Gly from Chain D (Figure 11).

ISSN: 2320-3528

The structure of FMN consists of two parts: the head formed by the isoalloxazine ring and the tail region is composed of the ribityl chain. Both these structural parts are involved in the interactions with the residues in the active site (Figure 12). O4 atom of isoalloxazine ring forms the hydrogen bond with NH2 atom of ^{337}Arg. A water molecule bridges N5 of isoalloxazine ring and OD2 atom of ^{331}Asp by forming a network of hydrogen bonds. The ribityl hydroxyl groups OH2*, OH4* and OH5* make the hydrogen bonds with NH2 atom of ^{131}Arg, OG atom of ^{133}Ser and NZ atom of ^{303}Lys, respectively. The ribityl phosphate group is also involved in the hydrogen bonding. The oxygen atoms of O2P and O3P form the hydrogen bonds with OD1 atom of ^{243}Asn and N atom of ^{288}Gly' from Chain D, respectively. The two water molecules also form the hydrogen bonds with the oxygen atoms of O2P and O3P of ribiyl phosphate group. There is also the formation of weak salt bridge between N atom of ^{246}Lys and the oxygen atom of O1P of ribityl phosphate group.

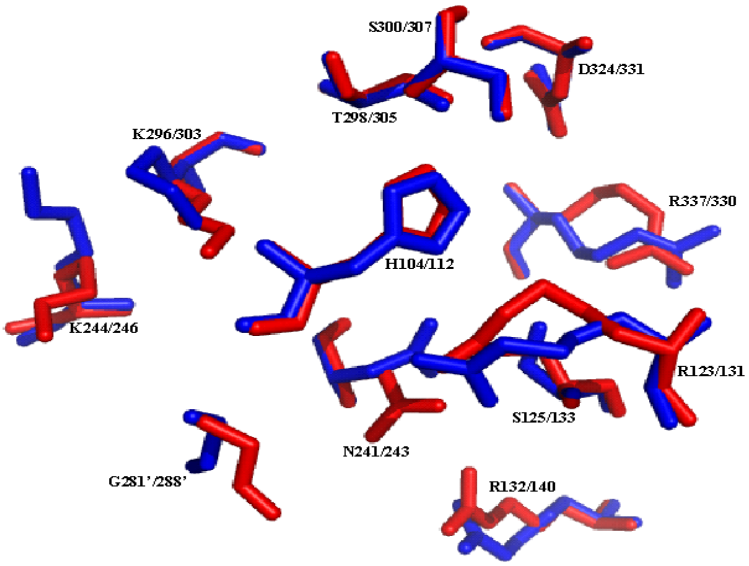


Figure 11: Conserved Amino Acids in Active Site of Template/Model (Template: Red and Model: Blue)

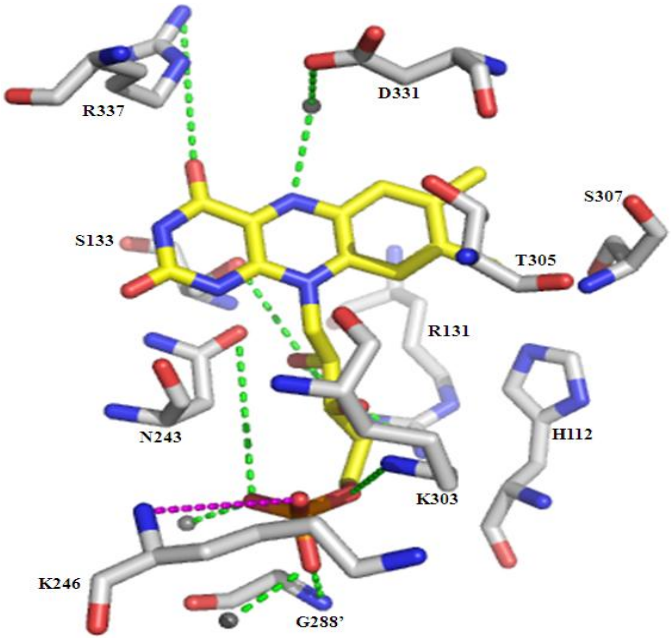


Figure 12: The Active Site Residues Interacting With FMN in A Particular Monomer (Hydrogen Bonds: Green, Water Molecules: Grey and Salt Bridge: Magenta)

The redox potential of FMN is strongly dependent on the charge present around it. It is proved that the positive charge increases, while the negative charge decreases the redox potential of FMN [34]. In our chorismate synthase model, the hydrophobic environment is present around the dimethyl-benzene moiety in isoalloxazine ring of FMN and three positively charged residues ^{131}Arg, ^{246}Lys and ^{303}Lys directly interacts with the ribityl chain and the ribityl phosphate group of FMN. Thus the electrostatic potential of FMN binding site in chorismate synthase seems to play a very important role in binding and activation of FMN (Figure 13).

ISSN: 2320-3528

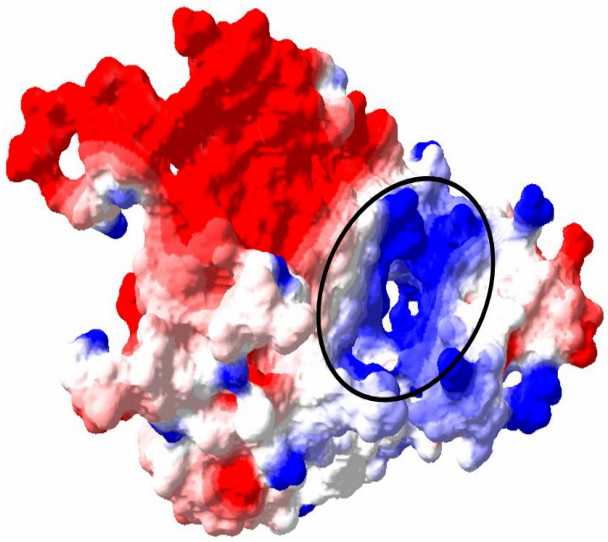


Figure 13: Positive Electrostatic Potential (Blue) of FMN Binding Pocket (Circled) for Ribityl Chain Interaction in Monomer of Model

A positive charge near N3-C4=O4 locus in the isoalloxazine ring stabilizes negatively charged due to N1 in the reduced flavin thus serving two purposes: to favor flavin binding and to regulate redox properties of the cofactor [29]. Therefore, ^{337}Arg might play an important role in binding and activation of reduced FMN in chorismate synthase.

CONCLUSION

For our knowledge we are the first to propose the molecular structure of chorismate synthase from *B. melitensis.* We have successfully generated a homology model of chorismate synthase with FMN for *B. melitensis*. Each monomer of 364 amino acids consists of newly emerged β-α-β sandwich fold in its core. The generated model is analyzed by various means and found to be a superior model in all respect. The conserved residues from the active site interact with FMN which is bound deeply within the protein; this information can be efficiently used in generating hypothesis regarding functional mechanism of chorismate synthase in *B. melitensis*. This study will surely prove to be useful in developing the specific therapeutic agent/s for the human use in the treatment of human brucellosis.

ACKNOWLEDGMENTS

We thank to Pravara Institute of Medical Sciences for providing the facility to carry out the study. We also thank to the staff members of Center for Biotechnology, Pravara Institute of Medical Sciences for their constant support and encouragement.

REFERENCES

1. Gwida M, Dahouk S, Melzer F, Rösler U, Neubauer H, Tomaso H. Brucellosis forum: Brucellosis - Regionally emerging zoonotic disease? Croat Med J. 2010; 51:289-95.
2. Mantur BG, Amarnath SK. Brucellosis in India - A review. J Biosci. 2008; 33:539-47.
3. Tan SY, Davis C. David Bruce (1855-1931): Discoverer of brucellosis. Singapore Med J. 2011; 52:138-9.
4. Ficht TA, Kahl-McDonagh MM, Arenas-Gamboa AM, Rice-Ficht AC. Brucellosis: The case for live, attenuated vaccines. Vaccine. 2009; 27:D40-3.
5. Smits HL, Kadri SM. Brucellosis in India: A deceptive infectious disease (review). Indian J Med Res. 2005; 122:375-84.
6. Bayram Y, Korkoca H, Aypak C, Parlak M, Cikman A, Kilik S, et al. Antimicrobial susceptibilities of Brucella isolates from various clinical specimens. Int J Med Sci. 2011; 8:198-202.
7. Sauret JM, Vilissova N. Human Brucellosis. J Am Board Fam Pract. 2002; 15:401-6.
8. Olsen SC, Stoffregen WS. Essential role of vaccines in brucellosis control and eradication programs for livestock. Expert Rev Vaccines. 2005; 4:915-28.
9. Hall WH. Modern chemotherapy for brucellosis in humans. Rev Infect Dis. 1990; 12:1060-99.
10. Mithani A, Preston GM, Hein J. Rahnuma: hypergraph-based tool for metabolic pathway prediction and network comparison. Bioinformatics. 2009; 25:1831-32.
11. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999; 27:29-34.
12. Cottret L, Wildridge D, Vinson F, Barrett MP, Charles H, Sagot MF, et al. MetExplore: A web server to link metabolomic experiments and genome-scale metabolic networks. Nucleic Acids Res. 2010, 38(suppl2):W132-7.

ISSN: 2320-3528

13. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, et al. The Swiss-Prot protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. 2003; 31:365-70.
14. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997; 25:3389-3402.
15. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, et al. Protein database searches using compositionally adjusted substitution matrices. FEBS J. 2005; 272:5101-9.
16. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res. 2000; 28: 235-42.
17. Šali A, Blundell TL. Comparative modelling by satisfaction of spatial restraints. J Mol Biol. 1993; 234:779-815.
18. Maiti R, Domselaar GHV, Zhang H, Wishart DS. SuperPose: A simple server for sophisticated structural superposition. Nucleic Acids Res. 2004; 32:W590-4.
19. Benkert P, Ku¨nzli M, Schwede T. QMEAN server for protein model quality estimation. Nucleic Acids Res. 2009; 37: W510-4.
20. Laskowski RA, MacArthur MW, Moss D, Thornton JM. ProCheck: A program to check the stereochemical quality of protein structures. J Appl Crystal. 1993; 26:283-91.
21. Wiederstein, Sippl. ProSA-Web: Interactive web service for the recognition of errors in three-dimensional structures of proteins; Nucleic Acids Res. 2007; 35:W407-10.
22. Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature. 1992; 356:83-5.
23. Castrignano T, De Meo PD, Cozzetto D, Talamo IG, Tramontano A. The PMDB protein model database. Nucleic Acids Res. 2006; 34:D306-9.
24. Meshram RJ, Jangle SN. Comparative metabolomic investigation of *Leishmania major* and *Homo sapiens*: An in-silico technique to discover and develop novel drug targets. Pharmacophore. 2011; 2:135-44.
25. Herrmann KM. The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. Plant Cell. 1995; 7:907-19.
26. Braus GH. Aromatic amino acid biosynthesis in the yeast *Saccharomyces cerevisiae*: A model system for the regulation of a eukaryotic biosynthetic pathway. Microbiol Rev. 1991; 55:349-70.
27. Combet C, Jambon M, Deléage G, Geourjon C. Geno3D: Automatic comparative molecular modeling of protein. Bioinformatics. 2002; 18:213-4.
28. Geourjon C, Combet C, Blanchet C, Deleage G. Identification of related proteins with weak sequence identity using secondary structure information. Prot Sci. 2001; 10:788-97.
29. Ahn HJ, Yoon HJ, Lee BI, Suh SW. Crystal structure of chorismate synthase: A novel FMN-binding protein fold and functional insights. J Mol Biol. 2004; 336:903-15.
30. Smith DK, Radivojac P, Obradovic Z, Dunker AK, Zhu G. Improved amino acid flexibility parameters. Prot Sci. 2003; 12:1060-72.
31. Vitkup D, Ringe D, Karplus M, Petsko GA. Why protein R-factors are so large: A self-consistent analysis. Proteins. 2002; 46:345-54.
32. Guex N, Peitsch MC. SWISS MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis. 1997; 18:2714-23.
33. Laskowski RA. PDBsum: Summaries and analyses of PDB structures. Nucleic Acids Res. 2001; 29:221-2.
34. Ghisla S, Massey V. Mechanisms of flavoprotein-catalyzed reactions. Eur J Biochem. 1989; 181:1-17.