

# STRUCTURAL STUDIES ON SOME METALLOPROTEINS: BIPHENYL DIOXYGENASE AND DAHP SYNTHASE

## A THESIS

*Submitted in partial fulfilment of the  
requirements for the award of the degree*

*of*

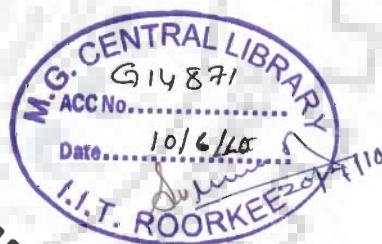
DOCTOR OF PHILOSOPHY

*in*

BIOTECHNOLOGY

*by*

**SUSHMITA BHATTACHARYA**

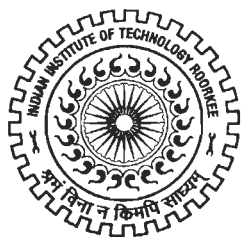


DEPARTMENT OF BIOTECHNOLOGY  
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE  
ROORKEE - 247 667 (INDIA)

FEBRUARY, 2009



**©INDIAN INSTITUTE OF TECHNOLOGY ROORKEE, ROORKEE, 2009  
ALL RIGHTS RESERVED**



# INDIAN INSTITUTE OF TECHNOLOGY ROORKEE ROORKEE

## CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in this thesis entitled, **STRUCTURAL STUDIES ON SOME METALLOPROTEINS: BIPHENYL DIOXYGENASE AND DAHP SYNTHASE** in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from January 2006 to February 2009 under the supervision of Dr. Pravindra Kumar, Assistant Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

The matter in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

(SUSHMITA BHATTACHARYA)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Pravindra Kumar)  
Supervisor

Dated: Feb. 13, 2009

The PhD Viva-Voce Examination of **Ms. Sushmita Bhattacharya**, Research Scholar has been held on ... ~~Jan 7~~ ... 30, ... 2009.

Signature of Supervisor

Signature of External Examiner

DEDICATED TO MY  
PARENTS



## ABSTRACT

Metalloproteins are a thrust area today because of the diverse roles that metal ions play in the stability and functional aspects of the protein. The work embodied in the thesis entitled '*Structural studies on some metalloproteins: Biphenyl dioxygenase and DAHP Synthase*' focuses on two metalloproteins that play important roles in their respective realms of function. The first metalloprotein, biphenyl dioxygenase (BPDO), is the enzyme that has significance in bioremediation strategies through bacterial degradation of polychlorinated biphenyls. Polychlorinated biphenyls (PCBs) are biologically recalcitrant compounds, toxic by nature, having carcinogenic potential and tendency to biomagnify at progressively higher levels in the food chain, posing some of the most serious health hazards worldwide. The PCB-degrading bacteria exhibit substantial differences in the range of degradation ability and in congener selectivity for PCB through the distinct structural difference of the BPDO enzymes they harbor. Therefore it is befitting that the three dimensional structures of BPDO enzymes from different species are known so that they could be engineered towards better degradation ability of PCBs. In this work, the three dimensional structures of variant BPDOs (derived from BPDO from *Burkholderia xenovorans LB400*) BPDO<sub>II9</sub> in substrate free form as well as complexed with biphenyl and BPDO<sub>III52</sub>, in substrate free form were determined. BPDO<sub>II9</sub> is  $\alpha_3\beta_3$  hexamer similar to Napthalene 1,2-Dioxygenases (NDOs) and other Biphenyl Dioxygenases (BPDOs).

The  $\alpha$  subunit can be divided into two distinct domains: a Rieske domain that contains the Rieske [2Fe-2S] center and the catalytic domain that contains the active site mononuclear non-heme ferrous iron. The ligation of the non heme iron (mononuclear centre) and Rieske (2Fe-2S) centres and the bridging between them in neighbouring catalytic subunits by

hydrogen bonds through a single amino acid (Asp 230) have been found similar to other BPDOs and NDOs (Naphthalene dioxygenases). The residues in the substrate binding pocket and possibly determining the substrate specificity have been identified. Also, the catalytic pocket of the variants had larger volumes as compared to other known BPDOs and NDOs. The docking studies of various PCBs into BPDO<sub>119</sub> shows that the amino acids mutated in BPDO<sub>119</sub> were more likely to accommodate greater variety of PCB congeners which could be responsible for the observed enhanced substrate specificity.

The second enzyme under study is the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS) synthase from *Arabidopsis thaliana* (*At*-DAHPS). DAHPS synthase is the first enzyme of the shikimate pathway which comprises a series of seven enzyme-catalyzed reactions responsible for the synthesis of aromatic compounds in microorganisms and plants. The absence of this pathway in animals makes the enzymes of this pathway attractive targets for the development of antibiotics and herbicides. In this study, the overexpression, purification, crystallization of *At*-DAHPS was performed and in the absence of suitably diffracting crystals (the crystals diffracted failed to diffract beyond 7 Å resolution), the structure prediction was done through homology modeling and phylogenetic inferences were drawn. Results have pointed out key residues involved in substrate binding and indicate a common phylogenetic ancestry for the two types of DAHPS synthases. The structure paves the way for design of inhibitors that may be used as herbicides.

## ACKNOWLEDGEMENTS

*At the outset I thank my guide, Dr. Pravindra Kumar for his unparalleled guidance and unstinted support. I express my heartfelt gratitude for his painstaking guidance during the sojourn of my research. His resourceful ideas, scientific attitude, stimulating discussions and valuable suggestions helped me all through my research work. I sincerely hope the lessons that I learnt under his benevolent guidance will hold me in good stead throughout my life.*

*More often than not, scientific research is an arduous task. It is based on hypothesis and collection of facts which either substantiate or negate the hypothesis. Today, when my research has almost culminated, down memory lane, I find a number of people without whose active support and help, it would have been difficult for me to complete the task.*

*Thus my sincere thanks go to Dr. Ritu Bathwal, Head, Department of Biotechnology, Indian Institute of Technology, Roorkee for her kind help whenever I needed and in providing me the necessary facilities to accomplish my research, Dr. Shailly Tomar for her painstaking guidance and support, Dr. Ashwani Kumar Sharma for his kind encouragement and guidance, Dr. Sanjoy Ghosh for his kind cooperation and Dr. Ramesh Chandra for his inspiration. I am also thankful to Dr. Partha Roy, Dr. H.S. Dhalwal, Dr. R. P. Singh, Dr. R. Prasad for their kind cooperation.*

*I owe special thanks to Dr. Arati Mishra, Scientist, Indian Institute of Technology, Kanpur and Prof. Pradip Sinha, Head, Department of Biosciences and Bioengineering, Indian Institute of Technology, Kanpur for their affectionate cooperation during my work in their laboratory and being a source of inspiration throughout the sojourn of my research.*

*I gratefully acknowledge the affectionate morale boosting by my fellow research scholars. I am also thankful to laboratory staff of the department for their cooperation.*

*Lastly, I express my gratefulness to my dear sisters for their boundless love, cooperation and moral support without which I could not have completed this work.*

FEBRUARY 2009

  
(SUSHMITA BHATTACHARYA)

## LIST OF PUBLICATIONS

### (i) Journals

- a) **Sushmita Bhattacharya**, Shailly Tomar, Ashwani Kumar Sharma and Pravindra Kumar (2009). “Structural Bioinformatics Studies on DAHP Synthase from *Arabidopsis thaliana*”, *Communicated*.
- b) **Sushmita Bhattacharya**, Michel Sylvestre and Pravindra Kumar (2009). “Exploring the efficacy of Biphenyl Dioxygenase variants in degrading polychlorinated biphenyls through structure studies.” *Communicated*.
- c) **Sushmita Bhattacharya**, Shailly Tomar, Ashwani Kumar Sharma and Pravindra Kumar (2008) “Combatting the Whooping Cough Bacteria: *in-silico* approaches”. *Communicated*.
- d) **Sushmita Bhattacharya**, Shailly Tomar, Ashwani Kumar Sharma and Pravindra Kumar. “Purification and characterization of a 14 kD protein from *Citrullus lanatus*”. *In preparation*.

### (ii) Conferences

- a) **Sushmita Bhattacharya**, Shailly Tomar, Ashwani Kumar Sharma and Pravindra Kumar (2008). “Studies on *Arabidopsis thaliana* DAHP (3-deoxy-D-arabino-heptulosonate 7-phosphate) Synthase”, Proceedings of the 3<sup>rd</sup> Uttarakhand State Science & Technology Congress, Uttarakhand State Council for Science & Technology, Indian Institute of Technology, Roorkee.
- b) Anindita Bhattacharya, Sagarika Bhattacharya and **Sushmita Bhattacharya** (2008). “Critical Loads Analysis of Acid Deposition in Kanpur City”, Proceedings of National Conference on Scientific and Legal Challenges of Global Warming, Brahmanand College, Kanpur, 54.
- c) **Sushmita Bhattacharya**, Shailly Tomar, Ashwani Kumar Sharma and Pravindra Kumar (2007). “Structural Bioinformatics Analysis of Shikimate Pathway Enzymes: A Promising Panacea for healthy living”, Proceedings of the International Conference on Bioinformatics and Drug Discovery (BioConvene 2007), University of Hyderabad, Hyderabad, 71.
- d) **Sushmita Bhattacharya**, Shailly Tomar, Ashwani Kumar Sharma and Pravindra Kumar (2007). “Structural bioinformatics analysis of Plant DAHP (3-deoxy-D-arabino-heptulosonate 7-phosphate) Synthase”, Proceedings of the National Symposium on Biophysics Biophysics in Medicine and Biology under the aegis of Indian Biophysical Society, Chandigarh.



- e) Anindita Bhattacharya, **Sushmita Bhattacharya** and Dipankar Saha (2007). “Critical Load of Acid Deposition in Agra, India”, Proceedings of the International Congress of Toxicology (ICT XI – 2007), Montreal, Canada, 10-15.
- f) **Sushmita Bhattacharya**, Anindita Bhattacharya and Subramanian Ganesh (2004). “Coding Length Polymorphism in human genes related to Polyglutamine Tracts” , Proceedings of the India Science Congress 2004, Chandigarh.
- g) Anindita Bhattacharya, Shaily Maloo and **Sushmita Bhattacharya**, (2003). “An Assessment of Critical Load in Agra Region”, Proceedings of the National Symposium on Biochemical Sciences: Health and Environmental Aspects (BSHEA – 2003), Dayalbagh Educational Institute, Agra, 50-53.



# CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	iii
LIST OF PUBLICATIONS	iv
LIST OF FIGURES	x
LIST OF TABLES	xii
ABBREVIATIONS USED	xiii
1.INTRODUCTION	5
2.REVIEW OF LITERATURE	5
2.1. BIPHENYL DIOXYGENASE	5
2.1.1. Polychlorinated Biphenyls as Persistent Environmental Pollutants	5
2.1.2. Bacterial Bioremediation of PCBs	7
2.1.3. The Biphenyl Dioxygenase Pathway	9
2.1.4. Rieske oxygenases	11
2.1.5. The Biphenyl Dioxygenase Enzyme System	15
2.1.6. Presence of metal in the catalytic centre	20
2.1.7. Regioselectivity of Biphenyl Dioxygenase towards different congeners	21
2.1.8. The Generation of Variants BPDO <sub>119</sub> and BPDO <sub>1152</sub>	25
2.1.9. Specific Aims of Study on Biphenyl Dioxygenase	29
2.2. 3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE	30
2.2.1. The Shikimate Pathway	30
2.2.2. Enzymes of the Shikimate Pathway	33
2.2.2.1. 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	33
2.2.2.2. 3-dehydroquininate synthase	33
2.2.2.3. 3-Dehydroquininate Dehydratase-Shikimate Dehydrogenase	34
2.2.2.4. Shikimate Kinase	35
2.2.2.5. 5-Enolpyruvyl shikimate 3-phosphate synthase	36
2.2.2.6. Chorismate Synthase	36

2.2.3. Subcellular Localization of the Shikimate Pathway	37
2.2.4. Regulation of 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) Synthase	38
2.2.5 Metal dependency of DAHP Synthase	39
2.2.6. <i>Arabidopsis thaliana</i> DAHP Synthase	40
2.2.7. DAHP Synthase from different organisms	42
2.2.7.1. <i>Pyrococcus furiosus</i>	42
2.2.7.2. <i>Escherichia coli</i>	43
2.2.7.3. <i>Saccharomyces cerevisiae</i>	43
2.2.7.4. <i>Thermotoga maritima</i>	44
2.2.7.5. <i>Mycobacterium tuberculosis</i>	44
2.2.8. Phylogenetic classification of DAHP synthase	45
2.2.9. Specific Aims of Study on 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase	47
3. STUDIES ON BIPHENYL DIOXYGENASE VARIANTS	49
3.1. ABSTRACT	49
3.2. INTRODUCTION	50
3.3. MATERIAL AND METHODS	52
3.3.1. Crystallization of BPDO <sub>119</sub> and BPDO <sub>11152</sub>	52
3.3.2. Preparation of Enzyme:Substrate (BPDO <sub>119</sub> :Biphenyl) Complex	52
3.3.3. Preparation of Protein Crystals for X-ray diffraction	53
3.3.4. X-ray Diffraction Intensity Measurements and Processing	54
3.3.5. Structure Determination by Molecular Replacement	55
3.3.6. Structure Refinement and Model Building	57
3.3.7. Structure Validation	57
3.3.8. Docking of substrates	57
3.4. RESULTS AND DISCUSSION	57
3.4.1. Crystallization of substrate free BPDO <sub>119</sub> , biphenyl complex of BPDO <sub>119</sub> and BPDO <sub>11152</sub>	57
3.4.2. Structure Determination	58
3.4.3. Overall Structure of Substrate free BPDO <sub>119</sub>	59
3.4.3.1. Structure of the $\beta$ -Subunit	60
3.4.3.2. Structure of the $\alpha$ -Subunit	61
3.4.3.3. Coordination of the Mononuclear Iron and the Rieske cluster	63
3.4.4. Structure of BPDO <sub>119</sub> with Substrate Biphenyl	65
3.4.5. Role of loop at the Active Site Entrance	66
3.4.6. Comparison with other Rieske non-heme oxygenases	66
3.4.7. Docking of Polychlorinated Biphenyls (PCBs).	70

3.5. CONCLUSIONS	71
4. STUDIES ON 3-DEOXY-D-ARABINO-HEPTULOSONATE 7- PHOSPHATE (DAHPS) SYNTHASE	73
4.1. ABSTRACT	73
4.2. INTRODUCTION	73
4.3. MATERIAL AND METHODS	76
4.3.1. Isolation, Purification and crystallization of <i>At</i> -DAHPS	76
4.3.1.1. Chemicals and Biochemicals	76
4.3.1.2. Expression of <i>Arabidopsis</i> DHS1 in <i>E. coli</i>	76
4.3.1.3. Purification of <i>Arabidopsis</i> DHS1	77
4.3.1.4. Protein estimation	77
4.3.1.5. Polyacrylamide gel electrophoresis(SDS-PAGE)	77
4.3.1.5.1. Preparation of Reagents	77
4.3.1.5.2. Preparation of sample buffer	79
4.3.1.5.3. Casting of gel	79
4.3.1.5.4. Sample Preparation	81
4.3.1.5.5. Electrophoresis	81
4.3.1.5.6. Staining and destaining of gel	81
4.3.1.6. DAHP Synthase Enzyme Assay	82
4.3.1.7. Crystallisation trials	82
4.3.2. Structure Prediction and Phylogenetic Studies	82
4.3.2.1. Sequence retrieval	82
4.3.2.2. Secondary structure prediction and sequence allignment	83
4.3.2.3. Molecular modeling	83
4.3.2.4. PDB viewers	84
4.3.2.5. Validation of the model	85
4.3.2.6. Phylogenetic studies	85
4.4. RESULTS AND DISCUSSION	
4.4.1. Isolation, Purification and Crystallisation of <i>At</i> -DAHPS	85

4.4.1.1. Expression and purification of <i>Arabidopsis</i> DHS1	85
4.4.1.2. SDS PAGE analysis	85
4.4.1.3. Crystallisation of <i>At</i> -DAHPS	86
4.4.2. Structure Prediction and Phylogenetic studies on <i>At</i> -DAHPS	86
4.4.2.1. Quality of the model	86
4.4.2.2. Overall description of the Model	87
4.4.2.3. Active site organization	89
4.4.2.4. Phylogenetic studies	91
4.4.2.5. Comparison of <i>At</i> -DAHPS with DAHPS Type I	92
4.5. CONCLUSIONS	93
5. BIBLIOGRAPHY	96



## LIST OF FIGURES

S.No.	Figure	Fig. No.
1	Chemical structure of PCBs.	2.1
2	The Biphenyl Dioxygenase Pathway	2.2
3	Ribbon diagram showing the groups of Rieske oxygenases (ROs), divided on the basis of phylogenetic analysis.	2.3
4	Substrates for Biphenyl dioxygenases	2.4
5	Strategies to generate variants with better PCB degrading capability	2.5
6	Sequence alignment of LB400 BphA with variants obtained from shuffling portions of the distal segment of LB400 bphA with B-356 bphA and P6 bphA1.	2.6
7	The Shikimate Pathway	2.7
8	DAHPS synthase from <i>Pyrococcus furiosus</i> (PDB code 1ZCO)	2.8
9	DAHPS synthase from <i>Escherichia coli</i> (PDB code 1KFL)	2.9
10	DAHPS synthase from <i>Saccharomyces cerevisiae</i> (PDB code 1OFQ).	2.10
11	DAHPS synthase from <i>Thermotoga maritima</i> (PDB code 1VR6)	2.11
12	DAHPS synthase from <i>Mycobacterium tuberculosis</i> (PDB code 2B7O)	2.12
13	Pictures and diffraction pattern of crystals of variant BPDOs	3.1
14	Ribbon diagram showing the overall structure of BPDO <sub>119</sub> .	3.2
15	Lys 95 of BPDO <sub>119</sub> from $\beta$ -subunit makes hydrogen bonds with Asp 387 from the $\alpha$ -subunit.	3.3
16	Electron Transfer between the Rieske center and the catalytic iron site.	3.4
17	Electron density contoured at 1.0 $\sigma$ showing the coordination of the catalytic iron in substrate free BPDO <sub>119</sub> .	3.5
18	Electron density figure contoured at 1.0 $\sigma$ showing the biphenyl bound state of the catalytic site in BPDO <sub>119</sub> .	3.6
19	Movement of the residues Asp 388, His 233 and His 239 in BPDO <sub>119</sub> upon substrate binding.	3.7
20	Access to the catalytic site in BPDO <sub>119</sub> and NDO.	3.8
21	The active site of BPDO <sub>119</sub> with biphenyl bound	3.9
22	Superimposition of BPDO <sub>119</sub> over NDO (PDB code 2B24)	3.10
23	PCBs docked in BPDO <sub>119</sub>	3.11
24	Sequence alignment of the <i>At</i> -DAHPS isozymes.	4.1
25	Expression and Purification of <i>At</i> -DAHPS	4.2
26	Photographs and diffraction pattern for crystals of <i>At</i> -DAHPS.	4.3
27	Ramachandran diagram $\phi$ - $\psi$ plots for the (A) <i>At</i> -DAHPS structure (B) <i>Mt</i> -DAHPS structure.	4.4

28	The sequence alignment of <i>At</i> -DAHPS encoded by <i>DHS1</i> (Accession no. NP_195708.1 ) and <i>Mt</i> -DAHPS (PDB Id 2B7O_B).	4.5
29	Superimposition of the backbone of <i>At</i> -DAHPS onto <i>Mt</i> -DAHPS (PDB code 2B7O).	4.6
30	Comparison of the core and non core elements of $(\beta/\alpha)_8$ barrel of DAHP synthase enzymes.	4.7
31	Predicted <i>At</i> -DAHPS structure	4.8
32	Active site of <i>At</i> -DAHPS.	4.9
33	Phylogenetic relationship of DAHP synthase family members and the <i>At</i> -DAHPS sequence.	4.10

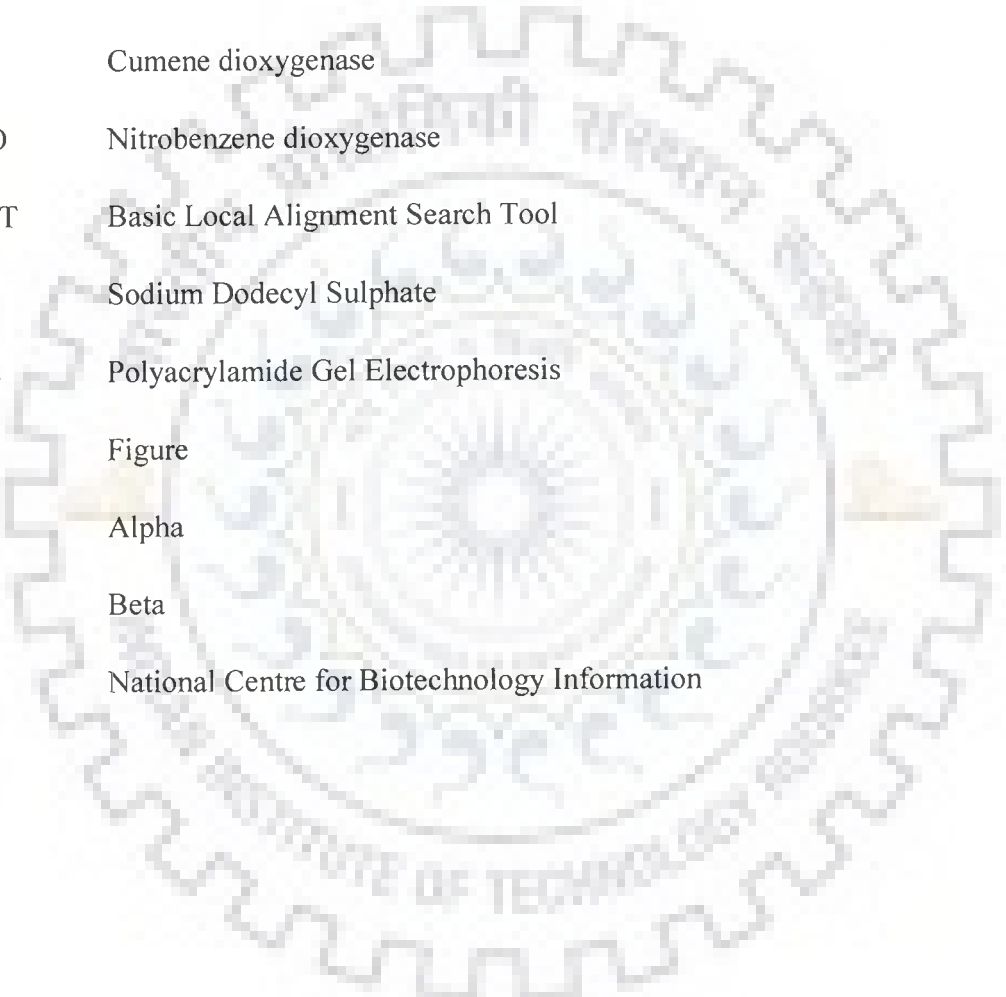


## LIST OF TABLES

S.No.	Table Title	Table No.
1.	Sequence identity (%) shared by variants BPDO <sub>119</sub> and BPDO <sub>1152</sub> with other Rieske oxygenases	2.1
2.	Sequence identity (%) shared between primary sequences of members of DAHP synthase family. The values were obtained using ClustalW	2.2
3.	Data collection statistics of BPDO <sub>119</sub> native crystal, BPDO <sub>119</sub> : Biphenyl complex and BPDO <sub>1152</sub>	3.1
4.	Sequence and structural statistics for various Rieske oxygenases compared to BPDO <sub>119</sub>	3.2
5.	Residues lining the active site in BPDO <sub>119</sub> and corresponding amino acids in other Rieske oxygenases	3.3.
6.	Distances between the amino acid residues and PEP in <i>At</i> -DAHPS	4.1



## ABBREVIATIONS USED



BPDO	Biphenyl dioxygenase
DAHP	3-deoxy-D- <i>arabino</i> -heptulosonate 7-phosphate
NDO	Naphthalene-1,2-dioxygenase
CDO	Cumene dioxygenase
NBDO	Nitrobenzene dioxygenase
BLAST	Basic Local Alignment Search Tool
SDS	Sodium Dodecyl Sulphate
PAGE	Polyacrylamide Gel Electrophoresis
Fig.	Figure
$\alpha$	Alpha
$\beta$	Beta
NCBI	National Centre for Biotechnology Information



# Chapter - 1

## *Introduction*

Metalloproteins are among the most efficient and diverse biocatalysts of nature. Metalloprotein is a generic term for a protein that contains a metal ion cofactor. Metal binding proteins are ubiquitous in bacteria, plants and animals and therefore represent a molecular group of major importance for living organisms. The role of metals in biological systems is important and manifold and is therefore characterised by a rapidly growing knowledge and data on metalloproteins. Metalloproteins serve a host of multifarious functions in diverse organisms. Indeed, about one quarter to one third of all proteins require metals to carry out their functions (Waldron and Robinson, 2009). The metal ion is usually coordinated by nitrogen, oxygen or sulfur atoms belonging to amino acids in the polypeptide chain and/or a macrocyclic ligand incorporated into the protein. The presence of the metal ion allows metalloenzymes to perform functions such as redox reactions that cannot easily be performed by the limited set of functional groups found in amino acids (Messerschmidt *et al.*, 2001).

Depending on the roles they play in the multitude of situations, metalloproteins are classified as proteins for metal ion transport and storage (transferrin, ferritin and aconitase), oxygen transport and storage (myoglobin, hemerythrin and hemocyanin), non-redox enzymes including transferases, hydrolases, lyases, isomerases, ligases (carbonic anhydrase, carboxypeptidase A, DAHP synthase), electron carriers (cytochrome C, rubredoxin, 2Fe-2S ferredoxin, 3Fe-4S and 4Fe-4S ferredoxin, copper blue) and redox enzymes (Rieske oxygenase, methane monooxygenase, nitrogenase, hydrogenase, Cu-Zn superoxide dismutase and dimethyl sulfoxide reductase). There are various techniques to elucidate protein structure and X-ray crystallography has become powerful tool to provide the three dimensional

structure for most of the proteins. The set of protein folds obtained by this method for the proteins provides the basis for the prediction of the three-dimensional structure of most of the remaining proteins using homology modeling techniques.

In this study, we have focused on two different metalloproteins, (i) biphenyl dioxygenase (BPDO) and (ii) 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase.

### **(i) Biphenyl dioxygenase**

Biphenyl dioxygenase (BPDO) is the first enzyme of the biphenyl degradation pathway. Biphenyl dioxygenases (BPDOs) have been extensively studied in terms of the degradation of toxic polychlorinated biphenyls (PCB) from several bacteria. These PCB-degrading bacteria exhibit substantial differences in the range of degradation ability and in congener selectivity for PCB. BPDOs are involved in the initial oxygenation of biphenyl and thereby the cometabolic degradation of PCB and other toxic aromatic compounds. BPDO is a three-component enzyme, which accommodates a mononuclear non-heme iron (II) center in the catalytic domain located in the terminal oxygenase component, where a Rieske [2Fe-2S] cluster is also hosted. The other two components of the enzyme, the reductase and the ferredoxin components, provide the electron transfer pathway, through which the two external electrons reach the active site. BPDOs display distinct ranges of PCB substrates from different species of microorganisms despite having high identity. It has been reported that some terminal dioxygenase variants exhibited enhanced abilities to degrade PCB and some biphenyl-related compounds. Therefore, it is necessary to determine three dimensional structures of terminal dioxygenases from different species and their variants with substrate

complexes. Here, the three dimensional structures of variants, BPDO<sub>119</sub>, in substrate free form as well as complexed with biphenyl and BPDO<sub>1152</sub>, in substrate free form were determined. Structure of BPDO<sub>119</sub> variant and its complex and comparison with other Rieske dioxygenases will be helpful in understanding the mechanisms by which the amino acids of catalytic pocket interact with the substrate to bind and orient it and to identify the critical residues to determine the substrate specificity and regiospecificity.

**(ii) 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHP) synthase**

3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (DAHP) synthase is the first metal activated enzyme in the shikimate pathway. This enzyme catalyzes an aldol-like condensation reaction between phosphoenolpyruvate (PEP) and D-erythrose 4-phosphate (E4P) to generate 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) and inorganic phosphate (Hermann and Weaver, 1999). The shikimate pathway comprises a series of seven enzyme-catalyzed reactions that result in the biosynthesis of chorismate, which is the precursor for many essential aromatic compounds. The shikimate pathway is found in microorganisms and plants and absent in animals, making the enzymes of this pathway attractive targets for the development of antibiotics and herbicides (Steinrucken and Amrhein, 1990). The study on DAHP synthase from a plant was undertaken with a view that though the structure of DAHP synthase is known from bacteria, fungi and archaea, there is no structure available from plant yet. The 3D structure from a model plant might aid in probing the plant DAHP synthase structure in general which might lead to design of effective herbicides in the future. Here the purification, crystallization and three dimensional

structure predictions of DAHP synthase from *Arabidopsis thaliana* are presented, the knowledge of which could aid in generation of efficient herbicides.

The focus on the first enzymes of two noteworthy biological pathways is with the impending view that the first enzyme of a pathway is generally the most important enzyme because the product of the first reaction serves as reactant for the second and so on goes the cascade. It has been observed in the case of BPDO that it is by varying the sequence of the first enzyme of the biphenyl degradation pathway, effective increase in the number of congeners degraded by the enzyme occurs.

Exciting progress has been made in the recent times in terms of the growing number of structures that are being solved of metalloproteins and those being predicted on the basis of these structures by computational biology. In view of the immense importance of metalloproteins, protein design too has gathered momentum and search is on to design small, stable and cost-effective model metalloproteins for biochemical, biotechnological and pharmaceutical applications. A lucid picture of metalloproteins is thus on the horizon, owing to the massive steps undertaken to study the structure of these metalloproteins by structure biologists and bioinformaticists alike and this thesis is a step in that direction. It is hoped that the work presented in this thesis through solving the crystal structure of biphenyl dioxygenase variants and predicting the three dimensional structure of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase, will contribute towards structure based enzyme engineering efforts on one hand and towards possible rational inhibitor design for DAHP synthase on the other. In the process it will also be helpful in accelerating the bridging of the sequence- structure gap by studying two biologically significant metalloproteins.



# Chapter - 2

## *Review of Literature*

## 2.1. BIPHENYL DIOXYGENASE

### 2.1.1. Polychlorinated Biphenyls as Persistent Environmental Pollutants

Polychlorinated biphenyls (PCBs) are a class of organic compounds containing 1 to 10 chlorine atoms attached to biphenyl which is a molecule composed of two benzene rings each containing six carbon atoms (Fig. 2.1). The chemical formula for all PCBs is  $C_{12}H_{10-x}Cl_x$ . There are 209 theoretically possible PCB congeners containing varying number of chlorine atoms at varying positions. Commercially available PCB mixtures, such as Aroclors, Clophens, or Kanechlors, typically contain between 60 and 80 different congeners differing in the number and position of chlorination. PCBs are stable, biologically recalcitrant compounds. It is estimated that our biosphere contains 750,000 tons of released PCBs.

Toxic by nature, they are widely dispersed in the environment (Francova *et al.*, 2003). These compounds were extensively produced industrially between 1929 and 1978 primarily because of their stable chemical and physical properties (Koppe and Keys, 2002). They are chemically inert and thermally stable and have low conductivity, which made them very useful for many applications (Hammond, 1972). PCBs were used as coolants and lubricating fluids for transformers and capacitors, stabilizing additives in flexible PVC coatings of electrical wiring and electronic components, pesticide extenders, cutting oils, flame retardants, hydraulic fluids, sealants (used in caulking, etc.), plastics, adhesives, wood floor finishes (Rudel *et al.*, 2008), paints, de-dusting agents and in carbonless copy paper (Boyle *et al.*, 1992).



Because of their widespread use, heterogeneity, persistence, toxicity, carcinogenic potential, and tendency to biomagnify at progressively higher trophic levels, PCBs are recognized to be some of the most serious environmental pollutants worldwide. (Abramowicz *et al.*, 1990; Hansen *et al.*, 1987; Silberhorn *et al.*, 1990). The production of PCB was banned in the 1970s in the USA in view of the impending alarm but the practice nevertheless continues, on a decreased scale. PCBs are omnipresent in sediments, soil and living organisms. The toxic effects of PCBs in the environment and on human health began to be visible about 30 years after production. Several studies found PCBs to be the cause of developmental disorders and infertility of marine animals and birds (Aoki, 2001; Fernie *et al.*, 2000; Fry, 1995; Koppe and Keys, 2002). Direct clinical evidence for the deleterious effect of PCBs on human health came from two accidents, the Yusho disease in Japan (1968) and the Yu-Cheng disease in Taiwan (1979), where people consumed rice oil contaminated with PCBs and developed darkened skin, eye discharge, severe acne and liver cirrhosis (Aoki, 2001; Hsu *et al.*, 1985). Other studies have proved that PCBs enhance the risk of developmental (Jacobson and Jacobson, 1996), neurological (Eriksson *et al.*, 2006), immune (Swanson *et al.*, 1995) and endocrine dysfunctions in humans (Aoki 2001; Toft *et al.*, 2004). Aggressive engineering methods, e.g. excavation, are suitable for “hot spot” removal of PCBs, but less-expensive containment and treatment technologies are required for remediation of surrounding areas or sites contaminated at a lower level.

PCBs can be degraded by high temperature (1200°C). However, one of the degraded products, dioxin, is more poisonous than the original PCB. Hence bioremediation may provide a safe and cost-effective alternative to current methods for PCB cleanup. Also, the high cost and public opposition to current physical remediation technologies have motivated

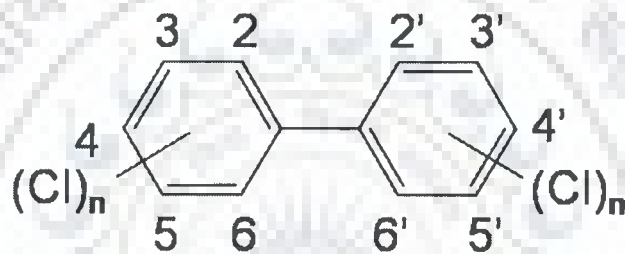


Fig. 2.1. Chemical structure of PCBs. The possible positions of chlorine atoms on the benzene rings are denoted by numbers assigned to the carbon atoms.

the surge for exploitation of ecofriendly biological breakdown systems employing microorganisms and recently also plants to clean-up PCB-contaminated sites.

### **2.1.2. Bacterial Bioremediation of PCBs**

For nearly 50 years prior to early 1970s, PCBs were perceived as immutable and completely refractile to microbial degradation owing to their super hydrophobicity, thermodynamic stability, structural diversity and toxicity. However, it was the early report of Ahmed and Focht (1973) and several other investigators (Furukawa *et al.*, 1978, 1979; Bedard *et al.*, 1986) that shattered the dogma of PCB recalcitrance. PCBs are transformed aerobically by the biphenyl pathway, which is found in a wide variety of soil organisms (Furukawa *et al.*, 1979; Bedard *et al.*, 1986). Bacteria have been shown to be able to degrade PCB compounds directly and cometabolically under anaerobic and aerobic conditions under both laboratory and field conditions (Hrywna *et al.*, 1999). However, bacterial bioremediation in the field is often confounded by numerous factors such as microbial competition, fluctuating environmental conditions or limited and sporadic metabolic resources. Plants have been shown to enhance or stabilize bacterial degradation of various organic pollutants, including PCBs and polycyclic aromatic hydrocarbons (PAHs). Enhanced dechlorination of PCBs was reported using genetically modified *Sinorhizobium meliloti* (Toure *et al.*, 2003).

Hydrophobic aromatic hydrocarbons may be metabolized through the initial activation of the aromatic ring by insertion of molecular oxygen using certain oxygenases found in the host organisms (Boyd *et al.*, 2001; Cerniglia *et al.*, 1979; Kim *et al.*, 1999). The consequences of such reactions result in the production of intermediates containing epoxide, phenol, or *cis*-

dihydrodiol functional groups that correspond to the parent aromatic compound from which it was derived (Cerniglia *et al.*, 1980; Cirino *et al.*, 2002, Ensley *et al.*; 1983, Guengerich *et al.*, 2003; Jerina *et al.*, 1971; Jerina *et al.*, 1968; Whited *et al.*, 1986). Based upon whether a eukaryote or prokaryote is performing the activity (Cerniglia *et al.*, 1980; Cerniglia *et al.*, 1979), there are fundamental differences in the initiation reactions for the metabolism of aromatic compounds. Eukaryotes produce aromatic epoxides, followed by the formation of *trans*-dihydrodiols (Oesch *et al.*, 1972). The key enzyme involved in the reaction has been shown to be cytochrome P450 monooxygenase, which operates over a wide spectrum of substrates (Guengerich *et al.*, 2003; Selander *et al.*, 1975). Prokaryotes, however, harness different metabolic routes for transforming aromatic compounds, producing either phenols by monooxygenases or *cis*-dihydrodiols by dioxygenases (Boyd *et al.*, 2001; Ensley *et al.*, 1983; Mitchell *et al.*, 2003; Suenaga *et al.*, 2001; Tao *et al.*, 2004).

Thus, there are two key processes for bacterial bioremediation of PCBs; aerobic biodegradation and anaerobic dehalogenation (Abramowicz, 1990). Aerobic biodegradation involves breakdown of the biphenyl core of PCBs with catabolism of the breakdown products by subsequent pathways. Anaerobic dehalogenation is a reductive dechlorination process where the chlorinated compound is an electron acceptor (Wiegel and Wu, 2000). The biphenyl core remains intact in this process, but the level of chlorination is reduced. As aerobic biodegradation is most successful against lightly chlorinated biphenyl cores, it is suggested that anaerobic dehalogenation may be utilized to enhance the aerobic degradability by lowering the chlorination in the biphenyl core (Furukawa, 2000). Compounds for which aerobic microbial degradative pathways have been identified include benzene, benzoate, biphenyl, naphthalene, phthalate, and toluene (Erickson *et al.*, 1992). In each case, the initial

step of aerobic degradation involves a ring-hydroxylating Rieske oxygenase (RO), which converts an aromatic substrate to a cyclic dienediol.

### 2.1.3. The Biphenyl Dioxygenase Pathway

The initial enzymic steps in the aerobic degradation of many aromatic compounds are very similar, suggesting that they have a common, if distant, evolutionary origin. The enzymes of the naphthalene-degradation pathway can transform other polycyclic aromatic hydrocarbons, including indole and biphenyl (Barriault *et al.*, 1999; Ensley *et al.*, 1983; Resnick *et al.*, 1996). However, 6-phenyl HODA hydrolase (in the biphenyl pathway), 2-hydroxy-6-oxohepta-2, 4- dienoic acid hydrolase (HOHDA hydrolase; in the toluene pathway) and cis-29-hydroxybenzalpyruvate aldolase (in the naphthalene pathway), are specific for their respective products.

The aerobic microbial degradation of biphenyl, PCBs, and PAHs starts with an “upstream” pathway, which catabolizes biphenyl into benzoate and 2-hydroxypenta-2,4-dienoate, and a “downstream” pathway, which processes 2-hydroxypenta-2,4-dienoate into Kreb’s cycle metabolites . The biphenyl dioxygenase pathway is a four step process (Fig. 2.2) with the initial oxygenase, product of the *bphA* gene, crucially responsible for recognition and binding of the substrate. The initial reaction is catalyzed by biphenyl dioxygenase (BPDO) (Haddock and Gibson, 1995; Hurtubise *et al.*, 1995). The terminal dioxygenase component of BPDO activates molecular oxygen to introduce it into the biphenyl molecule at the 2,3 position to obtain a 2,3-dihydro-2,3-diol, which then undergoes dehydrogenation to 2,3-dihydroxybiphenyl by dihydrodiol dehydrogenase (BphB). The second dioxygenase, 2,3-dihydroxybiphenyl dioxygenase (BphC), requires no any external reductant and cleaves the

2,3-dihydroxylated ring between carbon atoms 1 and 2 to produce 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid (HOPD, the ring *meta*-cleavage product), which is then hydrolyzed to benzoic acid and 2-hydroxypenta- 2,4-dienoate by a hydrolase (BphD). These upper pathway enzymes in biphenyl metabolism are encoded by the *bph* gene clusters, in which *bphA1* and *bphA2* encode a large and a small subunit of the terminal dioxygenase, *bphA3* encodes ferredoxin, and *bphA4* encodes ferredoxin reductase (Erickson and Mondello, 1992). The *bphB*, *bphC*, and *bphD* genes encode a dehydrogenase, a ring-cleavage dioxygenase, and a hydrolase, respectively. Out of these, the large subunit of terminal dioxygenase is critically involved in the substrate specificity of biphenyl dioxygenase (Kimura *et al.*, 1997). Therefore, evolutionary molecular engineering has been applied to large-subunit genes of different origins, creating novel dioxygenases. Evolved biphenyl dioxygenases thus obtained show enhanced and expanded degradation for not only PCBs, but also other related compounds (Barriault *et al.*, 2002; Furukawa *et al.*, 2004).

Each of the steps involved in the upstream pathway presents a unique bottleneck for the aerobic microbial degradation of PCBs and PAHs. Structural information is vital for identifying elements critical for the specificity responsible for catalytic limitations. Recent crystal structures of several ROs have contributed greatly to identifying structural elements responsible for specificity (Kauppi *et al.*, 1998; Dong *et al.*, 2005; Furusawa *et al.*, 2004; Martins *et al.*, 2005; Nojiri *et al.*, 2005; Friemann *et al.* 2005; Gakhar *et al.*, 2005).

Genes for catabolic functions have evolved in nature through various genetic events which include mutation, recombination, gene transfer, and assembly, resulting in a family of diverse but highly related sequences. As a consequence, the *bph* genes are present on bacterial

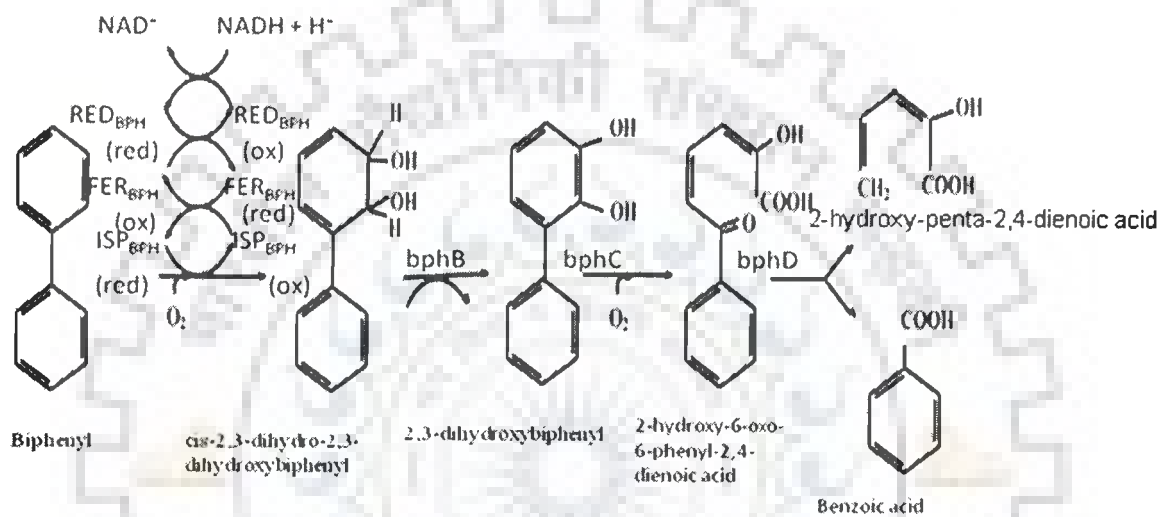


Fig. 2.2. The Biphenyl Dioxygenase Pathway

chromosomes, plasmids and transposons (Furukawa *et al.*, 2004). Genes that encode this pathway have been cloned and characterized (Erickson and Mondello, 1992; Taira *et al.*, 1992; Fukuda *et al.*, 1994; Sylvestre *et al.*, 1996). The enzymes of the bacterial biphenyl catabolic pathway are very versatile. Other than PCBs they can also catalyze reactions using several hydroxybiphenyls and chlorohydroxybiphenyls as substrate (Sondossi *et al.*, 1991). Genes encoding the BPH (upper) degradation pathway and benzoate (lower) pathways (Pieper, 2005) in isolates have been extensively characterized, and some genes have been directly detected in the environment using molecular tools (Erb and Wagner-Dobler, 1993; Ringelberg *et al.*, 2001; Baldwin *et al.*, 2003). However, until recently, methodological limitations precluded researchers from ascertaining which bacteria or genes are truly active in BP/PCB degradation in the soil community. The recent development of stable isotope probing (SIP) methods enables researchers to directly identify microorganisms involved in the degradation and/or assimilation of carbon from a specific compound without the limitations of cultivation (Friedrich, 2006; Whiteley *et al.*, 2006).

#### **2.1.4. Rieske oxygenases**

Proteins containing Rieske-type [2Fe-2S] clusters play important roles in many biological electron transfer reactions. Typically, [2Fe-2S] clusters are not directly involved in the catalytic transformation of substrate. The Fe-S clusters are generally redox active, and Fe-S proteins function over nearly the entire range of biologically relevant reduction potentials by modulating the Fe<sup>III</sup>/Fe<sup>II</sup> couple (Capozzi *et al.*, 1998). These proteins are widely distributed.

Proteins containing Rieske-type [2Fe-2S] clusters provide critical electron transfer reactions (Mason and Cammack, 1992; Trumpower and Gennis, 1994; Link 1999; Berry *et al.*, 2000)



in biological pathways. Rieske-type [2Fe-2S] clusters differ from the classical plant-type ferredoxin (Fd) [2Fe-2S] clusters in the nature of the amino acid ligands to the cluster. Although plant-type Fd [2Fe-2S] clusters are ligated by four cysteine residues, Rieske-type [2Fe-2S] clusters have a mixed coordination environment involving two histidines to one iron and two cysteines to the other iron. The cluster coordination has been firmly established by recent X-ray crystal structures of the Rieske iron–sulfur protein domains from the bovine mitochondrial *bc*<sub>1</sub> complex (Iwata *et al.*, 1996), the spinach chloroplast *b<sub>6</sub>f* complex (Carrell *et al.*, 1997), bacterial naphthalene dioxygenase (NDO) (Kauppi *et al.*, 1998), and bacterial Rieske-type Fd (BphF) (Colbert *et al.*, 2000). The di-histidine-coordinated iron becomes a localized ferrous site upon reduction of the cluster with the di-cysteine-coordinated iron remaining ferric (Fee *et al.*, 1984; Cline *et al.*, 1985; Kuila and Fee, 1986; Gurbiel *et al.*, 1996). Variations in typical protein bound Fe-S cluster occur in the form of construction of larger metallocofactors by elaboration or fusion of the basic units as well as the use of non-cysteiny protein ligands. Rieske-type Fe-S proteins provide an example of the latter kind. Both Fe atoms are ferric in the oxidized center, so that the core is  $[\text{Fe}_2\text{S}_2]^{2+}$  and the cluster-ligand complex is  $[(\text{Cys})_2\text{Fe}^{\text{III}}(\mu_2\text{-S})_2\text{Fe}^{\text{III}}(\text{His})_2]^0$ . The spectroscopic properties of the histidine-ligated Fe change upon reduction so that the reduced Rieske center is formally  $[\text{Fe}_2\text{S}_2]^{1+}$ , or  $[(\text{Cys})_2\text{Fe}^{\text{III}}(\mu_2\text{-S})_2\text{Fe}^{\text{II}}(\text{His})_2]^{1-}$ . Standard 2Fe ferredoxins exploit the same core oxidation levels. The different types of Rieske clusters differ in their reduction potential: one with higher reduction potentials ( $E_m = +150$  to  $+490$  mV) in cytochrome *bc*<sub>1</sub>/*b<sub>6</sub>f* complexes of the aerobic respiratory chain and photosynthesis, and the other with lower reduction potentials ( $E_m = -150$  to  $-50$  mV, NHE) in a diverse group of bacterial multicomponent terminal oxygenases and soluble Rieske-type ferredoxins (Kuila and Fee, 1986; Link *et al.*,

1992; Mason and Cammack, 1992; Riedel *et al.* 1995; Link *et al.* 1996; Brugna *et al.*, 1999). The degree to which structural variations in the vicinity of the clusters contribute towards functional properties is a matter of intense speculation, although correlations among reduction potentials, variations of hydrogen-bond networks (Denke *et al.*, 1998; Schröter *et al.*, 1998; Guergova-Kuras *et al.*, 2000) and polypeptide dipoles in the vicinity of the Rieske-type clusters (Colbert *et al.*, 2000) have been proposed. The midpoint reduction potentials of the clusters in archaeal Rieske-type ferredoxin (ARF) from *Sulfolobus solfataricus* strain P-1, a small, soluble, thermophilic protein of unknown function, and bacterial *Acinetobacter* sp. strain ADP1 anthranilate dioxygenase (AntDO). This catalyzes the conversion of anthranilate to catechol in the  $\beta$ -keto adipate pathway for biodegradation of aromatic compounds (Eby *et al.*, 2001), are examples of Rieske-type clusters that belong to the lower reduction potential class. The midpoint redox potentials of the clusters in ARF and AntDO are -155 mV and approximately -125 mV (estimated from redox potential of 2-halobenzoate 1,2-dioxygenase) (Correll *et al.*, 1992; Riedel *et al.*, 1995; Coulter *et al.*, 1999; Eby *et al.*, 2001), respectively.

It is the class of ROs called the Rieske nonheme iron oxygenases (ROs) that are the focus of the present study. The first step in the biocatalytic degradation of the toxic polyaromatic compounds PAHs is catalyzed by this class of enzymes. These enzymes catalyze *cis* dihydroxylation reactions and require an oxygen molecule and two electrons that are not from the aromatic substrate. The electrons are produced by a reductase and transported to the dioxygenase by a ferredoxin (Mason *et al.*, 1992). Besides bioremediation, the ability of these ROs to form enantiopure products and the potential to tailor their regio- and stereospecificities make them very useful in the chiral synthesis of precursor compounds (Boyd *et al.*, 2001).

There are four dioxygenase subfamilies in this large family (the toluene/biphenyl, naphthalene, benzoate, and phthalate subfamilies) based on sequence alignment of the catalytic components ( $\alpha$ -subunits) (Werlen *et al.*, 1996). The toluene/biphenyl subfamily includes enzymes for the degradation of toluene, benzene, cumene (isopropylbenzene), biphenyl and polychlorinated biphenyls (PCBs). The naphthalene subfamily consists of enzymes for the degradation of naphthalene and phenanthrene. The Rieske dioxygenases involved in bacterial hydrocarbon degradation comprise multicomponent enzyme systems (Wackett, 2002) in which reduced pyridine nucleotide is used as the initial source of two electrons for dioxygen activation. The electrons pass through a flavin cofactor and Rieske [2Fe-2S] centers into the mononuclear iron center of the terminal Rieske nonheme iron dioxygenase component. However recent classification efforts have divided ROs into four groups based on phylogenetic analyses (Gibson *et al.*, 2000 and Nam *et al.*, 2001). Homooligomeric ROs, consisting of only  $\alpha$ -subunits, are classified as Group I ROs, while heterooligomeric ROs, consisting of a trimer of a heterodimers of  $\alpha$  and  $\beta$  subunits, constitute Groups II, III and IV. Groups II, III, and IV are distinguished by the nature of the substrates of representative ROs, with benzoate/toluene consisting substrates for Group II, naphthalene/polycyclic aromatic hydrocarbons for Group III, and biphenyl/toluene/benzene for Group IV (Gibson *et al.*, 2000 and Nam *et al.*, 2001) (Fig. 2.3).

It is observed that RO structures have either an  $\alpha_3\beta_3$  or an  $\alpha_3$  structure. All these ROs have two metal centers (Gakhar *et al.*, 2005). The metal centers include one Rieske-type iron sulfur center and one mononuclear iron center. The mononuclear iron is located in the active site and binds oxygen. The residues that ligate the Rieske cluster (two cysteines and two histidines) as well as the 2-His-1-carboxylate facial triad residues that ligate the mononuclear



Figure 2.3. Ribbon diagram showing the groups of Rieske oxygenases (ROs), divided on the basis of phylogenetic analysis. (A) Group I RO consisting of the homotrimeric class of RO represented by Oxoquinoline monooxygenase (OxoMo<sub>8</sub>) (PDB code 1ZO3). Groups III and IV consist of the heterooligomeric ROs. (B) Group III RO represented by naphthalene dioxygenase (PDB code 2B1X). (C) Group IV RO represented by biphenyl dioxygenase (PDB code 2GBX). All figures were created using PyMOL.

iron center are conserved among ROs. Another conserved residue is an aspartic acid that bridges the ligands of the two metal centers across subunit interfaces. This aspartic acid has been shown to be important for the transfer of electrons between the Rieske iron cluster and the mononuclear iron in the neighboring subunit (Parales *et al.*, 1999) and to play a part in conformational change at the mononuclear iron active site via hydrogen bonds upon reduction of the Rieske cluster (Martins *et al.*, 2005).

### 2.1.5. The Biphenyl Dioxygenase Enzyme System

Many scientists have made efforts to characterize biphenyl dioxygenase (Haddock *et al.*, 1995, 1997; Hurtubise *et al.*, 1996, 1998; Broadus and Haddock, 1998; Imbeault *et al.*, 2000; Suenaga *et al.*, 2001, 2002). BPDO is a three-component enzymatic system. The first component which is an oxygenase, is an iron-sulphur protein (ISP<sub>BPH</sub>) that catalyzes the addition of molecular oxygen at positions 2 and 3 of biphenyl to produce *cis*-(2R,3S)-dihydroxy-1-phenylcyclohexa-4,6-diene(*cis*-biphenyl-2,3 dihydrodiol), commonly called *cis*-2,3-dihydro-2,3-dihydroxybiphenyl. Second and third components are a flavoprotein reductase (RED<sub>BPH</sub>) and ferredoxin (FER<sub>BPH</sub>) that are involved in the transfer of electrons from NADH to ISP<sub>BPH</sub>, which then activates molecular oxygen for insertion into the aromatic substrate. BPDO components are coded by *bphA* (a-subunit of ISP<sub>BPH</sub>), *bphE* (b-subunit of ISP<sub>BPH</sub>), *bphF* (FER<sub>BPH</sub>) and *bphG* (RED<sub>BPH</sub>) in *Burkholderia sp.* strain LB400 (Erickson and Mondello, 1992; Asturias and Timmis, 1993) and in *Comamonas testosteroni* B-356 (Sylvestre *et al.*, 1996; Francova *et al.*, 2003).

Although certain PCBs serve as substrate for biphenyl dioxygenase (Mckay *et al.*, 1997; Zielinski *et al.*, 2002), PCB-degrading organisms do not usually use PCBs as an energy source, but rather catabolize these pollutants in the presence of easily metabolizable substrates such as biphenyl. Therefore, there is a perpetually ongoing search for other organisms with possibly better PCB-degrading competence. Also, strains that degrade other biaryls provide an underexploited source of PCB-transforming enzymes. For example, DoxG, an extradiol dioxygenase involved in naphthalene degradation by *Pseudomonas* sp. Strain C18, catalyzes the extradiol cleavage of 1,2-dihydroxynaphthalene (1,2- DHN) (Denome *et al.*, 1993; Fortin *et al.*, 2005).

Several investigators believed that only monochlorobiphenyls (CBs) function as a sole source of carbon and energy for aerobic bacteria. Others were of the opinion that earlier claims of isolation of PCB-mineralizing or dichlorobiphenyl (diCB)-degrading organisms must be viewed as equivocal because the cultures are not available, the media are insufficiently described to be reproduced by others, the purity of PCBs (i.e., absence of biphenyl) was not determined and no accompanying data on growth curves and chloride liberation were given. The report of McCullar *et al.* (1994) demonstrating growth of *Pseudomonas acidovorans* M3GY, a recombinant bacterium on 3,4'-diCB was proof that some diCBs can at least support growth of aerobic microorganisms. This strain was reported to mineralize 67% of 3,4'-diCB in 29 days. However, Potrawfke *et al.* (1998) were the first to document unambiguous growth by the natural organism, *Burkholderia* sp. LB400 (now *Burkholderia xenovorans* LB400, hereafter referred to as LB400) on 2,3'- and 2,4'-diCB. Similarly, Kim and Picardal (2001) described for the first time growth on a diortho-substituted chlorobiphenyl (2,2'-diCB) by a novel natural bacterium capable of mineralizing both 4-CB

and 2,4'-diCB. It is noteworthy that the substitution pattern of these diCB congeners makes them particularly recalcitrant to aerobic and anaerobic degradation (Adebusoye *et al.*, 2008; Abramowicz, 1990).

Even the reported highly potent strains poorly transform congeners containing more than four chloro substituents and none completely degrades all lightly chlorinated congeners. For example, one of the best PCB-degrading strains, LB400, possesses the unusual ability to transform 2, 2', 5, 5'-tetrachlorobiphenyl; however, it is transformed to the corresponding 3,4-dihydroxybiphenyl, a dead-end metabolite. Efforts to improve the PCB-degrading capabilities of natural isolates have for some time now, included the engineering of biphenyl degradation pathway enzymes to transform a broader range of congeners and the introduction of downstream pathways to degrade the chlorobenzoates produced by *bph* (Gibson and Parales, 2000). The result of an X-ray structural analysis of naphthalene dioxygenase (Karlsson *et al.*, 2003), which is an enzyme with a topology slightly different from that of biphenyl dioxygenase, suggested that *bphA* carried a substrate-recognition site. Protein engineering of the Bph enzymes further proved that the substrate specificity of the biphenyl-degrading pathway was mainly dependent on *bphA* (Barriault *et al.*, 2002; Kimura *et al.*, 1997; Kumamaru *et al.*, 1998; Misawa *et al.*, 2002; Suenaga *et al.*, 2001; Furukawa *et al.*, 2006; Kagami *et al.*, 2008; Zielinski *et al.*, 2003). Therefore most of the engineering has focused on the catalytic component of biphenyl dioxygenase (*bphA1-2*), the first enzyme of the pathway (Barriault and Sylvestre, 2004, Furukawa *et al.*, 2004; Suenaga, 2001). Because, the large subunit of biphenyl dioxygenase harbors the principal determinants of the enzyme substrate specificity, the corresponding genes from divergent bacterial strains have been subjected to DNA shuffling, generating variants possessing improved rates of transforming

certain PCB congeners (Barriault *et al.*, 2002; Kumamaru *et al.*, 1998). Various aromatic compounds, in which heterocyclic aromatics are linked with phenyl groups, have been converted into their *cis*-dihydrodiol forms by the evolved *bphA* gene, *bphA1* (2072), which was developed from DNA family shuffling between the *bphA* genes of *Pseudomonas pseudoalcaligenes* KF707 (hereafter referred to as KF707) and LB400, in combined expression with *bphEFG* of strain KF707 in *E. coli* (Misawa *et al.*, 2002).

Comparing chemical structures of biphenyl and flavone (Fig. 2.4), flavones may be a substrate for biphenyl dioxygenase. Flavones have also been implicated in various physiological activities, such as antioxidants, phytoestrogens in controlling breast cancer incidence and osteogenesis, for example, and in chemotaxis of nitrogen-fixing bacteria (Hollman and Katan 1999a,b; Nijveldt *et al.*, 2001). Considering this, there has been an increasing interest in the production of novel compounds from flavones using microbial reactions (Hosny *et al.*; Ibrahim *et al.*, 1990; 2001). Previously, biphenyl dioxygenase from *P. pseudoalcaligenes* KF707 (hereafter referred to as BPDO<sub>KF707</sub>) was shown to produce *cis*-2, 3-dihydrodiol on the B ring of flavone (Mitchell *et al.*, 2003). During the biotransformation of flavanone, as a structural analog of flavone, the enzyme produced an epoxide functional group between C2' and C3' on the B ring of the flavanone, which has not been reported previously for prokaryotic dioxygenase (Kim *et al.*, 1999, Resnick *et al.*, 1996; Spain and Gibson, 1988). Han *et al.* (2005), reported a novel prokaryotic monooxygenase reaction that results in the production of an epoxide on the flavanone B ring by biphenyl dioxygenase rather than producing a *cis*-dihydrodiol (Han *et al.*, 2005).



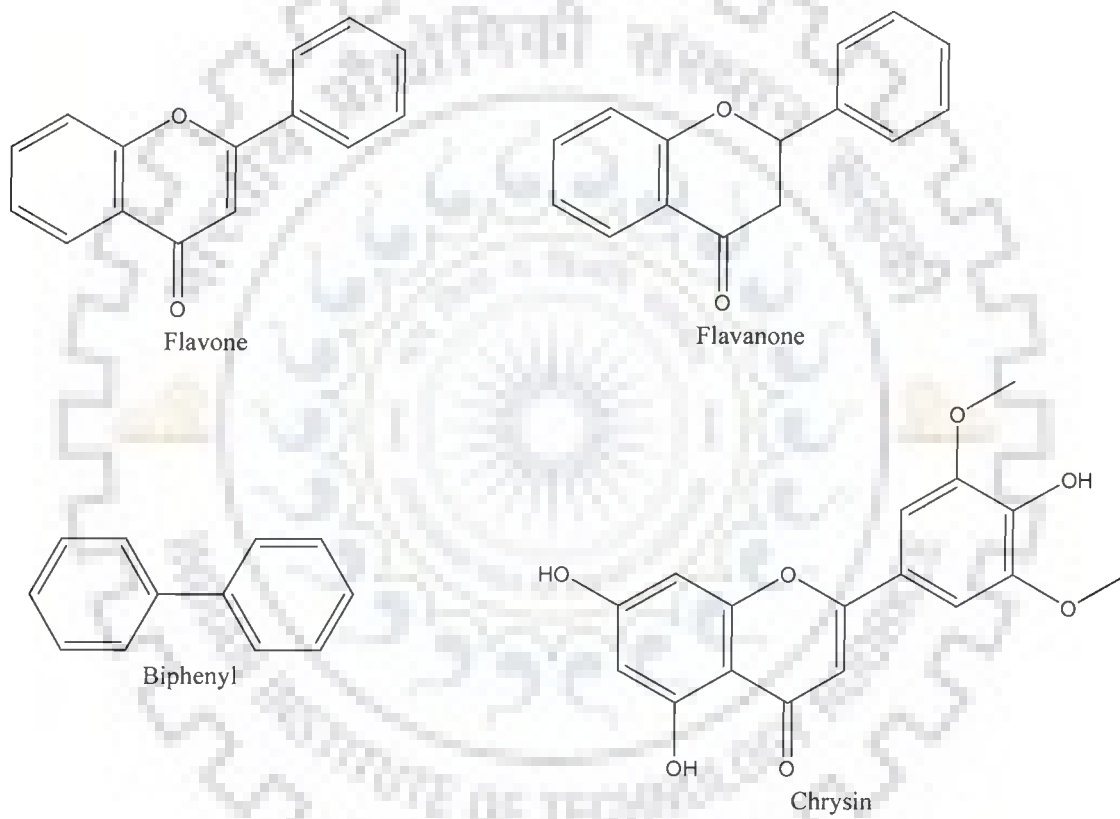


Fig. 2.4. Substrates for Biphenyl dioxygenases.

The *bphA1* gene has been used to convert various flavonoids, e.g., trans-chalcone, flavanone, flavones, 6-hydroxyflavanone, 6-hydroxyflavone, and 7-hydroxyisoflavone, into their respective vicinal diols when expressed in *E. coli* in combination with *bphEFG-bphB* from strain KF707 (Shindo *et al.*, 2003). However, it was not possible to convert 7-hydroxyflavone and 5,7-dihydroxyflavone (chrysin) with any unmodified and artificially-modified shuffled biphenyl dioxygenases including *BphA1* (Kagami *et al.*, 2008). Improved biphenyl dioxygenases were then generated which could convert 7-hydroxyflavone and chrysin, by hybrid formation among seven different *bphA1* genes followed by site-directed mutagenesis of amino-acid positions 324 and 325 (Kagami *et al.*, 2008).

The biphenyl dioxygenase-catalysed dihydroxylation of carbocyclic arenes, to yield the corresponding enantiopure cis-dihydrodiol metabolites, has provided a new pool of chiral compounds for use in chemoenzymatic synthesis. Recent reviews (Boyd and Sheldrake, 1998; Johnson, 2004; Boyd and Bug, 2006) list many examples of their applications in the synthesis of natural products, e.g. alkaloids and sugars, and non-natural products, e.g. carbasugars and isomeric dihydrodiols. The corresponding cis-dihydrodiol metabolites, derived from bicyclic azaarenes, and, more recently, monocyclic azaarenes, are also now available (Misawa *et al.*, 2002; Boyd *et al.*, 2008). It is noteworthy that both toluene dioxygenase (TDO) and biphenyl dioxygenase (BPDO) enzymes strongly favour catalytic cis-dihydroxylation of the carbocyclic rings rather than the pyridine rings in azaarenes. Although an increasing range of cis-dihydrodiol metabolites derived from mono- (Misawa *et al.*, 2002; Boyd *et al.*, 2008), di- (Boyd *et al.*, 1993; Boyd *et al.*, 2002), tri- (Boyd *et al.*, 2008, 1999; Shindo *et al.*, 2001) and tetracyclic (Boyd *et al.*, 1999) azaarenes have been reported, up to the present, few have been used as synthetic precursors (Boyd *et al.*, 2008).

### **2.1.6. Presence of metal in the catalytic centre**

Apart from the Rieske binding domain, the large subunit of the oxygenase component of ROs, contains the catalytic domain that has the nonheme iron. The production of cis-diols occurs at the catalytic mononuclear iron and is coordinated by two histidines and a single aspartic acid. The 2-his-1-carboxylate facial triad is found in many enzymes (Que, 2000). This motif is conserved in all known ROs (Ferraro et al., 2006). The mononuclear iron provides the platform for catalysis. Studies have revealed that the mononuclear iron binds to molecular oxygen in a side-on fashion, while the aromatic substrate binds in the active site above the iron-oxygen complex. Structural studies of ROs have demonstrated that the substrate binds in the active site with the atoms that would be oxidized closest to the mononuclear iron (Friemann et al., 2005; Furusawa et al., 2004). Coordination of this catalytic iron appears variable, involving between 4-6 ligands (Furusawa *et al.*, 2004; Dong *et al.*, 2005; Nojiri *et al.*, 2005) and is influenced by diverse factors, including RO dependent variation in the number and nature of the protein ligands, labile water ligands, presence or absence of substrate, as well as the influence of variations in the redox state of the rieske cluster. In cumene dioxygenase, nonheme iron in the active site was coordinated by His234, His240, and Asp388. Asp388 underwent monodentate coordination (Dong et al., 2005). In naphthalene dioxygenase, however, the corresponding Asp362 residue has been reported to undergo bidentate coordination (Kauppi *et al.*, 1998). At the positions corresponding to the sites of dioxygen in the binary and ternary complexes of naphthalene, electron density corresponding to two light atoms, such as oxygen, was observed. In nitrobenzene dioxygenase from *Comamonas* strain JS765, the mononuclear iron is coordinated by one carbonyl oxygen atom of Asp360, the NE atoms of His206 and His213 and two water

molecules (Friemann *et al.*, 2005). In the case of biphenyl dioxygenase from *Sphingomonas* strain B1, the mononuclear iron is coordinated by two histidines, His-207 and His-212 and one aspartate, Asp-360, at the rear of the active site. It was opined that water or molecular oxygen could be the fourth ligand on the iron (Ferraro *et al.*, 2007).

### 2.1.7. Regioselectivity of Biphenyl Dioxygenase towards different congeners

The three symmetrical 2,2', 3,3'-, and 4,4'-dichlorobiphenyls are not metabolized equally well by various BPDOs. *Comomonas. testosteroni* B-356 and *Rhodococcus globerulus* P6 BPDOs (hereafter referred to as BPDO<sub>B356</sub> and BPDO<sub>P6</sub> respectively), poorly catalyze the oxygenation of *para*-substituted congeners and degrade fairly well 3,3'-CB (Chebrou *et al.*, 1999; Hurtubise *et al.*, 1998; Barriault *et al.*, 2002). BPDO<sub>B356</sub> oxygenates 3,3'-dichlorobiphenyl much faster and more efficiently than it does 2,2'-dichlorobiphenyl and 4,4'-dichlorobiphenyl (Hurtubise *et al.*, 1998). On the other hand, the catalytic activity of BPDO from LB400 (hereafter referred to as BPDO<sub>LB400</sub>) toward 2,2'- dichlorobiphenyl is much greater than that of BPDO<sub>B356</sub> (Hurtubise *et al.*, 1998), but 3,3'-dichlorobiphenyl and 4,4'-dichlorobiphenyl are poor substrates for BPDO<sub>LB400</sub> (Mondello *et al.*, 1997). This enzyme, in fact, is among the few known BPDOs of natural occurrence that can efficiently metabolize the symmetrical *ortho*-substituted chlorobiphenyl 2,2'-CB and the symmetrical *ortho-meta*-substituted congeners 2,3,2',3'-CB and 2,5,2',5'-CB (Mondello *et al.*, 1997; Barriault *et al.*, 2004). BPDO<sub>LB400</sub> is regarded to be among the best characterized oxygenase and the one exhibiting the highest PCB degrading potency among the other BPDOs of natural occurrence. BPDO<sub>KF707</sub> is closely related to BPDO<sub>LB400</sub> (>98% amino acid identity) but their PCB degrading potency differs considerably. Among the differences, BPDO<sub>KF707</sub> oxygenates

4,4'-dichlorobiphenyl much more efficiently than BPDO<sub>LB400</sub> (Personal communication, Michael Sylvestre). The absence of 3,4 oxygenase activity in KF707 as compared to LB400 is an indication of significantly different functionality between the two oxygenases. All but 1 of the 21 amino acid differences between the dioxygenases of LB400 and KF707 are clustered within the 140-aminoacid region in the carboxy terminal half of the BphA subunit (Erickson and Mondello 1993). Minor structural variations on the C-terminal portion of the  $\alpha$  subunit of the oxygenase component are responsible for changes in specificities towards new substrates including higher chlorinated PCBs, chlorodibenzofurans, and chlorinated aliphatic solvents (Mondello *et al.*, 1997; Suenaga *et al.*, 2001; Barriault *et al.*, 2002; Mohammadi and Sylvestre, 2005). Studies showed that a stretch of seven amino acids (called region III) strongly influences the capacity of the enzyme to catalyze the oxygenation of wide range chlorobiphenyls and chlorodibenzofurans (Barriault and Sylvestre, 2004) including 4,4'-dichlorobiphenyl (Mondello *et al.*, 1997).

BPDOs are currently classified into two distinct types, KF707 and LB400, based on their substrate specificities for various PCB congeners). As already mentioned, double *para*-replaced congeners such as 4, 4' -dichlorobiphenyl are oxidized by KF707-type but not LB400-type BPDOs, while the reverse is true for *ortho-meta*-replaced congeners, such as 2, 5, 2', 5'-tetrachlorobiphenyl . BPDO<sub>RHA1</sub> was classified as a KF707-type based on its ability to degrade 4, 4' -dichlorobiphenyl, but not 2, 5, 2', 5'-tetrachlorobiphenyl . Mondello and coworkers (1997) identified four regions that are believed to confer substrate specificity between these two types of BPDOs. In order to determine how differences in BphA sequence influence PCB substrate specificity, a database of structure-function relationships was constructed by them for biphenyl dioxygenase enzymes. The database contained sequence

and congener specificity data for a total of 15 strains. The PCB-degradative abilities of the strains were found to fall into two categories. Six of the strains displayed a broad substrate specificity similar to that of LB400. Typically, these organisms degraded relatively highly chlorinated PCBs containing up to six chlorines per biphenyl. Congeners with chlorines at the 2 and 5 positions on the ring were readily attacked by these strains; however, they had relatively poor activity against congeners with chlorines at both *para* positions (e.g., 4,4'-CB and 2,4,4'-CB). All of the strains categorized as having broad substrate specificity extensively attacked a wide variety of tetrachlorobiphenyls, including 2, 5, 2', 5'-CB, which are oxidized at the 3, 4 position by LB400 and H850. This 3, 4-dioxygenase activity was found to be relatively rare among PCB-degrading bacteria, and no known strains lacking this ability were able to degrade 2, 5, 2',5'-CB. The degradation of this congener therefore served to distinguish bacteria that have LB400-type specificity from those that do not.

BPDO<sub>B356</sub> has been designated to be of type KF707 because it has a greater reactivity against several recalcitrant congeners, including 4, 4' -dichlorobiphenyl, relative to BPDO<sub>LB400</sub>. Strains designated as having a KF707 - type specificity attack a relatively narrow range of PCB substrates and are generally able to degrade only lightly chlorinated congeners. Many of them have poor activity against tetrachlorobiphenyls with chlorines at the 2 and 5 ring positions. None of these organisms were able to degrade 2, 5, 2',5'-CB or any of the pentachlorobiphenyls tested. In most cases, however, these organisms were better able to degrade the di-*para*-substituted congeners 4,4'-CB and 2,4,4'-CB than those of the LB400 group. From the work of Mondello and coworkers, several areas of the BphA proteins were identified in which broad or narrow substrate ranges were consistently associated with specific amino acid sequences (Mondello et al., 1997). In the first of these areas, located at

amino acid residues 237 and 238 of BPDO<sub>LB400</sub> (designated region I), each of the strains with a broad substrate range contains threonine at both positions. All of the strains with a narrow substrate range contain methionine at position 237 and serine (or alanine) at position 238. Two other very consistent associations between specificity and sequence occurred at amino acids 277 (region II) and 377 (region IV). All of the broad-specificity strains contained tyrosine and asparagine, respectively, at these sites, while all of the strains with narrow specificity contained phenylalanine and threonine. An association similar to that observed in region IV was also observed at position 338, within an area designated region III. Region III included the four amino acids at sites 335, 336, 338, and 341 which differed between LB400 and KF707. Broad specificity was associated with the amino acid asparagine at position 336, while a narrow substrate range was associated with threonine. But it was also found that there was indeed an influence of residues outside region III on the range of PCB substrates that BPDOs can oxygenate.

So BPDO<sub>LB400</sub> and BPDO<sub>KF707</sub> have been extensively engineered to improve their capabilities for environmental pollutant degradation by using various techniques, such as random mutagenesis, in vitro DNA shuffling, and subunit or domain exchange (Furukawa 2000; Suenaga *et al.*, 2001; Suenaga *et al.*, 2002; Zielinski *et al.*, 2002; Zielinski *et al.*, 2003). Recently a central part (amino-acid position 268–397 of 458 amino-acid residues) of the biphenyl dioxygenase large ( $\alpha$ ) subunit, BphA1 from KF707 was exchanged with the corresponding part of BphA1 from another biphenyl-degrading bacterium, *Pseudomonas putida* strain KF715, to construct hybrid BphA1, BphA1 (715-707). When expressed in *Escherichia coli* together with the bphA2A3A4BC genes from strain KF707, this enzyme was shown to possess activity for degrading both 1-phenylnaphthalene and 2-

phenylnaphthalene. Between central parts of BphA1 from strains KF707 and KF715, the difference of amino-acid residues resided only in position 324–325. Therefore the attempt was aimed at improving the substrate preference of BphA1 by applying random amino-acid substitutions at these positions to BphA1 (715-707). For members of the toluene/ biphenyl subfamily, however, crystallization of only one enzyme (BPDO from *Burkholderia* sp. strain RHA1 [BPDO RHA1]) has been reported (Nagarajan et al., 2003; Furusawa et al., 2004).

#### 2.1.8. The Generation of Variants BPDO<sub>119</sub> and BPDO<sub>1152</sub>

The PCB-transforming capabilities of the *bph* pathway are, as mentioned above, strain-dependent. Nevertheless, even the highly potent strains poorly transform congeners containing more than four chloro substituents and none completely degrades all lightly chlorinated congeners. For example, one of the best PCB-degrading strains, BPDO<sub>LB400</sub> possesses the unusual ability to transform 2, 2', 5, 5'-tetrachlorobiphenyl; however, it is transformed to the corresponding 3,4-dihydroxybiphenyl, a dead-end metabolite.

Efforts to improve the PCB-degrading capabilities of natural isolates have revolved around the engineering of *bph* enzymes to transform a broader range of congeners and the introduction of downstream pathways to degrade the chlorobenzoates produced by biphenyl pathway (Gibson and Parales, 2000). Most of the engineering has focused on the catalytic component of biphenyl dioxygenase (*bphA1-2*), the first enzyme of the pathway (Barriault and Sylvestre 2004; Furukawa *et al.*, 2004; Suenaga *et al.*, 2001). The large subunit of biphenyl dioxygenase contains the principal determinants of the enzyme substrate specificity. Therefore, the corresponding genes from divergent bacterial strains have been subjected to



DNA shuffling, generating variants possessing improved rates of transforming certain PCB congeners (Barriault *et al.*, 2002; Kumamaru *et al.*, 1998; Fortin *et al.*, 2005).

To overcome the hurdle of selective degradation of differently halogenated PCBs by different bacteria and increase the number of possible substrates for them, it is clear that newer approaches and their detailed analyses are essential to completely identify all features that influence substrate specificity. Engineering efficient biocatalysts for the degradation of biphenyl analogs will require that the turnover rate of reaction and regiospecificity of the enzyme toward these substrates is efficiently altered. These features are determined through interactions between amino acid residues and the carbons of the aromatic ring or the ring substituents. The configuration of the molecule such as co-planarity or stereoisomerism should also influence the catalytic reaction. Identification of the amino acids that interact with the substrate molecule and their mode of interaction should help design strategies to engineer evolved enzymes exhibiting desirable catalytic features.

There is 76% amino acid sequence identity between the B-356 and LB400 BphAs (broadly it is LB400 BphA that is referred to as  $BPDO_{LB400}$ , B-356 BphA that is referred to as  $BPDO_{B356}$  and P6 BphA1 that is referred to as  $BPDO_{P6}$  in the present work for it is the sequence of this constituent of the BPDO system that is chiefly engineered), 65% amino acid sequence identity between  $BPDO_{P6}$  and  $BPDO_{LB400}$ , and 64% amino acid sequence identity between  $BPDO_{P6}$  and  $BPDO_{B356}$ ). Variant BPDOs that had inherited the catalytic features of both parents and exhibiting extended activities toward several *tetra*- and *penta*-substituted congeners were obtained by shuffling the very closely related LB400 *bphA* with KF707 *bphA1* (Bruhlmann and Chen, 1999; Kumamaru *et al.*, 1998; Barriault *et al.*, 2002). Family

shuffling of genes of lesser homology increases the sequence diversity of the library, resulting in an accelerated rate of enzyme functional improvement (Cramer *et al.*, 1998; Barriault *et al.*, 2002). It is also a powerful tool to identify the major structural features of a protein family that confer a desired phenotype. However, as the sequence space to explore increases, a larger proportion of the progeny members are prone to be inactive (Voigt *et al.*, 2001), which implies that the screening assay needs to be strongly selective or based on a microarray design. In order to reduce the size of the library that needs to be explored in family shuffling of *bphA* of lesser homology, Barriault *et al.* (2002) targeted a portion of BphA that is critical for substrate specificity and selectivity. Structural features of the C-terminal portion of BphA influence the regioselectivity and regiospecificity of the enzyme (Bruhlmann and Chen, 1999; Kumamaru *et al.*, 1998; Mondello *et al.*, 1997). Furthermore, active BphA hybrids were recently obtained by replacing long stretches of LB-400 *bphA* encoding the C-terminal portion of the protein by the corresponding stretches of B-356 *bphA* (Barriault *et al.*, 2001).

Barriault *et al.* (2002) carried out an investigation towards shuffling targeted stretches of *bphA* genes of lesser homology in order to obtain better-performing BphA variants and to identify some of the major structural features of the C-terminal portion of BphA that contribute to relax the enzyme toward PCB congeners (Barriault *et al.*, 2002) (Fig. 2.5). In strategy I (Fig. 1), a 1,043-bp fragment of LB400 *bphA* amplified with primers external to the *Mlu*I and *Avr*II restriction sites was digested with DNase I and reassembled by primerless PCR in the presence of a DNase I-digested 513-bp fragment of B-356 *bphA*. The 1,043-bp fragment was amplified from pDB31[LB400-*bphA*], and it stretched from bp 368 of LB400 *bphA* to 30 bp downstream of the stop codon (in pDB31); the 513-bp fragment was amplified

from pDB31[B-356-*bphA*], and it extended from bp position 368 of B-356 *bphA* to the *ScaI* restriction site. In strategy II and BPDO<sub>II9</sub> was generated by this method, the same 1,043-bp DNA fragment of LB400 *bphA* was digested with DNase I and reassembled by primerless PCR in the presence of a DNase I-digested 717-bp fragment of B-356 *bphA*. The 717-bp fragment was amplified from pDB31[B-356-*bphA*], and it started at position 688 of B-356 *bphA* and ended 30 bp downstream of the stop codon in pDB31. In strategy III, the DNase I-digested fragments used to reassemble the 1,037- to 1,043-bp fragments included a 723-bp fragment amplified from pDB31[LB400-*bphA*] (from position 688 of LB400-*bphA* to 30 bp downstream of the stop codon), a 720-bp fragment amplified from pQE31[P6-*bphA1*] (from position 664 of P6 *bphA1* to the stop codon), and the 513-bp fragment of B-356 *bphA* that extended from bp position 368 of B-356 *bphA* to the *ScaI* restriction site. Variants generated by strategy III differed from parental enzymes. BPDO<sub>III52</sub> was obtained by this strategy. The reassembled 1,037- to 1,043-bp fragments obtained from each of the primerless PCR were amplified with appropriate internal primers to construct the libraries of *MluI-AvrII* fragments that were ligated to pDB31 [LB400-*bphAE*] previously deleted of its *MluI-AvrII* segment. The plasmid libraries were introduced into *E. coli* DH11S/pQE51 [LB400-*bphFGB*] to screen for clones exhibiting a similar or better degrading capacity than the parental LB400 BphA when exposed to ortho substituted bichlorinated biphenyl. It was observed that in a coupled reaction with BphB and BphC, BPDO<sub>II9</sub> was able to catalyze the oxygenation of benzene, naphthalene, and toluene. BPDO<sub>II9</sub> differed from the parental strain BPDO<sub>LB400</sub> only at amino acid residue positions 335, 336, 338, 341. The variations involved Thr335Gly, Phe336Ile, Asn338Thr and Ile341Thr. As for BPDO<sub>III52</sub>, at residue positions 178, 184, 188, 191 BPDO<sub>III52</sub> resembled BPDO<sub>B356</sub>. It had inherited the entire region III (Mondello *et*

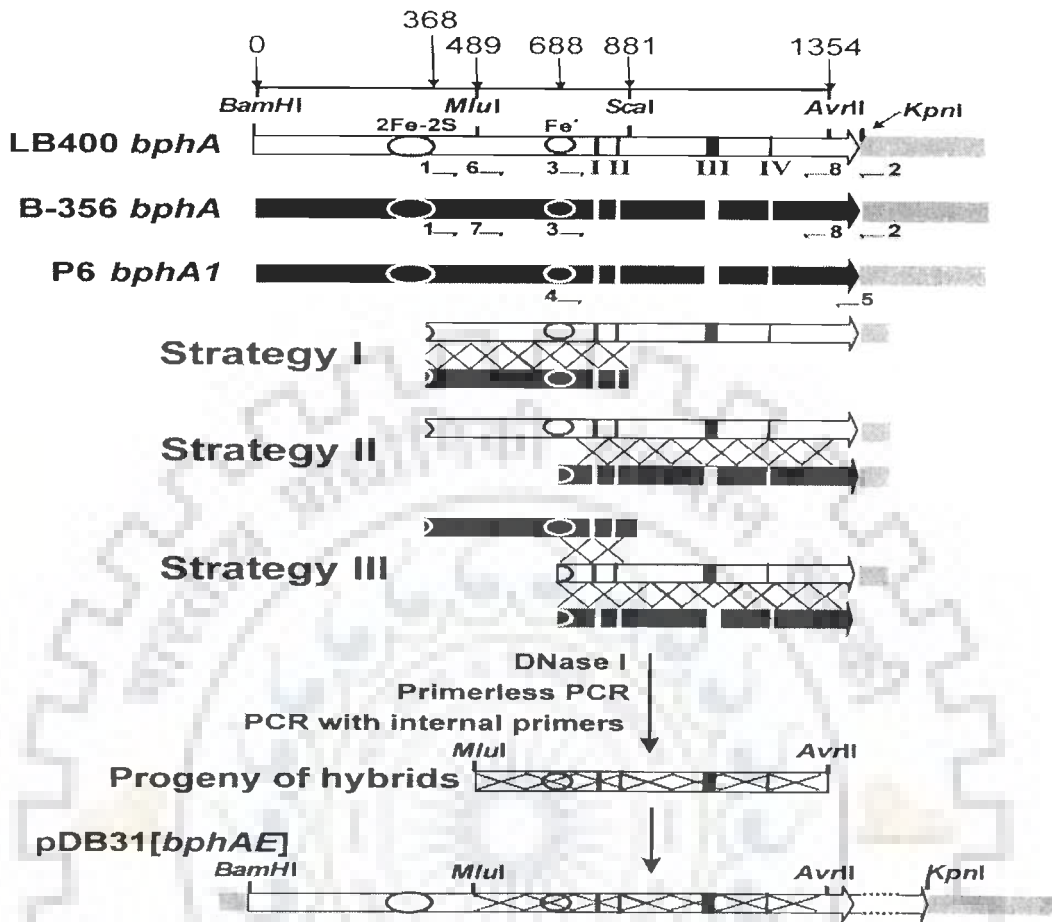


Fig. 2.5. Strategies to generate variants with better PCB degrading capability. The constructions used for shuffling are shown in the top part of the figure. These are LB400 and B-356 *bphA* cloned in pDB31 and P6 *bphA1* cloned in pQE31. Large arrows represent genes; the gray box on the right side of the genes represents the vectors. Portions of LB400, B-356 *bphA*, or P6 *bphA1* were PCR amplified with appropriate primers (indicated as small horizontal arrows) as described in the text. The appropriate DNA fragments for each of the three shuffling strategies were digested with DNase I and reassembled by primerless PCR. Primerless PCR products were amplified with internal primers (primers 6 or 7 and primer 8) to generate the libraries of *MluI*/*AvrII* hybrid fragments that were used to replace the corresponding fragment of LB400 *bphA* in pDB31[LB400-*bphAE*]. The numbers on top of the figure indicate the base pair positions in LB400 *bphA*. The indicated restriction sites are unique and common to both LB400 and B-356 *bphA*. Circles indicate the position of the 2Fe-2S Rieske center and the mononuclear Fe<sub>o</sub> active center. I, II, III, and IV refer to the localization of the designated *bphA* regions I, II, III, and IV, respectively (Mondello *et al.*, 1997). Figure adapted from Barriault *et al.*, 2002.



Table 2.1. Sequence identity (%) shared by variants BPDO<sub>I19</sub> and BPDO<sub>III52</sub> with other Rieske oxygenases (BPDO: Biphenyl dioxygenase; CDO: Cumene dioxygenase; NDO: Naphthalene dioxygenase; NBDO: Nitrobenzene dioxygenase, OxMo: Oxoquinoline monooxygenase. PDB codes are mentioned in parenthesis)

Terminal oxygenase $\alpha$ Subunit(PDB ID)	BPDO <sub>I19</sub>	BPDO <sub>III52</sub>	BPDO LB400	BPDO B-356	CDO <sub>IPO1</sub> (1WQL)	BPDO <sub>RHA1</sub> (1ULI)	BPDO <sub>B1</sub> (2GBW)	NDO <sub>12038</sub> (2B1X)	NDO <sub>9816-4</sub> (1NDO)	NBDO (1WW9)	BPDO KF707	OxMo <sub>ppu</sub> (1ZO3)	NBDO <sub>JS765</sub> (2BMQ)
BPDO <sub>I19</sub>	100	97	99	77	74	66	34	32	30	14	96	9	30
BPDO <sub>III52</sub>	97	100	98	78	74	66	34	31	30	12	93	4	30
BPDO LB400	99	98	100	76	73	65	34	31	30	13	95	7	29
BPDO B-356	77	78	76	100	75	64	34	30	31	13	77	8	28
CDO <sub>IPO1</sub> (1WQL)	74	74	73	75	100	64	31	30	30	12	74	11	30
BPDO <sub>RHA1</sub> (1ULI)	66	66	65	64	64	100	35	31	30	8	65	12	30
BPDO <sub>B1</sub> (2GBW)	34	34	34	34	31	35	100	31	42	8	35	5	40
NDO <sub>12038</sub> (2B1X)	32	31	31	30	30	31	31	100	29	10	32	6	27
NDO <sub>9816-4</sub> (1NDO)	30	30	30	31	30	30	42	29	100	9	28	6	81
NBDO (1WW9)	14	12	13	13	12	8	8	10	9	100	14	40	11
BPDO KF707	96	93	95	77	74	65	35	32	28	14	100	7	28
OxMo <sub>ppu</sub> (1ZO3)	9	4	7	8	11	12	5	6	6	40	7	100	9
NBDO <sub>JS765</sub> (2BMQ)	30	30	29	28	30	30	40	27	81	11	28	9	100

*et al.* 1997) of BPDO<sub>P6</sub> (P6 BphA1) (<sup>333</sup>GINTIRT<sup>339</sup>) and resembled BPDO<sub>II9</sub> in that respect. Table 2.1 shows the sequence similarity percentages between the variants under study, BPDO<sub>II9</sub> and BPDO<sub>III52</sub> and other Rieske oxygenases. BphA hybrids resulting from shuffling were found to have extended their PCB degrading capacity (Barriault *et al.* 2002). The numerous variants generated were screened for whether their PCB degrading capacity increased towards the variously substituted congeners or not. Fig. 2.6 shows the sequence alignment of the variants generated by these strategies.

Thus, BPDO<sub>II9</sub> was generated by the replacement of a stretch of LB400 *bphA* extending from bp positions 978 to 1028 by the corresponding fragment of B-356 *bphA* so that BPDO<sub>LB400</sub><sup>335</sup> TFNNIRI<sup>341</sup> was replaced by <sup>333</sup> GINTIRT <sup>339</sup> designated region III by Mondello *et al.* (1997). BPDO<sub>II9</sub> was more relaxed and exhibited a higher PCB degrading potency as compared to both LB400 and B-356. BPDO<sub>III52</sub> was made by replacing corresponding gene coding for amino acids 166-211 of BPDO<sub>P6</sub> with those of BPDO<sub>B356</sub>. Apart from effectively being able to use biphenyl as substrate, the enzyme was also seen to be able to oxygenate efficiently other biphenyl analogs such as naphthalene and toluene.

### 2.1.9. Specific Aims of Study on Biphenyl Dioxygenase

1. To carry out crystallization of BPDO<sub>II9</sub> (variant II-9) and BPDO<sub>III52</sub> (variant III-52).
2. Crystal structures determination of BPDO<sub>II9</sub> and BPDO<sub>III52</sub> and complex with substrate.
3. Analysis of structure to elucidate the mechanism for the better PAH and PCB degrading abilities of these organisms.

## 2.2. 3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE

### 2.2.1. The Shikimate Pathway

The shikimate pathway is the biosynthetic pathway essential for production of a plethora of aromatic compounds in plants, bacteria, and fungi. Seven enzymes of the shikimate pathway catalyze sequential conversion of erythrose 4-phosphate and phosphoenol pyruvate to chorismate (Fig. 2.7). Chorismate is then used as a substrate for other pathways that culminate in production of folates, ubiquinone, naphthoquinones, and the aromatic amino acids tryptophan, phenylalanine, and tyrosine. The shikimate pathway (SP) is absent from animals and present in the apicomplexan parasites *Toxoplasma gondii*, *Plasmodium falciparum*, and *Cryptosporidium parvum*. Inhibition of the pathway by glyphosate is effective in controlling growth of these parasites. These findings emphasize the potential benefits of developing additional effective inhibitors of the shikimate pathway. Such inhibitors may function as broad-spectrum antimicrobial agents that are effective against bacterial and fungal pathogens and apicomplexan parasites. The shikimate pathway functions at a critical interface between primary and secondary metabolism by channeling carbon from glycolysis and the pentose phosphate pathway towards the synthesis of a broad range of physiologically important aromatic compounds (Bentley, 1990). In plants these include the aromatic amino acids, phenylpropanoids, lignins, hormones, pigments, phytoalexins, alkaloids, UV protectants, and electron carriers (Birck and Woodard, 2001). Metabolites of the main trunk of the shikimate pathway are also considered branch point substrates for other secondary metabolic pathways (Hermann *et al.*, 1999, Fucile *et al.*, 2008).



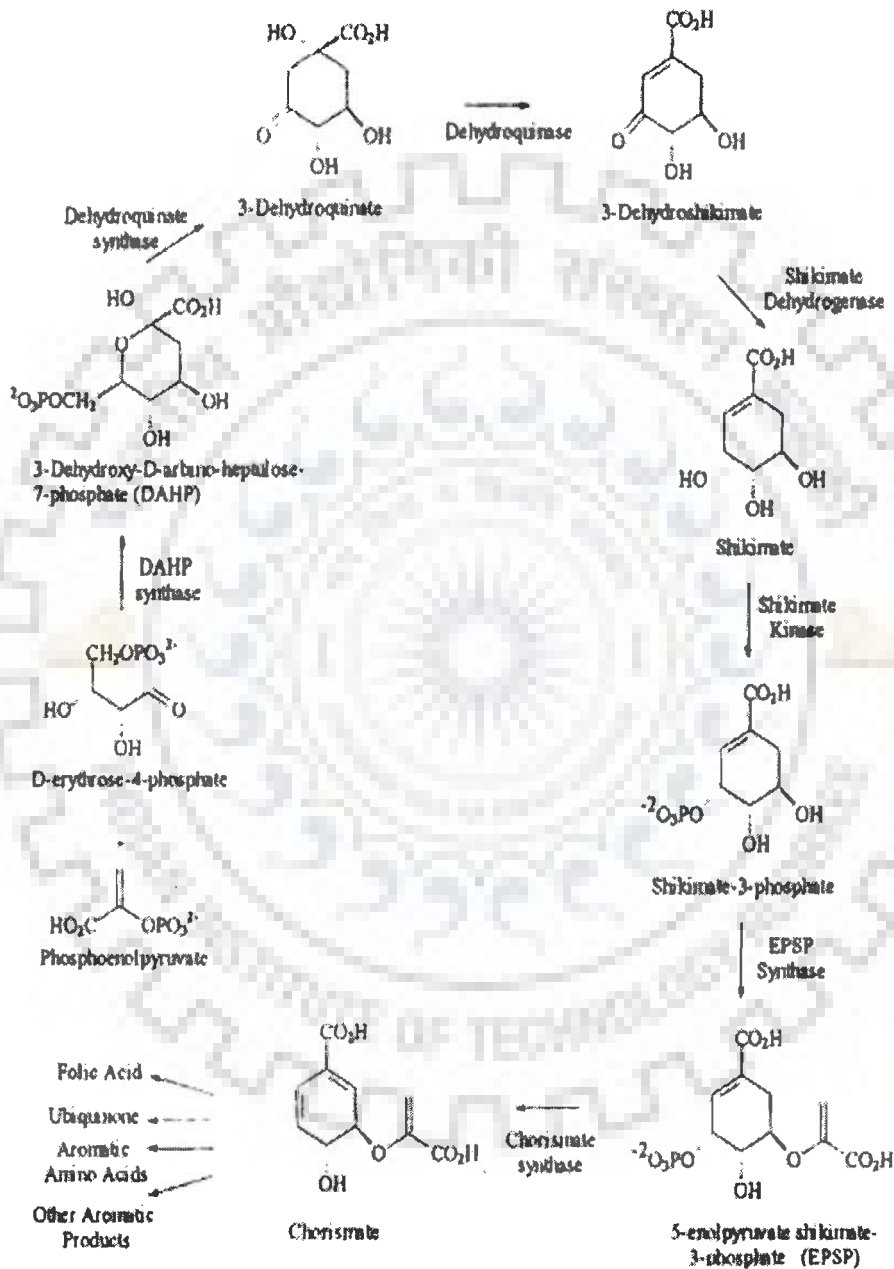


Fig. 2.7. The Shikimate Pathway.

In plants and microorganisms, all the key aromatic compounds, involved on the primary metabolism, are produced by SP. Twenty percent of the carbon derived from carbohydrate catabolism is utilized in the shikimate pathway in certain organisms; therefore, it is a major part of their metabolism. This pathway is essential for algae, higher plants, bacteria, fungi, apicomplexan parasites, whereas it is absent in mammals (Guo and Frost, 2002). It was discovered recently in the apicomplexan parasites *Plasmodium falciparum*, *Toxoplasma gondii*, and *Cryptosporidium parvum* (Roberts *et al.*, 1998; Campbell *et al.*, 2004). Therefore, these enzymes are potential targets for the development of non-toxic anti-microbial agents (Hartmann *et al.*, 2003; Hermann and Weaver, 1999). Also, the shikimate pathway is a viable herbicide target. Specifically, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase, the sixth enzyme in the pathway, is the target for glyphosate, a widely used herbicide (Roundup; Sigma) (Bentley, 1990, Herrmann and Weaver, 1999). Several enzymes of the SP have been submitted to structural studies (Hoffmann *et al.*, 1972; Howe *et al.*, 2003; Jensen *et al.*, 2002; Karlin *et al.*, 1992; König *et al.*, 2004).

All enzymes of this pathway have been obtained in pure form from prokaryotic and eukaryotic sources and their respective DNAs have been characterized from several organisms. The cDNAs of higher plants encode proteins with amino terminal signal sequences for plastid import, suggesting that plastids are the exclusive locale for chorismate biosynthesis. In microorganisms, the shikimate pathway is regulated by feedback inhibition and by repression of the first enzyme. In higher plants, no physiological feedback inhibitor has been identified, suggesting that pathway regulation may occur exclusively at the genetic level. This difference between microorganisms and plants is reflected in the unusually large variation in the primary structures of the respective first enzymes. Several of the pathway

enzymes occur in isoenzymic forms whose expression varies with changing environmental conditions and, within the plant, from organ to organ.

The penultimate enzyme of the pathway is the sole target for the herbicide glyphosate. Glyphosate-tolerant transgenic plants are at the core of novel weed control systems for several crop plants (Hermann and Weaver, 1999).

Bacteria spend more than 90% of their total metabolic energy on protein biosynthesis. Consequently, the bacterial shikimate pathway serves almost exclusively to synthesize the aromatic amino acids. In contrast, higher plants use these amino acids not only as protein building blocks but also, and in even greater quantities, as precursors for a large number of secondary metabolites, among them plant pigments, compounds to defend against insects and other herbivores (Dixon and Palva, 1995), UV light protectants, and, most importantly, lignin. Under normal growth conditions, 20% of the carbon fixed by plants flows through the shikimate pathway (Haslam, 1993). Globally, this amounts to about  $7 \times 10^{15}$  kg each year and most of it is used for the synthesis of the various secondary metabolites. The variation in shikimate pathway-derived secondary metabolites is substantial among plant species. The secondary metabolite makeup of a plant may be used for species classification.

Different plants not only synthesize different aromatic secondary metabolites but also synthesize varying amounts of them at specific times and in specific subcellular compartments. One would expect that regulation of the differential biosynthesis of sometimes very complex molecular structures might involve regulation of the supply of the precursors influencing the rate-limiting step for carbon flow through the shikimate pathway. Recent data on transgenic potatoes give some indication that this is indeed the case (Jones *et*

*al.*, 1995). The first step of the pathway is catalysed by DAHP synthase which will be taken up in detail after the other enzymes of the pathway in view of the enzyme being the focus of the current work.

## **2.2.2. Enzymes of the Shikimate Pathway**

### **2.2.2.1. 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase**

3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP) synthase is the first enzyme of the shikimate pathway that catalyses the condensation of phosphoenolpyruvate and erythrose-4-phosphate into 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP). This enzyme controls the flow of carbon into the shikimate pathway. The enzyme was discovered in *E.coli* three decades ago and has been isolated and purified from a number of sources but there is no plant DAHP synthase structure that has been elucidated so far. It is this enzyme that is the focus of the current study.

### **2.2.2.2. 3-dehydroquinate synthase**

The second enzyme of the shikimate pathway, 3-dehydroquinate synthase, catalyses the elimination of phosphate from DAHP to generate 3-dehydroquinate (DHQ). It requires catalytic amounts of NAD<sup>+</sup> and a divalent cation for activity. For the bacterial enzyme, Co<sup>2+</sup> and Zn<sup>2+</sup> are the most active metal ions. The *E. coli* enzyme (Herrmann and Weaver, 1999) catalyzes a seemingly very complex reaction involving an intramolecular oxidation reduction at C5 of DAHP with very tight binding of the NAD<sup>+</sup> cofactor, the elimination of phosphate, and an alicyclic ring formation (Bender *et al.*, 1989). However, it has been speculated that the enzyme may merely catalyze the oxidation and the reduction, with the other identified

steps in the overall reaction mechanism proceeding spontaneously. Unlike DAHP synthase, potato DHQ synthase activity does not change when cells are exposed to glyphosate (Pinto *et al.*, 1988). Plant DHQ synthases have been purified from *Phaseolus mungo* and *Pisum sativum*. Plant DHQ synthase is more closely related to the bacterial than to the fungal enzyme, unlike the case for DAHP synthase (Herrmann and Weaver, 1999).

### 2.2.2.3. 3-Dehydroquinate Dehydratase-Shikimate Dehydrogenase

The third step of the shikimate pathway, dehydration of DHQ to give 3-dehydroshikimate (DHS), is catalysed by DHQ dehydratase that exists in two forms distinguished on the basis of mechanistic differences: type I and type II in different bacteria. Type I DHQ dehydratase catalyses syn elimination and type II anti-elimination of water. The 3-dehydroquinate dehydratases of most bacteria and higher plants are type I enzymes. In the *E. coli* 3-dehydroquinate dehydratase, a histidine residue and a lysine residue are in the active site of the enzyme. A type I enzyme has been crystallized from *S. typhimurium*; the crystals refract to 2.3 Å resolution. A type II 3-dehydroquinate dehydratase from *Mycobacterium tuberculosis* has been crystallized; the crystals refract to 2.2-Å resolution (Gourley *et al.*, 1994).

The fourth step in the shikimate pathway is the reduction of DHS to shikimate. In *E. coli*, the reaction is catalysed by a NADP– dependent shikimate dehydrogenase. In higher plants, the third and fourth steps of the shikimate pathway are catalyzed by a bifunctional enzyme DHQ dehydratase-shikimate dehydrogenase (Koshiba, 1978; Polley, 1978). In the moss *Physcomitrella patens*, spinach, pea, and tobacco, the 3-dehydroquinate dehydratase and shikimate dehydrogenase activities reside on a single polypeptide (Mousedale *et al.*, 1987).

Such bi- or even multifunctional enzymes are found in several amino acid biosynthetic pathways, but homologous reactions are sometimes structurally organized in different ways in different organisms. For example, in *Neurospora crassa* (Case and Giles, 1968), *Aspergillus nidulans* (Charles *et al.*, 1986) and *Saccharomyces cerevisiae* (Duncan *et al.*, 1987), the second through sixth reactions of the shikimate pathway are catalyzed by a pentafunctional polypeptide encoded by the *arom* gene. In the bacteria *E. coli* and *Salmonella typhimurium*, the seven enzymes of the shikimate pathway are all individual polypeptides. The dehydratase activity is hosted in the amino terminal half of the polypeptide. A cDNA has been obtained for *Nicotiana tabacum* homologue. This sequence codes for a protein with the activity of dehydratase in the amino terminal and the dehydrogenase in the carboxy terminal half. The sequences for both plant enzymes are more similar to bacterial than to lower eukaryotic homologues. Shikimate dehydrogenase has been widely used as marker in genetics and in crop breeding for many years. Presently 5 dehydroquinate dehydratase and 12 shikimate dehydrogenase 3D structures are known from archaeon, bacteria, fungi and plants (Herrmann, 1995).

#### **2.2.2.4. Shikimate Kinase**

In the fifth step of the pathway, a kinase phosphorylates shikimate to yield shikimate 3-phosphate. Elimination of this phosphate two steps later leads to the second double bond of the benzene ring. Plant shikimate kinases have been described from pea and rice, and the activity has been purified to near homogeneity from spinach chloroplasts and pepper (Herrmann and Weaver, 1999). The enzyme is localized in the chloroplast. Presently 28 3D structures are known for shikimate kinase, all from fungi and bacteria.

#### 2.2.2.5. 5-Enolpyruvyl shikimate 3-phosphate synthase

The penultimate step in the shikimate pathway constitutes the reversible formation of 5-enolpyruvylshikimate 3-phosphate (EPSP) and inorganic phosphate from shikimate 3-phosphate and phosphoenolpyruvate. The reaction is catalyzed by EPSP synthase, the best studied out of all the enzymes in the pathway. This enzyme has been purified from both microbial and plant sources (Duncan *et al.*, 1984; Mousedale and Coggins, 1984). The enzyme-catalyzed reaction mechanism has been studied in great detail (Anderson *et al.*, 1990). Site directed mutagenesis and nuclear magnetic resonance studies have placed a histidine residue very close to the active site of the enzyme (Shuttleworth and Evans, 1994). The *E. coli* enzyme was crystallized and the X-ray structure analyzed to 3 Å resolution (Stallings *et al.*, 1991). This analysis shows that EPSP synthase has a two-domain structure, with the active site presumably near the interdomain crossover segment. EPSP synthase is the unique target for the widely used broad-spectrum herbicide glyphosate (Steinrücken and Amrhein, 1980). Glyphosate tolerant cell lines from several different organisms have elevated levels of DAHP synthase. The elevation is due to gene amplification (Herrmann and Weaver, 1999). Plant DNAs encoding EPSP synthases have been isolated from *Petunia*, *Arabidopsis*, tomato and *Brassica napus* (Gasser and Klee, 1990). 21 structures have been solved through X-ray crystallography upto now for bacterial and fungal EPSP synthases.

#### 2.2.2.6. Chorismate Synthase

The seventh and last step of the shikimate pathway is the concerted IP-trans elimination of phosphate from EPSP to yield chorismate (Balasubramanian *et al.*, 1990; Hawkes *et al.*, 1990). Chorismate synthase, which catalyzes this reaction, requires a reduced flavin

nucleotide (FMNHz) as a cofactor, even though the overall reaction is redox neutral. The same is true of step two of the pathway, which is catalyzed by 3-dehydroquinate synthase, an enzyme that requires NAD<sup>+</sup> for catalytic activity. Spectroscopic analysis of cofactor function suggested a radical as an intermediate in the reaction catalyzed by chorismate synthase (Ramjee *et al.*, 1992). The *N. crassa* (Welch *et al.*, 1974) and *Fuglena gracilis* enzymes are bifunctional; an associated NADPH driven flavin reductase generates the reduced cofactor. About 30 crystal structures have been solved for chorismate synthases till now ranging from archaeon to bacterial to fungal enzymes.

### 2.2.3. Subcellular Localization of the Shikimate Pathway

In plants, proteins are synthesized in three different compartments: in the cytoplasm, in the plastids and in the mitochondria. Therefore the aromatic amino acids must either be synthesised *in situ* in the respective protein-synthesizing compartments or synthesized outside the compartments and have to be imported.

Early work revealed that the isolated chloroplasts from spinach were able to synthesize aromatic amino acids from CO<sub>2</sub> or shikimate (Bickel *et al.*, 1978). This indicates the existence of an aromatic biosynthetic pathway in chloroplasts. Moreover, the localization of the SP in the plastids is supported by molecular analysis. In higher plants, the shikimate pathway enzymes are normally synthesized as precursors with a N-terminal extension, presumably a plastid-specific transit peptide. In addition, *in vitro* experiments demonstrated that some shikimate pathway enzymes (e.g. shikimate kinase and EPSPS) can be imported into chloroplasts (Della-Cioppa, *et al.* 1986, Schmid, *et al.* 1992). Thus, the shikimate pathway appears to be firmly established in plastids. However, it has been proposed that the



chloroplast localized biosynthetic activity does not count for 100% of the observed aromatic amino acid biosynthesis. Hence, a spatially separated pathway in the cytosol might exist. The dual pathway hypothesis proposed a plastid-located shikimate pathway responsible for the production of aromatic amino acids and a cytosolic pathway responsible for the production of secondary metabolites, because enzymes of phenylpropanoid and flavonoid biosynthesis are not present in plastids. The evidence for a dual shikimate pathway is further supported by isozyme analysis in subcellular fractions. Putative cytosolic isozymes have been described for DAHP synthase, DHD/SHD (Mousdale *et al.*, 1987), EPSP synthase (Mousdale and Coggins, 1985) and chorismate mutase. Moreover, since shikimate kinase (Schmid *et al.*, 1992) and EPSPS (Della-Cioppa *et al.*, 1986) are active with their transit peptides, these enzymes could also be a part of cytosolic shikimate pathway.

#### **2.2.4. Regulation of 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) Synthase**

The prokaryotic DAHP synthase is regulated by feedback inhibition. The genes encoding the Tyr- and Trp-sensitive isoenzymes are regulated by Tyr and Trp, respectively whereas the Trp-sensitive isoenzyme is regulated by both amino acids. In contrast to bacterial DAHP synthase, plant DAHP synthases are not subject to feedback inhibition by the aromatic amino acids. Unexpectedly, the purified enzymes from carrot and potato are activated by Trp and to a less degree by Tyr in a hysteretic fashion (Suzich *et al.*, 1985). Thus the aromatic amino acids cannot be considered to be feedback inhibitors of plant DAHP synthase. Based on this fact, they proposed a model for sequential feedback inhibition of the aromatic compounds biosynthesis. This model predicted a cessation of carbon flow into the shikimate pathway under a high concentration of arogenate, a rather unstable shikimate pathway derivative.

The first evidence of metabolic regulation of a plant DAHP synthase came from experiments with suspension-cultured potato cells which has been exposed to glyphosate (Pinto *et al.*, 1988). Glyphosate activated DAHP synthase by increasing the amount of this enzyme in vivo. This activation was specific because other shikimate pathway enzymes were not affected. The herbicide has no effect on DAHP synthase in vitro, indicating that inhibition of chorismate synthesis results in the production of a signal that presumably affects transcription and/or translation of the gene encoding DAHP synthase.

Both mechanical wounding and fungal elicitation induce DAHP synthase mRNA accumulation. The induction of DAHP synthase gene expression parallels a similar induction of PAL mRNA (Dyer *et al.*, 1989), suggesting that the synthesis of aromatic amino acids and secondary aromatic compound might be regulated in concert at the transcriptional level. In case of plants, light replaces the amino acids as a regulator, the regulator increases enzyme activity.

#### **2.2.5. Metal dependency of DAHP Synthase**

DAHP synthases isolated thus far are all metal-dependent. From the first time DAHP synthase was isolated from *E. coli* in 1969, there has been an ongoing controversy in the literature regarding the nature of the metal ion at the active site of this enzyme. A variety of metal ions have been proposed to reside at the active site of DAHP synthase under physiological conditions. They include cobalt, iron or copper. Today, after more than 30 years of research, the nature of the physiological metal ion at the active site of DAHP synthase remains an open question.

Potential metal coordinating residues, Cys61 and His268, have been shown to play an important role in the catalytic activity of DAHP synthase. Mn(II) (hard metal) and Cu(II) or Zn(II) [softer metals as compared with Mn(II)] coordinate different types of ligands, with Mn(II) preferring oxygen ligands and Cu(II) or Zn(II) preferring nitrogen and thiol ligands. The three DAHPS isozymes in *E.coli* regulated as catalytically competent systems contain a single divalent metal ion per subunit. The role of the metal maybe for catalysis and/or structural integrity. The identity of the metal ion in vivo remains elusive, and different metal ions (Cu<sup>2+</sup>, Co<sup>2+</sup>, Fe<sup>2+</sup>) have been found in the protein expressed from the same gene under identical or similar conditions (Jordan *et al.*, 2001). The native metal could be Cu, Zn or Fe- all in substoichiometric amounts. In vitro, the enzyme can be activated by a variety of divalent cations of Cd, Mn, Zn, Fe, Mg, Co, and Cu (Baasov and Knowles, 1989). The diverse nature of these metal activators suggests that they could either play a structural role or as a Lewis acid but is not based upon any redox cycle. No other cofactors have been identified for DAHP synthase from bacteria, although the major enzyme from plants, Mn activated DAHP synthase, is known to require reductants such as DTT.

#### **2.2.6. *Arabidopsis thaliana* DAHP Synthase**

*A. thaliana* contains three genes encoding DAHP synthases, called *DHS1*, 2 and 3. *DHS1* and *DHS2* are tandem genes on chromosome 2. *DHS3* is on chromosome 4 and maybe silent. The sequences of the three are fairly similar, all three genes contain five exons that are separated by four introns. The first three exons encode a polypeptide that aligns with the bacterial enzyme. *At*-DAHPS is a metalloenzyme, with Mn<sup>2+</sup> being the most effective metal ion judged by the enzyme specific activity. The Mn<sup>2+</sup> binding site contains a Cys residue as

ligand. The CYS needs to be reduced to keep the enzyme stable. *At*-DAHPS has at least two surface exposed Cys residues-one to bind  $Mn^{2+}$  and one or two to bind the TRP-CYS-GLY-PRO-CYS motif of the thioredoxin. The binding sites for metal ion and first substrate phosphoenol pyruvate overlap. *At*-DAHPS is an enzyme of the chloroplast stroma. In addition to  $Mn^{2+}$ , *At*-DAHPS has a separate binding site for  $Mg^{2+}$ .  $Mn^{2+}$  is absolutely required for activity;  $Mg^{2+}$  has a synergistic effect at pH 8, the pH of the illuminated stroma. *At*-DAHPS has 9 Cys residues, five in the domain that is considered to harbor the active site and four in the domain that is considered to be the binding site for thioredoxin. When these Cys were mutated, it was found that C139A and C328A had no activity; that could be because C328 was part of the CXXH motif that has been shown for the bacterial enzyme to be metal ion binding site and C139 maybe considered to be located nearby (personal communication, Klaus Herrmann). The heterologous expression of the *Arabidopsis thaliana* genes *DHS1*, *DHS2*, and *DHS3* in *Escherichia coli* yields three *A. thaliana* DAHP synthase isoforms that have a strict reducing agent dependence. The apparent dissociation constants for dithiothreitol (DTT) to the enzyme vary substantially among the isoforms. Half maximal enzyme activity for *DHS2* is reached with about 3  $\mu M$  DTT, whereas the corresponding values for *DHS3* and *DHS1* are about 13 and 43  $\mu M$ , respectively. Chemical modification of *DHS1* with iodoacetamide removes the dependence on reducing agents showing that enzymatic activity is regulated by the redox state of the enzyme through the formation of disulfide bridges.

Purified *A. thaliana* DAHP synthase is activated by reduced thioredoxins (TRX) from spinach, *Spirulina*, and *E. coli*. Recombinant spinach TRX activates *A. thaliana* DAHP

synthase with an apparent dissociation constant of about 0.2  $\mu\text{M}$ . *DHSI* is two fold more active in the presence of DTT and TRX than DTT alone.

*DHSI* has a strict  $\text{Mn}^{2+}$  dependence that is modulated by  $\text{Mg}^{2+}$ . The enzyme is inactive in submicromolar concentrations of  $\text{Mn}^{2+}$  or mM concentrations of  $\text{Mg}^{2+}$  alone. When the metals are present together they exhibit a synergistic effect, producing an active enzyme. This effect is strongly dependent on the pH. Under physiological conditions in the dark, *DHSI* is inactive. During the light cycle when the pH rises, the synergism between  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  produces an active enzyme (Entus, 2002).

### 2.2.7. DAHP Synthase from different organisms

#### 2.2.7.1. *Pyrococcus furiosus*

DAHP synthase from the hyperthermophile *Pyrococcus furiosus* requires a metal ion for activity. It is not inhibited by aromatic amino acids. Purified *P. furiosus* DAHP synthase is able to utilize the E4P and 2-deoxy-D-erythrose 4-phosphate and also the five-carbon phosphorylated monosaccharides A5P, D-ribose 5-phosphate. The structure of recombinant *P. furiosus* DAHP synthase in complex with PEP has been determined to 2.25 Å resolution. The asymmetric unit consists of a dimer of  $(\alpha/\beta)_8$ -barrel subunits. Analysis of the buried surfaces formed by dimerization and tetramerization, as observed in the crystal structure, gives insight into the oligomeric status in solution and the substrate ambiguity of *P. furiosus* DAHP synthase. *P. furiosus* DAHP synthase is both the first archaeal and the first “naked” DAHP synthase (without N-terminal extensions) to be fully characterized functionally and structurally. (Schofield *et al.*, 2005) (Fig. 2.8).

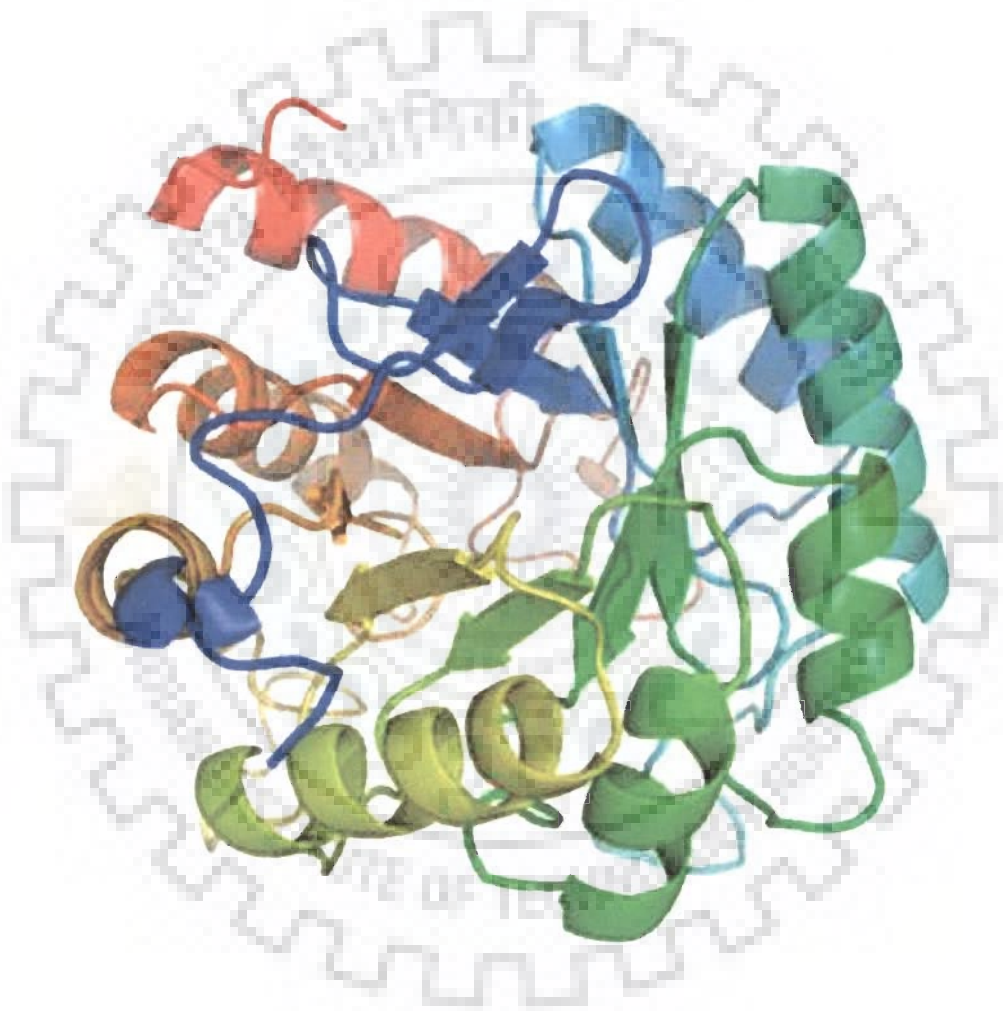


Fig. 2.8. DAHP synthase from *Pyrococcus furiosus* (PDB code 1ZCO).

### 2.2.7.2. *Escherichia coli*

The crystal structure of the phenylalanine-regulated 3-deoxy-D-arabinoheptulosonate- 7-phosphate synthase (DAHP synthase) from *Escherichia coli* has been solved (Fig. 2.9).  $Mn^{2+}$ , the most efficient metal activator, is coordinated by the four side-chains (Cys61, His268, Glu302 and Asp326) which also surround the poorly activating  $Pb^{2+}$ . The PGL, replaces the substrate, 2-phosphoenolpyruvate (PEP), in the active site. However, unlike PEP in the  $Pb^*PEP$  complex, PGL binds the  $Mn^{2+}$  via one of its carboxylate oxygen atoms. PEP binds in the same orientation as does PGL in the DAHP synthase\* $Mn^*PGL$  structure and the phosphate of E4P is tethered at the site of a bound sulfate anion (Wagner *et al.*, 2000).

### 2.2.7.3. *Saccharomyces cerevisiae*

The crystal structures of several complexes between the tyrosine-regulated form of DAHP synthase from *Saccharomyces cerevisiae* and different metal ions and ligands are available (Fig 2.10) (Konig *et al.*, 2004). The crystal structures gives evidence that the simultaneous presence of a metal ion and PEP leads to an ordering of the protein into a conformation that is prepared for binding the second substrate E4P. The site and binding mode of E4P has been derived from the 1.5 Å resolution crystal structure of DAHP synthase in complex with PEP,  $Co^{2+}$ , and the E4P analogue glyceraldehyde 3-phosphate. It is suggested that the oxygen atom of the reactive carbonyl group of E4P replaces a water molecule coordinated to the metal ion that strongly favours a reaction mechanism where the initial step is a nucleophilic attack of the double bond of PEP on the metal-activated carbonyl group of E4P. Mutagenesis experiments substituting specific amino acids coordinating PEP, the divalent metal ion or the

second substrate E4P, result in stable but inactive Aro4p-derivatives and have shown the importance of these residues for the catalytic mechanism (Konig *et al.*, 2004).

#### 2.2.7.4. *Thermotoga maritima*

Crystal structure of DAHP synthase from *Thermotoga maritima* (*Tm*-DAHPS) is available. The subunit of the enzyme homotetramer consists of an N-terminal ferredoxin-like (FL) domain and a  $(\alpha/\beta)_8$ -barrel domain (Fig 2.11). The active site located at the C-end of the barrel contains cadmium, PEP, and E4P, the latter bound in a non-productive conformation. The active site is nearly identical to the active sites of the other known DAHP synthase structures. But the secondary, tertiary, and quaternary structures of are more similar to the functionally related enzyme, 3-deoxy-D-manno-octulosonate-8-phosphate synthase (KDOPS) from *E. coli* and *Aquiflex aeolicus*, than to DAHP synthase from the other organisms. Although the enzyme is feedback-regulated by tyrosine and phenylalanine, it lacks the extra barrel segments that are required for feedback inhibition in *Ec*-DAHPS and *Sc*-DAHPS. It is understood that the mechanism of feedback regulation in *Tm*-DAHPS and other family I $\beta$  enzymes is different from that of family I $\alpha$  enzymes, and most likely being mediated by the FL domain (Shumilin *et al.*, 2004).

#### 2.2.7.5. *Mycobacterium tuberculosis*

The DAHP synthase from *M. tuberculosis* has found to be metal ion-dependent and subject to feedback inhibition by phenylalanine, tryptophan, tyrosine and chorismate, with a significant synergistic effect when tryptophan is used in combination with phenylalanine (Fig. 2.12). The crystal structure of *M. tuberculosis* DAHP synthase has been determined in



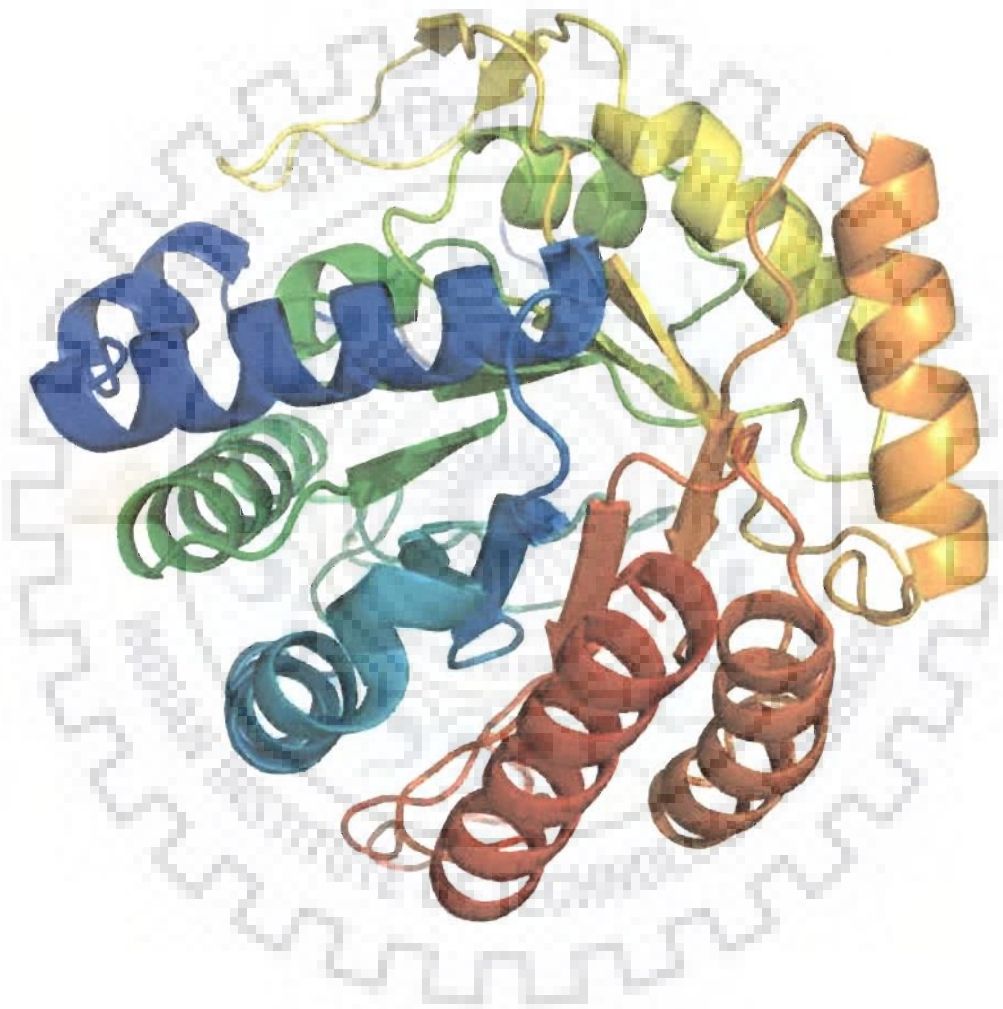


Fig. 2.9. DAHP synthase from *Escherichia coli* (PDB code 1KFL).

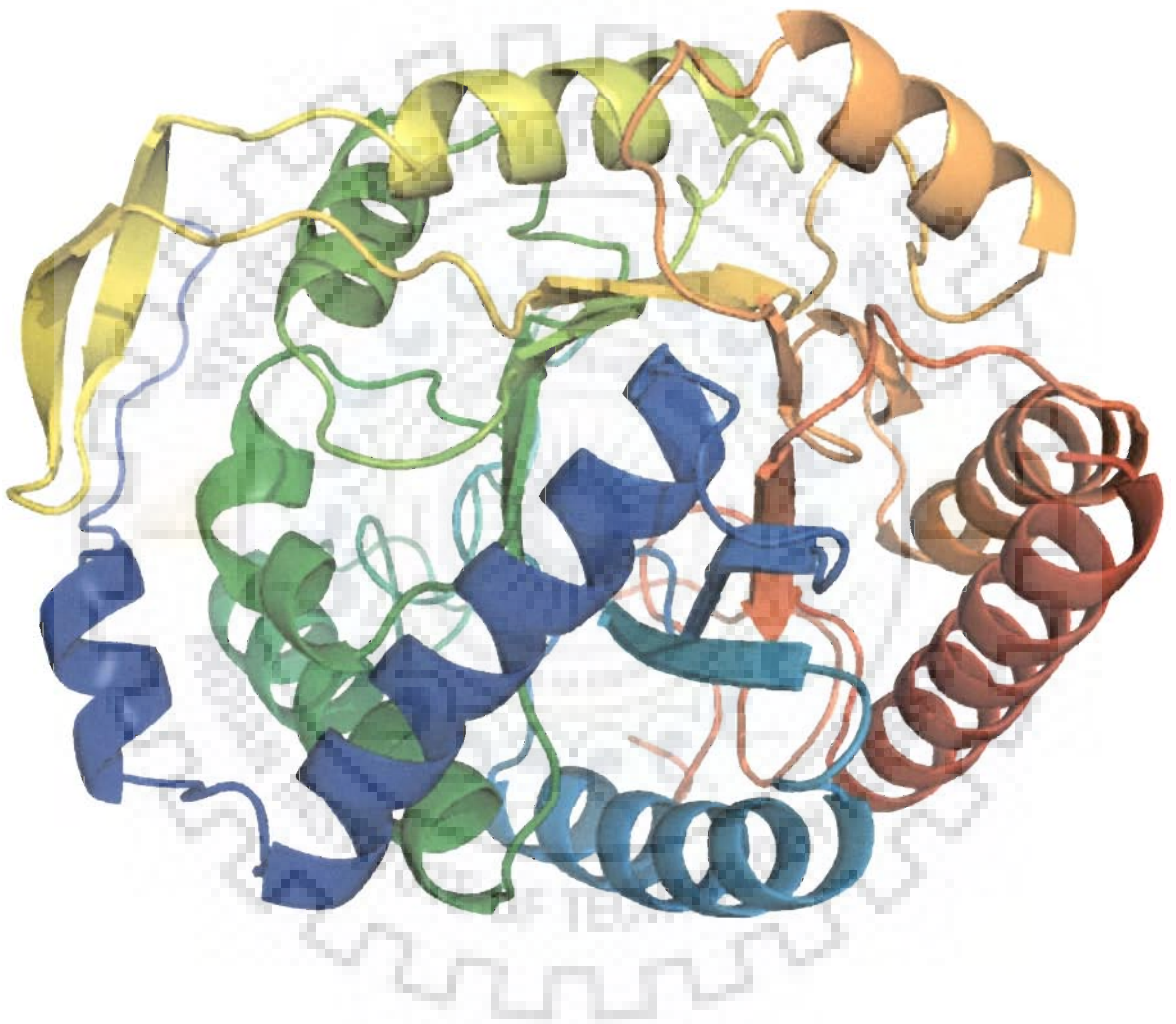


Fig. 2.10. DAHP synthase from *Saccharomyces cerevisiae* (PDB code 1OFQ).

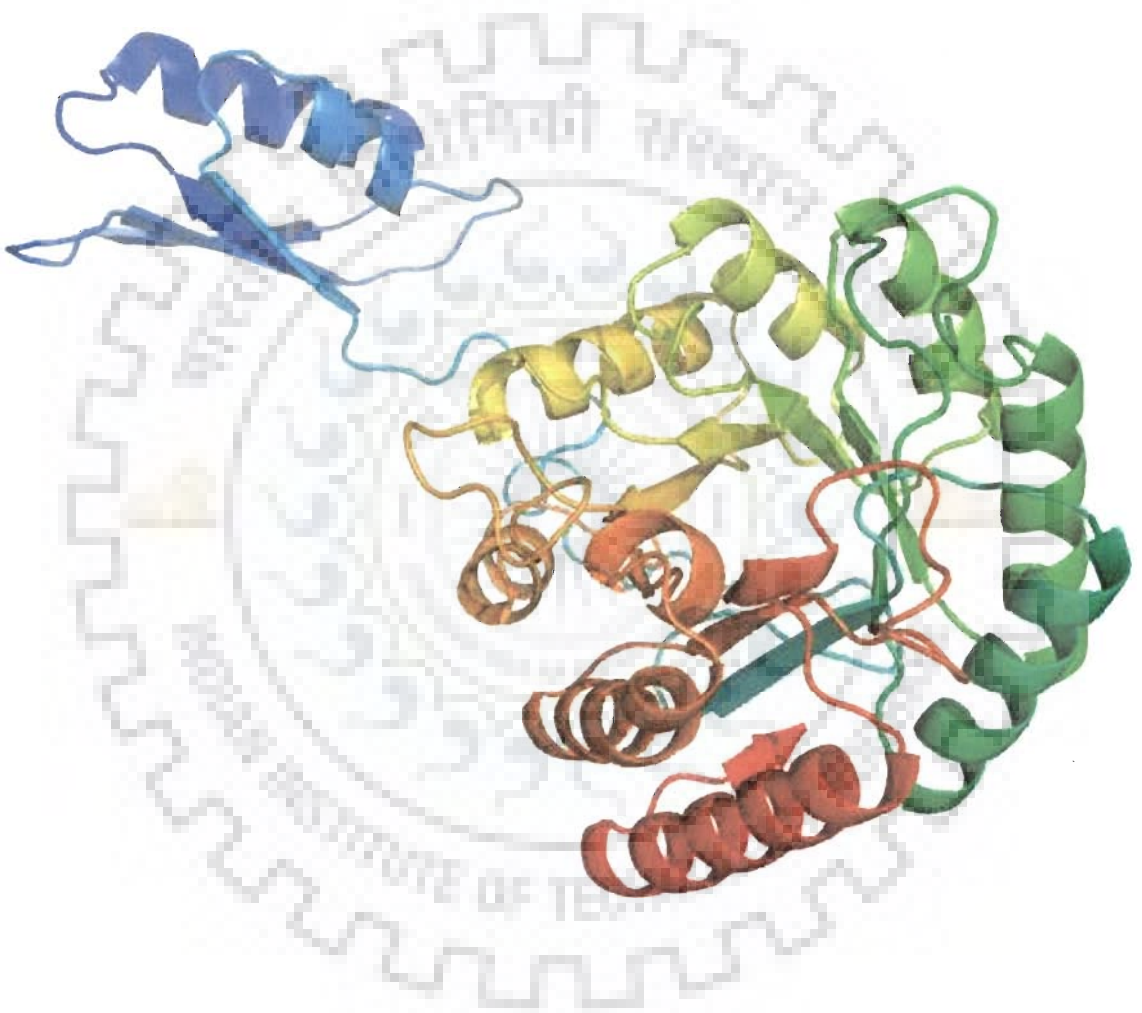


Fig. 2.11. DAHP synthase from *Thermotoga maritima* (PDB code 1VR6).

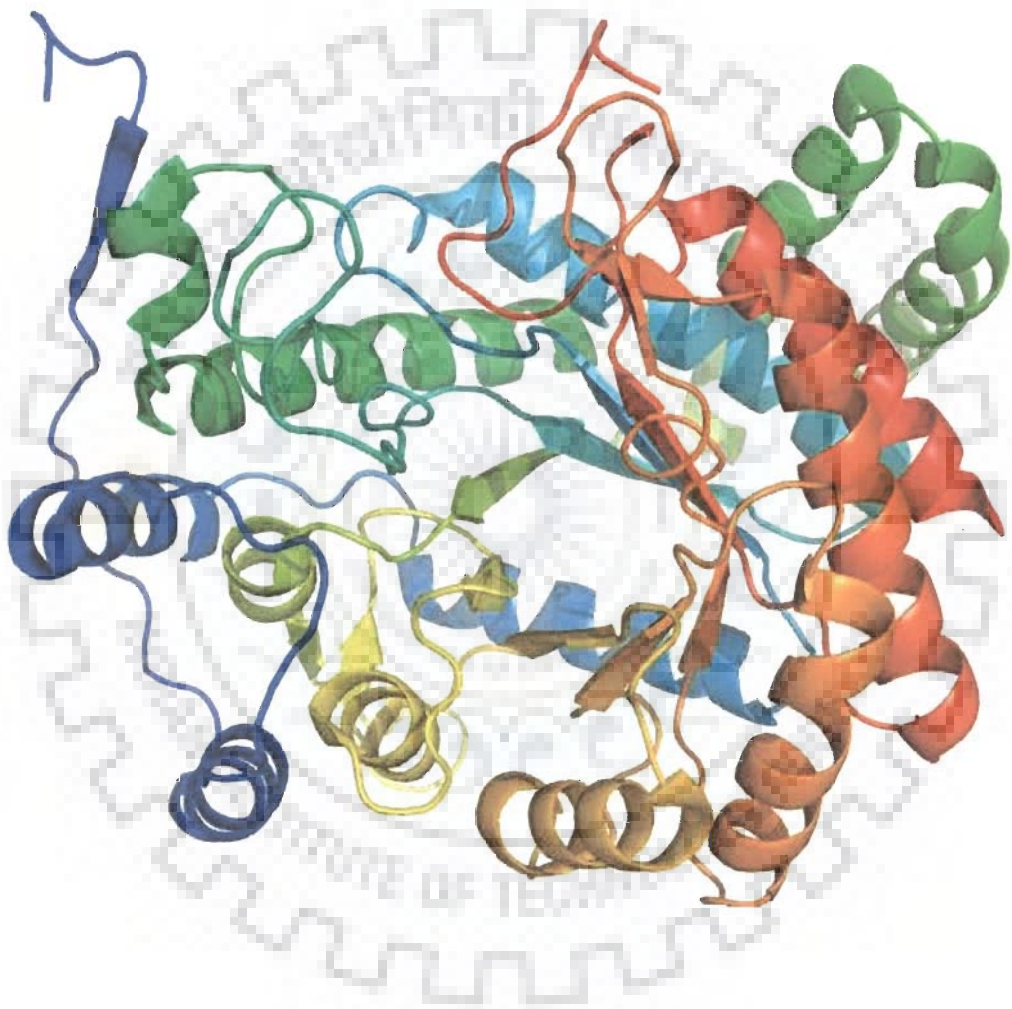


Fig. 2.12. DAHP synthase from *Mycobacterium tuberculosis* (PDB code 2B7O).

complex with substrate phosphoenolpyruvate and  $Mn^{2+}$ . The structure reveals a tightly associated dimer of  $(\alpha/\beta)_8$  TIM barrels. The monomer fold, the arrangement of key residues in the active site, and the binding modes of PEP and  $Mn^{2+}$ , resemble those of the type I enzymes, and point towards a common ancestry for the type I and type II DAHP synthase, in spite of their minimal sequence identity. In contrast, the structural elements that decorate the core  $(\alpha/\beta)_8$  fold differ from those in the type I enzymes, consistent with their different regulatory and oligomeric properties (Webby *et al.*, 2005).

### **2.2.8. Phylogenetic classification of DAHP synthase**

There can be two distinct apparently unrelated types of DAHP synthase proteins, classified on the basis of amino acid sequence homology and molecular mass (Konig *et al.*, 2004). Type I DAHP synthases are smaller than type II DAHP synthase, have molecular masses less than 40 kDa, and are further divided into two subfamilies, I $\alpha$  and I $\beta$  (Celia *et al.*, 2005).

Several other modes of classification for DAHP synthase have been adopted by various scientists so far. Birek and Woodard (Birek and Woodard, 2001) separated DAHP into Class I and Class II on the basis of their metal requirements. Presently, broadly accepted is the view that Type I DAHP synthase enzymes are mainly found in prokaryotic and archaeal organisms, although some eukaryotic examples have been identified [e.g., those from *S. cerevisiae* (Hartmann *et al.*, 2003) and *Neurospora crassa* (Nimmo *et al.*, 1981)]. Type II DAHP synthase proteins were originally identified in plants and are now known to encompass a diverse set of microbial proteins of which the plant proteins form a subcluster (Gosset *et al.*, 2001). Table 2.2 shows the sequence identity shared between the various members of type I and type II DAHP synthase family. Subfamily I $\alpha$  appears to consist

entirely of DAHP synthase proteins and includes the structurally and functionally characterized enzymes of *E. coli* and *S. cerevisiae*. All characterized I $\alpha$  DAHP synthase proteins are metalloenzymes and are sensitive to feedback inhibition by aromatic amino acids. Uptil now, the substrate specificity of subfamily I $\alpha$  enzymes has been examined only for the phenylalanine-sensitive isozyme from *E. coli* which in addition to its own substrate shows limited enzymic activity with a range of five-carbon monosaccharides, including A5P, 2-deoxy-D-ribose 5-phosphate (2dR5P), and D-ribose 5-phosphate (R5P) (Blundell *et al.*, 1987).

Subfamily I $\beta$  consists of both DAHP synthase (subfamily I $\beta_D$ ) and KDO8PS (subfamily I $\beta_K$ ) proteins. The DAHP synthase enzymes of subfamily I $\beta_D$  have more sequence similarity to KDO8PS enzymes of subfamily I $\beta_K$  than to the functionally equivalent I $\alpha$  subfamily of DAHP synthase (Subramaniam *et al.*, 2002). Subfamily I $\beta_D$  proteins are putative metalloenzymes, based on evidence from *Pyrococcus furiosus* DAHP synthase and *Tm*-DAHPS. However, out of the two, only *Tm*-DAHPS is subject to feedback inhibition. Moreover, *Tm*-DAHPS shows a narrow substrate specificity, utilizing E4P but not A5P or R5P as a substrate (Wu *et al.*, 2003). Subfamily I $\beta_K$  contains both metalloenzymes and non-metalloenzymes, Subfamily I $\beta_K$  enzymes are not subject to feedback inhibition and appear to be very substrate specific, utilizing only A5P and not R5P or E4P (Sheflyan *et al.*, 2002).

The ancestor of type I enzymes had a broad substrate specificity, had the ability to coordinate metal, was not subject to allosteric inhibition, and was most likely the I $\beta$  type of DAHP synthases, as this type is the most widely distributed in nature. Over time, a range of changes occurred to varying degrees: substrate specificity narrowed both in the stereospecificity of

Table 2.2. Sequence identity (%) shared between primary sequences of members of DAHP synthase family. The values were obtained using ClustalW.

Identity scores for DAHP synthase sequences	<i>A. thaliana</i>	<i>S. tuberosum</i>	<i>M.tuberculosis</i>	<i>T. maritima</i>	<i>P.furiosus</i>	<i>S.cerevisiae</i>	<i>E.coli</i>	<i>T.gondii</i>
<i>A. thaliana</i>	100	79	44	8	12	4	6	34
<i>S. tuberosum</i>	79	100	44	4	11	2	6	35
<i>M.tuberculosis</i>	44	44	100	8	6	5	9	34
<i>T. maritima</i>	8	4	8	100	61	14	6	10
<i>P.furiosus</i>	12	11	6	61	100	16	14	8
<i>S.cerevisiae</i>	4	2	5	14	16	100	56	8
<i>E.coli</i>	6	6	9	6	14	56	100	6
<i>T.gondii</i>	34	35	34	10	8	8	6	100

the reaction and in the specificity for the aldose, allosteric inhibition was acquired, and metal dependency was lost in some cases (Subramaniam *et al.*, 1998).

### **2.2.9. Specific Aims of Study on 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase**

1. Overexpression of DAHP synthase from the plant *Arabidopsis thaliana* (*At*-DAHPS) in *E.coli*.
2. Purification and crystallization of *At*-DAHPS. But crystals did not appear instead of extensive trials. So an attempt was made to predict its structure through bioinformatics approaches.
3. Prediction of structure and tracing the evolutionary history of *At*-DAHPS through bioinformatics approaches.





# Chapter - 3

## *Studies on Biphenyl Dioxygenase Variants*

### 3.1. ABSTRACT

Biphenyl dioxygenase (BPDO), the first enzyme of the biphenyl degradation pathway is responsible for the initial dioxygenation step during the metabolism of biphenyl. Variants II-9 (BPDO<sub>II9</sub>) and III-52 (BPDO<sub>III52</sub>) generated by Barriault *et al.* (2002) through family shuffling of a targeted region of *bphA* coding for the large subunit (BphA) of biphenyl dioxygenase from *Burkholderia xenovorans LB400* (BPDO<sub>LB400</sub>) exhibited enhanced degrading capabilities towards a wide range of polychlorinated biphenyl (PCB) substrates (Barriault *et al.*, 2002). In this work, the biphenyl dioxygenase from these variants, BPDO<sub>II9</sub> in native state, BPDO<sub>II9</sub> complexed with biphenyl and BPDO<sub>III52</sub> were crystallized to carry out structural studies towards identifying the underlying reason for their better degrading potential as compared to other well studied strains. The crystals were all rhombohedral in shape. BPDO<sub>II9</sub> native crystals belonged to space group R3 and diffracted to 2.50 Å resolution, BPDO<sub>II9</sub> crystals complexed with biphenyl belonged to P1 space group and diffracted to 1.95 Å resolution while the BPDO<sub>III52</sub> crystals belonged to space group R3 and diffracted to 4.50 Å resolution. Crystal structures of BPDO<sub>II9</sub> and BPDO<sub>III52</sub> were determined by using the coordinates of BPDO<sub>LB400</sub> (Kumar *et al.*, unpublished) as the search model, and twelve molecules were placed in the asymmetric unit of biphenyl complex of BPDO<sub>II9</sub> and 4 molecules were placed in BPDO<sub>II9</sub> substrate free and BPDO<sub>III52</sub>. After rigid body refinement, strict NCS was applied for the first rounds of refinement, followed by restrained NCS refinement. BPDO<sub>II9</sub> is  $\alpha_3\beta_3$  hexamer is similar to Naphthalene-1,2-Dioxygenases (NDOs) and other Biphenyl Dioxygenases (BPDOs). The  $\alpha$  subunit can be divided into two distinct domains: a Rieske domain that contains the Rieske [2Fe-2S] center and the catalytic domain that contains the active site mononuclear non-heme ferrous iron. The iron atom at the active

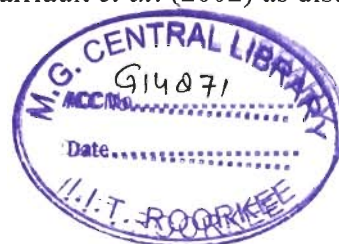
site is pentacoordinated by His 233, His 239, Asp 388 and by two water molecules and the geometry changes when it binds substrate and becomes tetracoordinated as only one water molecule coordinates in this form. The ligation of the non heme iron (mononuclear centre) and Rieske (2Fe-2S) centres and the bridging between them in neighbouring catalytic subunits by hydrogen bonds through a single amino acid (Asp230) are similar to BPDOs and NDOs (Naphthalene dioxygenases). The final models of BPDO<sub>119</sub> in substrate free form and BPDO<sub>119</sub> complex with biphenyl obtained after refinement were subjected to three-dimensional structure comparisons with the other related Rieske oxygenases. The entrance to the active site of BPDO<sub>119</sub> is different from the entrance to the site of other BPDOs and NDOs, as the loop forming the mouth of BPDO<sub>119</sub> entrance block the mouth of the entrance of other related dioxygenases. It was found out that the catalytic pocket of the variants had larger volumes as compared to other known BPDOs and NDOs. The docking studies of various PCBs shows that the amino acids mutated in BPDO<sub>119</sub> were more likely to accommodate greater variety of PCB congeners which could be responsible for the observed enhanced substrate specificity. In addition, BPDO<sub>119</sub> structure supports the report that  $\beta$  subunit play important role in determining substrate specificity (Hurtubise *et al.*, 1998). BPDO<sub>119</sub> displays a hydrogen bonding network between Lys 95 of  $\beta$  - subunit and Asp 387 of  $\alpha$  subunit, which is nearest neighbor of Asp 388 that coordinates with mononuclear iron in the active site.

### **3.2. INTRODUCTION**

Biphenyl dioxygenase (BPDO) catalyses the initial dioxygenation step during the metabolism of biphenyl. Biphenyl and polychlorinated biphenyls (PCBs) are oxidized to the dihydrodiol compound by biphenyl dioxygenase (BPDO). BPDO is a multicomponent enzyme consisting

of four subunits: a large ( $\alpha$ ) and a small ( $\beta$ ) subunit of terminal dioxygenase (encoded by *bphA* and *bphE*, respectively), ferredoxin (encoded by *bphF*), and ferredoxin reductase (encoded by *bphG*) (Suenaga *et al.*, 2001). BphA and BphE are associated as an  $\alpha_3\beta_3$  heterohexamer and catalyze the introduction of both atoms of molecular oxygen into the aromatic nucleus of biphenyl. BphA contains the motif Cys-Xaa-His-Xaa-17-Cys-Xaa-2-His, which forms a Rieske type [2Fe-2S] cluster involved in electron transfer. BPDO requires Fe (II) for activity, and oxygen activation is thought to occur at the mononuclear iron center. Ferredoxin and ferredoxin reductase act as an electron transfer system from NADH to reduce the terminal dioxygenase (Mason and Cammack, 1992).

Studies on the biphenyl dioxygenase revealed significant differences in the congener selectivity patterns and ranges of activity of various PCB-degrading bacteria (Bedard *et al.*, 1986; Furukawa, 1994; Suenaga *et al.*, 2001). It was also demonstrated that both the relative rates of the primary degradation of PCB and the choice of the ring attacked were dependent on the bacterial strain (Furukawa *et al.*, 1979). It is the large subunit (BphA) of BPDO that plays a crucial role in the recognition and binding of the substrate and hence determination of the substrate specificity of biphenyl-related compounds including polychlorinated biphenyls (PCBs). It was also reported that the  $\beta$  subunit influences the substrate specificity of BPDO for PCB (Hurtubise *et al.*, 1998). It has been observed that even small amino acid changes near the active site lead to altered substrate specificity or regiospecificity in some oxygenases. Variants BPDO<sub>II9</sub> and BPDO<sub>III52</sub> were evolved from *Burkholderia xenovorans* strain LB400 (BPDO<sub>LB400</sub>) by the strategies devised by Barriault *et al.* (2002) as discussed in Section 2.1.8.



### **3.3. MATERIAL AND METHODS**

#### **3.3.1. Crystallization of BPDO<sub>I19</sub> and BPDO<sub>I1152</sub>**

BPDO<sub>I19</sub> and BPDO<sub>I1152</sub> were purified in the laboratory of Prof. Michel Sylvestre (INRS Santé- Montreal), Canada (Barriault *et al*, unpublished). The critical steps of purification involved anion exchange chromatography (performed on Source 15Q anion-exchange resin), hydrophobic interaction chromatography (performed on phenyl-Sepharose resin), followed by cation exchange chromatography (performed on MonoQ HR10/10 resin). The enzymes were concentrated by ultrafiltration and shipped on dry ice to IIT Roorkee, India.

Crystals of BPDO<sub>I19</sub> and BPDO<sub>I1152</sub> were grown by vapour-diffusion using the sitting-drop method at 20 °C. The precipitating solution consisted of PEG 5000MME, 50 mM PIPES pH 6.0-7.0, 100 mM ammonium acetate and 6% glycerol.

#### **3.3.2. Preparation of Enzyme:Substrate (BPDO<sub>I19</sub>:Biphenyl) Complex**

Crystalline E:S complex was obtained by soaking crystals of the native enzyme at 20 °C for 3-5 days in artificial mother liquor solutions containing powder form of biphenyl, PEG 5000MME, 50mM PIPES pH 6.0, 100 mM ammonium acetate and 6% glycerol . Soaking of the crystals was performed at 20 °C in the sitting drop well of a 24- well Cryschem plate (Hampton Research, Aliso Viejo, CA) for 3 – 4 days. The well was sealed during the incubation period to prevent dehydration of the crystals.

#### **3.3.3. Preparation of Protein Crystals for X-ray diffraction**

Collection of X-ray diffraction data at cryogenic temperatures (ca. 100 K) improves the quality of diffraction and can also reduce the deleterious effects of X-ray radiation. For data collection of X-ray diffraction intensities at such temperatures, the crystals are transferred

from the growth drop onto a cryo-protectant solution. Thereafter, the crystals are mounted onto a cryo-loop and flash-frozen by immersion in liquid nitrogen. This technique holds the protein crystal within a vitrified cryo-protectant solution maintaining liquid around the protein crystal as an amorphous glass and prevents ice formation. During data collection the cryo-loop containing the protein crystal is exposed to a stream of gaseous nitrogen at 100 K. The composition of cryo-protectant solutions varies depending on the protein crystal and the crystallization condition. A suitable cryo-protectant condition is determined empirically by analyzing the diffraction patterns obtained in different cryo-protectants. The diffraction pattern should not involve contributions due to ice and ideally should contain well shaped spots. Small organic molecules such as glycerol, ethylene glycol, MPD, etc., are commonly added as cryo-protectants.

For our data collection on BPDO<sub>119</sub>, BPDO<sub>11152</sub> and BPDO<sub>119</sub>: Biphenyl complex at cryogenic temperatures, the crystals were transferred to a solution of 15% (v/v) glycerol and reservoir solution. In this study, a higher concentration of the precipitant, Peg 5000 MME and 15% glycerol, provided cryo-protection. The crystals were sequentially transferred into these solutions from the growth drop by a cryo-loop (Hampton Research, Aliso Viejo, CA). The crystals were flash frozen by immersion into liquid nitrogen.

#### **3.3.4. X-ray Diffraction Intensity Measurements and Processing**

Crystals were preliminarily screened for their diffraction properties, using a diffractometer with Cu-K $\alpha$  radiation produced by *Bruker Microstar-H* rotating anode generator and *MAR345-DTB* image plate rotating operated at 100 K and 45 kV and 60 mA equipped with helios optics. The diffraction patterns were analyzed for mosaicity and resolution. For R3 crystal, 100 frames were collected with a 1° rotation per frame with exposure times 10

minutes per degree while 180 degree data were collected for crystal belonged to P1 space group. BPDO<sub>III52</sub> data was collected with 1° rotation with exposure time 15 minutes. X-ray data were indexed, integrated and scaled using the HKL program suite (Otwinowski and Minor, 1997). The space group of the crystals was determined by analyzing the intensities of the measured reflections.

### **3.3.5. Structure Determination by Molecular Replacement**

Molecular replacement is a versatile method for determination of unknown protein structures using a known structure exhibiting similar primary and tertiary structures. It is useful to consider the method in terms of operations involving Patterson function. In this framework, the method involves placing the known structure into the unit cell of the unknown structure by superposing Patterson functions. The known and unknown structures of the proteins are represented by separate Patterson functions.

In this study, the molecular replacement method was performed by utilizing the program, MOLREP (Vagin, and Teplyakov, 1997), available within the CCP4 program suite (Bailey, 1994). For performing the rotational search the program utilizes Crowther Fast Rotation Function. The radius of integration for calculating the RF is determined from the search model as twice the radius of gyration. The rotational search can also be performed at values of radius of integration set by the user. The program generates a list of RF peaks, showing their respective polar angles, Eulerian angles and peak heights measured as deviation from the mean. After rotational search, translational search is performed. In case of crystals with more than one monomer in the asymmetric unit, RF and TF searches are performed on the first monomer and then fixing it before performing RF and TF searches on subsequent monomers. MOLREP performs the search for multiple copies of the monomer in the asymmetric unit

based on predictions from the Matthews coefficient, which is calculated before RF search. The previously determined crystal structure of BPDO<sub>LB400</sub> (Kumar *et al.*, unpublished) served as the search model. Reflections upto 4.0 Å resolutions were utilized to obtain a solution. The radius of integration determined from the search model was 20 Å for performing the RF. Consistent with the estimation from Matthews coefficient, the molecular replacement procedure detected four monomers in the asymmetric unit of the substrate free crystals; the R-factor and CC of the solution was 48% and 60%, respectively. In case of the BPDO<sub>I19</sub> complex, twelve monomers were detected in the asymmetric unit with R-factor  $\approx$  49% and CC  $\approx$  59 %. In all the solutions there was no packing conflict between the polypeptide chains. These models were used to obtain phase estimates for the respective diffracted datasets. A round of rigid body refinement was performed on the model prior to manual modeling.

### **3.3.6. Structure Refinement and Model Building**

Refinement is a crucial step in any structure determination and it involves adjusting some parameters of the atomic model so as to improve the agreement between observed and calculated structure factor amplitudes. During refinement, for each atom four parameters (x, y, z and B) are refined against the observed data while restraining stereochemical parameters such as bond lengths, bond angles, planarities and torsion angles to values obtained from high resolution structures of small molecules or peptides. This ensures improving the agreement of the model with the observed data without deviating from normal stereochemical parameters. Refinement is performed iteratively along with model building. Model building involves construction and fitting the components of the refined model into the electron density maps calculated during the refinement procedure. Validity of the refinement process is monitored by R and R<sub>free</sub>, both scoring the difference between



amplitudes of the observed and calculated structure factors. R is calculated from the data used in the refinement set (RS), whereas,  $R_{\text{free}}$  is calculated from 5 - 10 % data left out of the refinement set, referred to as the test set (TS) (Brünger, 1992).

The value of  $R_{\text{free}}$  is usually higher than R and for a successful refinement the values of both terms should decrease, which indicates improved agreement of the model with the experimental data. During refinement the difference between R and  $R_{\text{free}}$  should not be too large otherwise it indicates bias or over-fitting the data. Refinement is continued until the values of both R and  $R_{\text{free}}$  do not fall after an additional round of refinement. In this study, after obtaining a solution by molecular replacement, a round of rigid body refinement and iterative cycles of atomic parameter (x, y, z and B) refinement were performed using the program REFMAC (Murshudov et al, 1997), with 5% of the data randomly chosen for calculating  $R_{\text{free}}$ . After rigid body refinement, strict NCS was applied for the first rounds of refinement, followed by restrained NCS refinement. The approach of refinement was as follows: assigning isotropic B factors for each atom and restraining bond lengths, bond angles, planar groups and torsion angles. The calculated electron density was visualized using the molecular graphics program O (Jones et al, 1991) and COOT (Emsley and Cowtan, 2004). Residues and substrates were manually built into the electron density using these graphics programs. Coordinates of the biphenyl were generated by PRODRG (Schüttelkopf and van Aalten, 2004). The library files containing the refinement restraints were generated using LIBCHECK available within the refinement module of CCP4 package. The biphenyl molecule was added after adjusting the protein residues and adding water molecules. Water molecules were added both by manual inspection of the electron density and by using the

Table 3.1. Data collection statistics of BPDO<sub>119</sub> native crystal, BPDO<sub>119</sub>: Biphenyl complex and BPDO<sub>11152</sub>.

	BPDO <sub>119</sub> native	BPDO <sub>119</sub> : Biphenyl	BPDO <sub>11152</sub>
Space group	R3	P1	R3
Unit-cell parameters			
<i>a</i> (Å)	211.55	132.77	200.87
<i>b</i> (Å)	211.55	133.19	200.87
<i>c</i> (Å)	169.28	133.97	171.13
$\alpha$ (°)	90	102.30	90
$\beta$ (°)	90	102.50	90
$\gamma$ (°)	120	104.5	120
$V_M$ (Å <sup>3</sup> Da <sup>-1</sup> )	2.4 (48 % solvent)	2.3 (47% solvent)	2.2 (43% solvent)
Asymmetric unit content	4 $\alpha\beta$ heterodimers	12 $\alpha\beta$ heterodimers	4 $\alpha\beta$ heterodimers
Limiting resolution (Å)	2.50	1.95	4.50
Total reflections measured	356781	1229148	184165
No. of unique reflections	97519	634861	15134
$R_{sym}$ (%)	9.3 (23.0)	7.0 (19.0)	10.6 (29.0)
Completeness (%)	99.4 (99.9)	94.0 (92.0)	99.5 (99.3)
$I/\sigma(I)$ (Last resolution shell)	8.9 (2.0)	13.4 (2.2)	5.4 (1.5)
R-factor (%)	22.5	20.4	30.4
R-free (%)	26.4	25.4	35.6
rmsBOND	0.006	0.009	
rmsANGLE	0.913	1.139	

program ARP/wARP (Perrakis et al, 1997) invoked within the REFMAC program. Characteristics of the refined models are listed in Table 3.1.

### **3.3.7. Structure Validation**

The stereochemical properties of the refined models were analyzed by the program PROCHECK (Laskowski et al, 1993). Residues exhibiting sharp deviations of the geometric parameters (such as: bond length, bond angle and torsion angle) from the expected values were rebuilt before the final round of refinement. Stereochemistry of polypeptide backbone was also analyzed by using the Ramachandran plot (Ramakrishnan and Ramachandran, 1965), which is a plot of all observations of the backbone the dihedral  $\Phi$  and  $\Psi$  angles (generated by PROCHECK).

### **3.3.8. Docking of substrates**

The binding of the PCB's molecules with the protein molecule was analyzed using MOE, GLIDE and HEX docking program (Friesner et al., 2004; Ritchie *et al.*, 2003) to find the correct confirmation (with the rotation of bonds, structure of molecule is not rigid) and configuration of the ligand, so as to obtain minimum energy structure.

## **3.4. RESULTS AND DISCUSSION**

### **3.4.1. Crystallization of substrate free BPDO<sub>119</sub>, biphenyl complex of BPDO<sub>119</sub> and BPDO<sub>11152</sub>**

In all cases, the initial crystals appeared after 24 hours and grew to a maximum dimension of  $0.1 \times 0.05 \times 0.05$  mm in two weeks (Fig. 3.1A and 3.1B) for BPDO<sub>119</sub> and BPDO<sub>11152</sub> respectively. BPDO<sub>119</sub> substrate free crystals diffracted to 2.50 Å and belong to a

rhombohedral space group R3 with  $a = 211.55$ ,  $b = 211.55$ ,  $c = 169.28$  Å,  $\alpha = 90.0$ ,  $\beta = 90.0$  and  $\gamma = 120.0^\circ$  and biphenyl complex crystals diffracted to 1.95 Å, have the triclinic space group P1 with  $a = 132.77$ ,  $b = 133.19$ ,  $c = 133.97$  Å,  $\alpha = 102.3$ ,  $\beta = 102.5$  and  $\gamma = 104.5^\circ$ . BPDO<sub>III52</sub> variants in substrate free form belonged to space group R3 and diffracted to 4.5 Å resolution (Fig. 3.1C). Unit cell parameters are given in the table 3.1. Table 3.1 also delineates other statistics associated with the diffraction data and properties of the crystals. For BPDO<sub>II9</sub> models, > 98% of the residues are in favored or allowed regions of Ramachandran plots. The solvent content and Matthews coefficient (Matthews, 1968) of the BPDO<sub>II9</sub> substrate free crystals was 48 % and 2.4 Å<sup>3</sup>/Da, and BPDO<sub>II9</sub>: Biphenyl crystals was 47 % and 2.3 Å<sup>3</sup>/Da. For the BPDO<sub>III52</sub> crystals the solvent content was 43 % and Matthews coefficient was 2.2 Å<sup>3</sup>/Da.

### **3.4.2. Structure Determination**

The crystal structure of substrate free BPDO<sub>II9</sub> was solved by molecular replacement as described under the section materials and methods using the coordinates of BPDO<sub>LB400</sub> (Kumar *et al.* unpublished). The crystals have four  $\alpha\beta$  heterodimers in the asymmetric unit, corresponding to a solvent content of 48% (v/v). The structure has been refined with good stereochemistry to a final R-factor of 22.5 % ( $R_{\text{free}}$  26.4%) to 2.5 Å resolution (Table 3.1). The  $\alpha\beta$  heterodimer of the BPDO<sub>II9</sub> is composed of 459 residues in the  $\alpha$  subunit and 188 residues in the  $\beta$  subunit, and the model contains most of the residues except N-terminal residues 1–17, 143–152 in  $\alpha$  subunit, C-terminal residues 1–7 in  $\beta$  subunits. Thus all residues except for the first seventeen of the N-terminus of the  $\alpha$  subunit, residues 143–152 of the  $\alpha$  subunit and first seven of the N-terminus of the  $\beta$  subunit could be located in the electron density map.

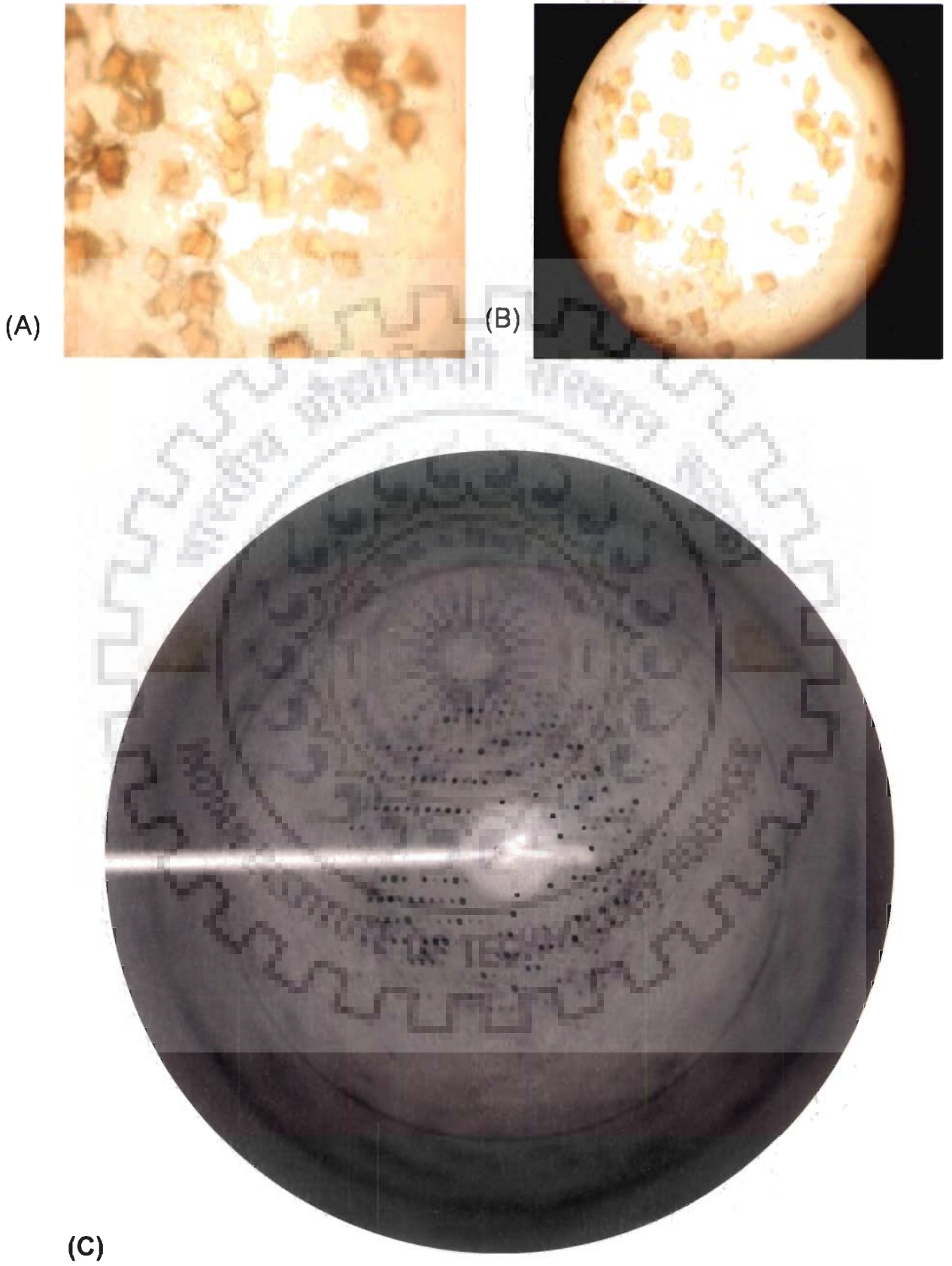


Fig. 3.1. Pictures and diffraction pattern of crystals of variant BPDOs. (A) Pictures of BPDO<sub>I19</sub> crystals. (B) Pictures of BPDO<sub>II52</sub> crystals. (C) Diffraction pattern of BPDO<sub>III52</sub> crystals.

Subsequently, for solving the crystal structure of BPDO<sub>119</sub> complexed with biphenyl, BPDO<sub>LB400</sub> (Kumar *et al.*, unpublished) coordinates served as the template for molecular replacement strategy. BPDO<sub>119</sub>: biphenyl complex had 12  $\alpha\beta$  heterodimers in the asymmetric unit, corresponding to a solvent content of 47% (v/v). Biphenyl complex of BPDO<sub>119</sub> structure was refined with good stereochemistry to a final R-factor of 20.4%. ( $R_{\text{free}}$  25.4%) to 1.95 Å resolution (Table 3.1). In this case too, the first seventeen residues of the N-terminus of the  $\alpha$  subunit and first seven of the N-terminus of the  $\beta$  subunit were disordered. The refined structure had Asn 18 - Pro 459 in the  $\alpha$  subunit and Ala 8 - Phe 188 in the  $\beta$  subunit. FeS occupied coordinate 900 and Fe<sup>2+</sup> occupied coordinate 901.

The crystal structure of BPDO<sub>1152</sub> also was solved by molecular replacement using the coordinates of BPDO<sub>LB400</sub> (Kumar *et al.*, unpublished). These crystals had 4  $\alpha\beta$  heterodimers in the asymmetric unit, corresponding to a solvent content of 43% (v/v). In this case, all residues except for the first seventeen of the N-terminus of the  $\alpha$  subunit and first seven of the N-terminus of the  $\beta$  subunit could be located in the electron density map. It had residues from Ala18-Pro459 in the  $\alpha$  subunit and Pro506-Phe 688 in the  $\beta$  subunit.

### **3.4.3. Overall Structure of Substrate free BPDO<sub>119</sub>**

BPDO<sub>119</sub> has  $\alpha_3\beta_3$  mushroom-shaped structural features (Fig. 3.2 A) where the three  $\alpha$  subunits form the cap and the three  $\beta$  subunits form the stem (Fig. 3.2B). The hexamer structure is believed to be the active biological unit of all known ROs. ROs can be divided into two classes based on their quaternary structure: the homotrimeric or  $\alpha_3$  class, represented by carbazole 1,9a-dioxygenase (CarDO<sub>J8</sub>) (Nojiri *et al.*, 2005) and oxoquinoline

8-monooxygenase (OxoMO<sub>8</sub>) (Martins *et al.*, 2005); and the heterohexameric or  $\alpha_3\beta_3$  class, which is represented by the prototypical heterohexameric RO, NDO from *Rhodococcus sp.* (PDB code 2B1X) (Kauppi *et al.*, 1998). In both classes, three protomers, either an  $\alpha$  monomer or the  $\alpha\beta$  dimer, are related by a three-fold axis to build the active trimer or heterohexamer. The final model of the BPDO<sub>119</sub> crystal structure contained residues Asn18 to Pro459 in the  $\alpha$  subunit, one Rieske [2Fe-2S] cluster, a nonheme iron atom and 334 water molecules.

#### **3.4.3.1. Structure of the $\beta$ -Subunit**

The  $\beta$  -subunit of BPDO<sub>119</sub> has 188 amino acids and has a similar structure to that first defined for the  $\beta$  -subunits of hexameric ROs by the structure of NDO<sub>9816</sub> (Kauppi *et al.*, 1998). The  $\beta$  -subunit consists of a central six-stranded mixed  $\beta$  sheet and three  $\alpha$  helices. The  $\beta$  -subunit has a cystatin-like fold, and belongs to the nuclear transport factor 2-like superfamily. Characteristically, the extensively curved central  $\beta$  sheet forms a “half-barrel” structure that is finished in “barrel-like” formation by the helices. In BPDO<sub>119</sub>, the 15 N-terminal residues adopt different conformations than other known ROs. Based on available structures, the  $\beta$  subunits of BPDO<sub>RHA1</sub>, Cumene dioxygenase from *Pseudomonas fluorescens* CDO<sub>IP01</sub> are more similar to BPDO<sub>119</sub>, while Biphenyl dioxygenase from *Sphingobium sp. B1* (BPDO<sub>B1</sub>) and Nitrobenzene dioxygenase from *Comamonas sp.* NBDO<sub>JS765</sub> is more similar to NDO<sub>RHA1</sub> and NDO<sub>9816-4</sub>.

Several studies have been carried out to understand whether the  $\beta$  -subunit affects substrate specificity and regioselectivity. The  $\beta$  -subunits of biphenyl and toluate dioxygenases has been shown to affect the substrate specificity. On the other hand, studies of benzene, 2-nitrotoluene, tetrachlorobenzene and 2,4-dinitrotoluene dioxygenases indicate that the  $\beta$  -

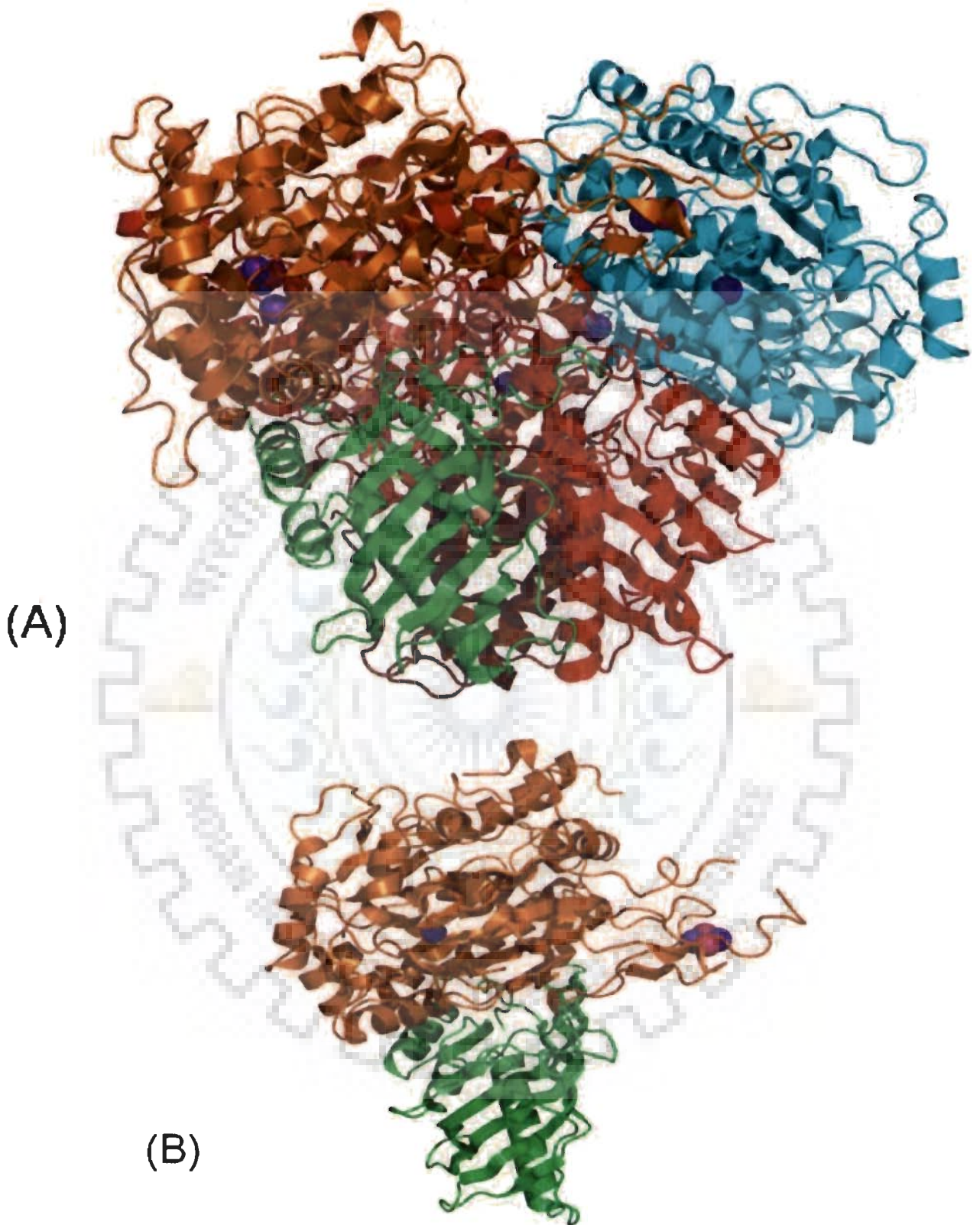


Fig. 3.2. Ribbon diagram showing the overall structure of BPDO<sub>119</sub>. (A) Orthogonal view showing protomers arranged around the crystallographic three-fold axis to form the active heterohexamer. (B) The  $\alpha\beta$  beta dimer of BPDO<sub>119</sub>. Also seen in the figure are the FeS cluster (purple spheres for iron and magenta sphere for sulphur) and the non heme iron (purple sphere). The ribbon diagram shows the C $\alpha$  trace and corresponding secondary structure elements for the respective subunits. All figures were created using Pymol.



subunit is essential for activity but the  $\alpha$ -subunit alone controls the substrate specificity and regioselectivity (Friemann *et al.*, 2005). Several groups have obtained evidence that the  $\beta$ -subunit plays a role in determining substrate specificity, either by N-methyl N'-nitro N-nitrosoguanidine (NTG) mutagenesis or by creating chimeric  $\alpha\beta$  dimers (Hurtubise *et al.*, 1998). In studies of chimeric enzymes, such as BPDO  $\alpha_{LB400}\beta_{B356}$  and BPDO  $\alpha_{B356}\beta_{LB400}$ , exchange of the  $\beta$ -subunit resulted in an extended substrate range relative to the parental proteins and/or a shift in substrate preference correlated with the source of the  $\beta$ -subunit (Hurtubise *et al.*, 1998).

Lys 95 of BPDO<sub>II9</sub> from  $\beta$ -subunit makes hydrogen bonds with Asp 387 from the  $\alpha$ -subunit, the Asp387 residue immediately preceding the active site Fe ligand D388 (Fig. 3.3). Compared to BPDO<sub>II9</sub>, in NDO<sub>9816-4</sub> and CDO<sub>IP01</sub> the corresponding loop is located approximately 2.0-3.0 Å further away from the  $\alpha$ -subunit, and is not involved in a similar hydrogen-bonding network. Therefore, in NDO<sub>9816-4</sub> and CDO<sub>IP01</sub> this loop is not expected to exert a similar influence on the adaptations of the active site during catalysis. These interface interactions may couple  $\alpha$  and  $\beta$  in a way that directly affects substrate specificity. For example, the contact between subunits might affect substrate specificity by restricting motion in the active site and the ability of the site to adjust to different substrates. Thus, variations in interactions at this interface could explain inconsistencies in substrate-profiling experiments using limited mutagenesis or subunit exchange to probe the role of the  $\beta$  subunit.

#### **3.4.3.2. Structure of the $\alpha$ -Subunit**

The catalytic or  $\alpha$ -subunit of ROs has two domains, a smaller Rieske ferredoxin domain and a larger mononuclear Fe(II)-containing catalytic domain. The  $\alpha$ -subunit contains Asn 18-Pro 459 residues. The N-terminal portion consists of a Rieske [Fe<sub>2</sub>S<sub>2</sub>] iron-sulfur cluster

domain. The Rieske domain consists of  $\beta$  strands and loops, forming an ISP domain fold. Four residues, two histidines and two cysteins, in the  $\alpha$  subunit coordinate the Rieske non-heme iron cluster  $[\text{Fe}_2\text{S}_2]$ . His 102 and His 123 coordinate one iron, while Cys 100 and Cys 120 coordinate the other. The assembly of the oligomer places the Rieske  $\text{Fe}_2\text{S}_2$  cluster of each subunit near the mononuclear iron site of an adjacent subunit. Thus, each  $\alpha$ -subunit interacts with an adjacent  $\alpha$ -subunit by extending its Rieske domain to clasp its neighbor (Fig. 3.4). While domain arrangements and the extent of interaction are similar, the orientation of the Rieske domain to the catalytic domain varies significantly between the trimeric and hexameric ROs. When the catalytic domains are superposed, the Rieske domains of the trimeric ROs are rotated by approximately 25 degrees relative to that of the hexameric ROs, which results in a 11 Å shift of the  $\text{Fe}_2\text{S}_2$  cluster towards the three-fold axis of the oligomer.

The C-terminal domain is a mix of helices and strands forming a TBP-like or helix-grip fold and is a member of the Bet v1-like superfamily. The structural conservation suggests that the electron transport in BPDO<sub>119</sub> is similar to the system described for NDO-O<sub>9816.4</sub> and benzoate dioxygenase.

Rieske domains constitute a unique, all  $\beta$  fold, which can be subdivided into a basal  $\beta$  - sandwich domain and a cluster-binding domain containing the Cys-X-His-(X)<sub>15-47</sub>-Cys-X-X-His  $\text{Fe}_2\text{S}_2$  cluster-binding motif (Colbert *et al.*, 2000).

As expected, the structures of the catalytic domains are also quite similar, with catalytically important residues being conserved in structurally equivalent positions. However, surface features of the  $\alpha$ -subunit catalytic domains are more variable than those of the  $\alpha$ -subunit Rieske domains or the  $\beta$ -subunits. Residues forming such variations are located on the

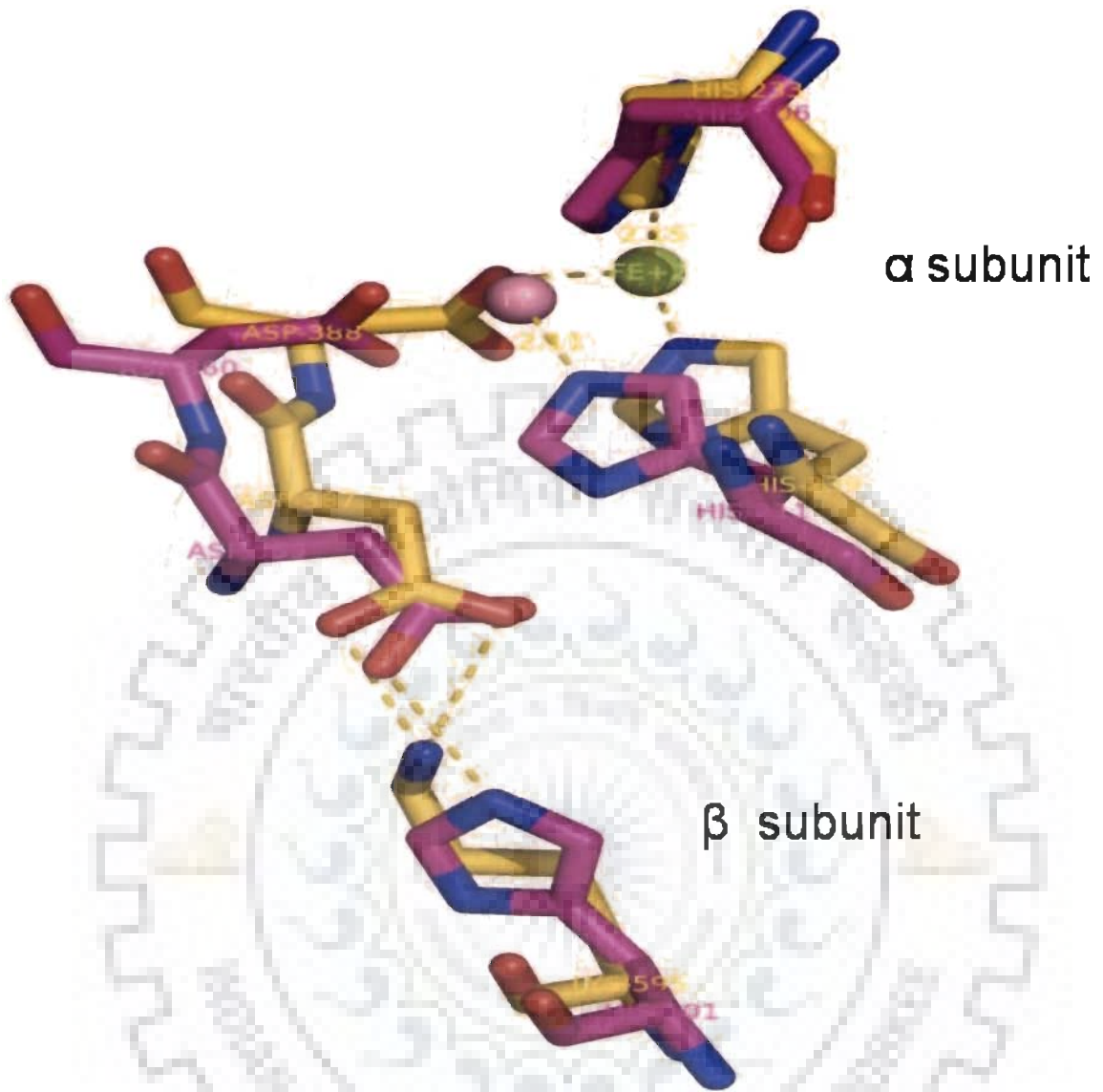


Fig. 3.3. Lys 95 of BPDO<sub>119</sub> from  $\beta$ -subunit makes hydrogen bonds with Asp 387 from the  $\alpha$ -subunit.

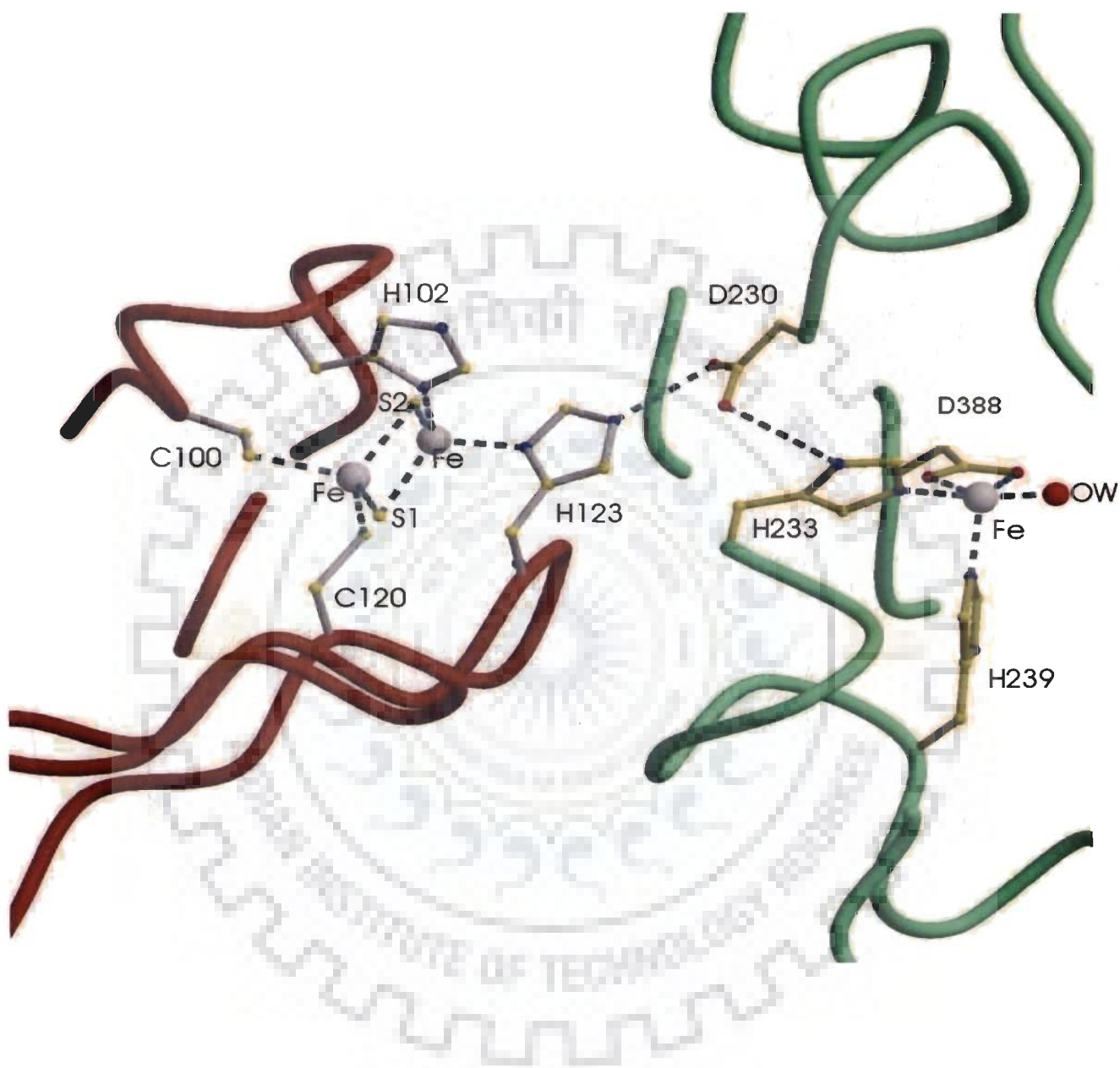


Fig. 3.4. Electron Transfer between the Rieske center and the catalytic iron site. The Rieske center and the iron at the active site are connected through hydrogen bonds to Asp230 in the symmetry-related  $\alpha$  subunit of BPDO. This residue binds to His233 at the active site, and to His123 at the Rieske center in a neighboring subunit. This is the hypothesized electron transfer pathway in BPDO and NDO.

external face of the  $\alpha$ -subunit, and hexameric RO structures can be subdivided into two groups based on these variable regions, with NDO<sub>9816-4</sub> and NBDO<sub>JS765</sub> clustered in one group, and BPDO<sub>II9</sub>, BPDO<sub>RHA1</sub> and CDO<sub>IP01</sub> in the other. The two groups correspond to Group III and Group IV ROs, respectively, of the classification scheme developed by Nam and coworkers (Nam *et al.*, 2001). The most important of these structural variations involves the residues corresponding to 240-264 in BPDO<sub>II9</sub>, which form the entrance to the active site.

#### **3.4.3.3. Coordination of the Mononuclear Iron and the Rieske cluster**

At the heart of the catalytic domain is a mononuclear Fe(II) atom. This non heme iron in the active site in BPDO<sub>II9</sub> was coordinated by NE2 His 239 at a distance of 2.11 Å, a monodentate bond to OD1 Asp388 at a distance of 1.91 Å, NE2 His 233 at a distance of 2.14 Å and two water molecules (OW1 and OW2) at a distance 1.91 and 2.13 ( Fig. 3.5). Coordination of this catalytic Fe(II) appears variable among the ROs, involving between 4-6 ligands (Furusawa *et al.*, 2004; Friemann *et al.*, 2005; Dong *et al.*, 2005; Gakhar *et al.*, 2005) and is influenced by diverse factors, including RO-dependent variation in the number and nature of the protein ligands, labile water ligands, presence or absence of substrate, as well as the influence of variations in the redox state of the Fe<sub>2</sub>S<sub>2</sub> cluster.

Both crystallographic and spectroscopic experiments agree that the protein ligands to the Fe(II) are two histidines and a single aspartate residue. All crystal structures show that the catalytic Fe(II) is coordinated via the N $\epsilon$ 2 atoms of the histidines, however, coordination by the aspartate residues is more variable, and may involve either bidentate or monodentate bonds to the O $\delta$  atoms (Coulter *et al.*, 1999; Kauppi *et al.*, 1998). The BPDO<sub>II9</sub> structures demonstrate that the Fe (II) coordination shell includes 2.1 Å bonds to the N $\epsilon$ 2 of both H233 and His239, where as monodentate bonds of 1.9 Å and 2.8 Å to O $\delta$ 1 and O $\delta$ 2 of D386,

respectively. This Fe(II) coordination agrees with BPDO<sub>RHA1</sub> and CDO<sub>IP01</sub> but differs with the NDO<sub>9816-4</sub> structures (Kauppi *et al.*, 1998). However, in some cases the distance to the O $\delta$ 2 atom clearly exceeds a typical Fe-O bond distance ( $2.1 \pm 0.2$  Å) with distances as long as 2.9 Å, indicating that the aspartate coordinates the Fe(II) in a monodentate manner (Dong *et al.*, 2005; Friemann *et al.*, 2005; Furusawa *et al.*, 2004). Extremely long bonds in the coordination sphere of the catalytic Fe appear to be a recurring theme as the distances observed for the NDO<sub>9816-4</sub> *cis*-dihydrodiol contained an Fe to -OH bond distance of 2.8 Å for both hydroxyl groups (Karlsson *et al.*, 2003).

Variation in mononuclear iron coordination may also arise because of the presence of an additional amino acid ligand or a differing number of water ligands. This residue is either a glutamine or asparagine in other ROs and is located at approximately the same distance from the Fe(II) in all available structures (Dong *et al.*, 2005; Furusawa *et al.*, 2004; Furusawa *et al.*, 2004; Gakhar *et al.*, 2005; Nojiri *et al.*, 2005; Friemann *et al.*, 2005). In BPDO<sub>I19</sub> the N $\epsilon$ 2 atom of the corresponding residue, Q 226, is 3.7 Å from the mononuclear iron. Therefore, despite their conservation and proximity to the Fe(II), these residues might not be ligands to the Fe(II), because the distances are too long for a bonded interaction. Thus, it is more likely that the variability in the number of water ligands is responsible for the discrepancy in the coordination number of the catalytic Fe(II).

Crystal structures of NDO<sub>9816-4</sub>, oxidized OxoMO<sub>8</sub>, and CarDO<sub>J3</sub> revealed a single water ligand (Kauppi *et al.*, 1998; Dong *et al.*, 2005; Furusawa *et al.*, 2004; Friemann *et al.*, 2005; Gakhar *et al.*, 2005; Nojiri *et al.*, 2005). We observe two water molecules (OW1) and (OW2) in the BPDO<sub>I19</sub> substrate free structure and one water molecule when complexed with biphenyl.

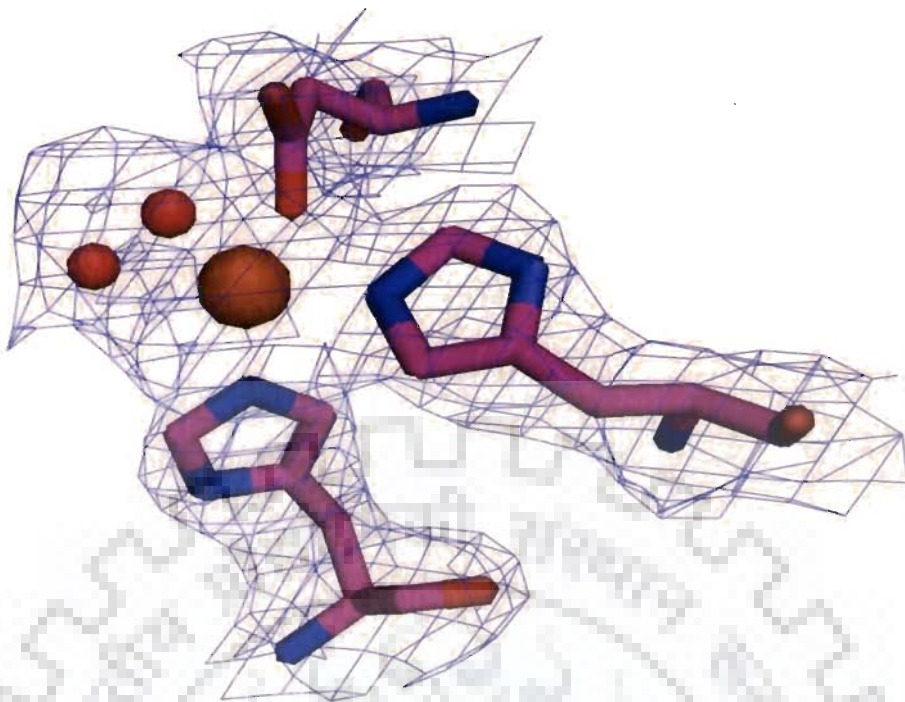


Fig. 3.5. Electron density contoured at  $1.0 \sigma$  showing the coordination of the catalytic iron in substrate free BPDO<sub>119</sub>. Mononuclear iron (chocolate sphere) coordinates with two water molecules (red sphere), Asp 388, His 233 and His 239 in the substrate free state are shown as magenta stick models.

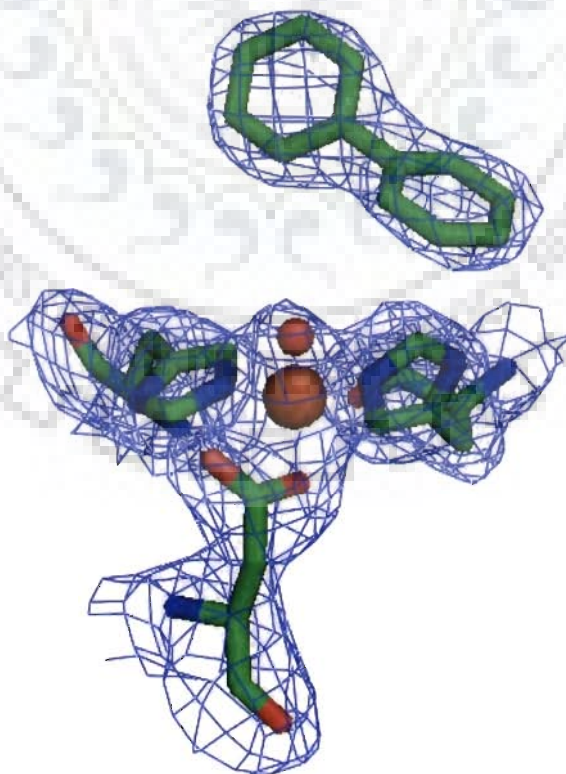


Fig. 3.6. Electron density figure contoured at  $1.0 \sigma$  showing the bipheryl bound state of the catalytic site in BPDO<sub>119</sub>. The mononuclear iron coordinates with one water molecule (red sphere) in the complex state. Asp 388, His 233 and His 239 in the substrate free state are shown as green stick models.

For BPDO<sub>II9</sub>, the Fe(II)-coordinating residues are His 233, His 239, and Asp 388, with the basal plane defined by His 233, Asp 388, and two water molecules (OW1) and (OW2). The coordination is changed in the presence of biphenyl. If we consider bidentate coordination of Fe (II) with Asp 388 in BPDO<sub>II9</sub> then it agrees with PDO hexacoordinate mononuclear iron coordination geometry in the substrate-free enzyme as deduced from spectroscopic studies of and with transformed pentacoordinate form upon substrate binding. Such a transformation would free a coordinating position for the O<sub>2</sub> molecule required for the reaction. This differs from the geometry revealed by the crystal structures of NDO<sub>9816-4</sub>, where a pentacoordinate Fe site was found in the aromatic substrate-free and substrate-bound enzyme, but a hexacoordinate Fe was observed in the presence of either O<sub>2</sub> or *cis*-dihydrodiol product (Karlsson *et al.*, 2003; Kauppi *et al.*, 1998). The structures of CDO<sub>IP01</sub> and BPDO<sub>RHA1</sub> revealed the presence of a bidentate external ligand that was tentatively modeled as a dioxygen molecule (Dong *et al.*, 2005; Furusawa *et al.*, 2004).

#### **3.4.4. Structure of BPDO<sub>II9</sub> with Substrate Biphenyl**

The biphenyl molecule could be identified clearly in the difference fourier maps in the active site of BPDO<sub>II9</sub> (Fig. 3.6). When biphenyl binds in the active site, the overall structure remained the same as found for the free enzyme. The closest atom (C2) of biphenyl to the Fe is 4.87 Å. Water molecule is directly on the line between the Fe and C2 atom and is at a distance of 2.56 Å from C2. Substrate binding makes conformational changes near the active site.

Residues Asp 388, His 233, His 239 have movement upon substrate binding (Fig. 3.7). Loop region from 243-261 exhibits a higher temperature factor, indicating the mobile character of the loop and shows significant movement when substrate binds. Ring 2 of the substrate was



oriented in a different way in comparison to CDO, NDO, NBDO, BPDO<sub>RHA1</sub>, while the ring 1 was in the similar position as it was in other BPDOs.

The analysis of the substrate-binding pocket shows that it is necessary to have some conformational movement to accommodate the biphenyl molecule. In BPDO<sub>I19</sub> has an access to active site via L-shaped tunnel (Fig. 3.8 A) whereas in NDO, the access to the catalytic site is a long narrow gorge (Fig. 3.8 B).

Residues in the catalytic domain in the  $\alpha$  subunit that interact with the substrate are hydrophobic, providing an appropriate environment for the binding of aromatic substrates.

Biphenyl binds in the same fashion as in other BPDOs and interact closely with the hydrophobic Gln 226, Phe 227, Asp 230, Met 231, His 233, Ala 234, Val 287, Gly 321, His 323, Leu 333, Ile 336, Phe 378, Phe 384. Shown in Fig 3.9A and Fig 3.9B are the superpositions of the active site of BPDO<sub>I19</sub> over BPDO<sub>RHA1</sub> (PDB code 1ULJ) and BPDO<sub>B1</sub> (PDB code 2GBX). The carbon atoms at positions 2 and 3 of biphenyl, which undergo the attack of the dioxygen, are located closest to the mononuclear iron with the distances 4.87 Å and 4.91 Å, respectively. The carbon atoms of naphthalene in NDO that are attacked by dioxygen are located at a similar distance from the iron (4.6 Å and 4.3 Å).

#### **3.4.5. Role of loop at the Active Site Entrance**

Differences between ROs at the active site entrance and cavity may dictate differences in substrate preferences. In BPDO<sub>I19</sub>, this entrance is formed by residues 235-237, residue 240, loop residues 250-263. Corresponding residues 250-263 of BPDO<sub>I19</sub> are key components of the active site entrance in all ROs. The entrance loop is topologically similar for BPDO<sub>I19</sub> and CDO<sub>IP01</sub> but it is different in NDO (Fig. 3.10 A and B). Although the corresponding

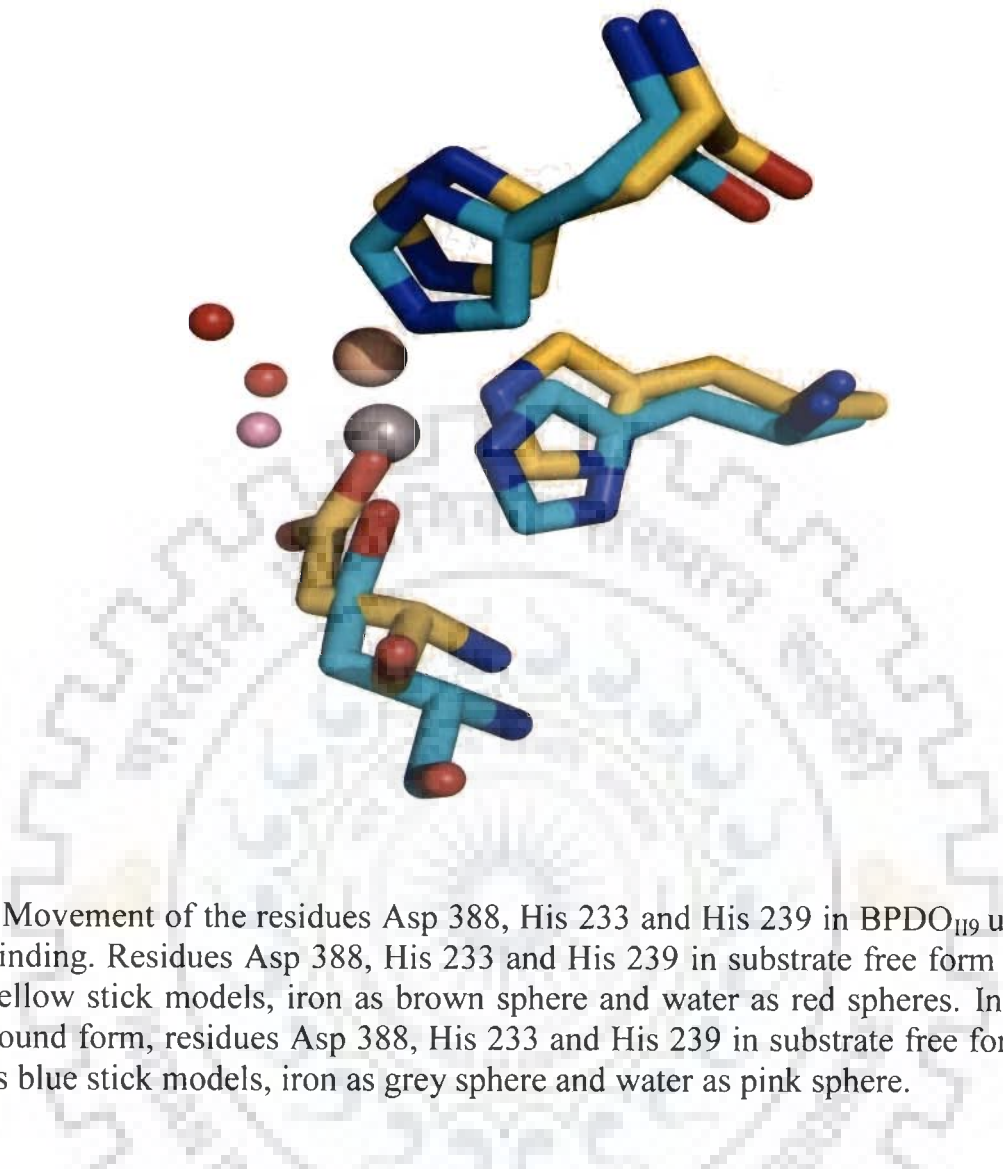


Fig. 3.7. Movement of the residues Asp 388, His 233 and His 239 in BPDO<sub>119</sub> upon substrate binding. Residues Asp 388, His 233 and His 239 in substrate free form are shown as yellow stick models, iron as brown sphere and water as red spheres. In the biphenyl bound form, residues Asp 388, His 233 and His 239 in substrate free form are shown as blue stick models, iron as grey sphere and water as pink sphere.

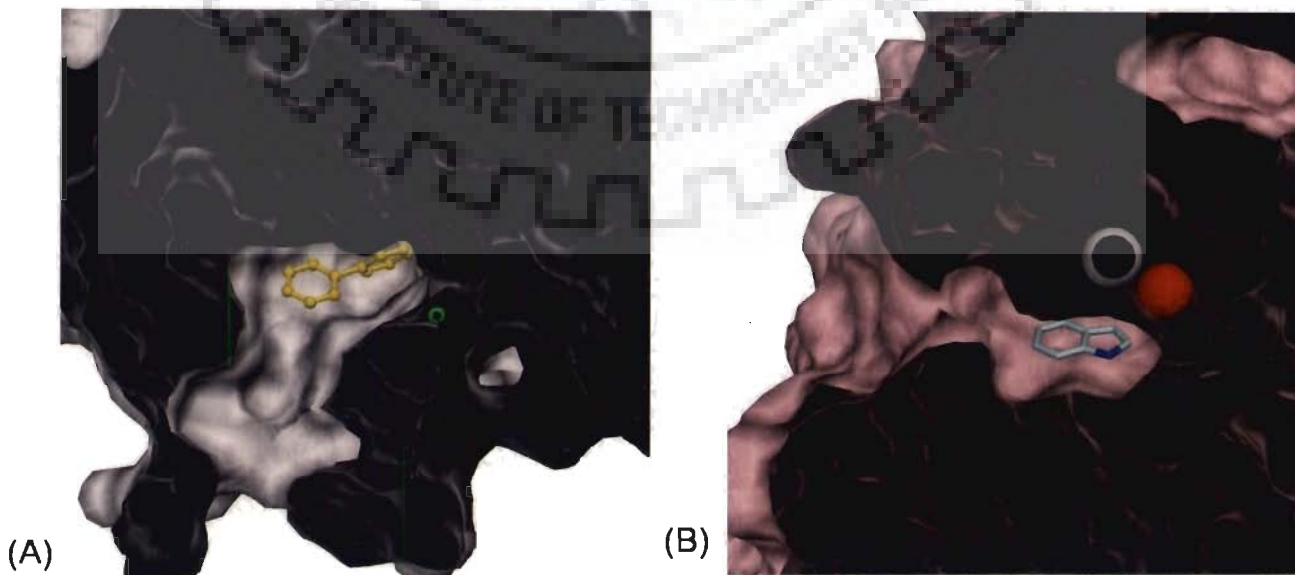


Fig. 3.8. Access to the catalytic site in BPDO<sub>119</sub> and NDO. (A) L-shaped tunnel in BPDO<sub>119</sub> (B) Long narrow gorge in NDO.

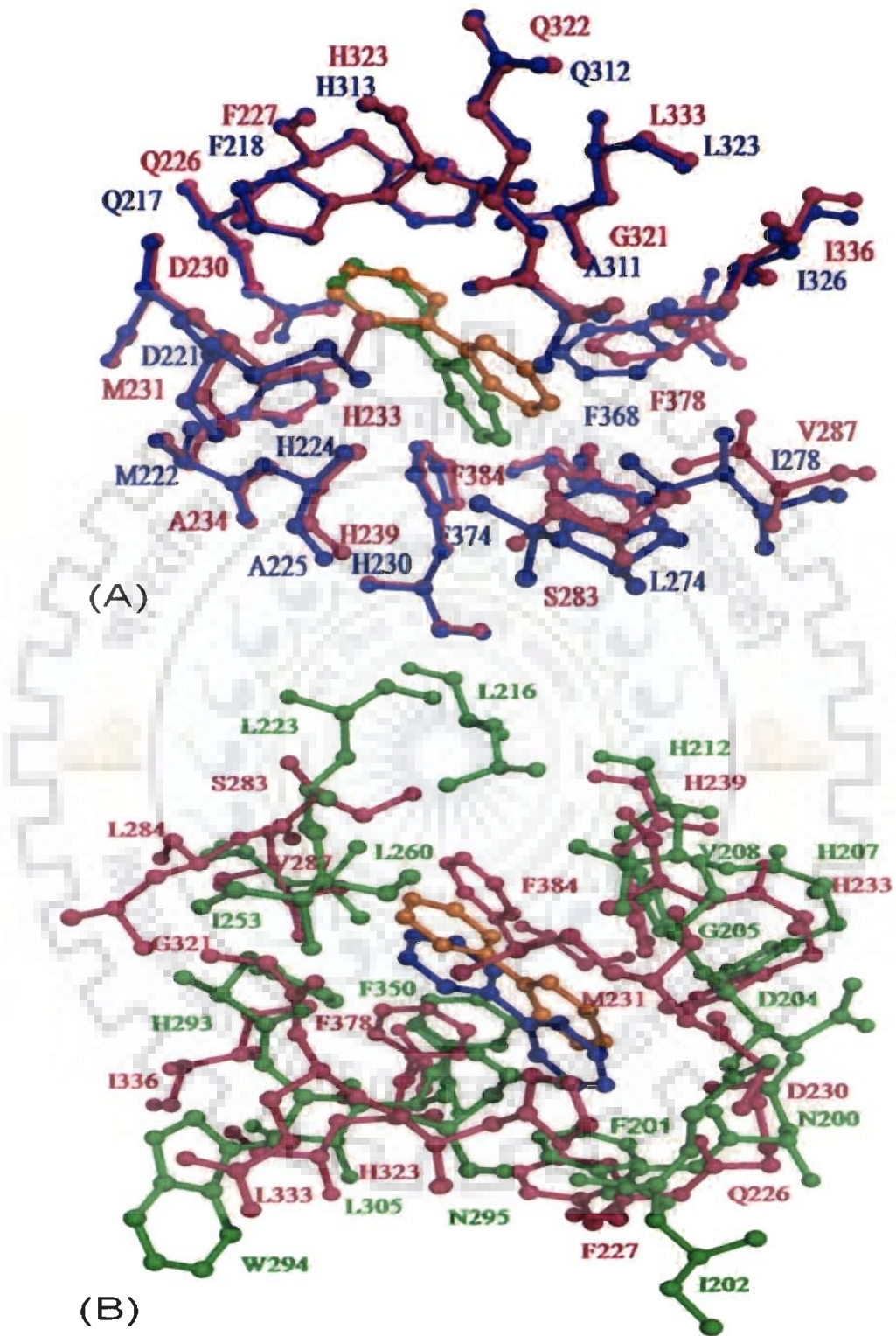


Fig. 3.9. The active site of BPDO<sub>119</sub> (magenta balls and sticks) with biphenyl bound (orange balls and sticks). (A) Overlaid is BPDO<sub>RHA1</sub> (PDB code 1ULJ) (blue balls and sticks) with biphenyl bound (green balls and sticks). (B) Overlaid is BPDO<sub>B1</sub> (PDB code 2GBX) (blue balls and sticks) with biphenyl bound (green balls and sticks).

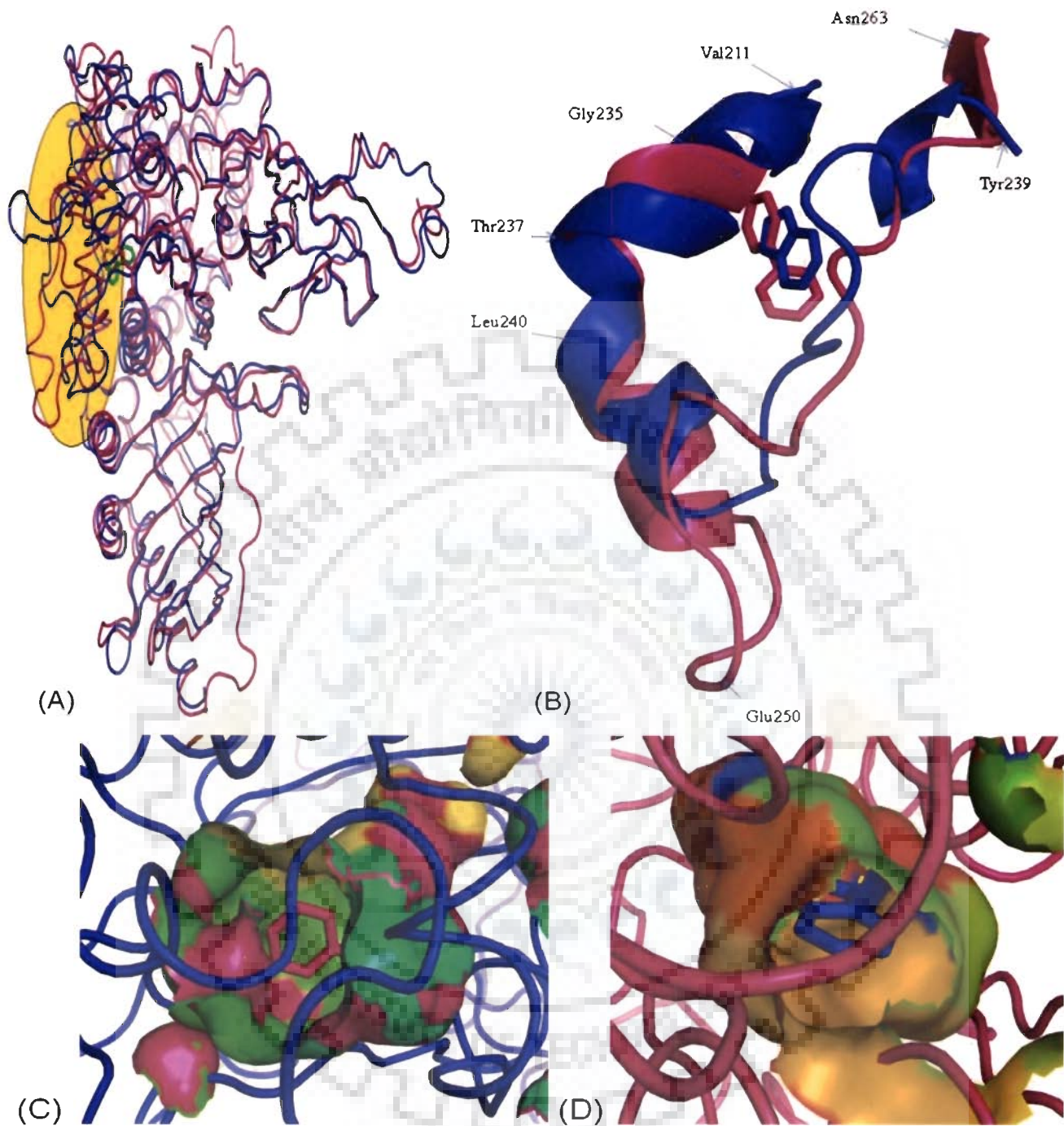


Fig. 3.10. Superimposition of BPDO<sub>119</sub> (magenta) over NDO (PDB code 2B24) (blue) (A) Backbone superposition of BPDO<sub>119</sub> (magenta) over NDO (PDB code 2B24) (blue). (B) Close view of the loop region at mouth of active site (Highlighted with yellow background in A). (C) The C-alpha trace of NDO (PDB code 2B24) is shown in blue over the active site cavity of BPDO<sub>119</sub>. Biphenyl is partially visible as magenta stick model in the binding pocket. (D) The C-alpha trace of BPDO<sub>119</sub> is shown in magenta over the active site cavity of NDO (PDB code 2B24) where indole is partially visible as blue stick model in the binding pocket.

residues are disordered in BPDO<sub>RHA1</sub>, phylogenetic analysis (Nam *et al.*, 2001) and a surface rendering of the BPDO<sub>RHA1</sub> active site cavity suggest that its entrance and cavity are more similar to BPDO<sub>I19</sub> than NDO. This loop adopts a similar conformation in all ROs of known structure. This has been called the “loop 1” conformation in the case of CDO<sub>IP01</sub> (Dong *et al.*, 2005). Further, these loop residues show evidence of conformational flexibility in all available Group IV RO structures. In BPDO<sub>I19</sub> and CDO<sub>IP01</sub> residues 250-263 have high B-factors (Dong *et al.*, 2005), while in BPDO<sub>RHA1</sub> these residues were disordered (Furusawa *et al.*, 2004). In contrast, in the case of the group III RO NDO (Kauppi *et al.*, 1998), the equivalent loop residues have a very different conformation, called “loop 2” by Dong and co-workers (Dong *et al.*, 2005), where the loop is located at the opposite side of the active site entrance. When NDO and BPDO<sub>I19</sub> active sites were superposed they seem to be blocking the mouth of the active site in each other thereby accounting for the difference in specificities ( Fig. 3.10 C and D). Thus, group III and IV ROs use equivalent loop residues differently to regulate access to the active site.

#### **3.4.6. Comparison with other Rieske non-heme oxygenases**

Table 3.2 gives information on sequence and structural similarity between BPDO<sub>I19</sub> and other ROs. BPDO<sub>I19</sub> shares the highest sequence identity (75%) with cumene dioxygenase (CDO<sub>IP01</sub>) and the structures are very similar. Superposition of the  $\alpha\beta$  heterodimers (588 Ca atoms) gives an RMS deviation of 0.70Å ( $\alpha$  subunit ; RMS deviation of 1.12 Å for 432 Ca atoms,  $\beta$  subunit ; RMS deviation of 0.69 for 169 Ca atoms). There are only minor structural differences between the enzymes and  $\alpha$  subunit differs more than the  $\beta$  subunit. The difference with CDO<sub>IP01</sub> lies in some areas in the N-terminal portion of the  $\beta$  subunit, while BPDO<sub>I19</sub> has a small  $\alpha$  helix composed of three residues and it extends away from the rest of

the barrel-like  $\alpha$  helix, CDO<sub>IPO1</sub> has no such  $\alpha$  helical structure at the N-terminal and it is more like a coil that lies close to rest of the  $\beta$  subunit throughout and yet slightly away from it. Then the residues 89-93 of the  $\beta$  subunit adopt a slightly different conformation in CDO<sub>IPO1</sub> as compared to BPDO<sub>I19</sub>. The fourth difference lies in the conformation adopted by the residues between 441-426 of the  $\alpha$  subunit.

It is interesting to note that though BPDO<sub>I19</sub> shares lower sequence identity with BPDO<sub>RHA1</sub> as compared to CDO<sub>IPO1</sub>, it is structurally more similar to it. BPDO<sub>RHA1</sub> shares 69% sequence identity with BPDO<sub>I19</sub>. A total of 585 C $\alpha$  atoms can be superimposed with an RMS deviation of 0.66 Å ( $\alpha$  subunit; RMS deviation of 0.63 Å for 413 C $\alpha$  atoms,  $\beta$  subunit ; RMS deviation of 0.63 for 172 C $\alpha$  atoms). The structures of BPDO<sub>I19</sub> and BPDO<sub>RHA1</sub> have three main differences between them. They occur in the loop region between  $\alpha$ 8 and  $\beta$ 15 in the  $\alpha$  subunit (which is believed to act as lid covering the channel to the active site, at the N-terminal end of the  $\beta$  subunit, where BPDO<sub>I19</sub> has a small  $\alpha$  helix which is absent in BPDO<sub>RHA1</sub>), and the loop between  $\alpha$ 13 and  $\alpha$ 14 in the  $\alpha$  subunit. The RMS deviation of the  $\alpha$  subunits between BPDO<sub>I19</sub> and naphthalene-1,2- dioxygenase from *Rhodococcus* species (NDO<sub>12038</sub>), nitrobenzene dioxygenase from *Comamonas* species strain JS765 (NBDO<sub>JS765</sub>) naphthalene-1,2- dioxygenase from *Pseudomonas putida*(NDO<sub>9816-4</sub>) and Biphenyl dioxygenase from *Sphingobium yanoikuyae* strain B1 (BPDO<sub>B1</sub>) ranges from 1.22 to 1.26 Å (Table 3.2 ) which suggests that the core structures are highly conserved in this family of enzymes. It is of significance that BPDO<sub>I19</sub> differs more from BPDO<sub>B1</sub> than CDO<sub>IPO1</sub> as far as the overall structure is concerned.

Further differences are observed in the active site cavity of Group III and IV ROs. In BPDO<sub>I19</sub> the active site cavity has a volume of 1410.7 Å<sup>3</sup>, as calculated by the USF program

Table 3.2. Sequence and structural statistics for various Rieske oxygenases compared to BPDO<sub>II9</sub>.

Terminal oxygenase $\alpha$ Subunit(PDB ID)	RMSD (C $\alpha$ )	Sequence identity (%)	Sequence similarity (%)
CDO <sub>IPO1</sub> (1WQL)	1.12(432)	75	85
BPDO <sub>RHA1</sub> (1ULI)	0.63(413)	69	82
NDO <sub>12038</sub> (2B1X)	1.26(334)	36	51
NBDO <sub>JS765</sub> (2BMQ)	1.23(352)	32	49
NDO <sub>9816-4</sub> (1NDO)	1.22(355)	32	48
BPDO <sub>B1</sub> (2GBW)	1.26(349)	39	53

Terminal oxygenase $\beta$ Subunit(PDB ID)	RMSD (C $\alpha$ )	Sequence identity (%)	Sequence similarity (%)
CDO <sub>IPO1</sub> (1WQL)	0.69(169)	59	71
BPDO <sub>RHA1</sub> (1ULI)	0.63(172)	61	80
NDO <sub>12038</sub> (2B1X)	0.89(157)	39	62
NBDO <sub>JS765</sub> (2BMQ)	1.12(147)	29	47
NDO <sub>9816-4</sub> (1NDO)	1.17(150)	24	42
BPDO <sub>B1</sub> (2GBW)	0.97(152)	31	52

VOIDOO using a probe radius of 1.20 Å. Compared to BPDO<sub>II9</sub>, the total volume of the active site cavity in CDO<sub>IP01</sub>, BPDO<sub>RHA1</sub>, NDO<sub>9816-4</sub> is significantly smaller (259.6, 1022.5 and 479.7 Å<sup>3</sup> respectively). This largest cavity of BPDO<sub>II9</sub> supports the observation that it has better substrate specificity and selectivity than other known BPDOs and NDOs. Biphenyl rings of BPDO<sub>II9</sub> are in closer proximity of some residues (Gln 226, Phe 227, His 233, His 323, Met 231, Ala 234, His 239, Phe 277, Ser 283, Val 287, Gly 321, Ile 336, Phe 378, and Phe 384). These residues are either different in different BPDOs (Table 3.3) and NDOs or adopt different conformations in all known BPDOs and NDOs structures. Therefore, they all might be responsible for substrate specificity and regio-specificity of BPDO<sub>II9</sub>. In NDO<sub>9816-4</sub> the active site is a flat, elongated cavity whereas it is L-shaped in BPDO<sub>II9</sub>.

Phylogenetically, BPDO<sub>II9</sub> and CDO<sub>IP01</sub> are more similar to known potent PCB metabolizing ROs in Group IV. The incongruity of grouping CDO<sub>IP01</sub>, which oxidizes cumene, essentially a truncated biphenyl molecule, with known PCB degrading enzymes is explained by a more detailed comparison of the active site cavities. The residue Ala 321 of CDO<sub>IP01</sub>, which is structurally analogous to Gly 321 in BPDO<sub>II9</sub>, obstructs access to the substrate pocket, resulting in a significantly smaller cavity than in BPDO<sub>II9</sub>. This constriction of the CDO<sub>IP01</sub> active site dictates its preference for cumene as a substrate. The BPDO<sub>II9</sub> active site cavity is also larger than that of BPDO<sub>RHA1</sub>. The orientation of the substrate, biphenyl, is quite similar in BPDO<sub>B1</sub> and BPDO<sub>RHA1</sub>. Biphenyl binding in the BPDO<sub>RHA1</sub> biphenyl binary complex places the 2,3-carbons approximately 4.5 Å from the Fe(II) with a dihedral angle between the two aromatic rings of 123.98° (Furusawa *et al.*, 2004). Compared to BPDO<sub>RHA1</sub>, binding of the biphenyl requires little adjustment of the substrate pockets in BPDO<sub>II9</sub>. Upon binding of biphenyl, the sidechain of Amino acid Phe 277, Ser 283, Gly 321, Phe 378 and Phe284 of



BPDO<sub>II9</sub> have significant large shift. Similarly, there is a large shift in BPDO<sub>RHA1</sub> upon biphenyl binding in the analogous residues. This movement is part of an overall shift in structure to accommodate the biphenyl substrate. Structure analysis shows how the mutations that occurred on residues T335 F336 of BPDO-LB400 to generate variant BPDO<sub>II9</sub>, which have Gly and Ile at 335 and 336 position respectively. These residues altered the shape of the substrate binding pocket of BPDO<sub>II9</sub> and these changes are presumed to allow binding of various PCBs more efficiently than parent enzyme. The larger active site cavity of BPDO<sub>II9</sub> explains its greater reactivity to a broad spectrum of substituted biphenyls, such as PCB congeners (Barriault *et al.*, 2002). Therefore, the arrangement of residues and overall dimensions of the BPDO<sub>II9</sub> active site cavities may provide a structural explanation of why Group IV ROs clustered with BPDO<sub>II9</sub> include the potent PCB degrading enzymes.

#### **3.4.7. Docking of Polychlorinated Biphenyls (PCBs).**

Four PCBs (2,2' di chlorobiphenyl, 3, 3' di chlorobiphenyl, 2,4,4' tri chlorobiphenyl and 2,2',5,5' tetra chlorobiphenyl) were docked in BPDO<sub>II9</sub> and BPDO<sub>LB400</sub>. The complex structures of BPDO<sub>II9</sub> with four PCB's were compared with the biphenyl complex of BPDO<sub>II9</sub> and BPDO<sub>LB400</sub> (Fig. 3.11). It was observed that structure of BPDO<sub>II9</sub> is able to accommodate all four PCB's without any hindrance whereas BPDO-LB400 and other known BPDOs have steric conflict to place all these PCB's in the active site. Docking data is consistent with the observation that BPDO<sub>II9</sub> exhibited high activity toward 2,2'-, 3,3'-, and 4,4'-dichlorobiphenyls unlike its parent enzyme (Barriault D., *et al.* 2002).

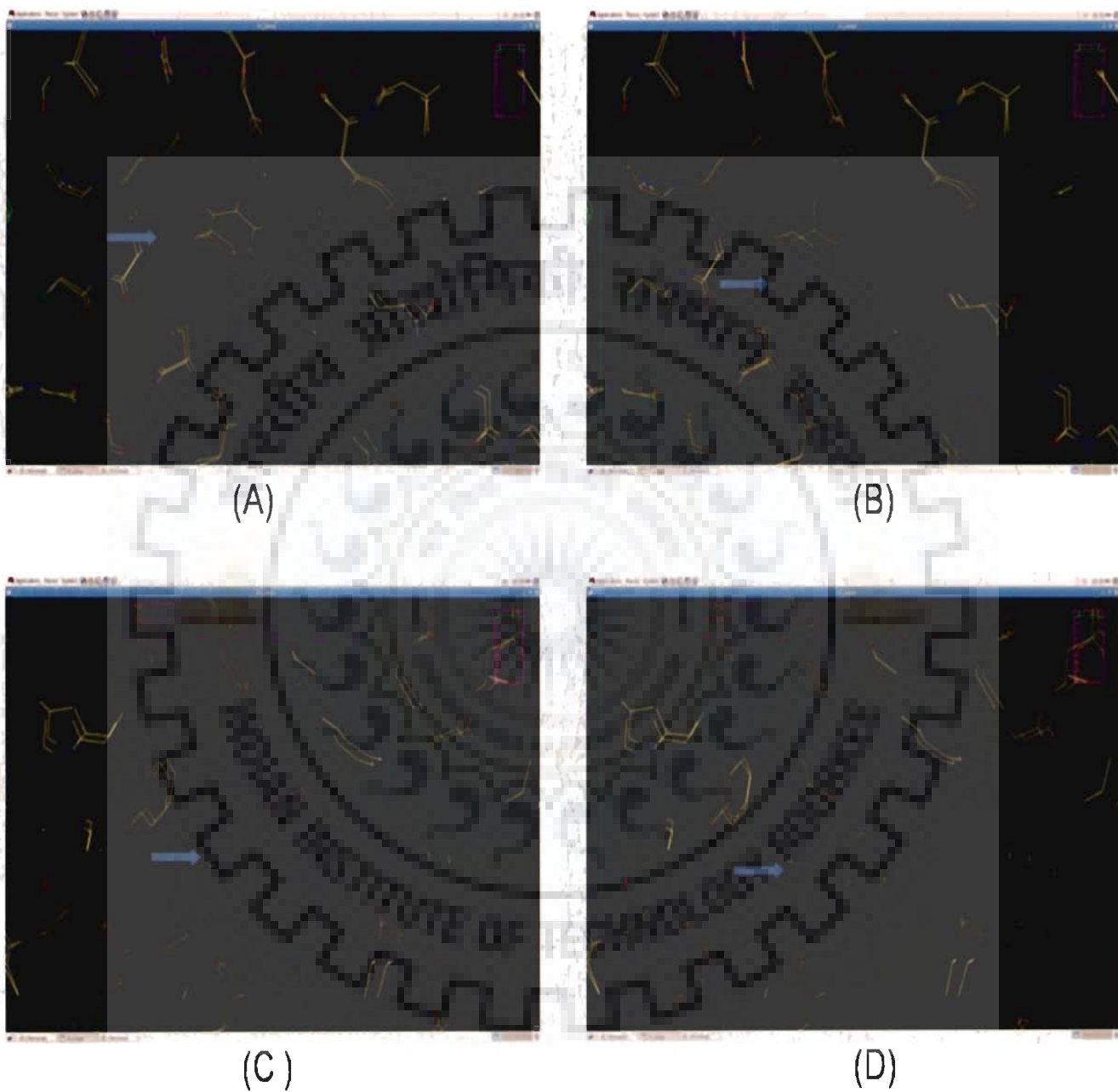


Fig. 3.11. PCBs docked in BPDO<sub>119</sub> (A) 2,2' di chlorobiphenyl, (B) 3, 3' di chlorobiphenyl, (C) 2,4,4' tri chlorobiphenyl and (D) 2,2',5,5' tetra chlorobiphenyl). The figures are snapshots of the dockings visualized in O molecular graphics program (Jones et al, 1991). The arrows in blue indicate the position of the docked molecules.



### 3.5. CONCLUSIONS

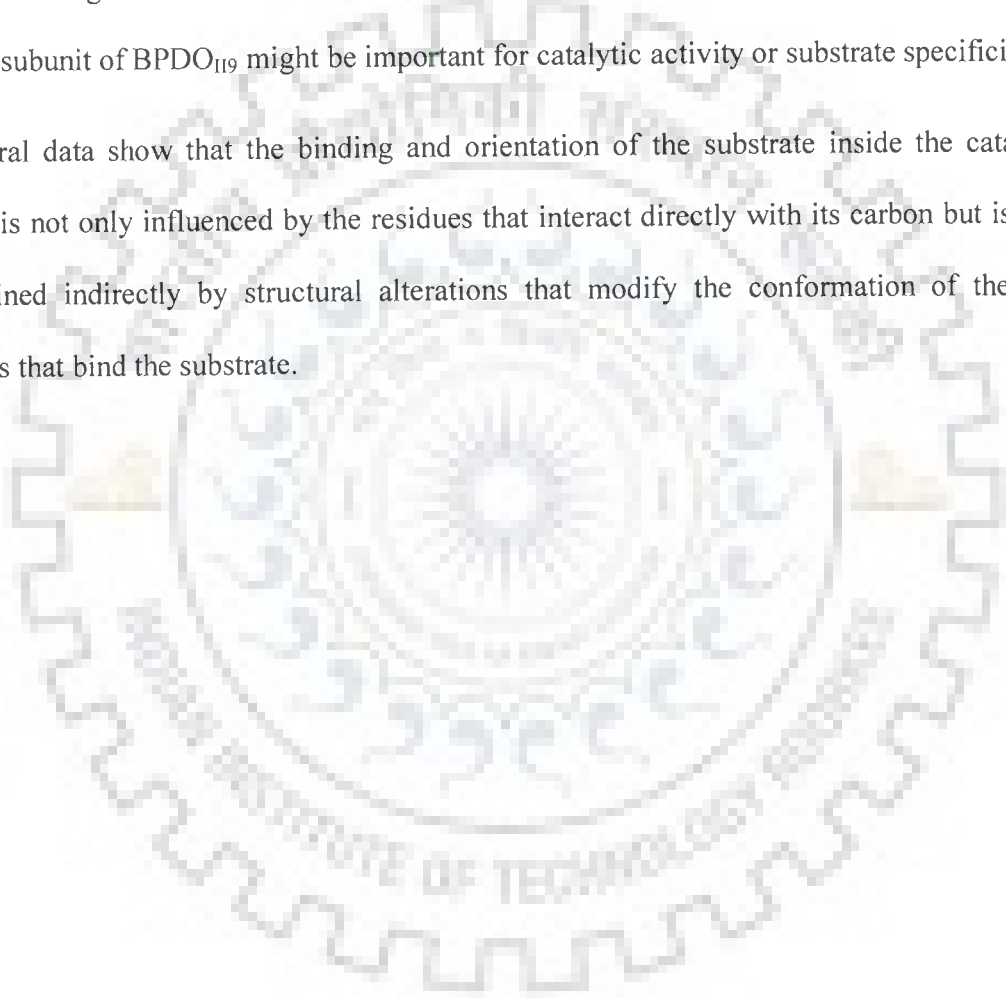
Crystal structures of BPDO<sub>119</sub> and its complex are presented and demonstrate strong structural conservation with other ROs.

- ❖ BPDO<sub>119</sub> is  $\alpha_3\beta_3$  hexamer similar to Napthalene 1,2-Dioxygenases (NDOs) and other Biphenyl Dioxygenases (BPDOs).
- ❖ The  $\alpha$  subunit can be divided into two distinct domains: a Rieske domain that contains the Rieske [2Fe-2S] center and the catalytic domain that contains the active site mononuclear non-heme ferrous iron.
- ❖ The iron atom at the active site is pentacoordinated by His233, His239, Asp 388 and by two water molecules and the geometry changes when it binds substrate and becomes tetracoordinated as only one water molecules coordinates in this form.
- ❖ It is proposed that the electron transfers from Fe-S cluster of one  $\alpha$ -subunit to the mononuclear iron of an adjacent  $\alpha$ -subunit like in other BPDO's and NSOs.
- ❖ The active site entrance residues for BPDOs and NDOs adopt different conformations.
- ❖ Differences in the substrate-binding pocket may be responsible for substrate specificity by affecting substrate orientation.
- ❖ BPDO<sub>119</sub> and shares about 34-75% identity with other BPDOs but they still have significant differences in their active site and some loop regions. Many amino-acids mainly, Gln 226, Phe 227, Asp 230, Met 231, His 233, Ala 234, Val 287, Gly 321, His 323, Leu 333, Ile 336, Phe 378, Phe 384, Ala 256, Pro 259, Tyr 277, Ser 283, Met

288, Gly 335, Ile, 336, Gly 321 and Phe 378 in BPDO<sub>119</sub> might have significant role in their substrate preference. Moreover, Identical amino-acids in the other BPDOs structures have different orientation.

- ❖ Some amino-acids at the interface of  $\alpha\beta$  subunits near the active site forms hydrogen bonding network in BPDO<sub>119</sub> whereas it is not present in NDO. This indicates that  $\beta$  subunit of BPDO<sub>119</sub> might be important for catalytic activity or substrate specificity.

Structural data show that the binding and orientation of the substrate inside the catalytic pocket is not only influenced by the residues that interact directly with its carbon but is also determined indirectly by structural alterations that modify the conformation of the key residues that bind the substrate.





# Chapter - 4

*Studies on 3-deoxy-D-  
arabino-heptulosonate  
7-phosphate Synthase*

#### 4.1. ABSTRACT

3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase is the first enzyme of the shikimate pathway, responsible for the synthesis of aromatic amino acids in microorganisms and plants. The pathway has been growing in interest in the recent past as the enzymes of the pathway are being targeted for antimicrobial drug design and herbicides. In the present work, the structure studies of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from the plant *Arabidopsis thaliana* (*At*-DAHPS) is described. The enzyme was overexpressed in *E.coli*, purified and crystallization attempts were made in order to grow high quality crystals, however all crystals only diffracted to 7.0 Å resolution limit. Various attempts were made to improve the crystal quality using different additives and gels but there was no improvement in crystal quality. Therefore, structure of *At*-DAHPS was predicted using bioinformatics approaches and also inferences regarding phylogenetic ancestry were drawn. The structure shows that the enzyme belongs to  $(\beta/\alpha)_8$  TIM barrel family and has residues in the active site which are conserved in the DAHP synthase type II enzymes. The structure of *At*-DAHPS provides a structural framework that may be useful to design specific inhibitors towards herbicide development.

#### 4.2. INTRODUCTION

The enzyme 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase (EC 4.1.2.15), the first enzyme of the shikimate pathway, catalyzes the condensation of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to form DAHP and inorganic phosphate. The biosynthesis of the intermediate compounds, chorismate and prephenate, that serve as precursors to the aromatic amino acids (Phe, Tyr, and Trp), catechols, and *p*-aminobenzoic acid (folic acid biosynthesis) and a number of other highly important microbial

## *Studies on 3-deoxy-D-arabino-heptulosonate 7-phosphate Synthase*

---

and plant compounds takes place via this pathway (Walsh *et al.*, 1996). The pathway intermediates also act as branch point compounds that may serve as substrates for other metabolic pathways (Hermann and Weaver, 1999). The shikimate enzymes have been considered as targets for drug development against a plethora of diseases caused by microorganisms and as targets for herbicide development for plants because the pathway occurs in microorganisms and plants but not in animals. Disruption of the shikimate pathway can lead to rapid death of the organism. This principle is made use of by the commercially available herbicide, *Roundup*, which contains N-phosphonomethylglycine (glyphosate), a powerful inhibitor of 5-enolpyruvyl shikimate 3-phosphate synthase (Steinrucken and Amrhein, 1990), an enzyme downstream of DAHP synthase (the sixth enzyme) in the shikimate pathway. Also, many of the apicomplexan parasites are serious pathogens of humans and domestic animals. Because these apicomplexan parasites are eukaryotes and share many metabolic pathways with their animal hosts, therapy against them is extremely difficult; a drug that harms an apicomplexan parasite is also likely to endanger its human host. Currently there are no effective vaccines or treatments available for most of the diseases caused by these parasites. Biomedical research on these parasites is challenging because it is often difficult, if not impossible, to maintain live parasite cultures in the laboratory and to genetically manipulate these organisms.

There can be two distinct apparently unrelated types of DAHP synthase proteins, classified on the basis of amino acid sequence homology and molecular mass (Konig *et al.*, 2004). Type I DAHP synthases are smaller than type II DAHP synthase, found mainly in prokaryotic and archaeal organisms, although some eukaryotic examples have been identified (e.g., those from *S. cerevisiae*) (Hartmann *et al.*, 2003; Paravicini *et al.*, 1989) and



*Neurospora crassa* (Hoffmann *et al.*, 1972; Nimmo and Coggins, 1981). Type II DAHP synthase proteins were originally identified in plants (Bentley, 1990) and are now known to encompass a diverse set of microbial proteins of which the plant proteins form a subcluster (Gosset *et al.*, 2001). DAHP synthase is ubiquitous in archaea, bacteria, fungi and plants. Uptil now, the structures of DAHP synthase from six organisms - *Saccharomyces cerevisiae*, *Escherichia coli*, *Thermotoga maritima*, *Aeropyrum pernix*, *Mycobacterium tuberculosis* and *Pyrococcus furiosus* have been solved. Knowledge of the structure of the enzyme in plants is of utmost importance since it is seen as a target for herbicide development. Any new inhibitor for the enzyme must have high specificity for the enzyme in the particular microorganism or unwanted weeds, since it is also present in the plants of interest. Based on the lack of structural information of a DAHP synthase from plants, an investigation of a DAHP synthase from a plant was initiated in order to provide further insight into the reason for the bifurcation in the DAHP synthase phylogenetic tree as well as to provide a structural framework on which specific structure based inhibitor design could be established in the direction of making herbicides.

*Arabidopsis thaliana* has three genes *DHS1*, *DHS2*, and *DHS3* that yield three *A. thaliana* DAHP synthase isoforms. The three isoforms have over 75% sequence identity. Fig. 4.1 shows the multiple sequence alignment of the three isoforms. Here, the overexpression, purification, crystallization, three dimensional structure prediction and phylogenetic analysis of the *At*-DAHPS enzyme encoded by *DHS1* is reported. Despite exhaustive crystallization trials, good diffracting crystals could not be obtained. Therefore, *in-silico* modeling was done to predict the structure. The results presented here show that *At*-DAHPS belongs to type II DAHP synthase enzyme, shares the  $(\beta/\alpha)_8$  TIM barrel fold as is characteristic for all other

DAHPS synthase, shares a common catalytic site with the type I DAHPS family and proves conclusively that it has some unique features which prevent it from being regulated by feedback inhibition.

### **4.3. MATERIAL AND METHODS**

#### **4.3.1. Isolation, Purification and crystallization of *At*-DAHPS**

##### **4.3.1.1. Chemicals and Biochemicals**

*N*-(2-Hydroxymethyl)piperazine-*N*-(3-propanesulfonic acid) (EPPS), DTT, isopropyl -d-thiogalactoside, PEP and E4P were from Sigma (St. Louis). *Escherichia coli* BL21 (DE3) and plasmid pET23d were from Novagen (Madison, WI). All other reagents were of the highest purity commercially available. The crystallization screens were from Hampton Research (USA).

##### **4.3.1.2. Expression of *Arabidopsis* DHS1 in *E. coli***

The expression vector pET23d harbouring the desired gene was a generous gift from Prof. Klaus Herrmann, Purdue University, USA. *E. coli* BL21 (DE3) was used as a host strain for the maintenance of plasmids and for expression of *DHS1* using *E. coli* BL21 (DE3) / pET23dDHS1. Eight milliliters of an overnight culture of this strain were inoculated into 1 L of Luria broth (Himedia, India) containing 50 mg of ampicillin. The culture was grown at 37°C to an optical density of 0.6 at 600 nm. At that point, isopropyl-d-thiogalactoside (final concentration 0.4 mM) was added and induced culture was grown for about 4 h.

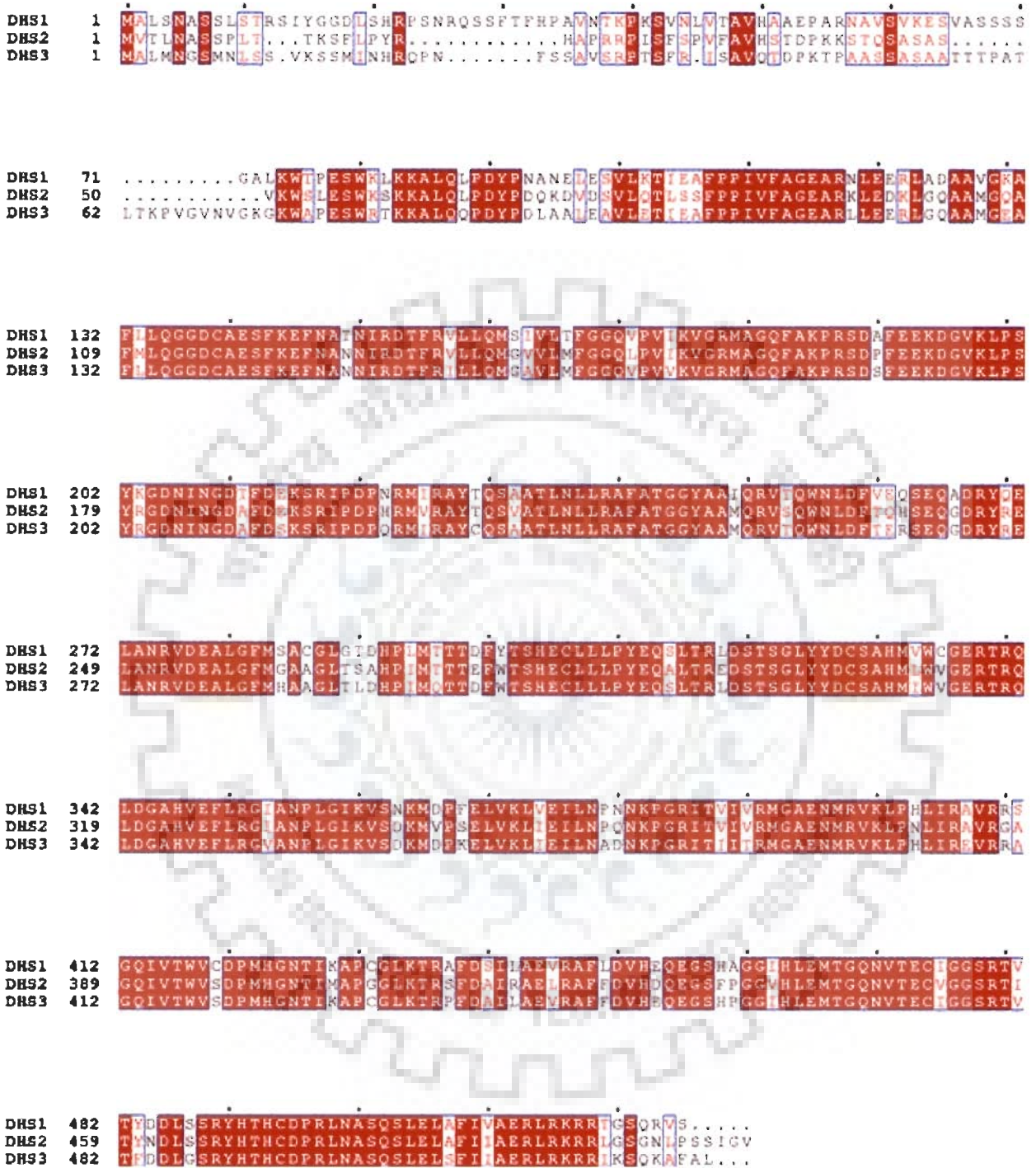


Fig. 4.1. Sequence alignment of the *At*-DAHPS isozymes. The sequence alignment of the three isozymes of *At*-DAHPS, encoded by DHS1 (Accession no. NP\_195708.1), DHS2 (Accession no. NP\_195077.1) and DHS3 (Accession no. NP\_173657.1) genes. Figure generated by EsPript (Gouet *et al.*, 1999).

#### **4.3.1.3. Purification of *Arabidopsis DHS1* encoded *At*-DAHPS**

All purification steps were carried out at 4°C. The cells were harvested by centrifugation, washed, and resuspended in buffer A (20 mM potassium phosphate, pH 7.4, containing 1 mM PEP and 0.2 mM DTT). The cells were lysed using ultrasonicator. The resulting cell extract was clarified by centrifugation for 20 min at 15,000g in an RC-5b centrifuge (Sorvall, Newton, CT). The supernatant was loaded onto a cation exchange Hi-Trap (SP HP) column (GE Healthcare, USA) equilibrated with buffer A. The protein was eluted with a linear gradient of 0.05 to 1 M potassium phosphate, pH 7.4, containing the buffer A supplements. Fractions containing enzyme activity were checked for purity on SDS PAGE gel, pooled, dialysed against EPPS buffer and concentrated to 8 mg/ml by ultrafiltration using Amicon filter (30 kDa cut off) (Millipore, USA). Concentrated protein was used for crystallization trials. The remaining fraction was stored at -80°C in aliquots for further use.

#### **4.3.1.4. Protein estimation**

The concentration of protein in the sample was estimated by the method of Bradford (Bradford, 1976). Bovine serum albumin (BSA) was used as protein standard.

#### **4.3.1.5. Polyacrylamide gel electrophoresis (SDS-PAGE)**

The purity test and further analysis of purified protein were performed using SDS and PAGE Laemmli (Laemmli, 1970) with little modifications. The various gel electrophoresis constituents and the compositions of the gel are listed below.

##### **4.3.1.5.1. Preparation of Reagents**

The stock solutions of various gel components were prepared as mentioned below.

## *Studies on 3-deoxy-D-arabino-heptulosonate 7-phosphate Synthase*

---

1. **Acrylamide solution (30%):** 30% w/v acrylamide solution containing 0.8% w/v N, N-methylene-bis-acrylamide: 29.2 g of acrylamide and 0.8 g of bisacrylamide was dissolved in 70 ml of deionised water. When acrylamide was completely dissolved, deionised water was added to make a final volume of 100 ml. Solution was filtered with a whatman No.1 paper and was stored at 4°C in a dark bottle.
2. **Resolving buffer (1.5 M Tris-HCl, pH 8.8):** 18.2 g of Tris base was dissolved in 80 ml of deionised water and the pH was adjusted to 8.8 with HCl and deionised water was added to make a final volume of 100 ml and was stored at 4°C.
3. **Stacking buffer (0.5 M Tris-HCl, pH 6.8):** 6.1 g of Tris base was dissolved in 80 ml of deionised water and pH was adjusted to 6.8 with HCl and deionised water was added to make a final volume of 100 ml and was stored at 4°C.
4. **Sodium dodecyl sulphate (SDS, 10% w/v):** 10 g SDS was dissolved in 60 ml of deionised water and kept at room temperature overnight without shaking. Then the volume was leveled at 100 ml by deionised water and was stored at room temperature.
5. **Catalyst: 10% ammonium per sulphate (APS):** 10 mg APS was dissolved in 100 µl of deionised water (Freshly prepared solution was used).
6. **TEMED (N, N, N, N-tetra methylethylenediamine):** It was used undiluted from the bottle stored at cool, dry and dark place.
7. **Electrode buffer:** (0.025 M Tris, 0.192 M glycine, 0.1% (w/v) SDS, pH 8.3) 0.3 g Tris base, 1.4 g glycine, 1 ml 10% SDS/100 ml electrode buffer. Electrode buffer was also prepared as stock solution 5X concentration, consisting of 15 g Tris

base, 72 g glycine and 5 g SDS/litre. It was stored at room temperature and was diluted to 5 times by adding 4 parts of deionised water before use.

8. **Stock sample buffer (2X)** (0.125 M Tris pH 6.8, 2.5% SDS, 20% glycerol, 0.002% Bromophenol blue, 10%  $\beta$ -mercaptoethanol).

The sample buffer (2X) was prepared by mixing the stock solutions as per the given composition.

#### 4.3.1.5.2. Preparation of sample buffer

Ingredient	Volume
Water	3.0 ml
0.5M Tris-HCl, pH 6.8	2.5 ml
10% SDS	2.5 ml
Glycerol	2.0 ml
0.5% Bromophenol blue (w/v)	0.5 mg

Stored at room temperature, SDS-reducing sample buffer was prepared by adding 100  $\mu$ l of  $\beta$ -mercaptoethanol to each 0.9 ml of stock sample buffer, before use.

#### 4.3.1.5.3. Casting of gel

Twelve percent denaturing discontinuous gel was prepared by mixing gel stock solutions as per given composition. The monomer solution was prepared for resolving gel by mixing all of the reagents given below except the ammonium per sulfate and TEMED.

#### Recipe for resolving gel (12%): (10 ml)

<b>Ingredient</b>	<b>Volume</b>
Acrylamide solution 30%	3.3 ml
1.5 M Tris-HCl pH 8.8	2.5 ml
10% SDS	0.1 ml
10% APS	0.1 ml
Water	4.0 ml
TEMED	4.0 $\mu$ l

**Recipe for stacking gel (5%): (5 ml)**

<b>Ingredient</b>	<b>Volume</b>
Acrylamide solution 30%	0.83 ml
0.5 M Tris-HCl pH 6.8	0.63 ml
10% SDS	0.05 ml
10% APS	0.05 ml
Water	3.40 ml
TEMED	5.00 $\mu$ l

The solution was degassed for at least 3-5 min. The APS and TEMED were gently mixed into the degassed monomer solution. The solution was well mixed uniformly and poured gently in between the plates. The resolving gel was cast up to 2/3 height on pre marked plates followed by layering of 200  $\mu$ l butanol overlaying solution. After 15 min, the demarcation occurred between the acrylamide layer and butanol layer indicated the complete polymerization of gel. Butanol was decanted and the space was washed with distilled water. Similarly 5% stacking gel was also layered on top of the resolving gel. The wells were cast in

stacking gel by placing the teflon comb in between and at the top of the two plates. After 15-20 min the comb was removed carefully.

#### **4.3.1.5.4. Sample Preparation**

Protein concentration 150 µg/50 µl were mixed with 1 volume of sample treatment buffer (0.125 M Tris pH 6.8, 2.5% SDS, 20% glycerol, 0.002% bromophenol blue, 10% β-mercaptoethanol) and heated to boiling point for 2 min.

#### **4.3.1.5.5. Electrophoresis**

Electrophoretic separation was done by using BIO-RAD Mini-PROTEAN® 3 Cell electrophoresis unit. The prepared samples were loaded into the wells with different protein concentration of each sample and electrophoresed at 80 V through stacking gel. Once the sample was concentrated at the interface of the stacking and separating gel as sharp blue line, the voltage was increased to 120 V and the electrophoresis was continued until the tracking dye reached at the bottom of the gel.

#### **4.3.1.5.6. Staining and destaining of gel**

After the run, gel was removed from plates and kept in staining solution (0.1% coomassie brilliant blue R-250 w/v in 40% methanol, 10% acetic acid w/w) for 4-6 h with mild shaking at room temperature. Then, the gel was destained with several changes of destaining solution I (40% methanol, 10% glacial acetic acid) and finally kept in destain II (10% glacial acetic acid).

Gels after electrophoresis were photographed and analyzed using Gel documentation system (BioRad, USA).



#### 4.3.1.6. DAHP Synthase Enzyme Assay

DAHP synthase was assayed (Suzich *et al.*, 1985) by measuring the  $A_{549}$  of the periodate degradation product of DAHP complexed with thiobarbiturate. The unit of activity is defined as the amount of protein catalyzing the appearance of 1 mmol DAHP  $\text{min}^{-1}$ .

#### 4.3.1.7. Crystallisation trials

Crystals of DAHP synthase were attempted by both hanging drop and sitting drop vapour diffusion methods using the crystal screens, Crystal Screen and Crystal Screen 2 (Hampton Research, USA). We got initial condition from Crystal Screen and it was further refined by using PEG 4000, 0.1 M HEPES pH 7.5 and 10% isopropanol.

#### 4.3.2. Structure Prediction and Phylogenetic Studies

##### 4.3.2.1. Sequence retrieval

Amino acid sequence of the experimentally characterized, DAHP synthase enzyme from *Arabidopsis thaliana* (Accession No. NP\_195708.1) was retrieved from the protein sequence database at National Centre for Biotechnology Information at <http://www.ncbi.nlm.nih.gov> as shown below.

MALSNASSLSTRSIYGGDLSHRPSNRQSSFTFHPAVNTKPKSVNLVTAVHAAEPARN  
AVSVKESVASSSSGALKWTPESWKLKALQLPDYPNANELESVLKTIEAFPPIVFAG  
EARNLEERLADAAVGKAFLQGGDCAESFKEFNATNIRDTRVLLQMSIVLTFGGQV  
PVIKVGRMAGQFAKPRSDAFEKDGVKLPSYKGDNINGDTFDEKSRIPDPNRMIRAY  
TQSAATLNLLRAFATGGYAAIQRVTQWNLDFVEQSEQADRYQELANRVDEALGFM  
SACGLGTDHPLMTTDFYTSHECLLLPYEQLTRLDSTSGLYYDCSAHMVWCGERT  
RQLDGAHVEFLRGIANPLGIKVS NKMDPFELVKLVEILNPNNKPGRITVIVRMGAEN

MRVKLPHLIRAVRRSGQIVTWVCDPMHGNTIKAPCGLKTRAFDSILAEVRAFLDVHE  
QEGSHAGGIHLEMTGQNVTECIGGSRTVTYDDLSSRYHTHCDPRLNASQSLELAFIV  
AERLRKRRTGSQRVS

To create a model of the *At*-DAHPS, first a BLAST search (both non-redundant and pdb databases) for proteins with similar sequence was performed using the sequence of the enzyme in *Arabidopsis thaliana* as query. The *Mt*-DAHPS sequence was identified to be showing highest identity with the query sequence amongst those whose 3D structure is available in database. The 3D structures of proteins were obtained from the protein data bank ([www.rcsb.org](http://www.rcsb.org)). Short segment from the N- and C- terminal end was eliminated and the sequence from residue 26-515 of *At*-DAHPS was used for further work.

#### **4.3.2.2. Secondary Structure Prediction and Sequence alignment**

Secondary Structure prediction and sequence alignment was done using ESPript (using *Mt*-DAHPS secondary structure as reference) (Gouet *et al.*, 2003) and ClustalW (Labarga *et al.*, 2007).

#### **4.3.2.3. Molecular modeling**

The three dimensional structure of *At*-DAHPS was predicted by the method of homology modeling. Homology modeling is used to construct an atomic model of the target protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (template). Homology modeling relies on the identification of one or more known protein structures, which shares considerable homology about 30% with the query sequence. The basis of this computational technique is that two proteins with similar sequences would adopt similar tertiary structures (Sali and Overington, 1994; Blundell *et al.*,

1987). There are two main approaches to homology modeling: (1) fragment-based comparative modeling (Blundell *et al.*, 1987, Blundell *et al.*, 1988) and (2) by satisfaction of spatial restraints (Sali and Blundell, 1993). For modeling of the *At*-DAHPS, we used the second approach using the program Modeller 9v1. It generates a refined 3D homology model of a protein sequence, utilising as input its alignment with a related structure, which serves as the template. Modeller uses the concept of satisfaction of the spatial restraints derived from the alignment which it expresses as probability density functions (PDFs) for the features restrained. The PDFs are derived analytically using statistical mechanics and empirically using a database of known protein structures. Then the spatial restraints and CHARMM energy terms enforcing proper stereochemistry (Brooks *et al.*, 1983) were combined into an objective function. Finally the model was obtained by optimizing the objective function in Cartesian space. To create the final 3D model, a set of 100 models for *At*-DAHPS were created by Modeller9v1 using the template *Mt*-DAHPS and the best one was selected on the basis of stereochemical quality. The crystal structure of *Mt*-DAHPS (PDB code 2B7O) was used as template in the modeling of *At*-DAHPS. The substrates were docked into the active site of *At*-DAHPS using Hex docking program (Ritchie *et al.*, 2003).

#### **4.3.2.4. PDB viewers**

Models were viewed and useful conclusions drawn after analysis with PyMOL (<http://pymol.sourceforge.net/index.html>) (DeLano, 2002) and COOT (<http://www.yasbl.york.ac.uk/~emsley/coot/>) (Emsley and Cowtan, 2004).

#### **4.3.2.5. Validation of the model**

Stereochemical quality of the model was assessed by Ramachandran plot analysis, which allows identification of the number of residues with non-ideal torsion angles. Ramachandran plot was retrieved from RAMPAGE (Lovell *et al.*, 2002).

#### **4.3.2.6. Phylogenetic studies**

Phylogenetic tree was constructed using MEGA 3.1 phylogenetic software (Tamura *et al.*, 2007).

### **4.4. RESULTS AND DISCUSSION**

#### **4.4.1. Isolation, Purification and Crystallisation of *At*-DAHPS**

##### **4.4.1.1. Expression and purification of *Arabidopsis DHS1* encoded *At*-DAHPS**

*At*-DAHPS was expressed in *E. coli* cells as mentioned in the material and methods section. The expression was checked on SDS PAGE. Fig. 4.2 (A) shows the expression profile. Total protein from *E. coli* was extracted as described in methods section and tested for DAHP synthase activity. The fractions having activity were pooled. Proteins were eluted in phosphate buffer with continuous gradient of 0 - 1 M Sodium Phosphate (pH 7.4). A peak was obtained with 50 mM Sodium Phosphate (pH 7.4). The elution profile is shown in Fig. 4.2 (B). The fractions of the peak containing DAHP synthase activity (P) (Fig. 4.2 (B)) were pooled. Pooled sample was dialyzed against 25mM EPPS buffer containing 1mM PEP and 0.2 mM DTT and concentrated using Amicon 30 kDa cut off membrane and was used for crystallization.

#### **4.4.1.2. SDS PAGE analysis**

To ensure the purity of the purified protein, peak P of SP-HP column which showed DAHP synthase activity was analyzed on SDS-PAGE. The coomassie blue stained 12% SDS-PAGE protein profiles of these fractions are shown in Fig. 4.2 (C). The crude preparation had several proteins but after purification, the coomassie blue stained gel profile of peak P with DAHP synthase activity has only single band of 57 kDa without any other contaminating protein. Thus the 57 kDa protein *At*-DAHPS was purified. It was clear from the gel that purified protein does not have any other contaminating protein even at higher concentration.

#### **4.4.1.3. Crystallisation of *At*-DAHPS**

Crystallisation of purified *At*-DAHPS was performed as described under the section material and methods. These crystals (Fig. 4.3) diffracted only to resolution 7.0 Å and we could not further proceed for structure solution.

### **4.4.2. Structure Prediction and Phylogenetic studies on *At*-DAHPS**

#### **4.4.2.1. Quality of the model**

The overall stereochemical quality assessment on the stereochemical quality of the 100 models was done by the Ramachandran plot from RAMPAGE (Lovell *et al.*, 2002). The best model amongst them in terms of stereochemical quality was selected as the final model. In the final model, 96% residues appeared to be in the Ramachandran plot preferred regions, 2.7 % residues in the allowed regions and 1.3% in the outliers (Fig. 4.4). Based on an analysis of various structures of resolution of at least 2.0 Angstroms and R factor no greater than 20%, a good quality model would be expected to have over 90% in the most favored regions. So, the values indicate that the model is in good agreement with the template. The model energy

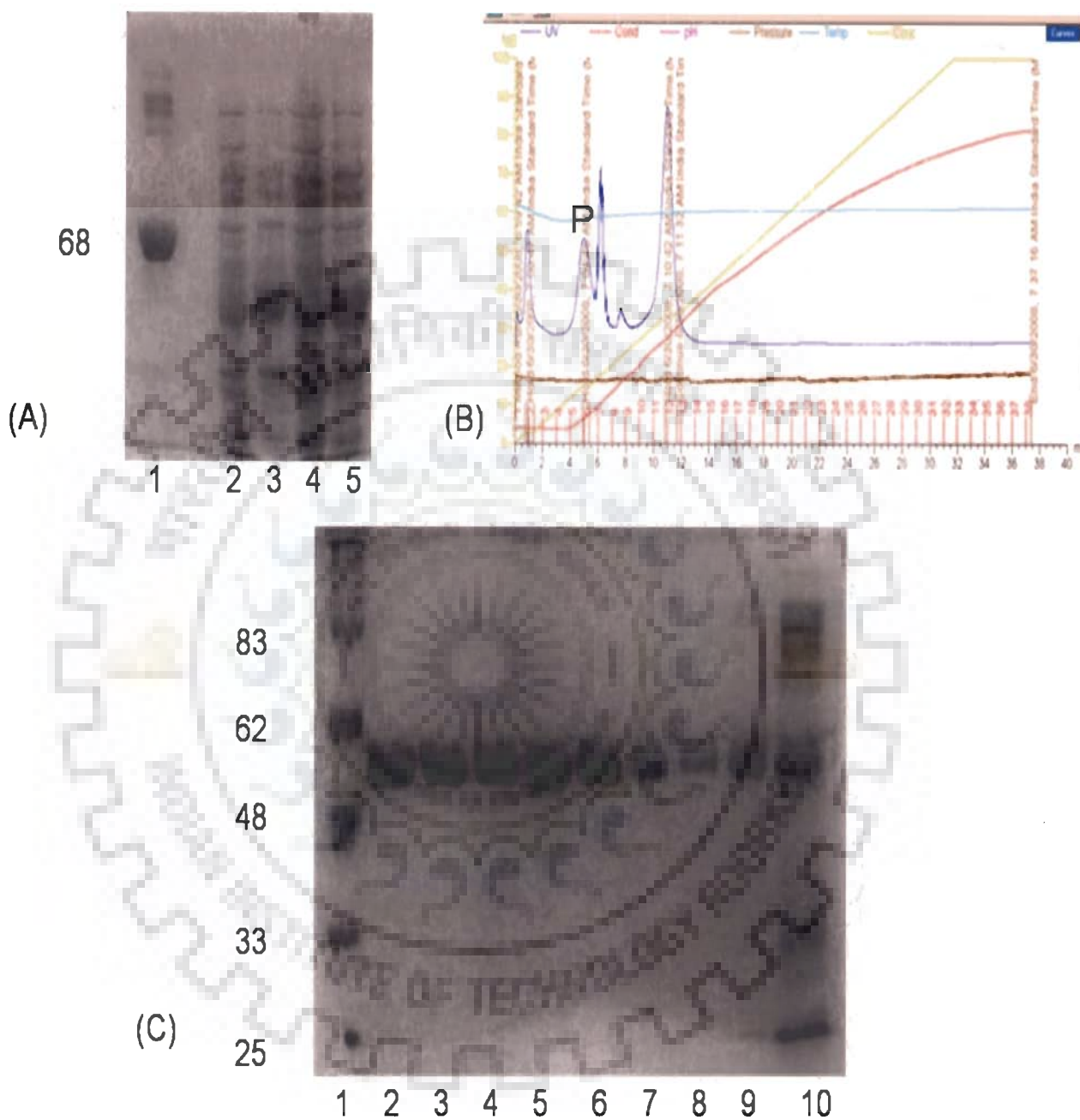


Fig. 4.2. Expression and Purification of *At*-DAHPS. (A) SDS PAGE gel for expression. Lane 1 shows the BSA (Mol. wt. 68 kDa) used as marker, lane 2 BL21(DE3) uninduced 10µg protein, lane 3 BL21(DE3) induced 10µg protein, lane 3 BL21(DE3) uninduced 20µg protein lane 4 BL21(DE3) induced 20µg protein (B) The chromatography profile for purification. 'P' represents the desired peak. (C) SDS PAGE gel for pure fractions. Lane 1 shows prestained  $M_r$  marker (BioRad, USA) with Mol. Wt. in kDa indicated alongside, lane 2-9 pure fractions and lane 10 shows elute in 2M buffer .

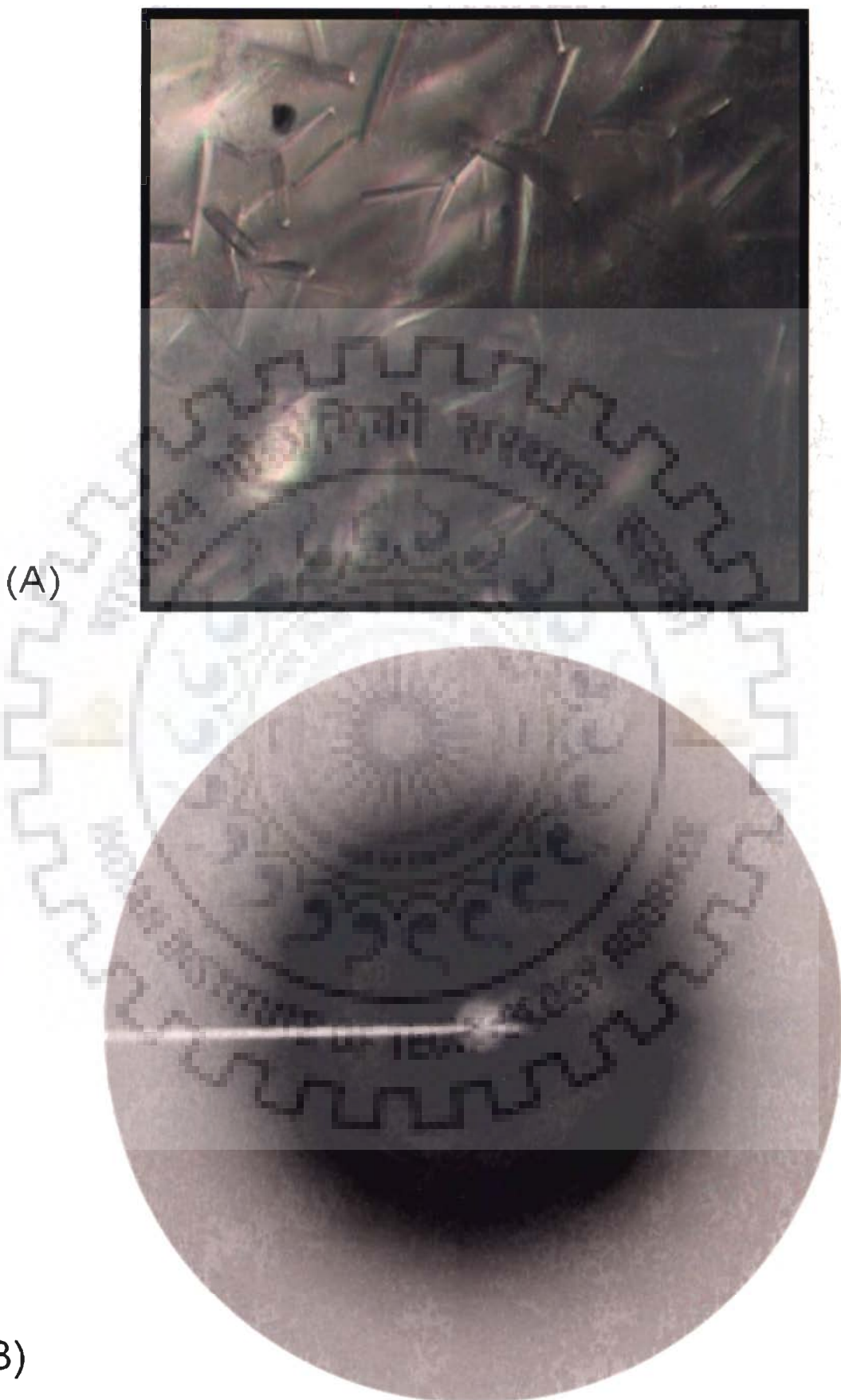
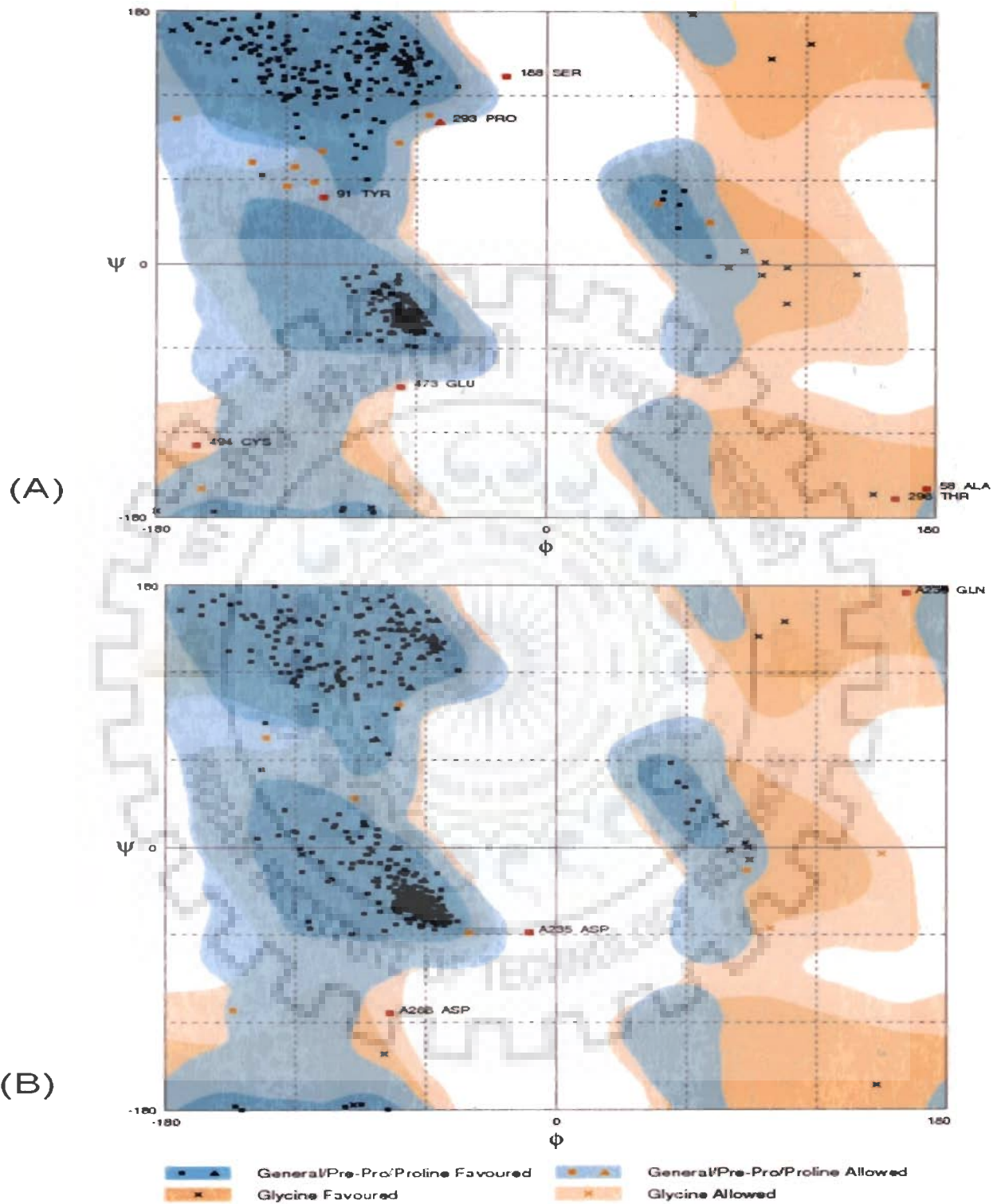


Fig. 4.3. Photographs and diffraction pattern for crystals of *At*-DAHPS. (A) Photographs of the crystals of *At*-DAHPS. (B) Diffraction pattern of crystals of *At*-DAHPS.



No. of Residues	(A)Target ( <i>At</i> -DAHPS)	(B)Template ( <i>Mt</i> -DAHPS)
In favoured region	470 (96.1)	436 (97.5%)
In allowed region	12 (2.5)	8 (1.8%)
In outlier region	7 (1.4)	3 (0.7%)

Fig. 4.4. Ramachandran diagram  $\phi$ - $\psi$  plots for the (A) *At*-DAHPS structure (B) *Mt*-DAHPS structure.



was  $-20,887.291 \text{ kCal mol}^{-1}$  and energy per residue was calculated to be  $-42.627 \text{ kCal Mol}^{-1}$ . The RMSD value for equivalent  $\text{C}\alpha$  atoms between *At*-DAHPS with *Mt*-DAHPS was evaluated to be  $0.31 \text{ \AA}$ .

#### 4.4.2.2. Overall description of the Model

*At*-DAHPS is 525 residue long and as stated above, shows 48% primary sequence identity with 464 residue long *Mt*-DAHPS. The sequence alignment of *At*-DAHPS with *Mt*-DAHPS is shown in Fig. 4.5. This is the highest percentage of sequence identity that it shows with DAHP Synthase from any organism and so it qualifies to be a suitable template for modeling *At*-DAHPS target. Model of target *At*-DAHPS was superposed on the structure of template *Mt*-DAHPS (Fig. 4.6 (A)). The *Mt*-DAHPS is a dimer so the model is only superposed on one subunit. Inspection of the superposition reveals that most of the region of the *At*-DAHPS is similar to the three-dimensional structure of the template. However, there are some regions found to be different than template. Secondary structure prediction reveals that the *At*-DAHPS like all other DAHPS structurally characterized till date consists of a core  $(\beta/\alpha)_8$  TIM barrel domain in which the eight parallel  $\beta$ -strands (residues Phe132–Gly137, Val173–Arg178, Tyr301–Glu305, Met332–Cys335, Leu357–Val361, Ile386–Val390, Thr416–Asp420 and Gly461–Glu465) are each followed by  $\alpha$ -helices (Pro148–Gly169, Pro221–Thr243, Leu309–Leu315, Ala345–Gly352, Pro367–Leu377, Met397–Ser411, Phe438–Glu455, Ala500–Arg523). At the N terminus of *At*-DAHPS, there are three helices  $\alpha_0a$  (Ala72–Lys83),  $\alpha_0b$  (Ala94–Ala106) and  $\alpha_0c$  (Ala113–Val128) instead of  $\beta_0$ -strand as it was in *Mt*-DAHPS. The next pair of helices ( $\alpha_2a$  and  $\alpha_2b$ ) is like the *Mt*-DAHPS, that extend the  $\alpha_2$ – $\beta_3$  connecting loop and pack against the outside of the barrel. Also, there are some small additional stretches of alpha helices within the barrel extending from Glu214–Arg217, Gly245–Ala247

## *Studies on 3-deoxy-D-arabino-heptulosonate 7-phosphate Synthase*

---

and  $\beta$ -strands stretching from Tyr325-Asp327 and Ile427- Lys428. The connecting loops at the C-terminal end of the barrel, where they help to form the active site, are markedly longer than those at the N-terminal end. Fig. 4.7 displays the core and non core structural elements in all DAHP synthase enzymes whose structures have been solved till date. There are two additional  $\beta$  strands between  $\alpha 5$  and  $\beta 6$  that alongwith  $\beta 0$  are involved in binding of feedback inhibitor in the case of type Ia DAHP synthase enzymes, *Ec*- DAHPS (*Escherichia coli* DAHP synthase) and *Sc*- DAHPS (*Saccharomyces cerevisiae* DAHP synthase) (Schofield *et al.*, 2005). These additional elements are absent in the case of *At*-DAHPS consistent with the observation of lack of feedback inhibition in the case of *At*-DAHPS. Instead the residues that surround this region are conserved in *Mt*- DAHPS and *At*-DAHPS hinting at a common mode of enzyme regulation in them if these residues are in fact involved in the enzyme regulation process or a motif that is possibly shared by type II DAHP synthase. Also the two stranded  $\beta$  hairpin that characterizes the base of the barrel opposite the active site in *Pf*-DAHPS (*Pyrococcus furiosus* DAHP synthase) and *Tm*-DAHPS (*Thermotoga maritima* DAHP synthase) (Schofield *et al.*, 2005) is absent in *Mt*- DAHPS and *At*- DAHP synthase. The ferredoxin-like domain that characterizes *Tm*-DAHPS (Schofield *et al.*, 2005) and is implicated in feedback inhibition is also absent in *At*-DAHPS suggesting that if enzyme regulation in *At*-DAHPS takes place then it does via a separate mechanism that could more closely resemble the mode in *Mt*-DAHPS. The additional  $\alpha$  helices at the N-terminal and  $\alpha 2a$  and  $\alpha 2b$  together might be responsible for the metabolic regulation of *At*-DAHPS.

The amino acid sequence of *At*-DAHPS qualifies it to be a type II DAHP synthase enzyme. Presently, *Mt*-DAHPS structure is only available type II DAHP synthase. Therefore, the *At*-DAHPS more closely resembles with *Mt*-DAHPS in comparison with *Pf*-DAHPS, *Tm*-



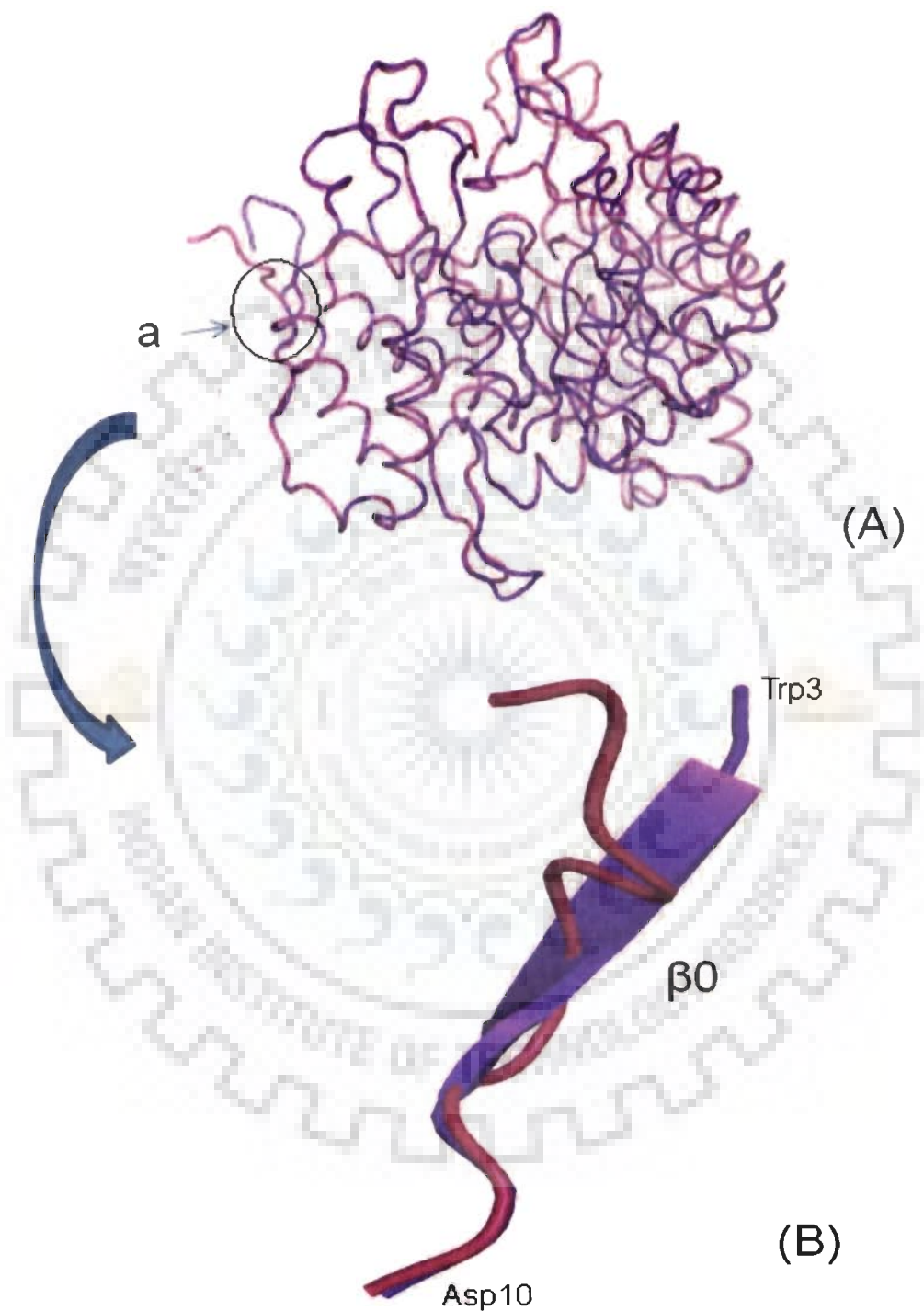


Fig. 4.6. Superimposition of the backbone of *At*-DAHPS (pink) onto *Mt*-DAHPS (purple) (PDB code 2B7O). Though most regions are similar in the two structures, there are few differences between the two. Of special note is the absence of  $\beta 0$  in the *At*-DAHPS structure which is present in *Mt*-DAHPS. The region marked 'a' in top figure (A) indicates this structural difference.

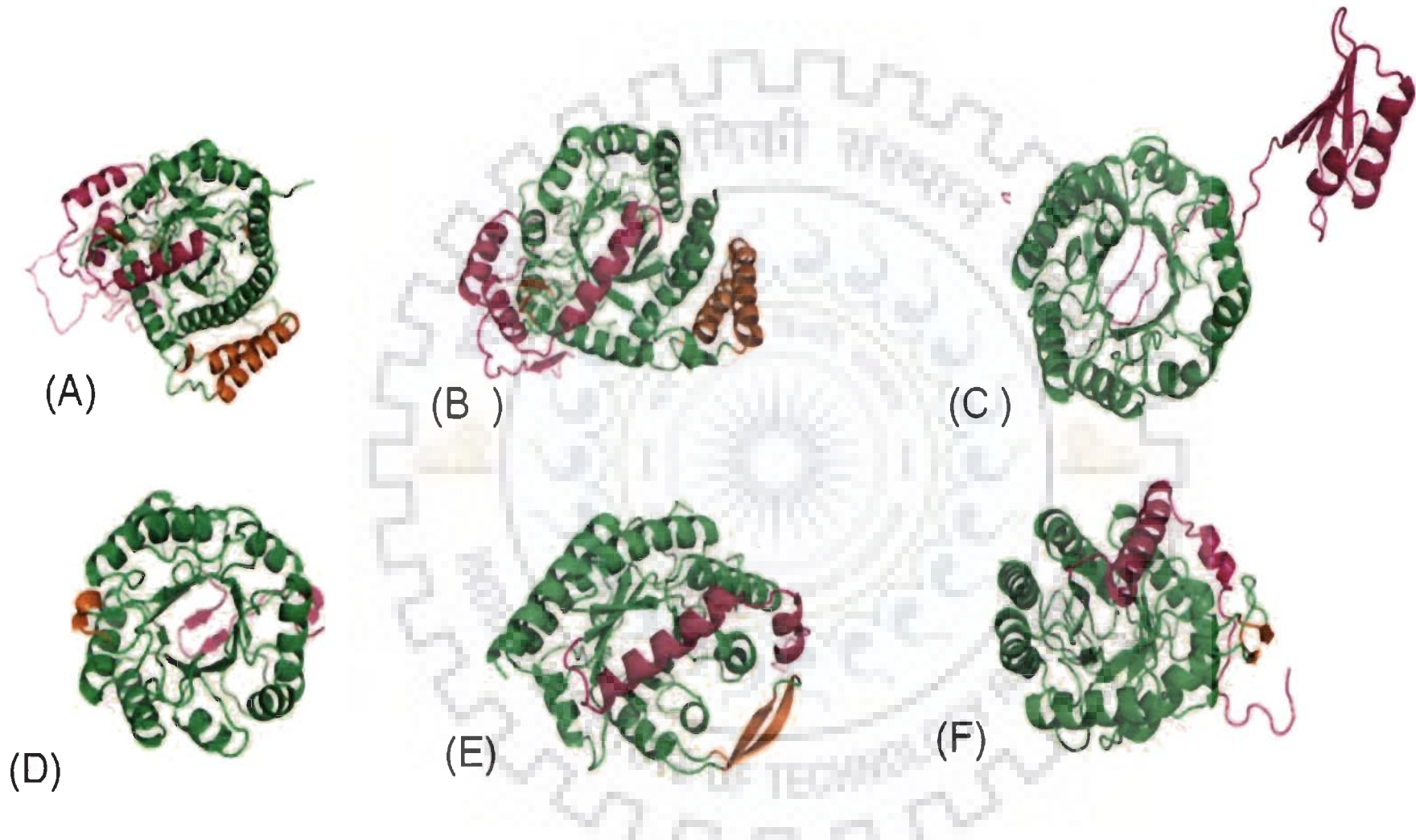


Fig. 4.7. Comparison of the core and non core elements of  $(\beta/\alpha)_8$  barrel of DAHP synthase enzymes. The core  $(\beta/\alpha)_8$ -barrel common to all DAHP synthases is in green. The N-terminal extensions are in magenta and the extra barrel (non core) elements are in orange. The structures shown are those of (A) *At*-DAHPS (predicted) (B) *Mt*-DAHPS (PDB code 2B7O), (C) *Tm*-DAHPS (PDB code 1VR6), (D) *Pf*-DAHPS (PDB code 1ZCO), (E) *Sc*-DAHPS (PDB code 1OFQ), and (F) *Ec*-DAHPS (PDB code 1KFL).

Table 4.1. Distances between the amino acid residues and PEP in *At*-DAHPS.

Phosphoenolpyruvate	<i>At</i> -DAHPS	Distance (Å)
O1	Arg 178 NH2	3.08
O2'	Arg 178 NH1	2.71
O1P	Lys 360 NZ	3.60
O1P	Arg 391 NE	2.88
O3P	Arg 391 NH2	4.79
O3P	Arg 338 NH1	2.38
O2P	Glu 337 N	2.99

DAHPS, *Sc*-DAHPS and *Ec*-DAHPS. This is in correlation with the rmsd values that are obtained when the *At*-DAHPS is superposed on any of these structures. Though there is a resemblance in the overall fold and core barrel structure in all other DAHP synthase enzymes in type I or type II and also the residues in active site but there are differences in the N-terminal region and the noncore  $\alpha$  and  $\beta$  elements that adorn the main barrel that account for the differences in the regulatory mode of the enzyme.

#### **4.4.2.3. Active site organization**

The subunit of *At*-DAHPS consists of a mixed beta sheet flanked by alpha helices as mentioned above. Fig. 4.8. It shows a view looking down the barrel with the rainbow colour scheme to highlight the progression of beta/alpha motifs around the barrel. The beta strands are flanked on either side by alpha helices. The active site of the enzyme is made up of residues located in a deep cleft at the C-end of the barrel as is the characteristic feature for all TIM barrel enzymes. The coordination of PEP and metal in the active site can be predicted with considerable accuracy because it exists as a ligand in the crystal structure of *Mt*-DAHPS which served as a template for modeling *At*-DAHPS. The active site is indicated by the presence of the bound metal ion cofactor, sulfate ion,  $Mn^{2+}$ , and PEP. The active site residues are contributed by the C-terminal ends of the  $\beta$ -strands and the connecting loops between  $\beta$ -strands and  $\alpha$ -helices. Fig. 4.9 shows the active site residues.

The metal binding for  $Mn^{2+}$  was predicted to be coordinated by the residues Cys139, His423, Glu465, Asp495 which are conserved across the DAHP synthase family. The neighboring residues of the active site residues are Cys 494, Arg 436 and Arg 178, which are conserved in the metal binding site.

## *Studies on 3-deoxy-D-arabino-heptulosonate 7-phosphate Synthase*

---

It is interesting to note that the location of Cys494 is close to metal ligand Cys139. A simple rotation about Cys494 C $\alpha$ -C $\beta$  to another rotamer would allow formation of a disulfide bond with Cys139, precluding metal binding and explaining the need for a reducing agent to maintain full activity of the enzyme as observed in the case of *Mt*-DAHPS residues Cys440 and Cys87.

The phosphate group of PEP is hydrogen bonded by the peptide NH of Glu337 and Arg338 (both located on the  $\beta$ 4- $\alpha$ 4 loop), Lys360 Nz (from the end of strand  $\beta$ 5), Arg391 NE and NH<sub>2</sub> (from the  $\beta$ 6- $\alpha$ 6 loop). The PEP carboxylate group forms a doubly hydrogen bonded ion pair with the conserved residue Arg178 (from the end of strand  $\beta$ 2), receiving hydrogen bonds from its NH1 and NH2 atoms, and is additionally hydrogen bonded to the 3-amino group of Lys360. Each carboxylate oxygen atom thus receives two hydrogen bonds of extremely favorable geometry, indicating exquisite molecular recognition (Table 4.1). Other residues surrounding the PEP in the binding pocket include Glu465, His463, Asp420, Cys139, Gln182, Trp334, Glu305, Gly336, Pro186 and His 423, all of which are conserved in *Mt*-DAHPS. Glu305 is invariant in all DAHP synthase. Trp334 is conserved in the DAHP synthase belonging to type II.

The active site is also likely to contain a binding pocket for the second substrate of the enzyme, E4P. The position where it can be predicted to be situated is near the opening of the barrel with its phosphate group situated at a distance of approximately 9.0 Å from the PEP. The sulphate ion, which marks the likely position of the E4P phosphate group, is bound to Arg187 N3, Arg338 NH<sub>2</sub>, Ser188 O and the peptide NH of Ser188. Arg338 provides a bridge between the PEP and sulfate sites, being hydrogen bonded to both the PEP phosphate group and the sulfate ion. Other residues in the active site surrounding the sulphate ion



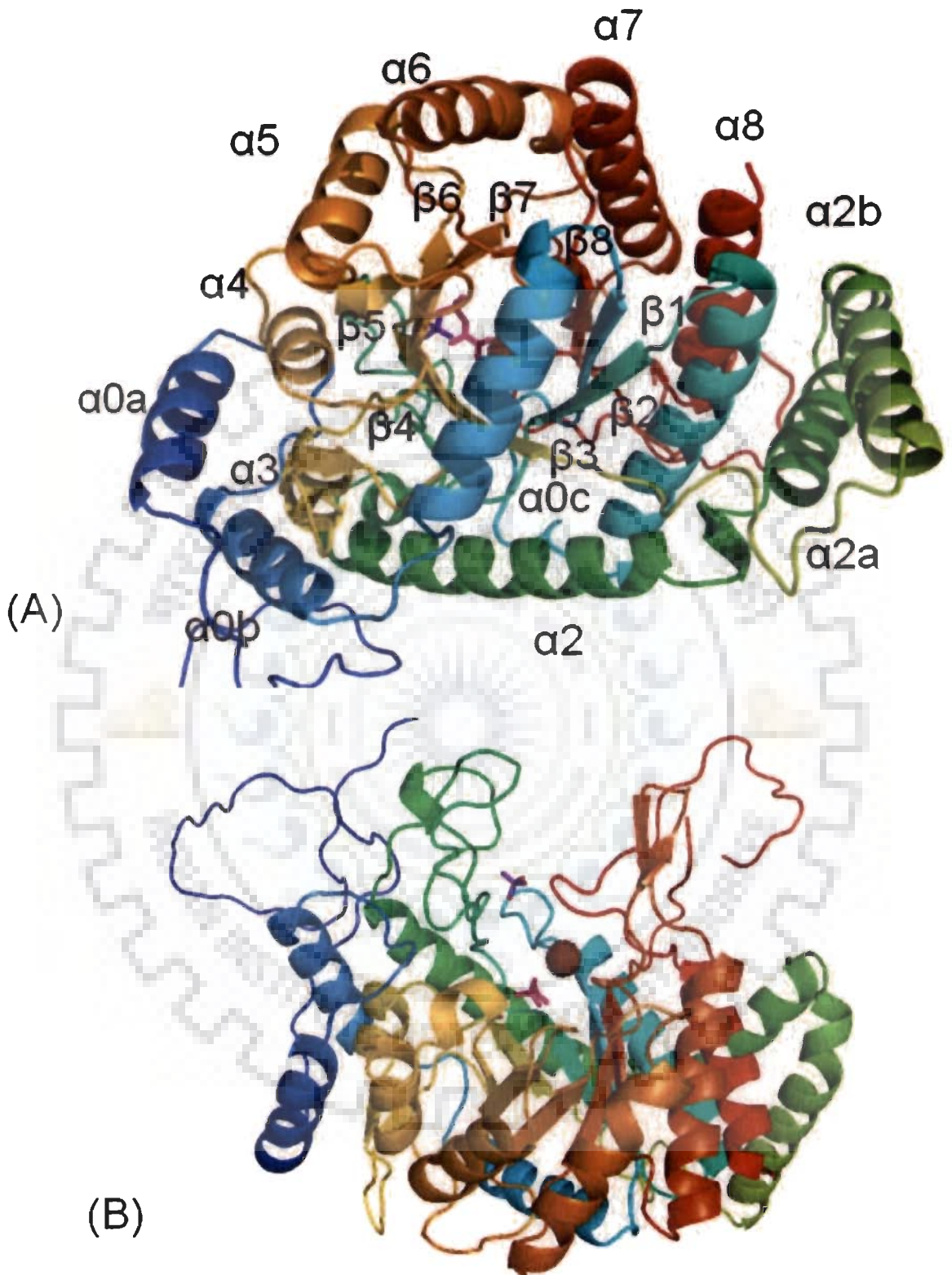


Fig. 4.8. Predicted *At*-DAHPS structure. (A) Stereo view of *At*-DAHPS showing the progression of helices and sheets while looking down into the (( $\beta/\alpha$ )<sub>8</sub>) TIM Barrel. Secondary structural elements are labeled to correspond with references in the text. Also shown bound is  $Mn^{2+}$  (chocolate sphere partially obscured by  $\alpha 0c$ ), PEP (magenta stick model),  $SO_4^{2-}$  (purple stick) that marks the location where E4P would bind. (B) Side-on view of the *At*-DAHPS structure, colour scheme and representation same as in (A).

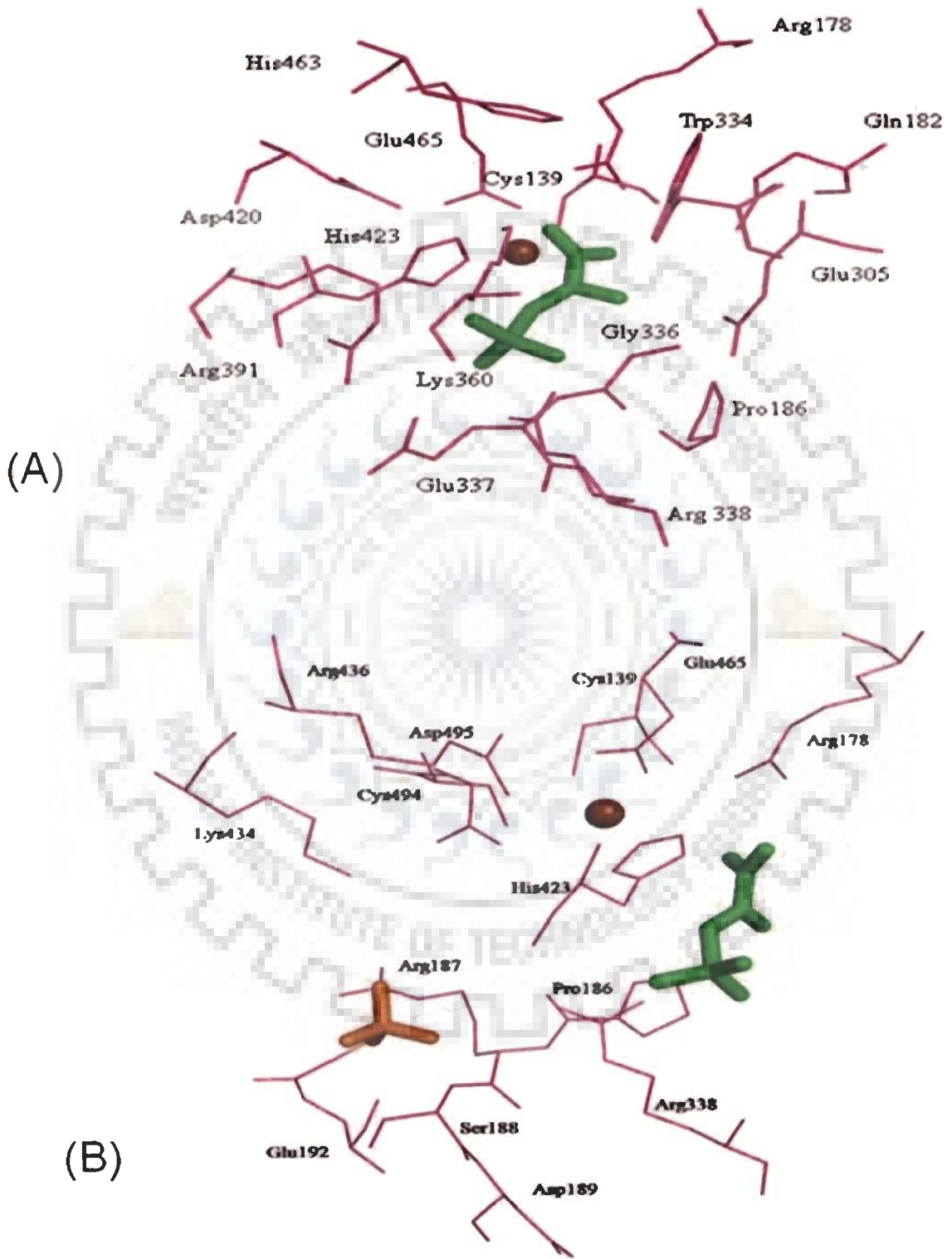


Fig. 4.9. Active site of *At*-DAHPS. (A) Residues lining the catalytic pocket where PEP binds. (B) Residues surrounding the sulphate ion and Mn ion in the binding pocket. Metal ion is shown as sphere (chocolate), PEP (green stick) and sulphate ion (orange stick).

include Pro186, Asp189, Glu 192, Lys434. Pro186 is conserved across the DAHP synthase family.

#### 4.4.2.4. Phylogenetic studies

Crystal structures of the type Ia enzymes from *Escherichia coli* (Shumilin *et al.*, 1999; 2003, Wagner *et al.*, 2000) and *Saccharomyces cerevisiae* (Hartmann *et al.*, 2003; König *et al.*, 2004) and the type Ib enzymes from *Thermotoga maritima* (Shumilin *et al.*, 2004) and *Pyrococcus furiosus* (Schofield *et al.*, 2005) reveal a common ( $\beta/\alpha$ )<sub>8</sub> TIM barrel fold. In some organisms, genes encoding both type I and type II DAHP synthases have been identified, with several of the type II enzymes apparently being required for the biosynthesis of specific secondary metabolites (Silakowski *et al.*, 2000; Guo and Frost, 2002). Sequence identity between the type I and type II DAHP synthases is less than 15 %, however, recent chemical characterization of the type II enzyme from *H. pylori* has indicated many prominent mechanistic features that are shared by both type I and type II DAHP synthases (Webby *et al.*, 2005). The presence of type II DAHP synthases in a number of important pathogenic bacteria makes them prime targets for study.

To establish the evolutionary history of DAHP synthase from *Arabidopsis thaliana*, phylogenetic studies were carried out on DAHP synthase sequences from 23 organisms belonging to diverse classes ranging from Protozoa (*Toxoplasma gondii*), bacteria (both gram -ve and gram +ve), fungi to plants. The resultant tree shows that *Arabidopsis thaliana* is evolutionarily more closely related to the type II DAHP synthase from *Mycobacterium tuberculosis* than to the type I DAHP synthase from *Pyrococcus furiosus*, *Thermotoga maritima*, *Escherichia coli* and *Saccharomyces cerevisiae* as expected. Also the phylogenetic tree places the DAHP synthase from *Toxoplasma gondii* (Fig. 4.10), an apicomplexan

parasite, close to the DAHP synthase type II. Thus studies on the structure of Type II DAHP synthase enzymes could also throw light on combatting these apicomplexan parasites.

#### **4.4.2.5. Comparison of *At*-DAHPS with DAHP synthase Type I**

Given the crucial role of DAHP synthase in the biosynthesis of many key aromatic molecules, the exploitation of differences between the two types of DAHP synthases may enable the development of narrow-spectrum antibiotics (Celia *et al.*, 2005). The predicted 3D structure for *At*-DAHPS shows that it belongs to the  $(\alpha/\beta)_8$  TIM barrel family of proteins as is also found in the type I DAHP synthase exemplified by the 3D structures of DAHP synthase from *E.coli*, *S.cerevisiae*, *P.furiosus* and *T.maritima* solved till date and the closely related 3-deoxy D-manno-octulosonate 8-phosphate (KDO8P) synthases (Radaev *et al.*, 2000). It is important to note that despite extremely low sequence similarity with type Ia DAHP synthase enzyme from *E.coli* and *S.cerevisiae* and type I $\beta_D$  DAHP synthase enzyme from *P.furiosus* and *T.maritima*, the sequence shows a remarkable conservation of metal binding sites and also the residues that surround the active site. All DAHP synthase studied structurally including *At*-DAHPS except the *P. furiosus* enzyme possess extra structural motifs (domains or subdomains) that are involved in allosteric regulation.

An interesting difference between the two types of DAHP synthase is the insertion of two residues within the RxxxxKPRT(S/T) phosphate binding motif of the  $\beta_2$ - $\alpha_2$  loop so that it becomes RxxxxxxKPRS. Motifs associated with allosteric inhibition of *Sc*- and *Ec*-DAHPS and *Tm*-DAHPS are absent in *At*-DAHPS, consistent with the observed absence of feedback inhibition. Also the (I/L)GAR motif in the type I DAHP synthase is substituted with xGxR in the case of type II DAHP synthase.

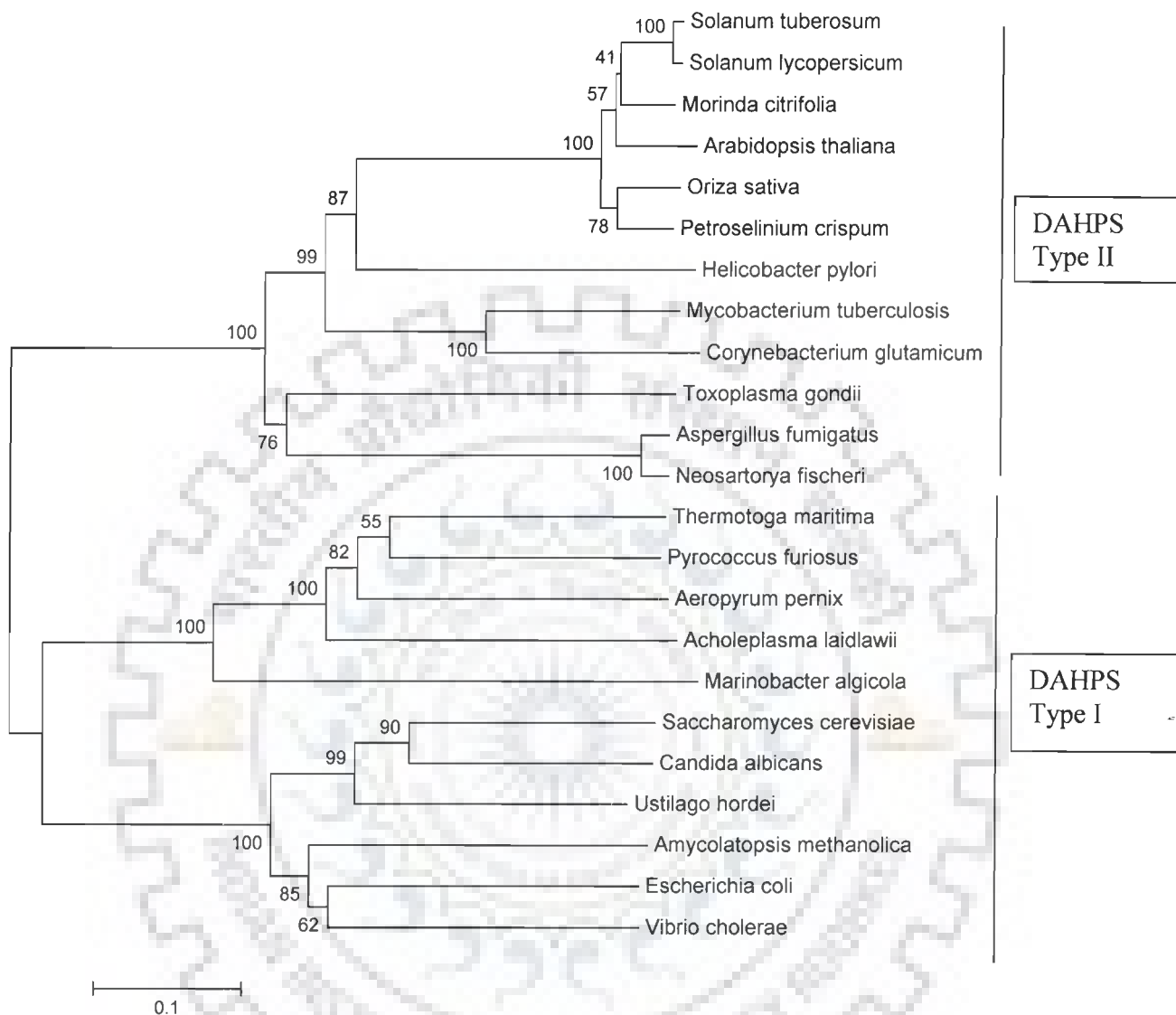


Fig. 4.10. Phylogenetic relationship of DAHP synthase family members and the *At*-DAHPS sequence. This is a bootstrap consensus tree based on 1000 replicates. The numbers on the nodes are bootstrap values. Figure generated by using the program MEGA 3.1 (Tamura *et al.*, 2007).

Another interesting point of difference is that between  $\alpha 5$  and  $\beta 6$ , the Ia DAHP synthase enzyme from *E. coli* and *S. cerevisiae* have a 18-residue extension that together with strand  $\beta 0$ , completes a three stranded sheet involved in binding of the feedback inhibitor. *Pf* and *Tm* feature instead a two stranded  $\beta$  hairpin that covers the base of the ( $\beta/\alpha$ ) barrel. This appears to be absent in the type II DAHP synthase. Instead a motif NK(I)PGR(K) is present in the sequences studied of DAHP synthase type II including *At*-DAHPS. The significance of this motif is yet to be elucidated.

Superimposition of  $C\alpha$  traces of *At*-DAHPS over *Mt*-DAHPS involves 1748 atoms with an rms difference of 0.31 Å. To test the extent of structural similarity despite the lack of recognizable sequence identity we also carried out the  $C\alpha$  superposition of *At*-DAHPS with *Pf*-DAHPS, *Tm*-DAHPS, *Sc*-DAHPS and *Ec*-DAHPS. In the case of *Pf*-DAHPS, 20 atoms matched with an rms deviation of 2.16, 28 atoms matched in the case of *Tm*-DAHPS with an rms deviation of 2.13 Å, and the number of atoms was 48 in the case of *Sc*-DAHPS with an rms deviation of 2.15 Å. The *Ec*-DAHPS could not be superposed over *At*-DAHPS for lack of sufficient structural similarity shared between the two.

#### **4.5. CONCLUSIONS**

In recent years, the methods used to predict the 3D structure of a protein starting from its sequence has improved in accuracy and statistical robustness (Thornton, 2001). Though the conservation of the structure can be inferred from the high sequence similarity, it is well known that in a single family the function and the fold can be retained even if proteins have a low sequence similarity. In this work, homology modeling of the DAHP synthase enzyme from *Arabidopsis thaliana* is presented using Modeller9v1.

## *Studies on 3-deoxy-D-arabino-heptulosonate 7-phosphate Synthase*

---

The knowledge of the 3D structures of the DAHP synthase enzyme from plants is important to better understand their biological function and to delineate structure-function relationships. Such knowledge can also be useful in explaining known properties of the enzymes in plants as well as in designing new experiments to study these properties further. For example, it would be helpful to have an outline of the structure of a protein before carrying out site-directed mutagenesis and such a structure could also help in interpreting results from such investigations. In light of these, the 3D structure model obtained in this study can be used to identify the residues that are likely to alter the donor and/ or acceptor substrate specificities, thereby facilitating their use in the chemoenzymatic synthesis of inhibitor molecules.

The modelled structure of *At*-DAHPS reveals an appealing correspondence of both fold and function in the two types of DAHP synthases, despite the low sequence identity, with a common catalytic apparatus on a shared  $(\beta/\alpha)_8$  TIM barrel fold in agreement with the crystal structure data of *Mt*-DAHPS. The model strongly indicates that the PEP binding domain residues and the metal binding residues constitute well-conserved motifs in DAHP synthase structures. Furthermore, the alignment of DAHP synthase sequences indicates that the main residues involved in intermolecular hydrogen bonds are conserved in all sequences. Such observation suggests that competitive inhibitors with PEP may be able to inhibit most of the DAHP synthases. Further inhibition experiments may confirm this prediction. However the efficiency of inhibition will depend on the design of specific inhibitors against the DAHP synthase of the targeted organism. There is no doubt that the shikimate pathway enzymes offer many important targets for the development of drugs against bacterial, fungal and apicomplexan parasitic diseases. All of the enzymic steps are of potential interest and the availability of structures for almost all of the pathway enzymes will allow rational inhibitor







# Bibliography

## BIBLIOGRAPHY

- Abramowicz D. A., (1990). "Aerobic and anaerobic biodegradation of PCBs: A review". *CRC Critical Reviews in Biotechnolog.* **10**, 241-251.
- Addison R. F., Ikonomou M. G. and Stobo W. T., (1999). "Polychlorinated dibenzo-p-dioxins and furans and non-ortho- and mono-ortho-chlorine substituted polychlorinated biphenyls in grey seals (*Halichoerus grypus*) from Sable Island, Nova Scotia, in 1995", *Mar. Environ. Res.*, **47** (3), 225–240.
- Adebusoye S. A., Picardal F. W., Ilori M. O., Amund O. O. and Fuqua C., (2007c.) "Characterization of multiple novel aerobic polychlorinated biphenyl (PCB)-utilizing bacteria strains indigenous to contaminated African soils", *Biodegradation*, [doi:10.1007/s10532-007-9122-x](https://doi.org/10.1007/s10532-007-9122-x).
- Adebusoye S. A., Picardal F. W., Ilori M. O., Amund O. O., Fuqua C. and Grindle N., (2008). "Metabolism of chlorinated biphenyls: Use of 3,3'- and 3,5-dichlorobiphenyl as sole sources of carbon by natural species of *Ralstonia* and *Pseudomonas*", *Chemosphere*, **70**, 656–663.
- Adebusoye S. A., Picardal F. W., Ilori M. O., Amund O. O., Fuqua C. and Grindle N., (2007a). "Aerobic degradation of di- and trichlorobenzenes by two bacteria isolated from polluted tropical soils", *Chemosphere*, **66**, 1939–1946.
- Adebusoye S. A., Picardal F. W., Ilori M. O., Amund O. O., Fuqua C. and Grindle N. (2007b). "Growth on dichlorobiphenyls with chlorine substitution on each ring by bacteria isolated from contaminated African soils". *Appl. Microbiol. Biotechnol.*, **74**, 484–492.

- Ahmed M. and Focht D.D., (1973). "Degradation of polychlorinated biphenyls by two species of *Achromobacter*", *Can. J. Microbiol.*, **19**, 47–52.
- Anderson K. S. and Johnson K. A., (1990). "Kinetic and structural analysis of enzyme intermediates: Lessons from EPSP synthase", *Chem. Rev.*, **90**, 1131-1149.
- Anderson K. S., Sammons R. D., Leo G. C., Sikorski J. A., Benesi A. J. and Johnson K. A., (1990). "Observation by <sup>13</sup>C NMR of the EPSP synthase tetrahedral intermediate bound to the enzyme active site", *Biochemistry*, **29**, 1460-1465.
- Aoki Y., (2001). "Polychlorinated biphenyls, Polychlorinated Dibenzo-*p*-dioxins, and Polychlorinated dibenzofurans as endocrine disrupters - what we have learned from Yusho Disease.", *Environmental Research Section A*, **86**, 2-11.
- Arnett C. M., Parales J. V. and Haddock J. D., (2000). "Influence of chlorine substituents on the rates of oxidation of chlorinated biphenyls by the biphenyl dioxygenase of *Burkholderia* sp. strain LB400" *Appl. Environ. Microbiol.*, **66**, 2928–2933.
- Asturias J. A. and Timmis K. N., (1993). "Three different 2,3-dihydroxybiphenyl-1,2-dioxygenase genes in the gram-positive polychlorobiphenyl-degrading bacterium *Rhodococcus globerulus* P6", *Journal of Bacteriology*, **175**, 4631–4640.
- Bailey S., (1994). "The CCP4 suite: programs for protein crystallography", *Acta Crystallographica, Section D* **50**, 760-763.
- Balasubramanian S., Abel, C. and Coggins J. R., (1990). "Observation of an isotope effect in the chorismate synthase reaction", *J. Am. Chem. Soc.*, **112**, 8581-8583.

- Baldwin B. R., Nakatsu C. H. and Nies L., (2003). "Detection and enumeration of aromatic oxygenase genes by multiplex and real-time PCR" *Appl Environ Microbiol*, **69**, 3350–3358.
- Barriault D., Plante M. M., and Sylvestre M., (2002). "Family shuffling of a targeted bphA region to engineer biphenyl dioxygenase", *J. Bacteriol.*, **184**, 3794–3800.
- Barriault D., Vedadi M., Powlowsk J. and Sylvestre M., (1999). "cis- 2,3 dihydroxybiphenyl dehydrogenase and cis-1,2-dihydroxynaphthalene dehydrogenase catalyze dehydrogenation of the same range of substrates", *Biochem Biophys Res Commun.*, **260**, 181–187.
- Bedard D. L., Unterman R., Bopp L. H., Brennan M. J., Haberl M. L. and Johnson C.,(1986). "apid assay for screening and characterizing microorganisms for the ability to degrade polychlorinated biphenyls", *Applied and Environmental Microbiology*, **51**, 761- 768.
- Bender S. L., Widlanski T. and Knowles J. R., (1989). "Dehydroquinase synthase: The use of substrate analogues to probe the early steps of the catalyzed reaction", *Biochemistry*, **28**, 7560-7572.
- Bentley R., (1990). "The shikimate pathway: A metabolic tree with many branches", *Crit. Rev. Biochem. Mol. Biol.*, **25**, 307-384.
- Berkaw M., Sowers K.R. and May H.D., (1996). "Anaerobic ortho dechlorination of polychlorinated biphenyls by estuarine sediments from Baltimore Harbor", *Appl. Environ. Microbiol.*, **62 (7)**, 2534–2539.
- Berry E. A., Guergova-Kuras M., Huang L. S. and Crofts A. R., (2000). "Structure and function of cytochrome bc complexes", *Annu. Rev. Biochem.*, **69**, 1005 – 1075.

Bickel H., Plame L. and Schultz G., (1978). "Incorporation of shikimate and other precursors into aromatic amino acids and prenylquinones of isolated spinach chloroplasts", *Phytochemistry*, **17**, 119-124.

Birck M R. and Woodard R W., (2001). "Aquifex aeolicus 3-deoxy-D-manno- 2-octulosonic acid 8-phosphate synthase: A new class of KDO 8-P synthase?", *J Mol Evol.*, **52**, 205–214

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

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

Boyd D. R., Sharma N. D. and Allen C. C., (2001). "Aromatic dioxygenases: molecular biocatalysis and applications", *Curr. Opin. Biotechnol.*, **12**, 564–573.

Boyd D. R., Sharma N. D., Sbircea L., Murphy D., Belhocine T., Malone J.F., James S.L., Allen C.C. R. Hamilton J.T.G., (2008). "Azaarene cis-dihydrodiol-derived 2,20-bipyridine ligands for asymmetric allylic oxidation and cyclopropanation" *ChemComm.* , DOI: **10.1039/b814678k**.

Bradford M., (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding", *Anal. Biochem.*, **72**, 248-254.

- Broadus R. M. and Haddock J. D., (1998). "Purification and characterization of the NADH:ferredoxinBPH oxidoreductase component of biphenyl 2,3-dioxygenase from *Pseudomonas* sp. strain LB400", *Archives of Microbiology*, **170**, 106–121.
- Brugna M., Nitschke W., Asso M., Guigliarelli B., Lemesle-Meunier D. and Schmidt C., (1999). "Redox components of cytochrome bc-type enzymes in acidophilic prokaryotes. II. The Rieske protein of phylogenetically distant acidophilic organisms", *J. Biol. Chem.*, **274**, 16766–16772.
- Brünger A. T., (1992). "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures", *Nature*, **355**, 472-475.
- Buchholz B., Reupke B., Bickel H. and Schultz G., (1979). "Reconstruction of amino acid synthesis by combining spinach chloroplasts with other leaf organelles", *Phytochemistry*, **18**, 1109-1111.
- Bush B., Streeter R. W. and Sloan R. J., (1990). "Polychlorobiphenyl (PCB) congeners in striped bass (*Morone saxatilis*) from marine and estuarine waters of New York State determined by capillary gas chromatography", *Arch. Environ. Contam. Toxicol.*, **19** (1), 49–61.
- Campbell S. A., Richards T. A., Mui E. J., Samuel B. U., Coggins J. R., McLeod R. and Roberts C. W., (2004). "A complete shikimate pathway in *Toxoplasma gondii*: an ancient eukaryotic innovation", *Int. J. Parasitol.*, **34**, 5–13.

Capozzi F., Ciurli S. and Luchinat C., (1998). "Coordination sphere versus protein environment as determinants of electronic and functional properties of iron-sulfur proteins", *Struct. Bonding*, **90**, 127–160.

Carrell C. J., Zhang H., Cramer W. A. and Smith J. L., (1997). "Biological identity and diversity in photosynthesis and respiration: Structure of the lumen-side domain of the chloroplast Rieske protein", *Structure*, **5**, 1613–1625.

Case M. E. and Giles N. H., (1968). "Evidence for nonsense mutations in the *amm* gene cluster of *Neurospora crassa*", *Genetics*, **60**, 49–58.

Cerniglia C. E. and Gibson D. T., (1980). "Fungal oxidation of (–)-9,10-dihydroxy-9,10-dihydrobenzo[*a*]pyrene: formation of diastereomeric benzo[*a*]pyrene 9,10-diol 7,8-epoxides", *Proc. Natl. Acad. Sci. USA*, **77**, 4554–4558.

Cerniglia C. E. and Gibson D. T., (1980). "Fungal oxidation of benzo[*a*]pyrene and (–)-*trans*-7,8-dihydroxy-7,8-dihydrobenzo[*a*]pyrene. Evidence for the formation of a benzo[*a*]pyrene 7,8-diol-9,10-epoxide", *J. Biol. Chem.*, **255**, 5159–5163.

Cerniglia C. E., Morgan J. C. and Gibson D. T., (1979). "Bacterial and fungal oxidation of dibenzofuran", *Biochem. J.*, **180**, 175–185.

Charles I. G., Keyte J. W., Brammar W. J., Smith M. and Hawkins A. R., (1986). "The isolation and nucleotide sequence of the complex AROM locus of *Aspergillus nidulans*", *Nucleic Acids Res.*, **14**, 2201–2213.

- Chiche L., Gregoret L. M., Cohen F. E. and Kollman P. A., (1990). "Protein model structure evaluation using the salvation free energy of folding". *Proc Natl Acad Sci USA*, **87**, 3240-3243.
- Cirino P. C. and Arnold F. H., (2002). "Protein engineering of oxygenases for biocatalysis", *Curr. Opin. Chem. Biol.*, **6**, 130–135.
- Cline J. F., Hoffman B. M., Mims W. B., LaHaie E., Ballou D. P. and Fee J. A., (1985). "Evidence for N coordination to Fe in the [2Fe-2S] clusters of Thermus Rieske protein and phthalate dioxygenase from Pseudomonas", *J. Biol. Chem.*, **260**, 3251–3254.
- Coggins J.R., Abell C., Evans L.B., Frederickson M., Robinson D.A., Roszak A.W. and Laphorn A.P., (2003) "Experiences with the shikimate-pathway enzymes as targets for rational drug design", *Biochemical Society Transactions*. **31(3)**.
- Colbert C. L., Couture M. M. J., Eltis L. D. and Bolin J., (2000). "A cluster exposed: Structure of the Rieske ferredoxin from biphenyl dioxygenase and the redox properties of Rieske Fe-S proteins", *Structure*, **8**, 1267–1278.
- Correll C. C., Batie C. J., Ballou D. P. and Ludwig M. L., (1992). "Phthalate dioxygenase reductase: A modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]", *Science*, **258**, 1604–1610.
- Coulter E. D., Moon N., Batie C. J., Dunham W. R. and Ballou D. P., (1999). "Electron paramagnetic resonance measurements of the ferrous mononuclear site of phthalate



dioxygenase substituted with alternate divalent metal ions: Direct evidence for ligation of two histidines in the copper(II)-reconstituted protein”, *Biochemistry*, **38**, 11062–11072.

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

Della-Cioppa G., Bauer S. C., Klein B. K., Shah D. M., Fraley R. T., and Kishore G. M., (1986). “Translocation of the precursor of 5-enolpyruvylshikimate 3-phosphate synthase into chloroplasts of higher plants *in vitro*”, *Proc.Natl.Acad.Sci.*, **83**, 6873-6877.

Della-Cioppa G., Bauer S. C., Klein B. K., Shah D. M., Fraley R. T. and Kishore G. M., (1986). “Translocation of the precursor of 5-enolpyruvylshikimate 3-phosphate synthase into chloroplasts of higher plants *in vitro*”, *Proc.Natl.Acad.Sci.*, **83**, 6873-6877.

Denke E., Merbitz-Zahradnik T., Hatzfeld O. M., Snyder C. H., Link T. A. and Trumpower B. L., (1998). “Alteration of the midpoint potential and catalytic activity of the **Rieske** iron–sulfur protein by changes of amino acids forming hydrogen bonds to the iron–sulfur cluster”, *J. Biol. Chem.*, **273**, 9085–9093.

Dixon R. A. and Palva N. L., (1995). “Stress-induced phenylpropanoid metabolism”, *Plant Cell*, **7**, 1085-1097.

Dong X., Fushinobu S., Fukuda E., Terada T., Nakamura S., Shimizu K., Nojiri H., Omori T., Shoun H. and Wakagi T., (2005). Crystal structure of the terminal oxygenase component of cumene dioxygenase from *Pseudomonas fluorescens*”, *J Bacteriol.*, **187**, 2483-2490.

- Duewel H. S., and Woodard R. W., (2000). "A metal bridge between two enzyme families. 3-Deoxy-D-manno-octulosonate- 8-phosphate synthase from *Aquifex aeolicus* requires a divalent metal for activity", *J. Biol. Chem.*, **275**, 22824-22831
- Duewel H. S., Radaev S., Wang J., Woodard R.W. and Gatti D. L., (2001). "Substrate and metal complexes of 3-deoxy-D-manno-octulosonate-8-phosphate synthase from *Aquifex aeolicus* at 1.9Å resolution: Implications for the condensation mechanism", *J. Biol. Chem.*, **276**, 8393–8402.
- Duewel H. S., Sheflyan G. Y. and Woodard R. W., (1999). "Functional and biochemical characterization of a recombinant 3-deoxy-D-manno-octulosonic acid 8-phosphate synthase from the hyperthermophilic bacterium *Aquifex aeolicus*", *Biochem. Biophys. Res. Commun.*, **263**, 346-351.
- Duncan K., Edwards R. M. and Coggins J. R., (1987). "The pentafunctional arom enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains", *Biochem. J.*, **246**, 375-386.
- Duncan K., Lewendon A. and Coggins J. R., (1984). "The complete amino acid sequence of *Escherichia coli* 5-enolpyruvylshikimate 3-phosphate synthase", *FEBS Lett.*, **170**, 59-63.
- Dyer W. E., Henstrand J. M., Handa A. K. and Herrmann K. M., (1989). "Wounding induces the first enzyme of the shikimate pathway in Solanaceae", *Proc Natl Acad Sci.*, **86**, 7370-7373.

- Eby D. M., Beharry Z. M., Coulter E. D., Kurtz D. M. and Neidle E. L., (2001). "Characterization and evolution of anthranilate 1,2-dioxygenase from *Acinetobacter* sp. strain ADP1", *J. Bacteriol.*, **183**, 109–118.
- Emsley P. and Cowtan K., (2004). "Coot: model-building tools for molecular graphics", *Acta Crystallographica, Section D-Biological Crystallography* **60**, 2126-2132.
- Ensley B. D. and Gibson D. T., (1983). "Naphthalene dioxygenase: purification and properties of a terminal oxygenase component", *J. Bacteriol.*, **155**, 505–511.
- Ensley B. D., Ratzkin B. J., Osslund T. D., Simon M. J., Wackett L. P. and Gibson D. T., (1983). "Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo", *Science*, **220**, 167–169.
- Erø R. W. and Wagner-Dobler I., (1993). "Detection of polychlorinated biphenyl degradation genes in polluted sediments by direct DNA extraction and polymerase chain reaction", *Appl Environ Microbiol.*, **59**, 4065–4073.
- Erickson B. D. and Mondello F. J., (1992). "2-Nucleotide sequencing and transcriptional mapping of the genes encoding biphenyl dioxygenase, a multicomponent polychlorinated-biphenyl-degrading enzyme in *Pseudomonas* strain LB400", *Journal of Bacteriology*, **174**, 2903–2912.
- Eriksson P., Fischer C. and Fredriksson A., (2006). "Polybrominated diphenyl ethers, a group of brominated flame retardants, can interact with polychlorinated biphenyls in enhancing developmental neurobehavioral defects.", *Toxicological Sciences*, **94**, 302-309.

- Facchiano A. M., Stiuso P., Chiusano M. L., Caraglia M., Guiberti G., Marra M., Abbruzzese A. and Colonna G., (2001). "Homology modeling of the human eukaryotic initiation factor 5A (eIF-5A)", *Protein Eng*, **14**, 881-890.
- Fee J. A., Findling K. L., Yoshida T., Hille R., Tarr G. E., Hearshen D. O., Dunham W. R., Day E. P., Kent T. A. and Münck E., (1984). "Purification and characterization of the **Rieske** iron-sulfur protein from *Thermus thermophilus*. Evidence for a [2Fe-2S] cluster having non-cysteine ligands", *J. Biol. Chem.*, **259**, 124-133.
- Fernie K. J., Bortolotti G. R., Smits J. E., Wilson J., Drouillard K. G. and Bird D. M., (2000). "Changes in egg composition of American kestrels exposed to dietary polychlorinated biphenyls", *Journal of Toxicology and Environmental Health Part A*, **60**, 291-303.
- Forster M. J., (2002). "Molecular modeling in structural biology", *Micron*, **33**, 365-384.
- Fortin P. D., Lo A. T., Haro M. A., Kaschabek S. R., Reineke W. and Eltis L. D., (2005). "Directed evolution of a ring-cleaving dioxygenase for polychlorinated biphenyl degradation", *J. Bacteriol.*, **187**, 415-421.
- Francova K., Mackova M., Macek T., Sylvestre M., (2004). "Ability of bacterial biphenyl dioxygenases from *Burkholderia sp.*LB400 and *Comamonas testosteroni* B-356 to catalyse oxygenation of ortho-hydroxychlorobiphenyls formed from PCBs by plants", *Environmental Pollution*, **127**, 41-48.
- Friedrich M. W., (2006). "Stable-isotope probing of DNA: insights into the function of uncultivated microorganisms from isotopically labeled metagenomes" *Curr Opin Biotechnol.*, **17**, 59-66.

Friemann R., Ivkovic-Jensen M. M., Lessner D. J., Yu C. L., Gibson D. T., Parales R. E., Eklund H. and Ramaswamy S., (2005). "Structural insight into the dioxygenation of Nitroarene compounds: the crystal structure of nitrobenzene dioxygenase." *J Mol Biol.*, **348**, 1139-1151

Friesner R.A., Banks J.L., Murphy R.B., Halgren T.A., Klicic J.J., Mainz D.T., Repasky M.P., Knoll E.H., Shelley M., Perry J.K., Shaw D.E., Francis P., Shenkin P.S. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem*;47: 1739–1749.

Fry D. M., (1995). "Reproductive effects in birds exposed to pesticides and industrial chemicals", *Environmental Health Perspectives*, **103**, 165-171.

Fucile G., Falconer S., Christendat D., (2008). "Evolutionary Diversification of Plant Shikimate Kinase. Gene Duplicates PloS", *Genetics*, **4**, 1-15.

Fujita M., Costas M. and Que L. J., (2003). "Iron-catalyzed olefin cis-dihydroxylation by H<sub>2</sub>O<sub>2</sub>: electrophilic versus nucleophilic mechanisms", *J. Am. Chem. Soc.*, **125**, 9912–9913.

Fukuda M., Yasukochi Y., Kikuchi Y., Nagata Y., Kimbara K., Horiuchi H., Takagi M. and Yano K., (1994). "Identification of the bphA and bphB genes of *Pseudomonas* sp. strain KKS102 involved in degradation of biphenyl and polychlorinated biphenyls", *Biochemical and Biophysical Research Communications*, **202**, 850–856.

Furukawa K., (2000). "Engineering dioxygenases for efficient degradation of environmental pollutants", *Curr. Opin. Biotechnol.*, **11**, 244–249.

- Furukawa K., (2006). "Oxygenases and dehalogenases: molecular approaches to efficient degradation of chlorinated environmental pollutants", *Biosci. Biotechnol. Biochem.*, **70**, 2335–2348.
- Furukawa K., Matsumura F. and Tonomura K., (1978). "Alcaligenes and Acinetobacter strains capable of degrading polychlorinated biphenyls", *Agric. Biol. Chem.*, **42**, 543–548.
- Furukawa K., Suenaga H. and Goto M., (2004). "Biphenyl dioxygenases: functional versatilities and directed evolution", *J. Bacteriol.*, **186**, 5189–5196.
- Furukawa K., Tomisuka N., Kabayashi A., (1979). "Effect of chlorine substitution on the bacterial metabolism of various polychlorinated biphenyls", *Appl. Environ. Microbiol.*, **38**, 301–310.
- Furusawa Y., Nagarajan V., Tanokura M., Masai E., Fukuda M. and Senda T., (2004). "Crystal structure of the terminal oxygenase component of biphenyl dioxygenase derived from *Rhodococcus* species strain RHA1", *J Mol Biol.*, **342**, 1041-1052.
- Gakhar L., Malik Z. A., Allen C. C. R., Lipscomb D. A., Larkin M. J., and Ramaswamy S., (2005). "Structure and increased thermostability of *Rhodococcus* sp Naphthalene 1,2-Dioxygenase", *J. Bacteriol.*, **187**, 7222-7231.
- Gasser C. S. and Klee H. J., (1990). "A brassica napus gene encoding 5-enolpyruvylshikimate 3-phosphate synthase", *Nucleic Acids Res.*, **18**, 2821.
- Gosset G., Bonner C. A. and Jensen R. A., (2001). "Microbial origin of plant-type 2-keto-3-deoxy-D-arabino-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and

tryptophan-regulated enzyme from *Xanthomonas campestris*", *J. Bacteriol.*, **183**, 4061-4070.

Gouet P., Courcelle E., Stuart D. I. and Metz F., (1999). "ESPrict: multiple sequence alignments in PostScript", *Bioinformatics*, **15**, 305-308.

Gourley D. G., Coggins J. R., Isaacs N. W., Moore J. D., Charles I. G. and Hawkins A. R., (1994). "Crystallization of a type II dehydroquinase from *Mycobacterium tuberculosis*", *J. Mol. Biol.*, **241**, 488-491.

Gracy J., Chiche L. and Sallantin J., (1993). "Improved alignment of weakly homologous protein sequences using structural information", *Protein Eng.*, **6**, 821-829.

Gruys K. J., Marzabadi M. R., Pansegrau P. D. and Sikorski J.A., (1993). "Steady-state kinetic evaluation of the reverse reaction for *Escherichia coli* 5-enolpyruvyl shikimate 3-phosphate synthase", *Arch. Biochem. Biophys.*, **304**, 345-351.

Guengerich F. P., (2003). "Cytochrome P450 oxidations in the generation of reactive electrophiles: epoxidation and related reactions", *Arch. Biochem. Biophys.*, **409**, 59-71.

Guergova-Kuras M., Kuras R., Ugulava N., Hadad I. and Crofts A. R., (2000). "Specific mutagenesis of the **Rieske** iron-sulfur protein in *Rhodobacter sphaeroides* shows that both the thermodynamic gradient and the pK of the oxidized form determine the rate of quinol oxidation by the bc(1) complex", *Biochemistry*, **39**, 7436-7444.

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

- Guo J. and Frost J. W., (2002). "Kanosamine biosynthesis: a likely source of the aminoshikimate pathway's nitrogen atom", *J. Am. Chem. Soc.*, **124**, 10642–10643.
- Gurbiel R. J., Doan P. E., Gassner G. T., Macke T. J., Case D. A., Ohnishi T., Fee J. A., Ballou D. P. and Hoffman B. M., (1996). "Active site structure of Rieske-type proteins: Electron nuclear double resonance studies of isotopically labeled phthalate dioxygenase from *Pseudomonas cepacia* and Rieske protein from *Rhodobacter capsulatus* and molecular modeling studies of a Rieske center", *Biochemistry*, **35**, 7834–7845.
- Haddock J. D. and Gibson D.T., (1995). "Purification and characterization of the oxygenase component of biphenyl 2,3-dioxygenase from *Pseudomonas* sp. strain LB400", *Journal of Bacteriology*, **177**, 5834–5839.
- Haddock J. D., Horton J. R. and Gibson D. T., (1995). "Dihydroxylation and dechlorination of chlorinated biphenyls by purified biphenyl 2,3-dioxygenase from *Pseudomonas* sp. strain LB400", *Journal of Bacteriology*, **177**, 20–26.
- Haddock J. D., Pelletier D. A. and Gibson D.T., (1997). Purification and properties of ferredoxin(BPH), a component of biphenyl 2,3-dioxygenase of *Pseudomonas* sp strain LB400", *Journal of Industrial Microbiology & Biotechnology*, **19**, 355–359.
- Hammond A. L., (1972). "Chemical pollution: Polychlorinated biphenyls." *Science*, **175**, 155-156.
- Han J., Kim S.Y., Jung J., Yoongho L., Ahn J. , Kim Su II and Hur H., (2005). "Epoxide Formation on the Aromatic B Ring of Flavanone by Biphenyl Dioxygenase of *Pseudomonas pseudoalcaligenes* KF707", *Applied and Environmental Microbiology*, **71**, 5354–5361



Hansen L. G., (1987). "Environmental toxicology of polychlorinated biphenyls", *Environ. Toxin Ser.*, **1**, 15–48.

Hartmann, M., Schneider, T. R., Pfeil, A., Heinrich, G., Lipscomb, W. N., and Braus, G. H., (2003). "Evolution of feedback-inhibited  $\alpha$ /R barrel isoenzymes by gene duplication and a single mutation", *Proc. Natl. Acad. Sci., U.S.A.*, **100**, 862-867.

Haslam E., (1993). "Shikimic Acid: Metabolism and Metabolites", (Chichester: John Wiley and Sons).

Hawkes T. R., Lewis T., Cogglins J. R., Mousedale D. M., Lowe D. J. and Thorneley R. N. F., (1990). "Chorismate synthase, pre-steadystate kinetics of phosphate release from 5-enolpyruvylshikimate 3-phosphate", *Biochem. J.*, **265**, 899-902.

Herman K. M. and Weaver L. M., (1999). "The shikimate pathway", *Annu. Rev. Plant Mol. Biol.*, **50** (3), 473–503.

Hoffmann P. J., Doy C. H. and Catcheside D. E. A., (1972). "Separation of three allosterically inhibitable 3-deoxy-D-arabinoheptulosonate 7-phosphate synthases from extracts of *Neurospora crassa* and the purification of the tyrosine inhibitable isoenzyme", *Biochim. Biophys. Acta*, **268**, 550-561.

Hollman P. C. and Katan M. B., (1999a). "Dietary flavonoids: intake, health effects and bioavailability", *Food Chem. Toxicol.*, **37**, 937–942.

Hollman P. C. and Katan M. B., (1999b). "Health effects and bioavailability of dietary flavonols", *Free Radic. Res.*, **31**(Suppl.), S75–S80.

- Hosny M., Dhar K. and Rosazza J. P., (2001). "Hydroxylations and methylations of quercetin, fisetin, and catechin by *Streptomyces griseus*", *J. Nat. Prod.*, **64**, 462–465.
- Howe D. L., Sundaram A. K., Wu J., Gatti, D. L. and Woodard, R. W., (2003). "Mechanistic insight into 3-deoxy-D-manno-octulosonate- 8-phosphate synthase and 3-deoxy-D-arabino-heptulosonate- 7-phosphate synthase utilizing phosphorylated monosaccharide analogues", *Biochemistry*, **42**, 4843-4854.
- Hrywna Y., Tsoi T. V., Maltseva O. V., Quensen J. F. III and Tiedje J. M., (1999). "Construction and characterization of two recombinant bacteria that grow on ortho- and para-substituted chlorobiphenyls", *Appl Environ Microbiol.*, **65(5)**, 2163–2169.
- Hsu S. T., Ma C. I., Hsu S. K., Wu S. S., Hsu N. H. M, Yeh C. C. and Wu S. B., (1985). "Discovery and epidemiology of PCB poisoning in Taiwan: a four-year followup", *Environmental Health Perspectives*, **59**, 5-10.
- Hurtubise Y., Barriault D. and Sylvestre M., (1996). "Characterization of active recombinant his-tagged oxygenase component of *Comamonas testosteroni* B-356 biphenyl dioxygenase", *Journal of Biological Chemistry*, **271**, 8152–8156.
- Hurtubise Y., Barriault D. and Sylvestre M., (1998). "Involvement of the terminal oxygenase beta subunit in the biphenyl dioxygenase reactivity pattern toward chlorobiphenyls", *Journal of Bacteriology*, **180**, 5828–5835.
- Ibrahim, A. R. and Y. J. Abul-Hajj Y. J., (1990). "Microbiological transformation of (+/-)-flavanone and (+/-)-isoflavanone", *J. Nat. Prod.*, **53**, 644–656.

iiKarlin S., Brendel P. and Bucher P., (1992). "Significant similarity and dissimilarity in homologous proteins", *Molecular Biology of Evolution*, **9**, 152–167

Imbeault N.Y., Powlowski J. B., Colbert C. L., Bolin J. T., Eltis L. D., (2000). "Steady-state kinetic characterization and crystallization of a polychlorinated biphenyl-transforming dioxygenase", *Journal of Biological Chemistry*, **275**, 12430–12437.

Inoue K., Habe H., Yamane H. and Nojiri H., (2006). "Characterization of Novel Carbazole Catabolism Genes from Gram-Positive Carbazole Degrader *Nocardioides aromaticivorans* IC177", *Applied and Environmental Microbiology*, 3321–3329.

Iwata S., Saynovits M., Link T. A. and Michel H., (1996). "Structure of a water soluble fragment of the "Rieske" iron sulfur protein of the bovine heart mitochondrial cytochrome bc(1) complex determined by MAD phasing at 1.5 angstrom resolution", *Structure*, **4**, 567–579.

Jacobson J. L. and Jacobson S.W., (1996). "Intellectual Impairment in Children Exposed to Polychlorinated Biphenyls in Utero", *The New England Journal of Medicine*, **335**, 783-789.

Jensen R. A., Xie G., Calhoun D. H. and Bonner C. A., (2002). "The correct phylogenetic relationship of KdsA (3-deoxy-D-manno-octulosonate 8-phosphate synthase) with one of two independently evolved classes of AroA (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase)", *J. Mol. Evol.*, **54**, 416–423.

Jerina D. M., Daly J. W., Jeffrey A. M. and Gibson D. T., (1971). "cis-1,2-Dihydroxy-1,2-dihydronaphthalene: a bacterial metabolite from naphthalene", *Arch. Biochem. Biophys.*, **142**, 394–396.

- Jerina D. M., Daly J. W., Witkop B., Zaltzman-Nirenberg P. and Udenfriend S., (1968). "The role of arene oxide-oxepin systems in the metabolism of aromatic substrates. 3. Formation of 1,2-naphthalene oxide from naphthalene by liver microsomes", *J. Am. Chem. Soc.*, **90**, 6525–6527.
- Jones D. T., (1999). "Protein secondary structure prediction based on position-specific scoring matrices", *J. Mol. Biol.*, **292**, 195-202.
- Jones J. D., Henstrand J. M., Handa A. K., Herrmann K. M. and Weller S.C., (1995). "Impaired wound induction of DAHP synthase and altered stem development in transgenic potato plants expressing a DAHP synthase antisense construct", *Plant Physiol.*, **108**.
- Jones T. A., Zou J. Y., Cowan S. W. and Kjeldgaard M., (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models", *Acta Crystallographica, Section A* **47**, 110-119.
- Jordan P. A., Scott Bohle D., Cecilia A. Ramilo and Jeremy Evans N. S., (2001). "New Insights into the Metal Center of 3-Deoxy-D-arabino-heptulosonate 7-Phosphate Synthase", *Biochemistry*, **40**, 8387-8396.
- Kagami O., Shindo K., Kyojima A., Kajuyo T., Ikenaga H., Furukawa K., Misawa N. (2008) "Protein Engineering on Biphenyl Dioxygenase for Conferring Activity to Convert 7-Hydroxyflavone and 5,7-Dihydroxyflavone (Chrysin)". *Journal of Bioscience and Bioengg.*, **106**, 121-127.

- Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2003) "Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron." *Science*, **299**, 1039–1042.
- Kauppi B., Lee K., Carredano E., Parales R. E., Gibson D. T., Eklund H. and Ramaswamy S., (1998). "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase", *Structure*, **6**, 571–586.
- Kauppi, B., Lee K., Carredano E., Parales R. E., Gibson D. T., Eklund H. and Ramaswamy S., (1998). "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase", *Structure*, **6**, 571–586.
- Kavi Kishor P. B., (1989). "Aromatic amino acid metabolism during organogenesis in rice callus cultures", *Physiol. Plant*, **75**, 395-398.
- Kim S. and Picardal F.W., (2001). "Microbial growth on dichlorobiphenyls chlorinated on both rings as a sole carbon and energy source", *Appl. Environ. Microbiol.*, **67**, 1953–1955.
- Kim, J. H., K. H. Stansbury K. H., N. J. Walker N. J., M. A. Trush M. A., P. T. Strickland P. T. and Sutter T. R., (1999). "Metabolism of benzo[a]pyrene and benzo[a]pyrene-7,8- diol by human cytochrome P450 1B1", *Carcinogenesis*, **20**, 515.
- Kimura N., Nishi A., Goto M. and Furukawa K., (1997). "Functional analyses of a variety of chimeric dioxygenases constructed from two biphenyl dioxygenases that are similar structurally but different functionally", *J. Bacteriol.*, **179**, 3936–3943.

- King T. L., Yeats P., Hellou J. and Niven S., (2002). "Tracing the source of 3,3'-dichlorobiphenyl found in samples collected in and around Halifax Harbour", *Mar. Pollut. Bull.*, **44** (7), 590–596.
- Kneller D. G., Cohen F. E. and Langridge R., (1990). "Improvements in Protein Secondary Structure Prediction by an Enhanced Neural Network", *J. Mol. Biol.*, **214**, 171-182.
- König 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.
- Koppe J. G. and Keys J., (2002). "PCBs and the precautionary principle", *EEA Environmental Issue, European Environmental Agency, Copenhagen*, **22**, 64-75.
- Koshiba T., (1978). "Purification of two forms of the associated 3-dehydroquinate hydrolyase and shikimate:NADP<sup>+</sup> oxidoreductase in *Phaseolus mungo* seedlings", *Biochim. Biophys. Acta.*, **522**, 10-18.
- Kuila D. and Fee J. A., (1986). "Evidence for a redox-linked ionizable group associated with the [2Fe-2S] cluster of *Thermus* Rieske protein", *J. Biol. Chem.*, **261**, 2768–2771.
- Kumamaru T., Suenaga H., Mitsuoka M., Watanabe T. and Furukawa K., (1998). "Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase", *Nat. Biotechnol.*, **16**, 663–666.
- Labarga A., Valentin F., Anderson M. and Lopez R., (2007). Web Services at the European Bioinformatics Institute. Nucleic Acids Research Web services.

Laemmli U. K., (1970). "Cleavage of structural proteins during assembly of the head of bacteriophage T4", *Nature*, **227**, 680-685.

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. Cryst.*, **26**, 283-291.

Link T. A., Hagen W. R., Pierik A. J., Assmann C. and Jagow G. V., (1992). "Determination of the redox properties of the Rieske [2Fe-2S] cluster of bovine heart bcl complex by direct electrochemistry of a water-soluble fragment", *Eur. J. Biochem.*, **208**, 685-691.

Link T. A., Hatzfeld O. M., Unalkat P., Shergill J. K., Cammack R. and Mason J. R., (1996). "Comparison of the "Rieske"[2Fe-2S] center in the bcl complex and in bacterial dioxygenases by circular dichroism spectroscopy and cyclic voltammetry", *Biochemistry*, **35**, 7546-7552.

Litten S., Fowler B. I. and Luszniak D., (2002). "Identification of a novel PCB source through analysis of 209 PCB congeners by US EPA modified method 1668", *Chemosphere*, **46 (9-10)**, 1457-1459

Lovell S. C., Davis I. W., Arendall W. B. III, de Bakker P. I., Word J. M., Prisant M. G., Richardson J. S., Richardson D. C., (2003). "Structure validation by C $\alpha$  geometry: phi, psi and C $\beta$  deviation", *Proteins*, **50(3)**, 437-450.

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

- Martins B. M., Svetlitchnaia T. and Dobbek H., (2005). "2-Oxoquinoline 8 monooxygenase oxygenase component: active site modulation by Rieske- [2Fe-2S] center oxidation/reduction", *Structure (Cambridge)*, **13**, 817–824.
- Mason, J. R. and R. Cammack. 1992. The electron-transport proteins of hydroxylating bacterial dioxygenases. *Annu. Rev. Microbiol.*, **46**, 277–305.
- McCandliss R. J., Poling M. D. and Herrmann K. M., (1978). "3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase. Purification and molecular characterization of the phenylalanine-sensitive isoenzyme from *Escherichia coli*", *J. Biol. Chem.*, **253**, 4259-4265.
- McClelland J. L. and Rumelhart D. E., (1988). "Explorations in Parallel Distributed Processing" , *MIT Press, Cambridge MA.*, **3**, 318-362.
- McCullar M. V., Brenner V., Adams R. H., Focht D. D., (1994). "Construction of a novel polychlorinated biphenyl-degrading bacterium: utilization of 3,40-dichlorobiphenyl by *Pseudomonas acidovorans* M3GY", *Appl. Environ. Microbiol.*, **60**, 3883–3889.
- McGuffin L.J., Bryson K. and Jones D.T., (2000). "The PSIPRED protein structure prediction server", *Bioinformatics*, **16**, 404-405.
- McKay D. B., Seeger M., Zielinski M., Hofer B., Timmis K. N., (1997). "Heterologous expression of biphenyl dioxygenase-encoding genes from a gram-positive broad-spectrum polychlorinated biphenyl degrader and characterization of chlorobiphenyl oxidation by the gene products", *J. Bacteriol.*, **179**, 1924–1930.



Misawa N., Nakamura R., Kagiya Y., Ikenaga H., Furukawa K. and Shindo K., (2005). "Synthesis of vicinal diols from various arenes with a heterocyclic, amino or carboxyl group by using recombinant *Escherichia coli* cells expressing evolved biphenyl dioxygenase and dihydrodiol dehydrogenase genes", *Tetrahedron*, **61**, 195–204.

Misawa N., Shindo K., Takahashi H., Suenaga H., Iguchi K., Okazaki H., Harayama S. and Furukawa K., (2002) "Hydroxylation of various molecules including heterocyclic aromatics using recombinant *Escherichia coli* cells expressing modified biphenyl dioxygenase genes", *Tetrahedron*, **58**, 9605–9612.

Mitchell K. H., Rogge C. E., Gierahn T. and Fox B. G., (2003). "Insight into the mechanism of aromatic hydroxylation by toluene 4-monooxygenase by use of specifically deuterated toluene and p-xylene", *Proc. Natl. Acad. Sci. USA.*, **100**, 3784–3789.

Millar G., and Coggins J.R., (1986). "The complete amino acid sequence of bdehydroquinase of *Escherichia coli* K12", *FEBS Lett.*, **200**, 11-17.

Morris A. L., MacArthur M. W., Hutchinson E. G. and Thornton J. M., (1992). "Stereochemical quality of protein structure coordinates", *Proteins*, **12**, 345-364

Mousdale D. M. and Coggins J. R., (1985). "Subcellular localization of the common shikimate pathway enzymes in *Pisum Sativum* L.", *Planta*, **163**, 241-249.

Mousdale D. M., Campbell M. S., and Coggins J. R., (1987). "Purification and characterization of bifunctional dehydroquinase-shikimate:NADP oxidoreductase from pea seedlings". *Phytochemistry*, **26**, 2665-2670.

- Mousedale D. A. and Coggins J. R., (1984). "Purification and properties of 5-enolpyruvylshikimate Sphosphate synthase from seedlings of *Pisum sativum L.*", *Planta*, **160**, 78-83.
- Mousedale D. A. and Coggins J. R., (1985). "Subcellular localization of the common shikimate pathway enzymes in *Pisum sativum L.*", *Planta*, **163**, 241-249.
- Mousedale D. M., Campbell M. S. and Coggins J. R., (1987). "Purification and characterization of bifunctional dehydroquinase-shikimate: NADP oxidoreductase from pea seedlings", *Phytochemistry*, **26**, 2665-2670.
- Murshudov G. N., Vagin A. A. and Dodson E. J., (1997). "Refinement of macromolecular structures by the maximum-likelihood method", *Acta Crystallographica, Section D* **53**, 240-255.
- Nakanishi N. and Yamamoto M., (1984). "Analysis of the structure and transcription of the *aro3* cluster gene in *Schizosaccharomyces pombe*", *MOI. Gen. Genet.*, **195**, 164-169.
- Nijveldt R. J., E. van Nood E., D. E. van Hoorn D. E., P. G. Boelens P. G., K. van Norren K. and P. A. van Leeuwen P. A., (2001). "Flavonoids: a review of probable mechanisms of action and potential applications", *Am. J. Clin. Nutr.*, **74**, 418-425.
- Nimmo G. A. and Coggins J. R., (1981). "Some kinetic properties of the tryptophan-sensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Neurospora crassa*", *Biochem. J.*, **199**, 657- 665.
- Oesch F., Jerina D. M., Daly J. W., Lu A. Y., Kuntzman R. and Conney A. H., (1972). "A reconstituted microsomal enzyme system that converts naphthalene to trans-1,2-dihydroxy-

1,2-dihydronaphthalene via naphthalene-1,2-oxide: presence of epoxide hydrase in cytochrome P-450 and P-448 fractions”, *Arch. Biochem. Biophys.*, **153**, 62–67.

Otwinowski Z. and Minor W., (1997). "Processing of X-ray diffraction data collected in oscillation mode", *Methods in Enzymology*, **276**, 20.

Parales R. E., Parales J. V. and Gibson D. T., (1999). "Aspartate 205 in the catalytic domain of naphthalene dioxygenase is essential for activity”, *J. Bacteriol.*, **181**, 1831–1837.

Paravicini G., Schmidheini T. and Braus G., (1989). "Purification and properties of the 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (phenylalanine-inhibitable) of *Saccharomyces cerevisiae*”, *Eur. J. Biochem.*, **186**, 361-366.

Perrakis A., Sixma T. K., Wilson K. S. and Lamzin V. S., (1997). "wARP: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models", *Acta Crystallographica, Section D* **53**, 448-455.

Pieper D. H., (2005). "Aerobic degradation of polychlorinated biphenyls”, *Appl Microbiol Biotechnol.*, **67**, 170–191.

Pinto J. E. B. P., Dyer W. E., Weller S. C. and Hermann K. M., (1988). "Glyphosate induces 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase in potato (*Solanum tuberosum L.*) cells grown in suspension culture”, *Plant Physiol.*, **87**, 891-893.

Pohl E., Holmes R. K. and Hol W. G. J., (1999). "Crystal structure of a cobalt-activated diphtheria toxin repressor-DNA complex reveals a metal-binding SH3-like domain”, *J. Mol Biol.*, **292**, 653-667.

- Polley L. D., (1978). "Purification and characterization of 3-dehydroquinate hydrolyase and shikimate oxidoreductase", *Biochim. Biophys. Acta*, **526**, 259-266.
- Potrawfke T., Lohnert T. H., Timmis K. N., Wittich R. M., (1998). "Mineralization of low-chlorinated biphenyls by Burkholderia sp. Strain LB400 and by a two-membered consortium upon directed interspecies transfer of chlorocatechol pathway genes", *Appl. Microbiol. Biotechnol.*, **50**, 440–446.
- Que, L. Jr., (2000). "One motif—many different reactions", *Nat. Struct. Biol.*, **7**, 182–184.
- Radaev S., Dastidar P., Patel M., Woodard R. W. and Gatti D. L., (2000). "Structure and mechanism of 3-deoxy-D-manno-octulosonate 8-phosphate synthase", *J. Biol. Chem.*, **275**, 9476–9484.
- Ramakrishnan C. and Ramachandran G. N., (1965). "Stereochemical criteria for polypeptide and protein chain conformations: II. Allowed conformations for a pair of peptide units", *Biophysics Journal*, **5**, 909-933
- Ramjee M. N., Balasubramanian S., Abel C, Cogglns J. R., Davies G. M., Hawkes T. R., Lowe D. J. and Thorneley R. N. F., (1992). "Reaction of (GR)-& F-EPSP with recombinant *Escherichia coli* chorismate synthase generates a stable flavin mononucleotide semiquinone radical", *J. Am. Chem. Soc.*, **114**, 3151-3153.
- Ray J. M. and Bauerle R., (1991). "Purification and properties of tryptophan-sensitive 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*", *J. Bacteriol.*, **173**, 1894- 1901.

- Ray P. H., (1980). "Purification and characterization of 3-deoxy- D-manno-octulosonate 8-phosphate synthetase from *Escherichia coli*", *J. Bacteriol.*, **141**, 635-644.
- Resnick S. M., Lee K. & Gibson D. T., (1996). "Diverse reactions catalyzed by naphthalene dioxygenase from *Pseudomonas* sp. Strain NCIB 9816", *J Ind Microbio.*, **17**, 438-457.
- Rhee G. Y., Sokol, R. C., Bethoney C. M. and Bush B., (1993). "A long-term study of anaerobic dechlorination of PCB congeners by sediment microorganisms: pathways and mass balance", *Environ. Toxicol. Chem.*, **12**, 1829-1834.
- Riedel, A., Fetzner S., Rampp M., Lingens F., Liebl U., Zimmermann J. L. and Nitschke W., (1995). "EPR, electron spin echo envelope modulation, and electron nuclear double resonance studies of the 2Fe2S centers of the 2-halobenzoate 1,2-dioxygenase from *Burkholderia (Pseudomonas) cepacia* 2CBS", *J. Biol. Chem.*, **270**, 30869-30873.
- Ringelberg D. B., Talley J. W., Perkins E. J., Tucker S. G., Luthy R. G. and Bouwer E. J., (2001). "Succession of phenotypic, genotypic and metabolic community characteristics during in vitro bioslurry treatment of polycyclic aromatic hydrocarbon-contaminated sediments", *Appl Environ Microbiol.*, **67**, 1542-1550.
- Ritchie D.W., Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins: Struct. Funct. Genet.* **52** (2003), pp. 98-106.
- Roberts F., Roberts C. W., Johnson J. J., Kyle D. E., Krell T. and Coggins J. R., (1998). "Evidence for the shikimate pathway in apicomplexan parasites", *Nature*, **393**, 801-805.

- Rudel R. A., Seryak L. M. and Brody J. G., (2008). "PCB-containing wood floor finish is a likely source of elevated PCBs in resident's blood, household air and dust: a case study of exposure". *Environmental Health*, **2**. [doi:10.1186/1476-069X-7-2](https://doi.org/10.1186/1476-069X-7-2).
- Sali A. and Blundell T. L., (1993). "Comparative protein modelling by satisfaction of spatial restraints", *J. Mol. Biol.*, **234**, 779–815.
- Sali A. and Overington J. P., (1994). "Derivation of rules for comparative protein modeling from a database of protein structure alignments", *Protein Sci.*, **3** (9), 1582–1596.
- Schmid J. and Amrhein N., (1995). "Molecular organization of the shikimate pathway in higher plants", *Phytochemistry*, **39**, 737–749.
- Schmid J., Schaller A., Leibinger U., Boll W. and Amrhein N., (1992). "The in-vitro synthesized tomato shikimate kinase precursor is enzymatically active and is imported and processed to the mature enzyme by chloroplasts", *Plant J.*, **2**, 375-383.
- Schmidt C. L., Danneel H. J., Schultz G. and Buchanan, (1990). "Shikimate kinase from spinach chloroplasts: Purification, characterization, and regulatory function in aromatic amino acid biosynthesis", *Plant Physiol.*, **93**, 758-766.
- Schofield L. R., Anderson B. F., Patchett M. L., Norris G. E., Jameson G. B. and Parker E. J., (2005). "Substrate ambiguity and structure of *Pyrococcus furiosus* 3-deoxy-D-arabinoheptulosonate 7-phosphate synthase: an ancestral 3-deoxy-D-ald-2-ulosonate phosphate synthase?", *Biochemistry*, **44**, 11950–11962.

Schofield L. R., Patchett M. L. and Parker E. J., (2004). "Expression, purification, and characterization of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*", *Protein Expression Purif.*, **34**, 17-27.

Schoner R. and Herrmann K. M., (1976). "3-Deoxy-D-arabinoheptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*", *J. Biol. Chem.*, **251**, 5440-5447.

Schröter, T., Hatzfeld, O.M., Gemeinhardt, S., Korn, M., Friedrich, T., Ludwig, B., and Link, T.A. 1998. Mutational analysis of residues forming hydrogen bonds in the Rieske [2Fe-2S] cluster of the cytochrome bc1 complex in *Paracoccus denitrificans*. *Eur. J. Biochem.*, **255**, 100–106.

Schultz S. C., Shields G. C. and Steitz T. A., (1991). "Crystal structure of a CAPDNA complex: the DNA is bent by 90 degrees", *Science*, **253**, 1001-1007.

Schuttelkopf A. W. and van Aalten D. M. F., (2004). "PRODRG: a tool for high-throughput crystallography of protein-ligand complexes", *Acta Crystallographica, Section D* **60**, 1355-1363.

Seeger, M., M. Gonzalez M., Camara B., L. Munoz L., E. Ponce E., L. Mejias L., C. Mascayano C., Y. Vasquez Y. and S. Sepulveda-Boza S., (2003). "Biotransformation of natural and synthetic isoflavonoids by two recombinant microbial enzymes", *Appl. Environ. Microbiol.*, **69**, 5045–5505.

- Selander H. G., Jerina D. M. and Daly J. W., (1975). "Metabolism of chlorobenzene. with hepatic microsomes and solubilized cytochrome P-450 systems", *Arch. Biochem. Biophys.*, **168**, 309–321.
- Sheehan D. and O'Sullivan S., (2006). "Homology modeling of milk enzymes using on-line resources: Insights to structure-function and evolutionary relationships", *International Dairy Journal*, **16**, 701–706
- Sheflyan G. Y., Sundaram A. K., Taylor W. P. and Woodard R. W., (2000). "Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae*", *J. Bacteriol.*, **182**, 5005-5008.
- Sheflyan, G. Y., Howe D. L., Wilson T. L. and Woodard R. W., (1998). "Enzymic synthesis of 3-deoxy-D-manno-octulosonate 8-phosphate, 3-deoxy-D-altru-octulosonate 8-phosphate, 3,5-dideoxy-D-gluco(manno)-octulosonate 8-phosphate by 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase", *J. Am. Chem. Soc.*, **120**, 11027-11032
- Shindo, K., Kagiya, Y., Nakamura, R., Hara, A., Ikenaga, H., Furukawa, K., and Misawa, N. (2003). Enzymatic synthesis of novel antioxidant flavonoids by *Escherichia coli* cells expressing modified metabolic genes involved in biphenyl catabolism. *J. Mol. Catal. B Enzym.*, **23**, 9–16
- Shindo, K., Osawa, A., Nakamura, R., Kagiya, Y., Sakuda, S., Shizuri, Y., Furukawa, K., and Misawa, N. (2004). Conversion from arenes having a benzene ring to those having a picolinic acid by simple growing cell reactions using *Escherichia coli* that expressed the six bacterial genes involved in biphenyl catabolism. *J. Am. Chem. Soc.*, **126**, 15042–15043



Shumilin I. A., Bauerle R. and Kretsinger 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., Wu J., Woodard R. W. and Kretsinger R. H., (2004). “Crystal structure of the reaction complex of 3-deoxy-D-arabino-heptulosonate- 7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation”, *J. Mol. Biol.*, **341**, 455–466.

Shumilin I. A., Kretsinger R. H. and Bauerle R. H., (1999). „Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*.”, *Structure*, **7**, 865–875.

Shuttleworth W. A. and Evans J. N. S., (1994). “Site-directed mutagenesis and NMR studies of histidine385 mutants of 5-enolpyruvylshikimate 3-phosphate synthase”, *Biochemistry*, **33**, 7062-7068.

Silakowski B., Kunze B. and Muller R., (2000). “*Stigmatella aurantiaca* Sg a15 carries genes encoding type I and type II 3-deoxy-D-arabino-heptulosonate-7-phosphate synthases: involvement of a type II synthase in aurachin biosynthesis”, *Arch. Microbiol.*, **173**, 403–411.

Silberhorn E. M., Glauert H. P. and Robertson L. W., (1990). “Carcinogenicity of polyhalogenated biphenyls: PCBs and PBBs”, *Crit. Rev. Toxicol.*, **20**, 439–496.

- Simon I., Fiser A. and Tusnady G. E., (2001). "Predicting protein conformation by statistical methods. Biochimica et Biophysica Acta— Protein Structure and Molecular", *Enzymology*, **1549**, 123–136.
- Sippl M. J., (1993). "Recognition of errors in three-dimensional structure of proteins", *Proteins*, **17**, 355-362.
- Sondossi M., Sylvestre M., Ahmad D. and Masse' R., (1991). "Metabolism of hydroxybiphenyl and chloro-hydroxybiphenyl by biphenyl/ chlorobiphenyl degrading *Pseudomonas testosteroni*, strain B-356", *Journal of Industrial Microbiology*, **7**, 77–88.
- Spain J. C. and Gibson D. T., (1988). "Oxidation of substituted phenols by *Pseudomonas putida* F1 and *Pseudomonas sp.* strain JS6", *Appl. Environ. Microbiol.*, **54**, 1399–1404.
- Stallings W. C., Abdel-Meguid S. S., Lim L. W., Shieh H. S., Dayrlnger H. E., Lelmgruber N. K., Stegeman R. A., Anderson K. S., Slbrskl J. A., Padgette S. R. and Klshore G. M., (1991). "Structure and topological symmetry of the glyphosate target 5-enolpyruvylshikimate 3-phosphate synthase: A distinctive protein fold", *Proc. Natl. Acad. Sci. USA*, **88**, 5046-5050.
- Steinrücken H. C. and Amrhein N., (1980). "The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid S phosphate synthase", *Biochem. Biophys. Res. Commun.*, **94**, 1207-1212.
- Stephens C. M. and Bauerle R., (1991). "Analysis of the metal requirement of 3 deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*", *J. Biol. Chem.*, **266**, 20810-20817.

Subramaniam P. S., Xie G., Xia T. and Jensen R. A., (1998). "Substrate ambiguity of 3-deoxy-D-mannoctulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases", *J. Bacteriol.*, **180**, 119–127.

Suenaga H., Mitsuoka M., Ura Y., Watanabe T. and Furukawa K., (2001). "Directed evolution of biphenyl dioxygenase: emergence of enhanced degradation capacity for benzene, toluene, and alkylbenzenes", *J. Bacteriol.*, **183**, 5441–5444.

Suenaga H., Watanabe T., Sato M., Ngadiman A. and Furukawa K., (2002). "Alteration of regiospecificity in biphenyl dioxygenase by active-site engineering", *Journal of Bacteriology*, **184**, 3682–3688.

Suzich J. A., Dean J. F. D. and Herrmann K. M., (1985). "3-Deoxy-d-arabinoheptulosonate-7-phosphate synthase from carrot root (*Daucus carota*) is a hysteretic enzyme", *Plant Physiol.*, **79**, 765–770

Suzich J. A., Dean J. F. D. and Herrmann K.M., (1985). "3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase from carrot root (*Daucus carota*) is a hysteretic enzyme", *Plant Physiol.*, **79**, 765-770.

Sylvestre M., Sirois M., Hurtubise Y., Bergeron J., Ahmad D., Shareck F., Barriault D., Guillemette I. and Juteau J. M., (1996b). "Sequencing of *Comamonas testosteroni* strain B-356-biphenyl/chlorobiphenyl dioxygenase genes: evolutionary relationships among Gram-negative bacterial biphenyl dioxygenases", *Gene*, **174**, 195–202.

- Taira K., Hirose J., Hayashida S., Furukawa K., (1992). "Analysis of bph operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707", *Journal of Biological Chemistry*, **267**, 4844–4853.
- Tamura K., Dudley J., Nei M. and Kumar S., (2007). "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0", *Molecular Biology and Evolution*, **24**, 1596-1599.
- Tao Y., Fishman A., Bentley W. E. and Wood T. K., (2004). "Oxidation of benzene to phenol, catechol, and 1,2,3-trihydroxybenzene by toluene 4-monooxygenase of *Pseudomonas mendocina* KR1 and toluene 3-monooxygenase of *Ralstonia pickettii* PK01", *Appl. Environ. Microbiol.*, **70**, 3814–3820.
- Thornton J. M., (2001). "From genome to function". *Science*, **292**, 2095-2097.
- Toft G., Hagmar L., Giwercman A. and Bonde J. P., (2004). "Epidemiological evidence on reproductive effects of persistent organochlorines in humans", *Reproductive Toxicology*, **19**, 5-26.
- Toure O., Chen Y. Q. and Dutta S. K., (2003). "Sinorhizobium meliloti electrotransformant containing ortho-dechlorination gene shows enhanced PCB dechlorination", *Fresenius Environ Bull*, **12(3)**, 320–322.
- Trumpower B. L. and Gennis R. B., (1994). "Energy transduction by cytochrome complexes in mitochondrial and bacterial respiration: The enzymology of coupling electron transfer reactions to transmembrane proton translocation", *Annu. Rev. Biochem.*, **63**, 675–716.

- Vagin A. and Teplyakov A., (1997). "MOLREP: an automated program for molecular replacement", *Journal of Applied Crystallography*, **30**, 1022-1025.
- Wackett L. P., (2002). "Mechanism and applications of Rieske non-heme iron dioxygenases", *Enzyme Microb. Technol.*, **31**, 577-587.
- Wagner T., Shumilin I. A., Bauerle R. and Kretsinger R. H., (2000). "Structure of 3-deoxy-D-arabino-heptulosonate- 7-phosphate synthase from *Escherichia coli*: Comparison of the Mn<sup>2+</sup>-phosphoglycolate and the Pb<sup>2+</sup>-phosphoenolpyruvate complexes and implications for catalysis", *J. Mol. Biol.*, **301**, 389–399.
- Webby C. J., Baker H. M., Shaun Lott J., Baker E. N., Emily J. and Parker, (2005). "The Structure of 3-Deoxy-D-arabino-heptulosonate 7-phosphate Synthase from *Mycobacterium tuberculosis* Reveals a Common Catalytic Scaffold and Ancestry for Type I and Type II Enzymes", *J. Mol. Biol.*, **354**, 927–939.
- Webby C. J., Patchett M. L. and Parker E. J., (2005). "Characterization of a recombinant type II 3-deoxy-D-arabino- heptulosonate-7-phosphate synthase from *Helicobacter pylori*", *Biochem. J.*, **390**, 223–230.
- Welch G. R., Cole K. W. and Gaertner F. H., (1974). "Chorismate synthase of *Neurospora crassa*: A flavoprotein", *Arch. Biochem. Biophys.*, **165**, 505-518.
- Werlen, C., Kohler H. P. and van der Meer J. R.. (1996). "The broad substrate chlorobenzene dioxygenase and cis-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain

P51 are linked evolutionarily to the enzymes for benzene and toluene degradation”, *J. Biol. Chem.*, **271**, 4009-4016.

Whited G. M., McCombie W. R., Kwart L. D. and Gibson D. T., (1986). “Identification of cis-diols as intermediates in the oxidation of aromatic acids by a strain of *Pseudomonas putida* that contains a TOL plasmid”, *J. Bacteriol.*, **166**, 1028 – 1039.

Whiteley A. S, Manefield M. and Lueders T., (2006). “Unlocking the ‘microbial black box’ using RNA-based stable isotope probing technologies”, *Curr Opin Biotechnol.*, **17**, 67–71.

Wiegel J. and Wu Q., (2000). “Microbial reductive dehalogenation of polychlorinated biphenyls.”, *FEMS Microbiology Ecology*, **32**, 1-15.

Wikipedia contributors, "Apicomplexa," *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/w/index.php?title=Apicomplexa&oldid=248901164> (accessed November 6, 2008).

William W.A., (1994). “Microbial reductive dechlorination of trichlorobiphenyls in anaerobic sediment slurries”, *Environ. Sci. Technol.*, **28**, 630–635.

Wdlanski T., Bender S. L. and Knowles J. R., (1989). “Dehydroquinase synthase: A sheep in wolfs clothing?”, *J. Am. Chem. Soc.*, **111(1)**, 2299-2300.

Wu J., Howe D. L. and Woodard R. W., (2003). “*Thermotoga maritima* 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase: The ancestral eubacterial DAHP synthase?”, *J. Biol. Chem.*, **278**, 27525-27531.

Zanaroli G., Perez-Jimenez J. R., Young L. Y., Marchetti L. and Fava, F., (2006). "Microbial reductive dechlorination of weathered and exogenous co-planar polychlorinated biphenyls (PCBs) in an anaerobic sediment of Venice Lagoon", *Biodegradation*, **17** (2), 121–129.

Zielinski M., Backhaus S. and Hofer B., (2002). "The principal determinants for the structure of the substrate-binding pocket are located within a central core of biphenyl dioxygenase alpha subunit", *Microbiology*, **148**, 2439–2448.

Zielinski M., Kahl S., Hecht H. J. and Hofer B., (2003). "Pinpointing biphenyl dioxygenase residues that are crucial for substrate interaction", *J. Bacteriol.*, **185**, 6976–6980.

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