

3D STRCUTURAL MODELLING FOR PROCATECHUATE 3, 4 DIOXYGENASE FROM PSEUDOMONAS CEPACIA BASED ON TEMPLATE OF M CHAIN FROM P. AERUGINOSA

D. JAYASREE¹, J.VENKATESHWARA RAO¹, D. VENKATA KRISHNA RAO¹, K. K. KUMAR¹, P. B. K. KISHORE³, H. A. NAGARAJARAM², Y.PRAMEELA DEVI¹, M. LAKSHMI NARASU¹,

The 3D model of Protocatethuate 3,4 dioxygenase beta chain from *Pseudomonas cepacia* was built, based on X-ray crystallographic structure of Protocatethuate 3,4 dioxygenase M chain from *Pseudomonas aeruoginosa*. Dioxygenase catalyses the cleavage of molecular oxygen with subsequent incorporation of both oxygen atoms into organic substrates. Some of the dioxygenase from bacteria catalyse critical ring-opening step in biodegradation of aromatic compounds. These bacterial enzymes contain non- heme ferric iron as the sole cofactor. Protocatethuate 3,4 dioxygenase (3,4PCD) was one of such enzyme recognized which catalyses the intradiol cleavage of protocatechuic acid by oxygen to produce â- carboxy –cis-cis-muconic acid. Studies on 3,4(PCD) found in *Pseudomonas aeruginosa* is an oligomer with relative molecular mass (587K). The homology modelling is done based on the programmes to annotate protein sequence aligment with 3D structural features which helps to understand the conservation of amino acids in the specific

¹ Centre for Biotechnology, Institute of Science & Technology, JNTU, Kukatpally-500 085, Hyderabad

² Centre for DNA and Finger Printing Diagnostics, Nacharam - Hyderabad

³ Department of Genetics, Osmania University, Hyderabad-500 007

local environment.

ABSTRACT

KEY WORDS				
Multiple sequence				
alignment				
Ramachandram Plot				
Joy alignment				
Steriochemical				
equality				
Received on : 21.10.06 Accepted on : 27.01.07				

 Corresponding author

INTRODUCTION

Protocatechuate 3,4-dioxygenase is a member of a family of bacterial enzymes that cleave the aromatic rings of their substrates between two adjacent hydroxyl groups, a key reaction in microbial metabolism of varied environmental chemicals.

Protocatechuate (3, 4-dihydroxybenzene, PCA) is an aromatic compound which is a key intermediate in the degradation of the plant biopolymer lignin and other aromatic compounds. The key step of PCA degradation is the ring-cleavage performed by dioxygenases adding both atoms from molecular oxygen to specific carbon atoms within the ring. This step can be performed by two distinct mechanisms; intradiol cleavage and extradiol cleavage. In intradiol cleavage the oxygen atoms are added to the carbons carrying the hydroxyl groups, producing two carboxylate groups. In extradiol cleavage the oxygens are added to one carbon carrying a hydroxyl group and another carrying a hydrogen, resulting in the formation of a carboxylate group and an aldehydic group.

The extradiol dioxygenases use Fe(II) to activate oxygen for nucleophilic attack on the aromatic substrate, while the intradiol dioxygenases use Fe(III) to activate the aromatic substrate for an electrophilic attack by oxygen. **Mechanism of action**

The aromatic-ring-cleavage dioxygenases open the aromatic ring by incorporating two atoms of dioxygen (O2) in their substrates, typically carrying two or more hydroxyl groups on the aromatic ring. If two of the hydroxyl groups of a substrate are in the ortho position, the ring fission by the intradiol aromatic-ring-cleavage dioxygenases (IARCD) occurs between the two hydroxyl groups (cf. extradiol aromatic-ring-cleavage dioxygenases):



ARCD enzymes, catechol 1,2-dioxygenase (1,2-CTD; EC 1.13.11.1) and protocatechuate 3,4-dioxygenase (3,4-PCD; EC 1.13.11.3), contain a single FellI as a prosthetic group. 1,2-CTD enzymes are oligomers composed of either heterodimers (alphaß)n or homodimers (alphaalpha)n. 3,4-PCD contain equal numbers of alpha and ß subunits and form different quaternary structures of (alphaß)n (n = 3 to 12) (Harayama et al., 1992). The sequence similarity between 1,2-CTD and the alpha and ß subunits of 3,4-PCD suggests common ancestry of IARCD (Que and Ho, 1996). The best-characterised IARCD, 3,4-PCD, catalyses the cleavage of an aromatic ring of 3,4-dixydroxybenzoate to form a dicarboxylic acid (Que and Ho, 1996). Each of the carboxylate groups contains one of the oxygen atoms from O2 (Harayama and Rekik, 1989).



The substrate activation mechanism for IARCD is represented in Fig. 1



Figure 1: The native enzyme contains a highspin, pentacoordinate Fe^{III} centre (I). Upon substrate binding, the solvent -derived ligand and Tyr_g147 are displaced by a bidentate catecholate dianion (cf. bidentate catecholate monoanion in extradiol aromatic-ring-cleavage reaction) and the Fe^{III} centre (II) remains pentacoordinate. The attack of O_2 on the semiquinone radical (III) yields a transient alkylperoxide radical (IV) which combines with Fe^{II} centre to generate a tridentate alkylperoxo- Fe^{III} complex (V). Decomposition of compound (V) by a Crigeetype rearrangement yields muconic anhydride and a native- like Fe^{III} centre (VI). Muconic anhydride is subsequently hydrolysed by an Fe^{III} -bound hydroxide derived from O_2 (Que and Ho, 1996).

MATERIAL AND METHODS Methods and software tools used for modelling: 1. Sequence retrieval : Pedant (http://pedant.gsf.de/)

The pedant genome database provides exhaustive automatic analysis of genomic sequences by a large variety of bioinformatics tools. For example the following precomputed analyses are available to analyse protein function: Blast similarity searches against the nonredundant protein sequence database, motif searches against the Pfam, BLOCKS and PROSITE databases

2. 3DPSSM : (http://www.sbg.bio.ic.ac.uk/ ~ 3dpssm/) Fast, web-based method for protein fold recognition using 1D and 3D sequence profiles coupled with secondary structure and solvation potential information was done.

3. Multiple sequence alignments & phylogenetic tree Clustal (www.molecularevolution.org/software/clustalx) The Clustal programme was employed as the Clustal programmes are the choice for the novice to make a decent phylogenetic analysis of obtained nucleotide or amino acid sequences. Clustal can align the sequences and produce output files for drawing trees

4. Modelling : SPDBVIEWER & SYBYL: (http://expasy.org spdbv/)

DeepView - Swiss-PdbViewer is an application that provides a user friendly interface allowing to analyze several

proteins at the same time. The proteins can be superimposed in order to deduce structural alignments and compare their active sites or any other relevant parts.

The purposes of any molecular modeling programme are to build, study and manipulate molecules. SYBYL provides powerful tools to accomplish these goals. The techniques, both command and menu driven, are simple enough for use on small molecules, yet powerful enough to manipulate large molecules.

5. Structural valuation : Procheck & Joy alignment

Procheck is a suite of programmes to check the stereochemical quality of protein structures.OY is a program to annotate protein sequence alignments with three-dimensional (3D) structural features. It was developed to display 3D structural information in a sequence alignment and help to understand the conservation of amino acids in their specific local environments.

RESULTS AND DISCUSSION

1. Target Sequence : A protein sequence with an accession number P15110 is selected as target sequence for modelling from *P.cepacia* with a protein length (a.a) 235 and molecular weight 26550 da.

The target sequence is given in Table 1.

Table 1: target sequence for modelling from P. cepacia

Template Sequence: The template selected and used for

protein modelling of P15110 3Pcc, is a X-ray crystallized structure of Protocatechuate 3,4 dioxygenase from Pseudomonas putida. 3pcc is having 238 residues with a molecular weight of 2662 da and 12 alpha, 12 beta chains with a resolution of 1.98 A⁰ 2. 3D PSSM results for fold recognition

Table 2: 3I	D PSSM resu	ults for flo	d recogni	tion	
Fold Library	Templat Length	PSSM E_value	Class	Fold	- 0
<mark>d3.ccm</mark> 47%i.d.	233	1.66 e -08	All beta proteins	Prealbumin-like	3pcc
<mark>cteob</mark> 46%i.d.	238	0.0027	All beta proteins	Prealbumin-like	1eoc
<mark>c1eoaa</mark> 31%i.d.	202	0.0027	All beta proteins	Prealbumin-like	1eoa
d3pcca 30%i.d.	200	0.0027	All beta proteins	Prealbumin-like	3pcc
d1dmha 23%i.d.	309	0.0027	All beta proteins	Prealbumin-like	1dmh

The target sequence was submitted to fold recognition server located at http://bmn.icnet.ac.uk/3dpssm . From the above results (Table 2) a template 3pcc with 47% identity of protocatechuate 3, 4 dioxygenase mchain from *Pseudomonas putida* with a resolution of 1.98 A⁰, R value of 0.163 and with out any bond breakages in the X-ray crystallography structure satisfying all the properties of a template was found.

3. PDB Blast

By using PDB Blast(http:// bioinformatics.ljcrf.edu/pdb blast) very remote homologues structures (Table 3) is detected which helps in constructiong sequence profiles. So the target sequence (P15110) is PDB blasted and results were found that the target sequence is close structural similarity with 3PCC M chain of 235 a.a length.

Table 3: Remote homologues structures using PDB Blast

- 1. gi|10121001|pdb|1EOA 266 1e-72 34% 176 ChainA,
- 2. gi|10120999|pdb|1EO9 266 1e-72 34% 176 ChainA,
- 3. gi|3212804|pdb|3PCC| 207 8e-55 47% 235 Chain M,
- 4. gi|3212780|pdb|3PCD| 205 4e-54 46% 235 Chain M,
- 5. gi|10121000|pdb|1EO9 200 9e-53 45% 226 Chain B,

4. Multiple sequence alignment

Based on above PDB blast results multiple sequence alignment was done by using CLUSTAL X programme and identified the conserved regions.(Table 4)(represented by using : marks). A phylogenetic tree was constructed to trace out the relationships between the sequences. We found that 3PCG and 3 PCD are the root structures which

Table 4: Multiple sequence alignment through Clustal X programme

3PCG B	PIE	LLPETPSOTAGPYVHIGLALEAAGNPTRDOEIWNRLA	40
3PCG_C	PTE	LLPETPSOTAGPYVHIGLALEAAGNPTRDOETWNRLA	40
3PCG D	 PTF	LLPFTPSOTAGPYVHTGLALFAAGNPTRDOFTWNRLA	40
SPCG F	DIE		40
SPCG_E	PIE	LI DET DSOTA C DVINTCI AL ENA CHETE DOE	40
SPCG_P		LIPETPOUROFIVITOIALEAAGNFIRDQEIWARAA	40
JFCG_A		ELFEIFSQIAGFIVHIGLERVONTEUEEUN IDDDUN	40
ILO9_A		ELKEIPSQIGGPIVHIGLERQANIEVPEHNLUNNLV	44
3PCG_P	PAQDNSRFVIRDRNWHPKALTPD	YKTSIARSPRQALVSIPQSISEITGPNFSHLGFGAHDHDLLLNFNN	69
3PCG_R	PAQDNSRFVIRDRNWHPKALTPD	YKTSIARSPRQALVSIPQSISETTGPNFSHLGFGAHDHDLLLNFNN	69
3PCG_Q	PAQDNSRFVIRDRNWHPKALTPD	YKTSIARSPRQALVSIPQSISETTGPNFSHLGFGAHDHDLLLNFNN	69
3PCG_N	PAQDNSRFVIRDRNWHPKALTPD	YKTSIARSPRQALVSIPQSISETTGPNFSHLGFGAHDHDLLLNFNN	69
3PCG_M	PAQDNSRFVIRDRNWHPKALTPD	YKTSIARSPRQALVSIPQSISETTGPNFSHLGFGAHDHDLLLNFNN	69
3PCG_0	PAQDNSRFVIRDRNWHPKALTPD	YKTSIARSPRQALVSIPQSISETTGPNFSHLGFGAHDHDLLLNFNN	69
1EOA_B	MSQIIWGAYAQRNTEDHPPAYAPG	YKTSVLRSPKNALISIAETLSEVTAPHFSADKFGPKDNDLILNYAK	70
1DMH_A	MEVKIFNTQDVQDFLRVASGLEQEGGNPRVKQIIHRVLSDLYKAIEDLNITSDEYWAGVA	YLNQLGANQEAGLLSPGLGFDHYLDMRMDAEDAALGIENATPRTIEGPLYVAGAPESV	GY 120
1DMH B	MEVKIFNTQDVQDFLRVASGLEQEGGNPRVKQIIHRVLSDLYKAIEDLNITSDEYWAGVA	YLNQLGANQEAGLLSPGLGFDHYLDMRMDAEDAALGIENATPRTIEGPLYVAGAPESV	GY 120
-			
3PCG_B	KPDAPGEHILLLGQVYDGNGHLVRDSFLEVWQADANGEYQDAYNLENAFN	SFGRTATTFDAGEWTLHTVKPGVVNNAAGVPMAPHINISLFARG	NI 137
3PCG_C	KPDAPGEHILLLGQVYDGNGHLVRDSFLEVWQADANGEYQDAYNLENAFN	SFGRTATTFDAGEWTLHTVKPGVVNNAAGVPMAPHINISLFARG	NI 137
3PCG D	KPDAPGEHILLLGQVYDGNGHLVRDSFLEVWQADANGEYQDAYNLENAFN	SFGRTATTFDAGEWTLHTVKPGVVNNAAGVPMAPHINISLFARG	NI 137
3PCG E	KPDAPGEHILLLGQVYDGNGHLVRDSFLEVWQADANGEYQDAYNLENAFN	SFGRTATTFDAGEWTLHTVKPGVVNNAAGVPMAPHINISLFARG	NI 137
3PCG F	KPDAPGEHILLLGQVYDGNGHLVRDSFLEVWQADANGEYQDAYNLENAFN	SFGRTATTFDAGEWTLHTVKPGVVNNAAGVPMAPHINISLFARG	NI 137
3PCG A	KPDAPGEHILLLGOVYDGNGHLVRDSFLEVWOADANGEYODAYNLENAFN	SEGREATEDAGEWELHTVKPGVVNNAAGVPMAPHINISLEARGI	NT 137
1E09 A	ODNTOGORIRLEGOVFDGLSLPLRDVLIEIWOADTNGVYPSOADTOGKOVDPNFL	GWGRTGADEGTGEWSENTIKEGAVEGRKGSTOAPHISLITEARG	NT 146
3PCG P	GGLPIGERITVAGRVVDOYGKPVPNTLVEMWOANAGGRYRHKNDRYLAPLDPNEG	CVCDCLTD_SDCVVSFDTTVDCDVDWDMCDNDWDDAHTHFGTSCDS1	AT 172
SPCG R	GGLPIGERITVAGRVVDOYGKPVPNTLVEMWOANAGGRYRHKNDRYLAPLDPNEG	CVCDCLTD_SDCVVSFDTTVDCDVDWDNCDNDWDDAHTHFCTSCDS1	AT 172
3800 0	CCI DIGEDITVACOVOQUEVENTIVEMWQANACCOVERNDOVI A DI DENEG	CUCDCLTD SDCVVCEDTINDC DVDWDWDWDWDWDWVTWECISCDS	AT 172
SPCG_N	CGI DIGED I TUX CDUUDOVCHUUDNTI VENWOANACCDVDHUNDVI A DI DDNEC	CUCCLID-SDGIISFRIIRFGFIFWRWGFNDWRFARINFGISGFS	AL 172
SPCG_N	CCI DICERTIVES COMPONENTI MEMONIA CONDUCT NUMBER AND A DI DRIFT	GVGRCLID-SDG115FRIIRFGFIFWRNGFNDWRFARIHFGISGFS	AI 172
SPCG_M	GGLFIGERIIVAGRVVDQIGRFVFNILVENWQANAGGRIRHRNDRILAFLDFNFG	GVGRCLID-SDGIYSFRIIKFGPIPWRNGPNDWRPAHIFGISGFS	AI 172
SPCG_0	GGLFIGERIIVAGRVVDQIGRFVFNILVENVQANAGGRIRRNDRILAFLDFNFG	GVGRCLID-SDGYYSFRIIKPGPYPWRNGPNDWRPAHIHFGISGPS	AI 172
IEOA_B	DGLPIGERVIVHGIVKDQFGRPVKNALVEVWQANASGRIRHPNDQIIGAMDPNFG	GCGRMLTD-DNGYYVFRTIKPGPYPWRNRINEWRPAHIHFSLIADGW	AQ 173
1DMH_A	ARMDDGSDPNGHTLILHGTIFDADGKPLPNAKVEIWHANTKGFYSHFDPTGEQQAF	NMRRSIITDENGQYRVRTILPAGYGCPPEGPTQQLLNQLGRHGNRPAHIHYFVSADG-	-HR 235
1DMH_B	ARMDDGSDPNGHTLILHGTIFDADGKPLPNAKVEIWHANTKGFYSHFDPTGEQQAF	NMRRSIITDENGQYRVRTILPAGYGCPPEGPTQQLLNQLGRHGNRPAHIHYFVSADG-	-HR 235
	. *. : : * : * . : : : * : * * : : * *	. * * :*: **** :	
ance p		LOCECETUREDE 200	
SPCG_D	IL HTDI VEDERQAMARCEVEN IL TEODODDETI INVOCENDO		
SPCG_C	ILTELUTED VERSEN ANALYSEN		
3PCG_D	HLHIRLYFDDEAQANAKCFVLNLIEQFQRREILIARRCEVDGKIAYRFDIR-	-IQGEGEIVEEDE 200	
3PCG_E	HLHTRLYFDDEAQANAKCPVLNLIEQPQRRETLIAKRCEVDGKTAYRFDIR-	-IQGEGETVFFDF 200	
3PCG_F	HLHTRLYFDDEAQANAKCPVLNLIEQPQRRETLIAKRCEVDGKTAYRFDIR-	-IQGEGETVFFDF 200	
3PCG_A	HLHTRLYFDDEAQANAKCPVLNLIEQPQRRETLIAKRCEVDGKTAYRFDIR-	-IQGEGETVFFDF 200	
1E09_A	GLHTRVYFDDEAEANAKDPVLNSIEWATRRQTLVAKREERDGEVVYRFDIR-	-IQGENETVFFDI 209	
3PCG_P	KLITQLYFEGDPLIP-MCPIVKSIANPEAVQQLIAKLDMNNANPMDCLAYRFDIV-	-LRGQRKTHFENC 238	
3PCG_R	KLITQLYFEGDPLIP-MCPIVKSIANPEAVQQLIAKLDMNNANPMDCLAYRFDIV-	-LRGQRKTHFENC 238	
3PCG_Q	KLITQLYFEGDPLIP-MCPIVKSIANPEAVQQLIAKLDMNNANPMDCLAYRFDIV-	-LRGQRKTHFENC 238	
3PCG_N	KLITQLYFEGDPLIP-MCPIVKSIANPEAVQQLIAKLDMNNANPMDCLAYRFDIV-	-LRGQRKTHFENC 238	
3PCG M	KLITQLYFEGDPLIP-MCPIVKSIANPEAVQQLIAKLDMNNANPMDCLAYRFDIV-	-LRGQRKTHFENC 238	
3PCG 0	KLITQLYFEGDPLIP-MCPIVKSIANPEAVQQLIAKLDMNNANPMDCLAYRFDIV-	-LRGORKTHFENC 238	
1EOA B	RLISOFYFEGDTLID-SCPILKTIPSEOORRALIALEDKSNFIEADSRCYRFDIT-	-LRGRRATYFENDLT 241	
1DMH A	KLTTOINVAGDPYTYDDFAYATREGLVVDAVEHTDPEAIKANDVEG-PFAFMVFDLKLTR	LVDGVDNOVVDRPRLAV 311	
1DMH B	KLTTOINVAGDPYTYDDFAYATREGLVVDAVEHTDPFATKANDVEG-PFAFMVFDLKLTR	LVDGVDNOVVDRPRLAV 311	
	*	*	

Cladogram



Figure 2 : Cladogram showing resemblance of structure of 3 PCG and 3 PCD with that of 3PCC

5. Alignment of Target with Template molecule 3DPSSM alignment is given below





jaya10a_01.ps

Figure 5: Ramachandran plot; darkest core regions representing more favourable phi & psi values

-----PROCHECX S U K M A R Y >>>----javaloa. 2.0 235 realdues Ramachandran plot: 88.4% core 11.6% allow .0% gener ÷ .Ok disa22 Gly & Pro Ramach: 2 labelled residues (out of 44) Chil-chi2 plots: ÷ 4 labelled residues (out of 138) Main-chain parama: 5 better 出る 0 inside 1 worse Side-chain parame: 5 hetter 0 inside 0 worse - 新 Residue properties: Max.deviation: 18.6 Bad contacts: 10 Bond len/angle: 7.9 Morris et al class: 1 1 2 ÷. 1 cis-peptides 1 D amino acida G-lactors Dihedrals: -.14 Covalent: -.15 Dveta[1: -.24 M/c bond lengths: 86.1% within limits 13.9% highlighted M/c bond angles: 33.14 within limits 16.9% highlighted Placer groups: 100.0% within limits .04 highlighted + May be worth investigating further. * Worth investigating further.

Figure 6: Verification of sterio-chemical equality by PROCHECK

D. JAYASREE et al.

have very close resemblance to the template structure (3PCC) (Fig. 2).

Based on 3DPSSM, results were found that template is 47% homologous to the target and the secondary structures like helices, coils and strands are clearly identified in target. So based on the alignment obtained further comparative modelling was done to the target molecule. **Model Builder**

The refined sequence – structure alignment as obtained by 3D-PSSM server was used to construct 3D models of Protocatachuate 3, 4 dioxygenase using MODELLER. The structure was refinied and some loop regions were corrected using modeller. The target molecule has been shown before and after the modelling in Figs. 3 & 4.



Figure 3 :Target molecule before modelling:



Figure 4 :Target molecule after modelling

Loop modeling

Loops often play an important role in defining functional specificity of a protein frame work. Loop modeling is major factor for determining usefulness of comparative models in application such as Ligand docking. Loop modelling provides the information about core and anchor regions.

Target: The anchor regions of the loop which we remodeled, is given in Table 5

Table 5 : The anchor regions of loop and resiedue number

Anchor regions	Resiedue no
30-45	15

The regions which are to be modelled are 30 to 45 residues were carefully selected in loop database SPDBV was scanned. After modelling loop is subjected to energy



Figure 7: Structure to structure alignment of the sequence model throught Joy alignment

minimization. This stereo chemical quality of the residue inthe remodelled loop as well as those in vicinity was checked by plotting Ramchandran plot (Fig. 5) and after remodelling was again submitted to verify 3D program.

Ram chandran plot

It shows phi-psi torsion angles for all residues. The darkest region corresponds to core regions representing more favourable combinations of phi and psi values. Percentage of core regions is one of the better guides to stereo chemical quality.

Compatibility score for the model is above zero. So model satisfies the environment.Next, the model was submitted to PROCHECK to verify stereo-chemical quality (Fig. 6). **Joy alignment**

Finally the sequence modelled is submitted to joy server, where joy shows structure to structure alignment (Fig. 7)., which is given by verifying secondary structure.

REFERENCES

Harayama, S., Kok, M. and Neidle, E.L. (1992) Functional and evolutionary relationships among diverse oxygenases. *Annu. Rev. Microbiol.* **46**: 565-601.

Harayama, S. and Rekik, M. (1989) Bacterial aromatic ring-cleavage enzymes are classified into two different gene families. *J. Biol. Chem.* 264: 15328-15333.

Lippard, S.J. and Berg, J.M. (1994) Principles of Bioinorganic Chemistry. University Science Books, Mill Valley.

Nishida, Y., Yoshizawa, K., Takahashi, S. and Watanabe, I. (1992) Reaction mechanism of protocatechuate 3,4-dioxygenase. Z. *Naturforsch.* C 47: 209-214.

Ohlendorf, D.H., Orville, A.M. and Lipscomb, J.D. (1994) Structure of protocatechuate 3,4-dioxygenase from Pseudomonas aeruginosa at 2.15 Å resolution. *J. Mol. Biol.* **244**: 586-608.

Que, L., Jr. and Ho, R.Y.N. (1996) Dioxygen activation by enzymes with mononuclear non-heme iron active sites. *Chem. Rev.* 96:2607-2624.

SPECIAL ISSUE OF THE BIOSCAN

The Editorial Board of The Bioscan is going to bring about special issue of the journal on

1. Physiology and Endocrinology

2. Ecological Productivity and Energetics

Interested Academicians, Researchers and Scientists are requested to contribute to the proposed issue. The pattern of preparation of manuscript will be the same as the instruction to the authors of this journal.