In silico Approach in the Prediction and Analysis of the Three-dimensional Structure of *Maleylacetate reductase*: A Biodegrading Protein

Shasank Sekhar Swain

Research & Development Center
Hi-tech Medical College and Hospital
Bhubaneswar, Odisha, India-751010
E-mail: swain.shasanksekhar86@gmail.com

Received: July 5, 2013 **Accepted: December 06, 2013**

Published: December 20, 2013

Abstract: *With the advent of biological research in the field of environmental science, several microbes were found to act as the most important biodegradable molecules. Maleylacetate reductase being a member of oxidoreductase is mostly found in Pseudomonas species. This enzyme participates in three metabolic pathways: gamma-hexachlorocyclohexane degradation, benzoate degradation which are higher alkyl compounds involved in chemical pollutions in metropolitan cities. Determining its three-dimensional structure will lead to the structure function analysis and also might be helpful for designing receptors for the degradation of new chemical compounds. For this purpose, the amino acid sequence of the protein had been imported from UniProtKB database and the template was searched in BLASTp2.2.27. This template was then used for the prediction of the three dimensional structure of the protein by using SWISS-MODEL Workspace. The predicted model was then validated using SAVES server which gave an almost result for better prediction.*

Keywords: *Maleylacetate reductase, BLASTp, SWISS-MODEL Workspace, UniProtKB, SAVES server.*

Introduction

Maleylacetate reductases (MAR) play a crucial role in the aerobic microbial degradation of aromatic compounds [33]. They catalyze the NADH- or NADPH-dependent reduction of maleylacetate or 2-chloromaleylacetate to 3-oxoadipate or of substituted maleylacetates to substituted 3-oxoadipates. In fungi, maleylacetate reductases contribute to the catabolism of very common substrates, such as tyrosine, gentisate, benzoate, 4-hydroxybenzoate, protocatechuate, vanillate, resorcinol and phenol [3, 8, 13, 16, 20, 21, 24, 28]. For bacteria, it has been shown that *Maleylacetate reductases* are involved in the degradation of resorcinol and 2,4-dihydroxybenzoate via hydroxyquinol [9, 10, 22, 30]. Other substrates whose catabolic routes be a substrate of this activity have a more complex structure, such as 2-hydroxydibenzo-p-dioxin and 3-hydroxydibenzofuran [4] or carry unusual substituents, such as nitro or sulfo groups [4, 15, 18, 26, 29] fluorine or chlorine atoms [14, 27, 32]. Our main objective is to develop a computational three-dimensional structure with reference to the sequence, which might play a vital role in increasing the efficiency of this biodegrading enzyme. This structure can be modified and validated further through homology modeling. The Ramachandran plot projects the residues must have favorable conformation for *psi* and *phi* angles that will help further in understanding the stability and rotation of amino acid in protein structure. The key objective is to explore a chemical-free environment through degradation of harmful deterrent chemical substances by the use of microorganism. Approaches can be made in identification of various active sites for the attachment

of receptors through servers and tools which may lead in identification of most catalytic site for the protein. The information may be of helps for the crystal structure prediction of this protein.

Materials and methods

Target selection

The amino acid sequence of *Maleylacetate reductase* protein of *Pseudomonas aeruginosa* was retrieved from the UniProtKB (Acc. No. O87612) database (http://www.uniprot.org/help/uniprotkb).

Template selection

A $BLAST_P$ [1] search with default parameter was performed against the Brook Heaven Protein Data Bank (PDB) [6] to find suitable template for homology modeling. A set of PDB structures i.e., 3JZD_A, 3HL0_A, 3IV7_A, 3OWO_A and 3BFJ_A, presented close similarity with the target sequence. The template search, based on the functional domains concept, was carried out in HHPred, PDBe, Atom2, and based on all of these searches and also on parameters like maximum identity with high positives and lower gap percentage (%) (Table 1). Crystal structure of putative alcohol dehedrogenase (3JZD_A) was selected as the template since the percentage of query coverage, sequence identity and positives between the template and the target was 98%, 53% and 70% respectively. Mod web (http://salilab.org/modweb) [11] and PS^2 http://ps2.life.nctu.edu.tw/) was also employed to determine the best template in terms of MODELLER base server.

Construction of homology model

SWISS-MODEL Workspace is used for homology or comparative modeling of protein three-dimensional structures. The user provides an alignment of a sequence to be modeled with known related structures and SWISS-MODEL Workspace automatically calculates a model containing all non-hydrogen atoms. SWISS-MODEL Workspace implements comparative protein structure modelling by satisfaction of spatial restraints, and can perform many additional tasks, including de novo modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, etc. It assists and guides the user in building protein homology models at different levels of complexity [5, 25, 31]. ClustalW (http://www.ebi.ac.uk/research) was used for alignment between template and query sequence.

The secondary structural features of *Maleylacetate reductase* protein were considered for this study by SOPMA server [17]. A new highly accurate secondary structure prediction method, SOPMA incorporates two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST (position specific Iterated BLAST) [2] and PSIPRED Server (http://bioinf.cs.ucl.ac.uk/psipred/) also used to predict the secondary structure of this protein. The PSIPRED Protein Sequence Analysis Workbench aggregates several UCL structure prediction methods into one location.

Assessment of homology model

The 3D structure prediction model was assessed by Verify3D (http://nihserver.mbi.ucla.edu/Verify_3D/). Structural analysis and verification server (SAVES) (http://nihserver.mbi.ucla.edu/SAVES) has been used. Visualization protein model

was carried out by PyMOL v I.5 software Package. Structural validation of protein model was done by Rampage (http://mordred.bioc. cam.ac.uk/~rapper/rampage.php) which determines stereo chemical aspects along with main chain and side chain parameters with comprehensive analysis [23]. The Ramachandran plot of *Maleylacetate reductase* protein shows that various residues are falling under allowed, favored and outer regions. The non-bonded atomic interactions were analyzed by using ERRAT program [12]. Finally, the model quality assessment program ProQ has been used to evaluate the output candidate model [7].

Active site prediction

The active site of both template and newly generated structure was carried out by CASTp [19] server.

Protein structure accession number

The predicted 3D structure of *Maleylacetate reductase* protein was submitted to the Protein Model Database (PMDB) (http://mi.caspur.it/PMDB/) and assigned the PMDB ID PM0079119.

Results and discussion

In this study the *Maleylacetate reductase* has been retrieved from UniProtKB database and the sequences of related proteins with known 3D structures were checked for their suitability for homology modeling using BLASTp analysis in Table 1.

Table 1. BLASTp report of *Maleylacetate reductase*

PDB ID	Query coverage	Max Identity	E value	Gap	Positive
3JZD_A	98%	53%	3e-123	1%	70%
3HL0_A	99%	43%	1e-83	0%	62%
3IV7_A	99%	42%	9e-81	2%	58%
3OWO_A	94%	28%	2e-32	10%	45%
3BFJ_A	97%	27%	2e-32	11%	45%

The alignment between template and query protein sequence is shown below (Fig. 1).

```
Template        GXKSSQPFIYEAHAARVVFGAGSSSQVAAEVERLGAKRALVLCTPNQQAEAERIADLLGP 60
Query           -----MNFIHDYRSPRVIFGPDSLARLPQELERLGIDRALVLTTPEQAPLGRQVAEPVIG 55
                      **:: ::.**:**..* :::. *:**** .***** **:* . ..::*: :

Template        LSAGVYAGAVXHVPIESARDATARAREAGADCAVAVGGGSTTGLGKAIALETGXPIVAIP 120
Query           HVAAFYDGATMHVPALVAEEACKIARTSEANGVIAIGGGSTIGLAKIVALRTELPIVAVP 115
                  *..* **. ***   *.:*   ** : *: .:*:***** **.* :**.*  ****:*

Template        TTYAGSEVTPVYGLTEAGTKRTGRDPRVLPRTVIYDPALTVGLPRGLSVTSALNAIAHAA 180
Query           TTYAGSEMTSIFGITEGGVKKTGRDARVMPRAVIYEPRLTLELPLSISVTSAINAIAHAV 175
                *******:*.:::*:**.*.*:****.**:**:***:* **: ** .:*****:******.

Template        EGLYARDANPVXSLXAEEGIRALAAGIPAVFNDPADLDARSQCLYGAWLCGTVLGGVGXA 240
Query           EGLYAPDATPLLTIMAQEGIAATVRAISRMYQSPRDLQARGDALYGAWLCASVVGNVSMA 235
                ***** **.*: :: *:*** * . .*. :::.* **:**.:.*******.:*:*.*. *

Template        LHHKLCHTLGGSFNLPHAETHTIVLPHALAYNAAAVPEAXARIRRATGAGEQSAAATLFD 300
Query           LHHKLCHTLGGTLDLPHAQTHTVVLPHALAYNARAVPDAMRVLRIALGHDD--PPTALYE 293
                **********:::****:***********  ***:*   :* * * .:  ..::*::

Template        LAQRHGAPVALRDIGXREEDLDRAADIALASPYWNPRPIEREPIRALLQAAYEGVRPD- 358
Query           LARDNGAPVALRDLGMREEDIEHVGDLALQDRYPNPRELDRDALLALLRDAYHGRPPSA 352
                **: :********:* ****:::..*:** . * *** ::*:.: ***: **.*  *.
```

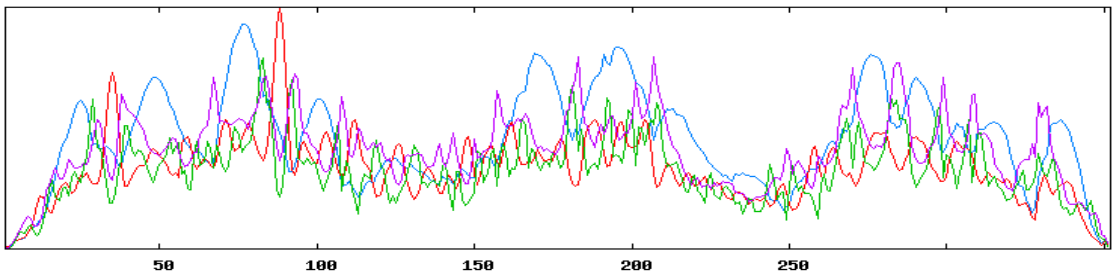
Fig. 1 Template-query sequnce alignment by ClustalW server

Secondary and tertiary structure prediction

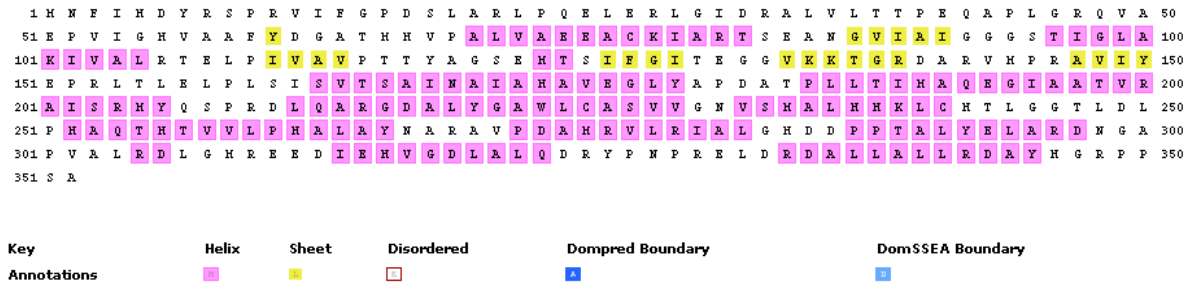
The secondary structures of proteins are the regularly repeating local structures stabilized by hydrogen bonds. The most common examples are the alpha-helix and beta-sheet. SOPMA view predicted the following secondary structures (Table 2): 175 (49.72%) in an alpha-helix, 35 (9.94%) in a beta-sheet, 119 (33.81%) in a coil and 23 (6.53%) in a beta-turn present at various positions in the *Maleylacetate reductase* protein structure of *Pseudomonas aeruginosa* (Fig. 2).

Table 2. Secondary structures predicted by SOPMA view

Protein structure, Unit	No. of amino acids	Percentage of structural, Unit
alpha-helix (Hh)	175.00	49.72
310-helix (Gg)	0.00	0.00
pi-helix (Ii)	0.00	0.00
beta-bridge (Bb)	0.00	0.00
extended-strand (Ee)	35.00	9.94
beta-turn (Tt)	23.00	6.53
bend-region (Ss)	0.00	0.00
random-Coil (Cc)	119.00	33.81
ambigous states	0.00	0.00
other states	0.00	0.00



(a)



(b)

Fig. 2 Secondary Structure prediction result of *Maleylacetate reductase* protein by SOPMA (a) and by PSIPRED server (b)

Tertiary structure

After the choice of a suitable template (3JZD_A), the model was constructed for the target protein using SWISS-MODEL Workspace (comparative Protein 3D modeling server. The predicted model was visualized under PyMOL visualization software. 3D structure of *Maleylacetate reductase* is shown below (Fig. 3).

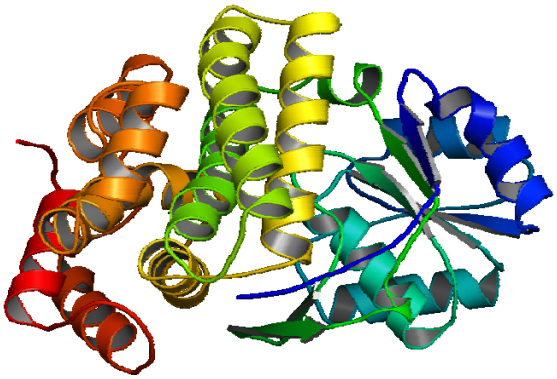


Fig. 3 3D structure of *Maleylacetate reductase*

Protein model validity

The geometrical and structural consistencies of both modeled and template proteins were evaluated by different approaches. The structural validation was carried out by PROCHECK, a well known protein structure checking program. The Φ and Ψ distributions of Ramachandran plot. This analysis revealed that only one residue (0.3%) in the Ramachandran plot of *Maleylacetate reductase* protein falls under the disallowed region. Overall, both homologues have nearly the same distribution in the steriochemically allowed main chain atoms (90.28%) (Fig. 4).

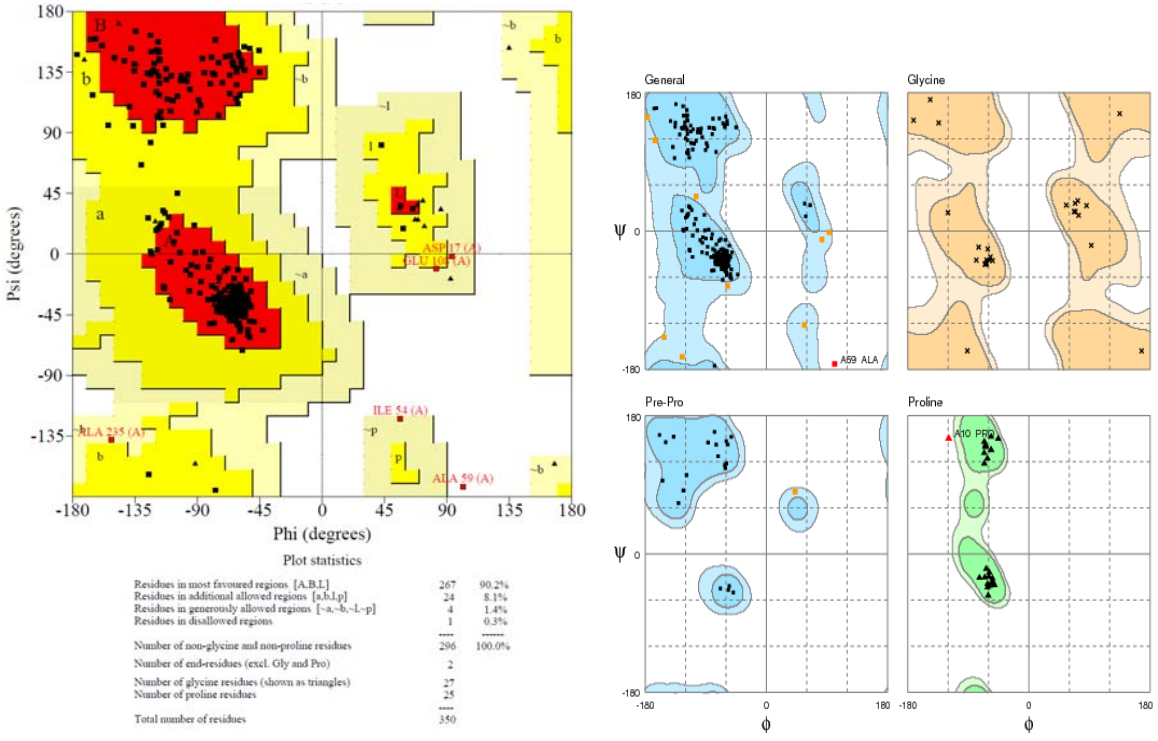


Fig. 4 Protein validation study by SAVE and Rampage Server

In addition, two more protein evaluation programs (Verify3D and ERRAT) were utilized to check the stereochemistry of our model. Verify3D (Fig. 5) scores the compatibility between the amino acid sequence and the environment of the amino acid side chains in the model. It assesses the environment of the side chain based on the solvent accessibility and the fraction of side chain covered by polar atoms. ERRAT assesses the arrangement of different types of atoms with respect to one another in the protein model. It is a sensitive technique, which is good at identifying incorrectly folded regions in preliminary protein models.

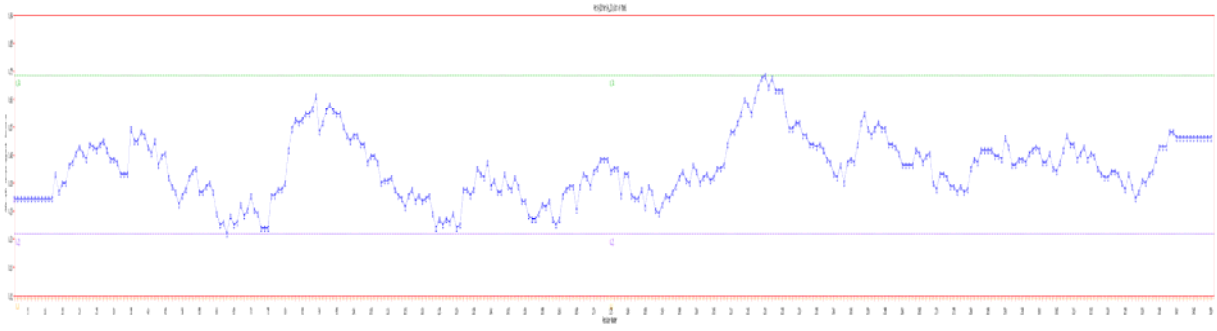


Fig.5 Verify3D Graph of *Maleylacetate reductase* protein

ERRAT plot (Fig. 6) shows that the developed structure of *Maleylacetate reductase* is acceptable (overall quality factor 90.80%).

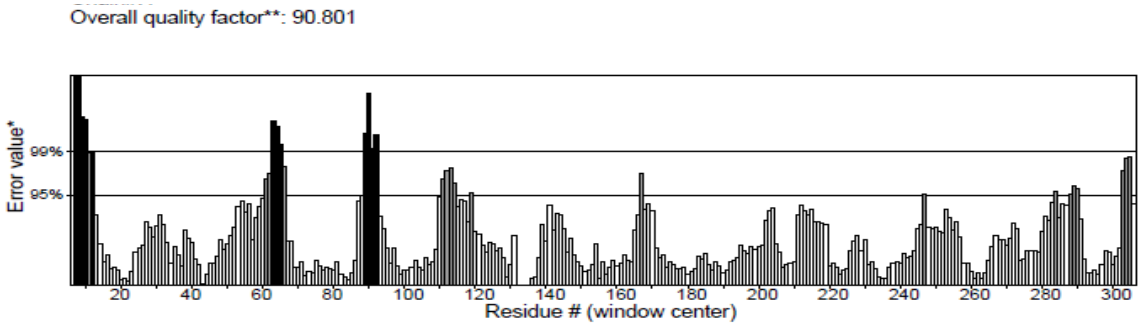
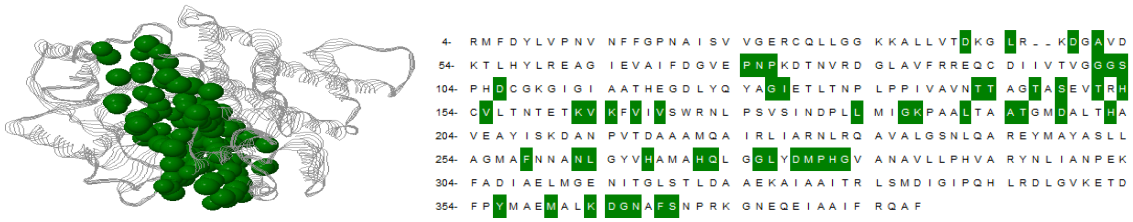


Fig. 6 ERRAT plot of *Maleylacetate reductase* protein

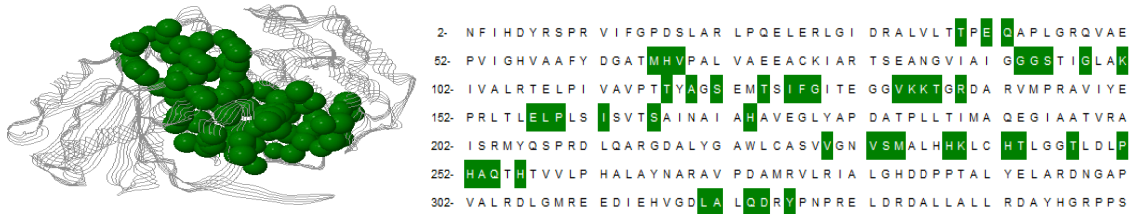
The modeled structure has been checked for quality by using ProQ. The result shows that predicted LG score: 7.776 (>4: extremely good model) and predicted MaxSub score: 0.747 (>0.5: very good model) were in acceptable range of a good model. Then the active site of the template and generated *Maleylacetate reductase* 3D protein structure was identified by CASTp server is shown below (Fig. 7).

The best pocket site was selected by the basis of area and volume of the protein active site. The area and volume of both template and *Maleylacetate reductase* 3D protein model structure is 1213.2, 2040.4 and 1179.6, 1879.9 respectively.

Finally, after its validation, the newly generated *Maleylacetate reductase* 3D protein model structure was then submitted in PMDB and got an authorship (PM0079119) of this protein structure.



(a)



(b)

Fig. 7 Active site prediction of template (a)
and *Maleylacetate reductase* (b) protein structure by CASTp server

Conclusion

The purpose of this study is to minimize the gap between *in silico* and wet lab determination of three-dimensional structure of a protein by molecular modeling. The three-dimensional structure model of *Maleylacetate reductase* protein was stable and proved reliable using the PROCHECK Verify3D and ProQ module. Most amino acids fall under α-helix secondary structure and this provides stability to the protein. The overall results provided the proofs that the predicted three-dimensional structure of *Maleylacetate reductase* protein is acceptable and of good quality. Predicting the structure for *Maleylacetate reductase* gives an idea of its active site and the active site residues which can be further analyzed by the use of software for the design of ligands and receptors.

Acknowledgements

We gratefully acknowledge the great encouragement and support of R&D Centre, Hi-Tech Medical College and Hospital research group. We are thankful to faculty members of BJB (A) College, Bhubaneswar for their help and encouragement for preparing the manuscript.

References

1. Altschul S. F., W. Gish, W. Miller, E. W. Myers, D. J. Lipman (1990). Basic Local Alignment Search Tool, J Mol Biol, 215, 403-410.
2. Altschul S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman (1997). Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, Nucl Acids Res, 25, 3389-3402.
3. Anderson J. J., S. Dagley (1980). Catabolism of Aromatic Acids in *Trichosporon cutaneum*, J Bacterial, 141, 534-543.
4. Armengaud J., K. N. Timmis, R.-M. Wittich (1999). A Functional 4-hydroxysalicylate/Hydroxyquinol Degradative Pathway Gene Cluster is Linked to the Initial Dibenzo-*p*-dioxin Pathway Genes in *Sphingomonas* sp. Strain RW1, J Bacteriol, 181, 3452-3461.

5. Arnold K., L. Bordoli, J. Kopp, T. Schwede (2006). The SWISS-MODEL Workspace: A Web-based Environment for Protein Structure Homology Modeling, Bioinformatics, 22, 195-201.
6. Berman H. M., J.Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne (2006). The Protein Data Bank, Nucl Acid Res, 28, 235-242.
7. Wallner B., A. Elofsson (2003). Protein Structure Prediction: Model Building and Quality Assessment, Protein Science, 12(5), 1073-1086.
8. Buswell J. A., K.-E. Eriksson (1997). Aromatic Ring Cleavage by the White-rot Fungus *Sporotrichum pulverulentum*, FEBS Letters, 104, 258-260.
9. Chapman P. J., D. W. Ribbons (1976). Metabolism of Resorcinylic Compounds by Bacteria: Alternative Pathways for Resorcinol Catabolism in *Pseudomonas putida*, J Bacteriol, 125, 985-998.
10. Chauhan A., G. Pandey, N. K. Sharma, D. Paul, J. Pandey, R. K. Jain (2010). P-nitrophenol Degradation via 4-nitrocatechol in *Burkholderia* sp. SJ98 and Cloning of Some of the Lower Pathway Genes, Environ Sci Technol, 44(9), 3435-3441.
11. Chen C. C., J. K. Hwang, J. M. Yang (2006). $(PS)^2$: Protein Structure Prediction Server, Nucleic Acids Res, 34, W152-W157.
12. Colovos C., T. O. Yeates (1993). Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions, Protein Sci., 2, 1511-1519.
13. Cámara B., P. Nikodem, P. Bielecki, R. Bobadilla, H. Junca, D. H. Pieper (2009). Characterization of a Gene Cluster Involved in 4-Chlorocatechol Degradation by *Pseudomonas reinekei* MT1, J Bacteriol, 191(15), 4905-4915.
14. Duxbury J. M., J. M. Tiedje, M. Alexander, J. E. Dawson (1970). 2, 4-D Metabolism: Enzymatic Conversion of Chloromaleylacetic Acid to Succinic Acid, J Agric Food Chem, 18, 199-201.
15. Feigel B. J., H.-J. Knackmuss (1993). Syntrophic Interactions During Degradation of 4 Aminobenzenesulfonic Acid by a Two Species Bacterial Culture, Arch Microbiol, 159,124-130.
16. Gaal A., H. Y. Neujahr (1979). Metabolism of Phenol and Resorcinol in *Trichosporon cutaneum*, J Bacteriol, 137, 13-21.
17. Geourjon C., G. Deleage (1995). SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments, Comput Appl Biosci, 11, 681-684.
18. Jain R. K., J. H. Dreisbach, J. C. Spain (1994). Biodegradation of p-nitrophenol via 1, 2, 4-benzenetriol by an *Arthrobacter* sp., Appl Environ Microbiol, 60, 3030-3032.
19. Dundas J., Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang (2006). CASTp: Computed Atlas of Surface Topography of Proteins with Structural and Topographical Mapping of Functionally Annotated Residues, Nucl Acids Res, 34, W116-W118.
20. Jones K. H., P. W. Trudgill, D. J. Hopper (1995). Evidence of Two Pathways for the Metabolism of Phenol by *Aspergillus fumigates*, Arch Microbiol, 163, 176-181.
21. Karasevich Y. N., V. S. Ivoilov (1977). Preparatory Metabolism of Para-hydroxybenzoic Acid in the Yeast *Candida tropicalis*, Microbiology, 46, 687-695.
22. Larway P., W. C. Evans (1965). Metabolism of Quinol and Resorcinol by Soil *Pseudomonas*, Biochem J, 95:52.
23. Lovell S. C., I. W. Davis, W. B. Arendall 3rd, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson (2003). Structure Validation by C Alpha Geometry: Phi, Psi and C Beta Deviation, Proteins, Structure, Function & Genetics, 50(3), 437-450.

24. Miyauchi K., Y. Adachi, Y. Nagata, M. Takagi (2009). Cloning and Sequencing of a Novel Meta-cleavage Dioxygenase Gene Whose Product is Involved in Degradation of γ-hexachlorocyclohexane in *Sphingomonas paucimobilis*, J Bacteriology, 181, 6712-6719.
25. Patchikolla S., A. A. Rao, S. K. Sangeeta, M. N. Babu, R. S. D. T. Grandhi (2010). Homology Modelling and Sequence Analysis of anxC3.1, International Journal of Engineering Science and Technology, 2(5), 1125-1130.
26. Rani N. L., D. Lalithakumari (1994). Degradation of Methyl Parathion by *Pseudomonas putida*, Can J Microbiol, 40, 1000-1006.
27. Schlomann M., P. Fischer, E. Schmidt, H.-J. Knackmuss (1990). Enzymatic Formation, Stability, and Spontaneous Reactions of 4-fluoromuconolactone, a Metabolite of the Bacterial Degradation of 4-fluorobenzoate, J Bacteriology, 172(9), 5119-5129.
28. Zhang S., W. Sun, L. Xu, X. Zheng, X. Chu, J. Tian, N. Wu, Y. Fan (2012). Identification of the Para-nitrophenol Catabolic Pathway and Characterization of Three Enzymes Involved in the Hydroquinone Pathway in *Pseudomonas* sp. 1-7, BMC Microbiology, 12, 27, http://www.biomedcentral.com/1471-2180/12/27.
29. Spain J. C., D. T. Gibson (1991). Pathway for Biodegradation of p-nitrophenol in a *Moraxella* sp., Appl Environ Microbiol, 57, 812-819.
30. Stolz A., H.-J. Knackmuss (1993). Degradation of 2,4-dihydroxybenzoate by *Pseudomonas* sp. BN9, FEMS Microbiology Letters, 108, 219-224.
31. Schwede T., J. Kopp, N. Guex, M. C. Peitsch (2003). SWISS-MODEL: An Automated Protein Homology-modeling Server, Nucleic Acids Research, 31, 3381-3385.
32. Mendez V., L. Agullo, M. Gonzalez, M. Seeger (2011). The Homogentisate and Homoprotocatechuate Central Pathways are Involved in 3- and 4-Hydroxyphenylacetate Degradation by Burkholderia xenovorans LB400, PLoS One, 6(3), e17583.
33. Seibert V., M. Thiel, I.-S. Hinner, M. Schlomann (2004) Characterization of A Gene Cluster Encoding the *Maleylacetate reductase* from *Ralstonia eutropha* 335^T, An Enzyme Recruited for Growth with 4-fluorobenzoate, Microbiology, 150, 463-472.

Shasank Sekhar Swain
E-mail: swain.shasanksekhar86@gmail.com,
swain.bioinformatics90@gmail.com

Shasank Sekhar Swain, Postgraduate from BJB (Autonomous) College, Bhubaneswar, Odisha under Utkal University, Odisha, India in 2012. Presently working as Research Scholar at R & D Centre, Hi-tech Medical College and Hospital, Bhubaneswar, Odisha. He has generic interest on structural bioinformatics, next generation sequence analysis, drug design, and cancer biology with identification and isolation of anti-diabetic compound from medicinal plants.