Global Journal of Science Frontier Research: G
Bio-Tech & Genetics
Volume 15 Issue 1 Version 1.0 Year 2015
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN:2249-4626&Print ISSN:0975-5896

In-Silico Characterization and Homology Modeling of Catechol 1,2 Dioxygenase Involved In Processing of Catechol- an Intermediate of Aromatic Compound Degradation Pathway

By Al-Hakim, Mahmudul Hasan, Rokib Hasan, Md. Hazrat Ali, Muhammad Fazle Rabbee, Marufatuzzahan, Hanif Mohammad Rejwan & Ziaul Faruque Joy

Shahjalal University of Science and Technology, Bangladesh

Abstract- Catechol 1, 2 dioxygenase (EC 1.13.11.1) is an enzyme intended to catalyze the degradation of catechol an intermediate of phenolic compound from *ortho*-mechanisms of the 3-oxoadipate pathway. Catechol 1, 2 dioxygenase plays a key role in the aerobic degradation of aromatic compounds, because it is the substrate for aromatic ring cleavage enzymes and as such it can be the starting point of many peripheral metabolic pathways. So, catechol 1, 2 dioxygenase is deliberated for a solution of environmental pollution occurred by aromatic compounds. In this study, we have focused on the *in-silico* characterization and homology modeling of catechol 1, 2 dioxygenase. The *in silico* analysis was performed by various computational tools and programmes. The physicochemical properties of the selected catechol 1, 2 dioxygenase were analyzed by using ExPASy'sProtParam tool and it was found that the molecular weight (M.Wt) ranges around 35000 Da. Isoelectric Points (pI) exhibits acidic nature and aliphatic index infers that 95% catechol 1, 2 dioxygenaseare stable. The negative value of GRAVY indicates that there will be better interaction with water. Motif analysis of the sequences was conducted by using MEME for predicting probable domain of catechol 1, 2-dioxygenase. Homology modeling of catechol 1, 2 dioxygenase taken from *Pseudomonas aeruginosa* MH38 (AC NO: CDH71767) was performed by I-TASSER. Various bioinformatics programmes and servers like RAMPAGE, PROCHECK and ERRAT were used for analysis and validation of final 3D structures created through homology modeling.

Keywords: *catechol 1, 2 dioxygenase, catechol, aromatic compounds, homology modeling.*

GJSFR-G Classification : *FOR Code: 060102*

Strictly as per the compliance and regulations of :

In-Silico Characterization and Homology Modeling of Catechol 1,2 Dioxygenase Involved in Processing of Catechol- an Intermediate of Aromatic Compound Degradation Pathway

Al-Hakim [α], Mahmudul Hasan [σ], Rokib Hasan [ρ], Md. Hazrat Ali [ω], Muhammad Fazle Rabbee[¥], Marufatuzzahan[§], Hanif Mohammad Rejwan [χ] & Ziaul Faruque Joy [ν]

Abstract- Catechol 1, 2 dioxygenase (EC 1.13.11.1) is an enzyme intended to catalyze the degradation of catechol an intermediate of phenolic compound from *ortho*-mechanisms of the 3-oxoadipate pathway. Catechol 1, 2 dioxygenase plays a key role in the aerobic degradation of aromatic compounds, because it is the substrate for aromatic ring cleavage enzymes and as such it can be the starting point of many peripheral metabolic pathways. So, catechol 1, 2 dioxygenase is deliberated for a solution of environmental pollution occurred by aromatic compounds. In this study, we have focused on the *in-silico* characterization and homology modeling of catechol 1, 2 dioxygenase. The *in silico* analysis was performed by various computational tools and programmes. The physicochemical properties of the selected catechol 1, 2 dioxygenase were analyzed by using ExPASy'sProtParam tool and it was found that the molecular weight (M.Wt) ranges around 35000 Da. Isoelectric Points (pI) exhibits acidic nature and aliphatic index infers that 95% catechol 1, 2 dioxygenaseare stable. The negative value of GRAVY indicates that there will be better interaction with water. Motif analysis of the sequences was conducted by using MEME for predicting probable domain of catechol 1, 2-dioxygenase. Homology modeling of catechol 1, 2 dioxygenase taken from *Pseudomonas aeruginosa* MH38 (AC NO: CDH71767) was performed by I-TASSER. Various bioinformatics programmes and servers like RAMPAGE, PROCHECK and ERRAT were used for analysis and validation of final 3D structures created through homology modeling.

Keywords: catechol 1, 2 dioxygenase, catechol, aromatic compounds, homology modeling.

I. Introduction

Phenol is one of the most widely used in the organic compounds in existence and is a basic structural unit for a variety of synthetic organic compounds including agricultural chemicals and pesticides. Phenols are aromatic compounds that are characteristic pollutants in waste water and effluents from chemicals, petrochemicals, pharmaceuticals, textiles, and steel industries [3]. The unwholesome and environmentally unacceptable pollution effects of the phenolic effluent have been reported worldwide [4]. Phenol contaminants are relatively soluble in water and accumulate in soil, resulting in extensive surface water, ground water, and soil contamination owing to its severe toxicity [5]. Currently removal of phenol effluents from contaminated sites has been a major environmental issue.

Different techniques have been applied to remove phenolic compounds from polluted areas [6–10]. However, among all, biodegradation process offers the more opportunities to completely destroy the pollutants if possible or at least to transform them to innocuous substance [11], it possess relatively low cost, no chemicals used, and high public acceptance [12]. Phenol and its derivatives are not easily biodegradable because they are toxic to most microorganisms. In higher concentrations, they can even inhibit the growth of microbial strains that are capable of assimilating them. Therefore, phenol is used in practice as an antimicrobial agent. It has been established that phenol can be toxic even in 0.05% concentrations [18]. The metabolism of aromatic compounds, phenol, and its derivatives in particular, is vigorously investigated in prokaryotic microorganisms [17]. A lot of information is accumulated on bacterial species from the *Pseudomonas* genus, which are known for their ability to utilize diverse aromatic compounds as a single carbon source and good degraders of phenol [19–21]. The ability of microorganisms to transform xenobiotics intocompounds that can enter the normal cycle of matter is due to specific microbial enzymes. Thus, the investigation of enzyme reactions including degradation and detoxification of phenol pollutants is the focus of attention for many researchers.

The metabolism of aromatic compounds and its regulation is extensively studied in prokaryotes. In 1973, Stainer reported that the 3-oxoadipate pathway was the main mechanism for degradation of aromatic compounds, including non-substituted phenol [13].

Author α: Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114 Bangladesh (Corresponding Author). e-mail: hakim@student.sust.edu
Author σ ρ ω ¥ § χ ν: Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh. e-mails: hasan_sust@yahoo.com, rokib29@gmail.com, hazratsust05@gmail.com, rabbi.biotech@gmail.com, marufatuzzahan@gmail.com, rejwan.sust@gmail.com, zfjoy-geb@sust.edu

There are both aerobic and anaerobic microorganisms that are able to complete the phenol degradation process [14]. Phenol-degrading aerobic bacteria have the ability to transform phenol into non-toxic intermediate compounds that enter the Tricarboxylic acid cycle through *ortho-* or *meta*-pathways of degradation [15].

The first step in both pathways is monohydroxylation at the *o*-position of the aromatic ring. The enzyme catalyzing these reactions and having a key role in the aerobic degradation of mono-aromatic compounds is a monooxygenase: phenol hydroxylase (EC 1.14.13.7) [16]. Phenol hydroxylase catalyzes the attachment of a hydroxyl group at the ortho-position of the aromatic ring, thus hydroxylating phenol to catechol. The first intermediate product of phenol degradation is catechol. The dioxygenase enzyme that catalyzes the aromatic ring cleavage of catechol and its derivatives realizes the critical step in the aerobic degradation of aromatic compounds in microorganisms. Two classes of such enzymes are identified on the basis of aromatic ring cleavage mechanisms: intradiol-dioxygenases, (intradiol, i.e., carbon bond between two hydroxyl groups) which use non-haem Fe (III) to cleave the aromatic ring at *ortho-* position regarding the hydroxyl substitutes; and extradiol- dioxygenases, (extra diol, i.e., between one of the hydroxyl groups and a nonhydroxylated carbon) which use non-haem Fe(II) or other two-valent metal ions to cleave the aromatic ring at meta-position with regard to hydroxyl groups.

The enzyme catechol 1,2-dioxygenase described in *Pseudomonas* is highly dependent on ferro- and ferri-ions and has high substrate specificity [22]. Recently, a new catechol 1,2-dioxygenase was isolated form a *Pseudomonas aeruginosa*TKU002 strain capable of assimilating benzoic acid as a single carbon source. The enzyme has unique characteristics, such as very low molecular mass (22 kD), highest activity against pyrogallol, high medium acidity for enzyme production, etc, which distinguishes it from other microbial catechol dioxygenases [23]. Catechol has a key role in the aerobic degradation of aromatic compounds, because it is the substrate for aromatic ring cleavage enzymes and as such it can be the starting point of many peripheral metabolic pathways. Only some of the catechol is subjected to direct aromatic ring scission during the cultivation of microorganisms on phenol as a single carbon source. The rest of the catechol can behydroxylated to pyrogallol, and after that transformed to α- or β-hydroxy-muconates [22,23]. In the previous study we did homology modeling and in silico structural analysis of phenol hydroxylase which is another enzyme strongly involved in aromatic compound degradation system[1].

The study is representing the computational study of catechol 1,2-dioxygenase and its homology modeling. Physiochemical properties, phylogenetic tree construction, motif election and homology modeling and model validation of catechol 1, 2-dioxygenase of *Pseudomonas aeruginosa*MH38(AC NO: CDH71767) are which will help us to better understand of catechol processing in aromatic compound degradation pathway.

II. Methods and Materials

a) Sequence retrieval

The sequences of catechol 1,2-dioxygenase were retrieved from NCBI (National Center for Biotechnology Information). Sequences retrieved was done by BLAST with *Pseudomonas aeruginosa* MH38(Accession numbers: CDH71767.1) in NCBI and 61 sequences are retrieved for further study.

b) Analysis of physicochemical parameters

The different physicochemical properties of catechol 1, 2-dioxygenase enzyme were computed using ExPASy'sProtParam tool and these properties can be deduced from a protein sequence. ProtParam tool is used to analyze various physiochemical properties of industrially important proteins[2]. The ProtParam includes the following computed parameters: Molecular weight (M.Wt), theoretical pI, instability index (II), aliphatic index (AI) and grand average of hydropathicity (GRAVY). The computed isolelectric point (pI) will be useful for developing buffer systems for purification by isoelectric focusing method [26]. The instability index provides an estimate of the stability of our protein. A protein whose instability index is smaller than 40 is predicted as stable; a value above 40 predicts that the protein may be unstable [27]. The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine). It may be regarded as a positive factor for the increase of thermo stability of globular proteins [28].

c) Motif election

Motif election is very important in case of predicting probable domain ofcatechol 1,2-dioxygenase. Motif election & domain analysis were done using MEME (http://meme-suite.org/) and Pfam (http://pfam.xfam.org/).

d) Construction of Homology Models of Catechol 1,2-dioxygenase

The amino acid sequence of catechol 1,2-dioxygenase of *Pseudomonas aeruginosa* MH38 (Accession numbers: CDH71767.1) was taken for homology modeling and p BLAST was done with PDB (Protein Data Bank). No significant similarity was found and for this why we have done homology modeling by I-TASSER(http://zhanglab.ccmb.med.umich.edu/I-TASSER/).

e) Validation of the models

Homology models generally contain errors in their initial structures. An essential part of homology

modeling is the verification or validation of the model. Several procedures were used to estimate errors in the 3D models [29]. Their stereo chemical quality and energetic parameters were evaluated to determine whether the bond lengths and angles were within normal ranges, or whether there were many bumps in the models (corresponding to high van der Waals energies). The structures were evaluated and validated by RAMPAGE, PROCHECK and ERRAT [30-32].

f) Rampage

Rampage is a program for visualizing and assessing the Ramachandran plot of a protein structure. It works on the basis of a manually curated set of high quality protein structures and a number of filters, reference phi/psi plots are derived for gly, pro, pre-pro and general residue type and subdivided into 'favoured', 'allowed' and 'outlier' regions will be listed, and a picture of the Ramachandran plotis displayed. The output, high-resolution multi-color Adobe PDF or PostScript file contains the general plot with critical data showing the percentage of residues that occurred in different regions[36].

g) Procheck

PROCHECK aims to assess how normal, or conversely how unusual, the geometry of the residues in a given protein structure, is compared with stereo chemical parameters derived from well-refined, high-resolution structures. The input to PROCHECK is a single file containing the coordinates of the protein structure. The outputs comprise a number of plots, together with a detailed residue-by-residue listing [36].

h) Errat

ERRAT is a program for verifying protein structures determined by crystallography. Error values will be plotted as a function of the position of a sliding 9-residue window. According to the analysis by ERRAT, the final model is significantly improved relative to the initial model. This program examines a PDB file, and generates a score based on the quality of the local structure surrounding each residue, based on the typical ranges of dihedral angles and side chain contacts observed in real proteins, generally speaking, the method is sensitive to smaller errors than 3-D profile analysis but is more forgiving than Procheck [36].

III. Result and Discussion

The physicochemical properties of catechol 1,2-dioxygenase were predicted by using ProtParam tool. The ProtParam includes the following computed parameters: Molecular Weight (M.Wt), theoretical pI, Instability Index (II), Aliphatic Index (AI) and grand average of hydropathicity (GRAVY) (**Table 1**). The physicochemical properties showed that molecular weight of maximum number of catechol 1, 2-dioxygenaseis around 40007.7 Da. The instability index showed that more than 50% catechol 1, 2-dioxygenaseis stable as their instability index stayed below 42. Isoelectric point (pI) is the pH at which the surface of protein is covered with charge but net charge of the protein is zero. The computed pI value showed that catechol 1, 2-dioxygenase are acidic in nature (pH<7). The instability index is used to measure in vivo half-life of a protein [32]. The proteins which have been reported as in vivo half-life of less than 5 h showed instability index greater than 40, whereas those having more than 16 h half-life [33] has an instability index of less than 40. Among the studied catechol 1, 2-dioxygenasemore than 95% protein showed stable nature having more than sixty hours of half-life as the contains instability index less than 40. In case of Aliphatic Index (AI) the studied catechol 1, 2-dioxygenaseshowed the tendency of having a wide range of temperature as showed Aliphatic Index (AI) above 70. GRAVY value of the studied catechol 1, 2-dioxygenaseshowed that maximum) most of them exhibit lower GRAVY value which indicates the better interaction of those proteins with water.

Table 1 : Physicochemical properties of catechol 1, 2-dioxygenase analyzed by Prot Param

Protien name	Organisms	Number of amino acids	Molecular Weight (M.W)	Theoretical pI	Instability index	Aliphatic index	Gravy
Catechol 1,2-dioxygenase	*Pseudomonas aeruginosa MH38*	310	34182.8	5.16	44.09	74.71	-0.594
3'-RNA processing protein	*Pseudomonas*	310	34159.8	5.24	43.95	74.71	-0.577
PA2507, partial	*synthetic construct*	311	34163.8	5.15	44.67	74.47	-0.554
Catechol 1,2-dioxygenase	*Pseudomonas aeruginosa*	309	34128.8	5.24	44.06	74.63	-0.589

Catechol 1,2-dioxygenase	*Pseudomonas aeruginosa PAO1H2O*	310	34217.0	5.16	46.20	74.06	-0.585
Catechol 1,2-dioxygenase	*Pseudomonas aeruginosa*	310	34189.9	5.24	43.68	75.97	-0.557
Catechol 1,2-dioxygenase	*Pseudomonas aeruginosa WS394*	360	40007.7	8.31	46.90	71.14	-0.687
Catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	282	31022.3	5.20	36.82	76.21	-0.539
catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	276	30404.7	5.18	36.95	77.50	-0.536
Catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	272	29977.2	5.18	37.94	77.57	-0.528
Catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	205	22005.2	5.10	40.48	76.29	-0.449
Catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	270	29750.0	5.08	38.14	77.78	-0.522
Catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	269	29664.8	5.08	38.56	77.32	-0.532
Catechol 1,2-dioxygenase, partial	*Pseudomonas aeruginosa*	272	29970.2	5.08	40.06	76.84	-0.543
Catechol 1,2-dioxygenase	*Pseudomonas denitrificans*	310	34006.8	4.84	40.74	81.26	-0.397
3'-RNA processing protein	*Pseudomonas sp. HYS*	310	34241.1	4.89	39.35	83.13	-0.396
Catechol 1,2-dioxygenase	*Pseudomonas alkylphenolia*	310	34139.1	4.82	38.64	82.81	-0.372
Catechol 1,2-dioxygenase	*Pseudomonas sp. LAIL14HWK12:I7*	311	34140.8	4.94	42.43	77.30	-0.395
Catechol 1,2-dioxygenase	*Pseudomonas sp. URMO17WK12:I8*	311	34243.1	5.07	41.03	79.81	-0.394
3'-RNA processing protein	*Pseudomonas fuscovaginae*	311	34293.1	5.29	40.56	78.23	-0.417

Catechol 1,2-dioxygenase, proteobacterial	*Pseudomonas stutzeri*	313	34481.2	5.09	39.02	78.91	-0.452
Catechol 1,2-dioxygenase	*Pseudomonas stutzeri*	312	34619.6	5.41	39.16	76.35	-0.537
Catechol 1,2-dioxygenase	*Pseudomonas plecoglossicida*	311	34181.8	4.95	40.85	77.27	-0.422
3'-RNA processing protein	*Pseudomonas putida group*	311	34077.0	5.35	35.54	80.45	-0.362
3'-RNA processing protein	*Pseudomonas stutzeri*	312	34331.1	5.35	47.09	77.28	-0.462
3'-RNA processing protein	*Pseudomonas sp. S9*	310	33856.6	4.81	43.47	78.10	-0.367
Catechol 1,2-dioxygenase	*Pseudomonas sp. HMP271*	312	34645.5	5.37	42.30	75.71	-0.542
3'-RNA processing protein	*Pseudomonas stutzeri*	312	34304.0	5.35	35.18	74.49	-0.526
3'-RNA processing protein	*Pseudomonas alcaligenes*	309	34066.8	4.81	40.24	79.00	-0.420
3'-RNA processing protein	*Pseudomonas putida*	311	34277.0	5.07	39.13	78.87	-0.422
3'-RNA processing protein	*Pseudomonas putida*	311	34279.1	5.21	40.23	78.55	-0.423
Catechol 1,2-dioxygenase	*Pseudomonas alcaligenes*	313	34337.5	5.43	32.99	83.29	-0.389
Catechol 1,2-dioxygenase	*Pseudomonas sp. 20_BN*	312	34391.1	5.03	40.83	79.49	-0.475
Catechol 1,2-dioxygenase	*Pseudomonas putida*	311	34263.0	5.07	40.25	80.13	-0.393
3'-RNA processing protein	*Pseudomonas stutzeri*	312	34577.5	5.21	41.40	80.10	-0.477
Catechol 1,2-dioxygenase	*Pseudomonas xanthomarina*	312	34584.5	5.49	42.06	74.78	-0.526

3'-RNA processing protein	*Pseudomonas putida group*	311	34320.1	5.14	41.85	79.16	-0.421
Catechol 1,2-dioxygenase	*Pseudomonas taeanensis*	312	34149.0	4.89	49.08	81.09	-0.343
3'-RNA processing protein	*Pseudomonas putida*	311	34264.1	5.14	38.44	79.49	-0.401
3'-RNA processing protein	*Pseudomonas chloritidismutans*	312	34614.6	5.49	43.74	74.78	-0.527
3'-RNA processing protein	*Pseudomonas putida*	311	34387.2	5.13	44.86	78.55	-0.440
3'-RNA processing protein	*Pseudomonas mendocina*	309	34088.1	5.22	41.94	78.35	-0.370
Catechol 1,2-dioxygenase	*Pseudomonas pseudoalcaligenes*	309	34240.2	4.86	34.88	81.81	-0.393
Catechol 1,2-dioxygenase	*Pseudomonas pseudoalcaligenes AD6*	309	34083.1	5.29	38.76	78.96	-0.390
Crystal Structure Of Catechol 1,2-dioxygenase	*Pseudomonas Arvilla C-1*	311	34312.2	5.20	40.18	79.49	-0.390
3'-RNA processing protein	*Pseudomonas putida*	311	34361.1	5.20	43.83	78.55	-0.446
3'-RNA processing protein	*Pseudomonas stutzeri*	312	34333.1	5.15	39.35	78.85	-0.482
Catechol 1,2-dioxygenase	*Pseudomonas cremoricolorata*	311	34140.9	4.90	43.46	78.23	-0.395
Catechol oxygenase	*Pseudomonas putida*	311	34301.2	5.36	39.11	80.74	-0.390
3'-RNA processing protein	*Pseudomonas monteilii*	311	34236.0	5.07	38.30	79.52	-0.391
Catechol 1,2-dioxygenase	*Pseudomonas*	311	34232.9	4.82	42.58	77.94	-0.435
3'-RNA processing protein	*Pseudomonas sp. EGD-AK9*	313	34514.6	5.42	40.83	82.01	-0.422

Catechol 1,2-dioxygenase	*Pseudomonas resinovorans*	309	33947.8	4.92	32.51	79.94	-0.388
3'-RNA processing protein	*Pseudomonas mendocina*	313	34684.8	5.33	44.03	81.73	-0.499
3'-RNA processing protein	*Pseudomonas sp. M1*	307	33845.7	5.33	37.28	82.41	-0.415
Catechol 1,2-dioxygenase	*Pseudomonas syringae*	309	34354.3	5.07	41.67	84.66	-0.397
Catechol 1,2-dioxygenase	*Pseudomonas sp. MOIL14HWK12:I2*	311	34290.9	4.78	42.64	77.94	-0.445
Catechol 1,2-dioxygenase	*Pseudomonas sp. URMO17WK12:I3*	309	34577.5	5.33	35.34	73.62	-0.537
3'-RNA processing protein	*Pseudomonas thermotolerans*	314	34952.1	5.22	44.83	82.13	-0.439
Catechol 1,2-dioxygenase	*Pseudomonas syringae*	309	34180.0	5.10	42.44	79.90	-0.475
Catechol 1,2-dioxygenase, partial	*Stenotrophomonasmaltophilia*	314	34513.4	5.23	41.63	79.65	-0.429

Motif helps to find out the functional domain of proteins and also motif represents the conserved pattern in protein sequences through which we can design degenerate primer of those protein sequences. Motif analysis of the sequences was conducted by using MEME. This is well known fact that E-value describes the statistical significance of the motif. According to Baker et al [34], by default, MEME looks for up to three motifs, each of which may be present in some or all of the input sequences.

Table 2 : MEME result showing sequence logos with width and respective E value

Motif No	Width	E Value	Sites	Sequence Logo
1	50	1.6e-2998	61	
2	50	1.0e-2920	61	
3	50	7.2e-2496	61	

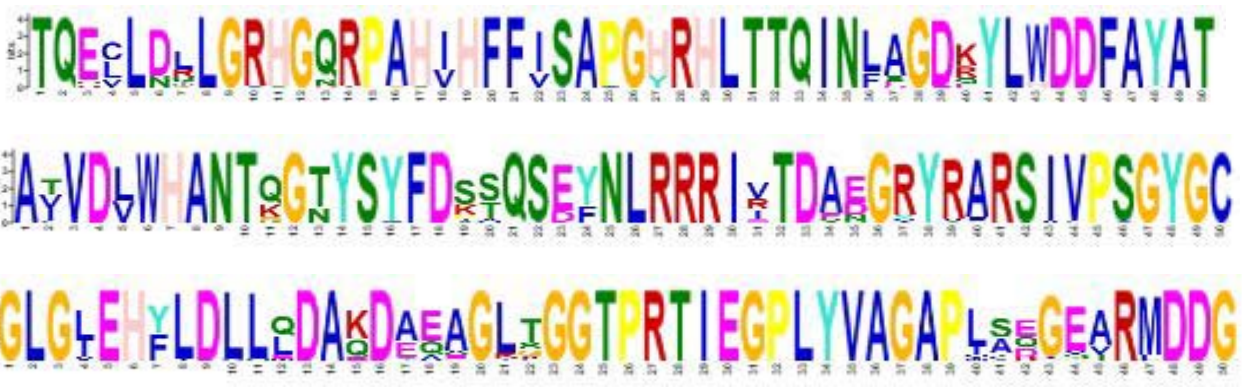

Figure 1 : Three motif of Catechol 1, 2-dioxygenase generated by MEME

MEME chooses the width and number of occurrences of each motif automatically in order to minimize the 'E-value' of the motif which increases the probability of finding an equally well-conserved pattern in random sequences. In our experiment, motif overview has shown in 'Table 3' describing 1.6e-2998 E-value of motif one, 1.0e-2920 E-value of motif two and 7.2e-2496 E-value of motif three. E-value, width, sites, sequence logo and regular expressions are given in 'Table 2'.

Catechol 1, 2-dioxygenase3D structure is very important in understanding it interactions, functions and their localization. Homology modeling is the most common structure prediction method. *Pseudomonas* species *Pseudomonas aeruginosa* MH38 (Accession numbers: CDH71767.1 was taken as a representative sequence in the experiment and we generated five 3D models (Figure: 2)of it by I-TASSER. Quality and reliability of structure was checked by several structure assessment methods including Rampage, ERRAT and Z-Score by various parameters and Ramachandran plot analysis.

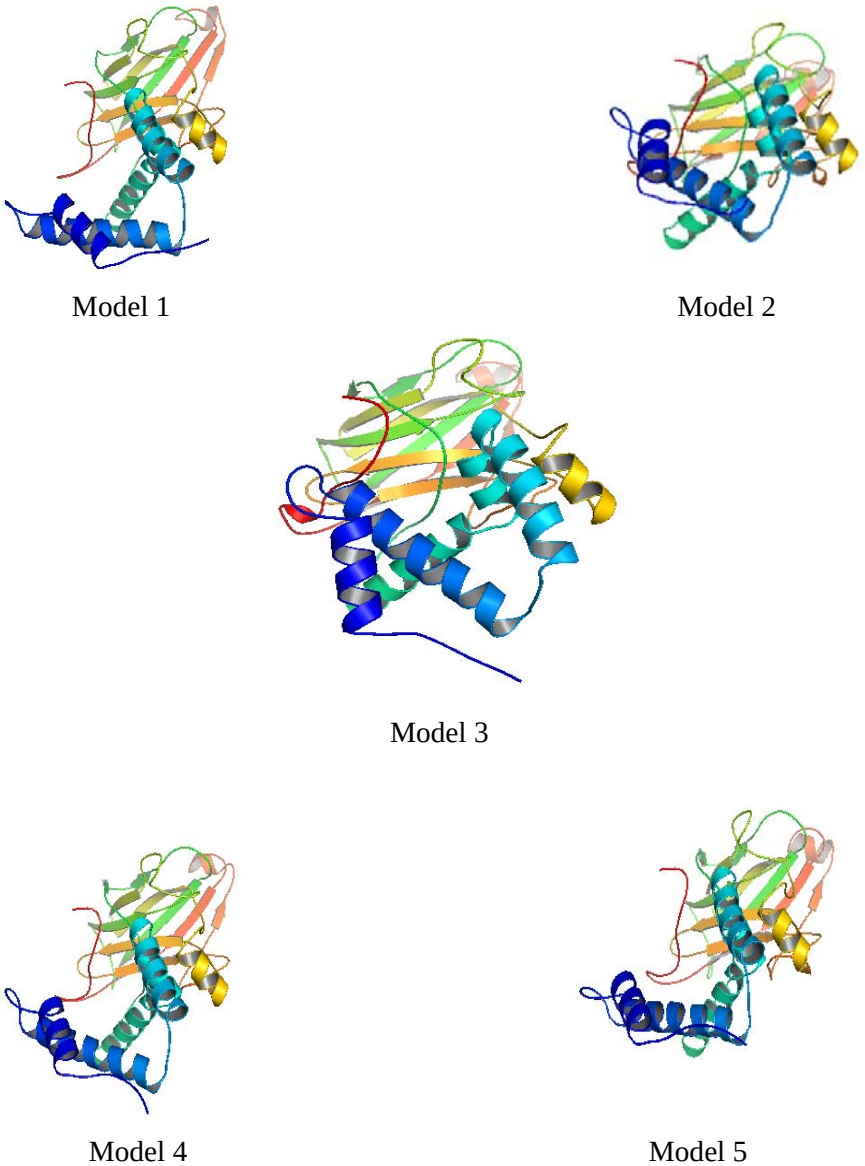

Figure 2 : 3D model of Catechol 1, 2-dioxygenase generated by I-TASSER

Among five models, model 4 has shown better conformational attribute in case of checking by several structure assessment methods. The stereochemical quality of the modeled protein was analyzed by RAMPAGE (Figure 3). Ramachandran plot analysis showed that only 1.0 % residues in outlier region, 2.6 % allowed region and 96.4 % in favored region, indicating that the models were of reliable and good quality.

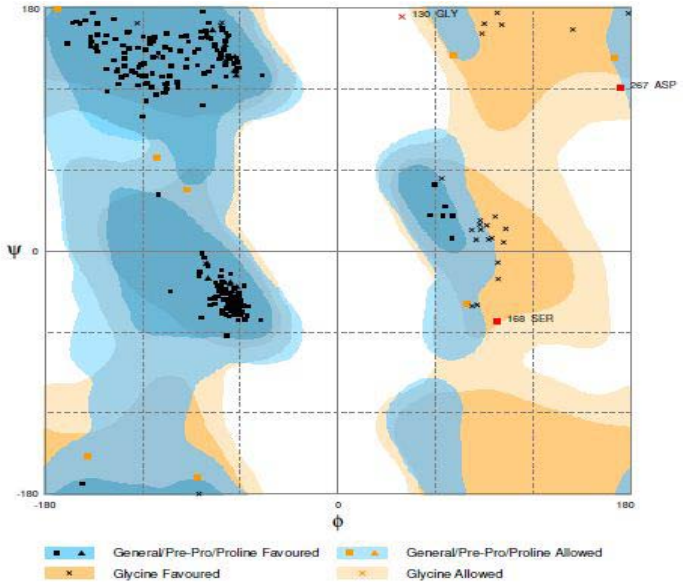

Figure 3 : Favoured amino acid residues in Ramachandran plot for 3-D structure of catechol 1, 2-dioxygenase generated by RAMPAGE

Procheck checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry. This tool was used to determine the Ramachandran plot to assure the quality of the model. The result of the Ramachandran plot showed that 77.7% of residues in favorable region and amino acid percentage in disallowed region as 1.2% (Figure 4).

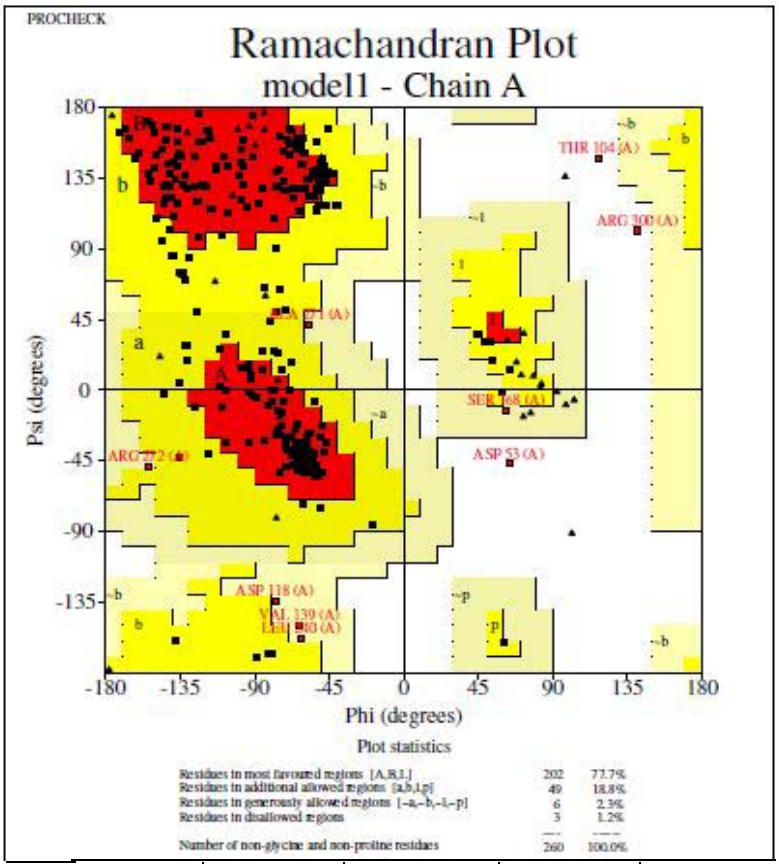

Figure 4 : Showing Ramachandran plot of phi versus psi angles of homology model of catechol 1, 2-dioxygenase generated by using PROCHECK. Color codes are Red color- most favorable regions, yellow color region- allowed region, and pale yellow generously allowed region and white color- disallowed regions

ERRAT is a protein structure verification algorithm that analyzes statistics of non-bonded interactions between different atom types based on characteristic atomic interaction [35]. ERRAT is a so-called "overall quality factor" for non-bonded atomic interactions, with higher scores indicating higher quality. The generally accepted range is >50 for a high quality model. For the current 3D model, the overall quality factor predicted by the ERRAT server was 86.093 (Figure-5).

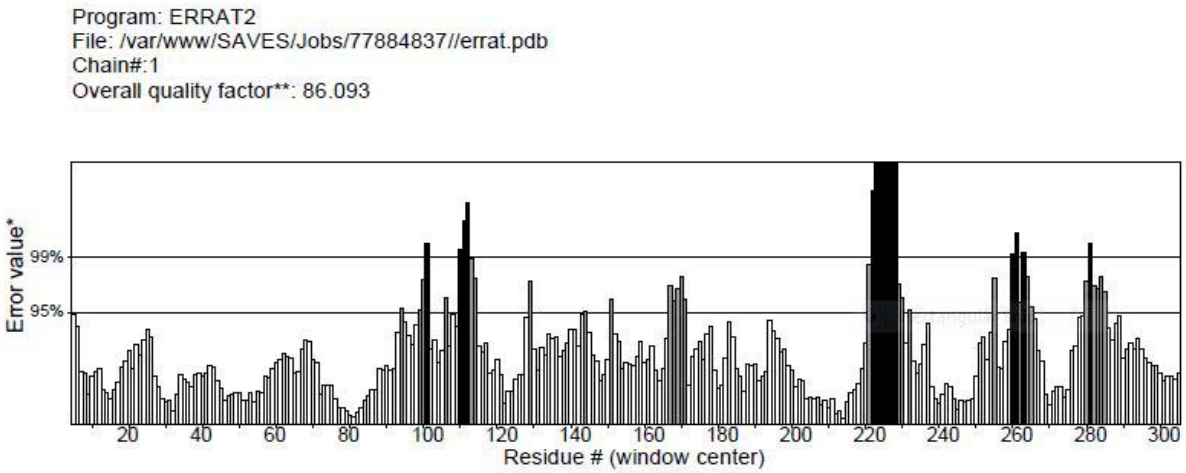

Figure 5 : ERRAT result showing amino acid distribution in error or non-error region for catechol 1, 2-dioxygenase. Black bars identify the misfolded region located distantly from the active site, gray bars demonstrate the error region between 95% and 99%, and white bars indicate the region with a lower error rate for protein folding

IV. Conclusion

In this study, catechol 1, 2-dioxygenase were selected and characterized from physicochemical perspectives.For this enzyme, molecular weight, theoretical isoelectric point, aliphatic index, instability index, Grand Average Hydropathy (GRAVY) was computed and that are essential and vital in providing data about catechol 1, 2-dioxygenase and their properties. Conserved sequences in motifs help us to culminate a significant insight of functional domain and also it may be utilized for designing specific degenerate primers for identification and isolation. Homology modeling generated five structural models and validation technique extols the best one. This 3-D structure may provide a new insight to better understand of structure and function. The study also focused on the future prospects of research in environmental safety concern.

References Références Referencias

1. Hasan M., Hakim A., Iqbal A., Farhana R.B., Kulsuma M.B. , Sharmin S., Abir R. A. (2015), "Computational Study and Homology Modeling of Phenol Hydroxylase: Key Enzyme for Phenol Degradation", *International Journal of Computational Bioinformatics and In Silico Modeling* . Vol. 4, No. 4, 691-698.
2. Hasan M., Joy Z.F., Elius H.B., Islam M. S. (2014), "In Silico Characterization and Motif Election of Neurotoxins from Snake Venom", *American Journal of Biochemistry and Biotechnology*, Volume 11, Issue 2 Pages 84-91.
3. Rocha LL, de Aguiar Cordeiro R, Cavalcante RM, et al. (2007), "Isolation and characterization of phenol-degrading yeasts from an oil refinery wastewater in Brazil," *My copathologia,* vol. 164, no. 4, pp. 183–188.
4. Ruiz-Ordaz N. Ruiz-Lagunez J.C., Castanon-Gonzalez J. A. et al., (2001),"Phenol biodegradation using a repeated batch culture of Candida Tropicalisina multistage bubble column," *Revista Latinoamericana de Microbiologia*, vol. 43, pp. 19–25.
5. Liu D., Liu H., Gao X., Leak D. J., and Zhou N. (2005),"Arg169 is essential for catalytic activity of 3-hydroxybenzoate 6-hydroxylase from *Klebsiell-apneumoniae* M5a1," *Microbiological Research*, vol. 160, no. 1, pp. 53–59.
6. Wu J., Rudy K., and Spark J. (2000), "Oxidation of aqueous phenol by ozone and peroxidase," *Advances in Environmental Research*, vol. 4, no. 4, pp. 339–346.
7. Lazarova Z. and Boyadzhieva S.(2004),"Treatment of phenol-containing aqueous solutions by membrane-based solvent extraction in coupled ultrafiltration modules," *Chemical Engineering Journal*, vol. 100, no. 1–3, pp. 129–138.
8. Kujawski W., Warszawski A., Ratajczak W., Pore T.¸bski, Capała W., and Ostrowska I.,(2004), "Removal of phenol from wastewater by different

separation techniques," *Desalination*, vol. 163, no. 1–3, pp. 287–296.
9. Hamad A., Aidan A., Fayed M., and Mehrvar M.(2005), "Experimental investigation of phenolic wastewater treatment using combined activated carbon and UV processes," *Clean Technologie and Environmental Policy*, vol. 7, no. 3, pp. 177–181.
10. Carmona M., de Lucas A. , Valverde J. L., Velasco B., and Rodr´ıguez J. F. (2006), "Combined adsorption and ion exchange equilibrium of phenol on Amberlite IRA-420," *Chemical Engineering Journal*, vol. 117, no. 2, pp. 155–160.
11. Vidali M. (2001), "Bioremediation: An overview" *Pure and Applied Chemistry*, vol. 73, no. 7, pp. 1163–1172.
12. Atlas R.M. and Unterman R., "*Bioremediation in: industrial Microbiology & Biotechno".*
13. Stainer, R. Y., Ornston, L. N. *(*1973), "The *β*-ketoadipate pathway". *Adv. Microb. Physiol. 9*, 89–151.
14. Agarry, S. E., Durojaiye, A. O., Solomon, B. O., "Microbial degradation of phenols: a review". *Int. J. Environ. Pollut.*,*32*, 12–28.
15. Bayly, C., Wigmore, J. *(*1973),"Metabolism of phenol and cresols by mutants of *Pseudomonas putida" J. Bacteriol. 113*, 1112–1120.
16. Dagley, S., Gibson, D. T. (1965), "The bacterial degradation of catechol". *Biochem. J. 95*, 46–74.
17. Watanabe, K., Yamamoto, S., Hino, S., Harayama, S. *(*1998), "Population dynamics of phenol degrading bacteria in activated sludge determined by *gyr* B-targeted quantitative PCR", *Appl.Environ. Microbiol. 64*, 1203**–**1209.
18. Kahru, A., Maloverjan, A., Sillak, H., Pollumaa, L. (2002),"The toxicity and fate of phenolic pollutants in the contaminated soils associated with the oil-shaleindustry", *Environ. Sci. Pollut. Res*. *1*, 27–33.
19. Hinteregger, C. R., Leinter, M., Loidl, A., Streichshier, F., (1992), "Degradation of phenol and phenolic compounds by *Pseudomonas* EKII", *Appl. Microbiol. Biotechnol.*,*37*, 252–295.
20. Kwon, K. H., Yeom, S. H.(2009), "Optimal microbial adaptation routes for the rapid degradation of high concentration of phenol", *Bioproc. Biosys. Eng32*, 435–442.
21. Seker, S., Beyenal, H., Salih, B., Tomas, A., (1997), "Multi-substrate growth kinetics of *Pseudomonas putida*for phenol removal", *Appl. Microbiol. Biotechnol*.,*47*, 610–614.
22. Nakai, C., Horiike, K., Kuramitsu, S., Kagamiyama, H. et al., (1990), "Three isozymes of catechol I, 2-dioxygenase (Pyrocatechase), *αα*, *αβ* and *ββ* from *Pseudomonas arvilla*C-I". *J. Biol. Chem.265*, 660–665.
23. Wang, C. L., You, S. L., Wang, S. L, (2006), "Purification and characterization of a novel catechol 1,2-dioxygenase from *Pseudomonas aeruginosa* with benzoic acid as a carbon source", *Proc. Biochem.*, *41*, 1594–1601.
24. Aldrich, T. L., Chakrabarty, A. M. *(*1988), "Transcriptional regulation, nucleotide sequence, and localization of the promoter of the cat BC Operon in *Pseudomonas putida"*. *J. Bacteriol. 170*, 1297–1304.
25. Neidle, E., Hartnett, C., Ornston, L. *(*1989), "Characterization of *Acinetobactercalcoaceticus*cat M, a repressor gene homologous in sequence to transcriptional activator genes". *J. Bacteriol. 171*, 5410–5421.
26. Sivakumar, K., Balaji S. and Gangaradhakrishnan.(2007), "In silico characterization of antifreeze proteins using computational tools and servers" *J.Chemical sci.* Volume 119, Issue 5, pp 571-579.
27. Guruprasad, K., Reddy B.V.B. and Pandit M.W., (1990), "Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting in vivo stability of a protein from its primary sequence", *Protein Eng.*, 4: 155-161. DOI: 10.1093/protein/4.2.155.
28. Walker, J.M., (2005), "The Proteomics Protocols Handbook". 1st Edn., Springer Science and Business Media, Totowa, N.J., ISBN-10: 1588293432, pp: 988.
29. Sippl, MJ (1993), 'Recognition of errors in three-dimensional structures of proteins', *Proteins*. 17(4):355–362.
30. http://mordred.bioc.cam.ac.uk/~rapper/rampage.php.
31. Colovos C1, Yeates TO, (1993), "Verification of protein structures: patterns of nonbonded atomic interactions", *Protein Science* Volume 2, Issue 9, pages 1511–1519.
32. 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. Applied Crystallography*, 26: 283-291. DOI:10.1107/S0021889892009944.
33. Rogers, S., Wells R. and Rechsteiner M., (1986), "Amino acid seq-uences common to rapidly degraded proteins: The PEST hypothesis" Science, 234: 364-368. DOI:10.1126/science.2876518.
34. Baker, E.N., Baker, H.M. and Kidd, R.D. (2002), "Lactoferrin and transferrin: Functional variations on a common structural framework"*, Biochem. Cell Biol*., 80: 27-34. DOI: 10.1139/o01-15.
35. Colovos VC, Yeates TO. (1993), "Verification of protein structures: Patterns of non-bonded atomic interactions"*, Protein Sci*, 2: 1511–1519.
36. J. Sarvanan., Sourav Dutta and Asit B. Mandal, (2014), "Modeling the interaction of Trehalose-6-Phosphate synthase and UDP-glucose bycomputational methods", *Indian Journal of Bioinformatics and Biotechnology*Vol 3 (2), 55-69.

This page is intentionally left blank