JOURNAL OF ENVIRONMENTAL SCIENCES 36 (2015) 144–151

Available online at www.sciencedirect.com

ScienceDirect

www.journals.elsevier.com/journal-of-environmental-sciences

JOURNAL OF ENVIRONMENTAL SCIENCES

www.jesc.ac.cn

Molecular level biodegradation of phenol and its derivatives through dmp operon of *Pseudomonas putida*: A bio-molecular modeling and docking analysis

Sujay Ray[1, 2,*], *Arundhati Banerjee*[3]

1. *Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, 741245 Nadia, West Bengal, India*
2. *Department of Biotechnology, Bengal College of Engineering and Technology, Shahid Sukumar Sarani, Bidhannagar, Durgapur-713212, West Bengal, India*
3. *Department of Biotechnology, National Institute of Technology, Mahatma Gandhi Avenue, Durgapur 713209, West Bengal, India*

ARTICLE INFO

Article history:
Received 16 December 2014
Revised 5 March 2015
Accepted 31 March 2015
Available online 14 July 2015

Keywords:
Biodegradation
dmp operon
Docking simulations
Modeling
Phenol
P. putida

ABSTRACT

Participation of *Pseudomonas putida*-derived methyl phenol (*dmp*) operon and DmpR protein in the biodegradation of phenol or other harmful, organic, toxic pollutants was investigated at a molecular level. Documentation documents that *P. putida* has DmpR protein which positively regulates *dmp* operon in the presence of inducers; like phenols. From the operon, phenol hydroxylase encoded by dmpN gene, participates in degrading phenols after *dmp* operon is expressed. For the purpose, the 3-D models of the four domains from DmpR protein and of the DNA sequences from the two Upstream Activation Sequences (UAS) present at the promoter region of the operon were demonstrated using discrete molecular modeling techniques. The best modeled structures satisfying their stereo-chemical properties were selected in each of the cases. To stabilize the individual structures, energy optimization was performed. In the presence of inducers, probable interactions among domains and then the two independent DNA structures with the fourth domain were perused by manifold molecular docking simulations. The complex structures were made to be stable by minimizing their overall energy. Responsible amino acid residues, nucleotide bases and binding patterns for the biodegradation, were examined. In the presence of the inducers, the biodegradation process is initiated by the interaction of phe50 from the first protein domain with the inducers. Only after the interaction of the last domain with the DNA sequences individually, the operon is expressed. This novel residue level study is paramount for initiating transcription in the operon; thereby leading to expression of phenol hydroxylase followed by phenol biodegradation.

* *Corresponding author*. E-mail: raysujay@gmail.com (Sujay Ray).

http://dx.doi.org/10.1016/j.jes.2015.03.035

Introduction

Phenol (C_6H_5OH) being a mildly acidic, volatile, organic pollutant is appreciably soluble in water (Weber et al., 2004). So, it is essential to completely remove phenol and its derivatives by biodegradation. A gram negative bacterium, *Pseudomonas putida* (*P. putida*), is found to be capable of consuming toxic compounds such as phenol, as their only carbon and energy source with optimum growth conditions being 30°C and pH = 6.8 (Şeker et al., 1997). Aerobic biodegradation of phenol is performed by gene encoding the N-fragment in *Pseudomonas putida*-derived methyl phenol operon (*DmpN* gene) that codes for phenol hydroxylase (Nordlund et al., 1990; Movahedyan et al., 2009).

Pseudomonas putida (CF-600) has DmpR protein, which acts as a regulator, positively controlling the expression of entire *dmp* operon (*dmpKLMNOPQBCDEFGHI*). This operon carries genes encoding certain enzymes necessary to breakdown phenols into pyruvate and acetyl-CoA (intermediates for Citric Acid Cycle) (Shingler et al., 1992; Sarand et al., 2001). DmpR protein comprises four domains namely Domain-A (the effector-sensing domain), Domain-B (a linker domain), Domain-C (transcriptional activator domain) and finally Domain-D (DNA binding domain). It is well documented through wet-laboratory research that in the presence of inducers like phenols, domain A interacts with inducers (Shingler and Pavel, 1995; Gupta et al., 2012). Then domain C interacts with domain A with the help of domain B and thereafter domain D binds to DNA structures present in two UAS in promoter (Shingler and Moore, 1994; Shingler and Pavel, 1995; Gupta et al., 2012). Interaction of DNA and domain D is aided by a DNA binding protein, Integration Host Factor (IHF) that introduces a sharp bend (>160°) in DNA, facilitating interaction between components in nucleoprotein array (Goosen and van de Putte, 1995). After these sequential interactions being successful, the operon opens up releasing phenol hydroxylase for biodegradation. It was also efficiently validated in the present study.

However, till date, detailed structural information regarding molecular level interactions between these proteins has not been dealt with. Although to understand thiosulfate oxidation, several molecular docking and interaction studies were examined in sox operon (*soxVWXYZABCDEFGH*) (Bagchi and Ghosh, 2011; Bagchi, 2011, Bagchi, 2012; Ray and Bagchi, 2013).

Present study is basically focused on residue level and molecular basis of degradation of toxic substances like phenols/phenolics with the involvement of dmp operon from *P. putida*, which thereby causes participation of gene encoding enzymes for the biodegradation. Current study therefore, includes analysis with description of 3D structures of four domains and two DNA structures of DmpR by discrete molecular modeling processes. Binding patterns and molecular level interactions were properly analyzed by performing several respective molecular docking simulations among domain A and inducers, among domains A–B–C, and domain D with two DNA structures independently. Involvement of the responsible amino acid residues and nucleotide bases was properly predicted and analyzed.

This is most probably an unexplored and novel description where the probe provides clear information regarding residue level interactions between domains of DmpR protein in the presence of inducers. It serves as a prime necessity for initiating transcription and thereby understanding biodegradation to create a sustainable environment.

1. Materials and methods

1.1. Sequence analysis and molecular modeling of domains A, B, C and D from DmpR protein of P. putida

1.1.1. Sequence analysis

The amino acid sequence of DmpR protein of *P. putida* was obtained from NCBI nucleotide database (Accession No.: X68033.1) and were verified from UniProtKB too. Pfam was used to identify the conserved domains (highly conserved regions especially responsible for protein functionality) (Punta et al., 2011). BLAST validated the results. Domains often serve as an important interaction site for proteins too (Jones et al., 1998; George and Heringa, 2002). Four domains were identified as XylR_N, V4R, Sigma54_activat and HTH_8 for domains A, B, C and D respectively. Total lengths of domains A, B, C and D were found to be 103, 62, 168 and 42 amino acid residues long respectively. Amino acid sequences of these four domains were identified. These amino acid sequences were used separately to build homology models by MODELLER.

1.1.2. Homology modeling of domains C and D

The amino acid sequences of these two domains were used separately to build homology models by MODELLER. For homology modeling, the widely accepted prerequisite lies that the sequence identity of the templates with target protein should be more than 30% (Sander and Schneider, 1991; Xiang, 2006). Results from HH-Pred (Söding et al., 2005) inferred that domains C and D had their templates in X-ray crystal structure from *Salmonella typhimurium* (PDB code: 1OJL, chain A) and *Escherichia coli* (PDB code: 1ETO, chain A); sharing 55% and 38% sequence identity respectively. BLAST against PDB corroborated the results from HH-Pred. HHpred is highly sensitive method for homology detection or structure prediction and quite often allows to make inferences from more remotely homologous relationships (Söding et al., 2005). Root Mean Square Deviations (RMSD) are then used to study globular protein conformation by superimposing C_α atoms of modeled structures on each of their respective original crystal templates. PyMOL (a *MOL*ecular viewer tool using *P*ython Language) yielded 0.691 Å and 0.977 Å RMSD for domains C and D respectively.

1.1.3. Modeling of domains A and B

No suitable templates were found for domains A and B from HH-Pred as well as from BLAST against PDB. Raptor-X (Källberg et al., 2012) was used to perform remote homology modeling for them. Remote homology modeling from protein sequence poses a challenging feature as structure is typically

more conserved than sequence over long evolutionary distances. Modeled structures from Raptor-X demonstrated larger percentage of disordered regions.

Furthermore, I-TASSER (Roy et al., 2010) and Phyre2 (Kelley and Sternberg, 2009) were utilized for fold recognition phenomena of modeling. I-TASSER executes programmed prediction for protein structure and function based on the sequence-to-structure-to-function prototype (Roy et al., 2010). Phyre2 is typical of many structure prediction systems using such profile–profile matching algorithms that can reliably detect up to twice as many remote homologies as per the usual sequence–profile searching (Kelley and Sternberg, 2009).

To select the most suitable structure for domains A and B, a comparative study was done by subjecting the domains to QUARK also. QUARK calculates Template Modeling(TM) score (algorithm to quantitatively assess accuracy of protein structure predictions relative to experimental structure, independent of protein lengths) for proteins. TM-score ranges in (0,1) (Xu and Zhang, 2010; Zhang and Skolnick, 2004). QUARK breaks the initial query sequences into fragments of 1–20 residues from which full-length structure models are assembled using replica-exchange Monte-Carlo simulation guided by composite knowledge-based force field. QUARK constructs 3D models of correct folds (Xu and Zhang, 2012). TM-scores for domains A and B were 0.3869 ± 0.0833 and 0.4933 ± 0.0833, respectively.

1.2. Best model selection

ERRAT values and Verify 3D values were compared for the best model selection in domains A and B individually. ERRAT is a program for authenticating protein structures determined by crystallography (Colovos and Yeates, 1993). Verify 3D examines the compatibility of an atomic tertiary model (3D) with its own amino acid sequence (1D) (Eisenberg et al., 1997). More is the ERRAT value and the Verify 3D value, the final model is brought nearer and is more improved relative to the initial native model (Eisenberg et al., 1997; Colovos and Yeates, 1993) (Suppl.). Table 1 shows that the ERRAT (Colovos and Yeates, 1993) and Verify 3D (Eisenberg et al., 1997) values for the models by QUARK are sufficiently more. So, the models by QUARK were selected for present study.

Table 1 – Protein–protein side chain–side chain hydrogen bonds.

POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM
4	C	GLN	NE2	38	X	GLY	O
4	C	GLN	NE2	38	X	GLY	O
11	C	VAL	N	32	X	TYR	OH
30	C	GLU	N	37	X	MET	SD
40	C	ARG	NE	57	A	SER	O
34	A	ARG	NH2	84	C	SER	O
34	A	ARG	NH2	84	C	SER	O
55	A	TYR	OH	49	C	ALA	O
69	A	ARG	NH1	32	C	GLY	O
69	A	ARG	NH1	32	C	GLY	O
69	A	ARG	NH2	32	C	GLY	O
69	A	ARG	NH2	32	C	GLY	O
4	X	SER	N	14	C	MET	SD
4	X	SER	OG	166	C	PRO	O
32	X	TYR	OH	10	C	THR	O

POS: position of the respective responsible amino acid residues, RES: responsible amino acid residues. CHAIN: represents the respective protein domain (for example, C: represents domain C from ACB protein domain complex, A: represents domain A from ACB protein domain complex, X: represents domain B from ACB protein domain complex) For example, 11 VAL N 32 X TYR OH, means nitrogen atom of Valine (position 11) from domain C interacts with hydroxyl group of tyrosine (position 32) from domain B (named as X here). Other atom names are used here such as SD, NE and NH2 as per PDB atomic nomenclature or record.

1.3. Loop optimisation using ModLoop

The modeled structures were subjected for loop optimisation. Often loop regions occupy disallowed regions in Ramachandran Plot. They need to be optimized and re-modeled using ModLoop for proper conformation of ψ–φ angles. ModLoop performs automated modeling of loops in protein structures using MODELLER. It calculates and analyses conformation in loops by the fulfillment of spatial restraints (Fiser and Sali, 2003).

1.4. Protein structure refinement

The loop optimized structures were further subjected to ModRefiner for their structure refinement.

ModRefiner follows an algorithm for the refinement of the protein structure model at its basic atomic-level. It can begin directly from C_α trace, main-chain model or full-atomic model. Here, the conformational search is jointly conducted by physics-based as well as knowledge-based force field. This algorithm draws the initial starting models of proteins, closer to their native state (where, maximum interaction between residues exists), not only in terms of hydrogen bonds but also in terms of topology of the backbone and positioning of the side-chains (Xu and Zhang, 2011).

1.5. Molecular modeling of two DNA sequences from DmpR protein

DmpR protein interacts with the promoter on two different UAS who has DNA sequences where domain D binds (Shingler et al., 1992). The DNA sequences are AGCATTTGCT and AAATGCTTA in 5′–3′ direction (Shingler et al., 1992). These DNA sequences were modeled using w3DNA-3D DNA Structure webserver. It is user-friendly web-based interface which builds 3D models of user uploaded nucleotide sequences. The visualization component presents illustrations of nucleic-acid structures involving images of bases and base pairs (Zheng et al., 2009).

1.6. Energy minimisation of the four domains and the two DNA models

Domains and DNA were energy minimized using CHARMM (Chemistry at HARvard Macromolecular Mechanics) force field to minimize the overall energy and to change the conformations of their dihedral angles so that they vary simultaneously (Brooks et al., 1983). Energy minimisation for domains and DNA structures was performed by steepest descend technique first, followed by conjugate gradient using Discovery studio

software, until they reached final RMS gradient of 0.0001 (Brooks et al., 1983).

1.7. Validation of models

Main chain properties of modeled structures of four domains of DmpR protein were verified by SAVeS server. To check amino acid residue profile of 3D models of domains, VERIFY3D was used (Eisenberg et al., 1997). Analysis of Ramachandran plots using PROCHECK webserver showed that there were no amino acid residues in the disallowed region (Laskowski et al., 1993; Ramachandran and Sashisekharan, 1968).

1.8. Molecular docking studies of domain A with inducers

For viewing conformational changes in domain A, in the presence of inducers like phenols and 2-chlorophenols, AutoDock Vina was utilized. It accepts PDBQT molecular structure file format as input and output (Trott and Olson, 2010). Only structures of molecules to be docked and specification of search space including the binding site were required. The best model is selected according to their respective values of binding energy (Gibbs free energy ΔG). The model with maximum negative ΔG value is selected as its binding affinity is maximum.

1.9. Molecular docking studies of domains A, B and C

The modeled domains A, C and then B were first docked using fully automated web-based program-ClusPro 2.0 server (Comeau et al., 2004; Kozakov et al., 2013). Initially, domain A was docked with domain C for proper interaction. Domain C, being longer, was uploaded as receptor whereas the latter was uploaded as ligand. Using advanced option of ClusPro 2.0, unstructured residues were removed. Model 0 was chosen and analyzed because it had the best cluster size among all possibly docked structures. Domain AC complex (receptor) and domain B (ligand) were then docked, in similar method by removing unstructured residues. Model 0 was again analyzed for finally docked protein complex. All protein–protein docking studies were also performed and validated using GRAMM (Vakser, 1995) and ZDOCK (Chen et al., 2003).

1.10. Energy minimisation of docked models

The docked models of the protein–protein complexes were also energy minimized in two steps. Firstly, all main and side chain atoms of proteins were allowed to be flexible. In the next step, side chains were allowed to remain flexible for proper interactions keeping backbones of proteins fixed. Energy minimisation was done by steepest descend technique first, followed by the conjugate gradient using Discovery studio software by CHARMM force field, until docked models reached final RMS gradient of 0.0001 (Brooks et al., 1983).

1.11. Molecular docking studies of DNA models with domain D

Binding of domain D with two DNA models, was executed using Hex server. It is Fast Fourier Transform (FFT)-based protein docking server (Macindoe et al., 2010). This server uses two graphic processors simultaneously. Thus, it completes 6D docking run within 15 sec. It is faster than conventional FFT-based docking approaches. Patch-Dock validated the results (Schneidman-Duhovny et al., 2005).

1.12. Calculation of protein-DNA and protein-ligand interaction

Domain D having helix–turn–helix motif conformations interacts with DNA structures individually. After performing protein–DNA and protein–ligand docking, interactions were examined using PyMOL (DeLano, 2002) and Discovery Studio Platform from Accelrys software. Covalent and hydrogen bonding for domain D with DNA structures and interaction patterns like π–π aromatic stacking between residues of domain A with inducers were studied in PyMOL (DeLano, 2002; Sivasakthi et al., 2013) and Discovery Studio Platform from Accelrys software.

1.13. Calculation of protein–protein interaction

Protein Interaction Calculator (P.I.C.) and Discovery Studio Platform from Accelrys software identified varied interactions, such as disulfide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic–aromatic interactions, aromatic–sulfur interactions and cation–π interactions in docked domain A–B–C complex (Tina et al., 2007). So, PDB file of the docked protein complex was uploaded for interaction calculation of domains A–C, accomplished by flexible linker-domain-B.

2. Results

2.1. Structural description of domains A, B, C, and D

Completely modeled structure of domain A (XylR_N) from DmpR protein of *P. putida* is 103 amino acid residues long as seen in Fig. S1. It starts with a tiny helix (residues 2–5), followed by three anti-parallel β-sheet regions (residues 6–9, 13–17 and 20–23) interspersed with coil regions. Three anti-parallel β-sheet regions were found to be further followed by two major helical regions (residues 27–40 and 43–68).

Modeled domain B (V4R) from DmpR protein is 62 amino acids long depicted by Fig. S2. It starts with two helical regions (residues 4–14 and 21–38) interconnected by small coils. Two anti-parallel β-sheet regions (residues 44–50 and 54–60) follow them.

Modeled domain C (Sigma54_activat) from DmpR protein is 168 amino acid residues long demonstrated by Fig. S3. Predicted structure is very similar to the structure of *Salmonella typhimurium* (PDB code: 1OJL, chain A) for domain C (Sigma54_activat). Domain C, a transcriptional activator domain, is the biggest among all domains of DmpR protein. Residues show conformational adaptability towards helix, β-sheet and coil conformations. It begins with helical region (residues 6–19), followed by parallel β-sheet regions and helical regions, interconnected with coil regions. Six helical and five parallel β sheet regions in the complete modeled structure (residues 6–19, 35–44, 63–68,107–118, 144–151,

156–162 as helical structures and 25–28, 54–57, 96–99, 136–140 and 165–168 as parallel β sheet structures) exist.

Modeled structure of domain D (HTH_8) from DmpR protein is 42 amino acid residues long, as depicted in Fig. S4. Predicted structure is very akin to the structure from *E. coli* (PDB code: 1ETO_A). Only three helical regions (residues: 3–17, 22–28, 32–40) with interconnected coil regions exist. No β-sheets were found. It is a DNA-binding protein and shares as helix–turn–helix motif.

2.2. Structural description of DNA1 and DNA2

DNA structure was observed to have (5′AGCATTTGCT3′) and (5′ AAATGCTTA3′) coiled with its complementary strand 3′ AGCAAATGCT5′ and 3′TAAGCATTT5′ (as in Figs. S5 and S6) respectively. Major and minor grooves were clearly visible.

2.3. Interaction of domain A with inducers (phenol and 2-chlorophenol)

Domain A was docked with inducers like phenol and 2-chlorophenol, individually, to study possible mode of their interaction using Autodock Vina. The model with maximum negative ΔG value is selected as its binding affinity is maximum. Interactions were observed and examined by PyMOL (DeLano, 2002; Söding et al., 2005; Sivasakthi et al., 2013) and Discovery Studio Platform from Accelrys software. An aromatic stacking interaction with phe-50 was analyzed in both cases. Benzene ring of phenol and 2-chlorophenol was perceived to interact with the aromatic ring of phe-50 by π–π stacking interactions. Electron-rich π system is often formed due to overlapping of p-orbitals in interaction with metal, anion, another molecule and even another π system (Anslyn and Dougherty, 2005). Generally, non-covalent interactions indulging π systems are crucial to biological events such as protein–ligand recognition (Meyer et al., 2003). Fig. 1 represents the interaction.

2.4. Interaction involving formation of domain AC complex and with domain B

In order to obtain the interactions between domain-A (the effector sensing domain), domain-C (transcriptional activation domain) and domain-B (linker domain), 3D structures of modeled domains were docked step-wise. They were found to interact strongly as shown in Fig. 2. Protein Interaction Calculator (P.I.C.) and Discovery Studio Platform from Accelrys software recognized various kinds of interactions like hydrogen-bond residues, hydrophobic interactions, aromatic–aromatic interactions etc. within a protein or between proteins in the complex. There exist extensive H-bonding interactions involving both, main and side chains of two proteins represented by Tables 1 and 2. Hydrophobic interactions also existed. Apart from this, protein–protein ionic interaction, aromatic–aromatic

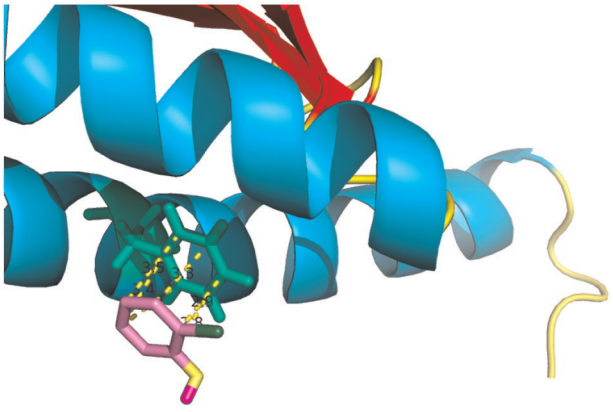

Fig. 1 – Protein (domain D)-ligand (2-chlorophenol), π–π stacking interaction (shown in dashed lines) between phenylalanine-50 (in green shade) and aromatic ring of 2-chlorophenol (in violet shade).

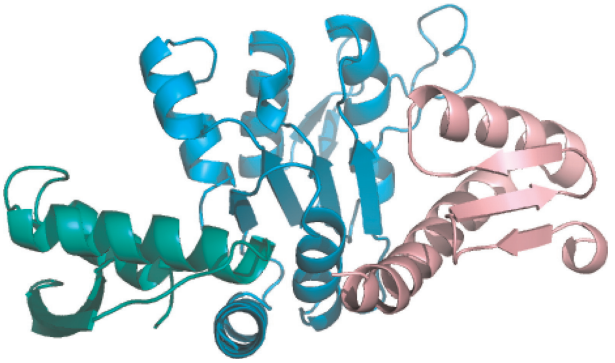

Fig. 2 – Interaction between domain A (in green shade), domain B (in pink shade) and domain C (in cyan shade).

Table 2 – Protein–protein main chain–side chain hydrogen bonds.

POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM
17	C	LYS	NZ	3	X	ASP	OD1
17	C	LYS	NZ	3	X	ASP	OD2
40	C	ARG	NH2	61	A	ASP	OD2
40	C	ARG	NH2	61	A	ASP	OD2
88	C	ARG	NE	30	A	MET	SD
88	C	ARG	NH2	30	A	MET	SD
88	C	ARG	NH2	30	A	MET	SD
91	C	ARG	NH1	30	A	MET	SD
91	C	ARG	NH1	30	A	MET	SD
163	C	ASN	ND2	8	X	GLU	OE1
163	C	ASN	ND2	8	X	GLU	OE1
34	A	ARG	NH1	83	C	GLN	OE1
34	A	ARG	NH1	83	C	GLN	OE1
34	A	ARG	NH2	83	C	GLN	OE1
34	A	ARG	NH2	83	C	GLN	OE1
66	A	ARG	NH1	8	C	TYR	OH
66	A	ARG	NH1	8	C	TYR	OH
4	X	SER	OG	14	C	MET	SD
8	X	GLU	OE1	163	C	ASN	ND2
8	X	GLU	OE1	163	C	ASN	ND2
32	X	TYR	OH	14	C	MET	SD

POS: position of the respective responsible amino acid residues, RES: responsible amino acid residues. CHAIN: represents the respective protein domain (for example, C: represents domain C from ACB protein domain complex, A: represents domain A from ACB protein domain complex, X: represents domain B from ACB protein domain complex) For example, 11 VAL N 32 X TYR OH, means nitrogen atom of Valine (position 11) from domain C interacts with hydroxyl group of tyrosine (position 32) from domain B (named as X here). Other atom names are used here such as SD, NE and NH2 as per PDB atomic nomenclature or record.

interaction, a pair of aromatic–sulfur interaction, and cation–pi interactions were perceived with few amino acid residues. No disulfide bonds were observed.

2.5. Interaction of domain D with two DNA sequences

Docking was performed to view stacking of two DNA structures with domain D individually using Hex Server (Macindoe et al., 2010) and Patch-Dock (Schneidman-Duhovny et al., 2005). PyMOL and Discovery Studio Platform from Accelrys software aided in interaction calculation in DNA-domainD complex (DeLano, 2002).

In DNA1, Cytosine-3 and Adenine-16 of the complementary strand form hydrogen bonds (shown in dotted lines in Fig. 3) with Ser8 and Glu7 of Domain-D respectively. Amino acid residues of domain D involved in polar contacts are Leu, Asn-Gln-Asn, Gly-Leu-Ser and Lys in the 2nd, 18th–20th, 29th–30th and 41st residues respectively.

In DNA2, Adenine-1 in DNA strand forms hydrogen and covalent bonds with Lys41 and only hydrogen bond with Glu6 of domain D. Guanine-13 and Thiamine-17 of complementary strand form hydrogen bonds with Leu2 and Ser31 of domain-D. Fig. 4 represents the interactions.

3. Discussion

Complete biodegradation of hazardous organic pollutants mainly phenols/phenolics, is essential for environmental sustainability or else, they might pose a threatening impact on earth. Here in this study, an attempt was made to elucidate the molecular basis of involvement of *dmp* operon from *P. putida* in phenol degradation.

Domains (functional, highly conservative and interactive regions) of DmpR protein were obtained from Pfam (Jones et al., 1998; George and Heringa, 2002). 3D structures of the four domains of DmpR protein were modeled and analyzed.

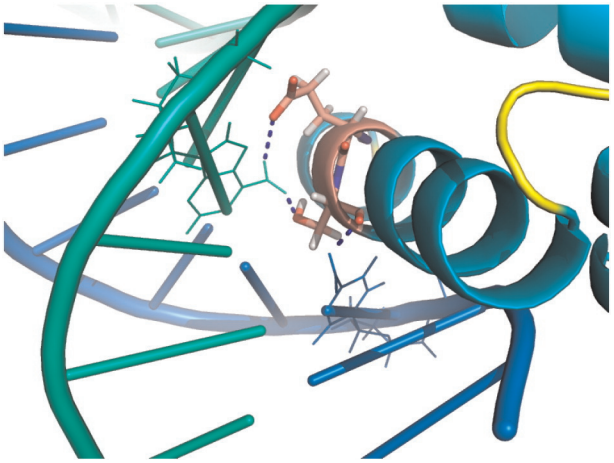

Fig. 3 – Interaction between DNA1 and domain D is shown with the hydrogen bonds indicated with blue dashed lines and the residues involved in interaction are in pink shade. The protein is in cyan shade. The strands of DNA are in blue and green shades and the polar regions are in yellow shade.

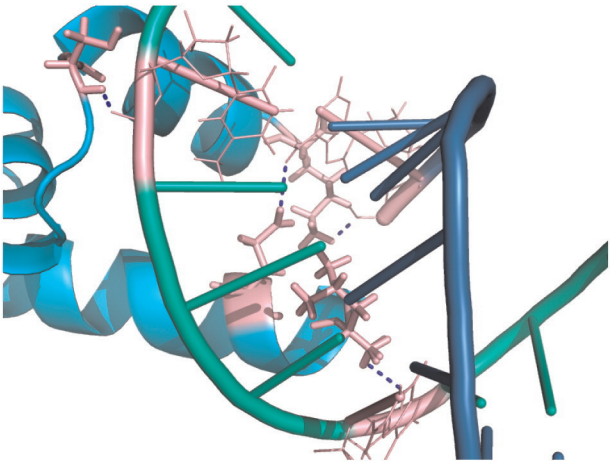

Fig. 4 – Interaction between DNA2 and domain D is shown with the hydrogen bonds indicated with blue dashed lines. The residues involved in the interaction as well as a covalent bond between adenine and Lys41 is in pink shade. The protein is in cyan shade and the strands of DNA in blue and green shades.

Protein–ligand, protein–protein and protein–DNA molecular docking examinations were accomplished. In the presence of inducers, domain A undergoes conformational alterations due to the formation of aromatic π–π stacking interaction of the benzene ring in phenol and 2-chlorophenol with phe-50 of domain A. Documentation documents that among all aromatic amino-acids, phenylalanine is extremely hydrophobic (hydrophobicity index = 100) at pH = 7 and in acidic pH of around 2, its hydrophobicity index is 95 approximately (Berg et.al., 2012). Thereby, this amino acid, phenylalanine plays a dominating role in protein interactions (Berg et al., 2012). Thus, contextually, in this present study, phenylalanine from domain A turns to be very prone to interact with the inducers. PyMOL unveiled the residues responsible for the formation of protein–ligand complexes. After the conformational changes, domain A was efficiently linked to the transcriptional activator-domain C by flexible linker-domain B. P.I.C. estimated the individual amino acid residues, along with their respective positions in the protein domains that were indulged in protein–protein hydrogen bonding like main chain–side chain and side chain–side chain interactions. Furthermore, residues responsible for hydrophobic, ionic, aromatic–sulfur, aromatic–aromatic and cation–π interactions for docked protein–protein complex were also revealed. Accomplishing this protein–protein interaction, domain D (DNA-binding domain with helix–turn–helix motif), then interacts with two DNA structures individually. Residues and nucleotide bases resulting in covalent and non-covalent bond formation were observed in docked DNA-domainD complexes through PyMOL. Interestingly, in case of DNA1-domainD complex, residues of domain D like Leu2, Asn18, Gln19, Asn20, Gly28, Leu39, Ser30 and Lys41 who were responsible for DNA-polar contacts assisted in stabilization of DNA-domain complex besides strengthening the hydrogen bond formation between Cytosine (3rd position) with Ser8 and Adenine (16th position) of the complementary strand with Glu7 of domain D. Even a point mutation in either of the domains or the modeled DNA structures, might alter the sequential and consecutive binding

process, leading to transcriptional inhibition, non-expression of the operon and thereby lack of phenol hydroxylase secretion from *dmpN* gene. This would lead to hamper the biodegradation. Earlier studies manifest contribution of several residues for microbial redox reaction for sulfur oxidation in sox operon (Bagchi, 2011; Bagchi, 2012; Ray and Bagchi, 2013).

Previously, there have been no reports regarding structural biology of these proteins and their molecular level participation for biodegradation for such toxic environmental pollutants. So, results from this study may shed light to discern the importance of individual residues in *dmp* operon for biodegradation of phenol and other toxic derivatives at their molecular level.

4. Conclusion and future scope

Exploration to molecular level contribution in the biodegradation of phenols and other toxic derivatives with the aid of DmpR protein from *P. putida* disclosed the residues and nucleotide bases undergoing covalent and non-covalent bonding. First and foremost, aromatic π–π stacking interaction by phe50 in domain A with inducers causes conformational alterations in the domain. Among the aromatic amino acids, only phenylalanine is extremely hydrophobic (hydrophobicity index = 100) at neutral pH and in acidic pH of around 2, its hydrophobicity index is 95 approximately. Thus, it has a general tendency to interact from proteins. In this context too, it readily formed aromatic π–π stacking interaction from the protein domain A with the inducer present at each time. Further, through several amino acid residues and interactions, domain A was efficiently linked to the transcriptional activator-domain C via flexible linker-domain B. This protein–protein interaction was then followed by the interaction of domain D with the two DNA sequences individually. Few residues of domain D like Leu2, Asn18, Gln19, Asn20, Gly28, Leu39, Ser30 and Lys41 were investigated from this study to be responsible for DNA-polar contacts thereby, assisting in the stabilization as well as the strengthening of the hydrogen bond formation in DNA1–domainD complex.

In this current study, an initial strive for inquest at residue level participation and importance of individual residues of *dmp* operon from *P. putida* provides a rational framework for designing experiments to determine the root variations for initiating transcription in the operon. Thus, it gradually leads to the secretion of phenol hydroxylase enzyme and thereby allowing the biodegradation of such toxic environmental pollutants for rendering a sustainable biota to dwell in.

Acknowledgments

Authors are deeply indebted for the immense help, paramount suggestions and continuous encouragement rendered by Dr. Angshuman Bagchi, Assistant Professor, Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia, India. Without his support and cooperation, it would not have been possible to complete the manuscript in a successful manner. Authors are deeply indebted to DST-PURSE program 2012–2015 going on in Department of Biochemistry and Biophysics, University of Kalyani for providing different equipments and essential infrastructural support. Deep gratitude is extended to DBT sponsored Bioinformatics Infrastructure Facility in the Department of Biochemistry and Biophysics, University of Kalyani for the necessary support. Gratefulness is rendered to the team members of the present research group who have made significant contributions to the work reported here. Finally, the authors are deeply grateful to the anonymous reviewers for their valuable comments as well as suggestions that helped to make the manuscript, a better one.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jes.2015.03.035.

REFERENCES

Anslyn, E.V., Dougherty, D.A., 2005. Modern Physical Organic Chemistry. University Science Books, Sausalito, CA.

Bagchi, A., 2011. Structural modeling of SoxF protein from *Chlorobium tepidum*: an approach to understand the molecular basis of thiosulfate oxidation. Biochem. Biophys. Res. Commun. 414 (2), 409–411.

Bagchi, A., 2012. Structural insight into the mode of interactions of SoxL from *Allochromatium vinosum* in the global sulfur oxidation cycle. Mol. Biol. Rep. 39 (12), 10243–10248.

Bagchi, A., Ghosh, T.C., 2011. Structural analyses of the interactions of SoxY and SoxZ from thermo-neutrophilic *Hydrogenobacter thermophilus*. J. Biophys. Chem. 2 (4), 408–413.

Berg, J., Tymoczko, J., Stryer, L., 2012. Protein composition and structure. Biochemistry 7th ed. W.H. Freeman and Company, New York.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M., 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. 4 (2), 187–217.

Chen, R., Li, L., Weng, Z.P., 2003. ZDOCK: an initial-stage protein docking algorithms. Proteins 51 (1), 80–87.

Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2 (9), 1511–1519.

Comeau, S.R., Gatchel, D.W., Vajda, S., Camacho, C.J., 2004. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 20 (1), 45–50.

DeLano, W.L., 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.

Eisenberg, D., Lüthy, R., Bowie, J.U., 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 277, 396–404.

Fiser, A., Sali, A., 2003. ModLoop: automated modeling of loops in protein structures. Bioinformatics 19 (18), 2500–2501.

George, R.A., Heringa, J., 2002. An analysis of protein domain linkers: their classification and role in protein folding. Protein Eng. 15 (11), 871–879.

Goosen, N., van de Putte, P., 1995. The regulation of transcription initiation by integration host factor. Mol. Microbiol. 16 (1), 1–7.

Gupta, S., Saxena, M., Saini, N., Mahmooduzzafar, K.R., Kumar, A., 2012. An effective strategy for a Whole-Cell biosensor based on

putative effector interaction site of the regulatory DmpR protein. PloS One 7 (8), e43527.
Jones, S., Stewart, M., Michie, A., Swindells, M.B., Orengo, C., Thornton, J.M., 1998. Domain assignment for protein structures using a consensus approach: characterization and analysis. Protein Sci. 7 (2), 233–242.
Källberg, M., Wang, H.P., Wang, S., Peng, J., Wang, Z.Y., Lu, H., et al., 2012. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. 7 (8), 1511–1522.
Kelley, L.A., Sternberg, M.J.E., 2009. Protein structure prediction on the web: a case study using the Phyre server. Nat. Protoc. 4 (3), 363–371.
Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S., Xia, B., Hall, D.R., et al., 2013. How good is automated protein docking? Proteins: Struct., Funct., Bioinf. 81 (12), 2159–2166.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26 (2), 283–291.
Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M.D., Ritchie, D.W., 2010. HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res. 38, W445–W449.
Meyer, E.A., Castellano, R.K., Diederich, F., 2003. Interactions with aromatic rings in chemical and biological recognition. Angew. Chem. Int. Ed. 42 (11), 1210–1250.
Movahedyan, H., Khorsandi, H., Salehi, R., Nikaeen, M., 2009. Detection of phenol degrading bacteria and *Pseudomonas putida* in activated sludge by polymerase chain reaction. Iran J. Environ. Health Sci. Eng. 6 (2), 115–120.
Nordlund, I., Powlowski, J., Shingler, V., 1990. Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. J. Bacteriol. 172 (12), 6826–6833.
Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., et al., 2011. The Pfam protein families database. Nucleic Acids Res. 40 (D1), D290–D301.
Ramachandran, G.N., Sashisekharan, V., 1968. Conformation of polypeptides and proteins. Adv. Protein Chem. 23, 283–438.
Ray, S., Bagchi, A., 2013. Structural analysis of the mode of interactions of sox B protein with sox YZ complex from *Allochromatium vinosum* in the global sulfur oxidation cycle. Comput. Mol. Biol. 3 (1), 1–5.
Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5 (4), 725–738.
Sander, C., Schneider, R., 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins 9 (1), 56–68.
Sarand, I., Skärfstad, E., Forsman, M., Romantschuk, M., Shingler, V., 2001. Role of the DmpR-mediated regulatory circuit in bacterial biodegradation properties in methylphenol-amended soils. Appl. Environ. Microbiol. 67 (1), 162–171.
Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. 33 (S2), W363–W367.
Şeker, Ş., Beyenal, H., Satin, B., Tanyolaç, A., 1997. Multi-substrate growth kinetics of *Pseudomonas putida* for phenol removal. Appl. Microbiol. Biotechnol. 47 (5), 610–614.
Shingler, V., Moore, T., 1994. Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas* sp. strain CF600. J. Bacteriol. 176 (6), 1555–1560.
Shingler, V., Pavel, H., 1995. Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. Mol. Microbiol. 17 (3), 505–513.
Shingler, V., Powlowski, J., Marklund, U., 1992. Nucleotide sequence and functional analysis of the complete phenol/3, 4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. J. Bacteriol. 174 (3), 711–724.
Sivasakthi, V., Anitha, P., Kumar, K.M., Bag, S., Senthilvel, P., Lavanya, P., et al., 2013. Aromatic–aromatic interactions: analysis of π–π interactions in interleukins and TNF proteins. Bioinformation 9 (8), 432–439.
Söding, J., Biegert, A., Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–W248.
Tina, K.G., Bhadra, R., Srinivasan, N., 2007. PIC: protein interactions calculator. Nucleic Acids Res. 35, W473–W476.
Trott, O., Olson, A.J., 2010. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry. 31 (2), 455–461.
Vakser, I.A., 1995. Protein docking for low-resolution structures. Protein Eng. 8 (4), 371–377.
Weber, M., Markus, W., Michael, K.B., 2004. Phenol, Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH, New York.
Xiang, Z.X., 2006. Advances in homology protein structure modeling. Curr. Protein Pept. Sci. 7 (3), 217–227.
Xu, J.R., Zhang, Y., 2010. How significant is a protein structure similarity with TM-score = 0.5? Bioinformatics 26 (7), 889–895.
Xu, D., Zhang, Y., 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophys. J. 101 (10), 2525–2534.
Xu, D., Zhang, Y., 2012. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins 80 (7), 1715–1735.
Zhang, Y., Skolnick, J., 2004. Scoring function for automated assessment of protein structure template quality. Proteins 57 (4), 702–710.
Zheng, G.H., Lu, X.J., Olson, W.K., 2009. Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. Nucleic Acids Res. 37 (S2), W240–W246.