IN SILICO CHARACTERIZATIONS OF DEGRADATIVE ENZYME FROM LANDFILL LEACHATE METAGENOME FOR POTENTIAL POLYCHLORINATED BIPHENYL (PCB) BIOREMEDIATION

SITI MARHAMAH DRAHAMAN [1], NOOR FAIZUL HADRY NORDIN [1*], HAMZAH MOHD SALLEH[2], HUSNA AHMAD TAJUDDIN[3]

[1] *International Institute for Halal Research and Training (INHART), International Islamic University Malaysia, Jalan Gombak, 53100, Selangor Malaysia.*
[2]*Halalan Thayyiban Research Centre, Universiti Islam Sultan Sharif Ali, Kampus Sinaut, Brunei Darussalam.*
[3] *Department of Chemical Engineering and Sustainability Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100, Selangor Malaysia.*

**Corresponding author: faizul@iium.edu.my*

ABSTRACT: The characterization of enzyme structure and function was essential for understanding biochemical pathways and developing effective biotechnological applications, particularly in environmental bioremediation. Traditional experimental methods for protein analysis were often labor-intensive and limited by the inability to culture certain microorganisms. In this study, an *in silico* approach was employed to predict the structure and function of a putative degradative enzyme identified from metagenomic analysis of landfill leachate. Using a combination of bioinformatics tools, including sequence alignment, domain annotation, secondary structure prediction and three-dimensional (3D) structural modeling, the target enzyme was analyzed for its catalytic potential and stability. Conserved motifs and active sites were identified, suggesting its involvement in the degradation of xenobiotic compounds such as polychlorinated biphenyls (PCBs). The 3D structure model revealed a typical fold associated with oxygenase or dehydrogenase activities, with predicted metal-binding sites critical for catalyses. These findings demonstrate the power of computational methods to accelerate the discovery and characterization of novel enzymes, especially from unculturable microbial communities. This approach provides a valuable foundation for future functional validation, protein engineering and the development of environmentally sustainable biocatalysts.

KEY WORDS: *Bioremediation, Enzyme characterization, In silico, Metagenomic*

1. INTRODUCTION

Enzymes play a crucial role in catalyzing biochemical reactions and are central to numerous biological and industrial processes, including the bioremediation of environmental pollutants. One class of such pollutants, PCBs, are synthetic organic compounds known for their chemical stability, environmental persistence and toxicity. Microorganisms capable of degrading PCBs typically produce specific enzymes, such as dioxygenases and dehydrogenases [1], which initiate and drive the degradation pathways.

Understanding the structure and function of these enzymes is essential for elucidating the mechanisms of pollutant breakdown and for enhancing their application in biotechnological and environmental settings.

Traditionally, the structural and functional characterization of proteins requires extensive laboratory-based experiments such as crystallography, NMR spectroscopy and enzymatic assays [2]. However, such methods are time-consuming, costly and often limited by the difficulty in culturing environmental microorganisms. Recent advances in computational biology have introduced *in silico* approaches as powerful tools for predicting protein structure and function directly from sequence data [3]. These methods enable researchers to model the 3D structure of proteins, identify conserved domains, predict active sites and infer potential enzymatic activities with high accuracy.

The *in silico* prediction of enzyme structure and function not only provides insight into protein stability and interactions but also guides experimental designs and the development of engineered enzymes with improved properties [4]. In the context of biodegradation, these predictions are particularly valuable in identifying novel enzymes from metagenomic data, especially from unculturable microbial communities in complex environments like landfills.

This study focuses on the *in silico* characterization of putative degradative enzymes identified from functional metagenomic analysis of landfill leachate. By using advanced bioinformatics tools, this study aims to predict the structure, active sites and functional domains of these enzymes, shedding light on their potential role in PCB degradation and their suitability for environmental bioremediation applications.

2. METHODOLOGY

2.1. Selection of the functional gene

From the functional annotation result generated from shotgun metagenomic sequencing, there were a total of 5382 functional genes annotated in the KEGG pathway. Among all, the gene numbers involved in the xenobiotics biodegradation and metabolism were selected to study PCB degradation. There were 80 functional genes involved in xenobiotic biodegradation and metabolism. From the listed genes (Table 1), gene number J2_444 was selected because it was an uncharacterized protein. Furthermore, this gene was involved in chlorobenzene degradation, benzoate degradation, fluorobenzene degradation, and toluene degradation (highlighted in black color). These degradations were involved in the PCB degradation pathway.

Table 1: Functional genes annotated from KEGG pathway that involved xenobiotics biodegradation and metabolism

KO	PATHWAY LEVEL 3	GENE NUM/PDB
ko00361	**Metabolism; Xenobiotics biodegradation and metabolism; Chlorocyclohexane and chlorobenzene degradation**	**1. J2_444** 2. J2_304
ko00362	**Metabolism; Xenobiotics biodegradation and metabolism; Benzoate degradation**	1. JB2_179 2. J2_2661 3. J2_1158 4. J2_789 5. J2_1679 6. J2_454 7. J2_281

		8. J2_2738 **9. J2_444** 10. JB2_1060 11. NOVO_MIX_368 12. JB2_943 13. NOVO_MIX_1492 14. NOVO_MIX_742 15. J2_1652 16. JB2_942 17. NOVO_MIX_587 18. J2_2531 19. J2_304 20. J2_1676 21. JB2_425 22. NOVO_MIX_509 23. NOVO_MIX_812 24. NOVO_MIX_241
ko00364	**Metabolism; Xenobiotics biodegradation and metabolism; Fluorobenzoate degradation**	**1. J2_444** 2. J2_304
ko00365	Metabolism; Xenobiotics biodegradation and metabolism; Furfural degradation	1. NOVO_MIX_1481 2. NOVO_MIX_1350
ko00621	Metabolism; Xenobiotics biodegradation and metabolism; Dioxin degradation	1. J2_1652 2. NOVO_MIX_366
ko00622	Metabolism; Xenobiotics biodegradation and metabolism; Xylene degradation	1. NOVO_MIX_368 2. J2_1652
ko00623	**Metabolism; Xenobiotics biodegradation and metabolism; Toluene degradation**	**1. J2_444** 2. J2_304
ko00625	Metabolism; Xenobiotics biodegradation and metabolism; Chloroalkane and chloroalkene degradation	1. NOVO_MIX_530 2. NOVO_MIX_544 3. J2_253 4. NOVO_MIX_1504 5. J2_1424 6. J2_832 7. J2_2338
ko00626	Metabolism; Xenobiotics biodegradation and metabolism; Naphthalene degradation	1. NOVO_MIX_544 2. J2_2338
ko00627	Metabolism; Xenobiotics biodegradation and metabolism; Aminobenzoate degradation	1. J2_2661 2. J2_75 3. JB2_1060 4. J2_1105
ko00633	Metabolism; Xenobiotics biodegradation and metabolism; Nitrotoluene degradation	1. JB2_176 2. J2_2995
ko00642	Metabolism; Xenobiotics biodegradation and metabolism; Ethylbenzene degradation	1. J2_170
ko00791	Metabolism; Xenobiotics biodegradation and metabolism; Atrazine degradation	1. J2_1736 2. J2_355
ko00930	Metabolism; Xenobiotics biodegradation and metabolism; Caprolactam degradation	1. JB2_1060 2. J2_2661 3. NOVO_MIX_587 4. J2_1866

ko00980	Metabolism; Xenobiotics biodegradation and metabolism; Metabolism of xenobiotics by cytochrome P450	1. JB2_1013 2. NOVO_MIX_544 3. J2_1962 4. J2_2338 5. NOVO_MIX-1387 6. JB2_620
ko00982	Metabolism; Xenobiotics biodegradation and metabolism; Drug metabolism - cytochrome P450	1. JB2_620 2. NOVO_MIX_1387 3. J2_2338 4. J2_1962 5. NOVO_MIX_544 6. JB2_1013
ko00983	Metabolism; Xenobiotics biodegradation and metabolism; Drug metabolism - other enzymes	1. J2_1271 2. J2_918 3. J2_1269 4. J2_2509 5. NOVO_MIX_848
ko00984	Metabolism; Xenobiotics biodegradation and metabolism; Steroid degradation	1. JB2_908 2. NOVO_MIX_866 3. J2_382

2.2. Bioinformatics analysis on the structural and functional

The potential domain or family of a query protein was predicted using the Pfam (https://pfam.xfam.org/) server. Domain and family were able to provide information about the query protein's potential function or interaction. The PSORTb server (http://www.psort.org/psortb/index.html) was utilized to forecast the query protein's subcellular location. Likewise, transmembrane helices and signal peptides were detected using the SignalP server, respectively. Knowing this information was crucial for classifying proteins as cytoplasmic, secretory, or membrane proteins. By forecasting its secondary structure, the query protein's initial structural annotation was established. The number of potential helices, strands, and loops that could have shaped the query protein is revealed by the prediction. The PDBsum server (https://omictools.com/pdbsum-tool) was used for this stage.

2.3. *In silico* structure building

The PDB was searched for homologous protein structures using the BLASTp utility software. The SWISS-MODEL workspace (https://swissmodel.expasy.org/) was used to create the template and model of the chosen protein.The Ramachandran plot from MolProbity version 4.4 was used to verify the model's dependability and for any steric conflicts. PyMOL v2.0.6 (https://pymol.org/2/) was used to view the modeled structure in order to identify the active sites that the COACH server (https://zhanggroup.org/COACH/) predicted and to determine whether the amino acid outliers were on the surface or buried in the structure. The STRING service (https://string-db.org/) was used to access the interactions between the surrounding proteins and the query protein. This makes it possible to determine the regulatory and functional relationships between proteins.

3. RESULT AND DISCUSSION

3.1. Protein characterization of J2_444

From the similarity search using Protein Databank (RCSB PDB), this uncharacterized protein with a sequence length of 331 has a molecular weight of 36.8 kDa. This uncharacterized protein was crucial as the biological function, structural features and physiological role remain unknown for this predicted protein. The molecular weight could be considered ideal because proteins with low molecular weight produce high copies of protein in the translational mechanism [5], as the rate of translation would affect the protein expression [6]. This was because proteins with molecular weights above 60 kDa resulted in a low probability of soluble protein expression [7]. The gene number of J2_444 was significant for the PCB degradation because from the Pfam server, it has a conserved domain of intradiol ring-cleavage dioxygenase at position 31-197 (Figure 1), which involved in PCB breakdown pathway. An essential enzyme in the aerobic breakdown of PCB and other aromatic chemicals was intradiol ring-cleavage dioxygenase. It catalyzed the ortho-cleavage of catechol or its derivatives by incorporating molecular oxygen into the substrate [8]. This reaction resulted in the breaking of the aromatic ring, forming intermediates that were further metabolized into simpler compounds. Other than that, intradiol dioxygenase prevents the accumulation of toxic chlorocatechol by breaking them into less harmful and metabolizable compounds. It also ensures the transformation of complex aromatic intermediates into compounds that can be fully degraded into CO_2 and H_2O, completing the biodegradation process. Intradiol ring-cleavage dioxygenase was crucial for the aerobic degradation of PCBs. It enabled the breakdown of aromatic and chlorinated intermediates, facilitating their complete mineralization. This not only detoxified harmful PCB metabolites but also play a central role in transforming these compounds into usable energy sources for the degrading microbes.

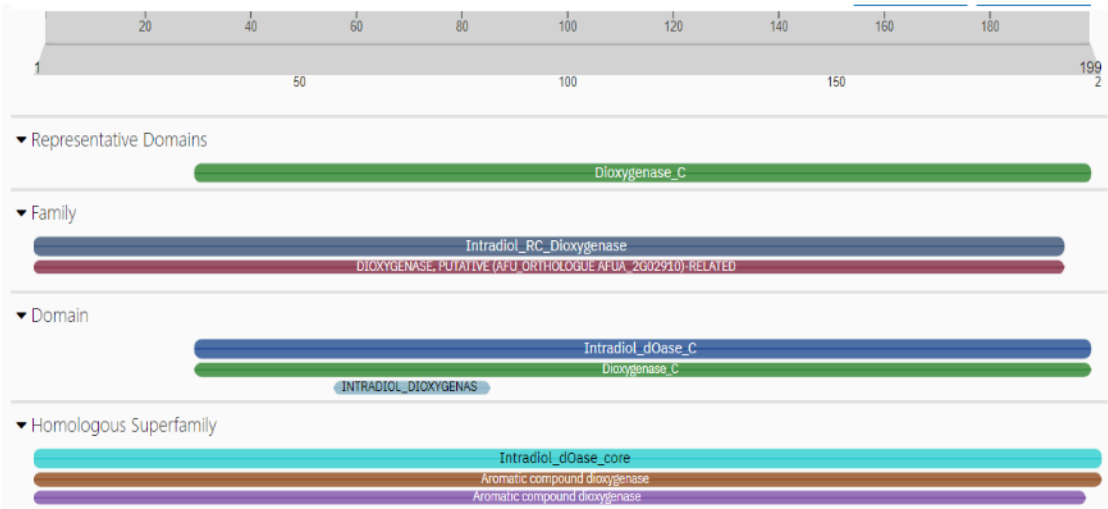


Fig. 1. Conserved domain of J2_444

3.2. Physicochemical properties of J2_444

The most prevalent amino acids in J2_444 were aspartic acid and glycine (10.1%), followed by alanine, valine, and leucine (8.5%, 8.5%, and 7.5%, respectively), according to the ProtParam parameter from Expasy. Because there were more negatively charged residues (aspartic acid and glutamic acids) than positively charged residues (arginine and

lysine), the physicochemical characteristics predicted that the protein would be negatively charged. Since arginine and lysine were positively charged and aspartate and glutamate were negatively charged at neutral pH, these four amino acids account for the majority of the total charge of protein. Negatively charged proteins have been explored for their potential in the bioremediation of PCB. Proteins with negatively charged surfaces can interact with hydrophobic contaminants like PCBs, potentially facilitating degradation or sequestration [9]. For example, a study explained that *Dehalococcoides mccartyi* expressed reductive dehalogenases specific to chlorinated compounds, with enzymes often functioning with negatively charged cofactors, enabling interaction with target compounds [10]. The protein's hydrophobicity, which has a major effect on its binding affinity to cells and receptors, was also influenced by the high aliphatic values [11]. Aliphatic amino acids contribute to the resistance of proteins in harsh environments, such as those encountered during the breakdown of hydrophobic pollutants [12]. This protein can be categorized as a stable protein as it was estimated to have a long half-life and low instability index (Table 2). This was crucial in bioremediation because it ensures consistent enzymatic activity under the diverse and often harsh environmental conditions encountered during pollutant degradation.

Table 2: Physicochemical properties of J2_444

Properties	**Value**
Number of amino acid	199
Mass (Da)	21949.74
Isoelectric point (pI)	4.84
Atomic composition	Carbon C 985 Hydrogen H 1503 Nitrogen N 267 Oxygen O 292 Sulfur S 6
Formula	$C_{985}H_{1503}N_{267}O_{292}S_6$
Estimated half-life (h)	5.5 hours (mammalian reticulocytes, in vitro). 3 min (yeast, in vivo). 2 min (*Escherichia coli*, in vivo)
Instability index (II)	33.65
Aliphatic index	82.31

3.3. Prediction of secondary structure

In order to predict the protein localization site, the protein sequence was submitted to the TargetP server. The result showed that it was categorized as a mitochondrial protein with a localization score of 0.0047 and a signal peptide score of 0.1125. As it was not a

membrane-bound protein, it was much easier to be purified in a soluble form as what had been observed from the purification experiments performed in this study. This type of protein was involved in various functions such as ATP production through oxidative phosphorylation, the regulation of metabolic pathways, apoptosis, and calcium homeostasis [13]. Among mitochondrial proteins that were involved in biodegradation were cytochrome P450 enzymes, which metabolized a range of xenobiotics, including organic pollutants [14], and ATP synthase enzymes that supported energy-intensive pathways, including active transport and bioremediation enzyme synthesis [15]. The secondary structure predicted by the PDBsum server showed that this protein consists of 7 α-helices, labelled as H1 to H7 (purple springs), and 16 β-strands (pink arrows) (Figure 2). The α-helix was a favored structure that removed the peptide backbone from hydrogen bonding to the solvent [16], and β-strands provided a stable scaffold for the active sites of enzymes or binding sites for ligands [17]. The presence of conserved secondary structures such as beta-sheets near loops suggested potential active sites for catalytic functions [18] in pollutant breakdown. The combination of helices and sheets indicated structural stability, essential for proteins functioning in harsh bioremediation environments under varying pH, temperature, or salinity [19]. Other than that, beta-sheets and loops could indicate which regions are involved in binding hydrophobic pollutants or cofactors required for degradation processes.

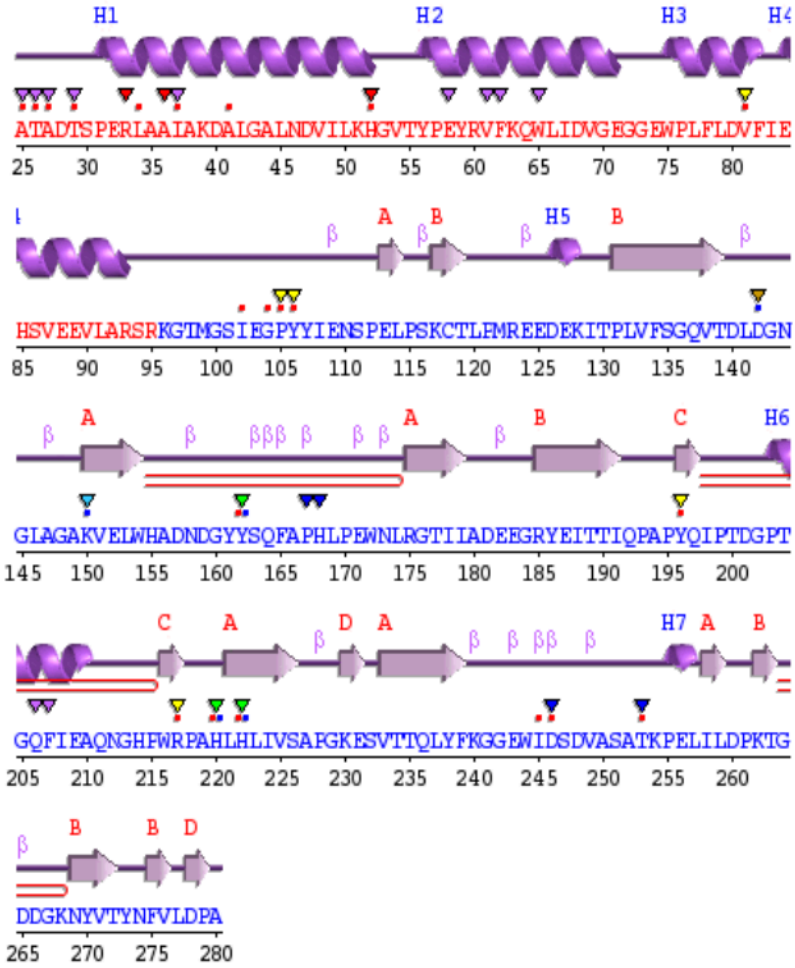


Fig. 2. Secondary structure prediction of J2_444 generated from PDBsum server

The topology diagram generated from the PDBsum server also showed red cylinders representing structured and stable alpha-helical regions (30–53, 55–72, 74–83, and 84–94),

making up the continuous domain, indicative of simpler and straightforward structures (Figure 3). Continuous sheets are more stable [20] and may fold quickly because of the direct connection between strands [21]. The red cylinders represent alpha-helical transmembrane segments. These regions likely anchor the protein in the membrane [22], which allows interaction with hydrophobic pollutants or mediates the transport of degradation products.

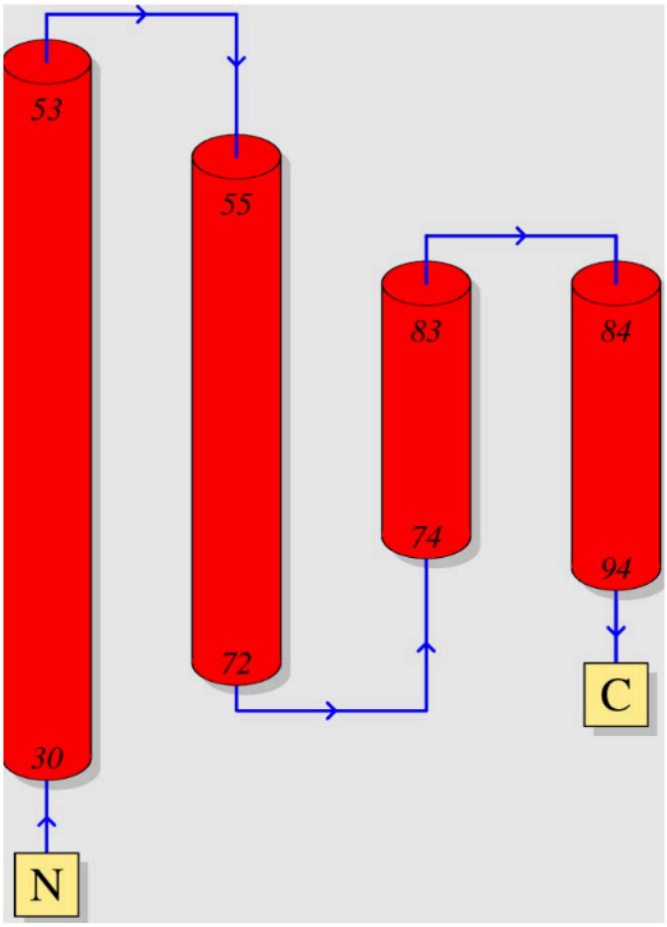
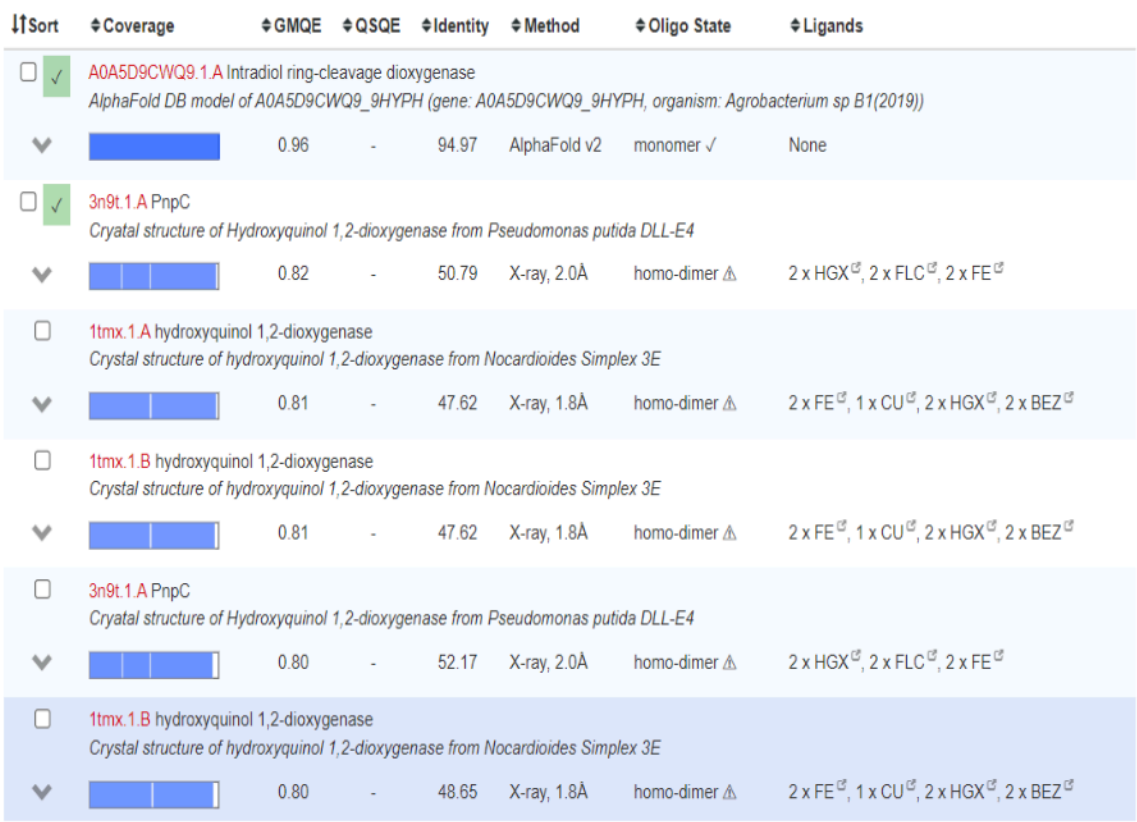


Fig. 3. Topology from secondary structure of J2_444 generated from PBDsum server

3.4. Prediction of tertiary structure

It was crucial to consider a protein's secondary and tertiary structure when annotating its functional properties. The amino acid sequence of a query protein can be used to predict its potential shape or folding (helices, strands, and loops), which may reveal crucial details about the active sites and ligand binding. From the SWISS-MODEL server, the top five templates used for modeling were listed (Figure 4).



Sort	Coverage	GMQE	QSQE	Identity	Method	Oligo State	Ligands
✓	A0A5D9CWQ9.1.A Intradiol ring-cleavage dioxygenase *AlphaFold DB model of A0A5D9CWQ9_9HYPH (gene: A0A5D9CWQ9_9HYPH, organism: Agrobacterium sp B1(2019))*	0.96	-	94.97	AlphaFold v2	monomer ✓	None
✓	3n9t.1.A PnpC *Cryatal structure of Hydroxyquinol 1,2-dioxygenase from Pseudomonas putida DLL-E4*	0.82	-	50.79	X-ray, 2.0Å	homo-dimer ⚠	2 x HGX, 2 x FLC, 2 x FE
	1tmx.1.A hydroxyquinol 1,2-dioxygenase *Crystal structure of hydroxyquinol 1,2-dioxygenase from Nocardioides Simplex 3E*	0.81	-	47.62	X-ray, 1.8Å	homo-dimer ⚠	2 x FE, 1 x CU, 2 x HGX, 2 x BEZ
	1tmx.1.B hydroxyquinol 1,2-dioxygenase *Crystal structure of hydroxyquinol 1,2-dioxygenase from Nocardioides Simplex 3E*	0.81	-	47.62	X-ray, 1.8Å	homo-dimer ⚠	2 x FE, 1 x CU, 2 x HGX, 2 x BEZ
	3n9t.1.A PnpC *Cryatal structure of Hydroxyquinol 1,2-dioxygenase from Pseudomonas putida DLL-E4*	0.80	-	52.17	X-ray, 2.0Å	homo-dimer ⚠	2 x HGX, 2 x FLC, 2 x FE
	1tmx.1.B hydroxyquinol 1,2-dioxygenase *Crystal structure of hydroxyquinol 1,2-dioxygenase from Nocardioides Simplex 3E*	0.80	-	48.65	X-ray, 1.8Å	homo-dimer ⚠	2 x FE, 1 x CU, 2 x HGX, 2 x BEZ

Fig. 4. Top 10 templates used by SWISS-MODEL for modeling J2_444 protein

The top template with a 95% identity sequence to intradiol ring-cleavage dioxygenase was chosen for modeling since this was one of the enzymes that was responsible for degrading PCB. The template showed the highest Global Model Quality Estimate (GMQE), 0.96, that described the model quality from the target-template alignment and the template structure. A high GMQE score suggested that the model was likely to have a structure close to the actual protein and can be used with more confidence for downstream analyses, such as studying protein functions or interactions in bioremediation applications. The template was searched using AlphaFold V2 since it has high sequence identity. This monomeric protein can be functional on its own, without needing to interact with other protein molecules. PCB was a persistent environmental pollutant that required enzymatic degradation for effective bioremediation. Certain monomeric proteins were highly efficient in degrading PCB due to their catalytic properties. For example, monomeric forms were biphenyl dioxygenases [23] that can improve substrate specificity for PCB degradation with fewer chlorines, laccases [24] that utilized oxygen as an electron acceptor to degrade phenolic and non-phenolic PCB derivatives, and hydrolases such as dehalogenases [25], which help remove chlorine atoms from PCB derivatives. The generated model was found to have a single domain, and the folding was simple with helices and strands (Figure 5). The overall 3D fold showed a mix of alpha-helices and beta-sheets, connected by loops, indicating a common protein fold that is possibly related to its biological function. This structural organization might play a role in substrate binding, catalytic activity, or structural stability, which was significant in proteins used for bioremediation. If this protein was involved in bioremediation, the active or binding site would typically be in a pocket or groove, often surrounded by conserved residues. The presence of both hydrophobic cores

and surface-exposed residues indicated potential stability under varying environmental conditions, which was crucial for bioremediation applications.

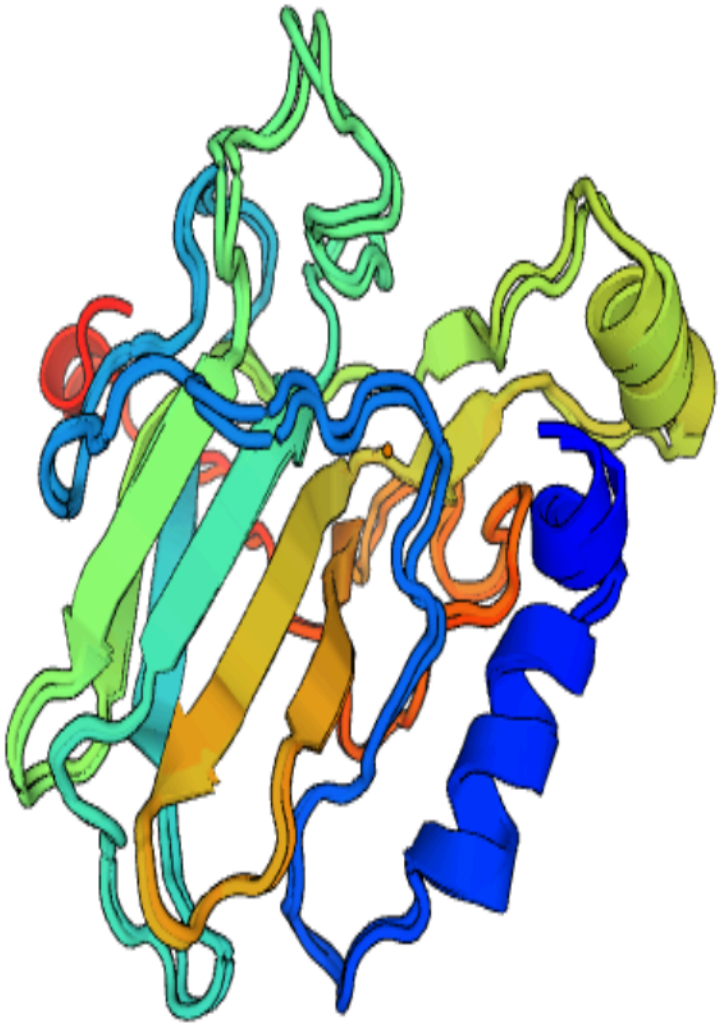


Fig. 5. Predicted three-dimensional structure model of J2_444, with N-terminal coloured blue and C-terminal coloured red; Red helices represent alpha-helices; Blue and green strands represent beta-sheets; Loops and turns represented by irregular connecting regions

The Ramachandran plot was used to structurally validate the created model. The Ramachandran plot produced the plots by calculating the phi/psi angles (ϕ,ψ) between the residues' N-Cα and Cα-C atoms. The amino acid residue-derived phi/psi plots were separated into "favored," "allowed," and "outlier" regions. Generally, a score greater than 98% indicates that the models were of high stereochemical quality. The protein model was deemed adequate for use *in silico* research if over 88% of the amino acid residues were in the preferred or permitted region [26]. The percentage of residues in the favored region was 98.5%, with no outliers' present (Figure 6). The clustering of residues in the alpha-helical

and beta-sheet regions confirmed the presence of secondary structural elements such as helices and sheets. Most residues in the plot fall into these regions, which correspond to the alpha-helix region (around ϕ=−60∘\phi = -60^\circϕ=−60∘ and ψ=−40∘\psi = -40^\circψ=−40∘) and the beta-sheet region (around ϕ=−120∘\phi = -120^\circϕ=−120∘ and ψ=120∘\psi = 120^\circψ=120∘). Proteins involved in bioremediation required stable structures for efficient substrate binding and catalysis. Good distribution in the favored regions supports the structural integrity needed for function. Any significant number of residues in disallowed regions could indicate structural issues or flexibility at specific sites, which might relate to dynamic interactions [27] during catalysis. Based on the Ramachandran plot, if most residues were in the favored regions, the modeled protein was of good quality, supporting its potential utility in applications [28], such as bioremediation of PCB.

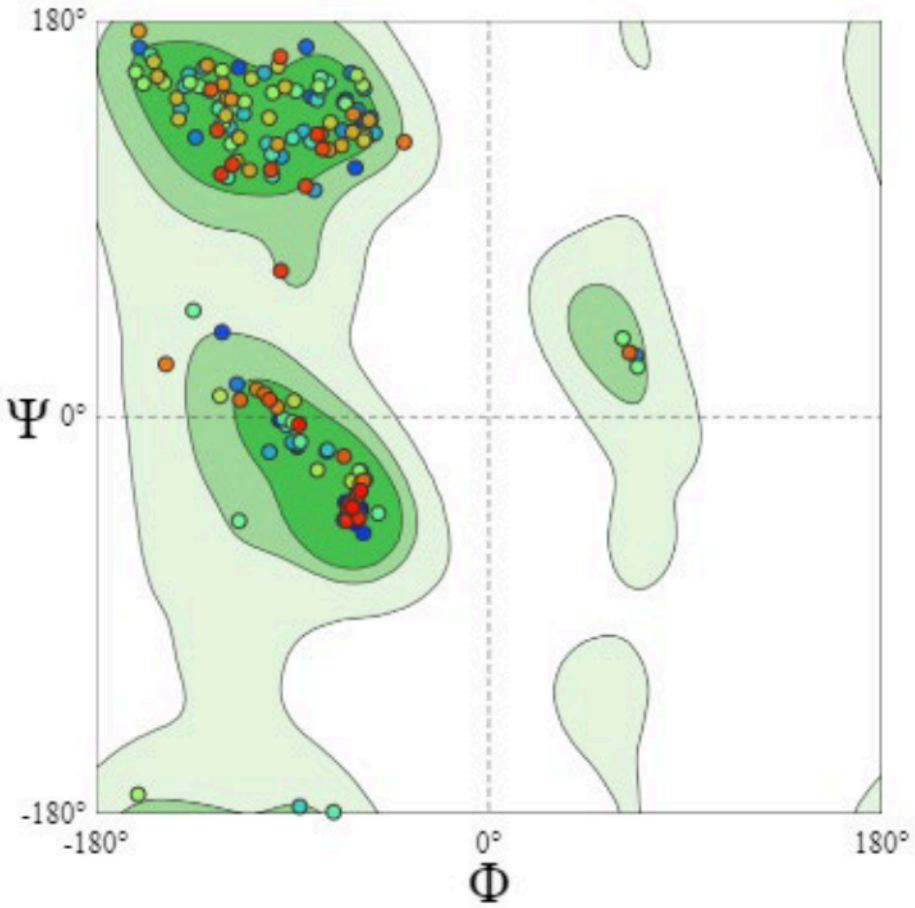


Fig. 6. Model validation studies of J2_444 by Ramachandran's plot using Molprobity in SWISS-MODEL server; Dark and light green were favored regions, pale green was allowed regions, white region was disallowed region

3.5. Interaction of protein-protein

Protein association was examined using the database STRING (Search Tool for the Retrieval of Interacting Genes) [29]. From the STRING analysis, there were interactions from curated databases and from experimentally determined interactions of some proteins with J2_444 (Figure 7). From the analysis, 6 nodes were observed (KJF73691, the red node as the central node) (Figure 7) with predicted interactions to 10 edges. The network cluster was described as dioxygenase and amidohydrolase with scores ranging from 0.99, 0.95, and 0.93 for the first three listed proteins: KJF73692, yellow node (maleylacetate reductase); KJF73693, leafy green node (amidohydrolase); and minty green node KJF71221 (cycloisomerase). It was worth noting that the joint protein shared a function of aromatic compound degradation, which primarily involved chlorocylohexane and chlorobenzene degradation with scores of 2.21. These degradations were closely related to PCB degradation [30] as they involved the breakdown of chlorinated organic compounds into simpler and less toxic intermediates. These similarities lie in their biochemical mechanisms, involving enzymes and ecological relevance. After dehalogenation, hydroxylation, and ring cleavage, the degradation intermediates were funneled into the TCA cycle for energy production and biomass synthesis [31]. The final degradation steps were similar for all three compounds, despite differences in their initial structures. Identifying such shared proteins can enhance the understanding of microbial metabolic versatility and help in engineering microorganisms for bioremediation. For example, bacteria expressing this joint protein could be deployed in sites contaminated with a mix of chlorinated aliphatic (e.g., chlorocyclohexane) and aromatic hydrocarbons (e.g., chlorobenzene). Additionally, shared proteins with broad functionality can be targeted in genetic engineering efforts to create more efficient biodegrading strains. The discovery of a shared protein with a strong function in aromatic compound degradation highlighted the connection of chlorocyclohexane and chlorobenzene degradation pathways. This protein likely played a vital role in dechlorination, hydroxylation, or ring cleavage, which were core steps in breaking down chlorinated compounds [32]. Its functional versatility underscored its importance in microbial metabolism and its potential application in environmental cleanup strategies.

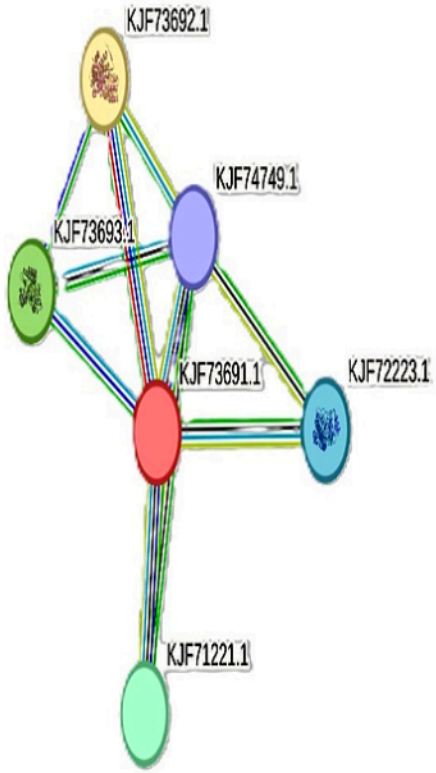


Figure 7. Protein Network analysis for J2_444 protein. The various line colors indicate the different types of evidence for association: the yellow line indicates text-mining evidence, the black line indicates co-expression evidence, the blue line indicates co-occurrence evidence, the purple line indicates experimental evidence, the green line indicates neighborhood evidence, the red line indicates fusion evidence and the light blue line indicates database evidence

3.6. Active site prediction

After validation, the modeled structure of J2_444 was fed into the COACH meta-server to generate ligand binding site predictions. The COACH server developed a correlative ligand binding site prediction by employing two related approaches, TM-site and S-site, which identified ligand binding configurations from the BioLiP protein function database through binding substructures. The highest anticipated ligand was iron (III), Fe^{3+}, with a C-score of 0.83. C-score predictions range from 0 to 1, with higher scores indicating more trustworthy predictions. The active site of J2_444, as predicted by the COACH server, was illustrated (Figure 8). The green color sphere indicated the active site predicted, which was located in a groove among the β-sheet domain at the center with consensus residues at TYR85, TYR120, ARG141, HIS144, and HIS146 that were labeled in purple. These residues (tyrosine, arginine, histidine) were predicted to be the binding site of the Fe^{3+} ligand for the structure that plays complementary roles in binding Fe^{3+} within that protein structure, particularly in systems involved in catalytic pathways. Ideally, Fe^{3+} serves as the catalytic center, facilitating redox reactions that cleave the PCB ring structure [33]. Histidine residues

may assist in activating Fe^{3+} for redox reactions [34], facilitating electron transfer needed to cleave PCB molecules or other aromatic compounds. Tyrosine residues can participate in electron transfer or stabilization of intermediate radical species [35] during PCB degradation. This may help with the breakdown of complex aromatic structures. Other than that, arginine may help position substrates or stabilize the active site geometry [36], enabling efficient interaction between Fe^{3+} and PCB molecules. It can also neutralize negatively charged reaction intermediates, ensuring proper progression of the catalytic cycle.

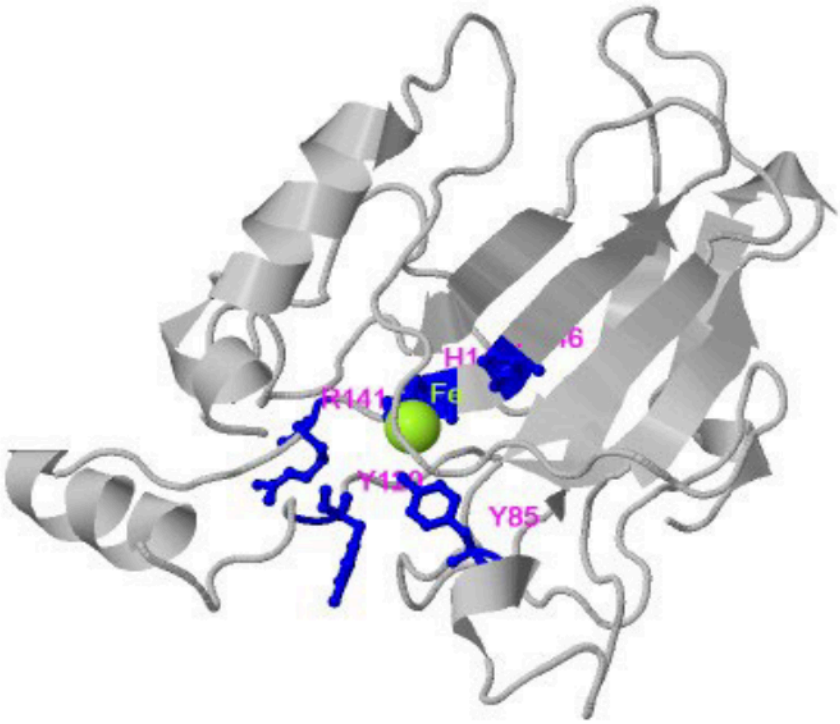


Fig. 8. Active sites of the anticipated three-dimensional structure of the target protein as identified by the COACH server.

4. CONCLUSION

This study successfully demonstrated the application of *in silico* tools to predict the structure and function of a computational degradative enzyme identified from metagenomic data derived from landfill leachate. The bioinformatics analyses revealed conserved functional domains and potential active sites that indicate the enzyme's role in the biodegradation of persistent pollutants such as polychlorinated biphenyls (PCBs). Structural

modeling suggested a stable 3D conformation with functional motifs commonly found in dioxygenases or dehydrogenases, supporting its potential catalytic activity. These findings highlight the efficiency and value of computational methods in accelerating enzyme discovery, particularly from unculturable microbial communities in complex environments. Furthermore, the study provides a strong foundation for future experimental validation, protein engineering, and the development of enzyme-based bioremediation strategies. Ultimately, *in-silico* characterization offers a cost-effective and rapid approach for exploring novel enzymes with significant environmental and industrial applications.

ACKNOWLEDGEMENT

We would like to acknowledge Ministry of Higher Education (MOHE), Malaysia, for funding this project under the Fundamental Research Grant Scheme (FRGS/1/2023/STG04/UIAM/02/3) which made this research possible.

REFERENCES

[1] Khalid, F., Hashmi, M. Z., Jamil, N., Qadir, A., & Ali, M. I. (2021). Microbial and enzymatic degradation of PCBs from e-waste-contaminated sites: a review. Environmental Science and Pollution Research, 28, 10474-10487.

[2] Kovermann, M., Rogne, P., & Wolf-Watz, M. (2016). Protein dynamics and function from solution state NMR spectroscopy. Quarterly reviews of biophysics, 49, e6.

[3] Bhat, G. R., Sethi, I., Rah, B., Kumar, R., & Afroze, D. (2022). Innovative in silico approaches for characterization of genes and proteins. Frontiers in Genetics, 13, 865182.

[4] Zhou, L., Tao, C., Shen, X., Sun, X., Wang, J., & Yuan, Q. (2024). Unlocking the potential of enzyme engineering via rational computational design strategies. Biotechnology Advances, 108376.

[5] Lu, P., Vogel, C., Wang, R., Yao, X., & Marcotte, E. M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nature biotechnology, 25(1), 117-124.

[6] Wang, K., Zhang, S., Weber, J., Baxter, D., & Galas, D. J. (2010). Export of microRNAs and microRNA-protective protein by mammalian cells. Nucleic acids research, 38(20), 7248-7259.

[7] Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in E. coli. Current protocols in protein science, 61(1), 5-24.

[8] Semana, P., & Powlowski, J. (2019). Four aromatic intradiol ring cleavage dioxygenases from Aspergillus niger. Applied and environmental microbiology, 85(23), e01786-19.

[9] Zhang, M., Tao, S., & Wang, X. (2020). Interactions between organic pollutants and carbon nanomaterials and the associated impact on microbial availability and degradation in soil: a review. Environmental Science: Nano, 7(9), 2486-2508.

[10] Islam, M. A., Waller, A. S., Hug, L. A., Provart, N. J., Edwards, E. A., & Mahadevan, R. (2014). New insights into Dehalococcoides mccartyi metabolism from a reconstructed metabolic network-based systems-level analysis of D. mccartyi transcriptomes. PLoS One, 9(4), e94808.

[11] Aruleba, R. T., Adekiya, T. A., Oyinloye, B. E., & Kappo, A. P. (2018). Structural studies of predicted ligand binding sites and molecular docking analysis of Slc2a4 as a therapeutic target for the treatment of cancer. International journal of molecular sciences, 19(2), 386.

[12] Kumari, S., & Das, S. (2023). Bacterial enzymatic degradation of recalcitrant organic pollutants: catabolic pathways and genetic regulations. Environmental Science and Pollution Research, 30(33), 79676-79705.

[13] Hüttemann, M., Lee, I., Samavati, L., Yu, H., & Doan, J. W. (2007). Regulation of mitochondrial oxidative phosphorylation through cell signaling. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1773(12), 1701-1720.

[14] Esteves, F., Rueff, J., & Kranendonk, M. (2021). The central role of cytochrome P450 in xenobiotic metabolism - a brief review on a fascinating enzyme family. Journal of xenobiotics, 11(3), 94-114.

[15] Roy, S., Schievano, A., & Pant, D. (2016). Electro-stimulated microbial factory for value added product synthesis. Bioresource technology, 213, 129-139.

[16] Pelton, J. T., & McLean, L. R. (2000). Spectroscopic methods for analysis of protein secondary structure. Analytical biochemistry, 277(2), 167-176

[17] Tyndall, J. D., Nall, T., & Fairlie, D. P. (2005). Proteases universally recognize beta strands in their active sites. Chemical Reviews, 105(3), 973-1000

[18] Giovanella, P., Vieira, G. A., Otero, I. V. R., Pellizzer, E. P., de Jesus Fontes, B., & Sette, L. D. (2020). Metal and organic pollutants bioremediation by extremophile microorganisms. Journal of hazardous materials, 382, 121024.

[19] Chakraborty, J., & Das, S. (2017). Application of spectroscopic techniques for monitoring microbial diversity and bioremediation. Applied Spectroscopy Reviews, 52(1), 1-38.

[20] Uversky, V. N., & Fink, A. L. (2004). Conformational constraints for amyloid fibrillation: the importance of being unfolded. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1698(2), 131-153.

[21] Jenkins, J., & Pickersgill, R. (2001). The architecture of parallel β-helices and related folds. Progress in biophysics and molecular biology, 77(2), 111-175.

[22] Nicolay, T., Vanderleyden, J., & Spaepen, S. (2015). Autotransporter-based cell surface display in Gram-negative bacteria. Critical reviews in microbiology, 41(1), 109-123.

[23] Furukawa, K., & Fujihara, H. (2008). Microbial degradation of polychlorinated biphenyls: biochemical and molecular features. Journal of Bioscience and Bioengineering, 105(5), 433-449.

[24] Lopes, J. M., Marques-da-Silva, D., Videira, P. Q., & Lagoa, R. L. (2022). Comparison of laccases and hemeproteins systems in bioremediation of organic pollutants. Current Protein and Peptide Science, 23(6), 402-423.

[25] Horsman, G. (2008). Characterization of BphD, a CC bond hydrolase involved in the degradation of polychlorinated biphenyls (Doctoral dissertation, University of British Columbia).

[26] Singh, S., Singh, H., Tuknait, A., Chaudhary, K., Singh, B., Kumaran, S., & Raghava, G. P. (2015). PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. Biology direct, 10, 1-19.

[27] Ravikumar, A., de Brevern, A. G., & Srinivasan, N. (2021). Conformational strain indicated by ramachandran angles for the protein backbone is only weakly related to the flexibility. The Journal of Physical Chemistry B, 125(10), 2597-2606.

[28] Gunasekaran, K., & Nussinov, R. (2007). How different are structurally flexible and rigid binding sites? Sequence and structural features discriminating proteins that do and do not undergo conformational change upon ligand binding. Journal of molecular biology, 365(1), 257-273.

[29] Von Mering, C., Jensen, L. J., Snel, B., Hooper, S. D., Krupp, M., Foglierini, M., & Bork, P. (2005). STRING: known and predicted protein–protein associations, integrated and transferred across organisms. Nucleic acids research, 33(suppl_1), D433-D437.

[30] Zhou, H., Yin, H., Guo, Z., Zhu, M., Qi, X., & Dang, Z. (2023). Methanol promotes the biodegradation of 2, 2′, 3, 4, 4′, 5, 5′-heptachlorobiphenyl (PCB 180) by the microbial consortium QY2: metabolic pathways, toxicity evaluation and community response. Chemosphere, 322, 138206.

[31] Pimviriyakul, P., Wongnate, T., Tinikul, R., & Chaiyen, P. (2020). Microbial degradation of halogenated aromatics: molecular mechanisms and enzymatic reactions. Microbial Biotechnology, 13(1), 67-86.

[32] Li, X., Chen, Y., Chen, Z., Guo, H., Yang, S., & Ma, X. (2022). The recent progress on gaseous chlorinated aromatics removal for environmental applications. Separation and Purification Technology, 296, 121364.

[33] Alvarez, P. J., & Illman, W. A. (2005). Biodegradation principles. Bioremediation and Natural Attenuation, 49-114.

[34] Miller, T. J. (2021). Small molecule activation and reduction using non-heme iron complexes (Doctoral dissertation, University of Illinois at Urbana-Champaign).

[35] Holliday, G. L., Mitchell, J. B., & Thornton, J. M. (2009). Understanding the functional roles of amino acid residues in enzyme catalysis. Journal of Molecular Biology, 390(3), 560-577.

[36] Nagagarajan, S., Xue, F., & MacKerell Jr, A. D. (2013). Impact of substrate protonation and tautomerization states on interactions with the active site of arginase I. Journal of chemical information and modeling, 53(2), 452-460.