Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 7, 2025

IN-SILICO CHARACTERIZATION OF TOLUENE DIOXYGENASE FROM TOLUENE DEGRADING *PSEUDOMONAS AERUGINOSA*

Nadia Hussain[1], Saboor Muarij Bunny[2], Amal H. I. Al Haddad[3], Riffat Iqbal[4], Fatima Muccee[*5], Aleena Ahmad Khan[6]

[1]*Department of Pharmaceutical Sciences, College of Pharmacy, Al Ain University, Al Ain Campus, 64141 Al Ain, United Arab Emirates.*
[1]*AAU Health and Biomedical Research Center, Al Ain University, Abu Dhabi Campus, P. O. Box 112612, Abu Dhabi, United Arab Emirates.*
[2]*Research Associate, Pakistan Biosafety Clearing House-Pakistan Environmental Protection Agency.*
[3]*Chief Operations Office, Sheikh Shakhbout Medical City (SSMC), PureHealth, Abu Dhabi, UAE.*
[4]*Department of Zoology, Government College University Lahore.*
[*5]*School of Biochemistry and Biotechnology, University of the Punjab, Lahore, Pakistan.*
[6]*Department of Biological Sciences, Virtual University of Pakistan, Lahore, Pakistan.*

[1]nadia.hussain@aau.ac.ae, [2]saboorbunny73@gmail.com, [3]ahhaddad@ssmc.ae, [4]riffatiqbal@gcu.edu.pk, [6]aleena.ahmad@vu.edu.pk

DOI: https://doi.org/10.5281/zenodo.16640393

Keywords
toluene, bioremediation, toluene 4-monooxygenase, conserved domains, interactions

Article History
Received: 28 April, 2025
Accepted: 13 June, 2025
Published: 29 July, 2025

Corresponding Author: *
Fatima Muccee
fatima.sbb@pu.edu.pk

Abstract

Toluene due to the associated health hazards and persistent nature has been declared a carcinogenic pollutant. Its mitigation from environmental sources is the major issue. Several toluene remediating bacteria have been reported which possess wide range of toluene degradation enzymes. To fully exploit the potential of these bacteria, there is need to get insights into these enzymes. Current study was initiated to characterize the toluene 4-monooxygenase (T4MO) enzyme from Pseudomonas aeruginosa. Initially, sequences of six subunits. i. e., TmoA, TmoB, TmoC, TmoD, TmoE and TmoF were retrieved from Uniprot database. These sequences were analyzed through different tools like PROTPARAM, SOPMA, ALPHAFOLD protein database and MEME SUITE. Analysis revealed the physicochemical attributes of components of enzymes as follows; number of amino acids (84-500), molecular weight (58103.54-9587.99), pI (4.38-5.57), aliphatic index (66.80-95.73) and GRAVY (-0.114-0.637). Secondary (2D) structure investigation demonstrated the alpha helix, extended strand and random coil ranging between 6-291, 74-6 and 194-34, respectively. TmoA and E subunits exhibited the highest level of complex folding. All the subunits were found to contain conserved motifs. Current study findings might be used for engineering of the enzyme using gene manipulation techniques in order to exploit its potential optimally.

INTRODUCTION

Due to the associated health and environmental risks, toluene (CAS No. 108-88-3) has been declared as persistent pollutant by Environmental Protection Agency (EPA). The environmental laws that EPA follows for its regulation include clean air act, clean water act, toxic substances control act (TSCA) and hazardous waste regulations (Agency 2005). In a study, the occupational exposure limit (OEL) for toluene has been recommended as 20 ppm 8-h with a short-term exposure limit (STEL) of 100 ppm for 15 min (Rooseboom et al. 2023).

It has tendency to get absorbed by gastrointestinal tract (GIT) and lungs. It affects central nervous system (CNS), causes deficiencies and lack of coordination in response time, skin and eyes irritation, nausea, vomiting, ventricular arrhythmias, memory loss and unconsciousness (Rajput et al. 2025). Its mitigation from environmental resources is crucial. Conventionally several techniques have been employed for this purpose including adsorption, incineration, condensation, advanced oxidation processes (AOPs), catalytic oxidation, dielectric barrier discharge (DBD) and soil vapor extraction (SVE) etc. (Chang et al. 2018, Mao et al. 2024, Wang et al. 2024, Zhou et al. 2019). These methods involve high cost, massive land disturbance, greater exposure to toluene by workers and incomplete degradation. So, there is need to replace these approaches with a better and more efficient strategy like Bioremediation. So far, wide range of toluene degrading bacteria have been isolated and characterized.

Genera of *Stenotrophomonas*, *Pseudomonas* and *Acinetobacter* isolated from waste water treatment plants were found associated with toluene degradation (Poyraz 2021). *Magentospirillum* sp. strain 15-1, *Acinetobacter* sp. Tol5, *Exiguobacterium mexicanum* M7, *Comamonas* sp. JB, *Burkholderia* sp. strain JS150, *Ideonella benzeniforans*, *Thauera* sp. strain DNT-1 and Cyperus (Barghoth et al. 2024, Bedics et al. 2022, Ishikawa et al. 2021, Jiang et al. 2015, Johnson and Olsen 1997, Meyer-Cifuentes et al. 2017, Ortega-González et al. 2013, Shinoda et al. 2004).

These bacteria exhibit diverse pathways for toluene breakdown like toluene 2-monooxygenase (T2MO) pathway has been reported in *Burkholderia cepacia* G4, toluene 3-monooxygenase (T3MO) pathway in *Pseudomonas pickettii* PKO1, toluene 4-monooxygenase (T4MO) pathway in *Pseudomonas mendocina* KR1, toluene dioxygenase (Tod) in *Pseudomonas putida*, toluene methylmonooxygenase (TOL) in *Alcanivorax* and ortho and meta ring cleavage pathways in *Burkholderia fungorum* and *Sphingomonas yanoikuyae* B1 (Busch et al. 2010, Dobslaw and Engesser 2015, Jindrova et al. 2002, Martínez-Lavanchy et al. 2015, Newman and Wackett 1995, Song et al. 2000, Zhang et al. 2025).

T4MO enzyme is composed of six components which catalyze the conversion of toluene into p-cresol through hydroxylation in the presence of NADH and O_2. This is the initial step in toluene 4-monooxygenase pathway (Figure 1) which is reported in *P. mendocina* KR1 and *Burkholderia* sp. JS150 (Parales et al. 2008). Keeping in view, the importance of T4MO enzyme and T4MO pathway in toluene degradation and the hazardous impact of toluene, we initiated current research documenting the characterization of this enzyme. This exploration of T4MO might help us in future in designing a potent enzyme through gene manipulation techniques.

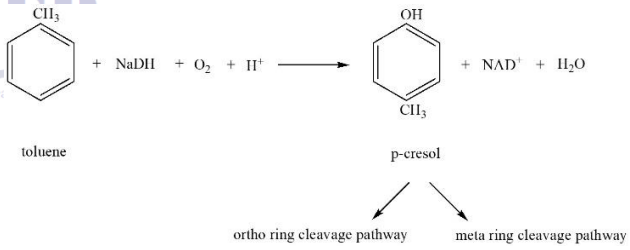

Figure 1: Function of toluene 4-monooxygenase (T4MO) in toluene degradation

2. Methodology

2.1 Retrieval of amino acid sequences of T4MO components:

UNIPROT DATABASE

Amino acid sequences and accession IDs of four components of T4MO were retrieved from UNIPROT DATABASE (Available at http://www.uniprot.org, accessed on April 2025) (Consortium 2015). This database is a protein sequence and protein annotation data resource. The components sequences were retrieved through searching for *Pseudomonas aeruginosa*.

2.2 Prediction of physicochemical attributes: PROTPARAM TOOL

To predict the physical and chemical properties of T4MO components documented in current study, EXPASY PROTPARAM TOOL (Available at http://web.expasy.org/protparam/, accessed on April 2025) was consulted (Garg et al. 2016). Tool helped in analyzing the parameters like number of amino acids, molecular weight, isoelectric point (pI), instability and aliphatic index and Grand Average of Hydropathicity (GRAVY). Input was provided as single letter code based protein sequence in FASTA format.

2.3 Prediction of secondary (2D) configuration: SOPMA TOOL

SOPMA Secondary Structure Prediction Method (Available at https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html, accessed on April 2025) (Geourjon and Deleage 1995). The input was provided as single letter code based protein sequence in FASTA format by selecting the output width of 70, number of conformations (4). i. e., helix, sheet, turn and coil. Similarity threshold was 8 and window width was 17.

2.4 Prediction of three (3D) configuration: ALPHAFOLD

ALPHAFOLD Protein Structure Database (Available at https://alphafold.ebi.ac.uk, accessed on April 2025) (Varadi et al. 2022). This database is based on AI system and is used to determine the 3D configuration of protein using amino acid sequence.

2.5 Prediction of conserved motifs: MEME SUITE

Motif Based Sequence Based Analysis Tool, The MEME SUITE (Available at https://meme-suite.org/meme/tools/meme, accessed on April 2025) (Nystrom and McKay 2021). This tool is employed to determine the fixed length patterns of ungapped novel motifs of proteins. For this analysis, classic mode of motif discovery was selected and sequence was Typed in as single letter code based amino acid sequence. For motif site distribution in sequences, Any Number of Repetitions (anr) was choosen. Number of motifs was 3. Each predicted motif was split into 2/3 sequences, each with separate p-values and single E-value, Motifs locations were provided in the form of diagramme with p-value.

3. Results

3.1 Sequences of T4MO enzyme components

IDs, accession numbers and sequences of subunits of T4MO enzyme. i. e., TmoA, TmoB, TmoC, TmoD, TmoE and TmoF, are described in detail in Table 1.

Table 1: Components of toluene 4-monooxygenase (T4MO), their Uniprot accession IDs and sequences

Enzyme subunit	Accession #	Sequence
Hydroxylase component subunit α (TmoA)	Q00456	MAMHPRKDWYELTRATNWTPSYVTEEQLFPERMS GHMGIPLEKWESYDEPYKTSYPEYVSIQREKDAGA YSVKAALERAKIYENSDPGWISTLKSHYGAIAVGEY AAVTGEGRMARFSKAPGNRNMATFGMMDELRHG QLQLFFPHEYCKKDRQFDWAWRAYHSNEWAAIAA KHFFDDIITGRDAISVAIMLTFSFETGFTNMQFLGLAA DAAEAGDYTFANLISSIQTDESRHAQQGGPALQLLIE NGKREEAQKKVDMAIWRAWRLFAVLTGPVMDYYT

		PLEDRSQSFKEFMYEWIIGQFERSLIDLGLDKPWYWD LFLKDIDELHHSYHMGVWYWRTTAWWNPAAGVTP EERDWLEEKYPGWNKRWGRCWDVITENVLNDRMD LVSPETLPSVCNMSQIPLVGVPGDDWNIEVFSLEHNG RLYHFGSEVDRWVFQQDPVQYQNHMNIVDRFLAGQ IQPMTLEGALKYMGFQSIEEMGKDAHDFAWADKCK PAMKKSA
Hydroxylase component subunit ϒ (TmoB)	Q00457	MSAFPVHAAFEKDFLVQLVVVDLNDSMDQVAEKVA YHCVNRRVAPREGVMRVRKHRSTELFPRDMTIAESG LNPTEVIDVVFEE
Ferredoxin componeny (TmoC)	Q00458	MSFEKICSLDDIWVGEMETFETSDGTEVLIVNSEEHG VKAYQAMCPHQEILLSEGSYEGGVITCRAHLWTFND GTGHGINPDDCCLAEYPVEVKGDDIYVSTKGILPNKA HS
Effector component (TmoD)	Q00459	MSTLADQALHNNNVGPIIRAGDLVEPVIETAEIDNPG KEITVEDRRAYVRIAAEGELILTRKTLEEQLGRPFNM QELEINLASFAGQIQADEDQIRFYFDKTM
Ferredoxin reductase (TmoE)	Q00460	MSFESKKPMRTWSHLAEMRKKPSEYDIVSRKLHYST NNPDSPWELSPDSPMNLWYKQYRNASPLKHDNWDA FTDPDQLVYRTYNLMQDGQESYVQSLFDQFNEREHD QMVREGWEHTMARCYSPLRYLFHCLQMSSAYVQQM APASTISNCCILQTADSLRWLTHTAYRTHELSLTYPDA GLGEHERELWEKEPGWQGLRELMEKQLTAFDWGEAF VSLNLVVKPMIVESIFKPLQQQAWENNDTLLPLLIDSQL KDAERHSRWSKALVKHALENPDNHAVIEGWIEKWRPL ADRAAEAYLSMLSSDILPAQYLERSTSLRASILTV

ferredoxin-NAD(+) reductase component (TmoF)	Q03304	MFNIQSDDLLHHFEADSNDTLLSAALRAELVFPYECNS GGCGACKIELLEGEVSNLWPDAPGLAARELRKNRFLA CQCKPLSDLKIKVINRAEGRASHPPKRFSTRVVSKRFLS DEMFELRLEAEQKVVFSPGQYFMVDVPELGTRAYSAA NPVDGNTLTLIVKAVPNGKVSCALANETIETLQLDGPY GLSVLKTADETQSVFIAGGSGIAPMVSMVNTLIAQGYE KPITVFYGSRLEAELEAAETLFGWKENLKLINVSSSVVG NSEKKYPTGYVHEIIPEYMEGLLGAEFYLCGPPQMINSV QKLLMIENKVPFEAIHFDRFF

3.2 PROTPARAM tool analysis

PROTPARAM tool analysis helped us to predict the physical and chemical attributes of enzyme. TmoA is the largest subunit with 500 amino acids and TmoB was the smallest one with 84 amino acids. TmoA is followed by TmoE, TmoF, TmoC and TmoD with 327, 326, 112 and 103 number of amino acids, respectively. The molecular weight of these components ranged between 9587.99 – 58103.54. Three subunits, TmoE (5.57), TmoA (5.20), TmoB (5.25) and TmoF (5.16) exhibited the highest pI values. Highest instability index was found in case of TmoD, TmoE and TmoB. i. e., 49.67, 47.77 and 41.98, respectively. The aliphatic index ranged between 95.73 and 66.80. All the subunits showed negative values for GRAVY ranging between -0.054 to -0.637 (Table 2).

Table 2: Prediction of physicochemical properties of components of T4MO using PROTPARAM tool

Component	No. of amino acids	Molecular weight	Isoelectric point (pI)	Instability index	Aliphatic index	GRAVY
TmoA	500	58103.54	5.20	39.27	66.80	-0.518
TmoB	84	9587.99	5.25	41.98	89.17	-0.114
TmoC	112	12325.73	4.38	27.78	77.41	-0.253
TmoD	103	11618.09	4.46	49.67	95.73	-0.359
TmoE	327	38346.28	5.57	47.77	74.92	-0.637
TmoF	326	35983.30	5.16	34.57	91.84	-0.054

3.3 SOPMA tool analysis

Alpha helix of T4MO subunits was highest in TmoA (291) and lowest in TmoC (6). In TmoF, 74 amino

acids were part of extended strand while in TmoE, the number of these amino acids was lowest. i. e., 6. Random coil ranged between 194 and 34 for TmoA and TmoB, respectively (Table 3).

Table 3: Prediction of secondary (2D) structure of components of T4MO using SOPMA tool

Component	Alpha helix (%)	Extended strand (%)	Random coil (%)
TmoA	291 (58.20)	15 (3.0)	194 (38.80)
TmoB	28 (33.33)	22 (26.19)	34 (40.38)
TmoC	6 (5.36)	36 (32.14)	70 (62.50)
TmoD	36 (34.95)	21 (20.39)	46 (44.66)
TmoE	215 (65.75)	6 (1.83)	106 (32.42)
TmoF	89 (27.30)	74 (22.70)	163 (50.00)

3.4 ALPHAFOLD: 3D configuration analysis

The prediction of 3D configuration of proteins by ALPHAFOLD revealed highest level of complexity in TmoA, TmoE and TmoF (Figure 2). While TmoB and TmoC and TmoD demonstrated least complex folding.

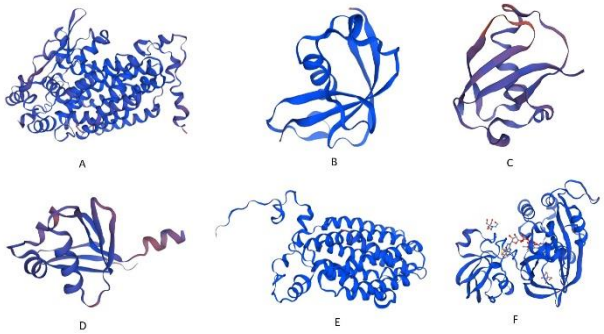

Figure 2: Prediction of three dimensional (3D) configuration of T4MO components based on SWISSMODEL
TmoA (B) TmoB (C) TmoC (D) TmoD (E) TmoE (F) TmoF

3.5 MEME SUITE Analysis

The six conserved motifs detected in case of TmoA were RMSGHMG, HHSYHMG, PSVCNM, PGNRNM, QTDESRHAQ and MMDELRHGQ with p-values 1.38e-9, 2.26e-9, 4.92e-9, 1.43e-8 And 2.77e-11, Respectively. In case of TmoB, six motifs identified were, AYHCVN, AESGLN, NDSMDQ, KDFLVQ, PRDMTI and PREGCM with p-values of 1.21e-8, 2.05e-7, 1.02e-7, 2.18e-7, 1.41e-7 And 3.23e-7, respectively. In TmoC, VKAYQAM, VITCRAH, LDDIWV, GDDIYV, LWTFND and METFET motifs were found with p-Values 2.67e-9, 3.14e-9, 7.48e-8, 5.43e-7, 1.30e-8 and 6.58e-7, respectively. In TmoD, domains identified included RKTLEEQLGRPFNM, MSTLADQALHNNNV, YFDKTM, SFAGQI, DRRAYN and DLVEPV. The p-values of these motifs were 1.08e-15, 8.29e-15, 2.73e-8, 6.48e-7, 3.64e-7 and 5.09e-7, respectively. In TmoE, VIEGWI, VREGWE, TISNCC, TMARCY, TAFDWG and TYPDAG motifs were identified with p-values 2.66e-9, 2.78e-8, 5.04e-9, 4.44e-8, 7.21e-9 and 1.28e-7, respectively. In case of TmoF, CGPPQM, CGACKI, KYPTGYVH, KVPFEAIH, KVPFEAIH, MFNIQSDD and HFEADSND domains were predicted with p-values 2.51e-8, 4.03e-8, 1.33e-10,

7.27e-10, 1.70e-10 and 5.19e-10, respectively (Table 4 And Figure 3).

Table 4: Prediction of conserved motifs of components of T4MO enzyme based on MEME suite

No.	E-value	p-value	Motif sequence
TmoA			
1	5.2e+000	1.38e-9	YVTEEQLFPE **RMSGHMG** IPLEKWESYD
2		2.26e-9	DLFLKDIDEL **HHSYHMG** VWYWRTTAWW
3	6.4e+000	4.92e-9	RMDLVSPETL **PSVCNM** SQIPLVGVPG
4		1.43e-8	EGRMARFSKA **PGNRNM** ATFGMMDELR
5	3.2e+001	2.77e-11	YTFANLISSI **QTDESRHAQ** QGGPALQLLI
6		2.77e-11	PGNRNMATFG **MMDELRHGQ** LQLFFPHEYC
TmoB			
1	6.4e+000	1.21e-8	DSMDQVAEKV **AYHCVN** RRVAPREGVM
2		2.05e-7	TELFPRDMTI **AESGLN** PTEVIDVVFE
3	5.7e+001	1.02e-7	FLVQLVVVDL **NDSMDQ** VAEKVAYHCV
4		2.18e-7	SAFPVHAAFE **KDFLVQ** LVVVDLNDSM
5	1.9e+002	1.41e-7	VRKHRSTELF **PRDMTI** AESGLNPTEV
6		3.23e-7	AYHCVNRRVA PREGCM RVRKHRSTEL
TmoC			
1	7.5e+000	2.67e-9	VLIVNSEEHG **VKAYQAM** CPHQEILLSE
2		3.14e-9	LLSEGSYEGG **VITCRAH** LWTFNDGTGH
3	1.7e+002	7.48e-8	MSFEKICS **LDDIWV** GEMETFETSD
4		5.43e-7	CLAEYPVEVK **GDDIYV** STKGILPNKA
5	9.7e+001	1.30e-8	EGGVITCRAH **LWTFND** GTGHGINPDD
6		6.58e-7	CSLDDIWVGE **METFET** SDGTEVLIVN
TmoD			

1	1.2e+001	1.08e-15	IAAEGELILT **RKTLEEQLGRPFNM** QELEINLASF
2		8.29e-15	**MSTLADQALHNNNV** GPIIRAGDLV
3	3.1e+001	2.73e-8	IQADEDQIRF **YFDKTM**
4		6.48e-7	NMQELEINLA **SFAGQI** QADEDQIRFY
5	1.4e+002	3.64e-7	DNPGKEITVE **DRRAYN** RIAAEGELIL
6		5.09e-7	NNVGPIIRAG **DLVEPV** IETAEIDNPG
TmoE			
1	3.2e+000	2.66e-9	HALENPDNHA **VIEGWI** EKWRPLADRA
2		2.78e-8	QFNEREHDQM **VREGWE** HTMARCYSPL
3	4.3e+000	5.04e-9	AYVQQMAPAS **TISNCC** ILQTADSLRW
4		4.44e-8	DQMVREGWEH **TMARCY** SPLRYLFHCL
5	3.7e+001	7.21e-9	GLRELMEKQL **TAFDWG** EAFVSLNLVV
6		1.28e-7	TAYRTHELSL **TYPDAG** LGEHERELWE
TmoF			
1	8.0e+001	2.51e-8	EGLLGAEFYL **CGPPQM** INSVQKLLMI
2		4.03e-8	VFPYECNSGG **CGACKI** ELLEGEVSNL
3	1.1e+002	1.33e-10	SSSVVGNSEK **KYPTGYVH** EIIPEYMEGL
4		7.27e-10	SVQKLLMIEN **KVPFEAIH** FDRFF
5	5.7e+001	1.70e-10	**MFNIQSDD** LLHHFEADSN
6		5.19e-10	FNIQSDDLLH **HFEADSND** TLLSAALRAE

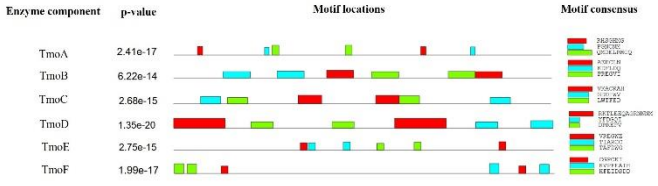

Figure 3: Prediction of conserved motifs, their locations and consensus of T4MO components based on MEME suite

4. Discussion

T4MO enzymes has previously been characterized for analysis of its crystal structure, functional study and for synthesis of chiral sulfoxides (Feingersch et al. 2008, Lountos et al. 2005). However, this is the first ever study documenting in-silico characterization of this enzyme. Among the subunits of T4MO, T4moD is the most crucial for catalysis as it binds TmoA

subunit and serves as catalytic effector. TmoB and C are associated with electron transfer (Hemmi et al. 2001).

The pI of protein effects its stability, activity, solubility, folding patterns and modifications (Schuurmans Stekhoven et al. 2008). TmoA, B, E and F exhibit slightly acidic character as compared to TmoC and D. Instability index below 40 shows stability of protein (Gamage et al. 2019). Hence, TmoB, D and E were found stable while TmoA, F and C were unstable. Aliphatic index is directly related with thermal stability of protein. TmoD was the highly thermostable followed by TmoF and TmoB. TmoA was the least thermostable. The less thermostable components may be engineered through modification of amino acids of their hydrophobic core (Ikai 1980). All the components of enzyme were hydrophilic with negative GRAVY values indicating their increase solubility in polar solvents.

Alpha helix content of proteins corresponds to the stability, DNA interactions and folding of membrane proteins. Only TmoA and E exhibited significant alpha helix character showing their more stable character and tendency of forming transmembrane domain (Li et al. 2019). The extended strand content represents tendency of enzyme catalysis and protein interactions of enzyme (Eswar et al. 2003). Only TmoF demonstrated good percentage of extended strand indicating its capacity of interacting other proteins.

Conserved motifs are the short sequences of amino acids that assist in understanding protein functions, its binding tendencies and active sites location. These sites may also be targeted to enhance the functioning of protein through engineering. All the subunits of T4MO documented in current study were found to exhibit the conserved motifs.

Current investigation enabled us to explore the different components of T4MO enzyme at structural, physical and chemical levels. This information might be exploited in future for the modification of enzyme for better toluene bioremediation, via genetic engineering. To enhance the stability of this enzyme, its unstable domains (TmoA, F and C) as well as conserved domains should be targeted.

Statements and Declarations:
Informed consent: N/A
Ethical approval: N/A

Competing interests: *The authors have no relevant financial or non-financial interests to disclose.*

Funding: *The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.*

Author contributions: N.H. write up, S.M.B. formal analysis, A.H.I.A.H. formal analysis, R.I. Data curation, F.M. conceptualization, supervision, A.A.K. formal analysis.

Acknowledgements: The authors would like to extend their gratitude to School of Biochemistry and Biotechnology, University of the Punjab, Lahore, Pakistan.

Submission declaration and verification: The manuscript has not been submitted anywhere else and is not under consideration by any other journal.

Data availability statement: The sequences of components of T4MO are available at uniprot database (http://www.uniprot.org).

References

Agency UEP. 2005. Integrated Risk Information System (IRIS) on Toluene. National center for environmental assessment, Office of Research and Development, Washington, DC.

Barghoth MG, Desouky SE, Radwan AA, Shah MP, Salem SS. 2024. Characterizations of highly efficient moderately halophilic toluene degrading *Exiguobacterium mexicanum* M7 strain isolated from Egyptian saline sediments. Biotechnology and Genetic Engineering Reviews 40:454-472.

Bedics A, Táncsics A, Tóth E, Banerjee S, Harkai P, Kovács B, Bóka K, Kriszt B. 2022. Microaerobic enrichment of benzene-degrading bacteria and description of *Ideonella benzenivorans* sp. nov., capable of degrading benzene, toluene and ethylbenzene under microaerobic

conditions. Antonie van Leeuwenhoek 115:1113-1128.

Busch A, Lacal J, Silva-Jímenez H, Krell T, Ramos JL. 2010. Catabolite repression of the TodS/TodT two-component system and effector-dependent transphosphorylation of TodT as the basis for toluene dioxygenase catabolic pathway control. Journal of Bacteriology 192:4246-4250.

Chang T, Shen Z, Huang Y, Lu J, Ren D, Sun J, Cao J, Liu H. 2018. Post-plasma-catalytic removal of toluene using MnO2–Co3O4 catalysts and their synergistic mechanism. Chemical Engineering Journal 348:15-25.

Consortium U. 2015. UniProt: a hub for protein information. Nucleic acids research 43:D204-D212.

Dobslaw D, Engesser KH. 2015. Degradation of toluene by ortho cleavage enzymes in *Burkholderia fungorum* FLU 100. Microbial biotechnology 8:143-154.

Eswar N, Ramakrishnan C, Srinivasan N. 2003. Stranded in isolation: structural role of isolated extended strands in proteins. Protein engineering 16:331-339.

Feingersch R, Shainsky J, Wood TK, Fishman A. 2008. Protein engineering of toluene monooxygenases for synthesis of chiral sulfoxides. Applied and environmental microbiology 74:1555-1566.

Gamage DG, Gunaratne A, Periyannan GR, Russell TG. 2019. Applicability of instability index for in vitro protein stability prediction. Protein and peptide letters 26:339-347.

Garg VK, Avashthi H, Tiwari A, Jain PA, Ramkete PW, Kayastha AM, Singh VK. 2016. MFPPI–multi FASTA ProtParam interface. Bioinformation 12:74.

Geourjon C, Deleage G. 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics 11:681-684.

Hemmi H, Studts JM, Chae YK, Song J, Markley JL, Fox BG. 2001. Solution structure of the toluene 4-monooxygenase effector protein (T4moD). Biochemistry 40:3512-3524.

Ikai A. 1980. Thermostability and aliphatic index of globular proteins. The Journal of Biochemistry 88:1895-1898.

Ishikawa M, Kojima T, Hori K. 2021. Development of a biocontained toluene-degrading bacterium for environmental protection. Microbiology spectrum 9:10.1128/spectrum. 00259-00221.

Jiang B, Zhou Z, Dong Y, Tao W, Wang B, Jiang J, Guan X. 2015. Biodegradation of benzene, toluene, ethylbenzene, and o-, m-, and p-xylenes by the newly isolated bacterium *Comamonas* sp. JB. Applied Biochemistry and Biotechnology 176:1700-1708.

Jindrova E, Chocova M, Demnerova K, Brenner V. 2002. Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene. Folia Microbiologica 47:83-93.

Johnson GR, Olsen RH. 1997. Multiple pathways for toluene degradation in Burkholderia sp. strain JS150. Applied and environmental microbiology 63:4047-4052.

Li B, Mendenhall J, Meiler J. 2019. Interfaces between alpha-helical integral membrane proteins: characterization, prediction, and docking. Computational and Structural Biotechnology Journal 17:699-711.

Lountos GT, Mitchell KH, Studts JM, Fox BG, Orville AM. 2005. Crystal structures and functional studies of T4moD, the toluene 4-monooxygenase catalytic effector protein. Biochemistry 44:7131-7142.

Mao Y, Yang L, Liu S, Song Y, Luo M, Guo Y. 2024. A theoretical study on toluene oxidization by OH radical. BMC chemistry 18:72.

Martínez-Lavanchy P, Chen Z, Lünsmann V, Marin-Cevada V, Vilchez-Vargas R, Pieper D, Reiche N, Kappelmeyer U, Imparato V, Junca H. 2015. Microbial toluene removal in hypoxic model constructed wetlands occurs predominantly via the ring monooxygenation pathway. Applied and environmental microbiology 81:6241-6252.

Meyer-Cifuentes I, Martinez-Lavanchy PM, Marin-Cevada V, Böhnke S, Harms H, Müller JA, Heipieper HJ. 2017. Isolation and characterization of *Magnetospirillum* sp. strain 15-1 as a representative anaerobic toluene-degrader from a constructed wetland model. PloS one 12:e0174750.

Newman LM, Wackett LP. 1995. Purification and characterization of toluene 2-monooxygenase from *Burkholderia cepacia* G4. Biochemistry 34:14066-14076.

Nystrom SL, McKay DJ. 2021. Memes: an R interface to the MEME Suite. bioRxiv:2021.2004. 2023.441089.

Ortega-González DK, Zaragoza D, Aguirre-Garrido J, Ramírez-Saad H, Hernández-Rodríguez C, Jan-Roblero J. 2013. Degradation of benzene, toluene, and xylene isomers by a bacterial consortium obtained from rhizosphere soil of *Cyperus* sp. grown in a petroleum-contaminated area. Folia Microbiologica 58:569-577.

Parales R, Parales J, Pelletier D, Ditty J. 2008. Diversity of microbial toluene degradation pathways. Advances in applied microbiology 64:1-73.

Poyraz N. 2021. Isolation of novel toluene degrading bacteria from waste water treatment plants and determination of their toluene tolerance and other biotechnological potential. Polish Journal of Environmental Studies 30:811.

Rajput SK, Singh S, Bhardwaj R. 2025. Toluene toxicity: Outline, management, and prognosis. Pages 391-404. Hazardous Chemicals, Elsevier.

Rooseboom M, Kocabas NA, North C, Radcliffe RJ, Segal L. 2023. Recommendation for an occupational exposure limit for toluene. Regulatory Toxicology and Pharmacology 141:105387.

Schuurmans Stekhoven F, Gorissen M, Flik G. 2008. The isoelectric point, a key to understanding a variety of biochemical problems: a minireview. Fish physiology and biochemistry 34:1-8.

Shinoda Y, Sakai Y, Uenishi H, Uchihashi Y, Hiraishi A, Yukawa H, Yurimoto H, Kato N. 2004. Aerobic and anaerobic toluene degradation by a newly isolated denitrifying bacterium, Thauera sp. strain DNT-1. Applied and environmental microbiology 70:1385-1392.

Song J, Sung J, Kim Y-M, Zylstra GJ, Kim E. 2000. Roles of the meta-and the ortho-Cleavage Pathways for the efficient utilization of aromatic hydrocarbons by *Sphingomonas yanoikuyae* Bl. Journal of Microbiology 38:245-249.

Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic acids research 50:D439-D444.

Wang X, Luo Y, Gao Y, Han D, Wang Z, Shen B, Wang X. 2024. Recent advances in catalysts for toluene elimination via catalytic oxidation. Chemosphere 368:143720.

Zhang B, Luo W, Liu Y, Chen H, Pan T, Wang B, Li K, Jia H, Zhou S, Meng S. 2025. The biodegradation pathway of HBCDs in the enriched microbial community.

International Biodeterioration & Biodegradation 201:106061.

Zhou W, Guan Z, Zhao M, Li J. 2019. Characteristics and mechanism of toluene removal from gas by novelty array double dielectric barrier discharge combined with TiO2/Al2O3 catalyst. Chemosphere 226:766-773.

.