

Mining of *Brevundimonas diminuta* MG genome to identify haloalkane dehalogenase with promiscuous epoxide hydrolase activity and its significance in green synthesis of bioactive molecules

Thesis submitted in partial fulfillment for the award of the degree
of

Doctor of Philosophy

by

**Annapoorni Lakshman Sagar
(15LAPH03)**

Supervisor

Prof. S. Dayananda



Department of Animal Biology

School of Life Sciences

University of Hyderabad-500046, India

APRIL 2022



University of Hyderabad
(A Central University established in 1974 by an act of the Parliament)
Hyderabad-500046, India

CERTIFICATE

This is to certify that the thesis entitled "Mining of *Brevundimonas diminuta* MG genome to identify haloalkane dehalogenase with promiscuous epoxide hydrolase activity and its significance in green synthesis of bioactive molecules" submitted by Ms. Annapoorni Lakshman Sagar, bearing registration number 15LAPH03 in partial fulfilment of the requirements for award of Doctor of Philosophy in the Department of Animal Biology, School of Life Sciences is a bonafide work carried out by her under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

Publications:

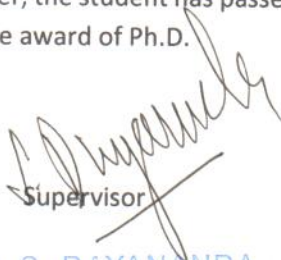
1. Microorganisms, 2020 Mar; 8 (3) (PMID: 32138166)
2. Genome Biol Evol, 2017 Jan; 9 (1) (PMID: 28175269)
3. Rev Environ Health, 2016 Mar; 31 (1) (PMID: 26953700)

Conference/Workshop:

1. Poster presentation at international conference of Association of Microbiologists of India (AMI-2019), held at University of Hyderabad, 9th-12th, December 2018.
2. Three-day workshop on "Exploring Gene Expression data using Transcriptome and Microarray" held at Computational Biology facility, Bio-Innovation Centre, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala, India, April 18-20, 2016.

Course Code	Name	Credits	Pass/Fail
AS 801	Seminar	1	Pass
AS 802	Research Ethics & Management	2	Pass
AS 803	Biostatistics	2	Pass
AS 804	Analytical Techniques	3	Pass
AS 805	Lab Work	4	Pass

Further, the student has passed the following courses towards the fulfillment of course work required for the award of Ph.D.


Supervisor

Prof. S. DAYANANDA
Dept. of Animal Biology,
School of Life Sciences
University of Hyderabad
Hyderabad-500 046, T.S.



Head of the Department

अध्यक्ष / HEAD
जंतु जैविकी विभाग
Department of Animal Biology


Dean of the School

DEAN
School of Life Sciences
University of Hyderabad
Hyderabad-500 046.



University of Hyderabad
(A Central University established in 1974 by an act of the Parliament)
Hyderabad-500046, India

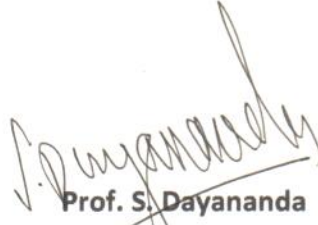
DECLARATION

This is to declare that the work embodied in this thesis entitled “Mining of *Brevundimonas diminuta* MG genome to identify haloalkane dehalogenase with promiscuous epoxide hydrolase activity and its significance in green synthesis of bioactive molecules” has been carried out by me under the supervision of Prof. S. Dayananda, Department of Animal Biology, School of Life Sciences. The work presented in this thesis is a bonafide research work and has not been submitted for any degree or diploma in any other University or Institute. A report on plagiarism statistics from the University Librarian is enclosed.

Date: 28-04-2022


Annapoorani Lakshman Sagar

15LAPH03


Prof. S. Dayananda
Supervisor

Prof. S. DAYANANDA
Dept. of Animal Biology.
School of Life Sciences
University of Hyderabad
Hyderabad-500 046, T.S.

ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor, **Prof. Dayananda Siddavattam**, for his constant encouragement, critical suggestions and insightful discussions throughout this study.

I thank my doctoral committee members Prof. Manjula Sritharan and Prof. Ch.Venkataramana for their kindness and timely advice. I am thankful to the present Dean, Prof. S. Dayananda and former Deans, School of Life Sciences, Prof. K.V.A Ramaiah, Prof. Reddanna, and Prof. A.S Raghavendra, for all the school facilities.

I thank the present Head, Prof. K.Srinivasulu, and former Heads of the Department of Animal Biology, Prof. Anita Jagota, Prof. Jagan Pongubala, and Prof. B. Senthilkumaran for all the department facilities.

I appreciate the help from Dr. Sarwar Azam, NIAB, Hyderabad, in analysing the whole genome sequencing data.

This work would have not been possible without the kindness of Prof. Yuji Nagata sensei ('teacher' in Japanese), Tohoku University, Japan, who hosted me in his lab for three months. I am grateful to his student Chen Nannan for her timely help with my research work. I am also grateful to Prof. Hiromi Kato sensei for demonstrating soil microcosm preparations. My stay in Japan was made very memorable by Prof. Y. Ohtsubo sensei and all the other students working in Prof. Nagata lab whom I wish a very bright future and happy life.

I am thankful to Dr. Santosh Padhi, Dept of Biochemistry, UoH, and his lab members for helping me with enzyme activity assays. I specially thank Ghufrana Abdus Sami who always patiently helped me with all the enzyme assays.

I thank all the faculty members in the School of Life Science, and allowing me to use their laboratory facilities whenever needed. I also thank the in-charges of common facilities of the department and school for their kind support.

I am thankful to my senior colleagues Dr. E.V. Paul, Dr. Swetha, Dr. Sunil, Dr. Ashok, Dr. Devyani, Dr. Vijay, Dr. Ramurthy, Dr. Harshitha, Dr. Hari and Dr. Aparna. I also take this opportunity to thank my current colleagues, Mr Akbar Pasha and Ms. Sandhya and especially Ms. Ganeshwari Dhurve in whom I also found a very dear friend. I am also thankful to all the project students of the lab. I thank my friends outside of work for many valuable memories.

The financial support provided by UGC (JRF & SRF), DST-JSPS, DBT, Biological E and Aurobindo Pharma and is greatly acknowledged. I also thank DBT, ICMR, UGC, UPE-II, UGC-SAP and DBT-CREBB, for providing the infrastructural facility of the school.

Special thanks to the non-teaching staff Mr. Rajendra, Mr. Upendra, Mr. Praveen, Mr. Jagan, Mr. Srinu, Mr. Nikhil and Mr. Sravan for their technical help. Our lab technician Mr Krishna is greatly acknowledged for all the help.

I am extremely grateful to my grandparents, my father (late) Shri. N.R. Lakshman Sagar, mother Smt. Seeta Lakshman Sagar, sister Ms. Bhavani Lakshman Sagar, brother Mr. N.L. Ramachandra Chudamani and my brother-in-law Dr. K.R. Gurukumar who have immensely supported me in all and my endeavours.

My deepest appreciation to my husband Dr. C.G. Venketasubramanian for being a wonderful life-partner and dearest friend. I also thank my in-laws, Shri. C.S. Gopalakrishnan, Smt. M.R. Seethalakshmi and my brother-in-law Dr. C.G. Sri Ranganath for their patience and understanding.

I am forever indebted to all my teachers for inspiring and believing in me. I thank Lord Ganapathi and all ancestors for the blessings.

LIST OF TABLES

Table 2.1 Antibiotics	16
Table 2.2 Chemicals	16
Table 2.3 Molecular biology	18
Table 2.4 Preparation of Antibiotics	19
Table 2.5 Components of a separating gel	27
Table 2.6 Components of a stacking gel	28
Table 2.7 Components of Phenol Red Dehalogenase Assay	31
Table 2.8 Components for Dehalogenase Assay	33
Table 2.9 List of halogenated substrates with concentration	34
Table 2.10 Components for Epoxide hydrolase Assay	35
Table 2.11 List of Epoxide substrates and diols used in this study	37
Table 3.1 General genome characteristics of <i>Sphingopyxis</i> strains used in Chapter III	41
Table 3.2 General genome features of <i>Sphingopyxis</i> sp MG.	46
Table 3.3 Comparison of <i>B. diminuta</i> genome with <i>Sphingopyxis</i> sps from NCBI database	47
Table 3.4 Predicted GI and their functions	53
Table 4.1 Characteristic properties of the E and NK soil types	78
Table 5.1 Bacterial strains used in Chapter V	87
Table 5.2 Plasmids used in Chapter V	88
Table 5.3 Primers used in Chapter V	89

LIST OF FIGURES

Fig. 1.1 Biotransformation of organophosphate pesticides. _____	3
Fig. 1.2 Schematic representation of various efflux pumps in bacteria. _____	4
Fig. 1.3 Complete degradation pathway for PCP (Pentachlorophenol) degradation in <i>Sphingobium chlorophenolicum</i> ATCC 39723. _____	5
Fig. 1.4 Relative distribution of k_{cat}/K_M ratios in degradative non-degradative enzymes. ____	7
Fig. 1.5 Comparison of structure and reaction mechanisms catalyzed by Phosphotriesterases and lactonases. _____	9
Fig. 1.6 Comparison of structure and reaction mechanisms catalyzed by methyl parathion hydrolase and β -lactamase. _____	10
Fig. 1.7 Comparison of structure and catalytic mechanism of melamine deaminase and atrazine chlorohydrolase. _____	11
Fig. 1.8 Reaction mechanism of pyrimidine deaminase, melamine deaminase and atrazine chlorohydrolase and their evolutionary trajectory. _____	12
Fig. 1.9 Schematic view highlighting the steps involved in synthesis of broad-spectrum antiviral drug Remdesivir (GS-5734) used in COVID-19 treatment. _____	13
Fig. 1.10 Structure of organophosphate pesticide Paraoxon (Panel A). Structure of S_p -stereoisomer of chemical warfare agent VX (Panel B). Structure of chiral precursor (R_p -stereoisomer) used in the synthesis of antiviral prodrug Remdesivir (Panel C). _____	15
Fig. 3.1 The circular map of the <i>Sphingopyxis</i> sp MG chromosome. _____	44
Fig. 3.2 The circular map of plasmid pCMS1. _____	45
Fig. 3.3 The circular map of plasmid pCMS2. _____	45
Fig. 3.4 Phylogenomic tree of the order Sphingomonadales. _____	48
Fig. 3.5 Pan/core genome analysis <i>Sphingopyxis</i> sp MG with other of <i>Sphingopyxis</i> species. _____	50
Fig. 3.6 Relationship of positively selected genes and the COG categories. _____	52
Fig. 3.7 Organization of ICE of <i>Sphingopyxis. granuli</i> MG. _____	56
Fig. 3.8 Organization of F_{420} biosynthetic island of <i>Sphingopyxis. granuli</i> MG. _____	58
Fig. 3.9 Organization of Arsenic resistance island of <i>Sphingopyxis. granuli</i> MG. _____	59
Fig. 4.1 Experimental design for CO ₂ trapping studies. _____	67

Fig. 4.2 Catabolic repertoire of <i>S. granuli</i> MG.	75
Fig. 4.3 In-silico predicted catabolic potential of <i>S. granuli</i> MG.	76
Fig. 4.4 Mineralization of ^{14}C -PNP by <i>S. granuli</i> MG.	77
Fig. 4.5 Survival plot of <i>S. granuli</i> MG in E and NK soil.	79
Fig. 4.6 Survival and growth of <i>S. granuli</i> MG in soil microcosm.	80
Fig. 4.7 Determination of in-situ mineralization of ^{14}C -PNP by measuring the trapped $^{14}\text{CO}_2$.	81
Fig. 5.1 The γ -HCH degradation pathway catalysed by LinA.	85
Fig. 5.2 Reaction mechanism of HLDs.	87
Fig. 5.3 Multiple sequence alignment of LinA_UT26 with <i>S. granuli</i> MG LinA homologs.	97
Fig. 5.4 Restriction digestion profiles of <i>sglinA</i> expression plasmids.	98
Fig. 5.5 Expression of LinA proteins.	99
Fig. 5.6 Dehalogenase activity of <i>sgLinA5</i> .	101
Fig. 5.7 Multiple sequence alignment and percent identity matrix of <i>sgHldA</i> and <i>sgHldB</i> .	103
Fig. 5.8 Phylogram showing the relationship of <i>sgHldA</i> and <i>sgHldB</i> with other HLD proteins.	104
Fig. 5.9 Cloning strategy and purification of <i>sgHldA</i> ^{N6His} and <i>sgHldB</i> ^{N6His} .	105
Fig. 5.10 Specific activity for <i>sgHldA</i> with halogenated substrates.	106
Fig. 5.11 Specific activity for <i>sgHldB</i> with halogenated substrates.	108
Fig. 5.12 Predicted tertiary structure of <i>sgHldA</i> .	109
Fig. 5.13 Predicted tertiary structure of <i>sgHldB</i> .	110
Fig. 5.14 Arrangement of catalytic site residues in <i>sgHldA</i> and <i>sgHldB</i> during dehalogenase activity.	112
Fig. 5.15 Arrangement of catalytic site residues in <i>sgHldA</i> and <i>sgHldB</i> during epoxide hydrolase activity.	113
Fig. 5.16 Specific activity of <i>sgHldB</i> towards epoxides.	114
Fig. 5.17 Sequencing, purification, HLD&EH assays of variant <i>sgHldB</i> ^{F37N} .	116
Fig. 6.1 GC-MS analysis of epoxide hydrolase activity of <i>sgHldB</i> .	123
Fig. 6.2 Assymetric hydrolysis of racemic styrene oxide.	123
Fig. 6.3 Chiral HPLC of R-, S- Styrene oxide.	124

LIST OF ABBREVIATIONS

Abbreviations	Full description
CEC	Contaminants of emerging concern
CFU	Colony forming units
°C	Degree centigrade
COG	Cluster of Orthologous groups
CPM	Counts per minute
DMSO	Dimethyl sulfoxide
DMT	Drug metabolite transporter
ECD	Electron capture detector
EC	Enzyme commission
GC	Gas chromatography
EH	Epoxide hydrolase
ELISA	Enzyme-linked Immunosorbent assay
FDR	False discovery rate
GI	Genomic Island
hr	hour
HAA	Hydroxyanthranilate
HCH	Hexachlorocyclohexane
HLD	Haloalkane dehalogenase
HPLC	High Performance Liquid Chromatography
HQ	Hydroquinone
ICE	Integrative conjugative element
IPA	Isopropyl alcohol
KBr	Pottasium Bromide
KCl	Pottasium Chloride
KI	Pottasium Iodide
kDa	Kilo dalton
kb	Kilo bases
µL	Microlitre
µM	Micromolar

μm	Micromoles
min	Minutes
MFS	Major Facilitator superfamily
MPF	Mating Pair formation
MPH	Methyl Parathion hydrolase
NP	Nitrophenol
OP	Organophosphate
OPAA	Organophosphate anhydrase
OPH	Organophosphate hydrolase
ORF	Open reading frame
PCCH	Pentachlorocyclohexene
PCP	Pentachlorophenol
PDB	Protein data bank
PNP	Paranitrophenol
POP	Persistent organic pollutant
PTE	Phosphotriesterase
PVDF	Polyvinylidene difluoride
RND	Resistance and Nodulation
SDS	Sodium dodecyl sulfate
SO	Styrene oxide
TCA	Tri carboxylic acid
TCB	Trichlorobenzene
TCDN	Tetrachlorocyclohexadiene
TCP	Trichloropropane
TEE	Translation enhancing element
TEV	Tobacco etch virus

TABLE OF CONTENTS

CERTIFICATE	ii
DECLARATION	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS	ix
CHAPTER 1.: Introduction	1
1.1. Environmental pollution	1
1.2. Bioremediation	1
1.2.1. Biotransformation	2
1.2.2. Xenobiotic efflux	3
1.2.3. Mineralization	4
1.3. Catalytic Promiscuity and Evolution of novel enzymes	5
1.4. Convergent - Divergent evolution and Enzyme promiscuity	7
1.4.1. Functional Convergence - Evolution of Phosphotriesterases	8
1.4.2. Functional Divergence - Evolution of Atrazine chlorohydrolase	10
1.5. Applications of enzyme promiscuity in green chemistry	12
1.5.1. Phosphotriesterase in synthesis of Remdesivir	13
1.6. Rationale of the study	15
1.7. Objectives	15
CHAPTER 2.: Materials and methods	16
2.1. Growth Media	18
2.1.1. Luria-Bertani (LB) media	19
2.1.2. Minimal salt medium for <i>Brevundimonas diminuta</i> MG	19
2.1.3. SOC media	19
2.2. Antibiotic & Chemical stock solutions	19
2.3. DNA methods	20
2.3.1. Preparation of Buffers and Solutions	20
2.3.2. Solutions for plasmid isolation	20
2.3.3. Isolation of plasmid DNA using Alkaline lysis method	21

2.3.4. Purification of plasmids using Miniprep kit method	22
2.3.5. DNA Quantification	22
2.3.6. Polymerase Chain Reaction (PCR)	22
2.3.7. Agarose gel electrophoresis	23
2.3.8. Molecular Cloning	23
2.4. Gene transfer methods	24
2.4.1. Preparation of ultra-competent E coli cells	24
2.4.2. Transformation	25
2.4.3. Electroporation	25
2.5. Protein Methods	25
2.5.1. Solutions for SDS-PAGE	25
2.5.2. Preparation of SDS-PAGE gels and Electrophoresis	27
2.5.3. Solutions for Western Blotting	28
2.5.4. Semi-Dry Western Blotting	29
2.5.5. Solutions for Protein Estimation	30
2.5.6. Protein Precipitation	30
2.6. Enzyme Assays and Preparation of reagents	30
2.6.1. Dehalogenase assay (Phenol-Red method) reagents	30
2.6.2. Dehalogenase Activity assay (Phenol-Red method)	31
2.6.3. Standard curve	32
2.6.4. Dehalogenase assay (Iwasaki method) reagents	32
2.6.5. Dehalogenase assay (Iwasaki method)	33
2.6.6. Substrates stock solutions	34
2.6.7. Epoxide hydrolase assay (Adrenaline Test) reagents	35
2.6.8. Epoxide hydrolase assay	36
2.6.9. Standard curve of diols	36
CHAPTER 3.: Structure and Phylogeny of <i>Brevundimonas diminuta</i> MG genome	38
3.1. Introduction	38
3.2. Objective specific methodology	38
3.2.1. Isolation of genomic DNA of <i>B. diminuta</i>	38
3.2.2. Structural and functional annotation of genome	39
3.2.3. Similar genome search	39
3.2.4. Orthology analysis	40

3.2.5. Phylogenomic and Phylogenetic analyses	41
3.2.6. Positive selection analyses	41
3.2.7. Genomic Island prediction	42
3.3. Results	42
3.3.1. Genomic sequencing & assembly	42
3.3.2. General features and Uniqueness of the genome	43
3.3.3. Comparative Genome Analysis	47
3.3.4. Pan /core genome analysis of <i>Sphingopyxis</i> sp MG	49
3.3.5. Positive selection analyses	51
3.3.6. <i>S. granuli</i> MG genome organisation: Evidence for Genomic Islands (GIs)	52
3.3.7. Largest Genomic Island is an Integrative Conjugative Element	54
3.3.8. Methane Metabolism-F ₄₂₀ biosynthesis Genomic Island	57
3.3.9. Arsenic Resistance Island	58
3.4. Discussion	60
CHAPTER 4.: Catabolic potential and adaptability of <i>Sphingopyxis. granuli</i> MG	63
4.1. Introduction	63
4.2. Objective specific Methodology	63
4.2.1. Gene Network/Pathway Analysis	64
4.2.2. Mineralization of ¹⁴ C-labelled PNP by <i>S. granuli</i> MG	64
4.2.3. Study design for soil microcosm	65
4.2.4. Study design for mixed culture microcosm	65
4.2.5. Study design for <i>in-situ</i> degradation of ¹⁴ C-labelled PNP in soil	66
4.3. Results	67
4.3.1. Genome of <i>S. granuli</i> MG is a treasure house of catabolic repertoire	67
4.3.2. Validation of <i>in-silico</i> predictions	77
4.3.3. <i>S. granuli</i> MG survives only in sterile E- type soils	78
4.3.4. Survival of <i>S. granuli</i> MG in soil microcosm	79
4.3.5. PNP degradation in soil microcosm by <i>S. granuli</i> MG	81
4.4. Discussion	82
CHAPTER 5.: Novel dehalogenases in the genome of <i>S. granuli</i> MG	84
5.1. Introduction	84
5.2. Objective specific methodology	87
5.2.1. Identification of LinA- HCH dehydrochlorinase homologues in <i>S. granuli</i> MG	90

5.2.2. Expression of <i>linA</i> gene in <i>E. coli</i>	90
5.2.3. Dehalogenase activity	91
5.2.4. Affinity Purification of <i>sgAL5</i> ^{N6His}	91
5.2.5. Dehalogenase activity of <i>sgLinA5</i> ^{N6His}	92
5.2.6. Multiple sequence alignment of haloalkane dehalogenase	93
5.2.7. Phylogenetic classification of haloalkane dehalogenases	93
5.2.8. Purification of <i>sgHldA</i> and <i>sgHldB</i>	93
5.2.9. Substrate specificity of <i>sgHldA</i> , <i>sgHldB</i> or <i>sgHldB</i> ^{F37N}	94
5.2.10. Homology modelling & visualizing active site architecture	95
5.2.11. Generation of <i>sgHldB</i> variant by site-directed mutagenesis	95
5.2.12. Purification of <i>sgHldB</i> ^{F37N}	95
5.2.13. Activity of <i>sgHldB</i> or <i>sgHldB</i> ^{F37N} with epoxides	96
5.3. Results	96
5.3.1. Putative <i>linA</i> -dehydrochlorinases in <i>S. granuli</i> MG	96
5.3.2. Heterologous expression and activity assays of <i>sgLinA</i>	98
5.3.3. Expression and purification of LinAs	99
5.3.4. Only <i>sgAL5</i> ^{N6His} shows weak HCH dehydrochlorinase activity	100
5.3.5. Haloalkane dehalogenases in <i>S. granuli</i> MG	102
5.3.6. Phylogenetic approach establishes relationship between <i>sgHldA</i> and <i>sgHldB</i>	104
5.3.7. Expression and Activity of <i>sgHldA</i> and <i>sgHldB</i>	105
5.3.8. Active site architecture of <i>sgHldA</i> and <i>sgHldB</i> for dehalogenation reaction	111
5.3.9. Active site architecture of <i>sgHldA</i> and <i>sgHldB</i> for epoxide hydrolase reaction	112
5.3.10. Epoxide hydrolase activity of <i>sgHldB</i>	113
5.3.11. Mutational studies with variant <i>sgHldB</i> ^{F37N}	115
5.4. Discussion	116
CHAPTER 6.: Enantioselectivity of epoxide hydrolase from <i>S. granuli</i> MG	120
6.1. Introduction	120
6.2. Objective specific methodology	121
6.2.1. Hydrolysis of epoxides using GC-MS analysis	121
6.2.2. Enantiospecific hydrolysis of racemic Styrene oxide	122
6.3. Results & Discussion	122
6.3.1. Formation of 1,2-phenyl-1-ethanediol	122
6.3.2. Resolution of (R-, S-) styrene oxide	124

CONCLUSIONS	xv
REFERENCES	xvii
PUBLICATIONS	xxviii
ANTI-PLAGIARISM REPORT	xxix

CHAPTER 1 : Introduction

1.1. Environmental pollution

Environmental pollution, be it soil, water or air pollution is a major challenge of the 21st century. Anthropogenic activities release pollutants at hazardously high concentrations into the environment causing deleterious impact on living organisms and ecosystems. These toxic contaminants, majority of which have industrial origin, have chemical structures that are alien to the biosphere (xenobiotics); for instance, persistent organic pollutants (POPs), endocrine disrupting chemicals (EDCs) and contaminants of emerging concern (CECs) are highly recalcitrant and resistant to degradation. Due to the recalcitrance they bioaccumulate in organisms and bio-magnify via food chains leading to adverse effects in human and wildlife in particular and environment in general.

Some of the major sources of the environmental pollutants include: (1) Pharmaceutical and chemical industries that produce multitude of synthetic polymers and xenobiotics; (2) Pulp and paper industries that are major sources of highly recalcitrant chlorinated organic compounds; (3) Incomplete combustion of fossil fuels like coal and petroleum which are responsible for the significant increase in CO₂ (green-house effect), acid rains and are sometimes released in huge amounts accidentally into the ecosystem (oil spill, etc); (4) Mining which causes heavy-metal pollution and (5) Intensive agriculture that release enormous amounts of pesticides and herbicides into the environment. Removal of the contaminants, restoration of the ecosystem and its conservation for the future generation should therefore be of utmost priority.

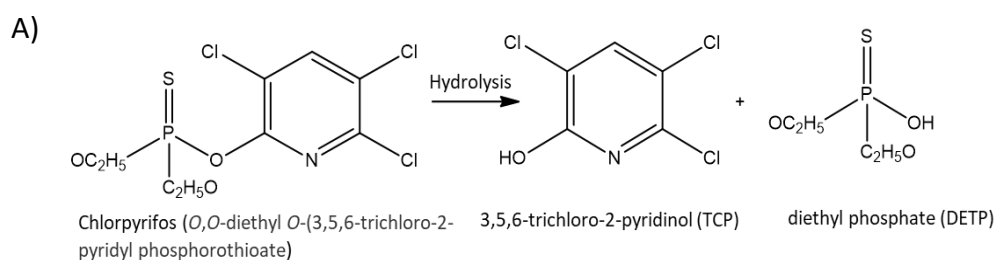
1.2. Bioremediation

Remediation of the contaminated sites is the only solution to rescue the future generation from environmental pollution. Compared to the slow and unpredictable way of encountering anthropogenic pollutants by natural physicochemical strategies, bioremediation; especially microbial bioremediation provides a promising alternative for environmental clean-up of toxic contaminants. Microbial transformation/bioremediation is often driven by diverse and versatile bacteria in response to the nutrient deficiency and pollutant induced stress. These conditions stimulate microbes to improvise their metabolic

potential and adapt to survive in presence of heterogeneous organic molecules, hitherto unknown to natural habitats. The microorganisms follow one or all of the three following strategies to survive in presence of these alien compounds. These are biotransformation, xenobiotic-efflux and mineralization.

1.2.1. Biotransformation

Microorganisms transform the toxic residues of the pollutants into non-toxic forms through enzyme-catalysed reactions. The process often referred as biotransformation alters characteristic property of the original compound affecting its solubility, mobility in the environment or toxicity of the pollutant. In nature, products of such bioconversions may be further transformed or degraded by other microorganisms in the microbial consortia eventually leading to complete degradation of the toxic pollutant. Biotransformation reactions are catalysed by mainly three types of enzymes i) hydrolases (Helbling, Hollender, Kohler, Singer, & Fenner, 2010) ii) oxido-reductases and iii) transferases (George & Haggblom, 2008; Larcher & Yargeau, 2011). Hydrolases convert pesticides like organophosphates (Lu et al., 2013), carbamates (Karns, Mulbry, Nelson, & Kearney, 1986), and pyrethroids (S. Chen et al., 2013) to less-toxic compounds by cleaving highly susceptible ester bonds (fig 1.1a). Likewise, detoxification of the carcinogenic azo dyes is done by bacterial oxidoreductases (Mahmood, Khalid, Arshad, Mahmood, & Crowley, 2016). The anthropogenic compounds containing nitro group (e.g., in explosives) (Annamaria, Manno, Strand, Bruce, & Hawari, 2010), cyano (e.g., herbicides) (Topp, Xun, & Orser, 1992) are transformed to less-toxic substrates by transferases (fig 1.1). The microorganisms catalysing this reaction may not necessarily degrade these substances further and more often possess broad substrate specificity.



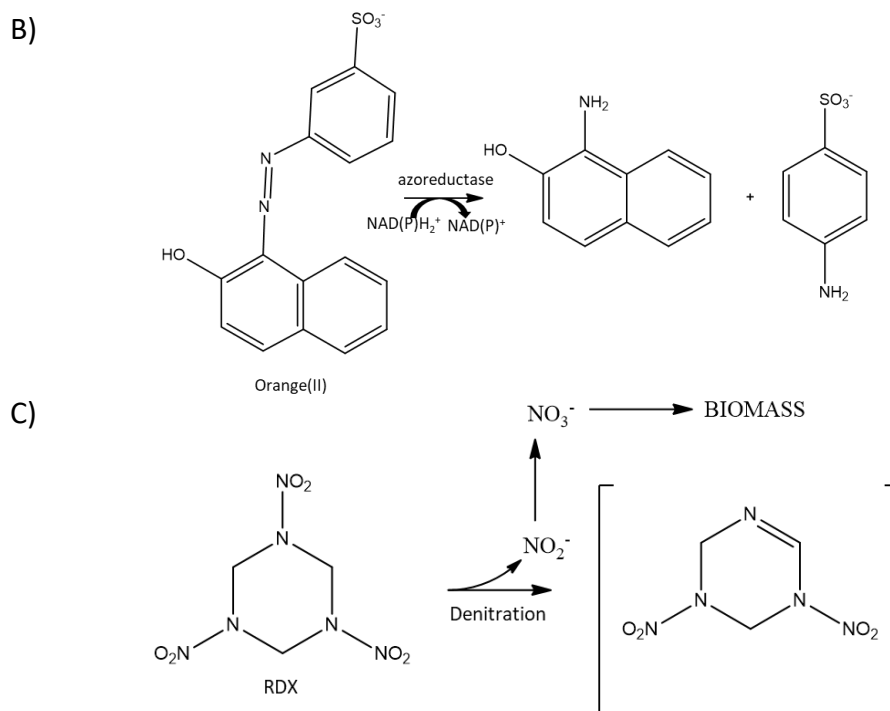


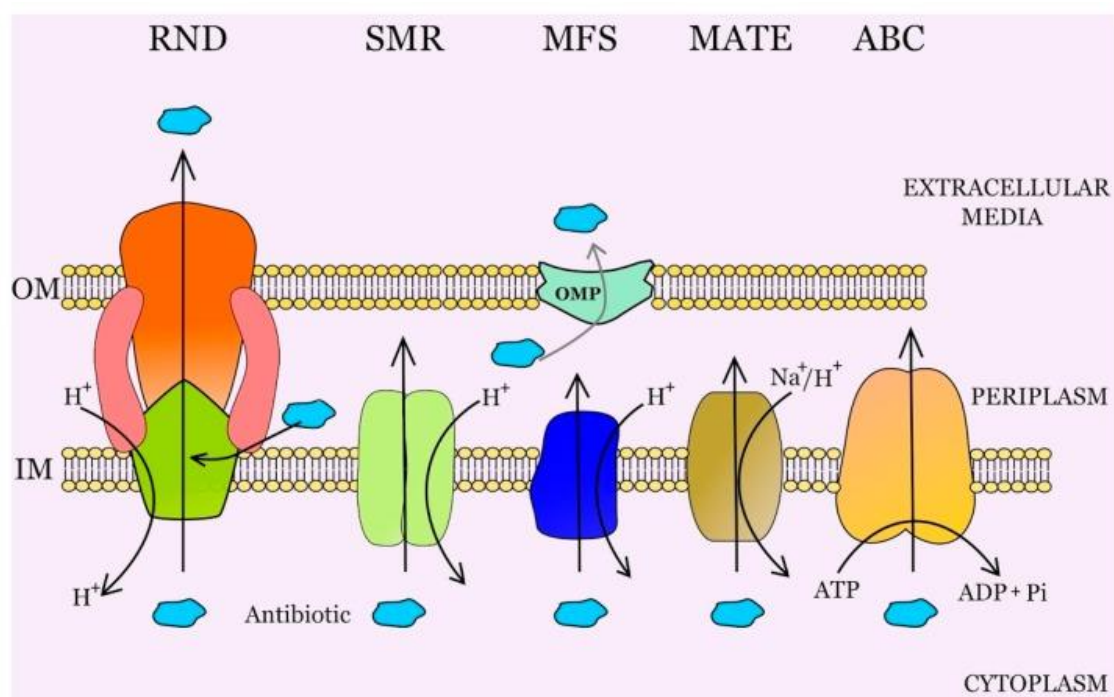
Fig. 1.1 Biotransformation of organophosphate pesticides.

a) Chlorpyrifos (Supreeth & Raju, 2017) b) Detoxification of Azo dye Orange II (H. Chen, 2006) c) Decomposition of explosive RDX (Annamaria et al., 2010).

1.2.2. Xenobiotic efflux

In response to selection pressure from antibiotics and heavy metals, microorganisms mainly respond through the mechanism of xenobiotic efflux via efflux pumps. Efflux pumps are transporters that are highly abundant in microorganisms and function not just to export xenobiotics but also have a myriad of physiological functions like solute uptake and export of cellular metabolites which contributes towards adaptation of microorganisms to challenging environments. Phylogenetically efflux transporters have been found to belong to one of the five transporter superfamilies - the major facilitator superfamily (MFS), the ATP-binding cassette superfamily (ABC), the drug-metabolite transporter superfamily (DMT), the resistance-nodulation-division superfamily (RND) and the multi-antimicrobial extrusion superfamily (MATE) (Saier & Paulsen, 2001) (fig 1.2). Physiological role of efflux pumps has been well demonstrated in quorum sensing (Alcalde-Rico, Hernando-Amado, Blanco, & Martinez, 2016), biofilm formation (Alav, Sutton, & Rahman, 2018), virulence and

pathogenicity, efflux of antibiotics and other exogenous molecules like heavy metals (Nies, 2003; Sun, Deng, & Yan, 2014) and recently in detoxification of endogenous substrates (Teelucksingh, Thompson, & Cox, 2020). In degradation of organic pollutants one of the well-studied efflux pumps is from *Pseudomonas putida* strain DOT-TE1 with a tripartite RND efflux pump involved in providing tolerance to toluene, an industrially important organic solvent and aromatic pollutant (Duque, Segura, Mosqueda, & Ramos, 2001).



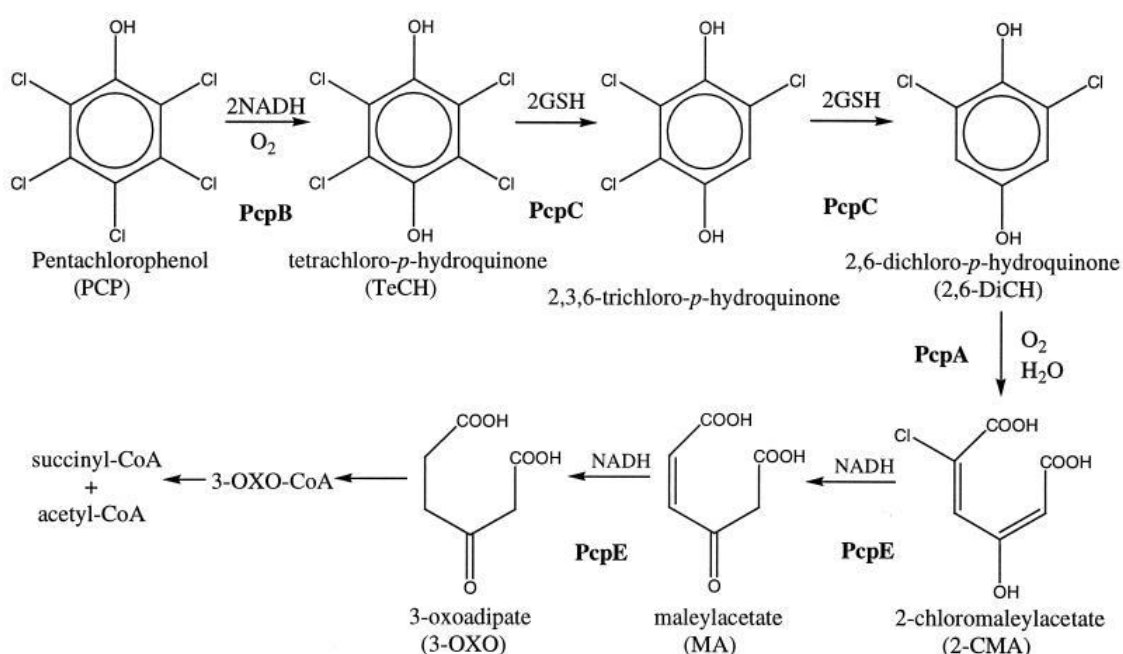
(Blanco et al., 2016)

Fig. 1.2 Schematic representation of various efflux pumps in bacteria.

1.2.3. Mineralization

Complete biodegradation or mineralization involves degradation of the organic compounds for consumption as sources of carbon and free energy. In *Pseudomonas sp.* strain ADP, widely used pesticide Atrazine is converted via three subsequent hydroxylation reactions to cyanuric acid which is then readily metabolized by this and other microorganisms (Shapir et al., 2007). In *Sphingobium chlorophenolicum* ATCC 39723, complete degradation pathway exists for degradation of wood preservative Pentachlorophenol (PCP). The enzymes in the pathway initially cleave three chlorine substituents followed by ring cleavage and

removal of two more chlorine substituents channelizing the toxic PCP to common TCA cycle intermediate β -ketoadipate (Cai & Xun, 2002; Dai, Rogers, Warner, & Copley, 2003) (fig 1.3).



(Cai & Xun, 2002)

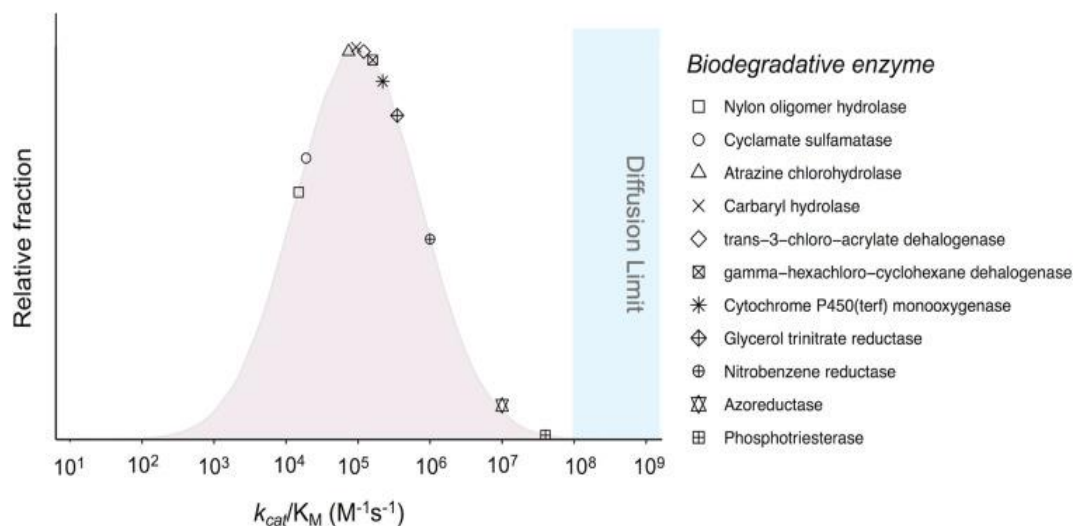
Fig. 1.3 Complete degradation pathway for PCP (Pentachlorophenol) degradation in *Spingobium chlorophenolicum* ATCC 39723.

Reactions catalysed by PcpA, 2,6-dichloroquinol-1,2-dioxygenase; PcpB, pentachlorophenol-4-monooxygenase; PcpC, tetrachloro-*p*-hydroquinone reductive dehalogenase and PcpE, maleylacetate reductase are shown

Selection pressure and use of novel substrates for adaptation and survival in presence of anthropogenic compounds has led to rapid evolution of new bio-degradative enzymes through enzyme plasticity. Enzyme plasticity suggests that under evolutionary selection pressure enzymes evolved as promiscuous biocatalyst efficiently transforming diverse substrates that are structurally similar to their cognate substrate (Kolvenbach, Helbling, Kohler, & Corvini, 2014; Wackett, 2009). Considering the diverse aromatic hydrocarbons often seen in polluted environments, catalytic promiscuity offers selective advantage and survival value for the microorganisms in these environments.

1.3. Catalytic Promiscuity and Evolution of novel enzymes

Enzyme promiscuity is aptly defined by Glasner et al (Glasner, Truong, & Morse, 2020) as the “coincidental ability of an enzyme to catalyse its native reaction and additional reactions that are not biological functions in the same active site”. An appealing hypothesis about the promiscuous function of enzymes was initially put forth in an inspiring review by Jensen et al (Jensen, 1976) proposing that primitive enzymes possessed a very broad substrate specificity. Survival and fitness of the ancestral organisms was attributed to these fewer but catalytically versatile enzymes that could perform a variety of functions albeit less efficiently. During the course of evolution, powered by selection pressure these generalist enzymes have been fine-tuned giving rise to modern specialist enzymes possessing specialized catalytic function with increased efficiency. Acquisition of enzyme proficiency may also proceed through specialist enzymes under selection pressure. In order to gain a new function for adaptation to selection pressure, specialist enzyme evolve spontaneously (without explicit selection against the old function) and diversify to become generalist exhibiting broad specificity. This phenomenon, also called “moonlighting”, is an excellent mechanism for generating complexity using existing protein without expanding the genome (Copley, 2003). In this route however, after a certain point the generalist enzyme may get re-specialized towards the new function with significant losses in its original activity (Khersonsky & Tawfik, 2010; Nam et al., 2012). Therefore, the selection pressure from adverse environments is crucial factor in shaping the evolution of novel enzymes. Microbial enzymes involved in biodegradation are exposed to the constant pressure from emerging environmental pollutants. The continuous presence of anthropogenic chemical promulgates the promiscuity in these enzymes which quickly evolve into new enzymes. In a recent study, 3500 well-studied enzymes (i.e, evolutionarily ancient) with their preferred substrate was compared for catalytic efficiency (K_{cat}/K_M) in the degradative enzymes. As seen in the fig 1.4, the dataset followed a Gaussian distribution (pink curve) with most enzymes having a median K_{cat}/K_M in the range of $10^4 - 10^6 \text{ M}^{-1} \text{ s}^{-1}$. Interestingly the median kinetic range of degradative enzymes, which have evolved recently in response to anthropogenic activities also show median in similar range, i.e, $10^4 - 10^7 \text{ M}^{-1} \text{ s}^{-1}$, indicating that new enzyme activities against man-made chemicals have evolved rapidly with substantial catalytic efficiency (Wackett & Robinson, 2020).



(Wackett & Robinson, 2020).

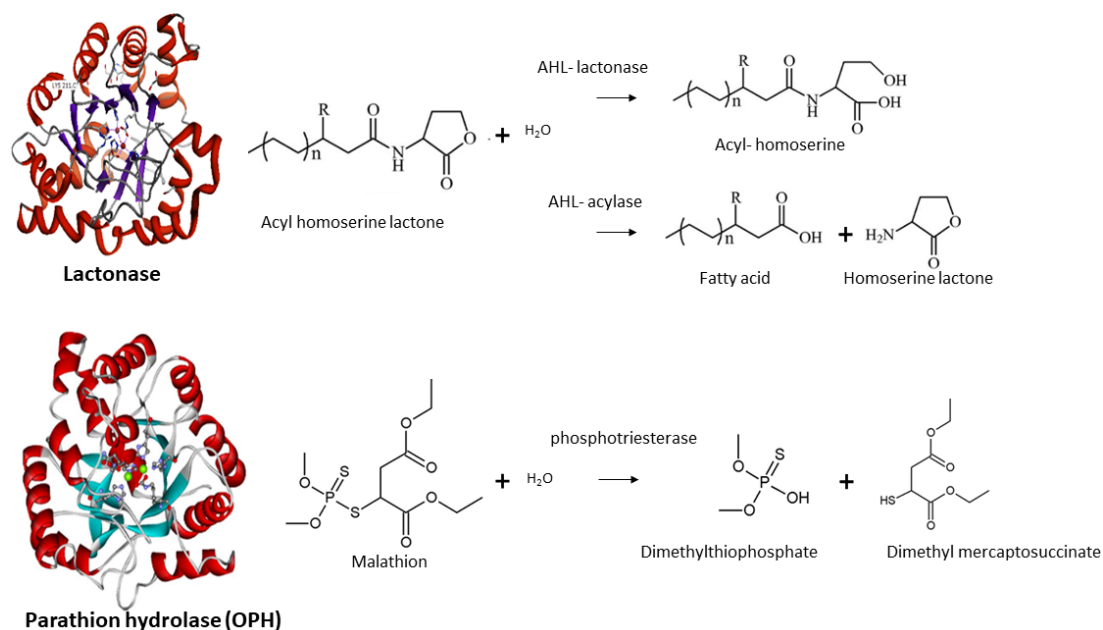
Fig. 1.4 Relative distribution of k_{cat}/K_M ratios in degradative non-degradative enzymes.

1.4. Convergent - Divergent evolution and Enzyme promiscuity

Evolution of enzymes is governed by two critical processes-functional convergence and functional divergence which give rise to convergent evolution or divergent evolution of enzymes (Zuckerandl & Pauling, 1965). In convergent evolution, more often enzymes with different secondary structure folds, or superfamilies, that have structural diversity share the same substrate and reaction mechanism (Elias & Tawfik, 2012). Such functional convergence results in members from distinct superfamily of proteins being recruited to catalyse the same metabolic reaction. In contrast, structurally similar enzymes act on different substrates to catalyse diverse biochemical reactions in divergent evolution (Galperin & Koonin, 2012). Functional divergence is a consequence of sequence and/or fold similarities in proteins typically belonging to the same superfamily. Enzyme promiscuity is reflected in shaping the evolutionary processes as it serves as evolutionary start point for novel enzymes (Copley, 2017; De Luca & Mandrich, 2020; Khersonsky & Tawfik, 2010). In the following section, classical examples especially for microbial enzymes involved in xenobiotic degradation with these modes of evolution is discussed elucidating key role of enzyme promiscuity in enzyme evolution.

1.4.1. Functional Convergence - Evolution of Phosphotriesterases

Phosphotriesterases are classes of bacterial enzymes that catalyse the hydrolysis of third ester linkage in toxic organophosphate pesticides like parathion, methyl parathion, fenusulfothion and nerve agents like sarin, soman with high catalytic efficiency (Tsai et al., 2012). The organophosphate compounds (OPs) are potent acetylcholinesterase (AChE) inhibitors and highly neurotoxic (Naughton & Terry, 2018; Sanchez-Santed, Colomina, & Herrero Hernandez, 2016). They were synthesized initially as chemical warfare agents during World War II but later used as pesticides (Epstein, 2014; Kloske & Witkiewicz, 2019). Therefore, the chemistry of OP compounds was novel for existing enzymes that attempted to degrade these compounds and have newly acquired catalytic functions. The enzymes involved in degradation of OP compounds also called triesterases are organized into three distinct structural families such as organophosphorous acid anhydrases (OPAA), phosphotriesterases (PTE), methyl parathion hydrolases (MPH). One of the best characterized enzymes of the three families of hydrolases is Organophosphate hydrolase (OPH) that belongs to the PTE family, initially isolated from *Pseudomonas sp diminuta* in the 1980s (Dumas, Caldwell, Wild, & Raushel, 1989). PTEs are highly efficient enzymes acting specifically on phosphotriesters with diffusion-limited kinetics is considered to be at its evolutionary endpoint. Moreover, no close homologs of these enzymes have been identified and their evolutionary origins are unknown (Russell et al., 2011). PTEs showed considerable promiscuous lactonase activity ($K_{cat}/K_M = 160\text{--}6.5 \times 10^5 \text{ s}^{-1} \text{ M}^{-1}$) and are considered to have evolved from an ancestral PTE-like lactonase with low level paraoxanase activity as a starting point (Afriat, Roodveldt, Manco, & Tawfik, 2006). The PTE-like lactonases which process quorum-sensing lactones also share both fold and active site architecture with PTEs (Del Vecchio et al., 2009; Hawwa, Aikens, Turner, Santarsiero, & Mesecar, 2009). Moreover, PTE-like lactonases with proficient lactonase activity ($K_{cat}/K_M = 6 \times 10^4 \text{--}1.7 \times 10^6 \text{ s}^{-1} \text{ M}^{-1}$) also show weak promiscuous paraoxanase activity. It is proposed that the paraoxanase activity might have been latent enzyme activity and selectively neutral initially but owing to the selection pressure from toxic OP compounds was positively selected and gave rise to the present day highly catalytically efficient PTEs (fig 1.5).

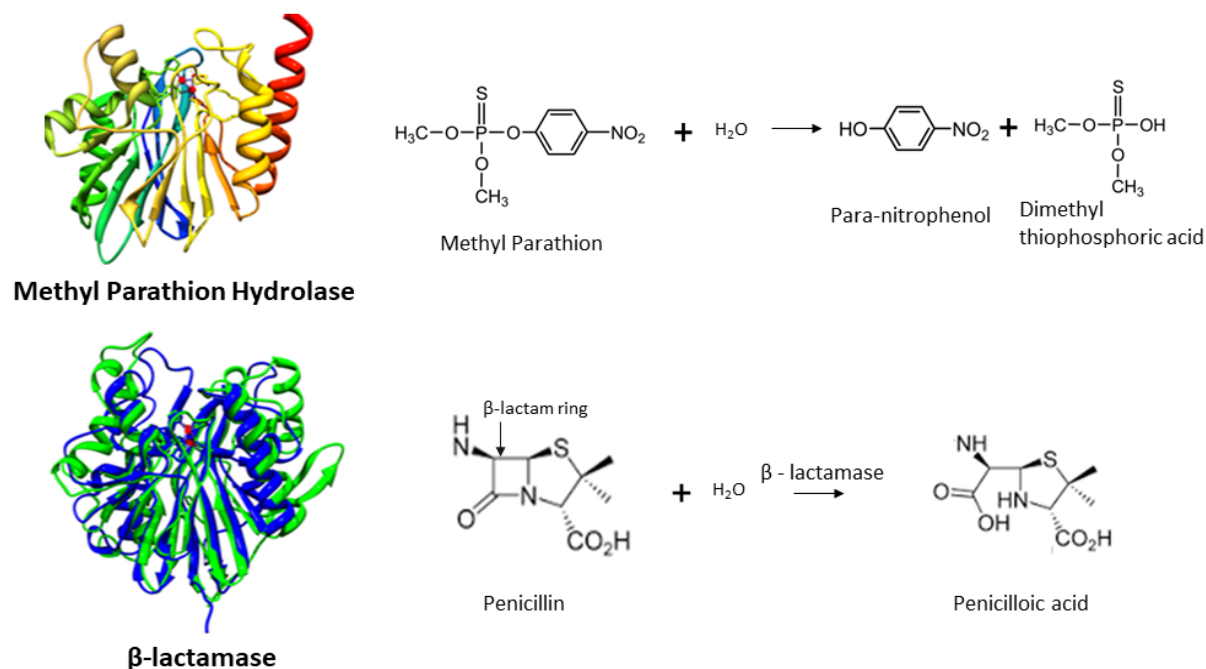


(Benning, Shim, Raushel, & Holden, 2001).

Fig. 1.5 Comparison of structure and reaction mechanisms catalyzed by Phosphotriesterases and lactonases.

The MPH and OPAA are structurally different from organophosphate hydrolase (OPH) but hydrolyze phosphotriesters, albeit at lower catalytic rates. OPAA has a broad substrate range and phylogenetically OPAA is distinct from both MPH and OPH and its evolutionary origins states that it is in fact a prolidase with dipeptidase activity. The hydrolysis of OP nerve agents by OPAA is a consequence of structural similarity with dipeptidases (Vyas, Nickitenko, Rastogi, Shah, & Quirocho, 2010). MPH specifically catalyzes hydrolysis of methyl parathion. Contrary to OPH, MPH shows structural similarity to β -lactamases that cleave the β -lactam ring present in a wide variety of antibiotics (fig 1.6).

Interestingly, a virtually identical catalytic mechanism of OP hydrolysis has evolved independently from two structurally distinct but mechanistically similar enzymes - OPH and MPH. This represents the classical example for structurally divergent enzymes with functional convergence. Thus, evolution of phosphotriesterases follows the convergent mode of evolution.



(Dong et al., 2005).

Fig. 1.6 Comparison of structure and reaction mechanisms catalyzed by methyl parathion hydrolase and β-lactamase.

1.4.2. Functional Divergence - Evolution of Atrazine chlorohydrolase

Atrazine is a heavily chlorinated pesticide introduced in 1959 and was considered to be poorly biodegraded in soils. However, after only a few decades of its introduction the soil microbes rapidly degraded them showing presence of gene *atzA*, encoding atrazine chlorohydrolase, the first enzyme in atrazine biodegradation (Udikovic-Kolic, Scott, & Martin-Laurent, 2012). The enzyme was found to be very specific for atrazine as it dehalogenated the fluorinated analog at much slower catalytic rates. The sequence of AtzA from *Pseudomonas* sp. strain ADP was found to be almost 98% identical and differed by just nine amino acids to the sequence of Triazine deaminase (TriA) from *Pseudomonas* sp. strain NRRL B-12227. TriA is a melamine deaminase involved in hydrolytic deamination of melamine to ammeline (4,6-diamino-1,3,5-triazin-2-ol) (Seffernick, de Souza, Sadowsky, & Wackett, 2001). The switch of chemospecificity from deamination to dehalogenation studied in an experiment recapitulating the original natural evolutionary pathway showed that just two point mutation were sufficient for the enzyme to catalyze dehalogenation indicating a very short evolutionary

trajectory for diversification of these enzymes (Seffernick & Wackett, 2001). Interestingly, TriA also showed promiscuous dehalogenase activity. However, atrazine chlorohydrolase did not show any promiscuous deaminase activity (Noor et al., 2012) (fig 1.7).

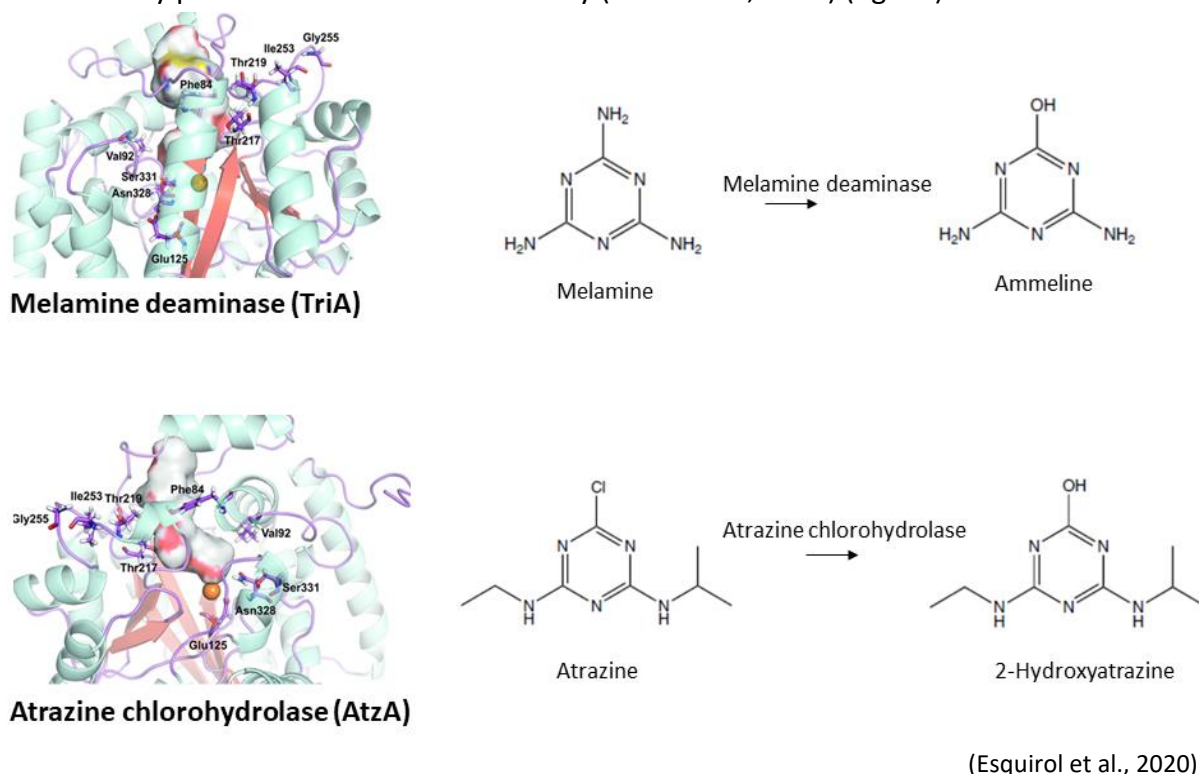
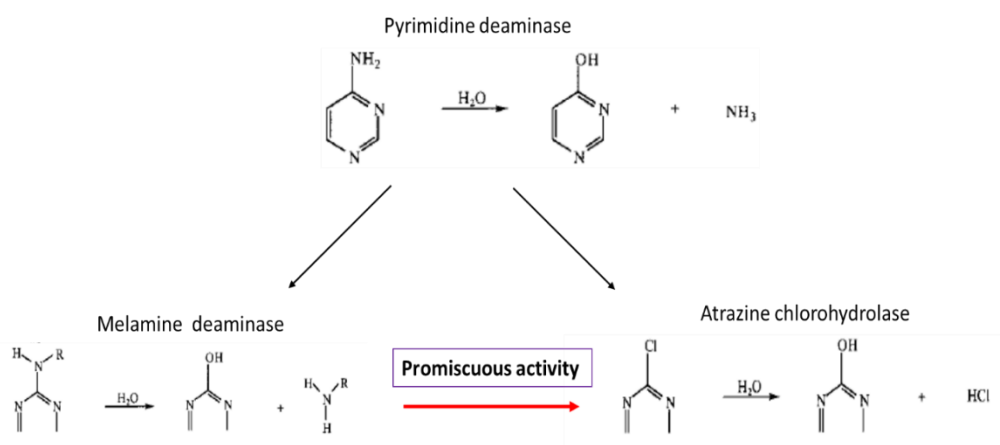


Fig. 1.7 Comparison of structure and catalytic mechanism of melamine deaminase and atrazine chlorohydrolase.

Structurally both atrazine and melamine belong to a broad group of compounds known as s-triazines comprising of a 3-nitrogen ring and 3-substituents. The compounds are synthetic and have been recently introduced into the environment however, they share extensive similarities with naturally occurring nucleotide building precursors - purines and pyrimidines. Moreover, the amidohydrolase family of enzymes catalysing purine and pyrimidine deamination are ancient and indispensable enzymes in biological systems. The reaction mechanism of deamination involves protonation of nitrogen adjacent to ring carbon undergoing nucleophilic substitution. Considering the fact that chloride ion is an excellent leaving group for nucleophilic substitution, the s-triazine compounds especially atrazine is more susceptible to activation than pyrimidine ring system suggesting plausible mechanistic similarities in deamination and atrazine dechlorination based on substrate similarity of these compounds (Gilchrist, 1992). The sudden discovery of enzymes capable of mineralizing s-

triazines suggests that selection pressure may have triggered the primitive pyrimidine deaminase that catalyse hydrolytic removal of amino groups from purine and pyrimidine rings to neo functionalize by point mutation and evolve into diverse enzymes such as TriA and AtzA that share the structural similarities but differ in their catalytic capabilities (fig 1.8).



(Seffernick & Wackett, 2001)

Fig. 1.8 Reaction mechanism of pyrimidine deaminase, melamine deaminase and atrazine chlorohydrolase and their evolutionary trajectory.

Such evolution of ancestral enzymes that share structural similarities but catalyse diverse biochemical reactions are an outcome of functional divergence. The evolution of atrazine chlorohydrolase thus follows a divergent mode of evolution.

These two modes of evolution suggest that the catalytic promiscuity and substrate ambiguity are critical factors in acquisition of novel biological functions. In addition to such crucial evolutionary insights, promiscuous enzymes also provide a platform for protein engineering and developing sustainable approaches to synthesize bioactive molecules. The earliest example of substrate and enzyme promiscuity exploited for industrial application was demonstrated in 1921 using pyruvate decarboxylase in C-C bond formation (Neuberg & Hirsch, 1921). This process is now used in a number of industrial activities suggesting that the understanding of substrates and mechanistic of enzyme promiscuity can be fine-tuned for development of new catalytic functions.

1.5. Applications of enzyme promiscuity in green chemistry

Destroying the strict substrate specificity by simple point mutations and/or enhancing the promiscuous activity of enzymes using protein engineering approaches is one of the most emerging techniques used in developing novel enzyme-substrate reactions and to increase turnover of required product. In addition to relaxing substrate specificity, promiscuity of enzymes can also be exploited via directed evolution to resolve enantioselectivity. Thus, synthesis of enantiopure industrially important bioactive molecules can be produced through enzyme engineering. The evolutionary insights provided by enzyme promiscuity has paved ways for rational designing of enzymes with new metabolic transformations and fine-tuned substrate profiles for improved catalytic efficiency.

1.5.1. Phosphotriesterase in synthesis of Remdesivir

Covid-19, a severe respiratory disease caused by the novel coronavirus (SARS-CoV-2) was identified in December 2019 and was declared as pandemic (Cucinotta & Vanelli, 2020). The severe outbreak called for an international health emergency and several therapeutic agents were rigorously evaluated for their efficacy. Remdesivir was found to inhibit the viral RNA-dependent, RNA polymerase of SARS-CoV-1 and was one of the first antiviral drugs identified with good clinical efficacy as effective therapeutic agent against SARS-CoV-2 infection (Liang et al., 2020).



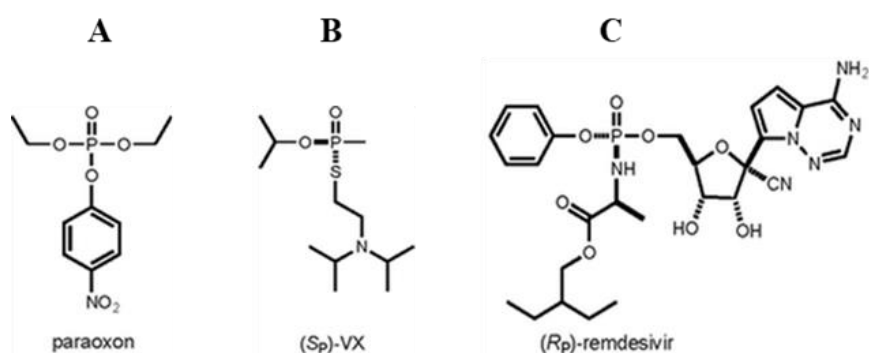
(Vargas, Larghi, & Kaufman, 2021).

Fig. 1.9 Schematic view highlighting the steps involved in synthesis of broad-spectrum antiviral drug Remdesivir (GS-5734) used in COVID-19 treatment.

Structurally, Remdesivir is a monophosphoramidate prodrug which contains an adenosine analogue with a chiral phosphorous centre. The S_p diastereomer shows higher selectivity and provides better therapeutic window for suitable antiviral therapy. However,

its synthesis requires highly pure R_p diastereomer precursor (Siegel et al., 2017). Raushel proposed a chemoenzymatic strategy for synthesis of the key intermediate required for the synthesis of Remdesivir by employing variant of the phosphotriesterase from *Pseudomonas diminuta* (Bigley, Narindoshvili, & Raushel, 2020) (fig 1.9).

Phosphotriesterases are best known for their role in detoxification of organophosphate pesticide and chemical warfare agents. The latter includes tabun (GA), sarin (GB), soman (GD), cyclosarin (GF), VX and VR all consisting of a chiral phosphorous and thereby each of these nerve agents have two stereoisomers. The differential enantioselectivity also effects the toxicity of the nerve agents for example, the S_P -stereoisomer of VX is 100 times more toxic than its R_P -stereoisomer. Interestingly, by manipulating the residues near the catalytic site by multiple mutations, inversion of the stereoselectivity with enhanced ability to hydrolyse the more toxic S_P -stereoisomer was identified. In a remarkable achievement, using rational and random mutagenesis H257Y/L303T variant of PTE was shown to hydrolyse the S_P -stereoisomer of sarin with a value of K_{cat}/K_M of $6 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ and engineered mutant was one of the best catalysts for hydrolysis of nerve agents (Bigley, Xu, Henderson, Harvey, & Raushel, 2013; Tsai et al., 2010). The mutants retained their catalytic power to hydrolyse organophosphates despite amendment at the active sites for manipulation of stereospecificity. This ability to reshape the active site architecture for differential hydrolysis of enantiomeric compounds from S_P - stereoisomer to R_P - stereoisomer was exploited to synthesize chiral precursor of antiviral prodrug Remdesivir (fig 1.10a-c).



(Holden & Raushel, 2021)

Fig. 1.10 Structure of organophosphate pesticide Paraoxon (Panel A). Structure of S_p-stereoisomer of chemical warfare agent VX (Panel B). Structure of chiral precursor (R_p-stereoisomer) used in the synthesis of antiviral prodrug Remdesivir (Panel C).

1.6. Rationale of the study

Soil microbes are considered as the treasure house of enzymes with novel catalytic properties. Our laboratory has been working on biodegradation of organophosphates using soil bacterium *Brevundimonas diminuta* MG as model organism. *B. diminuta* MG is an archetypical organophosphate degrader and thrives on OP insecticides and nitrophenols (Segers et al., 1994). Its quick adaptability and ability to survive on xenobiotics indicates existence of novel catabolic repertoire in its genome. However, due to lack of complete genome information, the strain has been identified only with the degradation of organophosphates. This study is specifically designed, with the following objectives, to elucidate the degradative potential of *B. diminuta* and to exploit the novel catalytic enzymes, if any, in bioremediation/ green synthesis of bioactive molecules.

1.7. Objectives

1. To determine the structure of *Brevundimonas diminuta* genome, genomic island landscape and core genome assisted phylogeny.
2. To gain genome guided insights on catabolic potential and adaptability of *Brevundimonas diminuta*.
3. To identify and characterize novel dehalogenases through investigative genome mining.

CHAPTER 2 : Materials and methods

List of reagents and manufacturer is presented from Table 2.1 to Table 2.3

Table 2.1 Antibiotics

Antibiotic	Manufacturer
Ampicillin sodium salt	HiMedia
Chloramphenicol	HiMedia
Kanamycin Sulphate	HiMedia
Gentamycin	HiMedia
Streptomycin	HiMedia
Polymyxin B	HiMedia
Tetracycline hydrochloride	HiMedia

Table 2.2 Chemicals

Reagents	Manufacturer
Acetic Acid (Glacial)	Sisco Research Laboratories (SRL)
Acetone	SRL
Acetonitrile	SRL
Acrylamide	Sigma Aldrich
Agar agar	HiMedia
Ammonium persulphate	GE Healthcare Lifesciences
Bovine serum albumin	HiMedia
Bromophenol blue	SRL
Butanol	SRL
β -mercaptoethanol	Sigma-Aldrich
Calcium chloride	SRL
Calcium nitrate	SRL
CHES (N-cyclohexyl-2-aminoethanesulfonic acid)	SRL
Coomassie Brilliant Blue R-250	HiMedia

Cobalt chloride	SRL
Clarity Western ECL Substrate	BIO-RAD
Diammonium hydrogen phosphate	SRL
Dipotassium hydrogen phosphate	SRL
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	SRL
Ethidium bromide	HiMedia
Ethanol	Hayman Laboratories
Ferrous sulphate	Sigma-Aldrich
Ficoll 400	Sigma-Aldrich
Glycerol	SRL
Glycine	SRL
Iso-amyl alcohol	Fischer Scientific
Iso-propyl alcohol	Fischer Scientific
Isopropyl thiogalactopyranoside (IPTG)	G-Biosciences
Imidazole	SRL
Magnesium sulphate	SRL
Methanol	SRL
Nickel sulphate heptahydrate	SRL
N, N'-Methylene bis acrylamide	GE Healthcare Lifesciences
Peptone	HiMedia
Potassium chloride	Qualigens
Potassium dihydrogen ortho phosphate	Merck
PMSF (Phenylmethylsulfonyl fluoride)	GE Healthcare
Protease inhibitor cocktail	Sigma Aldrich
Sodium chloride	SRL
Sodium dodecyl sulfate	SRL
Sodium hydroxide	SRL
Tetra ethyl methylene diamine (TEMED)	Sigma Aldrich
Tris-base	Merck

Triton X-100	Sigma Aldrich
Tryptone	HiMedia
Tween 20	Amersco
Yeast extract	HiMedia

Table 2.3 Molecular biology

Reagents	Manufacturer
GeneJET Plasmid Miniprep Kit	Thermo Scientific
DNeasy blood and tissue kit	Qiagen
GeneJet Gel Extraction kit	Thermo Scientific
Lysozyme	Sigma Aldrich
DNase I	Thermo Scientific
RNAse	Thermo Scientific
T4 DNA Ligase	Thermo Scientific
Phusion® High-Fidelity DNA polymerase	New England Biolabs
Taq DNA polymerase	TaKaRa
Unstained Protein Ladder, 10-200kDa	PurGene
HyperLadder 1kb DNA Ladder	Bioline
Nde1	Thermo Scientific
Xho1	Thermo Scientific
EcoR1	Thermo Scientific
BamH1	Thermo Scientific
BglII	Thermo Scientific

2.1. Growth Media

Liquid growth media were prepared and used for the either propagation of bacteria, production or expression of recombinant proteins. Sterilized solid media was prepared by adding 2 g of agar per 100 ml of broth. Sterilization of media was done by autoclaving at 15 psi for 20 min. When required appropriate concentrations of antibiotic and chemical stocks were added after cooling the media to 50°C-60°C.

2.1.1. Luria-Bertani (LB) media

The LB medium was made by dissolving 10 g peptone, 5 g yeast extract and 10 g NaCl per 1 L of deionised H₂O. The pH of the media was adjusted to 7.0 with 2 N NaOH and sterilized as described above.

2.1.2. Minimal salt medium for *Brevundimonas diminuta* MG

Minimal salt medium for *Brevundimonas diminuta* MG was prepared by dissolving 0.5 g of (NH₄)₂HPO₄ and 0.1 g of K₂HPO₄ in 1L of deionised H₂O and sterilized by autoclaving. After sterilization the solution was cooled to room temperature and independently sterilized stock solutions of 1% FeSO₄·7H₂O, 1% Ca (NO₃)₂, 20% MgSO₄·7H₂O were added to obtain final concentrations of 0.001% FeSO₄·7H₂O, 0.001% Ca (NO₃)₂, 0.01% MgSO₄·7H₂O. This constituted the complete minimal salt medium. When required, filter-sterilized carbon source and antibiotics were added at appropriate concentration to the complete minimal medium.

2.1.3. SOC media

The SOC media was made by dissolving 2 g tryptone, 0.5 g yeast extract, 0.5 g NaCl and 2.5 mM KCl per 100 ml of milli-Q water. The pH was adjusted to 7.0 and the solution was sterilized by autoclaving for 20 min at 15 psi. After cooling and just before use, 10 mM sterile autoclaved MgCl₂ and 20 mM filter sterilized glucose was added to the SOC medium.

2.2. Antibiotic & Chemical stock solutions

All the antibiotic chemical stock solutions were filter-sterilized using 0.22 µm syringe filter and stored as aliquots in - 20°C until use. The following table shows the stock and working concentration of routinely used antibiotics and chemicals.

Table 2.4 Preparation of Antibiotics

Name of Antibiotic	Solvent	Concentration (Stock)	Concentration (Working)
Ampicillin	Milli-q	100 mg/mL	100 µg/µL
Chloramphenicol	70% ethanol	30 mg/mL	30 µg/µL
Gentamycin	Milli-q	20 mg/mL	20 µg/µL

Kanamycin	Milli-q	30 mg/mL	30 µg/µL
Polymixin B	Milli-q	10 mg/mL	10 µg/µL
Streptomycin	Milli-q	20 mg/mL	20 µg/µL
Tetracycline	Milli-q	10 mg/mL	10 µg/µL
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Milli-q	1M	1 mM or 0.5 mM
X-Gal	N, N'-dimethylformamide	20 mg/ml	20 µg/µL

2.3. DNA methods

2.3.1. Preparation of Buffers and Solutions

2.3.1.1. Tris-Acetate- EDTA (TAE) buffer

A stock solution of 50x TAE buffer was using 242 g of Tris base, 100 ml of 0.5 M EDTA (pH 8.0) and 57.1 ml of glacial acetic acid dissolved in 900 ml of deionised H₂O. After complete dissolution of the contents, the solution was made up to 1L. When required, appropriate volume of 50x TAE was diluted to obtain 1x TAE.

2.3.1.2. Agarose gel loading buffer (6x)

Gel loading buffer was prepared by dissolving 0.25 g bromophenol blue, 0.25 g xylene cyanol FF and 15 g of Ficoll in 75 ml of deionised H₂O. The solution was finally made up to 100ml and stored at room temperature.

2.3.1.3. Ethidium bromide

A stock solution of was prepared by adding 10 mg ethidium bromide to 10 ml of deionised water in amber bottle and kept on magnetic stirrer overnight. From this stock solution, appropriate amount (5 µl) was added to 100 ml of gel solution to get a final concentration of 0.5 µg/ml in the agarose gel.

2.3.2. Solutions for plasmid isolation

2.3.2.1. Solution I (TEGL Buffer)

Solution I consist of 50 mM glucose, 25 mM Tris pH 8. 0, 10 mM EDTA dissolved in 25 ml deionised H₂O to final volume of 50 ml and sterilized by autoclaving. Whenever required

appropriate amount (10 µl) of stock solution of DNA free RNase A (10 mg/ml) was added to get a final concentration of 100 µg/ml in the solution.

2.3.2.2. Solution II (Lysis Buffer)

Freshly prepared equal volumes of 0.4 N NaOH and 2% SDS solution was used as Solution II.

2.3.2.3. Solution III (3M CH₃COONa, pH 4.8)

To prepare Solution III, 60 ml of 5 M CH₃COONa was prepared in deionised water and its pH was adjusted to 4.8 with glacial acetic acid. The volume was made to 100 ml and stored at 4°C.

2.3.2.4. Phenol: chloroform solution

Phenol chloroform solution was prepared by mixing 1:1 ratio of water saturated phenol and chloroform in amber bottle. The solution is stored at room temperature.

2.3.2.5. Chloroform: Isoamyl alcohol solution

Chloroform and Isoamyl alcohol were mixed in 24:1 ratio and stored at 4°C.

2.3.2.6. TE buffer

TE buffer was prepared by adding 10 mM Tris-Cl pH 8.0, 1 mM EDTA to 80 ml of deionised H₂O and finally made upto 100 ml. RNase A was added at final concentration of 20 µg/ml and stored at 4°C.

2.3.3. Isolation of plasmid DNA using Alkaline lysis method

Plasmid isolation was done by alkaline lysis method following standard procedures (Birnboim & Doly, 1979). A single colony containing desired plasmid was inoculated into 3 ml LB medium with required antibiotic and incubated for 12 hr at 37°C with shaking. Bacterial cells were harvested from 1 ml of overnight culture by centrifuging the culture at 13,000 rpm for 1 min at 4°C. The pellet was then resuspended by vigorous vortexing in 100 µl of TEGL buffer. Resuspended cells were lysed with 200 µl of Solution II and incubated on ice for 4-5 min before adding 150 µl of Solution III. The contents were then thoroughly mixed by inverting the tube and placed on ice for 3-5 min. The contents were centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was then transferred into a fresh eppendorf tube and extracted with equal volumes of phenol-chloroform and later with chloroform-isoamyl alcohol. The aqueous phase was added into new eppendorf tube and the plasmid DNA was precipitated by adding 1/10th volume of Solution III and double the volume of ethanol and incubated at -

80°C for 30 min. The plasmid was collected by centrifugation at 13,000 rpm for 20 min at 4°C. The traces of salts in the plasmid were removed by washing with cold 70% ethanol. The plasmid was further dried and redissolved in 100 µl of TE buffer.

2.3.4. Purification of plasmids using Miniprep kit method

Plasmid DNA was purified using Thermo Scientific GeneJET plasmid Miniprep kit as per manufacturer's protocol. Briefly, the overnight culture obtained from growing a single bacterial colony carrying plasmid was pelleted by centrifugation at 13,000 rpm for 1 min. The pellet was resuspended in 250 µl of resuspension buffer and was lysed by adding 250 µl of lysis buffer prior to mixing the tubes by inverting 4-6 times. After lysis of the cells the contents were neutralized by adding 350 µl of neutralization buffer prior to mixing of the contents by inverting immediately. Then tubes were centrifuged at 13,000 rpm for 10 min to pellet down the cell debris. After centrifugation the supernatant was transferred into a Thermo Scientific GeneJET spin column placed in a collecting tube. The supernatant was allowed to pass through the column for 1 min by centrifuging at 13,000 rpm. The column was washed with 750 µl of wash buffer. Finally, the DNA was eluted from column by adding 50 µl elution buffer, followed by brief centrifugation at 13,000 rpm. The plasmid DNA was stored at -20°C until further use.

2.3.5. DNA Quantification

Nucleic acids were quantified spectrophotometrically using NanoDrop® ND-1000 system. For a 1 mm path length cuvette OD at 260 nm equals 1 for 50 ng-cm/µl solution of double-stranded DNA. About 1 µl of sample was pipetted on the sample pedestal and the absorbance peak was measured at 260 nm for DNA. The purity of DNA was measured as ratio of 260/280 nm (A good purity ranges from 1.8-2.0). This was repeated for each sample.

2.3.6. Polymerase Chain Reaction (PCR)

PCR amplification reactions were carried out in a 25 µl reaction volume containing 2.5 mM MgCl₂, dNTP mix containing 200 µM each of dATP, dCTP, dGTP and dTTP mix, 10 pmol of each forward primer and reverse primer, 1 U of DNA polymerase, 10-20 ng of plasmid or 40-80 ng of genomic DNA as template. Amplifications were carried out in the Bio-Rad C1000

Touch™ thermal cycler by adjusting the PCR programme as per the amplicon size and melting temperature (T_m) of the primers. Amplicons were analysed on a 0.8- 1.0% agarose gel.

2.3.7. Agarose gel electrophoresis

A slab of 0.8% agarose was prepared by adding 0.8 g of agarose in 100 ml of 1X TAE. The contents were boiled till the suspended agarose particles got completely melted. This molten agarose was allowed to cool to approximately 65°C before adding 5 µl of ethidium bromide (EtBr). This mixture was poured into the casting unit inserted with 1 mm thick comb, and allowed to solidify. The solidified gel was then kept in electrophoresis tank and filled with 1X tank-buffer (~300 ml) keeping the wells towards anode. DNA samples mixed with loading dye were loaded into the wells and separated by electrophoresis applying constant voltage (100 V). The molecular size marker was loaded adjacent to the wells used to load DNA samples to measure the size of unknown DNA.

2.3.8. Molecular Cloning

2.3.8.1. Restriction Digestion

Restriction digestion was carried out with the required Thermo-Scientific FastDigest® restriction endonuclease (1 U) in a volume of 20 µl along with 2 µl of 10X Thermo-Scientific FastDigest® reaction buffer and 1 µg DNA. The reaction mixture was incubated for 1h at 37°C and the complete digestion of DNA fragments were ascertained by analysing them on agarose gel.

2.3.8.2. Gel extraction

After separation of DNA fragments generated either by restriction digestion, or by PCR amplification; agarose gel piece containing DNA fragment was extracted by using GeneJET gel extraction kit following manufacturer's protocols. Briefly, gel piece containing desired DNA fragment was excised using a sterile scalpel and weight of the gel piece was measured by placing it in a clean eppendorf tube. The gel slice was then mixed with equal volume (w/v) of binding buffer and the mixture was incubated at 50-60°C for 10 min or until the gel piece solubilized completely. The solubilized gel mixture was transferred on to GeneJET purification column and centrifuged at 13,000 rpm for 1 min. To it, 700 µl of wash buffer was added and centrifuged again 13,000 rpm till 1 min. In a clean microfuge tube the column was then placed

and 50 µl of elution buffer was added and centrifuged at 13,000 rpm for 1min. The eluted DNA was stored at -20°C until use.

2.3.8.3. DNA Ligation

The restriction digested vector and insert DNA were subjected to ligation by using Thermo-Scientific® T4 DNA ligase and incubated at 22°C for at least 1 hr as per to the manufacturer's protocol. About 100 ng of total DNA with a molar ratio (vector:insert) of 1:3 was used for ligation reaction. The following equation was used to calculate the appropriate amount of insert DNA for ligation.

$$mass_{insert}[ng] = \frac{mass_{vector}[ng] \times size_{insert}[bp]}{size_{vector}[bp]} \times 3$$

2.3.8.4. Colony PCR

Colony PCR was made to screen colonies containing insert in recombinant plasmids. Briefly, a single colony containing recombinant plasmid was carefully lifted with the help of a sterile tooth pick and resuspended the cells in TE buffer. The contents were then boiled for 5-10 min to ensure the lysis of the cells. The samples were then cooled and centrifuged at 13,000 rpm for 5 min. The supernatant (1 µl) obtained from each colony served as DNA template in a 10 µl colony PCR reaction that was set using EmeraldAmp® GT PCR Master Mix at 1X concentration with primers added at appropriate concentration. The reactions were carried on a Bio-Rad C1000 Touch™ thermal cycler with cycling parameters set depending on the amplicon size and melting temperature (T_m) of the primers. After PCR the samples were analysed on a 0.8 % agarose gel to ascertain the presence of insert in the colonies containing recombinant plasmid.

2.4. Gene transfer methods

2.4.1. Preparation of ultra-competent E coli cells

A single colony of E. coli DH5α was picked from a freshly streaked overnight LB plate and inoculated into 10 ml conical flask. The culture was incubated for 10-16 hrs at 37°C with vigorous shaking at 200 rpm. The above culture (1%) was used to inoculate fresh LB broth (250 ml) and incubated at 37°C with moderate shaking at 150 rpm till the cells reached mid-log phase (0.4-0.5 OD₆₀₀). The cells were collected by centrifugation at 3500 rpm for 10 min at 4°C. The cell pellet obtained was resuspended in 75 ml of ice-cold 80 mM MgCl₂-20 mM CaCl₂ solution by gently swirling the tube on ice. The cell suspension obtained was incubated on ice for 30 min and were subsequently centrifuged again at 3500 rpm for 10 min at 4°C.

Finally, the pellet obtained was resuspended in 10 ml of 100 mM CaCl_2 containing 15% glycerol. Quickly, these cells were distributed as 100 μl aliquots into sterile eppendorfs, snap freezed in liquid nitrogen and stored at -80°C until further use (J.F & Russell, 2001).

2.4.2. Transformation

The ultra-competent cells were initially thawed on ice for 2 min. The ligation mixture/plasmid of interest was added to the cells and incubated on ice for 30 min. After 30 min, the cells were subjected to heat shock at 42°C for exactly 90 sec and immediately chilled on ice for 2 min. Further, 750 μl of LB broth was added and incubated at 37°C for 1 hr. The cells were collected by centrifugation and resuspended in 100 μl of LB broth and plated on required antibiotic selective LB media. When required X-gal solution (20 $\mu\text{g}/\text{ml}$) and IPTG (1 mM) was added along with required antibiotic. The plates were then incubated at 37°C for 12-16 hr for colonies to appear.

2.4.3. Electroporation

Cells were harvested from mid-log phase cultures and made electro-competent by washing and incubating them with 10% ice-cold glycerol. About 1-2 μg of DNA was added to 100 μl of electro-competent cells. The suspension was mixed by tapping the tube. The cells/DNA mixture was placed in a pre-chilled cuvette, between the electrodes and a pulse of 2.5 kV was applied for 4-5 sec using Gene Pulser Xcell™ (Bio-Rad, USA). Following the pulse, immediately 1ml of sterile SOC medium was added to the cells. The mixture was transferred to sterile 1.5 ml eppendorf tube and incubated at 37°C for 1 hr with constant shaking. After incubation, the cells were diluted appropriately in SOC medium and plated on LB agar plates containing appropriate antibiotics.

2.5. Protein Methods

2.5.1. Solutions for SDS-PAGE

2.5.1.1. Acrylamide stock solution

Stock solution of acrylamide (30%) was prepared by mixing acrylamide (30 g) and N, N'-methylene-bis-acrylamide (0.8 g) in 70 ml of distilled H_2O in an amber bottle. After the contents were dissolved completely the volume of the solution was made up to 100 ml and stored at 4°C .

2.5.1.2. Resolving gel buffer

Tris-base (181.71 g) was dissolved in 800 ml of distilled H₂O and pH was made upto 8.8 by adding 30 ml of concentrated HCl. Then the volume was made up to 1000 ml with distilled H₂O. When required, SDS was added to the buffer to a final concentration of 0.3% by mixing adequate amount of 10% SDS stock solution. The resolving gel buffer was filter-sterilized and stored at room temperature until further use.

2.5.1.3. Stacking gel buffer

Tris-base (121.14 g) was dissolved in 800 ml of distilled H₂O and the pH was made upto 6.8 by adding 42 ml of conc. HCl and finally the volume of the buffer was made up to 1000 ml with distilled H₂O. This solution was autoclaved and allowed to cool to room temperature. When required SDS was added to the buffer to get a final concentration of 0.4% by mixing adequate amount of 10% SDS stock solution. The prepared stacking gel buffer was filter sterilized and stored at room temperature until further use.

2.5.1.4. 10x Tank buffer

30 g Tris base, 140 g Glycine and 10 g SDS were weighed and added to 700 ml double distilled water to prepare a 10X stock solution of Tris glycine buffer. After dissolving the contents, the final volume of the solution was adjusted to 1000 ml by adding water and stored at room temperature until subsequent use. When required the 10X stock was diluted to 1X using Milli-Q water which contains Tris base, Glycine and SDS at a concentration of 25 mM, 250 mM and 0.1% respectively.

2.5.1.5. 2x SDS gel loading buffer

Stock solution of 2x SDS loading buffer was prepared by mixing 1 ml of 1 M Tris-Cl pH 6.8, 2 g SDS, 10 mg of bromophenol blue and 10ml of glycerol in 25 ml of distilled H₂O and kept at 45°C for 10 min for complete solubilization of SDS. To this buffer 0.699 ml of β-mercaptoethanol was added and finally made up to 50 ml with distilled H₂O. The loading dye was then distributed in 10 ml aliquots and stored at -20°C. While running protein samples on SDS-PAGE, equal volumes of buffer and protein samples were mixed and boiled for 5 min, cooled and centrifuged at 13000 rpm for 1 min before loading on to the gel.

2.5.1.6. Staining solution

The staining solution was prepared by dissolving 0.25 g of Coomassie (R-250) in 50 ml of methanol. After dissolution, 15 ml of acetic acid was added and finally the volume was made up to 100 ml using distilled H₂O.

2.5.1.7. Destaining solution

Methanol (30 ml) was mixed with 10ml of glacial acetic acid before making up the volume to 100 ml using distilled H₂O. The destained gels were stored in 7% acetic acid solution.

2.5.2. Preparation of SDS-PAGE gels and Electrophoresis

The SDS-PAGE gels were prepared using MINI PROTEAN II™ Bio-Rad system. Initially, thin glass plates and the spacer plates of the required thickness (0.75, or 1.0, or 1.5 mm) were assembled into the casting unit. Then, 12.5% -15% resolving gel mixture was prepared and poured between glass plates leaving enough space for adding stacking gel. This was quickly followed by layering of resolving gel with water saturated n-butanol to remove air bubbles. The resolving gel was then left at room temperature for 30 min for polymerization. After polymerization of the resolving gel, the water saturated butanol layer was removed and the gel surface was washed 2-3 times using distilled water. Subsequently, 4.5 % stacking gel was prepared and layered on top of resolving gel and comb of appropriate thickness (0.75, or 1.0, or 1.5 mm) was placed quickly before polymerization to facilitate formation of wells. After polymerization of the stacking gel, the comb was slowly removed and the wells were cleaned with 1X tank buffer to remove acrylamide particles. The gels formed were then placed into the electrophoresis tank and the 1X tank buffer was poured upto the mentioned level in the running unit. The protein samples were prepared by mixing with equal volume of 2X SDS sample dye and boiled for 5-10 min in hot water bath. Samples were given a short spin and before loading in separate wells alongside sample containing molecular weight markers. The gel was run at a constant voltage of 120 V for 1-2 hr or until the dye front has reached the bottom of the gel. After the electrophoretic run, the gel was removed from the plates, stained with Coomassie stain for 1-2 hr and destained using destaining solution until gel background was clear. The results were documented by scanning the gel on a gel scanner.

Table 2.5 Components of a separating gel

COMPONENT	12.5 % gel (5ml)	15 % gel (5ml)
Acrylamide (30 %) (ml)	2	2.5
Solution II (1.5 M Tris, pH 8.8) (ml)	1. 25	1. 25

Ammonium per sulphate (APS) (10%) (μL)	50	50
TEMED (μL)	5	5
H ₂ O (ml)	1. 7	1. 25

Table 2.6 Components of a stacking gel

COMPONENT	Volume for 4.5 % gel (2. 5ml)
Acrylamide (30 %) (ml)	0. 375
Solution III (0.5 M Tris, pH 6.8) (ml)	0. 625
Ammonium per sulphate (APS) (10 %) (μl)	25
TEMED (μL)	2. 5
H ₂ O (ml)	1. 25

2.5.3. Solutions for Western Blotting

2.5.3.1. Protein transfer buffer (Towbin buffer)

Towbin buffer was prepared by dissolving Tris-base (3.03 g) and glycine (14.4 g) in 500 ml of distilled H₂O. After dissolving the contents, 200 ml of methanol was added and volume was made up to 1000 ml with distilled H₂O.

2.5.3.2. TBS-T Buffer

TBS-T buffer was prepared by adding Tris-HCl 1 M, pH 7.6 (20 ml), sodium chloride (8 g) and Tween-20 (1 ml) and the volume was made up to 1000 ml with distilled H₂O.

2.5.3.3. Blocking Reagent

The solution of skimmed milk powder (10%) was prepared in TBS-T buffer and was used for blocking the PVDF membrane.

2.5.3.4. Ponceau S reagent

Ponceau salt (100 mg) was dissolved in 100 ml of 5% acetic acid.

2.5.3.5. Membrane stripping solution

Membrane stripping solution was prepared by dissolving 1.5 g of glycine, 0.1 g of SDS and 1 ml of Tween-20 in 100 ml of distilled H₂O. The pH of this solution was adjusted to 2.2 with 6 N HCl.

2.5.3.6. Primary antibody solution

Primary antibody solution (2 µl) of anti-His mouse antibody (Amersham Biosciences) was added to 10 ml of blocking reagent to get a titre of 1:5000 and used for probing the proteins transferred onto PVDF membranes.

2.5.3.7. Secondary antibody solution

The secondary antibody solution of anti-mouse goat antibody (2 µl) conjugated with HRP was added to 10 ml of blocking reagent to get a titre of 1:5000 and used as secondary antibody solution.

2.5.4. Semi-Dry Western Blotting

After SDS-PAGE based separation of the protein samples, the polyacrylamide gels were soaked in transfer buffer for 5 min. In the meantime, PVDF membranes (Amersham Hybond-ECL, GE healthcare) were pre-treated by soaking in methanol for a 5 min followed by transfer buffer. Two layers of pre-soaked filter paper (in transfer buffer) were placed on the positive electrode of the transfer apparatus (Trans-Blot SD Semi-dry Transfer Cell™, Bio-Rad, USA) and the transfer membrane and the gel were sandwiched between two additional layers of filter paper on top. The transfer membrane was facing the negative electrode at the bottom of the transfer apparatus, while the gel was placed on top. The apparatus was assembled according to the instructions by the manufacturer and the transfer was performed at 18 V for 40 min in the transfer apparatus. After transfer of the proteins, the membrane was blocked in 10% skimmed milk suspension in TBST at room temperature for 1 hr, or at 4°C overnight, with shaking at 60 rpm. After blocking, the membrane was incubated with primary antibody solution using an appropriate dilution for 1 hr at room temperature with gentle shaking. The membrane was washed with TBST for 10 min and, if appropriate, incubated with the secondary antibody at the correct dilution for 1 hr with shaking. The membrane was finally washed three to five times with TBST and developed with Clarity Western™ ECL Western Blotting detection kit (Bio-Rad). The resulting chemiluminescence was captured by ChemiDoc™ Touch imaging system (Bio-Rad).

2.5.5. Solutions for Protein Estimation

2.5.5.1. Standard Bovine Serum albumin (BSA)

Stock solution (10 mg/ml) of BSA was prepared by dissolving 10 mg of BSA in 0.17M NaCl. This solution was stored at -20°C.

2.5.5.2. Bradford's reagent

Bradford's reagent was prepared by dissolving 10 mg of Coomassie brilliant blue G-250 in 50 ml of 95% ethanol and adding 10 ml of 85% (w/v) orthophosphoric acid. After complete solubilization of the dye the solution was made up to 100 ml with distilled H₂O. Finally, the reagent was filtered and stored in an amber bottle at 4°C until further use. While preparing standard graph appropriate amounts of BSA stock solution was added to various tubes containing 1 ml of Bradford's reagent to get a final concentration of BSA in the range of 2, 4, 6, 8, 10 µg. The contents were then adjusted to 1.1 ml by adding adequate amounts of distilled H₂O before storing the tubes in dark for 10 min. After completion of incubation period the OD of the reaction mix was measured at 595 nm and the readings were used to prepare a standard graph was prepared by taking concentration of BSA on X-axis and OD values on Y-axis. The protein concentration in the unknown sample was measured by treating the sample in a similar manner and comparing its OD values with that the values obtained for BSA standards.

2.5.6. Protein Precipitation

The protein sample was mixed with double the sample volume of methanol, one volume of chloroform and three volumes of distilled H₂O. The precipitated sample was centrifuged at 13,000 rpm for 5 min and the upper aqueous layer was discarded. The interphase containing the protein sample was vortexed vigorously with three sample volumes of methanol. The sample was centrifuged again at 13,000 rpm for 5 min. The supernatant was carefully removed and the pellet was air dried. The pellet obtained was finally resuspended in 50 mM potassium phosphate buffer (pH 7.5) or any other buffer of choice.

2.6. Enzyme Assays and Preparation of reagents

2.6.1. Dehalogenase assay (Phenol-Red Method) reagents

2.6.1.1. HEPES-NAOH Buffer (50 mM, pH 8.0)

Stock solution of 50 mM HEPES-NaOH buffer was prepared by adding 1.91 g of HEPES to about 40 mL of distilled H₂O. The pH of the solution was adjusted to 8.0 using 10 N NaOH

and the volume was made upto 100 mL. The solution was filtered with 0.45-micron filter and stored at 4°C until use.

2.6.1.2. EDTA (50 mM, pH 8.0)

Stock solution of EDTA was prepared by dissolving 1.46 g of EDTA in distilled water adjusting the pH to 8.0 and the volume was made upto 100 mL. The solution was filtered and stored at RT.

2.6.1.3. Sodium Sulphate (100 mM)

Stock solution was prepared by adding 1.42 g of Na₂SO₄ in 80 ml of distilled water. After dissolving the salt, the final volume was made up to 100 ml.

2.6.1.4. Phenol red (1 mg/mL)

The stock solution of Phenol red was prepared by dissolving 10 mg of Phenol red in 10 mL of 100% ethanol.

Table 2.7 Components of Phenol Red Dehalogenase Assay

COMPONENTS	TEST	CONTROL 1 (NO cells/protein)	CONTROL 2 (No substrate)
50 mM HEPES-NaOH buffer	20 µL	20 µL	20 µL
50 mM EDTA	20 µL	20 µL	20 µL
100 mM Sodium sulphate	100 µL	100 µL	100 µL
10 mg/ml Phenol Red	20 µL	20 µL	20 µL
H ₂ O	810 µL	830 µL	820 µL
Cells/Pure Protein*	20 µL	-	20 µL
Halogenated Substrates	10 µL	10 µL	-

*At appropriate concentration

2.6.2. Dehalogenase Activity assay (Phenol-Red Method)

The dehalogenase assay was performed following standard procedures (Holloway, Trevors, & Lee, 1998). The typical dehalogenase assay test reaction sample (190 µl) contained 20 µl of 50 mM HEPES (1 mM), 20 µL of 50 mM EDTA (1 mM), 100 µl of 100 mM sodium sulphate (20 mM), 20 µl of cell extract/purified protein at the appropriate concentration and 20 µl of 10 mg/ml phenol red (20 µg/mL) in each well of 96 well ELISA Microplate (Nunc)

microtitre plate. The reaction was initiated by adding 10 μ l of 100 mM halogenated substrate (1 mM) to the reaction sample. The control reactions sample contained the same reaction mixture but without substrate (substrate -ve control) or enzyme (enzyme -ve control). The samples were then incubated for specific time points at 30°C. Dehalogenase activity generated corresponding alcohols, halide ions and protons that caused reduction of pH of the weakly buffered HEPES solution. The change in pH was visualised by a colour change from red to yellow using absorbance at $\lambda=540$ nm in a MULTISKAN™ GO microplate spectrophotometer, Thermo-Scientific, USA. The reaction was quantified from a standard curve prepared as mentioned below. The experiment was performed in triplicates for each substrate. To calculate the dehalogenase activity at different pH, standard graphs were prepared using HEPES-NaOH buffered solution prepared at each of the pH ranging from 6.0-9.0. The specific enzyme activity was calculated as the amount of protein required to release 1 nmol of H^+ ions per min.

2.6.3. Standard curve

To produce a standard curve the reaction sample without enzyme and substrate was mixed with appropriate volume of 1 M HCl to obtain final concentrations between 0-10 mM in a total volume of 200 μ l. Using this standard curve, the concentration of protons released in the reaction were quantified.

2.6.4. Dehalogenase assay (Iwasaki method) reagents

2.6.4.1. Glycine buffer (0.5 M, pH 9.5)

The stock solution of Glycine buffer was prepared by dissolving 3.75 g of glycine in distilled water. The pH of the solution was adjusted to 9.5 using 10N NaOH and the volume was made up to 100 ml. The solution was filtered using 0.45-micron filter and stored at 4°C until use.

2.6.4.2. Mercury (II) thiocyanate Solution (Sol I)

This stock solution of Mercury (II) thiocyanate $Hg(SCN)_2$ was prepared by dissolving 0.3 g of salt in 100 ml of 100% ethanol. The solution is highly toxic and is prepared and disposed following Material Safety Data Sheet (MSDS). The solution was stored in a tightly closed container in well ventilated space under ambient temperature.

2.6.4.3. Ferric Ammonium Sulphate (FAS) Solution (Sol II)

About 12.32 g of Ammonium iron (III) sulphate, $\text{NH}_4\text{Fe}(\text{SO}_4)_2 \cdot 12 \text{H}_2\text{O}$ or ferric ammonium sulphate (FAS) or iron alum is dissolved in 128 mL distilled water. The solution is also used to stop the enzymatic reaction by including 72 mL of 70% nitric acid (HNO_3). The final volume of the solution was 200 mL. The solution was stored at RT.

2.6.4.4. KCl, KBr or KI standard solutions

The standard stock solutions of 50 mM are prepared by adding 0.476 g (KCl), or 0.297 g (KBr) or 0.41 g (KI) in 50 mL of milli-Q water. The dilutions were prepared from 10 nM -10 mM and calibration curve was prepared for volume of 200 μL for every concentration.

Table 2.8 Components for Dehalogenase Assay

COMPONENT	TEST	CONTROL 1	CONTROL 2
		(NO cells/protein)	(No substrate)
0.5M Glycine-NaOH buffer, pH 9.5	100 μL	100 μL	100 μL
H ₂ O	790 μL	890 μL	800 μL
Cells/Pure Protein*	100 μL	-	100 μL
Halogenated Substrates (1M)	10 μL	10 μL	-

*At appropriate concentration

2.6.5. Dehalogenase assay (Iwasaki Method)

Both kinetic (phenol-red) and end point quantitative and colorimetric methods were performed to assay dehalogenase activity. The Iwasaki assay was performed following standard procedures (Iwasaki, Utsumi, & Ozawa, 1952). In order to avoid false positive results, cells and protein samples prepared were made free of any halide ions. Initially 890-900 μL of 100 mM glycine solution was preincubated with 100 μL of crude protein extract/pure protein at 30°C. Then the reaction was started by adding 10 μL of substrates (10 mM). The contents were thoroughly mixed by vortexing before incubating the contents at 30°C. At appropriate time points (between 0 to 30 min) 200 μL of sample was withdrawn from the reaction mix and 20 μL of solution I ($\text{Hg}(\text{SCN})_2$) was added. The sample was again vortexed for 30 sec before adding 40 μL of solution II ($\text{NH}_4\text{Fe}(\text{SO}_4)_2 \cdot 12 \text{H}_2\text{O}$). After brief vortex the contents were

centrifuged at 13000 rpm for 5 min to remove the insoluble particles. The halide ions released due to dehalogenase enzyme activity were quantified by indirectly measuring the concentration of thiocyanate ions formed due to dissociation of $\text{Hg}(\text{SCN})_2$ after interacting with halide ions. The orange-coloured ferric thiocyanate complex formed due to interaction of thiocyanate with ferric ions were measured at $\lambda = 450 \text{ nm}$. While performing multiple reactions using different substrates about 200 μl of the supernatant was taken in a 96 well ELISA Microplate (Nunc) and absorbance was measured at 450 nm using a MULTISKAN™ GO microplate spectrophotometer from Thermo-Scientific, USA. The concentration of chloride, bromide or iodide was calculated from standard graph generated using known amounts of KCl, KBr or KI respectively. The specific enzyme activity was calculated as the amount of protein required to release 1 nmol halide ion per min.

2.6.6. Substrates stock solutions

The halogenated substrates and their concentrations used in this study are listed in the table given below. All the substrates were procured from Sigma-Aldrich, USA unless otherwise mentioned. The solutions were freshly prepared in acetonitrile before use.

Table 2.9 List of halogenated substrates with concentration

S.No	Halogenated Substrates	Stock concentration	Working concentration
1.	γ -Hexachlorocyclohexane (HCH)	34.3 mM	34.3 μM
2.	1-Iodohehexane	1 M or 100 mM	10 mM or 1 mM
3.	1-Iodobutane	1 M or 100 mM	10 mM or 1 mM
4.	1-Iodopropane	1 M or 100 mM	10 mM or 1 mM
5.	1,3-Diiodopropane	1 M or 100 mM	10 mM or 1 mM
6.	Bromocyclohexane	1 M or 100 mM	10 mM or 1 mM
7.	1-Bromohexane	1 M or 100 mM	10 mM or 1 mM
8.	4- Bromobutyronitrile	1 M or 100 mM	10 mM or 1 mM
9.	1-Bromobutane	1 M or 100 mM	10 mM or 1 mM
10.	1,2-Dibromopropane	1 M or 100 mM	10 mM or 1 mM

11.	1- Chlorohexane	1 M or 100 mM	10 mM or 1 mM
12.	1-Chlorobutane	1 M or 100 mM	10 mM or 1 mM
13.	1,2- Dichloropropane	1 M or 100 mM	10 mM or 1mM
14.	1,2- Dichloroethane	1 M or 100 mM	10 mM or 1mM

2.6.7. Epoxide hydrolase assay (Adrenaline Test) reagents

2.6.7.1. L-Adrenaline salt (1 M HCl) 10 mM stock solution

In a 10 ml volumetric flask, 18 mg of L-adrenaline ((-)-epinephrine) was added in 7 ml of water. Using a glass pipette, 98 μ l of HCl 37% was added to the volumetric flask. The solution was diluted to 10 ml. For storage, flask was sealed with a stopper and paraffin film and stored at 4°C for up to 8 weeks. The final concentration of L-adrenaline in the reaction is 1.5 mM.

2.6.7.2. Sodium periodate (10 mM)

About 21.4 mg NaIO₄ salt was dissolved in 10 ml distilled water. This solution is always freshly prepared and used immediately at 1 mM final concentration

2.6.7.3. Potassium Phosphate buffer (50 mM, pH 7.5)

Initially 1 M K₂HPO₄ was prepared by dissolving 87.09 g K₂HPO₄ in 500 ml of distilled water. Similarly, 1 M KH₂PO₄ was prepared by dissolving 68.05 g KH₂PO₄ in 500 ml distilled water. From these stocks, 42 mL of 1M K₂HPO₄ was mixed with 8 ml of 1 M KH₂PO₄ and the volume was made up to 1000 mL using distilled water. This solution now is 50 mM Potassium Phosphate buffer at pH 7.5. The solution was autoclaved and stored at RT until further use. Buffers like Tris, Bis-Tris and Triethanol amine are not suitable for this test. Also, glycerol which is used a stabilizer in enzyme preparation should be avoided as it gives false positive test.

2.6.7.4. Epoxides stock solutions (100 mM)

Epoxide substrate stocks (0.5 mmol) were dissolved in 5 ml of acetonitrile. Epoxide solutions were stored for up to 8 weeks at -20°C.

Table 2.10 Components for Epoxide hydrolase Assay

COMPONENT	TEST	CONTROL (NO cells/protein)
50 mM Phosphate buffer, pH7.5	70 μ L	70 μ L
Cells/Pure Protein*	10 μ L	-

Epoxides (100 mM)	10 μ L	10 μ L
NaIO ₄ (10 mM)	10 μ L	10 μ L
L-Adrenaline (10 mM)	15 μ L	15 μ L

*At appropriate concentration.

2.6.8. Epoxide hydrolase assay

The epoxides hydrolase activity was measured following standard protocols (Wahler & Reymond, 2002). Phosphate buffer (80 μ l) containing 1 mM of periodate (NaIO₄) was added to the wells of a 96 well microplate. In all these wells 10 μ l (1mg/mL) enzyme was added, the wells having no enzyme or substrate served as controls. The reaction was started after addition of 10 μ l substrate (various epoxides at 10 mM). The contents were then incubated at 30°C for 30 mins. When incubation process was in progress 15 μ L of 10 mM L-adrenaline stock solution was taken in a fresh microtiter plate. Soon after incubation 85 μ L of this solution reaction mix was taken and added to the wells containing L-adrenaline solution. The 1,2-diols formed due to epoxide hydrolase activity were oxidized by sodium periodate. The remaining unreacted periodate (NaIO₄) was estimated by titrating with standard solution of L-adrenaline. The red-coloured adrenochrome formed due to L-adrenaline interaction with periodate (NaIO₄) was used to indirectly quantify of the diols formed. The adrenochrome formed was measured at λ = 490nm using a MULTISKAN™ GO microplate spectrophotometer from Thermo-Scientific, USA.

2.6.9. Standard curve of diols

A standard curve was prepared by procuring diols listed in Table 2.10 to use while assaying epoxide hydrolase activity by using corresponding epoxides as substrates. Concentration of diols in the standard curve were ranging from 1 mM-10 mM. While preparing standard curve the reaction mixture contained sodium periodate in 100 μ l of phosphate buffer and various concentrations of diols. The standard reaction procedures described above was followed while preparing standard graph. The prepared standard graph was used to quantify diols formed from epoxides due to epoxide hydrolase activity.

Table 2.11 List of Epoxide substrates and diols used in this study

S. No	Epoxide Substrate	Stock concentration	Working concentration
1.	Styrene oxide	100 mM	10 mM
2.	1,2- Epoxy hexane	100 mM	10 mM
3.	1,2- Epoxybutane	100 mM	10 mM
4.	rac-Phenylethanediol	100 mM	Range from 1-10 mM
5.	rac- 1,2-hexanediol	100 mM	Range from 1-10 mM
6.	rac- 1,2-butanediol	100 mM	Range from 1-10 mM

CHAPTER 3 : Structure and Phylogeny of *Brevundimonas diminuta* MG genome

3.1. Introduction

Our lab has been using *Brevundimonas diminuta* MG as model organism to gain deeper understanding on the microbial metabolism of the organophosphate insecticides. In fact, *B. diminuta* was first isolated from sewage and soil samples contaminated with highly toxic and recalcitrant organophosphate (OP) insecticides (Munnecke & Hsieh, 1974). Based on its morphological and biochemical characteristics, initially the bacterium was classified as *Pseudomonas diminuta* MG (Serdar, Gibson, Munnecke, & Lancaster, 1982). Consequently, when the genus *Pseudomonas* was reclassified using polyphasic taxonomic approach, the organism was placed under the genus *Brevundimonas* and renamed *Brevundimonas diminuta* MG (Segers et al., 1994). *B. diminuta*, as a model organism, was used to understand the biology and genetics of organophosphate (OP) degradation (Gorla, Pandey, Parthasarathy, Merrick, & Siddavattam, 2009; Parapatla et al., 2020; Parthasarathy, Parapatla, Nandavaram, Palmer, & Siddavattam, 2016; Parthasarathy, Parapatla, & Siddavattam, 2017). It contains a large indigenous plasmid, pCMS1 (Pandeeti, Chakka, Pandey, & Siddavattam, 2011). The organophosphate degrading (opd) gene, coding organophosphate hydrolase (OPH) is localized on pCMS1. OPH initiates degradation of OP insecticides by hydrolyzing third ester linkage found in OP compounds having considerable structural diversity. *B. diminuta* also uses p-nitrophenol as sole carbon source, the toxic and recalcitrant degradation product generated from a number of OP insecticides after OPH activity. However, the complete degradation potential of this bacterium has not yet been explored due to unavailability of a complete genome sequence. In this chapter of the thesis complete, fully circularized genome sequence of *B. diminuta* MG is presented and its adaptive potential is discussed.

3.2. Objective specific methodology

3.2.1. Isolation of genomic DNA of *B. diminuta*

A 10 mL culture of LB medium containing polymixin B (10 µg/mL) was inoculated with a colony of *B. diminuta* and incubated 12-16 hr at 30°C and 180 rpm. One mL cells were pelleted by centrifugation at 13,000 g for 1 min. The isolation of genomic DNA was performed

using the Qiagen miniprep DNA isolation kit (Qiagen). The pellet was dissolved in 300 µL lysis solution in an eppendorf tube and incubated at 75°C for 5 min for complete cell lysis. Following this, RNase A (4 mg/mL) was added, tube was inverted for 30 secs and incubated at 37°C for 45 min. After incubation, 100 µL of Proteinase K Solution was added and sample was vigorously mixed for 20 sec and then centrifuged at 13,000 g for 5 min twice in succession to remove cellular debris. The supernatant containing DNA was precipitated by transferring it to a clean 1.5 mL tube containing 500 µL of 100% isopropanol. The precipitated DNA was collected by centrifugation at 8,000 g for 2 min. Further, the pellet containing DNA was washed with 500 µL 70% ethanol. After discarding supernatant, the residual ethanol was evaporated by drying the tubes inverted on a filter paper for 10-15 min. The pellet was resuspended in 100 µL 10 mM Phosphate buffer pH 7.5 at 65 °C for 10 min. The DNA quality was analysed on a 0.7% 1X TAE agarose gel, and the concentration was measured on a NanoDrop™ spectrophotometer.

3.2.2. Structural and functional annotation of genome

The protein coding genes were identified using three ab initio softwares the RASTtk pipeline (Brettin et al., 2015) Prodigal (Hyatt et al., 2010) of the Prokka pipeline (Seemann, 2014), Glimmer v3.02 (Salzberg, Delcher, Kasif, & White, 1998) and GeneMarkS v4.3 (Besemer, Lomsadze, & Borodovsky, 2001) along with a run of GenePRIMP v0.3 (Pati et al., 2010) for manual curation. The non-redundant databases from National Centre for Biotechnology Information (NCBI), UniProt, TIGRFam, Pfam, KEGG, COG and InterPro databases were used to search for the predicted CDS and combined data sources were used to allocate a gene product description for each of the predicted protein. The non-coding genes were predicted using RNAmmer (Lagesen et al., 2007) and ARAGORN (Laslett & Canback, 2004) for rRNA and tRNA respectively.

3.2.3. Similar genome search

The complete chromosome sequence of *B. diminuta* was considered and a meticulous workflow was followed to reveal the taxonomy and evolution of *B. diminuta*. Using MegaBLAST (Y. Chen, Ye, Zhang, & Xu, 2015) the representative genomes database downloaded from NCBI and whole genome of *B. diminuta* were subjected to similarity search.

Similar genomes were verified again using online MegaBLAST search and genomes with maximum sequence identity to the query sequence were ranked accordingly. Based on the results of MegaBLAST search, PATRIC (Wattam et al., 2014) database was used to download all the genomes belonging to Sphingomonadales to perform further analysis.

3.2.4. Orthology analysis

To identify homologous genes between *B. diminuta* genome and the other genomes of Sphingomonadales OrthoMCL v1.04 (L. Li, Stoeckert, & Roos, 2003) was run twice, once with lenient parameter of “Percent Match Cut-off at 0%”, “Percent Identity Cut-off at 0%” and “P-value Cut-off at 1×10^{-5} ”, on whole set of species. After calculating the total number of genes shared for each genome using perl scripts, Sphingomonadales dataset sharing less than 1500 genes were removed. A more stringent OrthoMCL was run with “Percent Match Cutoff at 80%”, “Percent Identity Cut-off at 40%” and “P-value Cut-off at 1×10^{-12} ” parameters on the curated dataset of Sphingomonadales. To find the core genome homology of the sequenced bacteria, the OrthoMCL was run for the third time using gene- sets of the *Sphingopyxis* genus and using *E. coli* as an out-group species in the analysis. The Shell scripts and the output file Perl scripts used at each stage was developed and processed in the lab of Dr Sarwar Azam, National Institute of Animal Biotechnology (NIAB), India.

Table 3.1 General genome characteristics of *Sphingopyxis* strains used in Chapter III

S. No	Organism	Accession Number	Source	Genome Size(Mbp)	Coverage
1.	<i>Sphingopyxis</i> sp MC1	NZ_AOUN01000000	Activated sludge from a waste water, USA	3.65	50x
2.	<i>Sphingopyxis</i> sp LC363	NZ_JNFC00000000	Cave surface layer, Lechuguilla Cave, New Mexico	4.21	28x
3.	<i>Sphingopyxis</i> sp. LC81	NZ_JNFD01000000	Cave surface layer, Lechuguilla Cave, New Mexico	4.39	104x
4.	<i>Sphingopyxis. fribergensis</i> Kp5.2	NZ_CP009122.1 NZ_CP009123.1	Soil sample Freiberg, Germany	5.2	210x
5.	<i>Sphingopyxis. alaskensis</i> RB2256	NC_008048.1 NC_008036.1	Marine heterotroph, Resurrection Bay, North Pacific	3.37	-
6	<i>Sphingopyxis</i> sp. MWB1	QFJ000000000	Crude oil contaminated seashore, South Korea	3.11	1046x
7.	<i>Sphingopyxis baekryungensis</i> DSM16222	NZ_ATUR00000000.1	Sea water Yellow Sea, Korea	3.07	-
8.	<i>Sphingopyxis</i> sp MG	NZ_CP026381.1, NZ_CP026382.1	Sewage and soil samples, USA	4.22	1361x

3.2.5. Phylogenomic and Phylogenetic analyses

For phylogenetic analysis, initially all single copy orthologs (1:1 ratio) present in at least 95% of the species were selected using customised Perl Scripts. Further, each orthologous group was independently aligned in multiple sequence alignment using MAFFT v7. 123b (Katoh & Standley, 2013). Furthermore, after removal of alignment columns with ambiguous sequence or gaps using G-Blocks (Castresana, 2000), the remaining alignments were concatenated using FASconCAT v1. 0 (Kuck & Longo, 2014). The refined alignments were then submitted for maximum likelihood phylogenetic analysis using PROTGAMMAWAGF model of substitution. The entire process was repeated using orthologous genes of only *Sphingopyxis* genus using PROTGAMMALG method of substitution and dendrograms were plotted with Dendroscope.

3.2.6. Positive selection analyses

Single copy orthologs of all the *Sphingopyxis* genus were aligned for constructing codon alignment using PAL2NAL (Suyama, Torrents, & Bork, 2006). Gene conversion events

between the aligned sequences was determined using GENECONV (Sawyer, 1989) and sequences with Global-P value greater than 0.05 were accepted. Phylogenetic tree was constructed using GTRGAMMA model of substitution. CodeML of the PAML package was used to compute ω and positively selected genes were identified with branch-site model where positive selection along the branch (selection test among lineages) and site (selection estimation at codon level within a protein sequence) were estimated by earmarking our genome sequence in foreground branch and other genomes in background branch. From null and alternate models of each set Likelihood Ratio test was performed. Multiple hypothesis testing was done using Q value package of R. To filter positively selected genes false discovery rate (FDR) cut off was set at 1% (0.01) level of significance.

3.2.7. Genomic Island prediction

The alignment tool Mauve v2.3.1 (Darling, Mau, Blattner, & Perna, 2004) was used to align two closely related genomes to predict the genomic islands. The predicted genomic islands were further verified using IslandViewer v3 (Langille & Brinkman, 2009) that uses an integrated approach from different softwares namely SIGI-HMM (Merkel, 2004) based on codon usage, IslandPath-DIMOB (Bertelli & Brinkman, 2018) which considers dinucleotide sequence composition bias and presence of mobile genes and IslandPick to further substantiate the genomic prediction.

3.3. Results

3.3.1. Genomic sequencing & assembly

The reference genome of *B. diminuta* MG was generated using a combination of Ion torrent and GAll of Illumina Platform and were assembled using Newbler and GAll Soapdenovo respectively. Contig level assembly was obtained by computational shredding of GAll into 300bp shreds with 200bp overlap and reassembled with Ion torrent data using Newbler. Despite good contig level assemblies, gaps could not be closed to generate good quality reference genome, therefore sequencing was performed again using Hiseq 2000 platform with a small insert size (300-700 bp) and Miseq platform of Illumina with a large insert library (~5kb, Mate-pair). High quality reads were processed and assembled using SPAdes V3.1.0 (Bankevich et al., 2012) and SeqQC V2.1 was used to remove low quality reads

and trim the contaminated adapter sequences. GapCloser and SSPACE-BASIC V2.0 (Boetzer & Pirovano, 2014) were used to close gaps between contigs to generate scaffolds after filtering and discarding contigs below 500bp length. Using BLAST tool and contigs from Ion torrent data the scaffolds were joined to make complete chromosome. The superscaffolds/pseudomolecules were further verified by sequencing randomly generated PCR amplicons. To raise the quality genome was subjected to PCR from the flanking forward and reverse sequence of chromosome and plasmid to amplify the junction and sub-cloned manually. This was then aligned with the assemblies to close gaps and circularize linear sequences thus obtaining a high-quality reference genome.

3.3.2. General features and Uniqueness of the genome

The assembled genome of *B. diminuta* generated a 6.5 Gbp data, showing almost 1361x coverage of the genome. Using Illumina data 28 scaffolds were obtained with N-50 of 1,786,567 bp which were then merged with data generated from Mate-pair library to obtain 3 super scaffolds. These super scaffolds were circularised by manual sequencing of the DNA from left and right flanks of assembled scaffold sequences. The largest super scaffold circularized with a 7.0 kb PCR amplicon. This 4,147,822 bp long DNA sequence was regarded as the chromosome sequence of *B. diminuta* (fig 3.1). The other two super scaffolds were circularized with 1.0 kb and 1.5 kb amplicon respectively. The sequence generated from these two amplicons facilitated circularization of respective scaffolds resulting in two plasmids with the size of 65,908 bp (fig 3.2) and 30,654 bp (fig 3.3) respectively. The 65,908 bp circular sequence perfectly matched with the physical map and partial sequence of pCMS1, a previously reported indigenous plasmid of *B. diminuta*. However, pCMS2 was a novel plasmid and was identified from the whole genome sequencing of *B. diminuta*. The complete annotated sequence has been deposited in the National Centre for Biotechnology Information (NCBI) under BioProject number PRJNA224116 and GenBank accession number CP026381. The general features of the genome are summarised in Table 3.2.

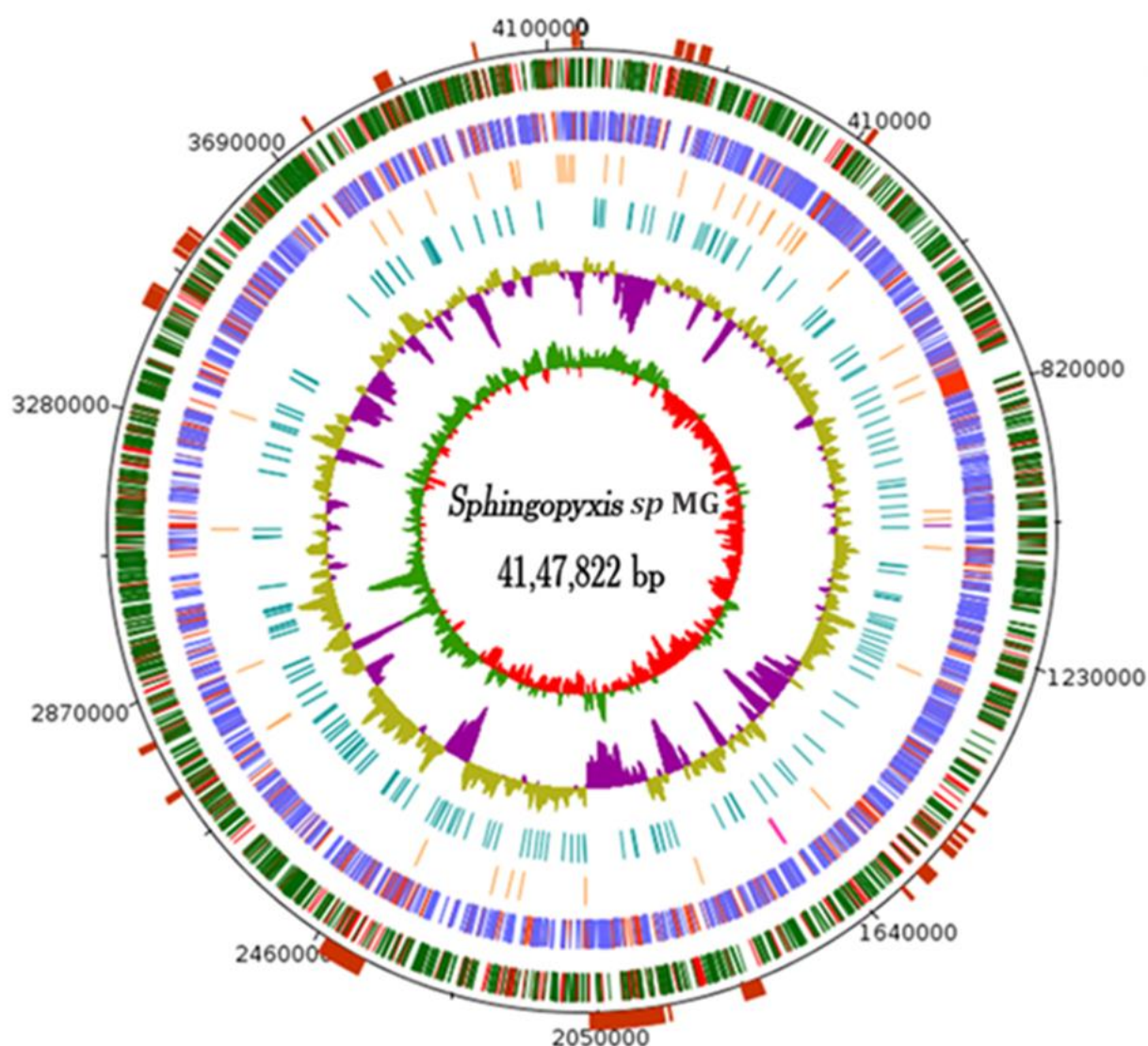


Fig. 3.1 The circular map of the *Sphingopyxis* sp MG chromosome.

The CDS on the forward strand (green for annotated, red for hypothetical) on the outermost circle and CDS on the reverse strand (blue for annotated, red for hypothetical) on the inner circle. The coordinates of the chromosomes are mentioned on the outermost circle. From the exterior towards the interior represented as circle 1, CDS on the forward strand (green for annotated, red for hypothetical); circle 2, CDS on the reverse strand (blue for annotated, red for hypothetical); circle 3, RNA genes (orange for tRNA, pink for rRNA and purple for tm RNA); circle 4, VNTRS (turquoise); circle 5, GC content (olive for positive and purple for negative); circle 6, GC skew (green for positive and red for negative). The maroon blocks shown above circle 1 represent genomic islands in the chromosome.

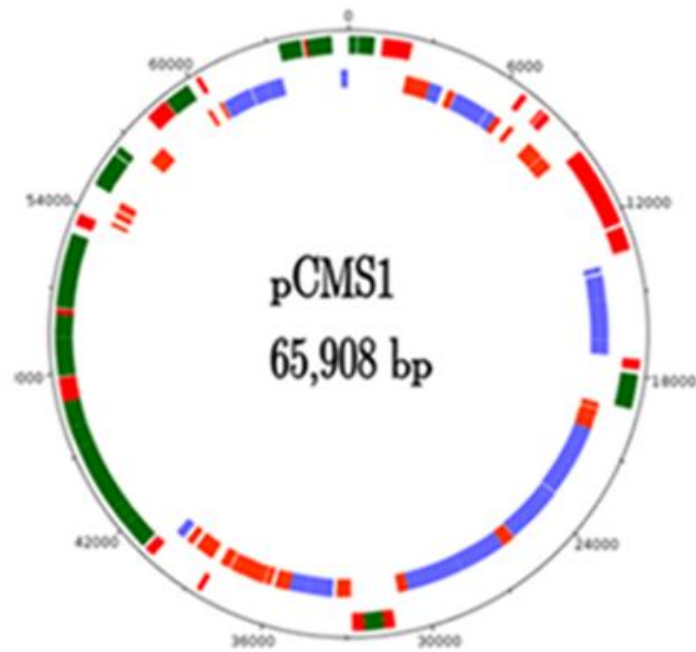


Fig. 3.2 The circular map of plasmid pCMS1.

The CDS on the forward strand (green for annotated, red for hypothetical) on the outermost circle and CDS on the reverse strand (blue for annotated, red for hypothetical) on the inner circle.

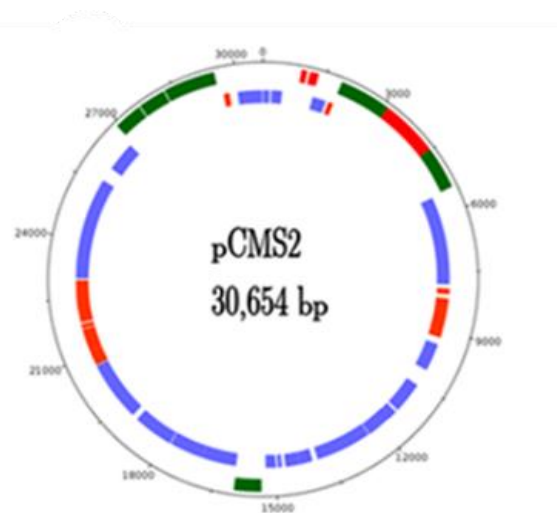


Fig. 3.3 The circular map of plasmid pCMS2.

The CDS on the forward strand (green for annotated, red for hypothetical) on the outermost circle and CDS on the reverse strand (blue for annotated, red for hypothetical) on the inner circle.

Table 3.2 General genome features of *Sphingopyxis* sp MG.

Characteristics	Total
Genome size (bp)	4147822
DNA coding region(bp)	3736976 (90.9%)
DNA (G+C content) (bp)	2759576 (66.53%)
Total genes	4018
RNA genes	52
Protein coding genes	3966
Hypothetical genes	1208
rRNA	3
tRNA	48
tmRNA	1

Since there exists taxonomic ambiguity in the classification of bacteria the whole genome sequence of the bacterium was initially used for obtaining a core genome assisted phylogenetic classification. Considering the current classification, the genome of *B. diminuta* was searched against the genomes of *Brevundimonas* genus available in the NCBI representative database. Surprisingly, none of the genomes of *Brevundimonas* available in the database showed any significant identity or coverage with *B. diminuta* genome sequence indicating that *B. diminuta* genome does not show any similarity to existing members of the genera. More surprisingly, the genome showed maximum identity with genome sequence of *Novosphingobium* sp PPIY, maximum similarity with the genome sequence of *Sphingomonas wittichii* and maximum coverage with the genome sequence of *Sphingopyxis alaskensis* (Table 3.3). In addition, none of the three genomes could cover the entire sequence of *B. diminuta* indicating that the whole genome sequencing has proclaimed the existence of a novel species possibly belonging the Sphingomonadaceae family of the order Sphingomonadales. The family consists of 4 genera namely *Sphingobium*, *Sphingomonas*, *Novosphingobium* and *Sphingopyxis*. To establish the taxonomic position of *B. diminuta* based on the whole genome, all the 120 then sequenced genomes of the order Sphingomonadales were considered for core genome-based phylogeny.

Table 3.3 Comparison of *B. diminuta* genome with *Sphingopyxis* sps from NCBI database

Species	Max Score	Total Score	Query cover	E value	Identity
<i>Sphingomonas wittichii</i> sp RW1	41422	1372000	40%	0	94%
<i>Novosphingobium</i> sp PPIY	23818	885900	33%	0	95%
<i>Sphingopyxis alaskensis</i> RB2256	18366	2556000	57%	0	90%
<i>Sphingobium</i> sp SYK-6DNA	8221	951300	35%	0	82%
<i>Novosphingobium aromaticivorans</i> DSM12444	7247	893300	34%	0	79%

3.3.3. Comparative Genome Analysis

Comparative genomic approaches facilitate both evolutionary and functional analyses. Conserved sequences and extent of homology can be useful to infer biochemical function, evolutionary history and also implied in phylogenetic classification (Pearson, 2013). All the 120 genome sequences of Sphingomonadales were analyzed in addition to genome sequence of *B. diminuta* and *E. coli* genome sequence, which was included as an out-group, to generate orthologous groups. Of the 122 genomes eight genomes of the 48 sister species of *Sphingomonas* showed very a smaller number of orthologs indicating either wrong classification or incomplete genome sequence and hence were excluded from the analysis. The other 114 genomes shared more than 1,500 orthologous groups and in fact of the 114, 103 genomes shared more than 3000 orthologous groups. To infer phylogenetic relationships, the curated set of 114 genomes with a total of 396 genes were selected to construct the phylogenetic tree showing taxonomic placement and evolutionary distance of all the genomes (fig 3.4). Interestingly, the organophosphate degrading bacterium, hitherto named as *B. diminuta*, clustered with the bacteria belonging to *Sphingopyxis* genus. As depicted in fig 3.4, the complete phylogenomic tree was divided into four major clades. Clade I & II, the major clades, contained strains that belong to *Sphingomonas* and *Sphingobium* respectively with few exceptions like *Sphingomonas paucimobilis* and *Sphingomonas* sp. SKA58 that were present in Clade II. The third clade branched into two groups i.e., *Zygomonas* and *Sphingomonas* members. The fourth clade also diverged separating the members of *Novosphingobium* and *Sphingopyxis*. Surprisingly, *Erythrobacter* which belongs to a separate family of Sphingomonadales actually clustered in this clade showing evolutionary closeness with *Novosphingobium*. Moreover, *Sphingomonas* sp 35-24ZXX clustered with *Blastomonas*

species. This result indicated that these bacteria are either wrongly classified or require further refinement to determine their exact taxonomic position. The phylogeny seen together with the reasonably low identity of its sequence with other members of *Sphingopyxis* genus supported reclassification of *B. diminuta* placing it *Sphingopyxis* genus. In accordance with the results of comparative genome analysis and phylogenomic classification, the bacterium was reclassified as *Sphingopyxis* sp.MG.

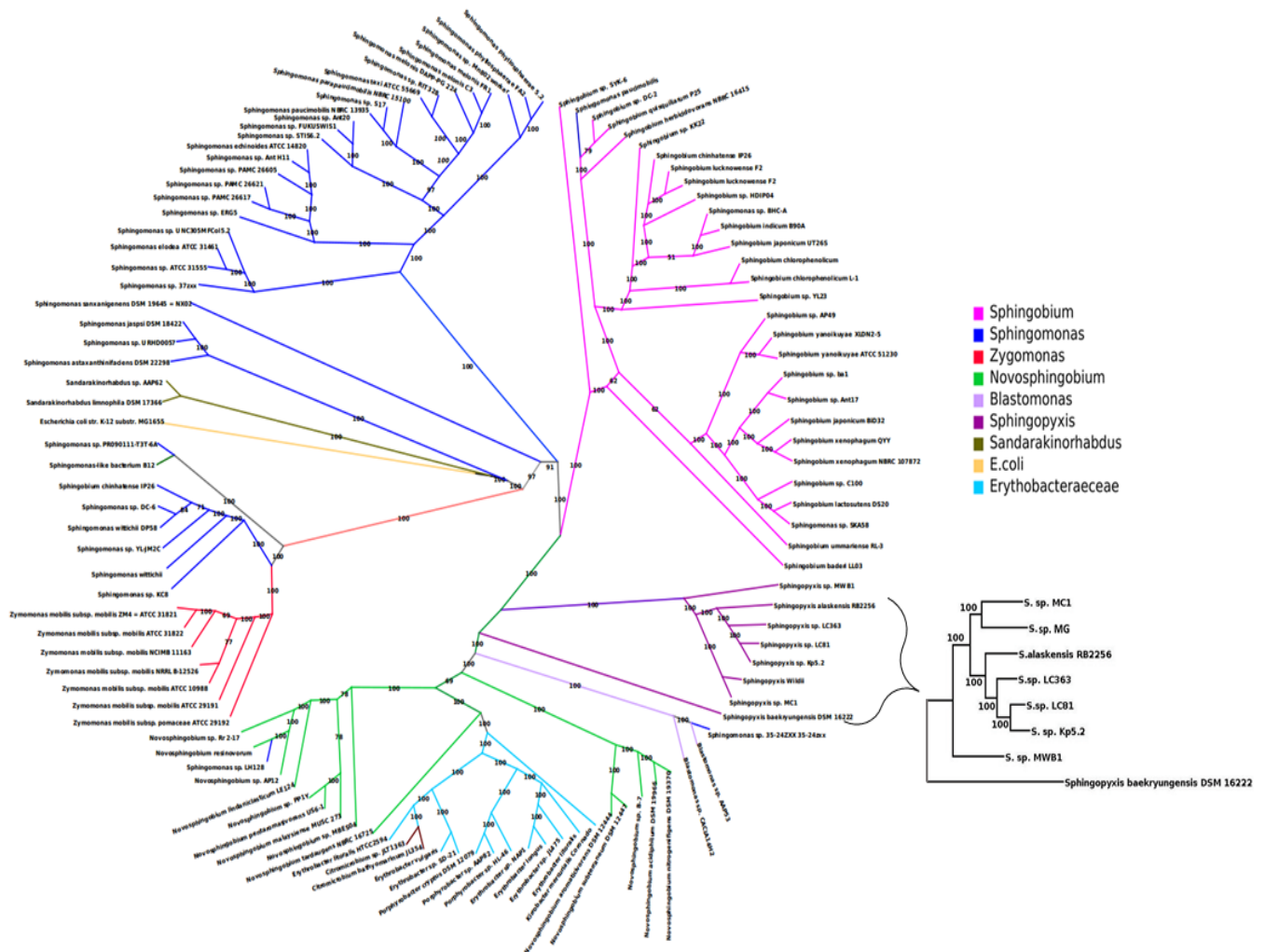


Fig. 3.4 Phylogenomic tree of the order Sphingomonadales.

Maximum likelihood tree of 114 whole genome sequenced bacterial taxa inferred from the concatenated, partitioned alignment of 396 core genes using RAXML. The support values for the branches were calculated from 100 bootstrap replicates. The branches of each genus of a Sphingomonadaceae are coloured uniquely, whereas the branches of Erythrobacteraceae are coloured in turquoise. This representation is a radial cladogram, in which branch length is not proportional to time, and some branches may be elongated such that the names of the taxa appear on the circumference of the circle. The tree in the inset depicts the phylogenetic relationships

of sequenced *Sphingopyxis* species. The scale bar indicates the number of substitutions per site. Each node of the tree is supported by a bootstrap value of 100.

3.3.4. Pan /core genome analysis of *Sphingopyxis* sp MG

A pan genome can be defined as being the entire gene content associated with the study group (Caputo, Fournier, & Raoult, 2019). Pan/core genome analysis provides better visualization of gene content differences, predict lifestyle of the bacteria, study the mobilome, resistome and also classify unique species based on their discontinuous genomic content (Medini, Donati, Tettelin, Masignani, & Rappuoli, 2005; Mira, Martin-Cuadrado, D'Auria, & Rodriguez-Valera, 2010). A pan/core genome analysis for the *Sphingopyxis* members (Table 3.1) including the sequence of *Sphingopyxis* sp MG was performed. About 4575 ortholog groups were obtained for all the *Sphingopyxis* genera with each ortholog containing 2-20 genes. Interestingly, the largest ortholog group containing 20 genes present in all *Sphingopyxis* genus was of RND efflux pumps. Moreover, except in *S. alaskensis* and *S. baekryungensis*, in all other members these set of genes were duplicated suggesting the extensive ability of *Sphingopyxis* species to mediate resistance to detrimental intermediates in metabolism. The other ortholog with 18 genes present in all species and duplicated in *Sphingopyxis baekryungensis* and *Sphingopyxis* sp MWB1 was Cobalt-Zinc-Cadmium metal resistance. Seen together with the RND efflux pumps, this indicated adaptation in all *Sphingopyxis* sps for metal ion resistance. Of the 4575 ortholog groups only 1515 clusters had at least one gene from each species (fig 3.5a). These single copy orthologs, constituting 33% of the total ortholog gene count, was regarded as the core genome of *Sphingopyxis* genus. The predicted core genome seemed to have achieved saturation indicating that addition of more *Sphingopyxis* genomes may not affect or lower the genetic content of the core genome (fig 3.5b). In addition, *Sphingopyxis* sp MG genome showed four ortholog groups having paralog duplications in its genome. Most of these genes were either hypothetical or either a highly similar transposase which was not unexpected. When compared with other genomes, *Sphingopyxis* sp MG showed lowest number of orthologs and minimum identity in ortholog gene set with the genome of *S. baekryungensis*. Phylogeny plotted from the whole genomes also showed that *S. baekryungensis* was evolutionarily distant from other species of *Sphingopyxis* (fig 3.5c). Consistent with this observation, recent studies by two independent groups also showed the need to reclassify *S. baekryungensis* based on its divergence from

other species of *Sphingopyxis* (Verma et al., 2020; Yang et al., 2020). More recently, we used about 100 genes from the core genome of 14 available complete genome sequences of *Sphingopyxis* species and performed a phylogenetic analysis to ascertain the species of *Sphingopyxis* sp MG. We identified that the genome shows maximum identity (98%) to the members of the species *granuli* (fig 3.5d). Core genome assisted phylogenetic classification shows that *Brevundimonas diminuta* MG belongs to *Sphingopyxis* genus and to the *granuli* species and hence the bacterium was appropriately reclassified in this study and renamed as *Sphingopyxis. granuli* MG.

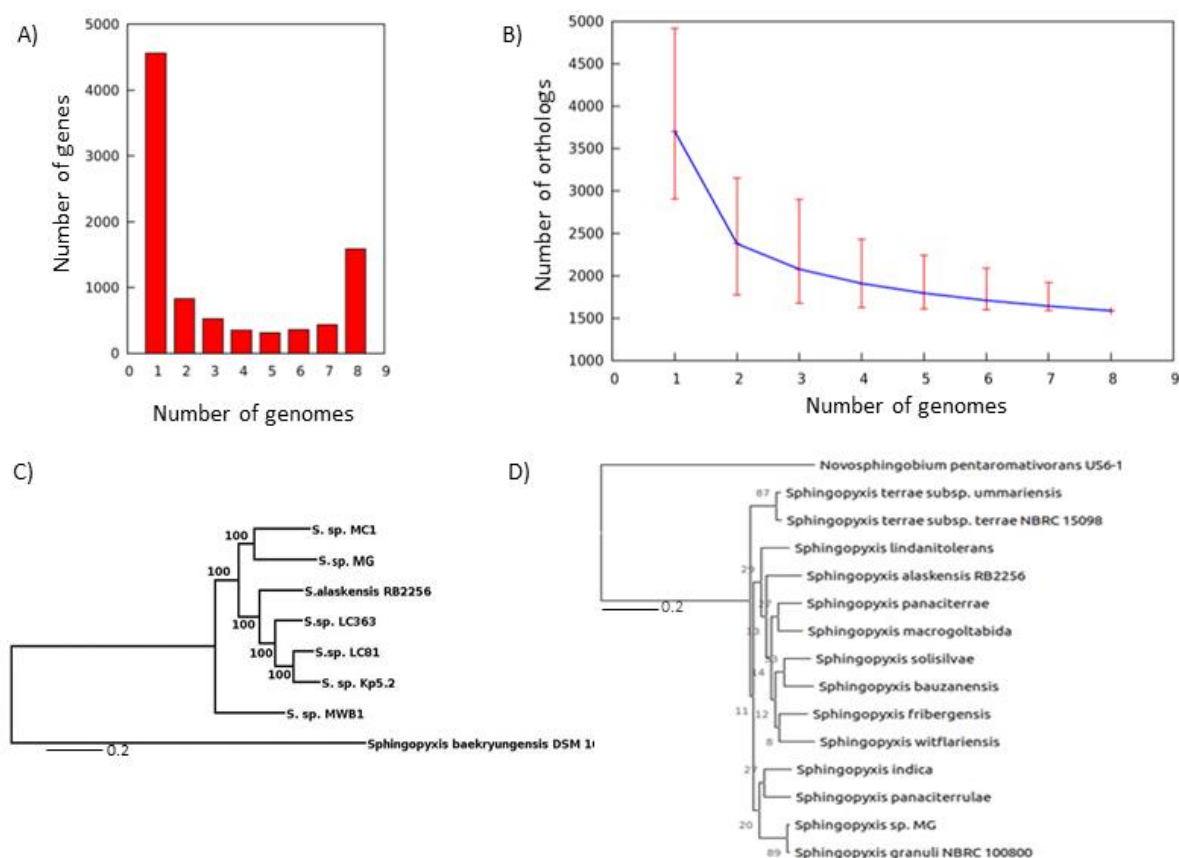


Fig. 3.5 Pan/core genome analysis *Sphingopyxis* sp MG with other of *Sphingopyxis* species.

Panel A shows the core genome size when different numbers of genomes are included. The point indicates the average number of orthologues in each comparison, whereas the error bars indicate the largest and smallest number of orthologues for all possible combinations of each number of genomes. Panel B shows the frequency of genes within the *Sphingopyxis* genome. Genes present in single genomes represent lineage-specific genes, while at the opposite end of the scale, genes found in all 8 species represent the core genome of *Sphingopyxis*. Panel C depicts the phylogenetic relationships of 8 *Sphingopyxis* species used in the study. The scale bar indicates

the number of substitutions per site. Each node of the tree is supported by a bootstrap value of 100. Panel D depicts the phylogenetic tree generated using Maximum likelihood method from all the recently sequenced *Sphingopyxis* species. The scale bar indicates the number of substitutions per site. The numbers at each node represent the bootstrap values.

3.3.5. Positive selection analyses

The core genes found in all the species of the genus *Sphingopyxis* were tested for positive selection. As recombination affects the analysis of positive selection, the genes that showed evidence of recombination were removed from further analysis (Anisimova, Nielsen, & Yang, 2003). The rest of the genes were selected for positive gene selection following the procedures described in materials and methods section. Almost 18% of the genes were positively selected with significant P value with significant (0.01) FDR level. All the genes in this positively selected group had at least one specific site positively selected with a minimum posterior probability >0.95. The fraction of positively selected genes, with respect to core genome is high (18%) and this value was comparable to the genus *Streptococcus* and was significantly less than the value obtained for *Campylobacter* genus (Lefebure & Stanhope, 2009). These selected genes belonged to all the COG categories, except Z category which has few core genes related with cytoskeleton. However, 7 categories seem to be enriched with the positively selected genes (fig 3.6). Genes of these categories, i.e., Extracellular structures (W), Defense mechanisms (V), Cell motility (N), wall/membrane/envelope biogenesis (M), Coenzyme transport and metabolism (H), Cell cycle control, cell division, chromosome partitioning (D), Energy production and conversion (C) were under pressure to evolve with their environment. Moreover, genes related to the 3 categories, cell-wall/membrane/envelope biogenesis (M), Cell cycle control, cell division, chromosome partitioning (D) and Energy production and conversion(C) were having proportionally more positively selected genes.

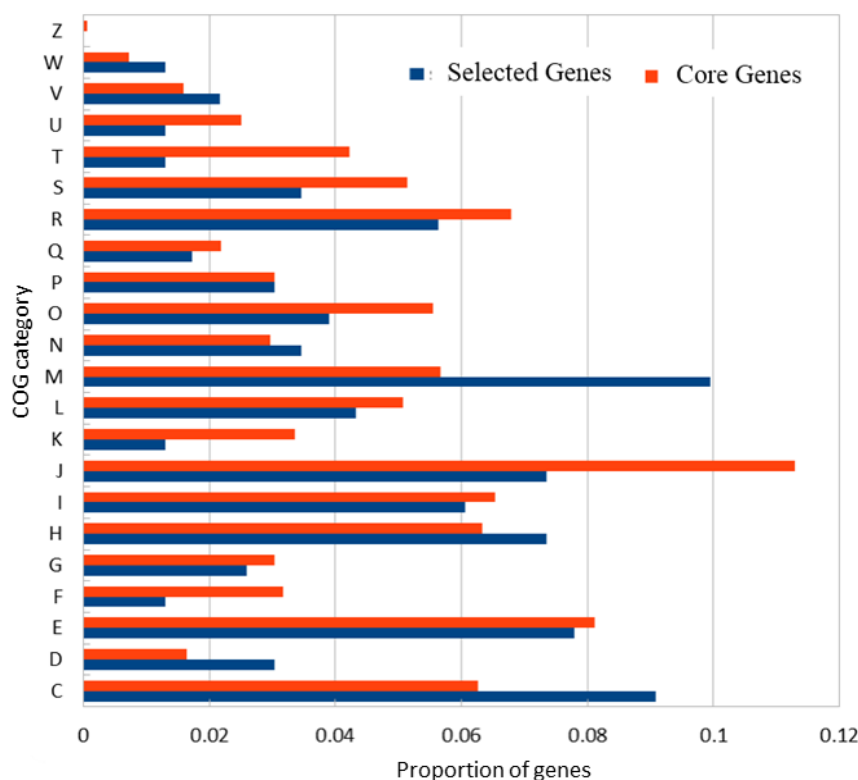


Fig. 3.6 Relationship of positively selected genes and the COG categories.

The X-coordinates stand for the diverse functional categories of COG, while the Y-coordinates stand for gene proportion within every functional category. Meanwhile, the blue and orange bars represent proportion of core genes of each COG, and that of positively selected genes (FDR < 1%), separately. Meanwhile, those COG categories are shown below: V, mechanisms of defense; U, vesicular transport, secretion and intracellular trafficking; T, mechanisms of signal transduction; S, unknown function; D, chromosome partitioning, cell division, cell cycle control; M, cell envelope/membrane/wall biogenesis; N, cell motility; O, modification at post-transcription level, chaperones and protein turnover; B, chromatin dynamics and structure; J, translation, such as biogenesis and ribosomal structure; K, transcription; L, repair, recombination and replication; C, energy conversion and production; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate metabolism and transport; H, coenzyme metabolism and transport; I, lipid metabolism and transport; P, inorganic ion metabolism and transport; Q, catabolism, transport, and biosynthesis of secondary metabolites; unknown proteins not collected in COG categories.

3.3.6. *S. granuli* MG genome organisation: Evidence for Genomic Islands (GIs)

The salient feature of Sphingomonads is their ability to degrade diverse recalcitrant aromatic compounds in the environment. Sphingomonads have been isolated from various toxic and recalcitrant xenobiotic polluted environments (Samantarrai, Lakshman Sagar, Gudla, & Siddavattam, 2020). Often bacteria that degrade xenobiotics carry catabolic genes as part of mobile genetic elements which are acquired through horizontal gene transfer. The mobile genetic elements conferring rapid adaptation in diverse harsh environmental

conditions occur as a part of the either genomic islands, plasmids or transposable elements. In contrast to plasmids, which are independently replicating the genomic islands (GI) are chromosomally integrated and show several striking features which are important in providing adaptive fitness to the organism (Juhas et al., 2009). In *Sphingopyxis. granuli* MG the organophosphate degrading *opd* was identified on a mobilizable plasmid (Table 3.4).

Table 3.4 Predicted GI and their functions

GI	Start position	End position	Length (bp)	No. of genes	GC (%)	Putative Function
1	129786	142214	12429	11	60.81	Iron acquisition
2	144098	155517	11420	10	58.3	Amino sugar and nucleotide sugar metabolism
3	163090	178352	15263	15	58.06	Haeme biosynthesis
4	418747	429099	10353	22	57.26	Not known
5	1432417	1442428	10012	7	60.06	Multiple functions
6	1462861	1468699	5839	6	59.09	Not known
7	1476275	1485294	9020	7	60.01	Iron acquisition
8	1487102	1495464	8363	5	62.63	Carbon metabolism
9	1499923	1510784	10862	11	59.76	Methane metabolism
10	1539509	1557959	18451	17	57.55	Not known
11	1583140	1587879	4740	5	61.56	aa metabolism
12	1818633	1850264	31632	29	58.83	Reverse Krebs cycle
13	1950635	1955342	4708	6	63.59	Two-component system
14	1959585	2063951	104367	113	60.97	Metal resistance
15	2382499	2447088	64590	72	60.32	Nucleotide sugar biosynthesis
16	2727706	2737614	9909	12	61.47	Not known
17	2800157	2812256	12100	14	58.35	Not known
18	3421235	3446824	25590	22	61.95	Not known
19	3447859	3454566	6708	8	60.91	Catechol degradation
20	3502348	3508809	6462	8	60.69	Arsenic resistance
21	3511382	3532998	21617	15	62.09	Not known
22	3535638	3543501	7864	7	64.11	Not known
23	3748585	3755388	6804	9	59.77	Iron acquisition
24	3858643	3880686	22044	21	59.79	Not known
25	3998557	4003774	5218	7	63.4	Hydroxypropionate- hydroxybutyrate cycle
26	4134170	4141529	7360	10	60.5	T4SS-related
27	4142593	4147677	5085	5	62.34	DNA repair

Using *in-silico* prediction software, it was identified that there are about 27 genomic islands in the genome of *S. granuli* MG. The size of the GI varied significantly with the largest GI of about 104kb with 113 ORFs and the smallest one of just 5kb with 5 ORFs. A summarised table showing the location and putative functions of each one of the GI is given in Table 3.4. The largest among them codes for heavy metal resistance and is shown in bold.

3.3.7. Largest Genomic Island is an Integrative Conjugative Element

Interestingly, the largest GI (GI:14) has also shown typical features of an Integrative Conjugative Element (ICE). An ICE is a mobile genetic element that is integrated into the host chromosome and encodes a functional conjugation system that mediates its transfer (Burrus, 2017; Johnson & Grossman, 2015; Wozniak & Waldor, 2010). The genes encoding the components of ICE can often be grouped into modules. A typical ICE therefore contains fixed or core module and variable or accessory module. The fixed modules of ICE mediate its integration, excision, conjugation and regulation. The integration and excision of ICE uses dedicated enzymes called integrase. The integrases recognize the attachment sites which are sites for integration and recombination of ICE with the recipient host. The attachment sites present in ICE are commonly 3'ends of the tRNA encoding genes. The transient replication of ICE in donor cells also requires a relaxase *tral* and *oriT* locus. In addition to replication genes, ICE also encodes all the genes involved in the synthesis of conjugative type-IV secretory system (T4SS). Moreover, to prevent its loss from the host, ICEs may most of the times also encode addiction systems like Toxin-Antitoxin (T-A) or Restriction-Modification (R-M) modules. The variable modules often confer benefits to the host in various circumstances. Based on its function, the variable module may be a resistance island conferring resistance to heavy metals, antibiotics or symbiotic island involved in photosynthetic function like N₂ fixation. It may sometimes be a xenobiotic degradation island providing adaptation in alternative carbon utilization or even virulence island conferring pathogenicity.

Careful analysis of the ICE (GI: 14) present in *S. granuli* MG showed the presence of genes encoding heavy metal resistance, multidrug resistance systems, toxin-antitoxin systems, restriction-modification systems and all modules for excision and integration. In case of heavy metal stress, uptake of toxic ions cannot be suppressed by simple downregulation of the transport activity. Therefore, active efflux systems are essential in adaptation of bacteria to heavy metal stress. These systems are usually encoded on plasmids. However, in *S. granuli* MG it was identified that genes *czcA* (C3E99_RS15220) encoding CzcA efflux pump antiporter along with gene *czcC* (C3E99_RS15215) encoding CzcC, efflux RND transporter which are involved in efflux of toxic heavy metals like Cadmium, Zinc and Cobalt were present in the ICE (Nies, 2003). The genes homologous to *ybhR* (C3E99_RS15240), *ybhS*

(C3E99_RS15245), *ybhF* (C3E99_RS15250), *ybhG* (C3E99_RS15255) and regulator *tetR* (C3E99_RS15260) identified are known to confer resistance to antibiotics chloramphenicol and cefoperazone (a third-generation cephalosporin) implying that *ybhRSFG*, an ABC-transporter maybe involved in efflux of antibiotics and conferring multidrug resistance to *S. granuli* MG (X. Z. Li, Plesiat, & Nikaido, 2015). Toxin-Antitoxin genetic elements ensure maintenance of the plasmid in bacteria by post-segregation killing by toxins in plasmid-less bacterial cells. Antitoxins regulate the expression of toxins thereby preventing the killing of cells. The gene encoding toxin element found in *S. granuli* MG ICE is *hipA* (C3E99_RS15440). Overexpression of HipA in *E. coli* has led to growth inhibition, persister formation and also tolerance to β -lactams and fluoroquinolone antibiotics (Bokinsky et al., 2013). As the antibiotics act on dividing cells, overexpression of HipA toxin in *S. granuli* MG ICE during antibiotic stress may generate a population of persister cells or non-dividing cells containing ICE which may decrease the susceptibility period in bacteria thereby enabling more bacteria to survive the toxic effect of antibiotics. Three antitoxin genes were also identified in the ICE. The two antitoxin genes homologous to *ygiT* (C3E99_RS15445, C3E99_RS15450) were located immediately downstream to the toxin *hipA* and another antitoxin homologous to the antitoxin belonging to the family that prevents host death, *phd* (C3E99_RS15665) was located further downstream within the ICE (fig 3.7).

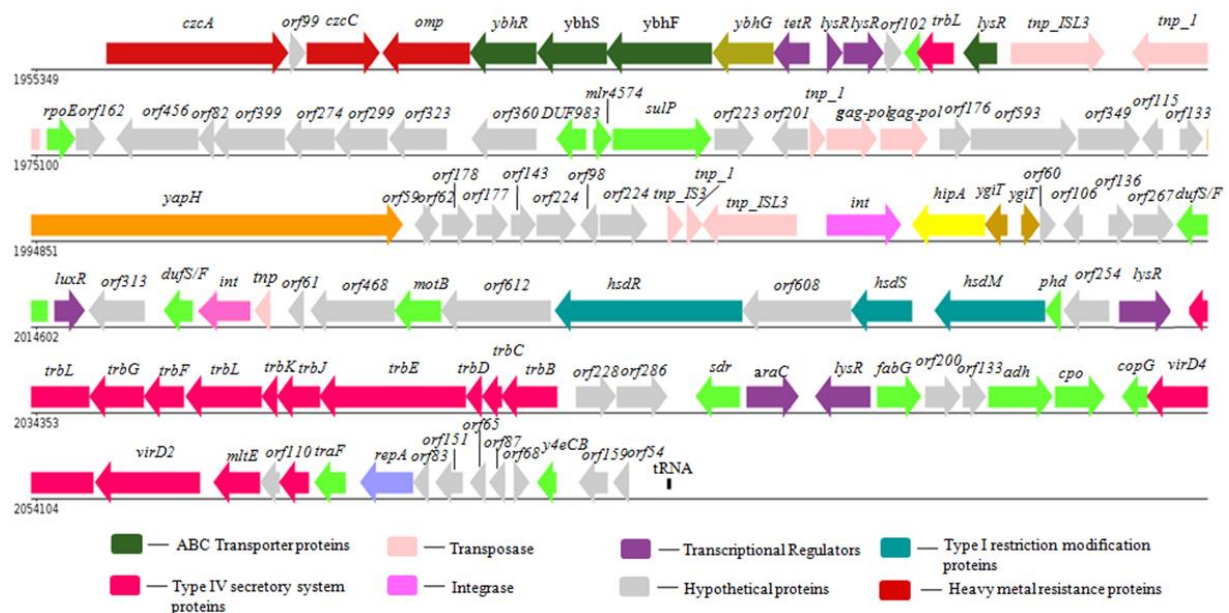


Fig. 3.7 Organization of ICE of *Sphingopyxis. granuli* MG.

The genes encoding fixed modules and other known functions are colour-coded. The genes encoding proteins with hypothetical functions are shown in grey. The numbers below the scale bars correspond to base pairs.

Furthermore, *su/P* (C3E99_RS15350), encoding a major facilitator superfamily-type sulphate transporter which is a transporter for mainly inorganic sulphate (Kertesz, 2001) was also part of ICE. In certain bacteria, Type I Restriction-Modification (R-M) system genes encode host specificity determinants. In addition to their role in defence, R-M systems also function to stabilize the host chromosome so that the genomic island acquired by horizontal gene transfer is not lost (Vasu & Nagaraja, 2013). Interestingly, *hsdM* (C3E99_RS15660), *hsdR* (C3E99_RS15645) and *hsdS* (C3E99_RS15655) encoding Type-I restriction-modification (R-M) system proteins were present in ICE of *S. granuli* MG. As seen in other typical ICEs, *S. granuli* MG ICE also encodes for all the proteins essential for mating pair formation (MPF) and its transfer by conjugation. The homologs of entire T4SS essential for conjugation including *trbI* (C3E99_RS15685), *trbG* (C3E99_RS15690), *trbF* (C3E99_RS15695), *trbL* (C3E99_RS15700), *trbK* (C3E99_RS15705), *trbJ* (C3E99_RS15710), *trbE* (C3E99_RS15715), *trbD* (C3E99_RS15720), *trbC* (C3E99_RS15725), *trbB* (C3E99_RS15730) and *traF* (C3E99_RS15815) were identified in the ICE of *S. granuli* MG. Vir D4 encoding gene (C3E99_RS15795) was also present in *S. granuli* MG ICE. This protein functions as a T4 coupling protein (T4CP) coupling the relaxosome to the T4SS and is known to harbor in most ICEs. In *Agrobacterium tumefaciens* VirD2, is a relaxase which cleaves one strand of T-DNA at the border sequence to generate ss-DNA (Pansegrau, Schoumacher, Hohn, & Lanka, 1993). Homolog of *virD2* (C3E99_RS15800) identified in *S. granuli* MG ICE may also be performing a similar function by nicking the DNA at the *oriT* generating the ss-ICE DNA for transfer into recipient host. In bacteria surviving in harsh environmental conditions, these large protein complexes are well known facilitators of horizontal gene transfer for adaptation of bacteria to various environmental conditions (Juhas, Crook, & Hood, 2008). For the excision from parent DNA and integration into recipient host DNA the ICEs encode specific recombinases known as integrases. The ICE of *S. granuli* MG showed the presence of two such integrase encoding *int* (C3E99_RS15435 and C3E99_RS15505) genes. After excision from the host chromosome the ICEs usually attach at specific attachment site known as *attB* on the recipient chromosome which is often at the 3' end of tRNA encoding gene and the attachment site on the ICE is known

as *attP* or *attI*. As expected, we also identified a tRNA encoding gene (C3E99_RS15870) at the border of the *S. granuli* MG ICE. A graphical map of the largest genomic island of *S. granuli* MG is represented in figure 3.7.

3.3.8. Methane Metabolism-F₄₂₀ biosynthesis Genomic Island

The cofactor F₄₂₀ was first identified for its role as redox cofactor in methanogenic bacteria and plays a key role in the central metabolism of bacteria and archaea (Hagemeier et al., 2003). Following this, its role was also elucidated in activation of prodrugs against tuberculosis infection (Stover et al., 2000). More recently, widespread distribution and significant role for F₄₂₀ was implied in aerobic bacterial metabolism and in soil ecosystem (Ney et al., 2017). These genes *cofC*, *cofD* and *cofE* are involved in the biosynthesis of F₄₂₀ cofactor from its precursor F₀. The F₀ precursor (7,8-didemethyl-8-hydroxy-5-deazariboflavin) is also redox active but is uncharged and therefore easily diffuses through the membrane. The *cof* genes help in addition of charged moieties to this precursor making which helps to retain the cofactor inside the cell. In soil bacteria, *Sphingopyxis granuli* MG we have identified F₄₂₀ biosynthetic genes *cofC*, *cofD* and *cofE* as part of the 10.4 kb genomic island (GI:9, Table 3.4) (fig 3.8a). The gene *cofC* (C3E99_RS13155) encodes for an enzyme that produces phosphorylated intermediates from substrates like 2-Phosphoenolpyruvate, 2-phospho-L-lactate or 3-phospho-D-lactate. These phosphorylated intermediates attach to F₀ precursor using *cofD* (C3E99_RS13160) which form F₄₂₀-O derivatives. The *cofE* (C3E99_RS13165) encoding γ -glutamyl ligase adds a poly- γ -glutamate tail to finally form a mature F₄₂₀ cofactor (Bashiri et al., 2016; Forouhar et al., 2008) (fig 3.8b). In *S. granuli* MG, the *cof* genes were associated with a complete Type IV secretory system (T4SS) genes. As seen with the metal-resistance island, T4SS genes in the island may help in conjugation and lateral transfer of *cof* genes amongst soil bacteria. The F₄₂₀-H₂-dependent reductases have been shown to reductively degrade diverse xenobiotic compounds like poly nitroaromatic compounds (Bashiri et al., 2016) and detoxification of environmental contaminants like malachite green (Jones & Falkinham, 2003). The diverse degradative role F₄₂₀ cofactor of seen together with the association of F₄₂₀ cofactor and T4SS in *S. granuli* MG shows that *S. granuli* MG may confer selective advantage to itself other soil bacteria by horizontal gene transfer of *cof* genes.

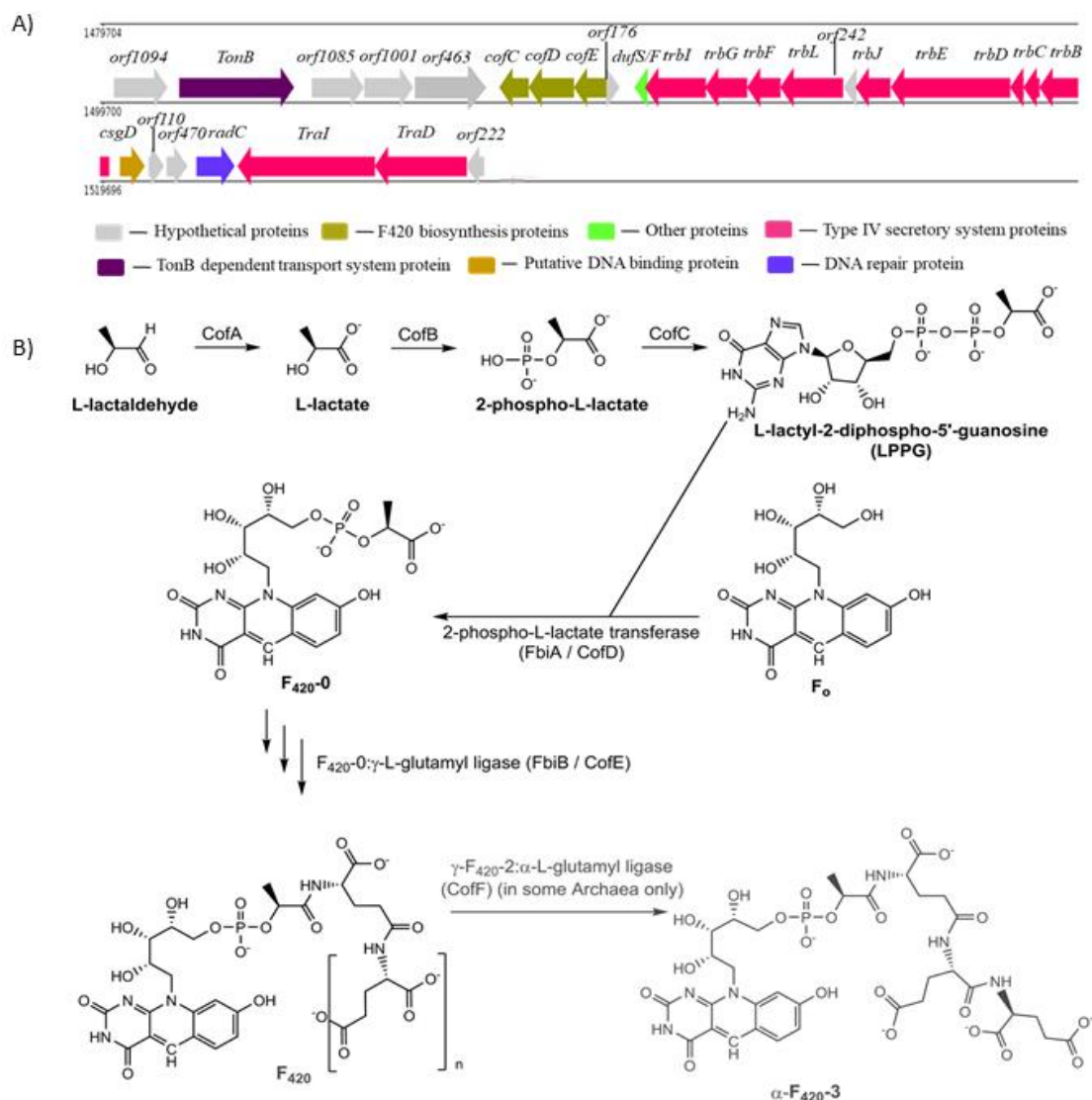


Fig. 3.8 Organization of F₄₂₀ biosynthetic island of *Sphingopyxis granuli* MG.

Panel A shows organization of F₄₂₀ biosynthetic Island. Genes encoding F₄₂₀ cofactor biosynthesis are highlighted in coloured rectangle arrows. The numbers below the scale bars correspond to base pairs. Panel B shows detailed pathway involved F₄₂₀ cofactor biosynthesis starting from precursor F₀.

3.3.9. Arsenic Resistance Island

Arsenic is found throughout the world, from natural and anthropogenic sources. Arsenic prevails in environment as trivalent species, As (III) (oxyanion arsenite (AsO₂⁻), and pentavalent species, As (V), or arsenate (AsO₄³⁻). Consumption of water with arsenic is the leading cause of poisoning in humans, leading to serious concerns worldwide (Abernathy, Thomas, & Calderon, 2003; Raju, 2022). Compared to As (V), As (III) has higher toxicity due to

its ability to bind strongly to sulfahydryl groups of proteins and thiol groups like glutathione and cysteine. Toxicity of As (V) is due to its ability to compete with phosphate for energetics and transport functions and also due to its transformation to As (III) (Ben Fekih et al., 2018).

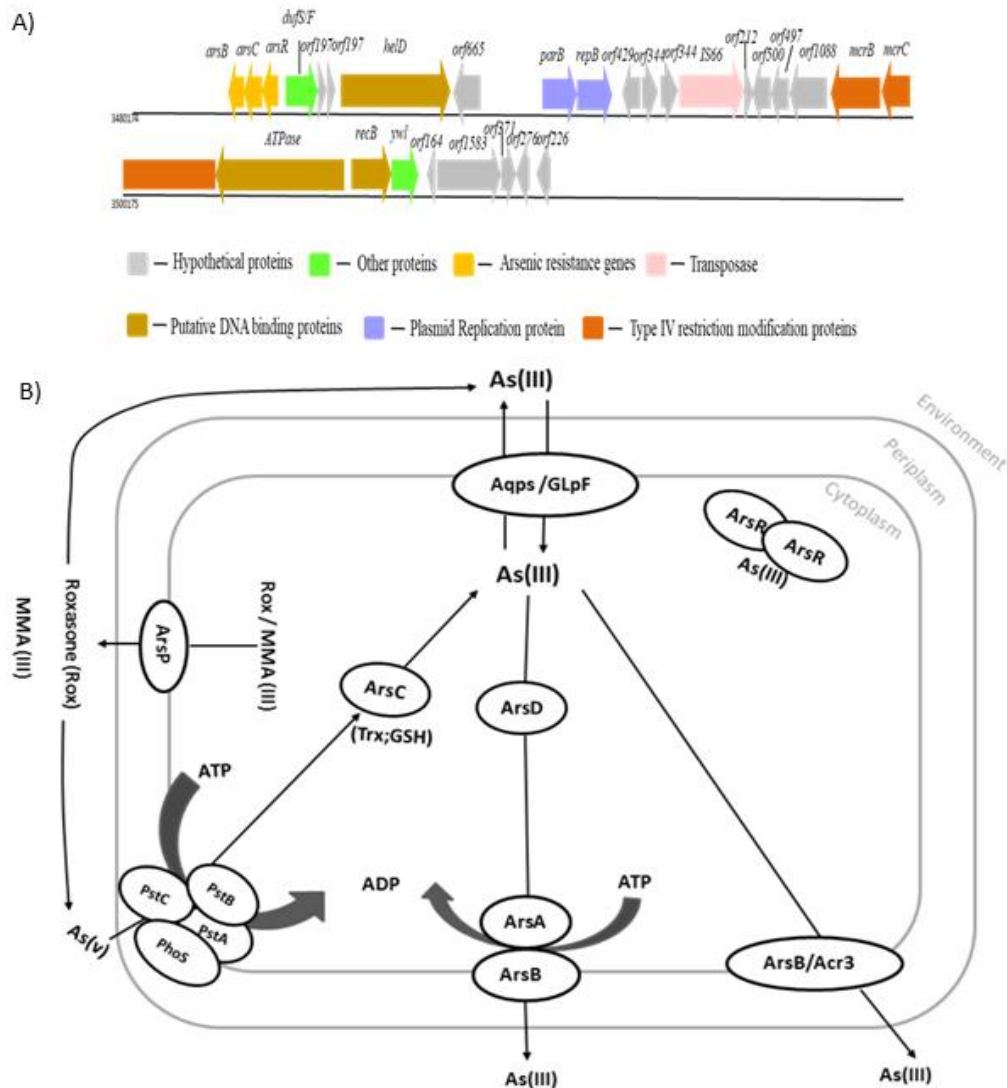


Fig. 3.9 Organization of Arsenic resistance island of *Sphingopyxis granuli* MG.

Panel A shows the organization of Arsenic biosynthesis island. The genes encoding Arsenic resistance are highlighted in a red coloured rectangle. Panel B shows the canonical pathway for Arsenic resistance in microorganisms.

In *S. granuli* MG, a 6.4 kb genomic island (GI:20, Table 3.3) showed the presence of canonical operon consisting of *arsRBC* (fig 3.9a) which is also seen in many other bacteria including *E. coli* (Carlin, Shi, Dey, & Rosen, 1995). The *arsR*(C3E99_RS03275) gene encoded

ArsR which is a transcriptional repressor and binds to the promoter region of ars operon. The Arsenite (AsO_2^-) and the trivalent species (As(III)), interacts with the AsrR dissociating it from the DNA leading to transcription of the ars operon. The gene *arsB* (C3E99_RS03265) identified coded for ArsB which is an integral membrane protein with ability to expel arsenite, thus reducing arsenite concentration in cell. The gene *arsC* (C3E99_RS03270), encoded ArsC is an arsenate reductase enzyme; that transforms arsenate to arsenite before expelling it out of the cell (fig 3.9b). The location of arsenic resistance in *S. granuli* MG genomic island emphasized the involvement of horizontal gene transfer in efficient gene dissemination mechanisms and also a strategy for *S. granuli* MG to survive in extreme heavy-metal polluted environment.

3.4. Discussion

The whole genome sequencing of *B. diminuta* MG has unravelled a plethora of information regarding the taxonomy and adaptive potential of the organism. The extensive coverage of the whole genome using multiple sequencing methods and manual PCR has given good quality complete circularized map of chromosome and plasmids. The sequencing has also revealed existence of a novel plasmid. The taxonomic ambiguities were resolved using this whole genome and the organism was rightfully classified based on core genome assisted phylogeny and appropriately renamed as *Sphingopyxis granuli* MG.

The *Sphingopyxis* genus was identified in 2001 and were characterized as being strict aerobes and chemo-organotrophic bacteria (Takeuchi, Hamana, & Hiraishi, 2001). At the time of sequencing this organism, sequence information was available for only seven other isolates of which only two organisms namely *Sphingopyxis alaskensis* RB2256 isolated from marine environments (Eguchi et al., 2001) and *Sphingopyxis fribergensis* sp. Kp5.2 capable of degrading styrene and phenylacetic acid (Oelschlagel et al., 2015) had completely assembled genomes. Scaffold genomes available for the other organisms were mostly isolated from triclosan contaminated environment (D. G. Lee, Zhao, Rezenom, Russell, & Chu, 2012) or crude oil contaminated sites (J. Kim et al., 2014). The versatile metabolic potential and adaptation of *Sphingopyxis* sp in general was also evident in the pan/core genome analysis. Core genome of all the *Sphingopyxis* members encode genes for membrane transport like

RND efflux systems increasing the survival chances of this species in toxic environments. Stress resistance especially heavy metal resistance is also seen in the core genome of most *Sphingopyxis* members indicating the core genes contribute to not only to the essential lifestyle of these species but is indicative of versatile metabolic profile in all or most of the *Sphingopyxis* spp. Furthermore, the compactness of these species is evident from the identification of an evolutionarily distant *S. baekryungensis* DSM16222. With the identification of new isolates and availability of more complete genomes the species level classification was also performed and *Sphingopyxis* sp MG was renamed appropriately as *Sphingopyxis. granuli* MG.

The adaptive potential was further elucidated from the core genome based positive selection analyses and presence of genomic islands. In *Sphingopyxis. granuli* MG seven COG categories of genes were enriched and found to be positively selected. Of these the Energy production and conversion (C) category is found to be one of the most enriched categories in almost all *Sphingopyxis* species (Yang et al., 2020). The genomic islands show several interesting features of *Sphingopyxis. granuli* MG. The largest genomic island was identified as an Integrative Conjugative Element (ICE). The ICE mainly contains gene involved in encoding heavy metal resistance. It also has genetic repertoire required to codes Toxin -antitoxin systems and antibiotic resistance suggesting that stress resistance is an important feature the ICE and is acquired through horizontal gene transfer.

The F₄₂₀ biosynthesis island has been reported for the first time in *Sphingopyxis* species. Cofactor F₄₂₀ catalyzes complex redox reactions and provides immense competitive advantage by enhancing endogenous metabolic processes and broadens the metabolic scope and degradative ability soil bacteria (Ney et al., 2017). Moreover, recombinant expression of F₄₂₀ biosynthesis have also been explored for bioremediation and biocatalytic applications (Greening et al., 2017; Taylor et al., 2010). Existence of genetic information for F₄₂₀ biosynthesis on a genomic island is undoubtedly beneficial to the *S. granuli* but its lateral mobility has the potential to enhance the soil microbial community.

The genome of *S. granuli* has arsenic resistance island and codes for one of the simplest and compact arsenic resistance pathways seen among microorganisms (Ben Fekih et

al., 2018). Due to its presence in genomic island, its lateral mobility to other bacterial strains is possible through horizontal gene transfer. The genome of *S. granuli* is tailor made to code for metabolic functions required to survive under harsh environmental conditions. It also has potential to share these critical traits with the other members of the soil community enabling them to coexist along with it in a highly polluted environment

CHAPTER 4 : Catabolic potential and adaptability of *Sphingopyxis. granuli* MG

4.1. Introduction

As discussed in the previous chapter, the Sphingomonads in general, and *Sphingopyxis* strains in particular, survive in xenobiotic contaminated environmental habitats ranging from waste water treatment plant (Kampfer, Witzemberger, Denner, Busse, & Neef, 2002), anaerobic sludge blankets (M. K. Kim, Im, Ohta, Lee, & Lee, 2005) to hydrocarbon contaminated soils (D. C. Zhang et al., 2010). Additionally, a number of *Sphingopyxis* strains have been shown to harbour the ability to efficiently degrade toxic anthropogenic pollutants like styrene (Oelschlagel, Zimmerling, Schlomann, & Tischler, 2014), tetralin (Garcia-Romero, Forstner, Santero, & Floriano, 2018), triclosan (D. G. Lee et al., 2012), microcystin (Massey, Zhang, & Yang, 2018), crude oil (J. Kim et al., 2014) and Hexachlorocyclohexane (Jindal, Dua, & Lal, 2013). *Sphingopyxis. granuli* MG was also isolated from highly polluted organophosphate contaminated environment (Munnecke & Hsieh, 1974). The xenobiotic pollutants are majorly aromatic compounds. Their degradation is initiated by peripheral oxygenases which have broad substrate range. These peripheral oxygenases transform the xenobiotic aromatic compounds into restricted set of catabolic intermediates. The ring containing intermediates are then cleaved by mono- or dioxygenases and resulting organic acids are subsequently funnelled into the Krebs's cycle (Habe & Omori, 2003; Nesvera, Rucka, & Patek, 2015; Perez-Pantoja, De la Iglesia, Pieper, & Gonzalez, 2008).

The ability of any microorganisms to survive in such a complex, harsh and highly polluted environment, suggests the existence of a robust catabolic repertoire. Since *S. granuli* MG is one such organism, in this chapter the whole genome sequence of the strain was used as a platform to predict the degradation potential and adaptability of *S. granuli* MG. The *in-silico* predictions were further validated using ¹⁴C- labelled PNP. Further, the soil microcosm studies were also performed to assess the *in-situ* degradation potential of *S. granuli* MG in soils fortified with ¹⁴C- labelled PNP.

4.2. Objective specific Methodology

4.2.1. Gene Network/Pathway Analysis

Before generating the metabolic map, an online database for microbial biocatalytic reactions and biodegradable pathways, EAWAG-BBD (Gao, Ellis, & Wackett, 2010) tool was obtained. Using this, the genome of *S. granuli* MG was searched for putative proteins of various aromatic compound degradation pathways. Each of the pathway protein domain were initially identified from the data base and the corresponding domain/s in the *S. granuli* MG were earmarked as homologs. In cases where more than one protein of the *S. granuli* MG showed similar domain, the ORFs of the predicted homologous were subjected to BLAST analysis. The ORF of protein from pathway tool was used as query sequence with filtering parameters $e\text{-value} < 1 \times 10^{-5}$. The closest homolog of *S. granuli* MG genome with most significant similarity was assigned as predicted protein for the pathway.

4.2.2. Mineralization of ^{14}C -labelled PNP by *S. granuli* MG

Wild type *S. granuli* MG cells were initially grown in minimal medium with sodium acetate (2%) as the only carbon source until OD 0.5. The cells were then washed thrice with minimal media without carbon source to remove any traces of carbon source and re-dissolved in fresh minimal medium to get a final OD of 0.2. Unlabelled PNP (50 μM , sterilized with 0.2-micron filter) was then added as the only carbon source in control (unlabelled) and test (labelled) samples. For radiolabelling of test samples, an aliquot of ^{14}C -labelled PNP was added to the culture medium such that the final count obtained was 1000 Bq. Upon consumption of this PNP, a fresh PNP was added in samples maintaining the concentration of labelled PNP at 1000 Bq and unlabelled PNP not exceeding 50 μM . The cultures were allowed to grow till OD of the culture reached to 0.4 units. The protein samples were taken from 1 ml of cultures grown using ^{14}C -labelled and unlabelled PNP and analysed by SDS-PAGE. The gels were dried, and the autoradiogram was developed following standard procedures (Dormer, 1981). The total RNA isolated from cells grown in labelled and unlabelled PNP were also analysed. The denaturing agarose gel containing RNA was transferred onto nylon membrane and visualized by autoradiogram.

4.2.3. Study design for soil microcosm

While designing the microcosm two types of soils were used. The brown forest soil (designated as E soil) used for the inoculation of *S. granuli* MG cells was sampled from a field at the Ehime Agricultural Experiment Station (Matsuyama, Japan) and the Andisol type of soil (designated as NK soil) sample from Tokyo University of Agriculture and Technology (Tokyo, Japan). The physicochemical properties of the soil [e.g., pH, total carbon content, cation exchange capacity, and exchangeable cations (Ca^{2+} , Mg^{2+} , K^{+} and Na^{+}) have been described in Table 4.1. The soil samples upto depth of 10 cm was collected by eliminating surface plant and other materials. The soil samples were sieved and sterilized by gamma radiation with Co-60 (60 kGy). The soil microcosm was prepared as follows under aseptic conditions. Initially, five loops of culture from a freshly streaked plate of *S. granuli* MG incubated at 30°C was collected using sterile inoculation loop into 5 mL of sterile double distilled water. The resuspended cells were thoroughly mixed and washed thrice and the cells were pelleted after each wash by centrifugation at 7000 g at room temperature. The cell pellet obtained was finally adjusted with sterile double-distilled water to obtain a final OD of 1.0. Meanwhile sterile 50 mL test-tubes were filled with 10 g of sterile soil per tube. Each one of the soil sample test tubes were prepared in triplicates and one of them served as control (uninoculated soils) and rest of the two served as test samples. The inoculations were performed ensuring water holding content of soil not exceeding 60% of its maximum water holding capacity. The inoculum was added to the sterile soil to obtain a final count of 10^6 colony forming units (cfu) per g of soil. After addition of inoculum the sample was left standing still for 30 min for better penetration. The samples were then thoroughly mixed with sterile spatulas and incubated in the dark at 30°C. At different time points, 1 g of soil sample was aliquoted and added to 9 ml PBS in a 50 mL test tube. The samples were vigorously vortexed and mildly sonicated to separate the bacterial cells adhered to the soil particles. The supernatant was then dilution plated onto LB to measure cfu per g soil.

4.2.4. Study design for mixed culture microcosm

The mixed culture microcosm was prepared by mixing with *S. granuli* MG with three different bacteria. Sterile soils samples were inoculated with 1) *S. granuli* MG + *Burkholderia*

Bcrs1W, 2) *S. granuli* MG + *Cupriavidus* sp TKC or 3) *S. granuli* MG + *Paenibacillus* sp NK-L2. The protocol followed for mixed culture microcosm studies was same as mentioned above with slight modifications. At the time of inoculation into soils both the bacteria were inoculated as a mixture such that each bacterium had a final count of $\sim 10^6$ cfu/g of soil. For measuring cfu the cells were distinguished based on the time taken to obtain cfu after plating (all the other bacteria formed single colony in 24 hrs whereas *S. granuli* MG formed single colonies in 3 days).

4.2.5. Study design for *in-situ* degradation of ^{14}C -labelled PNP in soil

The degradation of ^{14}C -labelled PNP ($50\ \mu\text{M}$ /g soil containing 10^6 dpm ^{14}C -PNP) was monitored in mixed culture microcosm soil by indirectly measuring the $^{14}\text{CO}_2$ released due to degradation of ^{14}C -PNP. The released $^{14}\text{CO}_2$ was trapped by placing an open vial containing solid $\text{Ca}(\text{OH})_2$ (fig 4.1). Initially the caps of 50 ml falcon tubes were drilled at the centre through which a stainless-steel wire tied with the vial containing $200\ \mu\text{L}$ of $0.1\ \text{M}$ $\text{Ca}(\text{OH})_2$ was inserted. The other end of the wire was passed through the lid and the length of the wire was adjusted in such a way that the vial containing $\text{Ca}(\text{OH})_2$ hanging above soil sample. The falcon tubes were then used to add sterile soils containing ^{14}C -labelled PNP and microcosm. Uninoculated soils containing ^{14}C -labelled PNP served as control. The cap was then screwed tightly onto the tubes. Any $^{14}\text{CO}_2$ released as a result of catabolism of ^{14}C -PNP was trapped as $^{14}\text{CaCO}_3$. When required, the cap of the falcon tube was unscrewed and the CO_2 trap was quickly removed and quickly replaced with a new one. The vial was removed and scintillation fluid (5mL) was added to the vial. The amount of CO_2 trapped was calculated by indirectly measuring the amount of $^{14}\text{CaCO}_3$.

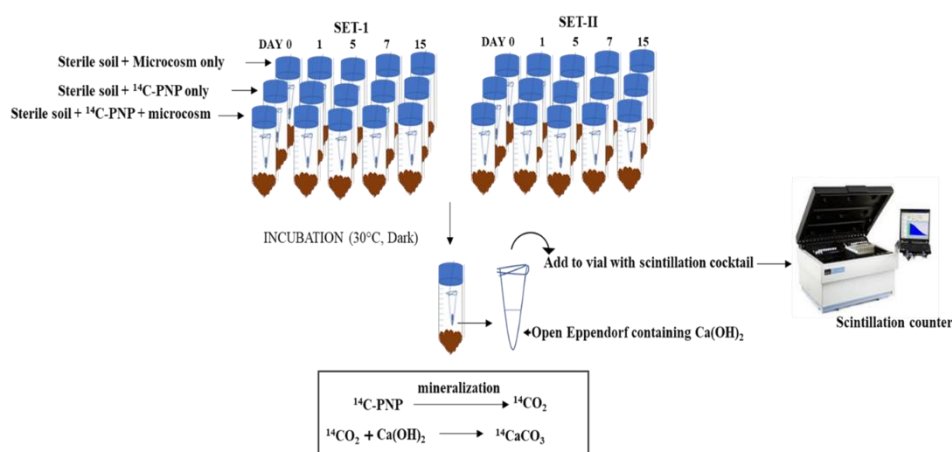


Fig. 4.1 Experimental design for CO₂ trapping studies.

The figure depicts the method used for *in-situ* mineralization of ¹⁴C-PNP by *S. granuli* MG. The radiolabelled ¹⁴CaCO₃ was measured using liquid scintillation counter.

4.3. Results

4.3.1. Genome of *S. granuli* MG is a treasure house of catabolic repertoire

S. granuli MG is known for degradation of organophosphates owing to the presence of *opd* gene in its indigenous plasmid (Pandeeti et al., 2011). The *opd* gene was present on a 65 kb plasmid named as pCMS1 (Parthasarathy, Azam, et al., 2017). The plasmid pCMS1 showed existence of T4SS along with *opd* gene, replication and partition modules. The pCMS1 has also showed existence of beta-ketoadipate pathway genes. Homologs of 3-oxoadipate CoA-transferase, *pcaIJ*, (C3E99_RS19340, C3E99_RS19345) and 3-oxoadipyl-CoA thiolase, *pcaF*, (C3E99_RS19335) involved in the conversion of beta-ketoadipate or 3-oxoadipate to TCA cycle intermediates were identified in pCMS1. Additionally in plasmid pCMS1, homolog of intradiol dioxygenase *catA*, (C3E99_RS19355) coding for catechol 1, 2 dioxygenase was also identified. Interestingly, as seen in other soil bacteria, in *S. granuli* MG also the beta-ketoadipate pathway genes are organized as one operon. In *S. granuli* MG, organophosphate hydrolase, the product of *opd* gene, initiates degradation of methyl parathion (MP) by hydrolysing triester linkage. The hydrolysis of MP generates p-nitrophenol (PNP) as one of the degradation products (Walker, 1993). PNP is more toxic to the soil microbial communities than parent compounds (Chakka et al., 2015). Therefore, the genome sequence of *S. granuli* MG was searched to know if it contains genetic repertoire (*pnpA*, and *pnpB*) to code for p-nitrophenol monooxygenase (*pnpA*) and benzoquinone reductase (*pnpB*) (W. M. Zhang, Zhang, Jiang, Chao, & Zhou, 2015). The genome search revealed existence of both the genes on chromosome indicating the involvement of plasmid pCMS1 coded *opd* and chromosome coded *pnpA* (C3E99_RS19035), and *pnpB* (C3E99_RS11405) in mineralization of PNP. Thus, *in-silico* search revealed that *S. granuli* MG has the potential to completely mineralize PNP. Encouraged by these findings, the genome was further used to unravel the catabolic potential of *S. granuli* MG.

Aromatic pollutants are degraded by two pathways namely, the upper pathway consisting of monooxygenases and dioxygenases which convert the complex and toxic

aromatic pollutants to simpler catechol or its degradation products (Dubinkina, Fridman, Pandey, & Maslov, 2019). The lower pathway consists of genes required for converting catechol into TCA cycle intermediates. Since *S. granuli* MG survives in heavily polluted soils, its genome was exhaustively searched to identify upper pathway genes involved in degradation of highly recalcitrant and toxic aromatic pollutants. The genome revealed that it has ability to degrade aromatics like 3-nitrophenol (3-NP), benzene and its derivatives like toluene and xylene, nitroaromatic derivatives like 3-hydroxyanthranilate, aromatic compounds 3-hydroxyphenyl propionate, benzoate, organochloride 3-chlorobenzoic acid and recalcitrant pesticide hexachlorocyclohexane (Lindane) (fig 4.3).

4.3.1.1. Degradation pathway for 3-Nitrophenol (3-NP)

The first enzyme in 3-NP degradation is nitrophenol nitro reductase (NbzA) catalysing reaction converting 3-NP to 3-hydroxylaminophenol. It is well characterized in *Pseudomonas alcaligenes* (Meulenberg, Pepi, & de Bont, 1996). The nitro reductase domain is from 11-201 amino acid (aa) in these 227 aa containing proteins with molecular wt 25.9 kDa. In *S. granuli* MG genome, 5 proteins were identified with this domain of which one protein C3E99_RS18520 having 233 aa and 25.7 kDa in size shows 90% sequence coverage and 33% identity with NbzA of *Pseudomonas alcaligenes*. The second enzyme is 3-hydroxyl aminophenol mutase converting 3-hydroxylaminophenol to amino hydroquinone. In domain similarity search, the N-terminal domain of hydroxyl aminophenol mutase was found to be most similar to those of glutamine synthetases from different species. The glutamine synthetase (GlnA) is well characterized in *E. coli*. The GlnA contains 469 aa with a mass of 61.9 kDa. Only one protein (C3E99_RS12095) in *S. granuli* MG genome shows size (containing 470 aa and 51.8 kDa molecular weight) and sequence (98%) coverage with 63% sequence identity with GlnA of *E. coli*. The next 3 enzymes namely amino hydroquinone oxidase, 1,2,4-benzenetriol dehydrogenase, 2-hydroxy-1,4-benzoquinone 2-reductase are characterized but domains for these enzymes were not identified to match with *S. granuli* MG genome. Therefore, the enzyme catalysing reactions from amino hydroquinone to benzoquinone were not identified. In the lower degradation pathway, the next enzyme is benzoquinone reductase catalysing reaction that converts benzoquinone to hydroquinone. It is well characterized in *Pseudomonas* (207 aa, 22 kDa). It has the FMN reductase domain from 5-196 aa. In *S. granuli*

MG genome, five such domain containing protein were identified of which one protein (C3E99_RS11405) containing 200 aa and 20 kDa in size shows 98% sequence coverage and 50% identity to PnpB of *Pseudomonas*. The product formed is hydroquinone which is channelled to the β -ketoadipate pathway as discussed earlier.

4.3.1.2. Toluene and Xylene degradation pathways

Aerobic degradation of toluene degradation which serves as important pathway for studying bacterial benzene ring metabolism. Its degradation begins by either methyl group oxidation (Shaw & Harayama, 1992), oxidation (mono-) at ring at positions 2, 3, or 4 (Olsen, Kukor, & Kaphammer, 1994; Shields, Reagin, Gerger, Campbell, & Somerville, 1995; Yen et al., 1991), di-oxidation at ring position 2,3 (Wackett, Kwart, & Gibson, 1988). In *S. granuli* MG genome, of these three mechanisms, the oxidation of methyl group appears to be the initiation step in degradation of both xylene and toluene. However, we could not identify the first enzyme with significant similarity. This could also be attributed to the species specificity for the monooxygenase enzymes. The toluene/xylene degrading enzyme were mostly from *Pseudomonas* species which are taxonomically distant from the *Sphingopyxis* species. Nevertheless, the enzymes for the subsequent steps were identified. The first enzyme in toluene/xylene degradation pathway is toluene side chain monooxygenase which belongs to class of oxidoreductases (EC 1.14.13.-). This enzyme catalyses the conversion of toluene to benzyl alcohol. Previous studies have shown that monooxygenase class consists of two different proteins, products of the *xylA* (in toluene degradation) and *xylM* (in xylene degradation) genes, which function as an electron-transfer protein and a terminal hydroxylase, respectively. Upon looking for the proteins containing these domains in *S. granuli* MG genome, one protein (C3E99_RS08385) was identified containing the Rieske (2Fe-2S) domain similar to domain in the electron transfer protein XylA from *Pseudomonas*. However, no significant sequence similarity was observed. The search for the second protein similar to XylM led to the identification of protein (C3E99_RS19035) with aromatic ring hydroxylase domain C-terminal end. However, as stated earlier, it also did not show any significant similarity to XylM of *Pseudomonas*. The second enzyme is benzyl alcohol dehydrogenase catalysing conversion of benzyl alcohol to benzaldehyde. The aryl alcohol dehydrogenase of *Pseudomonas putida* is well characterized. It has 366 aa and molecular

weight of 38.5 kDa and consists of alcohol dehydrogenase Gro ES like domain from 1-366 aa. In *S. granuli* MG genome ten such domain containing proteins were identified of which one protein (C3E99_RS02415) of 375 aa and 39.1 kDa shows a sequence coverage of 99% and maximum percent identity (47.76) with the XylB of *Pseudomonas putida*. The third enzyme is benzaldehyde dehydrogenase catalysing conversion of benzaldehyde to benzoate. This enzyme is also well characterized in *Pseudomonas putida* and consists of 487 aa with molecular mass of 51.8 kDa. It consists of an aldehyde dehydrogenase family domain from 1-487 aa. Upon searching for this domain in *S. granuli* MG genome, 17 such domain containing proteins were identified of which one protein (C3E99_RS02420) having 477 aa and molecular weight of 50.8 kDa shows 92% sequence coverage with maximum percent identity of 33.19 to the XylC protein of *Pseudomonas putida*. These reactions convert toluene finally to benzoate from where the benzoate degradation pathway begins.

4.3.1.3. Benzoate, 3-Chlorobenzoic acid and Benzene degradation pathways

Two pathways for aerobic metabolism of benzoate have been proposed. One is the dioxygenation to form catechol found in many bacteria (Harayama, Rekik, Bairoch, Neidle, & Ornston, 1991). The other pathway is via monooxygenation to form protocatechuate, mostly used in fungi (Sahasrabudhe & Modi, 1985). The first enzyme in the benzoate/chlorobenzoic acid degradation pathway is benzoate 1,2-dioxygenase or benzoate hydroxylase. This dioxygenase system consists of three proteins: the two subunits of the hydroxylase component BenA and BenB, and an electron transfer component BenC. There was no protein found in *S. granuli* MG genome showing homology either in domain or in sequence upon comparing with BenA or BenB of well characterized enzymes of *Acinetobacter baylyi*. BenC is also well characterized in *Acinetobacter baylyi*. It consists of 348 aa and mol wt of 38.7kDa. It consists of mainly 2 domains i.e., 2Fe-2S ferredoxin type domain from 14-109 aa and FAD binding domain from 116-217 aa. Upon searching for these domains, two proteins were found containing FAD binding domain but not the other domain. Of the two, one protein (C3E99_RS14190) having 271 aa and 29.7 kDa molecular weight is showed 92% sequence coverage and 22% identity with BenC of *Acinetobacter baylyi*. Another pathway for benzoate degradation is seen in most fungi. However, domain signatures for even this enzyme could be found. Benzoate 4-monooxygenase is well characterized in *Aspergillus niger*. It's a 517 aa

containing protein of 58 kDa. PFAM database shows that it contains p450 domain from 1-517 aa. Upon searching for this domain in *S. granuli* MG genome, eight protein sequences were found showing this domain of which one protein (C3E99_RS03660) containing 463 aa and molecular weight of 51.7kDa showed sequence coverage of 56% and a maximum percent identity of 25 to the benzoate 4-monoxygenase of *Aspergillus niger*. This enzyme catalyzes conversion of benzoate to 4-hydroxy benzoate. After the first step the product obtained is either cis-dihydrobenzenediol (in case of benzene upon action of benzoate 1,2-dioxygenase) or chlorobenzoate dihydrodiol (in case of chlorobenzoic acid upon action of benzoate 1,2-dioxygenase) or 4-hydroxybenzoate (in case of benzoate upon the action of benzoate 4 monoxygenase). In case of benzoate, if the reaction is catalyzed by benzoate 1,2- dioxygenase then the product is cis-1,6-dihydroxy-2,4-cyclohexadiene-1-carboxylic acid. The products formed by benzoate-4-monoxygenase follow a different pathway. The next enzyme in the benzoate degradation pathway which is similar for benzene, benzoate and 3-chlorobenzoic acid is BenD. BenD acts on the intermediates formed by the action of BenABC. It is well characterized in *Burkholderia pseudomallei* strain K96243. The protein consists of 264 aa with molecular weight of 27.3 kDa. It consists of short chain dehydrogenase reductase family domain. Out of the 49 short chain dehydrogenases domain containing protein in *S. granuli* MG genome one protein (C3E99_RS03875) containing 311 aa and molecular weight of 32.8 kDa showed 80% sequence coverage and maximum percent identity of 37 upon comparison with BenD of *Burkholderia pseudomallei*. The product formed is either catechol (benzene) or chlorocatechol (3-chlorobenzoic acid). In case of formation of 4-hydroxy benzoate the next enzyme is 4-benzoate-1-hydroxylase. The systematic name of this enzyme class is 4-hydroxybenzoate, NAD(P)H:oxygen oxidoreductase (1-hydroxylating, decarboxylating). This enzyme is also called 4-hydroxybenzoate 1-monoxygenase. It employs one cofactor, FAD. Since it is not well characterized in other bacteria upon looking for FAD binding domain in *S. granuli* MG, 12 such domains were identified of which one protein (C3E99_RS19035) having 406 aa and 43.4 kDa molecular weight showed 99% sequence coverage and 75% sequence identity to the multispecies monoxygenase of *Sphingopyxis*. The product formed is hydroquinone (HQ). The action of BenD on its substrates produces either catechol (benzene or benzoate degradation) or chlorocatechol (chlorobenzoic acid degradation) as stated earlier. This is acted upon by enzyme CatA (catechol 1,2- dioxygenases). CatA of *Acinetobacter*

baylyi is well characterized and has 311 aa and is of 34.3 kDa. It belongs to family of intradiol ring cleavage dioxygenases domain from 1-311 aa. Upon searching for proteins that can have similar function one protein was identified in the plasmid pCMS1 (C3E99_RS19355). The protein consists of 296 aa and its molecular weight is 33 kDa. It showed sequence coverage of 95% identity of 30% with the CatA of *Acinetobacter baylyi*. The next enzyme is muconolactone cycloisomerase CatB. It is well characterized in *Pseudomonas putida*. It contains 375 aa and is of 40.5 kDa. It belongs to the mandelate racemase/muconate lactonizing enzyme family with domain region spanning from 2-374 aa. Upon searching for this domain, two proteins were identified containing this domain of which one protein (C3E99_RS04580) containing 335 aa and mol weight of 35.8 kDa showed 70% query coverage and 29% identity with CatB of *Pseudomonas putida*. When chlorocatechols are substrates, the chloromuconate cycloisomerase not only catalyze the cycloisomerization after the ring cleavage but also are responsible for the dehalogenation reaction. The product thus formed are enol lactones i.e diene lactone (for chlorobenzoic acid) or β -keto adipate enol lactone (for benzene) The formation of maleylacetate from dienelactone is an additional step seen only in chlorinated aromatics like chlorobenzoic acid. The enol lactone in case of benzene is channelled to the beta keto adipate pathway. In case of chlorobenzoic acid maleylacetate is the substrate for the action of acetyl coA and thiolase. The next enzyme dienelactone hydrolase (PcaL) catalysing conversion of dienelactone to maleylacetate. It is well characterized in *Rhodococcus equi*. It consists of 395 aa and 41.5kDa mol wt. It has two domains mainly the $\alpha\beta$ - hydrolase and the carboxymuconolactone decarboxylase (CMD) domain from 20-243 and 300-379 aa respectively. Upon searching for proteins with these domains in *S. granuli* MG genome three proteins of which one protein (C3E99_RS08460) containing 289 aa and 31 kDa molecular weight showed 32% sequence coverage and 55% identity with PcaL of *Rhodococcus equi*. The product formed maleyl acetate is funnelled into the beta-keto adipate pathway as discussed earlier.

4.3.1.4. 3- Hydroxyanthranilate degradation pathway

3-HAA is formed as an intermediate in the aromatic pollutant 2-nitrobenzoate degradation pathway (Muraki, Taki, Hasegawa, Iwaki, & Lau, 2003). In 3-hydroxyanthranilate degradation the first enzyme is hydroxyanthranilate 3,4-dioxygenase NbaC. It is well

characterized in *Cuprivadus mettaldurans*. It consists of 174 aa and molecular weight of 17 kDa. It has an HAO domain. Upon searching for this domain in *S. granuli* MG genome, one protein (C3E99_RS12565) was identified containing 180 aa and of 20 kDa that shows 96% query coverage and 76% identity with NbaC of *Cuprivadus mettaldurans*. The second enzyme is 3-amino-2-carboxymuconate-6-semialdehyde decarboxylase. It is well characterized in *Pseudomonas fluorescens* and consists of 334 aa and 37 kDa molecular weight. It consists of an amidohydrolase domain. Upon searching for this domain in *S. granuli* MG genome, one such domain containing protein (C3E99_RS12630) was identified which showed 100% coverage and 65% identity to NbaD of *Pseudomonas fluorescens*. The third enzyme is 2-aminomuconate 6-semialdehyde dehydrogenase NbaE. It is also well characterized in *Pseudomonas fluorescens* and consists of 500 aa and 53 kDa molecular weight. It has the aldehyde dehydrogenase family domain and upon searching in *S. granuli* MG genome 17 proteins containing this domain were found of which one protein (C3E99_RS12580) showed 97% query coverage and 63% identity to NbaE of *Pseudomonas fluorescens*. The next enzyme is 2-aminomuconate deaminase NbaF. it is also well characterized in *Pseudomonas fluorescens*. It consists of 143 aa and 15.5 kDa molecular weight. It has endoribonuclease L-PSP domain. Upon mining *S. granuli* MG genome for this domain four such domain containing proteins were identified of which one protein (C3E99_12570) containing 148 aa and 15.6 kDa molecular weight showed 89% query coverage and 65% identity to NbaF of *Pseudomonas fluorescens*. The next enzyme is 4-oxalocrotonate decarboxylase encoded by nbaG, well characterized in *Pseudomonas fluorescens*. It consists of 251 aa and 21.7kDa with FAA hydrolase domain. In *S. granuli* MG genome six proteins containing this domain were identified of which one protein C3E99_RS12575 containing 259 aa and 26.6 kDa molecular weight shows 89% sequence coverage and 47% identity to NbaG of *Pseudomonas fluorescens*. The next enzyme is 2-oxopent-4-dienoate hydratase. It is well characterized in *Pseudomonas fluorescens* with 263 aa and 27.9 kDa molecular weight and FAA domain. In *S. granuli* MG genome, of six proteins with this domain, one protein (C3E99_RS03965) containing 259 aa and 26.7 kDa molecular weight showed 99% query coverage and 47% identity to NbaH of *Pseudomonas fluorescens*. The last enzyme in this reaction is 4-hydroxy-2-oxovalerate aldolase encoded by DpMG. It is well characterized in *Pseudomonas sp* CF600 with 345 aa, 37.4 kDa and DmpG like communication domain. One protein (C3E99_RS03970) was

identified containing 345 aa and 31.1 kDa that showed 97% coverage and 73% identity to DmpG of *Pseudomonas* sp CF600.

4.3.1.5. 3- Hydroxyphenylpropionate degradation pathway

The first enzyme in this pathway is 3-(2,3-dihydroxyphenyl propionate)-1,2-dioxygenase encoded by mhpA. The protein MhpA is well characterized in *Burkholderia cenocepacia* HI 2424, with 542 aa, 60.2kDa and FAD binding domain. Of the three proteins with this domain in *S. granuli* MG genome, one protein (C3E99_RS15890) with 489 aa and 43 kDa mol weight showed 76% query coverage and 29% identity to MhpA of *Burkholderia cenocepacia* HI2424. The second enzyme is 3-hydroxyphenyl propionate hydroxylase MhpB. It is well characterized in *Burkholderia cenocepacia* HI2424, with 321 aa, 34.3kDa mol weight and LigB domain. One protein (C3E99_RS11410) with 271aa and 28.1 kDa mol weight shows 45% query coverage and 27% identity to MhpB of *Burkholderia cenocepacia* HI2424. The third enzyme is 2-hydroxy-6-ketone,2,4-dienedioic acid hydrolase MhpC. It is well characterized in *Cupriavidus taiwanensis* strain DSM 1734 with 300 aa, 32.3kDa and $\alpha\beta$ -hydrolase domain. Of the 20 proteins containing this domain in *S. granuli* MG genome, one protein (C3E99_RS12505) with 310 aa and 34.3 kDa mol weight showed 89% query coverage and 28% identity to MhpC of *Cupriavidus taiwanensis* DSM 17343. The next enzyme MhpD is the similar in domain to NbaH and has already been discussed earlier. The identified degradative genes were found to be similar to the reported counter parts isolated from soil microflora (Perez-Pantoja et al., 2008; Perez-Pantoja et al., 2012). The potential degradation pathway for each of the aromatic compound was first identified and then searched for homologous genes in *S. granuli* MG genome. Further, the *in-silico* predicted similarities between *S. granuli* MG catabolome and well characterized degradative enzymes were linked to draw putative aromatic compound degradation pathways (fig 4.2). The generated pathway revealed existence of coordination between the chromosome encoded peripheral or upper pathway oxygenases and plasmid coded lower β -ketoadipate pathway enzymes (fig 4.3).

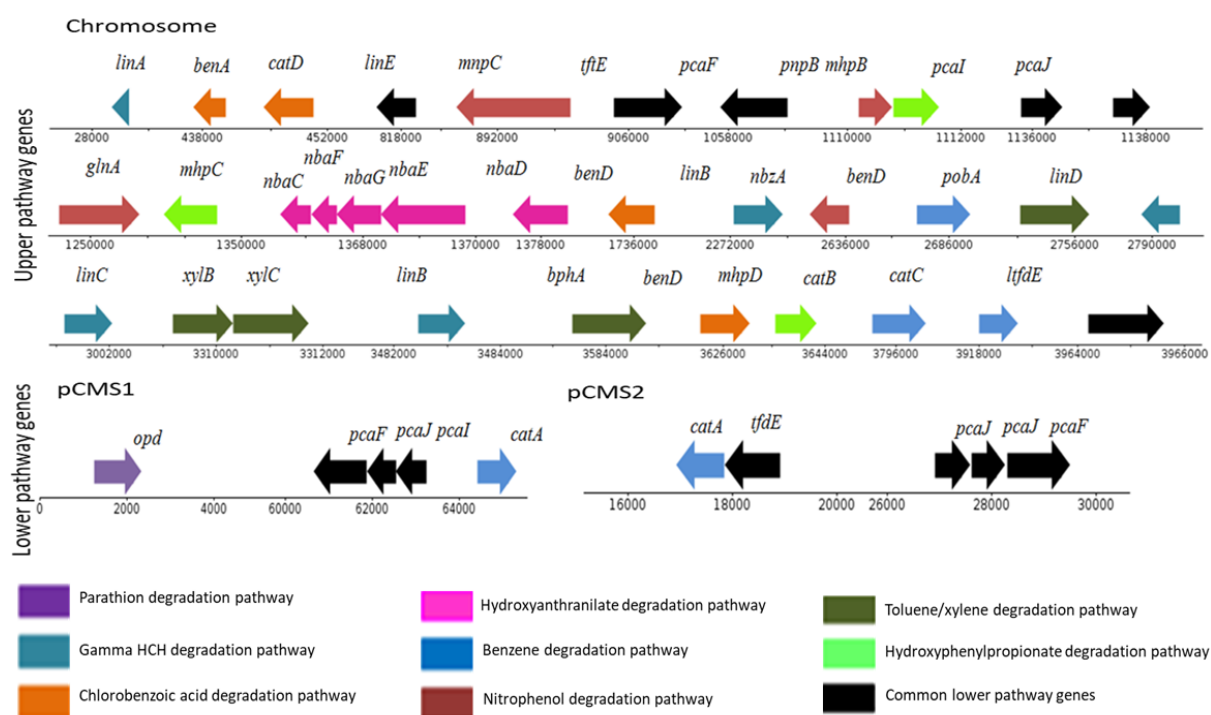


Fig. 4.2 Catabolic repertoire of *S. granuli* MG

The aromatic compound degradation genes in the chromosome and plasmids pCMS1 and pCMS2 of *Spingopyxis. granuli* MG. The genes are color-coded according to their catabolic pathway.

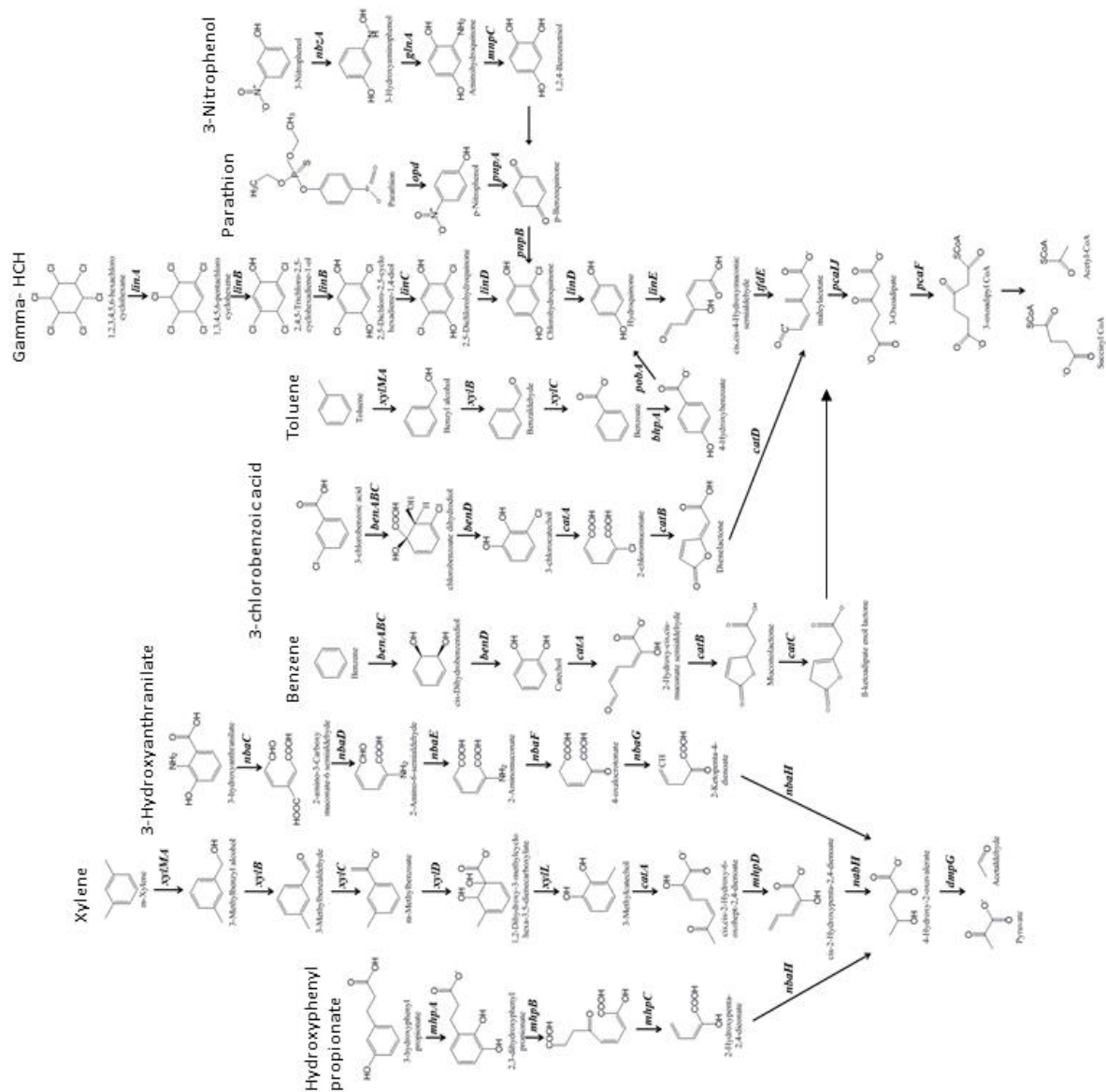


Fig. 4.3 *In-silico* predicted catabolic potential of *S. granuli* MG

The aromatic compounds used for the degradation studies selected based on the genomic context. Aromatic compound degradation pathways as revealed based on similarities with the well characterized degradative enzymes are shown. The products generated by the upper pathway monooxygenases/dioxygenases are funnelled into plasmid coded lower (β-ketoadipate pathway) pathway.

4.3.2. Validation of *in-silico* predictions

The *in-silico* predictions were validated by conducting *in-vitro* experiments, where *S. granuli* MG cells were grown in presence of either cold PNP or ^{14}C -labelled PNP. After growing the overnight, the cells were harvested and the total RNA and proteins were extracted as per the protocol discussed in the methods. If the ^{14}C -labelled PNP supported the growth of *S. granuli* MG, the ^{14}C -labelled carbon should become part of carbon backbone. Providing strong evidence for the mineralization of ^{14}C -labelled PNP in *S. granuli* MG cells, the ^{14}C -label of PNP, became assimilated into the protein molecules and RNA (fig 4.4). The proteins and RNA isolated from ^{14}C -labelled PNP grown *S. granuli* MG emitted radiation suggesting that the radiolabelled carbon found in PNP is now part of proteins and RNA. This showed direct evidence that ^{14}C -labelled PNP has degraded to form catabolic intermediates that have incorporated into the central metabolic pathways generating labelled precursor molecules apparent in the synthesis of macromolecules like protein and RNA.

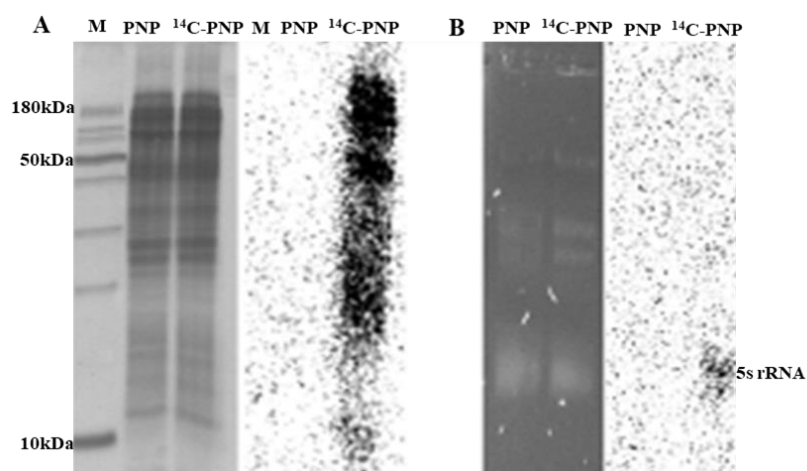


Fig. 4.4 Mineralization of ^{14}C -PNP by *S. granuli* MG.

Panel A shows proteins extracted from *S. granuli* MG cells grown in presence of normal and ^{14}C -labeled PNP analysed on SDS-PAGE and incorporation of ^{14}C into proteins is shown in corresponding autoradiogram. Panel B shows the RNA extracted from similarly grown cultures and the corresponding autoradiogram. The sharp signal shown by the arrow indicates incorporation ^{14}C into 5s rRNA.

4.3.3. *S. granuli* MG survives only in sterile E- type soils

The bioremediation potential of *S. granuli* MG can only be realized based on its ability to survive in soil. The ability of *S. granuli* MG, to proliferate and survive in soil was tested by inoculating the cells in two types of soils namely, brown forest soil (designated as E soil) and the Andisol type of soil (designated as NK soil). Before inoculating the *S. granuli* the physicochemical properties of the soil were determined and they are listed in Table 4.1.

Table 4.1 Characteristic properties of the E and NK soil types

Properties	E soil	NK Soil
pH (H ₂ O)	6.4	6.7
Electrical conductivity (mS/cm)	0.1	0.2
Cation exchange capacity (meq/100 g)	10.6	34.5
Degree of base saturation	107	73.1
Total C (g/100 g)	2.2	5.3
Total N (g/ 100 g)	0.2	0.5
Ammonium –N (mg/ 100 g)	0.9	1.2
Nitrate- N (mg/100 g)	1.2	5.4
Available Phosphate (mg/ 100 g)	108.3	13
Phosphate absorption coefficient	20	2087
Exchangeable K (mg/ 100 g)	30.3	88
Exchangeable Ca (mg/ 100 g)	263.3	517
Exchangeable Mg (mg/ 100 g)	24.3	97
Exchangeable Mn (mg/ 100 g)	1.3	0.6
Available Fe (mg/ kg)	4.5	6.2
Available Cu (mg/ kg)	0.1	0.9
Available Zn (mg/kg)	3.1	3.7
Ca/Mg	7.7	3.8
Mg/K	1.9	2.6
Boron (mg/kg)	0.5	1.0

After testing samples physico-chemical properties, the *S. granuli* MG cells were added to the sterile soil as described in methods section and the ability to survive in the soils was assessed by measuring the growth of cells by counting CFUs. At various intervals the cells were recovered and the cfu for each type of soil g⁻¹ was counted (fig 4.5). During the first 15 days, the cfu of *S. granuli* MG cells grown in NK soil declined constantly indicating that NK soil is not suitable for the survival or proliferation of *S. granuli* MG. In contrast, E soil was able to support the survival of *S. granuli* MG. However, in both the soil samples, the bacteria did not show any proliferation or growth indicating that as monoculture *S. granuli* MG was able to survive in E soil but did not show any growth until the period tested.

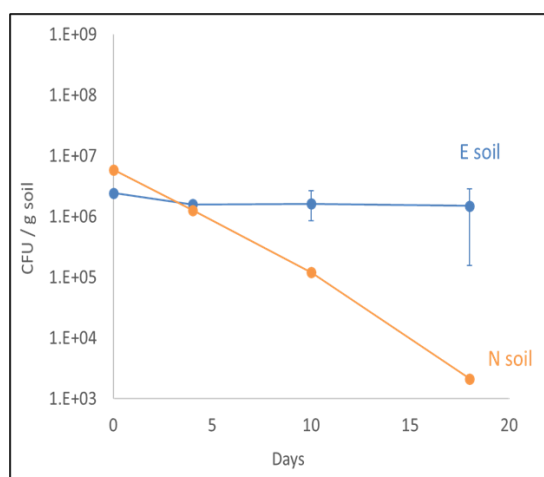


Fig. 4.5 Survival plot of *S. granuli* MG in E and NK soil.

The experiment was performed in triplicate. Mean values of cfu on the LB are plotted. Error bars indicate the maximum and the minimum value. E soil is brown forest soil and N soil is Andisol.

4.3.4. Survival of *S. granuli* MG in soil microcosm

The growth and survival of *S. granuli* MG in microcosm was assessed by co-culturing *S. granuli* MG with *Cupriavidus* sp TKC, *Burkholderia* sp Bcrs 1W, and *Paenibacillus* sp NK L2. As shown in fig 4.6a, both *S. granuli* MG and *Cupriavidus* have grown exponentially upto 5 days. Subsequently, their number remained constant till 10 days without ever observing decline in their number. There was no decline in *S. granuli* MG population even after 10 days of inoculation. When *S. granuli* MG was co-inoculated with *Burkholderia* (fig 4.6b), both of them have exponentially grown. Even at the end of 10th day, no stationary phase was

observed suggesting the potential to grow these two cells beyond 10 days of inoculation. In contrast to these two observations, in co-cultures having *Paenibacillus* sp, *S. granuli* MG has overgrown (fig 4.6c). When compared to the co-cultured *Paenibacillus* sp that had 1×10^6 CFU at the end of 18th day, at the same time point, approximately 1×10^9 *S. granuli* MG cells were found per gram of soil. This clearly indicates that *S. granuli* MG shows significant growth and survivability when its co-cultured with soil bacteria *Paenibacillus* sp. Therefore, the bioremediation potential of *S. granuli* MG can be fully exploited if it is co-cultured with soil bacteria as a microcosm.

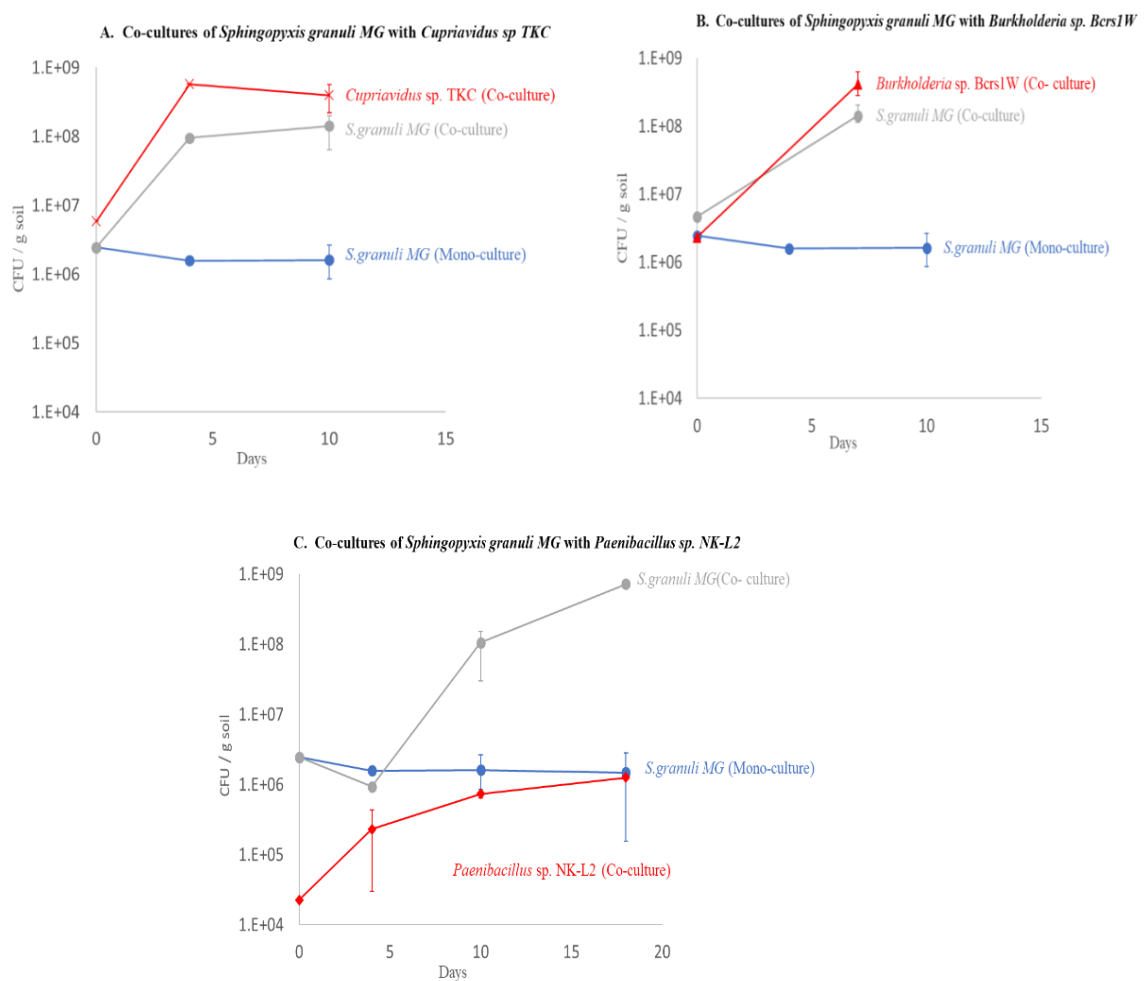


Fig. 4.6 Survival and growth of *S. granuli* MG in soil microcosm.

The panel A shows the growth of *S. granuli* MG (cfu/ml) in microcosm along with *Cupriavidus* sp TKC over time period of 15 days. The panel B shows the growth of *S. granuli* MG (cfu/ml) in microcosm along with *Burkholderia* sp Bcrs1W over time period of 15 days. The panel C shows the growth of *S. granuli* MG (cfu/ml) in microcosm along with *Paenibacillus* sp NK-L2 over time period of 15 days. The experiment was performed in triplicate. Mean values of cfu on the LB are plotted. Error bars indicate the maximum and the minimum value.

4.3.5. PNP degradation in soil microcosm by *S. granuli* MG

The degradation of ^{14}C -labelled PNP in soil microcosm was assessed by monitoring the increase in the amount of trapped CO_2 in the falcon tubes containing soil microcosm. If PNP added to soil microcosm is degraded the catabolic activity should contribute for the release of $^{14}\text{CO}_2$. As seen in fig 4.7, on day 0, soils containing *S. granuli* MG and *Paenibacillus* sp microcosm (10^6 cfu/g soil) only (S+M), soils containing only ^{14}C -labelled PNP (50 μM /g soil) without microcosm (S+P), and soils with both ^{14}C -labelled PNP and microcosm of *S. granuli* MG and *Paenibacillus* sp have shown not considerable counts for $^{14}\text{CO}_2$. However, with increasing days of incubation in soil upto the 15th day, mineralization of ^{14}C -labelled PNP is detected by the increase in counts per minute (cpm) of trapped $^{14}\text{CO}_2$ only in soils containing *S. granuli* MG and *Paenibacillus* sp microcosm with ^{14}C -labelled PNP. The sterile soils with microcosm (S+M) or soils with only ^{14}C -labelled PNP (S+P) did not show significant increase in the cpm of trapped $^{14}\text{CO}_2$ indicating the *in-situ* mineralization of PNP in the soil microcosms prepared by mixing *S. granuli* MG and *Paenibacillus* sp with the sterile soil.

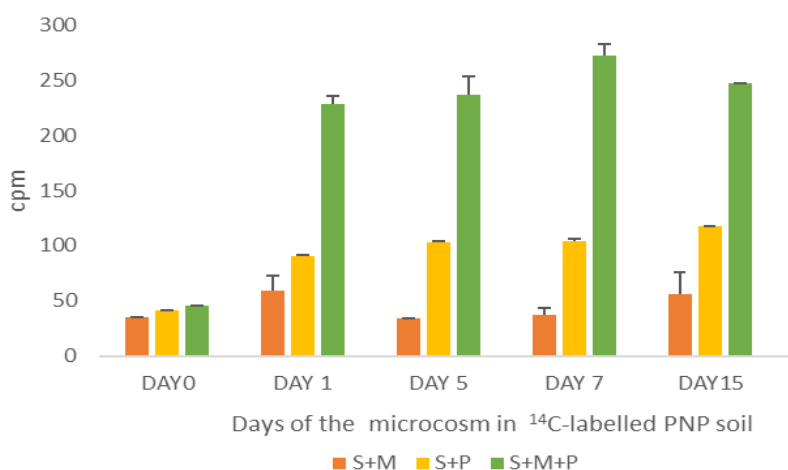


Fig. 4.7 Determination of *in-situ* mineralization of ^{14}C -PNP by measuring the trapped $^{14}\text{CO}_2$

The orange bars (S+M) indicate the cpm from soils containing only microcosms at different days, from day 0 to day 15. The yellow bars (S+P) indicate the cpm from soils with only ^{14}C -PNP without microcosm at different time points from day 0 to day 15. The green bars (S+M+P) indicate the cpm from soils with ^{14}C -PNP and microcosm at different time points from day 0 to day 15. Error bars represent standard error from duplicates.

4.4. Discussion

After using the genome sequence to reclassify the bacteria and identifying the genomic island landscape, the whole genome sequence was used as the platform to determine the catabolic potential of *S. granuli* MG by using *in-silico* tools. The organophosphatase degrading *Sphingopyxis. granuli* MG shows a versatile catabolic repertoire encoding a number of monooxygenases, dioxygenases, oxidoreductases, dehydrogenases and dehalogenases converting the complex and toxic aromatic compounds into simplest catechol. The *in-silico* predicted catabolic potential indicated that the chromosome of *Sphingopyxis. granuli* MG encoded for the peripheral pathway genes involved in degradation of toluene, xylene, benzene, 3 and 4-nitrophenol, hexachlorocyclohexane and chlorobenzoate. The genome sequence has also showed potential to degrade aromatic compounds like hydroxyanthranilate and hydroxyphenyl propionate. The indigenous plasmids pCMS1 and pCMS2 code for lower pathway enzymes. Therefore, the chromosome and plasmid borne catabolic genes show potential to degrade aromatic pollutants like organochlorides, nitrophenols, benzene derivatives and nitroaromatics. The peripheral oxygenases coded by chromosome convert the aromatics to catechol and the plasmid coded ring cleavage dioxygenases convert catechol into TCA cycle intermediates. The *in-silico* predicted degradative potential suggests the ability of *S. granuli* MG in mineralization of a variety of aromatic compounds.

Studies performed to validate *in-silico* predictions using ^{14}C -labelled PNP has demonstrated the ability of *S. granuli* MG in mineralization of PNP. Incorporation labelled carbon into proteins and RNA in cultures grown in ^{14}C -labelled PNP suggests entry of PNP in central carbon catabolic pathway. The soil microcosms studies clearly shown that *S. granuli* MG fails to survive in NK soil as pure culture. It can only survive and grow only in presence of other cultures. Bacterial survival depends on both abiotic and biotic factors. In the sterile soils no biological activities are possible. Therefore, secretion of organic acids, which serve as carbon source for bacteria will not be present. In absence of such carbon source growth of inoculum is not possible. Interestingly, *S. granuli* MG proliferated, only in the presence of *Cupriavidus* sp TKC, *Burkholderia* Bcrs1W and *Paenibacillus* sp NK-L2. This indicated that remediation of aromatic pollutants like nitrophenols in soil using *S. granuli* MG may be

successful with a mixed microcosm containing soil. *S. granuli* MG has the potential to mineralize nitrophenols as seen from the radioactivity experiments, however, its survival and growth is best seen in soils with other supporting bacteria.

The ability of microcosms of *S. granuli* MG with *Paenibacillus* sps was further used to investigate *in-situ* degradation of ^{14}C labelled PNP. The incorporation of evolved CO_2 as labelled $\text{Ca}^{14}\text{CO}_3$ demonstrates mineralization of the labelled PNP in the soil environment. This shows that the parent compound PNP was completely transformed and the radioactivity was distributed to various unidentified intermediates and also to the final by product CO_2 . Analysis of microbial degradation also showed that *in-situ* PNP transformation is majorly contributed by bacterial activity of the microcosm included in this study. The studies described in this chapter thus indicate the bioremediation potential of PNP contaminated soils.

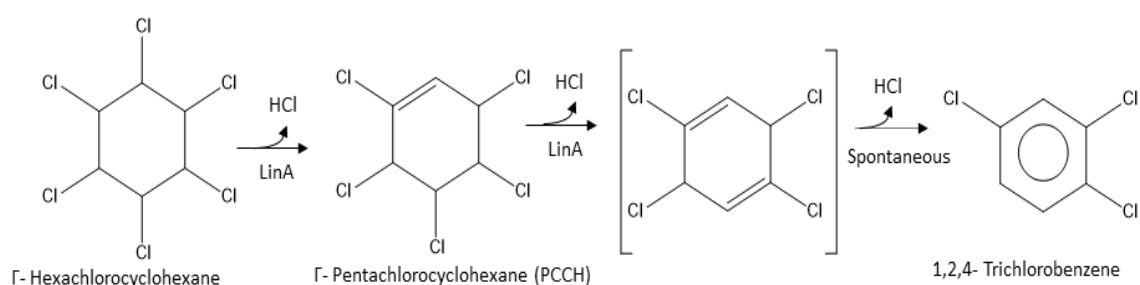
CHAPTER 5 : Novel dehalogenases in the genome of *S. granuli* MG

5.1. Introduction

As highlighted in Chapter IV, genome guided catabolic pathways have revealed existence of a highly versatile catabolic repertoire in *S. granuli* MG. In support of the *in-silico* predictions, the organism mineralizes toxic nitrophenols and survives on para-nitrophenol as lone source of carbon. The genome mining strategy employed in this study also revealed existence of novel dehalogenases in the genome of *S. granuli* MG. Bacterial dehalogenases catalyse cofactor independent dehalogenation and remove halogen substituent from many recalcitrant and toxic pollutants and transform them into their non-toxic counter parts. Based on the mechanism of dehalogenation, the reaction catalysed by dehalogenases are divided into different classes viz. i) reductive dehalogenation, ii) oxygenolytic dehalogenation, iii) dehydrohalogenation and iv) hydrolytic dehalogenation. Among them the hydrolytic dehalogenases, especially members of the haloalkane dehalogenases have also been recruited in several practical applications like i) bioremediation of environmental pollutants (Dvorak, Bidmanova, Damborsky, & Prokop, 2014; Lal *et al.*, 2010; Nagata, Ohtsubo, & Tsuda, 2015; Stucki & Thueer, 1995) ii) removal of warfare chemicals (Prokop, Oplustil, DeFrank, & Damborsky, 2006) iii) biosensing of toxic contaminants (Bidmanova, Chaloupkova, Damborsky, & Prokop, 2010; Campbell, Muller, & Reardon, 2006) iv) biocatalytic synthesis of enantiopure chemicals (Schober & Faber, 2013; Szymanski, Westerbeek, Janssen, & Feringa, 2011) and also v) in cell/protein imaging analysis and protein labelling (Koudelakova *et al.*, 2013; Los *et al.*, 2008).

Eight dehalogenases were identified in the genome of *S. granuli* MG. Of these six were dehydrohalogenases similar to the γ -hexachlorocyclohexane dehydrochlorinase also designated as LinA. The LinA catalyzes first two dechlorination steps in the biodegradation of γ -hexachlorocyclohexane (γ -HCH) widely popular as Lindane. Lindane is widely used pesticide in agriculture and in control of vector borne diseases as public health measure (Lal *et al.*, 2010; Nagata, Endo, Ito, Ohtsubo, & Tsuda, 2007). However, its rampant production and use over the years has led to accumulation of γ -HCH and other isomers in soils as a recalcitrant pollutant (Berger *et al.*, 2016; Y. F. Li, Scholtz, & Van Heyst, 2003). The dechlorination is the

first and very significant step as it facilitates further degradation of Lindane into other non-toxic intermediates or to its mineralization. The LinA-dehydrochlorinase is approximately 17 kDa protein that catalyses conversion of γ -HCH to γ -pentachlorocyclohexane (γ -PCCH) (Nagata, Futamura, Miyauchi, & Takagi, 1999). As shown in figure 5.1, the dehydrochlorination reaction by LinA removes HCl group from γ -HCH giving rise to the stable product 1,2,4 Trichlorobenzene (1,2,4-TCB) containing double bond. LinA enzyme was identified in the aerobic γ -HCH degrader *S. japonicum* UT26 which has the complete γ -HCH degradation pathway genes (Nagata, Miyauchi, & Takagi, 1999). In this chapter an attempt is made to identify a novel γ -HCH- dehydrochlorinase in the genome of *S. granuli* MG.



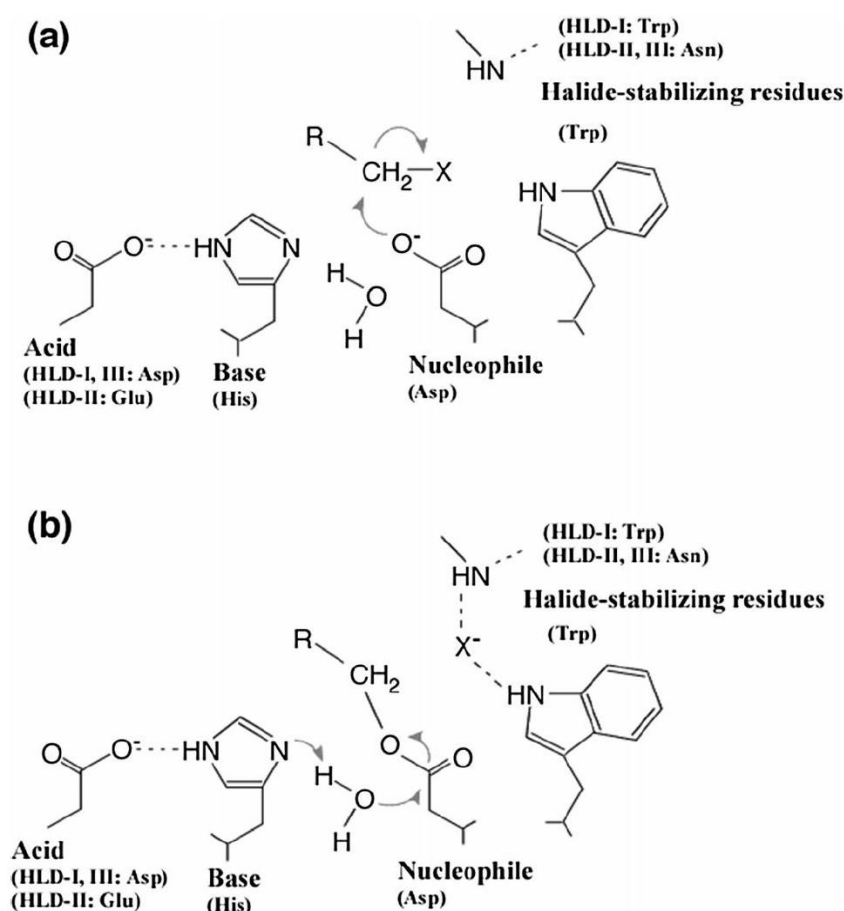
(Imai, Nagata, Fukuda, Takagi, & Yano, 1991).

Fig. 5.1 The γ -HCH degradation pathway catalysed by LinA

Out of the eight dehalogenases identified in the genome of *S. granuli* MG, two dehalogenases belonged to the class-haloalkane dehalogenases (HLDs). Haloalkane dehalogenases (HLDs) catalyse the hydrolytic cleavage of carbon-halogen bond in various halogenated compounds. They belong to a large superfamily of enzymes that contain typical α/β -fold hydrolase fold that perform epoxide hydrolase, esterase, hydroxy nitrile lyases, peroxidases and proteases apart from dehalogenase activities (Koudelakova *et al.*, 2011). Phylogenetically HLDs have been classified into three HLD-I, HLD-II and HLD-III subgroups (Chovancova, Kosinski, Bujnicki, & Damborsky, 2007). The classification is based on variations in the active site catalytic triad and two halide stabilization residues. The Asp-His-Asp serves as catalytic triad and Trp-Trp as halide stabilizing residues in HLD-I. In HLD-II Asp-His-Glu and Asn-Trp serve as catalytic and halide stabilization residues, whereas in HLD-III Asp-His-Asp and Asn-Trp residues were identified at similar positions. In contrast to HLD-I and HLD-II, the subgroup HLD-III is not well distinguished (Chovancova *et al.*, 2007).

Structurally HLDs consists of a catalytic central core domain and a cap domain. The core domain has the α/β -fold structure and cap domain extend as an R-helical protrusion between the 6th and 7th β -strand of the core domain. At the crossing between these two domains is an active site cavity which is hydrophobic and possesses the catalytic triad, Aspartate (Asp-nucleophile), Histidine (His-base), Aspartate/Glutamate (Asp/Glu-acid catalyst) involved in dehalogenation. In contrast to the core domain that has high sequence similarity across subgroups, there exists significant sequence and structure variation in the cap domain of HLDs. Therefore, the cap domain is the crucial element determining specificity of substrates HLDs (Koudelakova *et al.*, 2013).

The reaction mechanism of HLDs follows SN_2 nucleophilic substitution with subsequent hydrolysis requiring water as cofactor (fig 5.2). Initially, from the catalytic triad residues the carboxylate oxygen of the nucleophile Asp begins an attack on the carbon of the carbon-halogen substrate ($R-CH_2X$). This produces an ester bond containing alkyl-enzyme intermediate and a halide ion (X^-). Subsequently the acid- base pair, His-Asp/Glu hydrolyze the water molecule producing OH^- ions that attacks the carbon of the ester bond. A transition state tetrahedral intermediate ($R-CH_2O^-$) is generated that reacts with proton (H^+) of nucleophile forming a stable alcohol ($R-CH_2-OH$). The halide stabilization residues are a unique feature of haloalkane dehalogenases that bind to the halide ion during the formation of transition complex (Damborsky & Koca, 1999). Here, in this chapter the experiments conducted to assess the catalytic functions of *S. granuli* MG haloalkane dehalogenases ($_{sg}HldA$ and $_{sg}HldB$) are described.



(Nagata et al., 2015).

Fig. 5.2 Reaction mechanism of HLDs

The part (a) depicts the ester formation by S_N2 nucleophilic substitution with the stabilization of transition state by halide stabilizing residues. In (b) water molecule activated by His-Asp pair attacks the ester intermediate to form alcohol and halide ion

5.2. Objective specific methodology

Table 5.1 Bacterial strains used in Chapter V

Strains	Characteristics	Reference
<i>E. coli</i> DH5 α	<i>supE44, lacU169, (lacZ, M15) hsdR17 recA1 endA1 gyrA96 thi1 rel A1</i>	(Hanahan, 1983)
<i>E. coli</i> BL21(DE3)	<i>F ompT hsdSB (rB-mB-) gal ompT, λ(DE3)</i>	(Studier & Moffatt, 1986)
<i>Sphingopyxis. granuli</i> MG	Wild type Strain, PolB r, <i>opd</i> ⁺	(Parthasarathy, Azam, et al., 2017)

Table 5.2 Plasmids used in Chapter V

Plasmid Name	Genotype or Phenotype	Reference
pRSETA	Ap ^r , expression vector, cloned genes are expressed from a T7 promoter with N- terminal 6xHis-tag	Invitrogen
pAL1	Ap ^r , generated by cloning 450 bp <i>linA1</i> gene of <i>S. granuli</i> MG as <i>Bam</i> H1- <i>Pst</i> 1 fragment under the control of T7 promoter of pRSETA. Codes for <i>sgAL1</i> ^{N6His}	This study
pAL2	Ap ^r , generated by cloning 450 bp <i>linA2</i> gene of <i>S. granuli</i> MG as <i>Bam</i> H1- <i>Pst</i> 1 fragment under the control of T7 promoter of pRSETA. Codes for <i>sgAL2</i> ^{N6His}	This study
pAL3	Ap ^r , generated by cloning 396 bp <i>linA3</i> gene of <i>S. granuli</i> MG as <i>Bam</i> H1- <i>Eco</i> R1 fragment under the control of T7 promoter of pRSETA. Codes for <i>sgAL3</i> ^{N6His}	This study
pAL4	Ap ^r , generated by cloning 399 bp <i>linA4</i> gene of <i>S. granuli</i> MG as <i>Bam</i> H1- <i>Eco</i> R1 fragment under the control of T7 promoter of pRSETA. Codes for <i>sgAL4</i> ^{N6His}	This study
pAL5	Apr, generated by cloning 561 bp <i>linA5</i> gene of <i>S. granuli</i> MG as <i>Sac</i> 1- <i>Eco</i> R1 fragment under the control of T7 promoter of pRSETA. Codes for <i>sgAL5</i> ^{N6His}	This study
pAL6	Ap ^r , generated by cloning 438 bp <i>linA6</i> gene of <i>S. granuli</i> MG as <i>Bam</i> H1- <i>Pst</i> 1 fragment under the control of T7 promoter of pRSETA. Codes for <i>sgAL6</i> ^{N6His}	This study
pWD5	Ap ^r , <i>sghldA</i> gene as <i>Nde</i> I- <i>Xho</i> I fragment in expression plasmid pET22b (+). Codes for <i>sgHldA</i> ^{N6His}	Gift from Prof. Nagata, Tohoku University, Japan
pWD6	Ap ^r , <i>sghldB</i> gene cloned as <i>Nde</i> I- <i>Xho</i> I fragment in expression plasmid pET22b (+). Codes for <i>sgHldB</i> ^{N6His}	Gift from Prof. Nagata, Tohoku

		University, Japan
pWD6 ^{F37N}	Ap ^r , variant of <i>sgHldB</i> gene generated by site directed mutagenesis substituting phenylalanine at 37th position with asparagine. Codes for <i>sgHldB</i> ^{F37N}	This study

Table 5.3 Primers used in Chapter V

Primer	Sequence 5'-3'	Purpose
pAL 1 FP	ACTATCGGATCCATGAACGAACTC GACAAGATGCTCATC	Forward primer amplifying gene <i>linA1</i> as <i>Bam</i> H1 and <i>Pst</i> 1 fragment
pAL 1 RP	ACTATCCTGCAGCTAGCGTGCGTCC AGATCGATCAC	Reverse primer amplifying gene <i>linA1</i> as <i>Bam</i> H1 and <i>Pst</i> 1 fragment
pAL 2 FP	ACTATCGGATCCATGCGGAGCGAA GACGAACTGCTCA	Forward primer amplifying gene <i>linA2</i> as <i>Bam</i> H1 and <i>Pst</i> 1 fragment
pAL 2 RP	ACTCTAGAATTCTCAGGCAGGCGG CGCGGATTTC	Reverse primer amplifying gene <i>linA2</i> as <i>Bam</i> H1 and <i>Pst</i> 1 fragment
pAL 3 FP	ACTATCGGATCCATGGCAGAACAG GACGATATCC	Forward primer amplifying gene <i>linA3</i> as <i>Bam</i> H1 and <i>Eco</i> R1 fragment
pAL 3 RP	ACTATCGAATTCTCAGGCGTCGGC ATATTTGCG	Reverse primer amplifying gene <i>linA3</i> as <i>Bam</i> H1 and <i>Eco</i> R1 fragment
pAL 4 FP	ACTATCGGATCCATGAGCGGTCCT ACGGGGCT	Forward primer amplifying gene <i>linA4</i> as <i>Bam</i> H1 and <i>Eco</i> R1 fragment
pAL 4 RP	ACTATCGAATTCTCAGCCCGCCTGC GCG	Reverse primer amplifying gene <i>linA4</i> as <i>Bam</i> H1 and <i>Eco</i> R1 fragment
pAL 5 FP	ACTATCGAGCTCATGGACAAGCGG ATCGAAGACCT	Forward primer amplifying gene <i>linA5</i> as <i>Sac</i> 1 and <i>Eco</i> R1 fragment

pAL 5 RP	ACTCTAGAATTCCTATCGCCTCCCG ACCTCGG	Reverse primer amplifying gene <i>linA5</i> as <i>Sac1</i> and <i>EcoR1</i> fragment
pAL 6 FP	ACTATCGGATCCATGAGCGACCTC GACAAACGGCGG	Forward primer amplifying gene <i>linA6</i> as <i>BamH1</i> and <i>Pst1</i> fragment
pAL 6 RP	ACTATCCTGCAGTCAGAACAGGTC GAGCGCCTTCATCA	Reverse primer amplifying gene <i>linA6</i> as <i>BamH1</i> and <i>Pst1</i> fragment
F37N FP	CTCCACGGAAATCCCGAATCCACC GCACCTGGCGC	Primers used to generate HldB variant HldB ^{F37N} by performing site directed mutagenesis. The codon introduced to substitute phenylalanine at position 37 to asparagine is shown in bold case
F37N RP	CGATTGCGGATTTCCTGGAGAAG ACCAGCGCCGG	

5.2.1. Identification of LinA-HCH dehydrochlorinase homologues in *S. granuli* MG

The reference strain used for searching LinA-HCH dehydrochlorinase homologues was *Sphingobium japonicum* UT26, a soil isolate known for possessing the complete γ -HCH degradation pathway genes (Camacho-Perez, Rios-Leal, Rinderknecht-Seijas, & Poggi-Varaldo, 2012; Nagata, Miyauchi, *et al.*, 1999). Initially, all proteins deduced from the genome sequence of *S. granuli* MG were downloaded from the NCBI database and the proteins containing LinA-dehydrochlorinase like domain, as seen in *S. japonicum* UT26, were taken in a separate folder. Further the proteins containing dehydrochlorinase specific fold were aligned with LinA sequence of *S. japonicum* UT26 LinA to ascertain percent similarity. Particularly, the active site residues were marked in the sequences of predicted LinAs of *S. granuli* MG.

5.2.2. Expression of *linA* gene in *E. coli*

After identifying *linA* homologs in *S. granuli* MG further studies were conducted to identify their catalytic functions. Initially *linA* gene was cloned in expression vector pRSETA to express LinA in *E. coli* with an N-terminal histidine tag. The plasmid pRSETA was digested with appropriate restriction enzymes according to the insert to be cloned. Five putative *linA* genes named as *linA1*, *linA2*, *linA3*, *linA4*, *linA5* and *linA6* were amplified and the amplicons were

ligated to the pRSETA. The resulting recombinant plasmids were named as pAL1, pAL2, pAL3, pAL4, pAL5 and pAL6 respectively. These plasmids were introduced independently in to *E. coli* BL21 DE3 cells and the expression of cloned genes were performed following standard procedures. The recombinant strains carrying the respective plasmids were incubated at 37°C till the OD of the culture is reached to 0.5 OD ($A_{600\text{ nm}}$) and 1ml of culture was taken out from the culture and the cells were harvested and used as uninduced sample. The remaining culture was induced with 1 mM IPTG and the cells were further grown for 3-4 hr at 30°C. Following induction, the cells were harvested from 1 ml culture and the cells were suspended in appropriate amounts of 2X SDS loading dye and boiled for 10-15 min before analysing the sample on 12.5% SDS-PAGE. The cell pellet collected from uninduced cells served as controls.

5.2.3. Dehalogenase activity

The dechlorinase activity in the whole cells expressing *sg*LinAs (*sg*AL1^{N6His}, *sg*AL2^{N6His}, *sg*AL3^{N6His}, *sg*AL4^{N6His}, *sg*AL5^{N6His} or *sg*AL6^{N6His}) was performed as described in methods section. Initially the whole cells (30 mg) were preincubated with 890 μ l of 100 mM glycine pH 8.2 at 30°C. Then 10 μ l of γ -HCH (34.3 mM) dissolved in DMSO was added to start the reaction. The contents were thoroughly mixed by vortexing before incubating the contents at 30°C. At appropriate time points (0, 60, 360 and 1440 min) a sample of 200 μ l was withdrawn from the reaction mix and 20 μ l of solution I - Hg (SCN)₂ was added. The sample was again vortexed for 30 sec before adding 40 μ l of solution II - (NH₄Fe (SO₄)₂·12 H₂O). After brief vortex, the mixture was thoroughly centrifuged at 12,000 rpm for 10 min to remove the insoluble particles. About 200 μ l of the supernatant was taken in a 96 well ELISA Microplate (Nunc) and absorbance was measured at 450 nm using a MULTISKAN™ GO microplate spectrophotometer from Thermo-Scientific, USA. The concentration of chloride ion released from whole cells was calculated from standard graph generated using known amounts of KCl.

5.2.4. Affinity Purification of *sg*AL5^{N6His}

The *sg*AL5^{N6His} was purified using nickel sepharose affinity column following standard protocols. Briefly, the induced cell pellet was washed and resuspended in 10 ml of equilibration buffer (20 mM potassium phosphate, 200 mM NaCl pH 7.4). The resuspended solution was then incubated for 1 hr on an ice bucket after adding 10 μ l of lysozyme (50

mg/mL). The cells were then lysed using sonication for 13 cycles with a 20 sec ON and 40 sec OFF cycle at 4°C. The resulting homogenate was centrifuged at 15,000 rpm for 20 min to separate cells debris and the supernatant obtained was used for purification of protein. The supernatant sample was passed through the pre-equilibrated Ni-NTA affinity column. The protein loaded column was then extensively washed (about 4-5 column volumes) with wash buffer (20 mM potassium phosphate, 500 mM NaCl, 20 mM imidazole) before eluting the bound proteins with elution buffer (20 mM potassium phosphate, 500 mM NaCl, 300 mM imidazole). The eluate *sgAL5*^{N6His} protein collected was dialyzed against 50 mM Phosphate buffer (pH 7.5). The dialyzed protein sample was concentrated using Vivaspin 2 (Sartorius) following the manufacturer's protocol. Briefly the Vivaspin tubes were washed with Milli-Q water (5000 g, 5 min) after which the tubes were washed with the concentration buffer which was the same as dialysis buffer (5000 g, 5 mins). The sample was concentrated to 200 µL and stored as aliquots of 10 µL in 50 mM Phosphate buffer (pH 7.5) with 15% glycerol at -20°C until further use. The concentration of pure *sgAL5*^{N6His} was determined using Bradford method using BSA as standard.

5.2.5. Dehalogenase activity of *sgLinA5*^{N6His}

The purified *sgAL5*^{N6His} protein was used to also quantify the reaction using Gas Chromatograph containing a ⁶³Ni electron capture detector (GC-ECD). Initially, 890 µL of 100 mM glycine pH 8.2 solution was preincubated with pure protein (0.05 mg/mL) at 30°C. Then the 10 µL of γ-HCH (34.3 mM) was added starting the reaction. The reaction mixture was incubated at 30°C for time points 0 and 12 hr. At the given time points 100 µL of reaction mixture was extracted with the 100 µL of ethyl acetate with dieldrin (2 ppm, internal standard). The samples were centrifuged at 12,000 rpm for 10 min and the upper organic layer was subjected to GC-ECD analysis. GC equipped with a Rtx-1 capillary column (30 m×0.25 mm×0.25 µm; Restek) was operated under following conditions. The temperature at injection was 260°C and the temperature of column was increased from 160°C to 280°C at 20°C/ min. The flow rate of gas was maintained at 30 mL/min. Using standard graph from commercially pure chemicals, the concentrations of γ-HCH and products were determined.

5.2.6. Multiple sequence alignment of haloalkane dehalogenase

The two haloalkane dehalogenases *sg*HldA and *sg*HldB identified by genome mining in *S. granuli* MG were subjected to multiple sequence alignment using Clustal Omega (Madeira et al., 2019). For each subgroup one representative HLD was considered for multiple sequence alignment. Further, the catalytic triad and halide binding residues typical of haloalkane dehalogenases were also identified in both *sg*Hlds.

5.2.7. Phylogenetic classification of haloalkane dehalogenases

Based on the phylogenetic classification of other known haloalkane dehalogenase, representative proteins of three subfamilies were used to identify the subfamily of *sg*Hlds. All the protein sequences were aligned using Clustal W of MEGA11 (Kumar, Stecher, Li, Knyaz, & Tamura, 2018). The aligned sequences were then used to construct a phylogenetic tree using Neighbourhood-Joining (NJ) Method.

5.2.8. Purification of *sg*HldA and *sg*HldB

The expression plasmids pWD5 expressing *sg*HldA and pWD6 expressing *sg*HldB were used for protein purification respectively. The N-terminal his-tagged plasmids were constructed by using pETWD1 (modified pET22b (+) vector with an insert translational enhancing element (TEE) and tobacco etch virus (TEV) protease sequence). The recombinant plasmids were then transformed into *E. coli* BL21 NICO strain. The transformants were then incubated at 37°C until OD reached 0.5, induced with 0.5 mM IPTG and incubated further for 12 hr at 23°C for expression of proteins. The N-terminal his-tagged proteins were purified using nickel sepharose affinity column using the standard protocol. Briefly, after induction the cell pellets were washed and resuspended in buffer with 50 mM potassium phosphate, 200 mM NaCl pH 7.5. The resuspended cells, were incubated with 50 µg/ml lysozyme at 4°C for 1 hr on rotation. The cells were then disrupted by sonication for 13 cycles with a 30 sec 'ON' and 30 sec 'OFF' cycle at 4°C. The resulting homogenate was centrifuged at 13,000 rpm for 15 min to remove cells debris. The supernatants thus obtained were used for purification of protein. The supernatant samples were passed through the pre-equilibrated Ni-NTA affinity columns and then washed (4-6 column volumes) with wash buffer (50 mM potassium

phosphate, 500 mM NaCl, 20 mM imidazole). The *sgHldA* bound to column was eluted in 50 mM potassium phosphate, 500 mM NaCl, 300 mM imidazole containing elution buffer. Similarly, the elution buffer with 50 mM potassium phosphate, 500 mM NaCl, 200 mM imidazole was used to elute *sgHldB*. The eluted samples of *sgHldA* and *sgHldB* were collected and independently dialyzed against 50 mM Phosphate buffer (pH 7.5). The dialyzed protein samples were then separately concentrated using Vivaspin 2 (Sartorius) following the manufacturer's protocol. Briefly the Vivaspin tubes were washed with milli-q water (5000 g, 5 min) followed by concentration buffer (5000 g, 5 mins). The samples were concentrated to 200 μ L and stored as aliquots of 10 μ L in 50 mM phosphate buffer (pH 7.5) with 15% glycerol at -20°C until further use. The concentrations of pure *sgHldA*^{N6His} and *sgHldB*^{N6His} was determined using Bradford method using BSA as standard.

5.2.9. Substrate specificity of *sgHldA*, *sgHldB* or *sgHldB*^{F37N}

To assess the specific dehalogenase activity of pure *sgHldA*, *sgHldB* or *sgHldB*^{F37N} against halogenated compounds, continuous spectrophotometric dehalogenase assay using phenol red was used as described in the materials and methods. Initially 190 μ L of reaction sample consisting of 20 μ L of 50 mM HEPES (1 mM) pH 8.0, 20 μ L of 50 mM EDTA (1 mM), 100 μ L of 100 mM sodium sulphate (20 mM), and 20 μ L of 10 mg/ml phenol red (20 μ g/mL) was added in each well of 96 well ELISA Microplate (Nunc) and preincubated with appropriate concentration of 20 μ L of *sgHldA* (1 mg/ml), *sgHldB* or *sgHldB*^{F37N} (10 mg/ml) at 30°C in triplicates. Then 10 μ L of substrates (10 mM for *sgHldA* and 1 mM for *sgHldB*) was added to start the reaction. The *sgHldB*^{F37N} was tested against 1-Bromobutane (1 mM). The substrates had varying chain length, halides and halide position varied from substrate to substrate (Table 2.8). The microtitre plate was then inserted into the given slot of a MULTISKAN™ GO microplate spectrophotometer, Thermo-Scientific, US and absorbance at was continuously monitored for one minute at 30°C. The H⁺ ions released into the HEPES buffer by the was visualized at absorbance λ = 540 nm and quantified using a standard graph containing 0-1000 nmol H⁺ ion concentration. The specific activity for each substrate was calculated as the amount of protein consumed to release 1 nmol of H⁺ ions in 1 min.

5.2.10. Homology modelling & visualizing active site architecture

The homology modelling was conducted using the SWISS-MODEL web server under the Automated mode for both dehalogenases. Additionally, for *sgHldA* manual mode was also used. The PDB structure of *sgHldA* and *sgHldB* generated were visualized using PyMOL to predict the active site architecture.

5.2.11. Generation of *sgHldB* variant by site-directed mutagenesis

The variant of *sgHldB* was generated by performing site-directed-mutagenesis using Q5-site-directed mutagenesis kit (NEB). The plasmid pWD6 encoding *sgHldB*^{WT} was used as the template to perform the PCR and the primer pair F37N FP/F37N RP corresponding to the *hldB* gene that specify insertion of desired mutation that substitutes phenylalanine to asparagine was at 37th position was designed (Table 5.3). The PCR amplified products were treated with KLD enzyme mix (containing a kinase, a ligase and *Dpn* I) and incubated at room temperature for 30 min. The PCR products were then transformed using chemically competent *E. coli* DH5 α cells and spread on 100 μ g/ml ampicillin containing LB plates. The plasmid DNA was isolated from the individual colony of transformants and confirmed by sequencing. The plasmid that contains AAT instead of TTT was designated as pWD6^{F37N} and was used for further studies. The plasmid pWD6^{F37N} encodes for the variant *sgHldB*^{F37N}.

5.2.12. Purification of *sgHldB*^{F37N}

In the section 5.1.7, detailed protocol used for purification of *sgHldB*^{WT} were described. The same protocol was followed for purification of *sgHldB*^{F37N}. Fresh overnight cultures of *E. coli* BL21 NICO containing plasmid pWD6^{F37N} were inoculated in 500 mL of LB medium containing 100 μ g/ml ampicillin at 37°C. At 0.5 OD, 0.5 mM IPTG was added and further incubated for 12 hr at 23°C. The induced cell was washed, pellets were resuspended in 50 mM potassium phosphate, 200 mM NaCl buffer pH 7.5. Lysozyme (50 μ g/ml) was added and resuspended cells were kept at 4° C for 1 hr on rotation. The cells were then disrupted by sonication for 13 cycles with a 20 sec 'ON' and 40 sec 'OFF' cycle at 4°C. The resulting homogenate were centrifuged at 13,000 rpm for 15 min to remove cells debris. The supernatants thus obtained were used for purification of protein. The supernatant samples

were passed through the pre-equilibrated Ni-NTA affinity columns. The column was then extensively washed (50 mM potassium phosphate, 500 mM NaCl, 10 mM imidazole). The bound protein was eluted with elution buffer containing 50 mM potassium phosphate, 500 mM NaCl, 200 mM imidazole and dialysed against 50 mM Phosphate buffer (pH 7.5). The dialyzed protein samples were then separately concentrated using Vivaspin 2 (Sartorius) as described elsewhere. The samples were concentrated to 200 μ L and stored as aliquots of 10 μ L in 50 mM Phosphate buffer (pH 7.5) with 15% glycerol at -20°C. The concentration of pure *sg*HldB^{F37N} was determined using Bradford method using BSA as standard.

5.2.13. Activity of *sg*HldB or *sg*HldB^{F37N} with epoxides

The epoxide hydrolase activity of purified *sg*HldB or its variant were determined by performing spectrometric assay as described in methods. Initially, 80 μ L of 0.5 M phosphate buffer containing 1 mM of periodate (NaOI₄) was added to control and test wells of a 96 well microplate. Further, in all test wells 10 μ L (0.05 mg) purified *sg*HldB or *sg*HldB^{F37N} was added, the wells having no enzyme or substrate served as controls. The addition of 10 μ L of aromatic or aliphatic epoxides (10 mM) started the reaction. The reactions were performed at 30°C for 30 min. When incubation process was in progress 15 μ L of 10 mM L-adrenaline stock solution was taken in a fresh microtiter plate. Soon after incubation 85 μ L of the reaction mix was taken and added to the wells containing L-adrenaline solution. The red-coloured adrenochrome formed due to L-adrenaline interaction with periodate (NaOI₄) was used to indirectly quantify of the diols formed. The adrenochrome formed was measured at λ = 490 nm using a MULTISKAN™ GO microplate spectrophotometer from Thermo-Scientific, USA. The specific activity was calculated as the μ mol of diols formed per minute per mg of enzyme from a standard plotted using different diols.

5.3. Results

5.3.1. Putative *linA*-dehydrochlorinases in *S. granuli* MG

LinA-HCH dehydrochlorinase of *Sphingobium japonicum* UT26 (LinA_UT26) shows a distant relationship with scytalone dehydratase, nuclear transport family-2 and 3-oxo-steroid isomerase protein domains, indicating a common superfamily of proteins that have diverged extensively and evolved to perform different functions but have retained general design of

the active site cavity (Nagata, Mori, Takagi, Murzin, & Damborsky, 2001). Moreover, docking experiments using crystal structure of LinA_UT26 with its inhibitor was compared with scytalone dehydratase and its substrate to predict the common fold and conserved catalytic His(H)-73/ Asp(D)-25 of LinA_UT26 (Okai et al., 2010). Therefore, the genome of *S. granuli* MG was searched for the presence of genes with similar domains and reported catalytically active residues. Genome mining led to the identification of six *linA* homologs based on domain similarity. These were named as *sglinA1-sglinA6* respectively. In LinA isolated from *S. japonicum* UT26, the residue R129 along with H73 and D25 is also catalytically important for LinA activity (Nagata et al., 2001). However, D25 and H73 are highly conserved and form the catalytic dyad. To gather prima-facie evidence multiple sequence alignment of all putative LinA amino acid sequences from *S. granuli* MG was performed with LinA_UT26. Interestingly, *sglinA1*, *sglinA3* and *sglinA5* of *S. granuli* MG showed the presence of all three active site residues D25, H73 and R129 (fig 5.3).

<i>sglinA4</i>	-----MSGPTGLAAWHAYMA-----GGDPQAL-	23
<i>sglinA2</i>	-MAEQDDIR--AMAQRFF-DAIEAGDIETMKASFAPGAEIWHNSDEAI-VTPEQTAATL-	54
<i>sglinA_UT</i>	-MSDLDRLASRAAIQDLYSDKLI---AVDKRQEGRLASIWDDAEWT-IEGIGTYKGP-	53
<i>sglinA6</i>	-MSDLDKRRIETVKRHMALE-----IAQDWD-GVL-ATFDHPRYELLG	40
<i>sglinA5</i>	MDKRIEDLLDKQALYENMMTYCR---AIDRMDIELMKSTYWPDGTD---DHGRFAGN-	51
<i>sglinA3</i>	-MRSEDELLIEAACRKLMTTRYCV---GLDTRDADLFSLVFSADAIWELVNDTPR-TLA-	53
<i>sglinA1</i>	-MNELDKMLIERECHRLMVDYNI---YNDTLDIEAFLTLWTEDCVFSRVVPEPVFEVS-	54
.		
<i>sglinA4</i>	-----RDLLA--EDAVFHSPVVHTDQGRDVFAYLHAASHVLGG-EHFRYLREIVD	72
<i>sglinA2</i>	-----TGMVARIKDRAYAERQL-HVFPG-----GFVQQ-----HVLKGRRVH	90
<i>sglinA_UT</i>	-----EGALDLANNVLWPMFHE-CIHVGTNLRLEFVSAD-KVNGI-GDVLLLGNLVE	102
<i>sglinA6</i>	PGTVFDGEAAVRSYFASSREP-----FPDQANEIIA-IAADAATDTVLVEFWLTG-THL	92
<i>sglinA5</i>	-----AHEWCEVAIKS--REALVS-CNHLSNVLIE-LDGDRAQRES-MFLVV-TTYKD	99
<i>sglinA3</i>	-----GQDAIRSYFES--RQKKVF-NUHLLQNPLVT-VTGPTAEAGYCIALIVDGPWGD	103
<i>sglinA1</i>	-----GHAGLRAAMDQIIIVHSDRL-RRHLLVNPRIT-VHGPDDAEGFCIGLAITYGPAGD	106
:		
<i>sglinA4</i>	G-----DQACLEFASELD---GIHING	91
<i>sglinA2</i>	DDVAVRLPCAIV-----CRVEDGKIARLDEYFDSAHVAAFRKYADA-----	131
<i>sglinA_UT</i>	GNQSILI-----AAVFTDEYERRDGVWKFSSRNACTNYFTPLA---GIHFAP	146
<i>sglinA6</i>	G--PLRLGLRTVEPTGKAFRIRMAASFEPAGSDRIVCER---PYDQSAVMKALDLF--	145
<i>sglinA5</i>	D--P-----KHQFLG-----GRYRDLCERRHGEWRVLRATCIWDWNRKIVAEPGWDIMQA	147
<i>sglinA3</i>	G--ALPVLRLGTEMV-----TEYRDRFELRPEGWRIVRRITMRKLIDKKSA-PPA-----	149
<i>sglinA1</i>	R--SLPVAMGGVELV-----GEYRDQYRRRTSEGWKIARRELTRVIDLDAR-----	149
:		
<i>sglinA4</i>	VDIIRWNDEGKIQDFKVMVRPLKAINKVWEKMAAMLAAQAG-----	132
<i>sglinA2</i>	-----	131
<i>sglinA_UT</i>	PGIH- FAPSGA-----	156
<i>sglinA6</i>	-----	145
<i>sglinA5</i>	PEMSNWGQFHP-----KDPIYAADWGAIEPTSAASGRPEVGRR	186
<i>sglinA3</i>	-----	149
<i>sglinA1</i>	-----	149

Fig. 5.3 Multiple sequence alignment of LinA_UT26 with *S. granuli* MG LinA homologs

Six homologs of *S. granuli* MG were aligned with LinA_UT26 using CLUSTAL OMEGA. The active site residues D25, H73, R129 from LinA_UT26 aligned with the corresponding active site amino acid residues in LinA homologs are shown in black boxes.

5.3.2. Heterologous expression and activity assays of *sgLinA*

The bioinformatic analysis revealed that out of six LinA homologues only three have shown conservation of active site residues. These three LinAs are, *sglinA1*, *sglinA3* and *sglinA5* of *S. granuli* MG. Nevertheless, all the lin genes were cloned and expressed to identify functional *linA* gene in *S. granuli*. The *linA1-linA6* genes were amplified from the genome using the primers as mentioned in Table 5.3 and the amplified gene products were digested with *Bam*H1 and *Pst*1 for *sglinA1*, *sglinA2* and *sglinA6*; *Bam*H1 and *Eco*R1 for *sglinA3* and *sglinA4*; *Sac*1 and *Eco*R1 for *sglinA5* and cloned in pRSETA. The resulting recombinant plasmids are named as pAL 1, pAL 2, pAL 3, pAL 4, pAL 5, and pAL 6. The details pertaining to the construction of expression plasmids coding *sglinAs* are shown in fig 5.4.

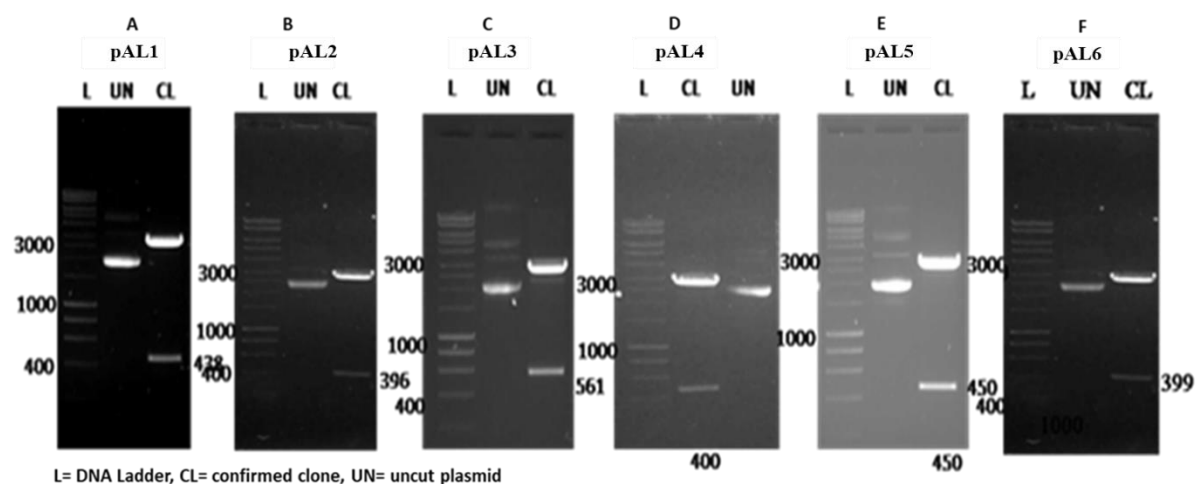


Fig. 5.4 Restriction digestion profiles of *sglinA* expression plasmids.

The panel A-F (lane-CL) shows release of *linA* genes from vector pRSETA after digesting them with restriction endonucleases. Panel A shows release of *sglinA1* after digesting the recombinant plasmid pAL1 with *Bam*H1 and *Pst*1. Panel B shows release of *sglinA2* after digesting the recombinant plasmid pAL2 with *Bam*H1 and *Pst*1. Panel C shows release of *sglinA3* after digesting the recombinant plasmid pAL3 with *Bam*H1 and *Eco*R1. Panel D shows release of *sglinA4* after digesting the recombinant plasmid pAL4 with *Bam*H1 and *Eco*R1. Panel E shows release of *sglinA5* after digesting the recombinant plasmid pAL5 with *Sac*1 and *Eco*R1. Panel F shows release of *sglinA6* after digesting the recombinant plasmid pAL6 with *Bam*H1 and *Eco*R1. Lanes L and UN represent ladder and uncut plasmid respectively.

5.3.3. Expression and purification of LinAs

After successful construction of recombinant plasmids, they were independently transformed into *E. coli* BL21 DE3 and induced their expression following the protocols described in methods section. The expression of LinA homologues were checked by analysing the proteins of the induced cultures using western blot with anti-his antibodies (fig 5.5). The western blot profile clearly showed the expression of $sgAL1^{N6His}$ (17.04 kDa) in fig 5.3 panel A lane 3&4; $sgAL2^{N6His}$ (16.9 kDa) fig 5.3 panel B lane 4&5; $sgAL3^{N6His}$ (14.6 kDa) fig 5.3 panel C lane 2&3; $sgAL4^{N6His}$ (14.4 kDa) fig 5.3 panel D lane 4&5; $sgAL5^{N6His}$ (21.6 kDa) fig 5.3 panel E lane 1&2; and $sgAL6^{N6His}$ (16.2 kDa) fig 5.3 panel F lane 2.

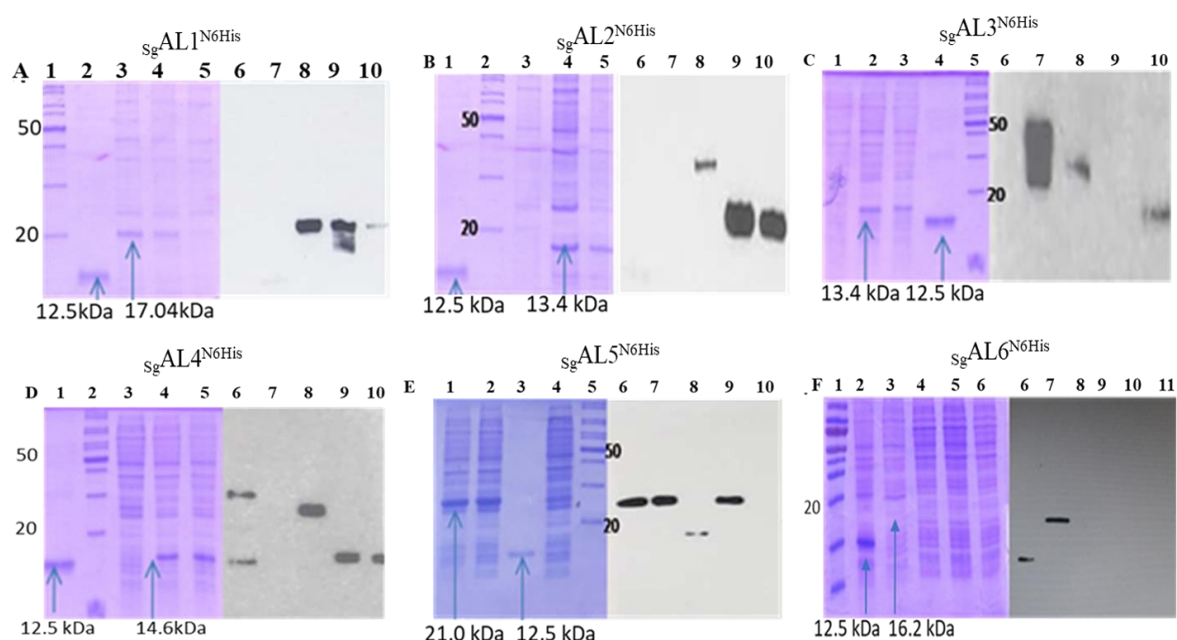


Fig. 5.5 Expression of LinA proteins

The SDS-PAGE and corresponding western blots developed by probing with anti-His antibodies indicate expression of $sgAL1^{N6His}$ (Panel A Lanes 3&4, 8&9 show expression and western blot profile respectively), $sgAL2^{N6His}$ (Panel B lanes 4&5, 9&10 show expression and western blots respectively), $sgAL3^{N6His}$ (Panel C lanes 2&3, 7&8 show expression and western blots respectively), $sgAL4^{N6His}$ (Panel D lanes 4&5, 9&10 show expression and western blots respectively), $sgAL5^{N6His}$ (Panel E lanes 2&3, 6&7 show expression and western blots respectively), and $sgAL6^{N6His}$ (Panel F lanes 3&7 show expression and western blots respectively).

5.3.4. Only *sgAL5*^{N6His} shows weak HCH dehydrochlorinase activity

After confirming the expression and stability, the cell lysates containing *sgAL1*^{N6His}, *sgAL2*^{N6His}, *sgAL3*^{N6His}, *sgAL4*^{N6His}, *sgAL5*^{N6His} and *sgAL6*^{N6His} were used to assay dechlorinase activity. The LinA activity was spectrophotometrically quantified by evaluating the amount of chloride ion released. Interestingly, of all LinA homologues, only *sgAL5*^{N6His} expressing *sgAL5*^{N6His} showed release of 0.03 mM chloride ion/mg of cells after 24 hrs of incubation (fig 5.6a). The rest of the LinA homologues failed to show any dechlorinase activity when γ -HCH was used as a substrate. The *sgAL5*^{N6His} was further purified using Nickel-NTA chromatography (fig 5.6b). However, the specific activity was below detection limit for purified protein using spectrophotometric method. To reinforce the result, the γ -HCH degradation activity was further examined using GC analysis with purified *sgAL5*^{N6His}. As shown in Fig 5.1, γ -HCH degradation by produces γ -PCCH, 1,4 TCDN

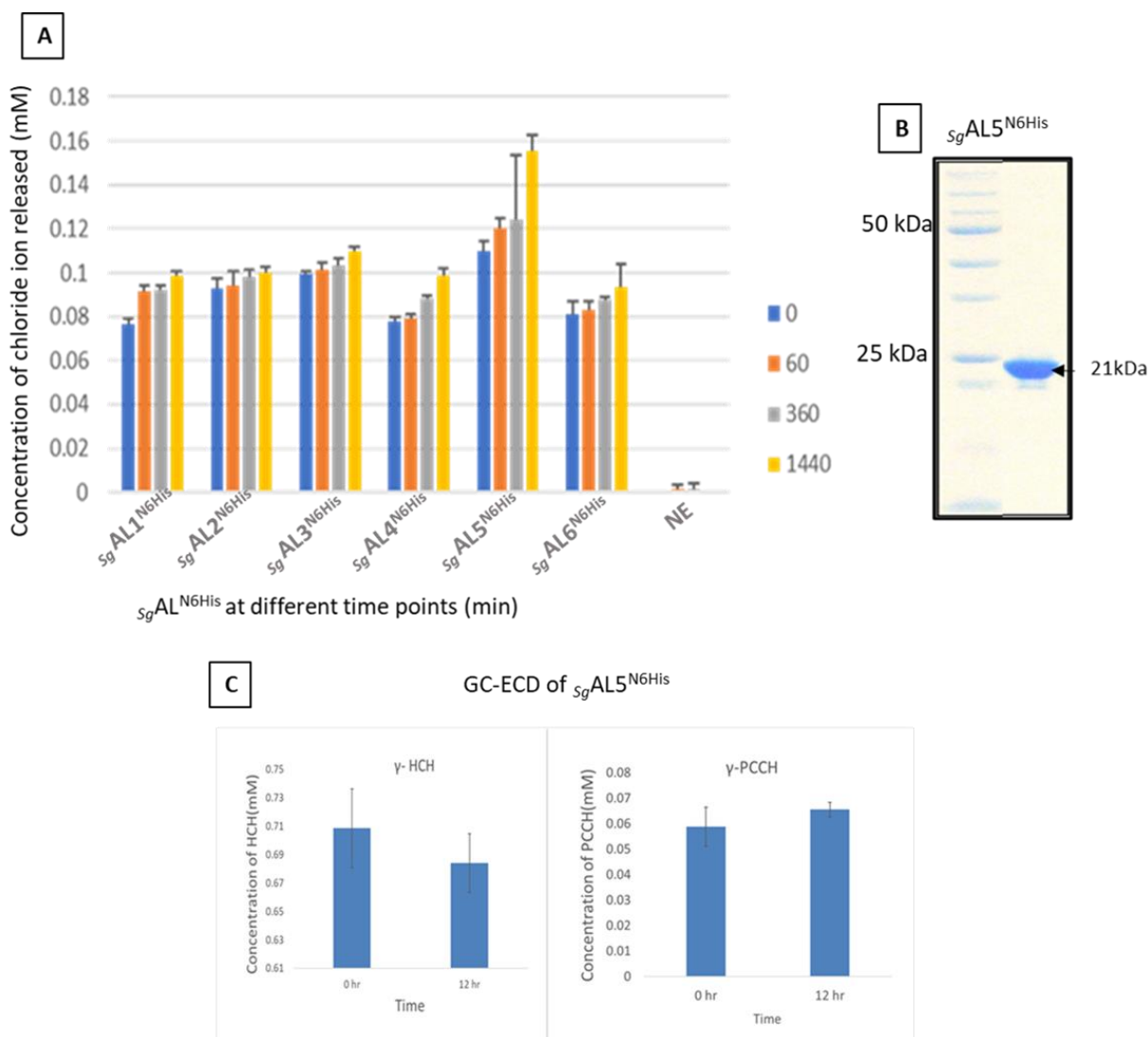


Fig. 5.6 Dehalogenase activity of $sgLinA5$

Panel A shows the chloride ion release when *E. coli* cells expressing $sgAL1$ proteins were incubated with 34.3 mM γ -HCH. Error bars represent standard error from triplicates. Panel B shows purified $sgAL5^{N6His}$ using NI-NTA affinity chromatography. Panel C shows the GC-analysis of products generated in reaction mix containing purified $sgAL5^{N6His}$ and 34.3 mM γ -HCH. The error bars represent standard error from triplicates.

and 1,2,4-TCB. Of these 1,4 TCDN is unstable and hence not detected in GC. But GC is a highly sensitive method for detecting dechlorinase activity as it can detect both decrease of γ -HCH and increase of the metabolites γ -PCCH and 1,2,4-TCB. As depicted in fig 5.6c, the reaction mix containing $sgAL5^{N6His}$ showed γ -HCH degradation. The reaction showed depletion of γ -HCH and production of γ -PCCH after 12 hrs of incubation, indicating existence of weak dechlorinase activity in $sgAL5^{N6His}$.

5.3.5. Haloalkane dehalogenases in *S. granuli* MG

As highlighted in the introduction part of this chapter, the genome mining approach has identified two haloalkane dehalogenases in the genome of *S. granuli* MG and named them as *sgHldA* and *sgHldB*. There are three well characterized haloalkane dehalogenases, DhIA isolated from *Xanthomonas autotrophicus* GJ10 and the second LinB isolated from *Sphingobium japonicum* UT26. The DhIA of *Xanthomonas autotrophicus* GJ10 is considered as a representative enzyme of HLDI subfamily, whereas LinB is considered as a representative of HLDII subfamily. The haloalkane dehalogenase DmbC of *Mycobacterium tuberculosis* belongs to HLDIII group. Among these representative dehalogenases of three different subfamilies, *sgHldA* showed maximum sequence identity (33%) with HLD-I subfamily, whereas *sgHldB* showed maximum sequence identity (29%) with HLD-III subfamily (fig 5.7b). Although catalytic pentad composition varies within the subfamilies, in all subfamilies of HLDs three residues of the pentad are highly conserved which include nucleophile, catalytic base and one halide stabilizing residue. The multiple sequence alignment showed that both *sgHldA* and *sgHldB* have the highly conserved catalytic triad (fig 5.7a). The *sgHldA* has Asp-119, His-273 and Trp-120. In *sgHldB* the Asp-105, His-276 and Asp-105 form the nucleophile, catalytic base and halide binding residue respectively. This result indicates that both *sgHldA* and *sgHldB* are putative haloalkane dehalogenases.

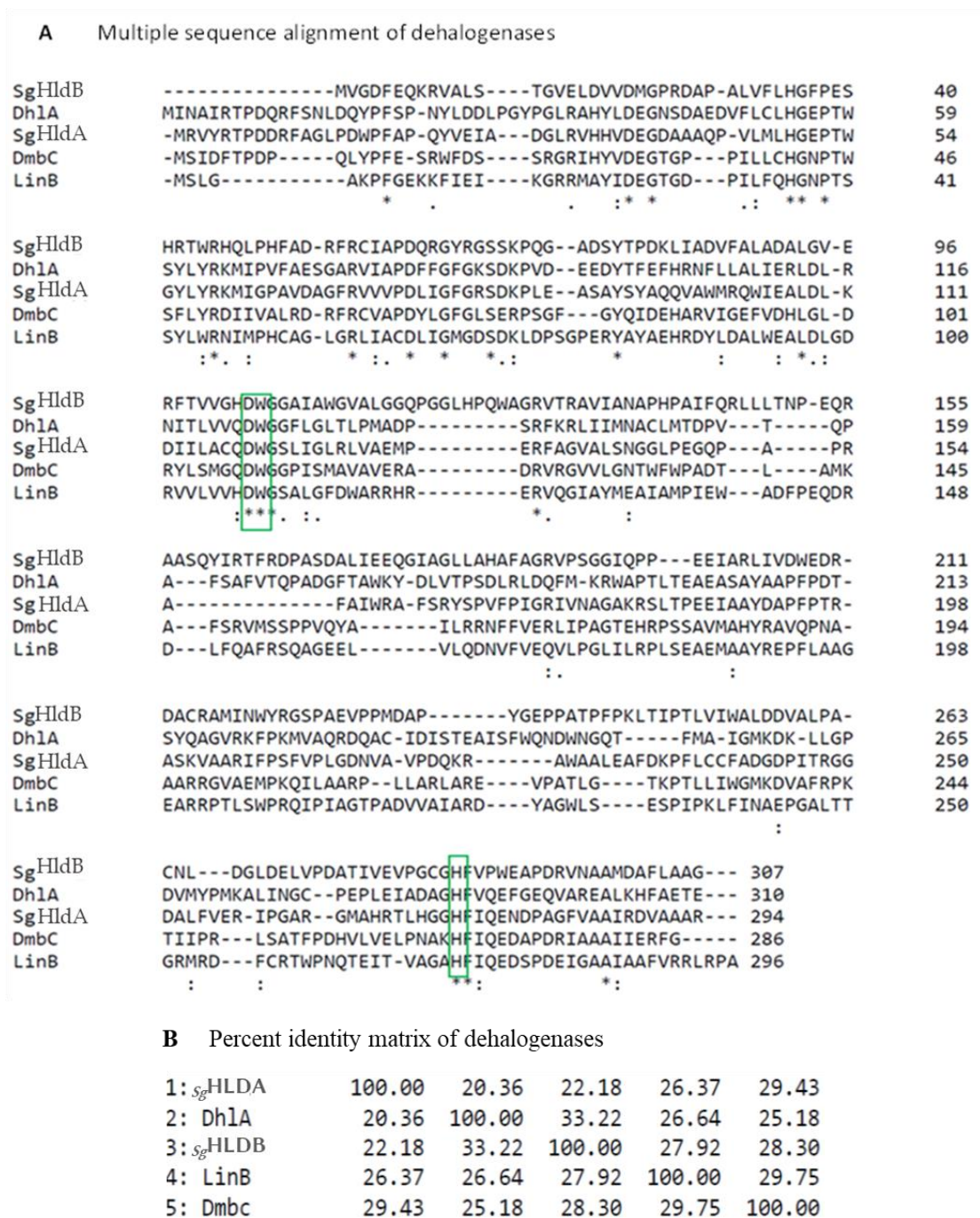


Fig. 5.7 Multiple sequence alignment and percent identity matrix of *SgHldA* and *SgHldB*

In panel A three conserved residues of *SgHldA* and *SgHldB* are indicated in green boxes. Multiple sequence alignment was generated using CLUSTAL W. Panel B shows the comparison of percent identity matrix of all HLDs.

5.3.6. Phylogenetic approach establishes relationship between *sg*HldA and *sg*HldB

There are about eighteen proteins designated as HLDs and kept in different subfamilies. All these sequences were used along with *sg*HldA and *sg*HldB to construct a phylogenetic tree (fig 5.8). As revealed by the phylogram, the *sg*HldA belongs to the HLD-I subfamily, whereas *sg*HldB shows evolutionary relationship with HLD-III subfamily. This result indicates that *sg*HldA belongs to HLD-I subfamily and *sg*HldB belongs to HLD-III subfamily of enzymes.

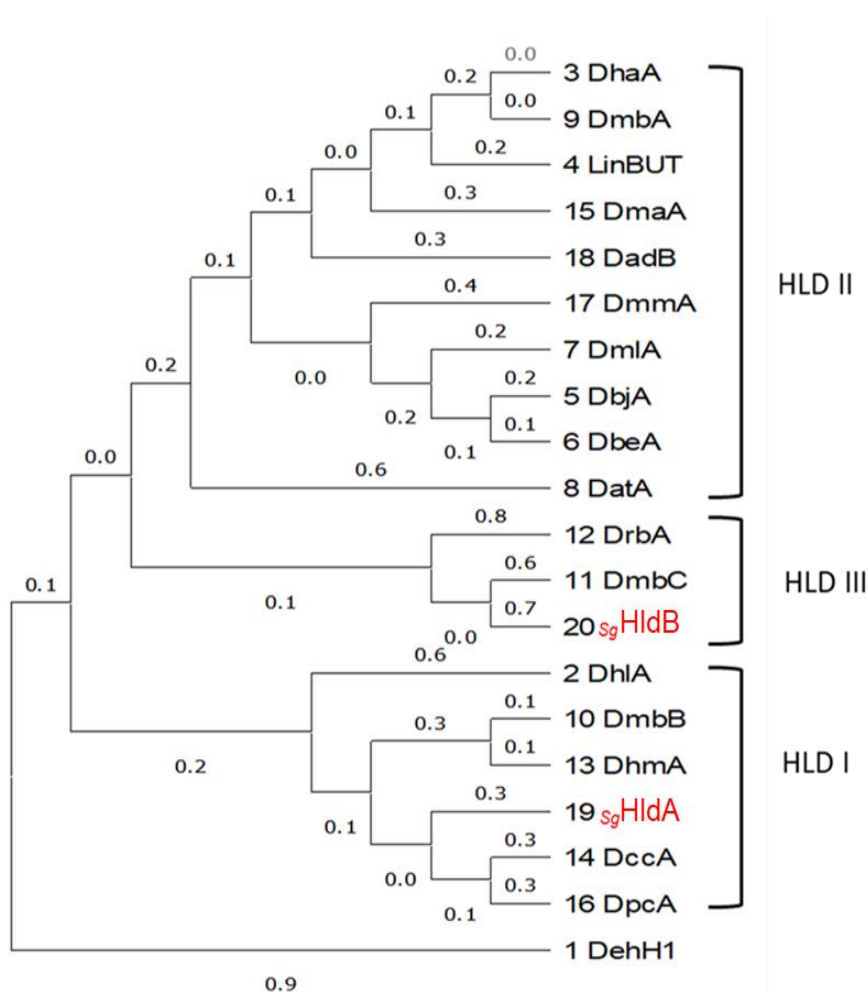


Fig. 5.8 Phylogram showing the relationship of *sg*HldA and *sg*HldB with other HLD proteins.

The multiple sequence alignment of all dehalogenases was performed using MUSCLE and phylogenetic tree was constructed using neighbour joining method with MEGA v 11.0. DehH1, a haloacetate dehalogenase was used as outgroup.

5.3.7. Expression and Activity of *sgHldA* and *sgHldB*

After classification of *sgHldA* and *sgHldB* the gene fragments *sghldA* and *sghldB* were independently cloned into expression plasmid pET22b (+) to express them in *E. coli* with an N-terminal histidine tag. The vector was linearized by digesting with restriction enzymes *NdeI* and *XhoI*. Further, gene fragments *sghldA* and *sghldB* were amplified independently using appropriate primers, ligated and transformed using Gibson Assembly method (NEB) (fig 5.9a). The recombinant plasmids were pWD5 and pWD6 encoding *sgHldA*^{N6His} and *sgHldB*^{N6His} respectively. The recombinants were expressed in *E. coli* NICO and the proteins were purified as described in the methods (fig 5.9b).

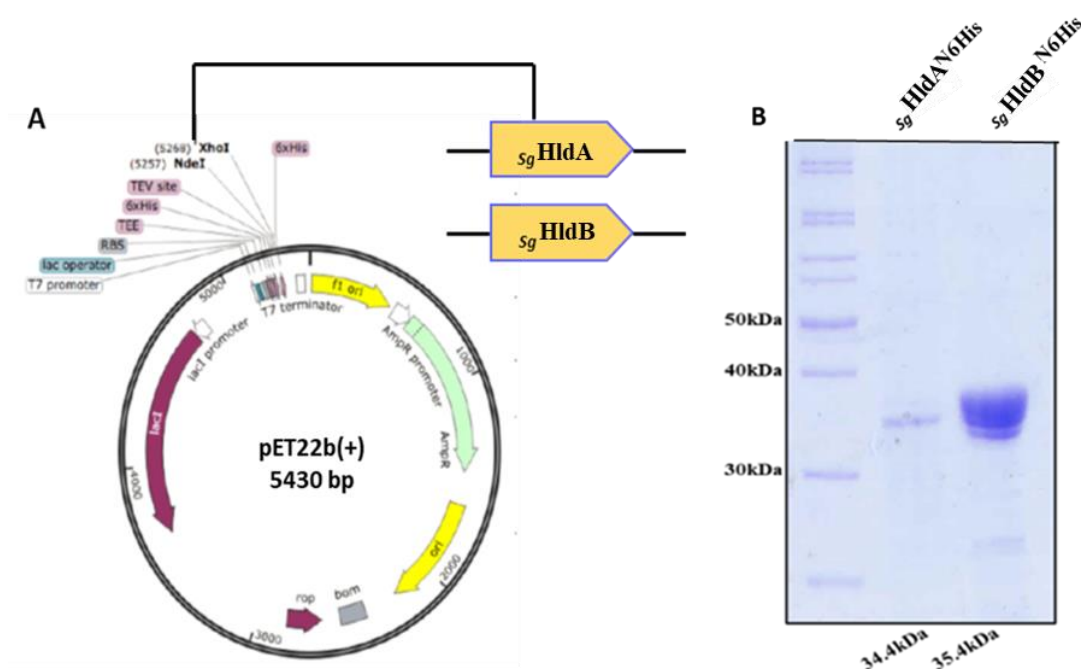


Fig. 5.9 Cloning strategy and purification of *sgHldA*^{N6His} and *sgHldB*^{N6His}

Panel A shows the cloning strategy employed for obtaining recombinants. Panel B shows SDS-PAGE of purified proteins. *sgHldA*^{N6His} is 34.4 kDa and *sgHldB*^{N6His} is 35.4 kDa.

After purification both *sgHldA* and *sgHldB* were tested against a panel of 13 halogenated substrates with varied chain lengths (fig 5.10 and 5.11). The enzyme *sgHldA*

appears to be a robust haloalkane dehalogenase and shows activity towards all the tested substrates fig 5.10 a,b&c.

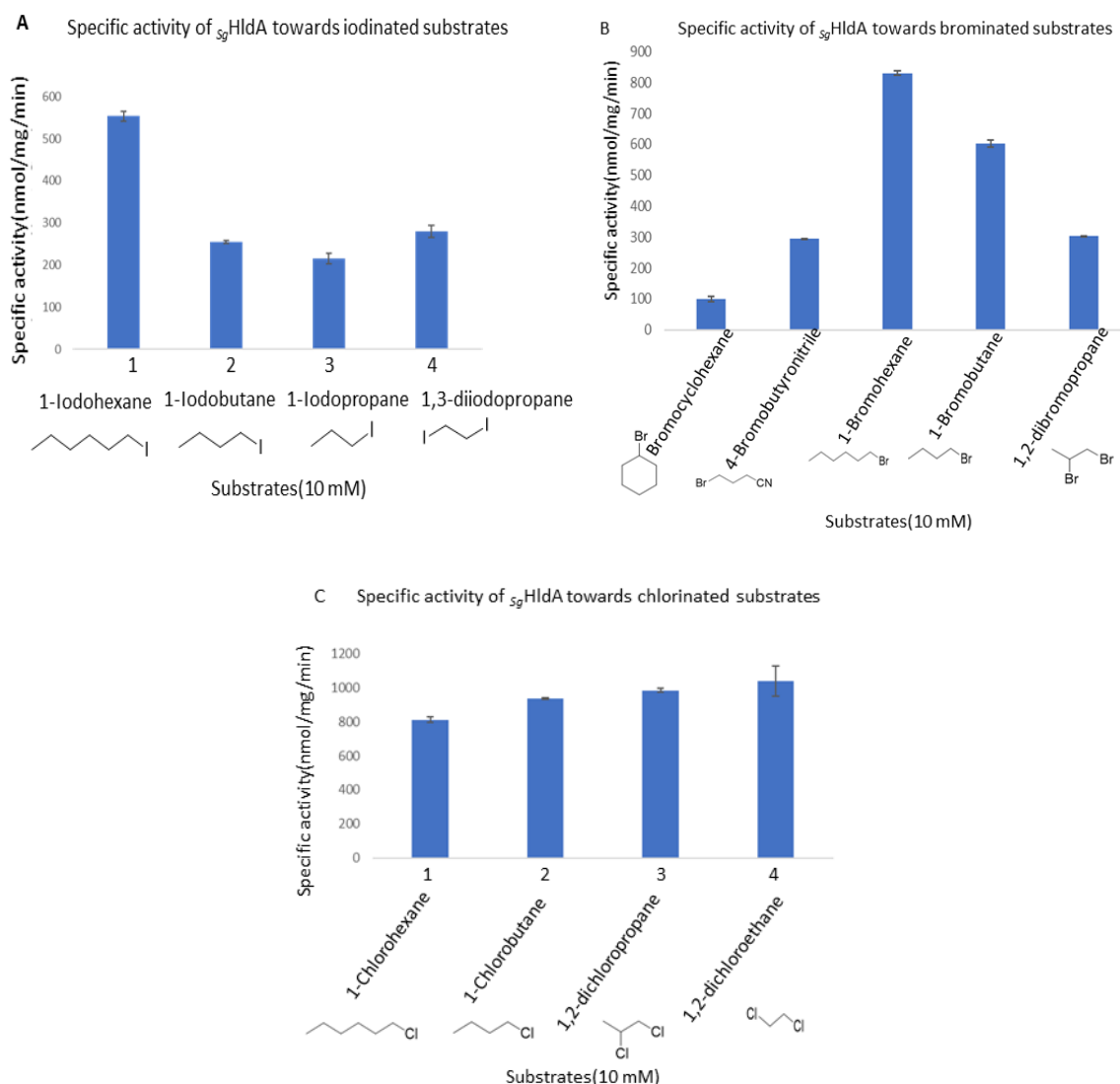


Fig. 5.10 Specific activity for *sg*HldA with halogenated substrates

Panel A Lanes 1-4 show specific activity for *sg*HldA with iodinated haloalkanes, Panel B Lanes 1-5 show specific activity for *sg*HldA with brominated haloalkanes and Panel C Lanes 1-4 show specific activity for *sg*HldA with chlorinated haloalkanes. Error bars represent standard error from triplicates.

Amongst the iodinated substrates, the activity of *sg*HldA is directly proportional to the chain length of iodinated substrates tested, with *sg*HldA showing highest activity towards 1-iodohexane at 545 nmol/mg/min (fig 5.10a, lanes 1-4). Moreover, *sg*HldA showed activity against both linear and branched iodinated substrates. In a comparison of similar length

substrates 1-iodopropane which is terminally halogenated and branched halogenated substrate 1,3-diiodopropane, *sgHldA* showed higher specific activity towards the branched substrate 1,3-diiodopropane (fig 5.10a, lane 3&4). Amongst the brominated substrates, highest activity of *sgHldA* was seen against 1-bromohexane at 815 nmol/min/mg (fig 5.10b, lane 3). Even amongst the brominated substrates specific activity was directly proportional to the chain length (fig 5.10b, lanes 1-5). Moreover, *sgHldA* showed least activity towards bromocyclohexane compared to all halogenated substrates indicating that cyclic haloalkanes were least preferred substrates for this enzyme (fig 5.10b, lane 1). Interestingly, among the chlorinated haloalkanes highest activity was seen for 1,2-dichloroethane at 840 nmol/mg/min with specific activity decreasing as chain length of halogenated substrates increased (fig 5.10c, lanes 1-4).

The enzyme *sgHldB* had a significantly lower specific activity than *sgHldA* towards all the substrates tested. Amongst the iodinated substrate, *sgHldB* showed highest specific activity towards 1-iodobutane (40 nmol/mg/min) (fig 5.11a, lane 2). This activity was fifty-fold lower than the activity seen in *sgHldA* for the same substrate even at 10- fold lower substrate concentration (1 mM) (fig 5.11 a, b&c). Moreover, the *sgHldB* did not show any activity for long chain iodinated haloalkanes like 1-iodohexane (fig 5.11a, lane 1). Also, *sgHldB* showed much lower activity towards branched halogenated iodinated substrates 1,3-diiodopropane (10 nmol/mg/min) compared to *sgHldA* (fig 5.11a, lane 3&4). Among the brominated substrates too, the *sgHldB* showed no activity towards 1-bromohexane (fig 5.11b, lane 3). Unlike *sgHldA*, *sgHldB* shows some specific activity towards 1-bromocyclohexane, albeit weakly (15 nmol/mg/min) (fig 5.11b, lane 1). The specific activity of *sgHldB* was also poor against all the tested brominated substrates. Surprisingly, compared to iodinated and brominated substrates (fig 5.11b, lanes 2-5), *sgHldB* showed better specific activity for chlorinated haloalkanes (fig 5.11c, lanes 1-4). Among the chlorinated haloalkanes, the specific activity of *sgHldB* was best for 1, 2 dichloroethane (42 nmol/mg/ml) and similar towards both 1,2- dichloropropane (35 nmol/mg/ml), and 1-chlorobutane (32 nmol/mg/ml).

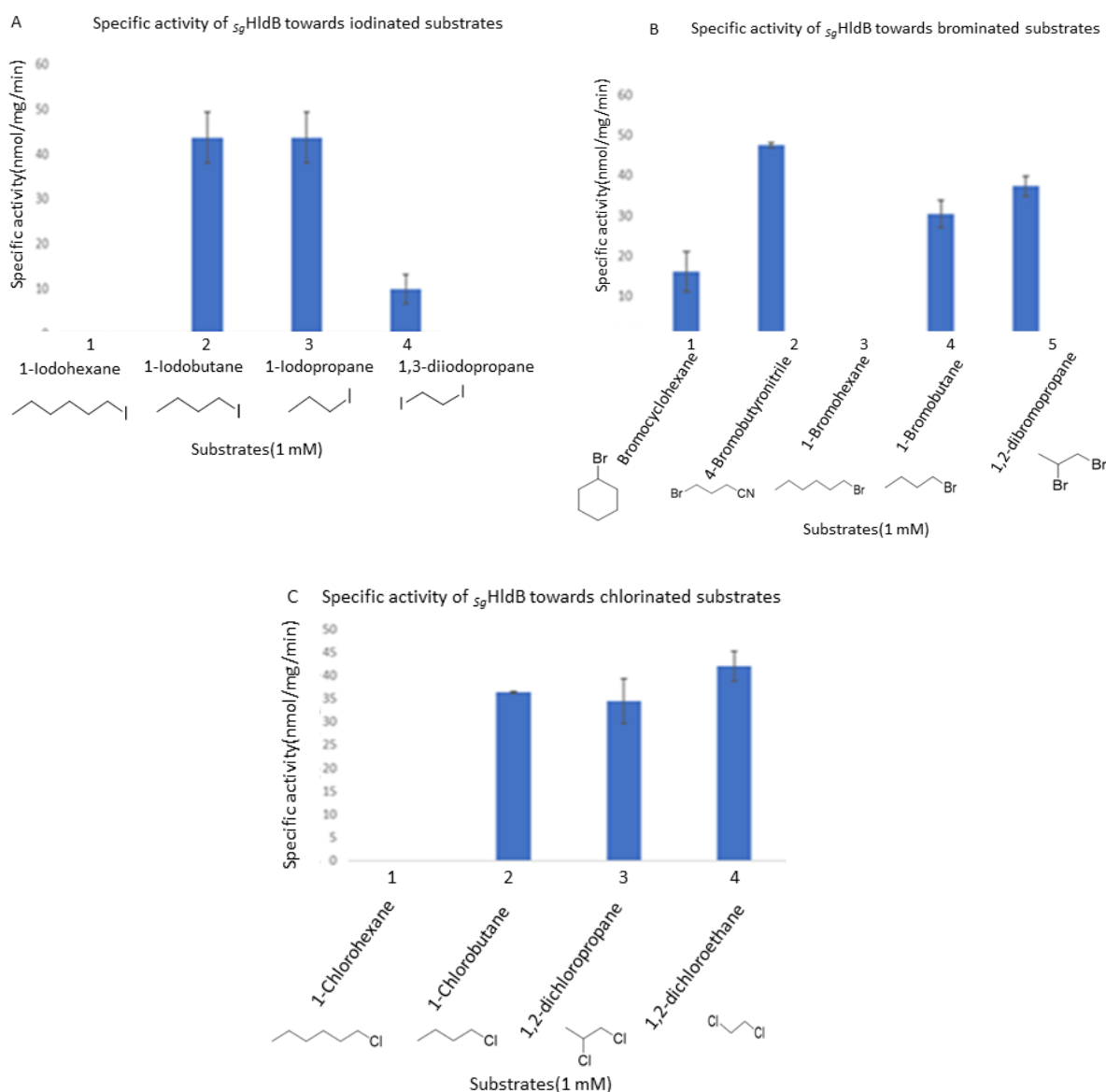


Fig. 5.11 Specific activity for *sg*HldB with halogenated substrates

Panel A lanes 1-4 show specific activity for *sg*HldB with iodinated haloalkanes, Panel B lanes 1-5 show specific activity for *sg*HldB with brominated haloalkanes and Panel C lanes 1-4 show specific activity for *sg*HldB with chlorinated haloalkanes. Error bars represent standard error from triplicates.

As *sg*HldB showed weak haloalkane dehalogenase activity against all tested halogenated aliphatic substrates, it is assumed that these are not cognate substrates to *sg*HldB. The promiscuous haloalkane dehalogenase activity of *sg*HldB might be due to structural similarities between its cognate substrate and halogenated substrates. In order to understand substrate specificities of both the dehalogenases, structural analysis was

performed using homology modelling. The enzyme *sg*HldA showed structural identity of 55% with another haloalkane dehalogenase DccA (PDB:5esrA). DccA is a haloalkane dehalogenase isolated from *Caulobacter crescentus* and belongs to HLD-I subfamily. Crystal structure indicates that DccA has a significant structural similarity to other well characterized haloalkane dehalogenase (Carlucci, Zhou, Malashkevich, Almo, & Mundorff, 2016).

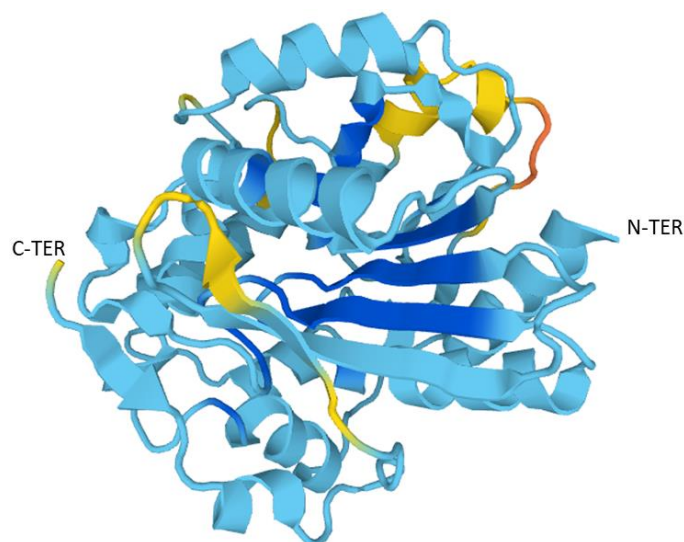


Fig. 5.12 Predicted tertiary structure of *sg*HldA

The N-terminal and C- terminal domains are marked. The model was generated by SWISS-MODEL with amino acid sequence of *sg*HldA as template. The dark blue areas indicate regions with very high structural identity (confidence > 0.9), light blue indicates regions with high structural identity (0.7-0.9), yellow regions indicate low structural identity (0.5-0.6), and red regions represent very low homology (< 0.5).

An overall comparison of structural homology of the predicted tertiary structure as seen in fig 5.12 shows that most of the *sg*HldA has very high (dark blue) to high (light blue) structural similarity to DccA. The regions of low (yellow) homology or very low homology (red) are also indicated. Therefore, homology modelling of *sg*HldA with DccA as template suggests that the structurally the *sg*HldA doesn't deviate significantly from other structurally characterized haloalkane dehalogenase. The high enzyme activity of *sg*HldA towards haloalkane dehalogenase also supports this hypothesis. This analysis shows that DccA serves as a good template to deduce the structure of *sg*HldA and that *sg*HldA is a functional haloalkane dehalogenase of *S. granuli* MG.

Using the same bioinformatic tool, a homology model was also generated for *sg*HldB. Unfortunately, *sg*HldB shows very low amino acid level identity to most haloalkane

dehalogenases. Of structurally characterized haloalkane dehalogenases, *sg*HldB showed maximum structural identity of only 26.52% with haloalkane dehalogenase DhaA (PDB:4HZG). DhaA is a haloalkane dehalogenase from *Rhodococcus rhodochrous* NCIMB 13064. DhaA was the first haloalkane dehalogenase with detectable activity against a highly toxic and carcinogenic pollutant 1,2,3-Trichloropropane (TCP). It belongs to HLD-II subgroup and wild type enzyme slowly ($k_{cat} = 0.08^{s^{-1}}$) converts toxic TCP into nontoxic 2,3-dichloropropan-1-ol (Lahoda et al., 2014).

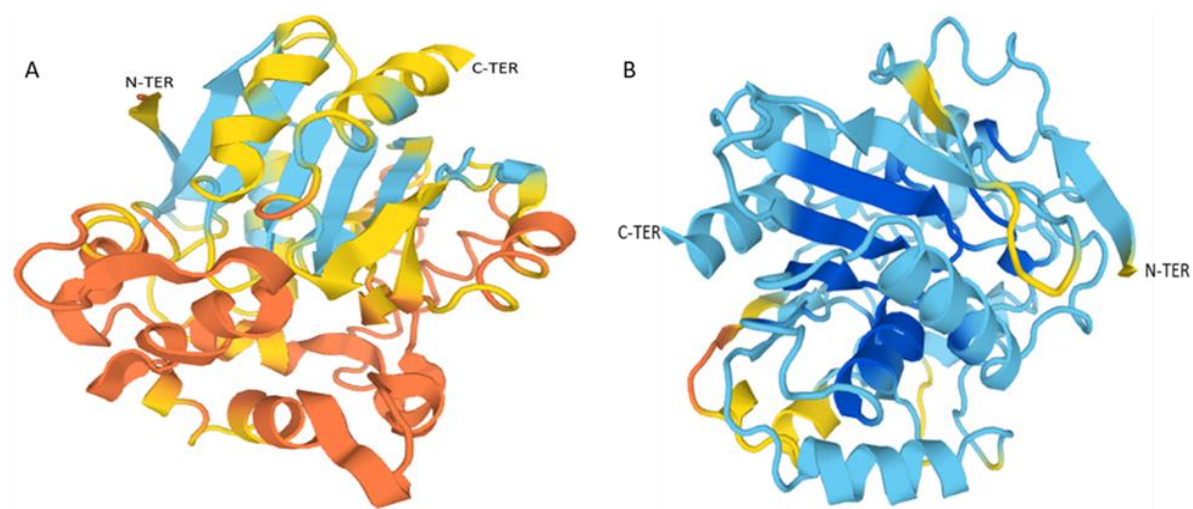


Fig. 5.13 Predicted tertiary structure of *sg*HldB

Panel A shows the homology model of *sg*HldB deduced from DhaA (PDB: 4HZG) and panel B shows homology model of *sg*HldB deduced from Sibe-EH (PDB: 5NFQ). The N- and C-terminal domains are marked. The model was generated by SWISS-MODEL with amino acid sequence of *sg*HldB as template. The dark blue areas indicate regions with very high structural identity (confidence > 0.9), light blue indicates regions with high structural identity (0.7-0.9), yellow regions indicate low structural identity (0.5-0.6), and red regions represent very low structural homology (< 0.5).

An overall comparison of the homology model generated using DhaA as template as seen in fig 5.13a shows that *sg*HldB has more regions with low structural identity (yellow) and very low homology (red) to the structure of haloalkane dehalogenase DhaA compared to those with high structural identity (light blue). This indicates that *sg*HldB may not be structurally similar to haloalkane dehalogenases. This also supports the assumption that haloalkanes may not be cognate substrates for this enzyme.

Surprisingly, the homology model of *sg*HldB showed structural identity of 67% to an epoxide hydrolase Sibe-EH (PDB: 5NFQA). The Sibe-EH is a thermophilic epoxide hydrolase isolated from metagenome of hot spring with broad substrate specificity and enantioselectivity (Ferrandi et al., 2018). An overall comparison of homology model of *sg*HldB generated using Sibe-EH as template seen in fig 5.13b shows that *sg*HldB has more regions of very high (dark blue) and high structural identity (light blue) to epoxide hydrolase Sibe-EH. This result shows that *sg*HldB may be an epoxide hydrolase with promiscuous haloalkane dehalogenase activity due to structural similarities between their cognate substrates.

5.3.8. Active site architecture of *sg*HldA and *sg*HldB for dehalogenation reaction

In order to gain further evidence for the reasons for ambiguity in structure and function of *sg*HldB, the active site residues of both *sg*HldA and *sg*HldB were compared using computational methods to a well characterized haloalkane dehalogenase DhIA isolated from *Xanthomonas autotrophicus* GJ10. The reaction mechanism of DhIA is well established both by elucidating its structure and mutating the active site residues (Franken, Rozeboom, Kalk, & Dijkstra, 1991; Krooshof, Kwant, Damborsky, Koca, & Janssen, 1997). Using DhIA as a template, the active site architecture was determined both for *sg*HldA and *sg*HldB. The catalytic reaction in DhIA begins by the nucleophilic attack of the carboxylate oxygen of Asp (124) on the carbon atom of halogenated substrate (fig 5.14a). This residue was identified both in *sg*HldA and *sg*HldB as Asp (119) and Asp (104) respectively (fig 5.14b&c). Following the nucleophilic attack in DhIA, an alkyl enzyme tetrahedral intermediate is formed. This is stabilized by amide nitrogens of Glu (56) and Trp (175) occurring in oxyanion pocket. In the active site of *sg*HldA both these amino acids are present as Glu (54) and Trp (120). Moreover, at equivalent positions, in *sg*HldB also shows existence of Glu (36) and Trp (105) residues. In DhIA, the negative charge that develops on halogen atom is stabilized by side chains of Trp (125) and Trp (175). Of these Trp (125) is fully conserved, however the other halide stabilizing residues is structurally non-conserved regions and occurs at different position in HLD-II and HLD-III subfamilies. Using computational methods, the homology model of *sg*HldA, Trp (159) was identified as probable halide stabilization residues for *sg*HldA. However, for *sg*HldB no residue could be assigned at equivalent positions. At a non-equivalent position, however, a Trp (111) was identified but its orientation was different compared to other Trp residues. In

contrast to DhIA that belongs to HLD-I subgroup, *sg*HldB belongs to HLD-III subgroup and therefore may have different secondary halide stabilizing residue. In DhIA, the alkyl-enzyme intermediate formed is subsequently hydrolyzed by His (289) and the catalytic acid Asp (260). Both *sg*HldA and *sg*HldB possess the catalytic acid Asp (244, 258) and His (273, 286) respectively. This result indicates that, despite poor overall homology of *sg*HldB with HLDs, the active site residues and the active site architecture appears to be conserved. This prediction also supports the dehalogenation function for *sg*HldB. Moreover, this analysis reinforces that *sg*HldA is a typical haloalkane dehalogenase.

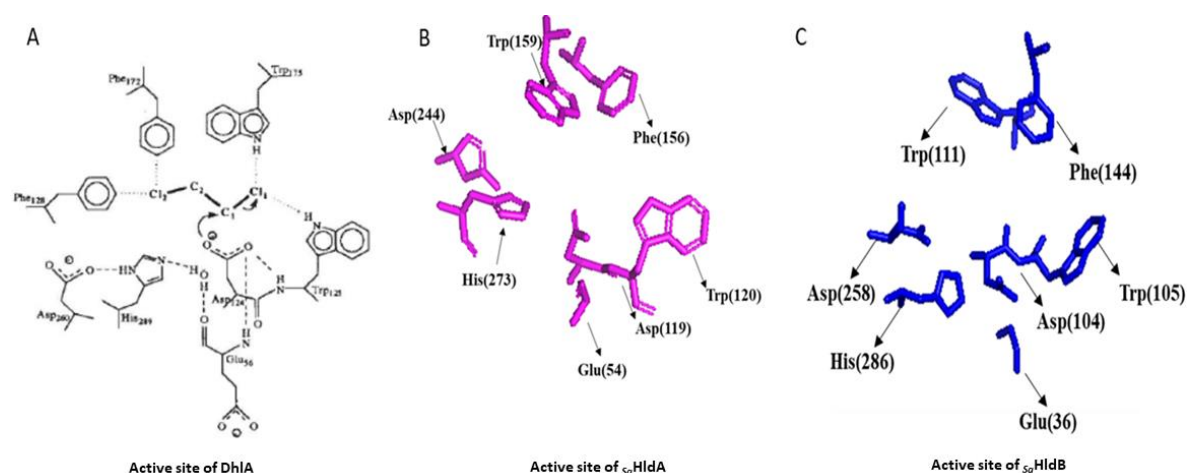


Fig. 5.14 Arrangement of catalytic site residues in *sg*HldA and *sg*HldB during dehalogenase activity

Panel A shows amino acid residues in DhIA. Panel B shows the amino acid residues in *sg*HldA and Panel C shows amino acid residues in *sg*HldB. Arrows indicate the residues and the numbers indicate their position.

5.3.9. Active site architecture of *sg*HldA and *sg*HldB for epoxide hydrolase reaction

Since homology modelling showed *sg*HldB to be an epoxide hydrolase, active site architecture of *sg*HldA and *sg*HldB were also compared with a well-characterized epoxide hydrolase, EchA from *Agrobacterium radiobacter* AD1 (Rui, Cao, Chen, Reardon, & Wood, 2005). As a part of α/β -hydrolase family, epoxide hydrolase too follows S_N2 reaction mechanism and contains catalytic triad (Asp- His-Asp/Glu) like in haloalkane dehalogenases. In EchA, carboxylate oxygen of nucleophile Asp (107) launches an attack on the least hindered carbon atom of epoxides (fig 5.15a). This residue is identical to the nucleophilic attack residues seen in dehalogenation reaction as the Asp (119) and Asp (104) of *sg*HldA and *sg*HldB

respectively (fig 5.15b&c). The ester-intermediate formed in this reaction is hydrolysed by Asp (246)- His (275) pair in EchA. These residues are also seen as Asp (244)-His (273) pair in *sgHldA* and Asp (258) and His (286) pair in *sgHldB* respectively. In the crystal structure of EchA, it is clear that the hydroxyl groups of two tyrosine residues, namely Tyr (152) and Tyr (215) are proton donors during the ring opening of epoxide hydrolase. Moreover, of these two, the Tyr (215) is completely conserved at the C-terminal of all epoxide hydrolases (Nardini et al., 1999). Absence of this residue in *sgHldA* indicated that its active site architecture cannot support an epoxide hydrolase activity. However, both the tyrosine residues were predicted from homology model of *sgHldB* as Tyr (160) and Tyr (221) indicating that *sgHldB* may also show epoxide hydrolase activity.

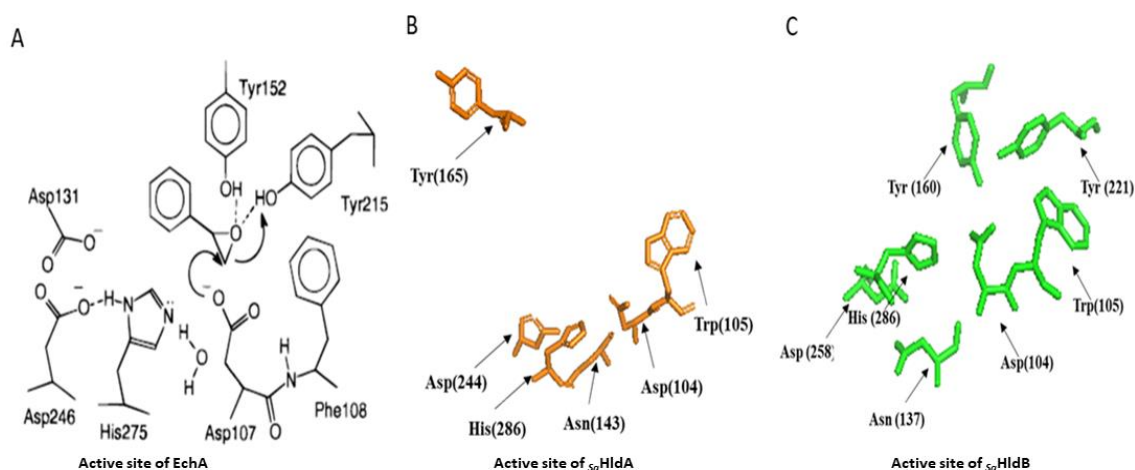


Fig. 5.15 Arrangement of catalytic site residues in *sgHldA* and *sgHldB* during epoxide hydrolase activity

Panel A shows amino acid residues in EchA. Panel B shows the amino acid residues in *sgHldA* and Panel C shows amino acid residues in *sgHldB*. Arrows indicate the residues and the numbers indicate their position.

5.3.10. Epoxide hydrolase activity of *sgHldB*

After establishing the active site similarities between epoxide hydrolases and *sgHldB*, further experiments were conducted to determine epoxide hydrolase activity of *sgHldB*. The activity assay was performed using methods described earlier. The enzyme *sgHldB* showed significant hydrolysis of aromatic epoxide styrene oxide showing specific activity of 0.6 $\mu\text{mol}/\text{min}/\text{mg}$ and 1-epoxybutane showing a specific activity of 0.1 $\mu\text{mol}/\text{min}/\text{mg}$. The highest specific activity was observed for 1-epoxyhexane at 1.6 $\mu\text{mol}/\text{min}/\text{mg}$ (fig 5.16). These results indicate that *sgHldB* is a functional epoxide hydrolase.

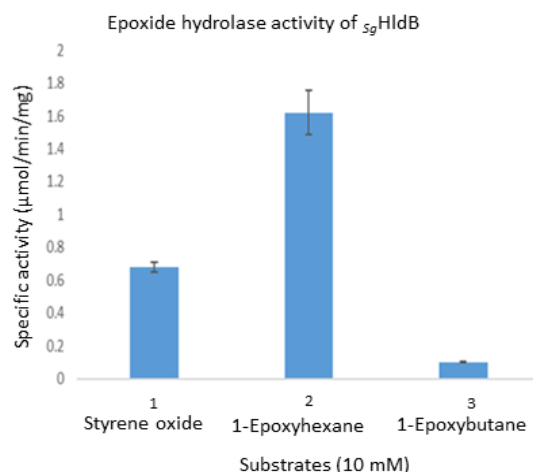


Fig. 5.16 Specific activity of *sgHldB* towards epoxides.

The lanes 1-3 show specific activity of *sgHldB* towards styrene oxide, 1-epoxyhexane and 1-epoxybutane respectively. The error bars show standard deviation in triplicates.

Epoxide hydrolases (EH) and haloalkane dehalogenases (HLD) belong to the same superfamily of enzymes and are phylogenetically related. Moreover, both enzymes also share mechanistic similarities by using nucleophile aspartate for an S_N2 attack on substrate. Therefore, it may be possible that *sgHldB* is a unique enzyme present in genome of *S. granuli* MG which is at the evolutionary crossroads between HLD and EH activity.

Both EH and HLD have the same Asp-His-Asp/Glu catalytic triad. In both the enzymes nucleophile aspartate performs an S_N2 attack on either the epoxide (in case of EH) or halogenated carbon (in case of HLD) to generate an alkyl enzyme intermediate which is then hydrolyzed by the His-Asp/Glu pair. In HLD-I, III and in epoxides the Asp/Glu acid and histidine base are in close proximity. In HLD-II subfamily, however, catalytic nucleophile and acid residue are located near to one another. Therefore, the catalytic site residues are quite conserved in both epoxide hydrolase and haloalkane dehalogenase. Additionally, both EH and HLD have conserved supporting residues. In epoxide hydrolases a Tyr-Tyr pair that is conserved in most epoxide hydrolases is required for ring opening reaction. In haloalkane dehalogenase subgroup I, as seen for DhIA and *sgHldA*, Trp-Trp pair form the halide stabilization residues. In HLD-II and III, however, the primary halide stabilization Trp is coupled to an asparagine present in the HGXP motif of the protein sequence where X is any residue.

Interestingly, in *sgHldB*, a phenylalanine residue (F37) instead of an asparagine is present in the conserved HGXP pattern as seen in other haloalkane dehalogenases. Such Trp-Phe pair at these two positions is a unique feature of *sgHldB* and is not seen in any other known dehalogenases identified till date. It is assumed that the unique Phe (F37) residue may be responsible for the promiscuous haloalkane dehalogenase activity of *sgHldB*. To investigate the role of Phe (F37) residue in HLD and EH activity of *sgHldB*, mutational studies using site directed mutagenesis was performed.

5.3.11. Mutational studies with variant *sgHldB*^{F37N}

The variant *sgHldB*^{F37N} was generated using site directed mutagenesis as described earlier. The chromatogram of the sequenced plasmid containing the mutated residues is shown in fig 5.17a. The substitution of phenylalanine residue with an asparagine residue in the HGXP motif was confirmed by sequencing plasmid pWD6^{F37N}. The Ni-NTA affinity purification of variant *sgHldB*^{F37N} yielded pure protein as seen in the SDS-PAGE (fig 5.17b) which was further used in activity assays. The dehalogenase activity assay was performed as described earlier with 1-Bromobutane (1 mM) as the substrate. As shown in fig 5.17c the variant *sgHldB*^{F37N} did not show any change in color in the dehalogenase activity assay using 1-Bromobutane as the substrate (fig 5.17c lanes 7-9) compared to the controls (fig 5.17c lanes 1-6) indicating that variant *sgHldB*^{F37N} has lost the dehalogenase activity. This result suggests that the phenylalanine residue may be important for promiscuous dehalogenase activity of *sgHldB*. The effect of variant *sgHldB*^{F37N} on the epoxide hydrolase activity was also determined with styrene oxide (10 mM) and 1-epoxyhexane (10 mM) as substrates. Interestingly, the variant has shown reduced activity with 1-epoxyhexane, retaining only 18% of the activity compared to the wild type enzyme (fig 5.17d lane 3). Moreover, the enzyme has completely lost activity against the aromatic epoxide styrene oxide (fig 5.17d lane 4). The variant targeting the secondary halide stabilizing residue phenylalanine, substitutes the large non-polar residue by another small polar residue asparagine resulting in change of hydrophobicity that may have resulted in loss of epoxide hydrolase activity (Visser, de Bont, & Verdoes, 1999).

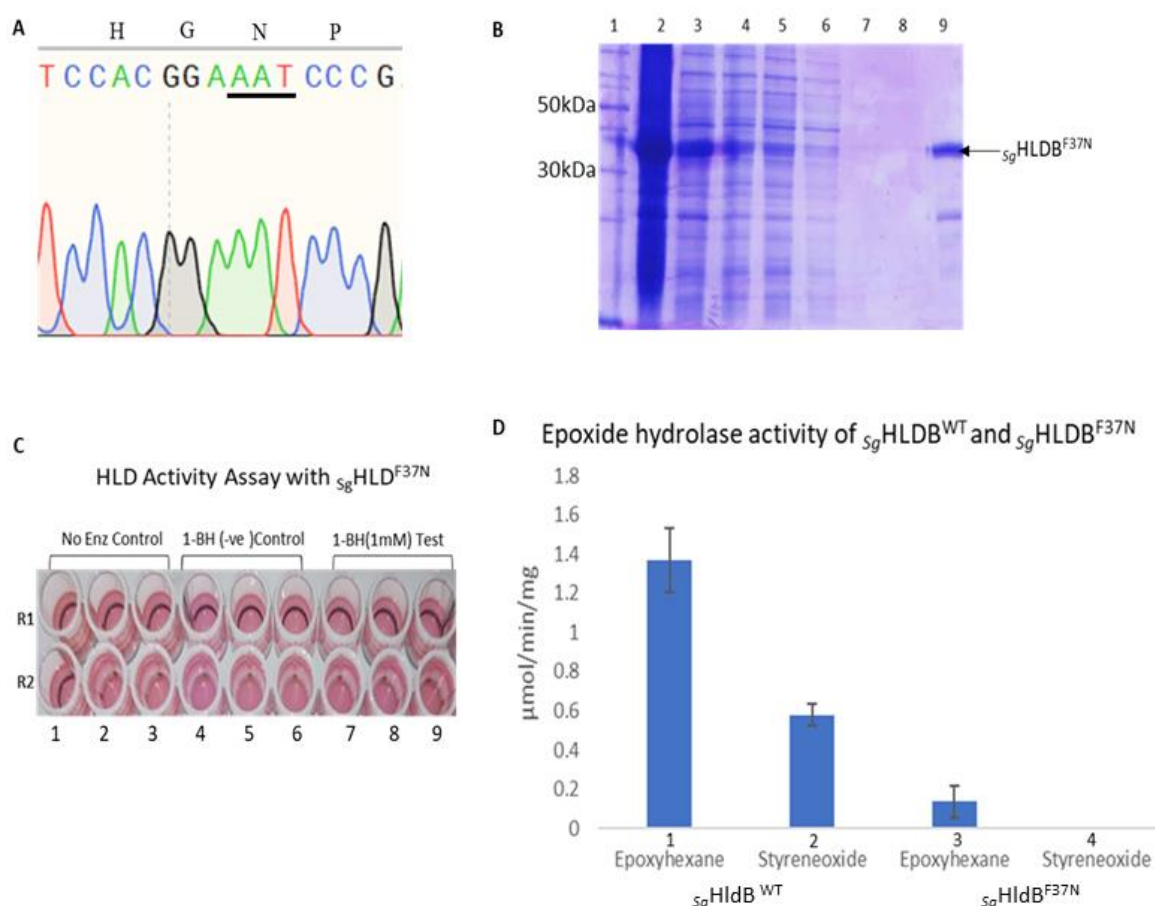


Fig. 5.17 Sequencing, purification, HLD&EH assays of variant *sgHldB*^{F37N}

Panel A shows the chromatogram generated after the sequencing of pWd6^{F37N} plasmid. Insertion of codon AAT which specifies asparagine, in place of TTT that specifies phenylalanine is underlined. Panel B shows the purification of *sgHldB*^{F37N} using Ni-NTA chromatography. The lane 1 shows Ladder. Lanes 2&3 show the pellet and supernatant fractions. The lanes 4-5 show the flowthrough and 6-8 show the wash fractions respectively. The elution of purified *sgHldB*^{F37N} is shown in lane 9. Panel C shows the dehalogenase activity assay of *sgHldB*^{F37N} using 1-bromobutane as substrate (1-BB, 1 mM). Lanes 1-3 represent enzyme negative controls, 4-6 are the substrate negative controls and lanes 7-9 are the test samples. R1 and R2 represent two biological replicates. Panel D shows the epoxide hydrolase activity of *sgHldB*^{WT} and *sgHldB*^{F37N} using 1-epoxyhexane and styrene oxide as substrates (10 mM each). Lanes 1-2 show the epoxide activities of *sgHldB*^{WT} with 1-epoxyhexane and styrene oxide respectively. Lanes 3-4 show the epoxide activities of *sgHldB*^{F37N} with 1-epoxyhexane and styrene oxide respectively. Error bars represent standard error from triplicates.

5.4. Discussion

In this chapter, based on the genome guided catabolic pathways, three unique dehalogenases were identified in the genome of *S. granuli* MG. The first dehalogenase identified is a LinA-HCH dehydrochlorinase or *sgLinA5*^{N6His}. From six homologs predicted using domain-based homology, *sgLinA5*^{N6His} has shown to be a functional enzyme with a weak activity

towards γ -HCH. The other five homologs of LinA failed to show dehydrochlorinase activity when γ -HCH was used as a substrate. With respect to lindane (γ -HCH) degradation, emerging evidence from Sphingomonad genomes suggests that a number of strains involved in degradation have considerable variation in the recruitment of genes. In some of them the known *lin* genes are either missing or show existence of multiple copies of the complete set of *lin* genes. However, as seen in *S. granuli* MG, several strains invariably show the presence of *linA* gene, despite missing all of the upstream genes (*linB-F*) involved in lindane degradation. Since most *lin* genes are with mobile elements, it is suggested that these strains that have acquired some but not all *lin* genes during lateral gene transfer. Indeed, these strains serve as a valuable resource for understanding the evolution pathway of HCH degradation (Pearce, Oakeshott, & Pandey, 2015). Even *linAs* appear to have evolved due to extensive gene shuffling. Supporting this notion, nearly 74% of functional *linAs* have been successfully reconstituted under *in-vitro* condition by reshuffling metagenomic DNA isolated from non-HCH contaminated soils (Boubakri, Beuf, Simonet, & Vogel, 2006). Since the origin of *linA* is unknown, it is suspected that *linA* may have originated in the environment as a consequence of bacterial adaptability to survive in soils contaminated with HCH, perhaps by DNA shuffling process.

Two more dehalogenases were also identified by genome mining in *S. granuli* MG. One of it named as *sgHldA* is an active haloalkane dehalogenase (HLD) enzyme. It exhibits substantial dehalogenase activity like many other HLD-I subgroup enzymes and takes up broad variety of haloalkanes as substrates. DhIA, the typical HLD-I subgroup enzyme, consists of unique feature. It contains a 10 amino acid insertion loop between the entrance and slot tunnel and this inserted sequence is believed to be the key reason for its activity against chlorinated substrate (Pikkemaat & Janssen, 2002). Interestingly, *sgHldA* also shows activity against chlorinated haloalkanes despite lacking any such insertion sequence. Elucidation of the crystal structure of *sgHldA* in complex with chlorinated substrates might reveal the true mystery behind this novel catalytic reaction mechanism. Furthermore, the broad substrate specificity of this enzyme makes it an interesting target towards a number of biotechnological applications.

The unique enzyme of *S. granuli* MG identified in this study is *sg*HldB. The *sg*HldB is an EH with promiscuous HLD activity. Both EH and HLD are members of the same superfamily- the α/β - hydrolase fold superfamily. This superfamily of enzymes consists of large group of enzymes that share common catalytic framework with diverse catalytic functions. Thus, this superfamily enzymes are considered as an excellent model system to study divergent evolution (Holmquist, 2000). Apart from dehalogenases and epoxide hydrolases, the superfamily includes enzymes with high sequence variability and broad catalytic activities. They include acyltransferases, amidases, lipases, esterases, thioesterases, hydroxynitrile lyases, diene lactone hydrolases, peroxidases and proteases (Mindrebo, Nartey, Seto, Burkart, & Noel, 2016; Nardini & Dijkstra, 1999). Moreover, many of these enzymes depict catalytic promiscuity, sometimes with changes in the position of one or two amino acid position. Intriguingly, an esterase was switched to hydroxynitrile lyase (Padhi et al., 2010), hydrolase to acyltransferase (Larsen et al., 2010), esterase to amidase (Kourist, Bartsch, Fransson, Hult, & Bornscheuer, 2008), lipase to an epoxide hydrolase (Svedendahl et al., 2008), or esterase to epoxide hydrolase (Jochens et al., 2009) with changes in the position of one or two amino acids.

This study has shown the existence of EH activity and HLD activity in *sg*HldB. Despite the mechanistic similarities very recently only one other enzyme, epoxide hydrolase from *Corynebacterium* sp C12 has shown a EH with promiscuous HLD activity (Schuiten et al., 2021). Inspired by the striking overall similarities and dual catalytic activity of *sg*HldB, further analysis was performed to understand the catalytic promiscuity of this enzyme. The existing knowledge suggests that the catalytic triad Asp-His-Asp/Glu was indispensable for both the activities. Therefore, mutations in these residues can potentially lead to the loss of both activities. A major difference in the reaction mechanism of EH and HLD is that most epoxide hydrolase enzymes employ canonical Tyr-Tyr pair for positioning of the epoxide for nucleophilic attack and protonation of the epoxy oxygen. Intriguingly, in HLDs, rather than tyrosine residues, Trp-Trp or Trp-Asn residues are involved in stabilizing the halide substrate for catalysis. In *sg*HldB, however, non-canonical Tyr-Phe pair functions as the halide stabilizing residue. Substitution of the phenylalanine residue to asparagine resulted in the total loss of promiscuous dehalogenase activity of the enzyme. The *sg*HldB variant retained only 18% of

the epoxide hydrolase activity (fig 5.17d) indicating that in *sg*HldB^{F37N} there was a significant reduction in epoxide hydrolase activity. This may be due to loss of the overall conformation of the molecules or it could be that the phenylalanine residue may be important for catalysis in both HLD and EH activities. Further investigations by computational modelling of Michael-Menton complexes using crystal structure of *sg*HldB may help resolve lacunas in this study.

The *sg*HldB can be considered an evolutionary intermediate between haloalkane dehalogenase and epoxide hydrolase. As a promiscuous enzyme *sg*HldB can be used a model enzyme to study enzyme evolution and diversification. The fact that both EH and HLD activity exists in *sg*HldB, this enzyme can be possibly an ancestral enzyme with both EH and HLD activities. The *sg*HldB appears to be at the crossroads of evolution and it can serve as a progenitor for evolution of novel enzymes accepting novel substrates for performing either dehalogenase or epoxy hydrolase activities.

CHAPTER 6 : Enantioselectivity of epoxide hydrolase from *S. granuli*

MG

6.1. Introduction

The epoxide hydrolases are enzymes that catalyse the conversion of epoxides into their corresponding vicinal diols. The simple mechanism employed by this enzyme catalysed biotransformation demonstrates high superiority over conventional chemical synthesis routes. The high stereo- or regio- selectivity which stem out of the epoxide hydrolase activity yields product with very high purity (Orru & Faber, 1999). Apart from the significant economic benefit over the organic synthesis routes, approaches using biocatalysts provide numerous advantages like catalytic activities at ambient temperature, pH and atmospheric pressure. The major advantage is however, the sustainability of the approach as use of biocatalysts introduces no damage to the environment (Schmitz, Rosenthal, & Lutz, 2019; Schwarz, Rosenthal, Snajdrova, Kittelmann, & Lutz, 2020).

Epoxides are reactive due to the presence of polar C-O bond and strain on the three membered epoxide ring, thus becoming easy target for a nucleophile attack. Using conventional methods selective epoxides or vicinal diols are derivatized from the natural chiral pool or by asymmetric synthesis of the prochiral substrate. Bio catalysis of these substrates using epoxide hydrolases yield commercially valuable products with excellent enantiomeric excess (ee), especially in cases where substrates are poorly or not amenable to organic synthesis. Vicinal diols and enantiopure epoxides have high commercial value as chiral precursors in the synthesis of fine chemicals. Additionally, they also form important precursors in active pharmaceutical ingredients of drugs like β -blockers (Genzel, Archelas, Broxterman, Schulze, & Furstoss, 2002; Monterde et al., 2004), antibiotics (Ueberbacher, Osprian, Mayer, & Faber, 2005), HIV-protease inhibitors (Pedragosa-Moreau, Archelas, & Furstoss, 1996), antifungals (Monfort, Archelas, & Furstoss, 2004), 11-heterosteroids (Bottalla et al., 2007) and non-steroidal anti-inflammatory drugs (Cleij, Archelas, & Furstoss, 1999). It has also been used to synthesize certain agrochemicals and natural products (W. J. Choi, Puah, Tan, & Ng, 2008; Kroutil, Osprian, Mischitz, & Faber, 1997).

As seen in chapter 3, the epoxide hydrolase *sgHldB* belongs to α/β -hydrolase fold superfamily with the catalytic triad (Asp-His-Gln) wherein the aspartate nucleophile attack on the carbon of the oxirane ring displaces the oxygen forming an alkyl-ester intermediate. The hydrolysis of the covalent intermediate formed is supported by the water molecule in presence of histidine residue. The enzyme is thus regenerated and the product, a vicinal diol is formed (Won Jae Choi, Choi, & Engineering, 2005; de Vries & Janssen, 2003). The homology modelling depicted that *sgHldB* shows greater identity to other epoxide hydrolases. Moreover, owing to mechanistic similarities *sgHldB* also shows promiscuous haloalkane dehalogenase activity. In this chapter, the epoxide hydrolase activity of *sgHldB* was confirmed by hydrolysis of substrate and formation of products. Further the enantioselectivity of the *sgHldB* was determined using racemic (S-, R-) styrene oxide as substrate.

6.2. Objective specific methodology

6.2.1. Hydrolysis of epoxides using GC-MS analysis

All the reactions were performed in 2.5 mL mininert vials to prevent evaporation of volatile products. Briefly, 480 μ L of 50 mM phosphate buffer (pH 7.5) was incubated with 1 mg/mL of purified *sgHldB* (10 μ L) for 5 min. The reaction mixture was then added with 10 μ L substrate containing 100 mM racemic styrene oxide, or 100 mM 1,2-epoxyhexane, or 100 mM 1,2-epoxybutane. The reaction was incubated at 37°C and reactions without substrate or enzyme served as controls. After 16 hr of the reaction, an aliquot of 100 μ L was withdrawn and extracted with equal volume of ethyl acetate. The product and substrate were extracted by vigorous vortexing of the mixture. The dried organic layer obtained using anhydrous Na_2SO_4 and subjected for analysis using GC-MS equipped with Rtx-1 capillary column with dimensions 30 m $\times$ 0.25 mm $\times$ 0.25 μ m. The reaction products were detected by keeping injection temperature at 260°C and the column temperature in the range of 80°C to 300°C at 20°C/min. The gas flow rate was maintained at 1 mL/min. While detecting 1,2-epoxyhexane, and 1,2-hexanediol, the column temperature was ramped at 40°C for 1 min, 40-100°C (10°C/min), 100-150°C (20°C/min) 150-300°C (10°C/min). The products were detected using the mass spectra data obtained and confirmed based on standard mass spectrum retrieved from NIST library for each compound.

6.2.2. Enantiospecific hydrolysis of racemic Styrene oxide

The set-up of the reaction for detecting enantiospecific products obtained after hydrolysis of racemic styrene oxide were same as described in method 6.1.1. After the reaction was incubated for 1 hr, the reaction was stopped by adding equal volume of hexane: IPA (1:1 v/v). The product and substrate were extracted by vigorous vortexing of the mixture. The dried organic layer was subjected for analysis using chiral-HPLC. The HPLC column was CHIRACEL-OB (250 mm× 4.6 mm× 5 µm) equipped with a UV-detector. The mobile phase used for chromatographic separation was Hexane: IPA (95:5 v/v) with flow rate of 0.4 ml/min. The column temperature was maintained at 30°C and the UV-detector set at 220 nm. In this reaction, commercially procured R- or S- phenyl ethanediol were used as standards (1 mM).

6.3. Results & Discussion

6.3.1. Formation of 1,2-phenyl-1-ethanediol

After the indirect spectrophotometric epoxide hydrolase assay of *sgHldB*, the formation of the product 1,2-diol was detected using GC-analysis. After 16 hr of reaction, the conversion of racemic R-S- styrene oxide to (R-)1,2-phenyl-1-ethanediol or (S-)1,2-phenyl-1-ethanediol was detected as seen in fig 6.1. The peak for racemic styrene oxide was detected at a retention time (RT) of 4.46 min and phenyl-ethanediol was detected at 6.83 min (fig 6.1a). The structure and mass spectra depicted in the inset also corroborated with the standard mass spectra of the compounds. Similarly, reaction was also performed for 1-epoxyhexane and 1-epoxybutane. In 1-epoxyhexane, the substrate 1-epoxyhexane was detected as a clear peak at an RT of 5.17 min. The corresponding diol i.e, 1-hexanediol was detected at a RT of 9.29 min (fig 6.1b). The 1,2-hexanediol and corresponding mass spectra are shown. In case of 1-epoxybutane, the substrate peak could not be resolved under these experimental conditions. However, the product 1,2- butanediol was detected at RT of 4.89 min (fig 6.1c). Formation of the products using GC-MS analysis is therefore a clear confirmation of the epoxide hydrolase activity of *sgHldB*.

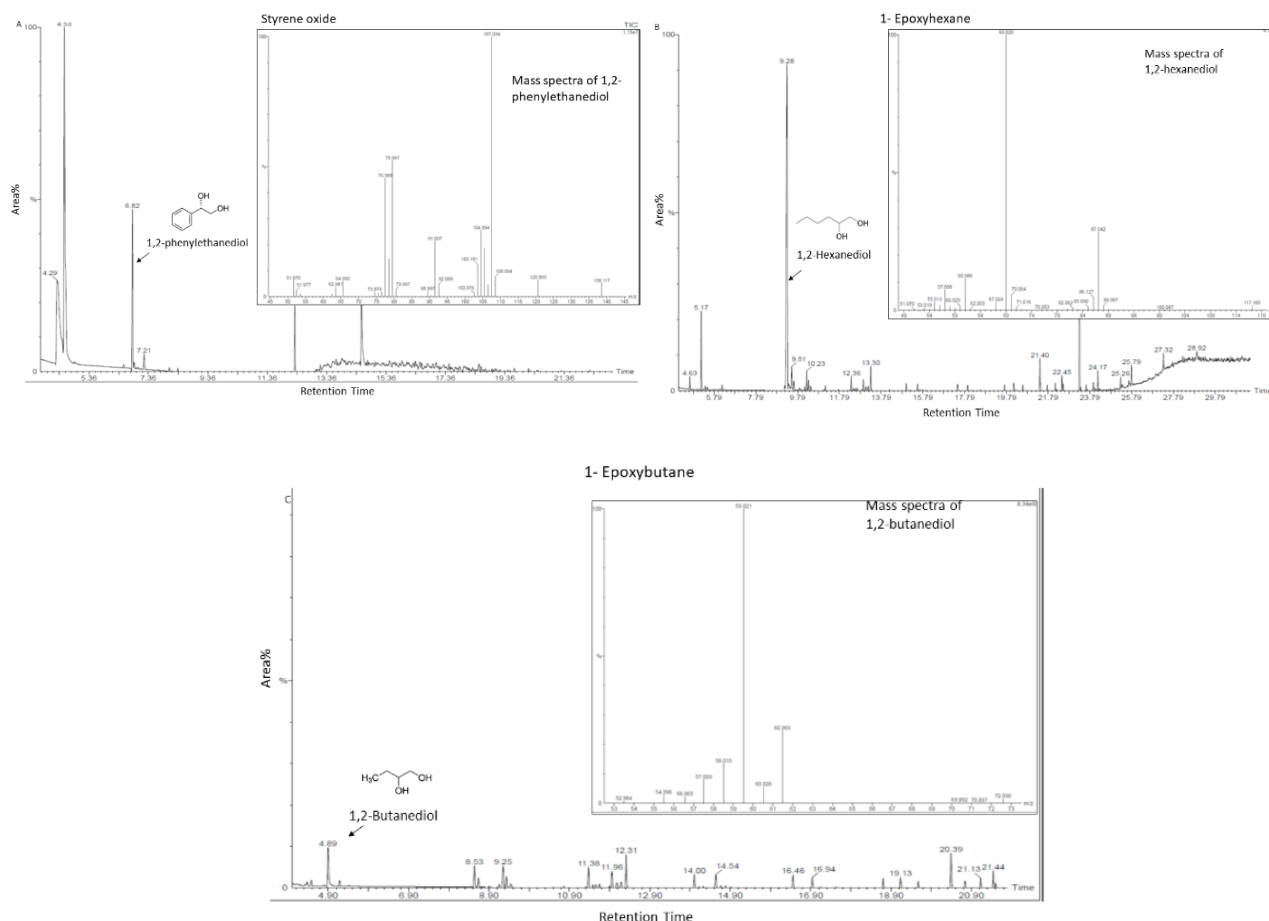


Fig. 6.1 GC-MS analysis of epoxide hydrolase activity of *sgHldB*.

The Panel A shows the GC-peak of styrene oxide and 1-phenyl-1-ethanediol. The arrows show the RT and name of the product. In the inset the its mass spectrum of the product formed is indicated. The Panel B shows the GC-peak of 1,2-epoxyhexane and 1,2-hexanediol. The arrows show the RT and name of the product. In the inset the mass spectrum of product is indicated. The Panel C shows the GC-peak of 1,2-butanediol. The arrow shows the RT and name of the product. The inset indicates its mass spectrum.

After confirming the epoxide hydrolase activity of *sgHldB*, its enantiospecificity was determined. Following is the standard reaction catalyzed by epoxide hydrolase for the substrate racemic (*R*-, *S*-) styrene oxide.

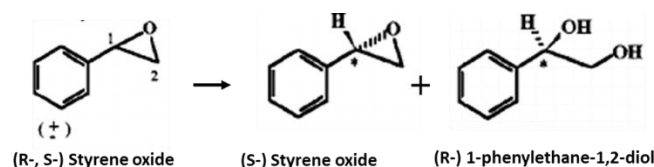


Fig. 6.2 Asymmetric hydrolysis of racemic styrene oxide

The figure shows the hydrolysis of styrene oxide yields enantiopure styrene oxide and (R-) or (S-) 1-phenylethane-1,2-diol. Both the enantiomers of phenylethandiol are considered valuable building blocks in commercial fine chemical industry. As GC-MS cannot differentiate enantiospecific R- or S- isomers of chiral epoxide substrates and products, chiral- HPLC was employed to determine the resolution of substrate racemic styrene oxide during the epoxide hydrolase reaction.

6.3.2. Resolution of (R-, S-) styrene oxide

To elucidate enantioselectivity, progress of hydrolysis was analysed after 60 min of reaction incubation performed according to methods described in section 6.1.2. The commercially available enantiopure R or S- 1-phenylethane-1,2-diol standard was used to distinguish the enantiomers. At the start of the reaction both R and S- phenylethandiol were found with almost same area% indicating no enantiomeric excess (ee) (fig 6.3a). After 1 hr of reaction, R- enantiomer was found to be retained with an area of 65.34% at a retention time (RT) of 34.9 min. The S- enantiomer was found at RT 37.2 min with an area of 34.65% (fig 6.3b). This indicates that the epoxide hydrolase *sgHldB* favours the formation of R-phenylethandiol with an enantiomeric excess (ee) of ~30% in 1 hr from the racemic styrene oxide in the given condition.

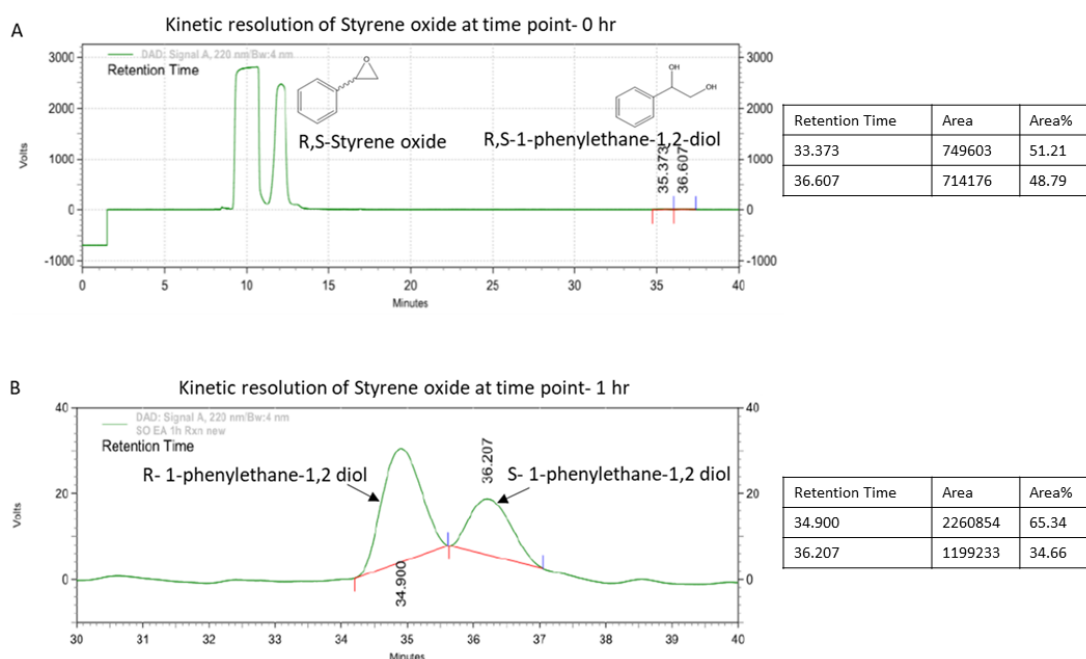


Fig. 6.3 Chiral HPLC of R-, S- Styrene oxide.

Panel A indicates the HPLC peaks found at start of the reaction. Panel B indicates HPLC peaks found after 1 hr of reaction. The arrows indicate the retention time of R and S enantiomers of phenylethanedol distinguished by chiral HPLC.

Amongst the bacterial EH enzymes, so far kinetic resolution using styrene oxide has been well-studied in only a handful of microorganisms including *Aspergillus sp.*, *Rhodotorula glutinis* and *Sphingomonas sp* with significant enantioselectivity (J. W. Lee et al., 2003; Morisseau et al., 1997; Wu, Li, Chin, & Li, 2013). The study shows a novel, enantioselective epoxide hydrolase of *S. granuli* MG i.e *sgHldB* that shows enantiospecific hydrolysis of styrene oxide. As discussed earlier the enzyme *sgHldB* showed a wide substrate acceptance with both aromatic epoxides like styrene oxide and aliphatic epoxides like 1-epoxyhexane and 1-epoxybutane as substrates. The epoxide hydrolase *sgHldB* from *S. granuli* MG can be considered as a potential biocatalyst to selectively synthesize industrially valuable vicinal diols. Furthermore, with processed parameters and by employing rational engineering methods, a robust catalyst can be generated.

CONCLUSIONS

The genome sequence of *B. diminuta* revealed the presence of a single chromosome and two plasmids pCMS1 and pCMS2. The plasmid pCMS2 is a novel plasmid and is identified in this study.

The genome sequence shows similarity with the members of Sphingomonadaceae. The core genome assisted phylogenetic tree reclassified *B. diminuta* as *Sphingopyxis* sp MG. It belongs to the *granuli* sps. Therefore *B. diminuta* is renamed as *Sphingopyxis. granuli* MG.

The genome shows presence of 27 Genomic Island (GI). The largest GI contains 113 ORF and the smallest has just 5 ORFs. The largest GI contains genetic repertoire to code complete type IV secretary system (T4SS) along with genes that contribute for metal resistance. It is an Integrative Conjugative Element. Novel co-factor F₄₂₀ biosynthesis genes and Arsenic resistance genes are also found as part of GI.

Sphingopyxis. granuli MG chromosome contains upper pathway genes to convert Nitrophenols like PNP, 3-Nitrophenol, halogenated organic compounds like γ -HCH, phenol derivatives like Toluene and Xylene into catechol. Ortho and meta-cleavage pathways are encoded by plasmids pCMS1 and pCMS2.

The bioremediation potential of *Sphingopyxis. granuli* MG was assessed by measuring *in-situ* mineralization of ¹⁴C-labelled p-nitrophenol in soil microcosm. *Sphingopyxis. granuli* MG survived and effectively eliminated fortified PNP when present as a mixed culture. The *Sphingopyxis. granuli* MG failed to grow in microcosm as a monoculture.

Eight dehalogenases were identified in *Sphingopyxis. granuli* MG of which six show sequence similarity with Lindane (HCH)- dehydrochlorinase LinA. However, only one (*sg*LinA5) shows HCH-dechlorinase activity.

Two haloalkane dehalogenases were also identified in the genome of *Sphingopyxis. granuli* MG. One of them (*sg*HldA) is similar to HLD I subfamily and has good specific activity against a broad range of haloalkanes. The second HLD designated as *sg*HldB is structural homologue of epoxide hydrolase and shows promiscuous haloalkane dehalogenase activity.

The *sg*HldB displayed both haloalkane dehalogenase and epoxide hydrolase activity. Site directed mutagenesis shows involvement of a non-canonical Phe-Tyr pair in promiscuous dehalogenase activity.

The *sg*HldB is considered to be an evolutionary intermediate between haloalkane dehalogenases and epoxide hydrolases. It can be used as a molecular platform to engineer substrate specific haloalkane dehalogenases/ epoxide hydrolases.

The *sg*HldB showed good specific activity for both aromatic and aliphatic epoxides. It converts them into respective diols and it is an R- enantioselective enzyme with respect to styrene oxide with ee value of 30% for R-diol.

.....

REFERENCES

- Abernathy, C. O., Thomas, D. J., & Calderon, R. L. (2003). Health effects and risk assessment of arsenic. *J Nutr*, 133(5 Suppl 1), 1536S-1538S. doi:10.1093/jn/133.5.1536S
- Afriat, L., Roodveldt, C., Manco, G., & Tawfik, D. S. (2006). The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. *Biochemistry*, 45(46), 13677-13686. doi:10.1021/bi061268r
- Alav, I., Sutton, J. M., & Rahman, K. M. (2018). Role of bacterial efflux pumps in biofilm formation. *J Antimicrob Chemother*, 73(8), 2003-2020. doi:10.1093/jac/dky042
- Alcalde-Rico, M., Hernando-Amado, S., Blanco, P., & Martinez, J. L. (2016). Multidrug Efflux Pumps at the Crossroad between Antibiotic Resistance and Bacterial Virulence. *Front Microbiol*, 7, 1483. doi:10.3389/fmicb.2016.01483
- Anisimova, M., Nielsen, R., & Yang, Z. (2003). Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics*, 164(3), 1229-1236. doi:10.1093/genetics/164.3.1229
- Annamaria, H., Manno, D., Strand, S. E., Bruce, N. C., & Hawari, J. (2010). Biodegradation of RDX and MNX with *Rhodococcus* sp. strain DN22: new insights into the degradation pathway. *Environ Sci Technol*, 44(24), 9330-9336. doi:10.1021/es1023724
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., . . . Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*, 19(5), 455-477. doi:10.1089/cmb.2012.0021
- Bashiri, G., Rehan, A. M., Sreebhavan, S., Baker, H. M., Baker, E. N., & Squire, C. J. (2016). Elongation of the Poly-gamma-glutamate Tail of F420 Requires Both Domains of the F420:gamma-Glutamyl Ligase (FbiB) of *Mycobacterium tuberculosis*. *J Biol Chem*, 291(13), 6882-6894. doi:10.1074/jbc.M115.689026
- Ben Fekih, I., Zhang, C., Li, Y. P., Zhao, Y., Alwathnani, H. A., Saquib, Q., . . . Cervantes, C. (2018). Distribution of Arsenic Resistance Genes in Prokaryotes. *Front Microbiol*, 9, 2473. doi:10.3389/fmicb.2018.02473
- Benning, M. M., Shim, H., Raushel, F. M., & Holden, H. M. (2001). High Resolution X-ray Structures of Different Metal-Substituted Forms of Phosphotriesterase from *Pseudomonas diminuta*. *Biochemistry*, 40(9), 2712-2722. doi:10.1021/bi002661e
- Berger, M., Löffler, D., Ternes, T., Heininger, P., Ricking, M., & Schwarzbauer, J. (2016). Hexachlorocyclohexane derivatives in industrial waste and samples from a contaminated riverine system. *Chemosphere*, 150, 219-226. doi:10.1016/j.chemosphere.2016.01.122
- Bertelli, C., & Brinkman, F. S. L. (2018). Improved genomic island predictions with IslandPath-DIMOB. *Bioinformatics*, 34(13), 2161-2167. doi:10.1093/bioinformatics/bty095
- Besemer, J., Lomsadze, A., & Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res*, 29(12), 2607-2618. doi:10.1093/nar/29.12.2607
- Bidmanova, S., Chaloupkova, R., Damborsky, J., & Prokop, Z. (2010). Development of an enzymatic fiber-optic biosensor for detection of halogenated hydrocarbons. *Anal Bioanal Chem*, 398(5), 1891-1898. doi:10.1007/s00216-010-4083-z
- Bigley, A. N., Narindoshvili, T., & Raushel, F. M. (2020). A Chemoenzymatic Synthesis of the (RP)-Isomer of the Antiviral Prodrug Remdesivir. *Biochemistry*, 59(33), 3038-3043. doi:10.1021/acs.biochem.0c00591
- Bigley, A. N., Xu, C., Henderson, T. J., Harvey, S. P., & Raushel, F. M. (2013). Enzymatic neutralization of the chemical warfare agent VX: evolution of phosphotriesterase for phosphorothiolate hydrolysis. *J Am Chem Soc*, 135(28), 10426-10432. doi:10.1021/ja402832z
- Birnboim, H. C., & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*, 7(6), 1513-1523. doi:10.1093/nar/7.6.1513
- Blanco, P., Hernando-Amado, S., Reales-Calderon, J. A., Corona, F., Lira, F., Alcalde-Rico, M., . . . Martinez, J. L. (2016). Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants. *Microorganisms*, 4(1). doi:10.3390/microorganisms4010014
- Boetzer, M., & Pirovano, W. (2014). SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics*, 15, 211. doi:10.1186/1471-2105-15-211

- Bokinsky, G., Baidoo, E. E., Akella, S., Burd, H., Weaver, D., Alonso-Gutierrez, J., . . . Keasling, J. D. (2013). HipA-triggered growth arrest and beta-lactam tolerance in *Escherichia coli* are mediated by RelA-dependent ppGpp synthesis. *J Bacteriol*, 195(14), 3173-3182. doi:10.1128/JB.02210-12
- Bottalla, A. L., Ibrahim-Ouali, M., Santelli, M., Furstoss, R., Archelas, A. J. A. S., & Catalysis. (2007). Epoxide Hydrolase-Catalysed Kinetic Resolution of a Spiroepoxide, a Key Building Block of Various 11-Heterosteroids. 349(7), 1102-1110.
- Boubakri, H., Beuf, M., Simonet, P., & Vogel, T. M. (2006). Development of metagenomic DNA shuffling for the construction of a xenobiotic gene. *Gene*, 375, 87-94. doi:10.1016/j.gene.2006.02.027
- Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., . . . Xia, F. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep*, 5, 8365. doi:10.1038/srep08365
- Burrus, V. (2017). Mechanisms of stabilization of integrative and conjugative elements. *Curr Opin Microbiol*, 38, 44-50. doi:10.1016/j.mib.2017.03.014
- Cai, M., & Xun, L. (2002). Organization and regulation of pentachlorophenol-degrading genes in *Sphingobium chlorophenolicum* ATCC 39723. *J Bacteriol*, 184(17), 4672-4680. doi:10.1128/JB.184.17.4672-4680.2002
- Camacho-Perez, B., Rios-Leal, E., Rinderknecht-Seijas, N., & Poggi-Varaldo, H. M. (2012). Enzymes involved in the biodegradation of hexachlorocyclohexane: a mini review. *J Environ Manage*, 95 Suppl, S306-318. doi:10.1016/j.jenvman.2011.06.047
- Campbell, D. W., Muller, C., & Reardon, K. F. (2006). Development of a fiber optic enzymatic biosensor for 1,2-dichloroethane. *Biotechnol Lett*, 28(12), 883-887. doi:10.1007/s10529-006-9014-x
- Caputo, A., Fournier, P. E., & Raoult, D. (2019). Genome and pan-genome analysis to classify emerging bacteria. *Biol Direct*, 14(1), 5. doi:10.1186/s13062-019-0234-0
- Carlin, A., Shi, W., Dey, S., & Rosen, B. P. (1995). The ars operon of *Escherichia coli* confers arsenical and antimonial resistance. *J Bacteriol*, 177(4), 981-986. doi:10.1128/jb.177.4.981-986.1995
- Carlucci, L., Zhou, E., Malashkevich, V. N., Almo, S. C., & Mundorff, E. C. (2016). Biochemical characterization of two haloalkane dehalogenases: DccA from *Caulobacter crescentus* and DsaA from *Saccharomonospora azurea*. *Protein Sci*, 25(4), 877-886. doi:10.1002/pro.2895
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*, 17(4), 540-552. doi:10.1093/oxfordjournals.molbev.a026334
- Chakka, D., Gudla, R., Madikonda, A. K., Pandeeti, E. V., Parthasarathy, S., Nandavaram, A., & Siddavattam, D. (2015). The Organophosphate Degradation (opd) Island-borne Esterase-induced Metabolic Diversion in *Escherichia coli* and Its Influence on p-Nitrophenol Degradation. *J Biol Chem*, 290(50), 29920-29930. doi:10.1074/jbc.M115.661249
- Chen, H. (2006). Recent advances in azo dye degrading enzyme research. *Curr Protein Pept Sci*, 7(2), 101-111. doi:10.2174/138920306776359786
- Chen, S., Dong, Y. H., Chang, C., Deng, Y., Zhang, X. F., Zhong, G., . . . Zhang, L. H. (2013). Characterization of a novel cyfluthrin-degrading bacterial strain *Brevibacterium aureum* and its biochemical degradation pathway. *Bioresour Technol*, 132, 16-23. doi:10.1016/j.biortech.2013.01.002
- Chen, Y., Ye, W., Zhang, Y., & Xu, Y. (2015). High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic Acids Res*, 43(16), 7762-7768. doi:10.1093/nar/gkv784
- Choi, W. J., Choi, C. Y. J. B., & Engineering, B. (2005). Production of chiral epoxides: epoxide hydrolase-catalyzed enantioselective hydrolysis. 10(3), 167-179.
- Choi, W. J., Puah, S. M., Tan, L. L., & Ng, S. S. (2008). Production of (R)-ethyl-3,4-epoxybutyrate by newly isolated *Acinetobacter baumannii* containing epoxide hydrolase. *Appl Microbiol Biotechnol*, 79(1), 61-67. doi:10.1007/s00253-008-1405-6
- Chovancova, E., Kosinski, J., Bujnicki, J. M., & Damborsky, J. (2007). Phylogenetic analysis of haloalkane dehalogenases. *Proteins*, 67(2), 305-316. doi:10.1002/prot.21313
- Cleij, M., Archelas, A., & Furstoss, R. (1999). Microbiological Transformations 43. Epoxide Hydrolases as Tools for the Synthesis of Enantiopure alpha-Methylstyrene Oxides: A New and Efficient Synthesis of (S)-Ibuprofen. *J Org Chem*, 64(14), 5029-5035. doi:10.1021/jo982101+
- Copley, S. D. (2003). Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr Opin Chem Biol*, 7(2), 265-272. doi:10.1016/s1367-5931(03)00032-2
- Copley, S. D. (2017). Shining a light on enzyme promiscuity. *Curr Opin Struct Biol*, 47, 167-175. doi:10.1016/j.sbi.2017.11.001

- Cucinotta, D., & Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta Biomed*, 91(1), 157-160. doi:10.23750/abm.v91i1.9397
- Dai, M., Rogers, J. B., Warner, J. R., & Copley, S. D. (2003). A previously unrecognized step in pentachlorophenol degradation in *Sphingobium chlorophenolicum* is catalyzed by tetrachlorobenzoquinone reductase (PcpD). *J Bacteriol*, 185(1), 302-310. doi:10.1128/JB.185.1.302-310.2003
- Damborsky, J., & Koca, J. (1999). Analysis of the reaction mechanism and substrate specificity of haloalkane dehalogenases by sequential and structural comparisons. *Protein Eng*, 12(11), 989-998. doi:10.1093/protein/12.11.989
- Darling, A. C., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*, 14(7), 1394-1403. doi:10.1101/gr.2289704
- De Luca, V., & Mandrich, L. (2020). Enzyme Promiscuous Activity: How to Define it and its Evolutionary Aspects. *Protein Pept Lett*, 27(5), 400-410. doi:10.2174/0929866527666191223141205
- de Vries, E. J., & Janssen, D. B. J. C. O. i. B. (2003). Biocatalytic conversion of epoxides. 14(4), 414-420.
- Del Vecchio, P., Elias, M., Merone, L., Graziano, G., Dupuy, J., Mandrich, L., . . . Manco, G. (2009). Structural determinants of the high thermal stability of SsoPox from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Extremophiles*, 13(3), 461-470. doi:10.1007/s00792-009-0231-9
- Dong, Y. J., Bartlam, M., Sun, L., Zhou, Y. F., Zhang, Z. P., Zhang, C. G., . . . Zhang, X. E. (2005). Crystal structure of methyl parathion hydrolase from *Pseudomonas* sp. WBC-3. *J Mol Biol*, 353(3), 655-663. doi:10.1016/j.jmb.2005.08.057
- Dormer, P. (1981). Quantitative carbon-14 autoradiography at the cellular level: principles and application for cell kinetic studies. *Histochem J*, 13(2), 161-171. doi:10.1007/BF01006877
- Dubinkina, V., Fridman, Y., Pandey, P. P., & Maslov, S. (2019). Multistability and regime shifts in microbial communities explained by competition for essential nutrients. *Elife*, 8. doi:10.7554/eLife.49720
- Dumas, D. P., Caldwell, S. R., Wild, J. R., & Raushel, F. M. (1989). Purification and properties of the phosphotriesterase from *Pseudomonas diminuta*. *J Biol Chem*, 264(33), 19659-19665.
- Duque, E., Segura, A., Mosqueda, G., & Ramos, J. L. (2001). Global and cognate regulators control the expression of the organic solvent efflux pumps TtgABC and TtgDEF of *Pseudomonas putida*. *Mol Microbiol*, 39(4), 1100-1106. doi:10.1046/j.1365-2958.2001.02310.x
- Dvorak, P., Bidmanova, S., Damborsky, J., & Prokop, Z. (2014). Immobilized synthetic pathway for biodegradation of toxic recalcitrant pollutant 1,2,3-trichloropropane. *Environ Sci Technol*, 48(12), 6859-6866. doi:10.1021/es500396r
- Eguchi, M., Ostrowski, M., Fegatella, F., Bowman, J., Nichols, D., Nishino, T., & Cavicchioli, R. (2001). *Sphingomonas alaskensis* strain AFO1, an abundant oligotrophic ultramicrobacterium from the North Pacific. *Appl Environ Microbiol*, 67(11), 4945-4954. doi:10.1128/AEM.67.11.4945-4954.2001
- Elias, M., & Tawfik, D. S. (2012). Divergence and convergence in enzyme evolution: parallel evolution of paraoxonases from quorum-quenching lactonases. *J Biol Chem*, 287(1), 11-20. doi:10.1074/jbc.R111.257329
- Epstein, L. (2014). Fifty years since Silent Spring. *Annu Rev Phytopathol*, 52, 377-402. doi:10.1146/annurev-phyto-102313-045900
- Esquirol, L., Peat, T. S., Sugrue, E., Balotra, S., Rottet, S., Warden, A. C., . . . Scott, C. (2020). Bacterial catabolism of s-triazine herbicides: biochemistry, evolution and application. *Adv Microb Physiol*, 76, 129-186. doi:10.1016/bs.ampbs.2020.01.004
- Ferrandi, E. E., Sayer, C., De Rose, S. A., Guazzelli, E., Marchesi, C., Saneei, V., . . . Monti, D. (2018). New Thermophilic alpha/beta Class Epoxide Hydrolases Found in Metagenomes From Hot Environments. *Front Bioeng Biotechnol*, 6, 144. doi:10.3389/fbioe.2018.00144
- Forouhar, F., Abashidze, M., Xu, H., Grochowski, L. L., Seetharaman, J., Hussain, M., . . . Tong, L. (2008). Molecular insights into the biosynthesis of the F420 coenzyme. *J Biol Chem*, 283(17), 11832-11840. doi:10.1074/jbc.M710352200
- Franken, S. M., Rozeboom, H. J., Kalk, K. H., & Dijkstra, B. W. (1991). Crystal structure of haloalkane dehalogenase: an enzyme to detoxify halogenated alkanes. *EMBO J*, 10(6), 1297-1302.
- Galperin, M. Y., & Koonin, E. V. (2012). Divergence and convergence in enzyme evolution. *J Biol Chem*, 287(1), 21-28. doi:10.1074/jbc.R111.241976
- Gao, J., Ellis, L. B., & Wackett, L. P. (2010). The University of Minnesota Biocatalysis/Biodegradation Database: improving public access. *Nucleic Acids Res*, 38(Database issue), D488-491. doi:10.1093/nar/gkp771

- Garcia-Romero, I., Forstner, K. U., Santero, E., & Floriano, B. (2018). SuhB, a small non-coding RNA involved in catabolite repression of tetralin degradation genes in *Sphingopyxis granuli* strain TFA. *Environ Microbiol*, 20(10), 3671-3683. doi:10.1111/1462-2920.14360
- Genzel, Y., Archelas, A., Broxterman, Q. B., Schulze, B., & Furstoss, R. (2002). Microbiological transformations 50: selection of epoxide hydrolases for enzymatic resolution of 2-, 3- or 4-pyridyloxirane. *Journal of Molecular Catalysis B: Enzymatic*, 16(5), 217-222. doi:https://doi.org/10.1016/S1381-1177(01)00064-9
- George, K. W., & Hagglom, M. M. (2008). Microbial O-methylation of the flame retardant tetrabromobisphenol-A. *Environ Sci Technol*, 42(15), 5555-5561. doi:10.1021/es800038q
- Gilchrist, T. L. (1992). *Heterocyclic chemistry*. Harlow, Essex, England: Longman Scientific & Technical.
- Glasner, M. E., Truong, D. P., & Morse, B. C. (2020). How enzyme promiscuity and horizontal gene transfer contribute to metabolic innovation. *FEBS J*, 287(7), 1323-1342. doi:10.1111/febs.15185
- Gorla, P., Pandey, J. P., Parthasarathy, S., Merrick, M., & Siddavattam, D. (2009). Organophosphate hydrolase in *Brevundimonas diminuta* is targeted to the periplasmic face of the inner membrane by the twin arginine translocation pathway. *J Bacteriol*, 191(20), 6292-6299. doi:10.1128/JB.00824-09
- Greening, C., Jirapanjawan, T., Afroze, S., Ney, B., Scott, C., Pandey, G., . . . Warden, A. C. (2017). Mycobacterial F420H2-Dependent Reductases Promiscuously Reduce Diverse Compounds through a Common Mechanism. *Front Microbiol*, 8, 1000. doi:10.3389/fmicb.2017.01000
- Habe, H., & Otori, T. (2003). Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. *Biosci Biotechnol Biochem*, 67(2), 225-243. doi:10.1271/bbb.67.225
- Hagemeyer, C. H., Shima, S., Thauer, R. K., Bourenkov, G., Bartunik, H. D., & Ermler, U. (2003). Coenzyme F420-dependent methylenetetrahydromethanopterin dehydrogenase (Mtd) from *Methanopyrus kandleri*: a methanogenic enzyme with an unusual quaternary structure. *J Mol Biol*, 332(5), 1047-1057. doi:10.1016/S0022-2836(03)00949-5
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol*, 166(4), 557-580. doi:10.1016/S0022-2836(83)80284-8
- Harayama, S., Rekik, M., Bairoch, A., Neidle, E. L., & Ornston, L. N. (1991). Potential DNA slippage structures acquired during evolutionary divergence of *Acinetobacter calcoaceticus* chromosomal benABC and *Pseudomonas putida* TOL pWW0 plasmid xylXYZ, genes encoding benzoate dioxygenases. *J Bacteriol*, 173(23), 7540-7548. doi:10.1128/jb.173.23.7540-7548.1991
- Hawwa, R., Aikens, J., Turner, R. J., Santarsiero, B. D., & Mesecar, A. D. (2009). Structural basis for thermostability revealed through the identification and characterization of a highly thermostable phosphotriesterase-like lactonase from *Geobacillus stearothermophilus*. *Arch Biochem Biophys*, 488(2), 109-120. doi:10.1016/j.abb.2009.06.005
- Helbling, D. E., Hollender, J., Kohler, H. P., Singer, H., & Fenner, K. (2010). High-throughput identification of microbial transformation products of organic micropollutants. *Environ Sci Technol*, 44(17), 6621-6627. doi:10.1021/es100970m
- Holden, H. M., & Raushel, F. M. (2021). From the Three-Dimensional Structure of Phosphotriesterase. *Biochemistry*, 60(46), 3413-3415. doi:10.1021/acs.biochem.1c00311
- Holloway, P., Trevors, J. T., & Lee, H. J. J. o. m. m. (1998). A colorimetric assay for detecting haloalkane dehalogenase activity. 32(1), 31-36.
- Holmquist, M. (2000). Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr Protein Pept Sci*, 1(2), 209-235. doi:10.2174/1389203003381405
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119. doi:10.1186/1471-2105-11-119
- Imai, R., Nagata, Y., Fukuda, M., Takagi, M., & Yano, K. (1991). Molecular cloning of a *Pseudomonas paucimobilis* gene encoding a 17-kilodalton polypeptide that eliminates HCl molecules from gamma-hexachlorocyclohexane. *J Bacteriol*, 173(21), 6811-6819. doi:10.1128/jb.173.21.6811-6819.1991
- Iwasaki, I., Utsumi, S., & Ozawa, T. J. B. o. t. C. S. o. J. (1952). New colorimetric determination of chloride using mercuric thiocyanate and ferric ion. 25(3), 226-226.
- J.F, S., & Russell, D. (2001). *Molecular Cloning: A Laboratory Manual (3-Volume Set)* (Vol. 1).
- Jensen, R. A. (1976). Enzyme recruitment in evolution of new function. *Annu Rev Microbiol*, 30, 409-425. doi:10.1146/annurev.mi.30.100176.002205

- Jindal, S., Dua, A., & Lal, R. (2013). *Sphingopyxis indica* sp. nov., isolated from a high dose point hexachlorocyclohexane (HCH)-contaminated dumpsite. *Int J Syst Evol Microbiol*, 63(Pt 6), 2186-2191. doi:10.1099/ijs.0.040840-0
- Jochens, H., Stiba, K., Savile, C., Fujii, R., Yu, J. G., Gerassenskov, T., . . . Bornscheuer, U. T. (2009). Converting an esterase into an epoxide hydrolase. *Angew Chem Int Ed Engl*, 48(19), 3532-3535. doi:10.1002/anie.200806276
- Johnson, C. M., & Grossman, A. D. (2015). Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu Rev Genet*, 49, 577-601. doi:10.1146/annurev-genet-112414-055018
- Jones, J. J., & Falkinham, J. O., 3rd. (2003). Decolorization of malachite green and crystal violet by waterborne pathogenic mycobacteria. *Antimicrob Agents Chemother*, 47(7), 2323-2326. doi:10.1128/AAC.47.7.2323-2326.2003
- Juhas, M., Crook, D. W., & Hood, D. W. (2008). Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol*, 10(12), 2377-2386. doi:10.1111/j.1462-5822.2008.01187.x
- Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev*, 33(2), 376-393. doi:10.1111/j.1574-6976.2008.00136.x
- Kampfer, P., Witzemberger, R., Denner, E. B., Busse, H. J., & Neef, A. (2002). *Sphingopyxis wiflariensis* sp. nov., isolated from activated sludge. *Int J Syst Evol Microbiol*, 52(Pt 6), 2029-2034. doi:10.1099/00207713-52-6-2029
- Karns, J. S., Mulbry, W. W., Nelson, J. O., & Kearney, P. C. (1986). Metabolism of carbofuran by a pure bacterial culture. *Pesticide Biochemistry and Physiology*, 25(2), 211-217. doi:https://doi.org/10.1016/0048-3575(86)90048-9
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, 30(4), 772-780. doi:10.1093/molbev/mst010
- Kertesz, M. A. (2001). Bacterial transporters for sulfate and organosulfur compounds. *Res Microbiol*, 152(3-4), 279-290. doi:10.1016/s0923-2508(01)01199-8
- Khersonsky, O., & Tawfik, D. S. (2010). Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem*, 79, 471-505. doi:10.1146/annurev-biochem-030409-143718
- Kim, J., Kim, S. J., Kim, S. H., Kim, S. I., Moon, Y. J., Park, S. J., . . . Chung, Y. H. (2014). Draft Genome Sequence of *Sphingopyxis* sp. Strain MWB1, a Crude-Oil-Degrading Marine Bacterium. *Genome Announc*, 2(6). doi:10.1128/genomeA.01256-14
- Kim, M. K., Im, W. T., Ohta, H., Lee, M., & Lee, S. T. (2005). *Sphingopyxis granuli* sp. nov., a beta-glucosidase-producing bacterium in the family Sphingomonadaceae in alpha-4 subclass of the Proteobacteria. *J Microbiol*, 43(2), 152-157.
- Kloske, M., & Witkiewicz, Z. (2019). Novichoks - The A group of organophosphorus chemical warfare agents. *Chemosphere*, 221, 672-682. doi:10.1016/j.chemosphere.2019.01.054
- Kolvenbach, B. A., Helbling, D. E., Kohler, H. P., & Corvini, P. F. (2014). Emerging chemicals and the evolution of biodegradation capacities and pathways in bacteria. *Curr Opin Biotechnol*, 27, 8-14. doi:10.1016/j.copbio.2013.08.017
- Koudelakova, T., Bidmanova, S., Dvorak, P., Pavelka, A., Chaloupkova, R., Prokop, Z., & Damborsky, J. (2013). Haloalkane dehalogenases: biotechnological applications. *Biotechnol J*, 8(1), 32-45. doi:10.1002/biot.201100486
- Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M., Fortova, A., Jarkovsky, J., & Damborsky, J. (2011). Substrate specificity of haloalkane dehalogenases. *Biochem J*, 435(2), 345-354. doi:10.1042/BJ20101405
- Kourist, R., Bartsch, S., Fransson, L., Hult, K., & Bornscheuer, U. T. (2008). Understanding promiscuous amidase activity of an esterase from *Bacillus subtilis*. *Chembiochem*, 9(1), 67-69. doi:10.1002/cbic.200700521
- Krooshof, G. H., Kwant, E. M., Damborsky, J., Koca, J., & Janssen, D. B. (1997). Repositioning the catalytic triad aspartic acid of haloalkane dehalogenase: effects on stability, kinetics, and structure. *Biochemistry*, 36(31), 9571-9580. doi:10.1021/bi971014t
- Kroutil, W., Osprian, I., Mischitz, M., & Faber, K. J. S. (1997). Chemoenzymatic synthesis of (S)-(-)-Frontalin using bacterial epoxide hydrolases. 1997(02), 156-158.
- Kuck, P., & Longo, G. C. (2014). FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front Zool*, 11(1), 81. doi:10.1186/s12983-014-0081-x

- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*, 35(6), 1547-1549. doi:10.1093/molbev/msy096
- Lagesen, K., Hallin, P., Rodland, E. A., Staerfeldt, H. H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*, 35(9), 3100-3108. doi:10.1093/nar/gkm160
- Lahoda, M., Mesters, J. R., Stsiapanava, A., Chaloupkova, R., Kutý, M., Damborsky, J., & Kuta Smatanova, I. (2014). Crystallographic analysis of 1,2,3-trichloropropane biodegradation by the haloalkane dehalogenase DhaA31. *Acta Crystallogr D Biol Crystallogr*, 70(Pt 2), 209-217. doi:10.1107/S1399004713026254
- Lal, R., Pandey, G., Sharma, P., Kumari, K., Malhotra, S., Pandey, R., . . . Oakeshott, J. G. (2010). Biochemistry of microbial degradation of hexachlorocyclohexane and prospects for bioremediation. *Microbiol Mol Biol Rev*, 74(1), 58-80. doi:10.1128/MMBR.00029-09
- Langille, M. G., & Brinkman, F. S. (2009). IslandViewer: an integrated interface for computational identification and visualization of genomic islands. *Bioinformatics*, 25(5), 664-665. doi:10.1093/bioinformatics/btp030
- Larcher, S., & Yargeau, V. (2011). Biodegradation of sulfamethoxazole by individual and mixed bacteria. *Appl Microbiol Biotechnol*, 91(1), 211-218. doi:10.1007/s00253-011-3257-8
- Larsen, M. W., Zielinska, D. F., Martinelle, M., Hidalgo, A., Jensen, L. J., Bornscheuer, U. T., & Hult, K. (2010). Suppression of water as a nucleophile in *Candida antarctica* lipase B catalysis. *ChemBiochem*, 11(6), 796-801. doi:10.1002/cbic.200900743
- Laslett, D., & Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res*, 32(1), 11-16. doi:10.1093/nar/gkh152
- Lee, D. G., Zhao, F., Rezenom, Y. H., Russell, D. H., & Chu, K. H. (2012). Biodegradation of triclosan by a wastewater microorganism. *Water Res*, 46(13), 4226-4234. doi:10.1016/j.watres.2012.05.025
- Lee, J. W., Lee, E. J., Yoo, S. S., Park, S. H., Kim, H. S., Lee, E. Y. J. B., & Engineering, B. (2003). Enantioselective hydrolysis of racemic styrene oxide by epoxide hydrolase of *Rhodospiridium kratochvilovae* SYU-08. 8(5), 306-308.
- Lefebure, T., & Stanhope, M. J. (2009). Pervasive, genome-wide positive selection leading to functional divergence in the bacterial genus *Campylobacter*. *Genome Res*, 19(7), 1224-1232. doi:10.1101/gr.089250.108
- Li, L., Stoeckert, C. J., Jr., & Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*, 13(9), 2178-2189. doi:10.1101/gr.1224503
- Li, X. Z., Plesiat, P., & Nikaido, H. (2015). The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin Microbiol Rev*, 28(2), 337-418. doi:10.1128/CMR.00117-14
- Li, Y. F., Scholtz, M. T., & Van Heyst, B. J. (2003). Global gridded emission inventories of beta-hexachlorocyclohexane. *Environ Sci Technol*, 37(16), 3493-3498. doi:10.1021/es034157d
- Liang, C., Tian, L., Liu, Y., Hui, N., Qiao, G., Li, H., . . . Zhao, X. (2020). A promising antiviral candidate drug for the COVID-19 pandemic: A mini-review of remdesivir. *Eur J Med Chem*, 201, 112527. doi:10.1016/j.ejmech.2020.112527
- Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., . . . Wood, K. V. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol*, 3(6), 373-382. doi:10.1021/cb800025k
- Lu, P., Li, Q., Liu, H., Feng, Z., Yan, X., Hong, Q., & Li, S. (2013). Biodegradation of chlorpyrifos and 3,5,6-trichloro-2-pyridinol by *Cupriavidus* sp. DT-1. *Bioresour Technol*, 127, 337-342. doi:10.1016/j.biortech.2012.09.116
- Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., . . . Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*, 47(W1), W636-W641. doi:10.1093/nar/gkz268
- Mahmood, S., Khalid, A., Arshad, M., Mahmood, T., & Crowley, D. E. (2016). Detoxification of azo dyes by bacterial oxidoreductase enzymes. *Crit Rev Biotechnol*, 36(4), 639-651. doi:10.3109/07388551.2015.1004518
- Massey, I. Y., Zhang, X., & Yang, F. (2018). Importance of bacterial biodegradation and detoxification processes of microcystins for environmental health. *J Toxicol Environ Health B Crit Rev*, 21(6-8), 357-369. doi:10.1080/10937404.2018.1532701

- Medini, D., Donati, C., Tettelin, H., Massignani, V., & Rappuoli, R. (2005). The microbial pan-genome. *Curr Opin Genet Dev*, 15(6), 589-594. doi:10.1016/j.gde.2005.09.006
- Merkl, R. (2004). SIGI: score-based identification of genomic islands. *BMC Bioinformatics*, 5, 22. doi:10.1186/1471-2105-5-22
- Meulenbergh, R., Pepi, M., & de Bont, J. A. (1996). Degradation of 3-nitrophenol by *Pseudomonas putida* B2 occurs via 1,2,4-benzenetriol. *Biodegradation*, 7(4), 303-311. doi:10.1007/BF00115744
- Mindrebo, J. T., Nartey, C. M., Seto, Y., Burkart, M. D., & Noel, J. P. (2016). Unveiling the functional diversity of the alpha/beta hydrolase superfamily in the plant kingdom. *Curr Opin Struct Biol*, 41, 233-246. doi:10.1016/j.sbi.2016.08.005
- Mira, A., Martin-Cuadrado, A. B., D'Auria, G., & Rodriguez-Valera, F. (2010). The bacterial pan-genome: a new paradigm in microbiology. *Int Microbiol*, 13(2), 45-57. doi:10.2436/20.1501.01.110
- Monfort, N., Archelas, A., & Furstoss, R. J. T. (2004). Enzymatic transformations. Part 55: Highly productive epoxide hydrolase catalysed resolution of an azole antifungal key synthon. 60(3), 601-605.
- Monterde, M. I., Lombard, M., Archelas, A., Cronin, A., Arand, M., & Furstoss, R. J. T. A. (2004). Enzymatic transformations. Part 58: Enantioconvergent bihydrolysis of styrene oxide derivatives catalysed by the *Solanum tuberosum* epoxide hydrolase. 15(18), 2801-2805.
- Morisseau, C., Nellaiah, H., Archelas, A., Furstoss, R., Baratti, J. C. J. E., & technology, m. (1997). Asymmetric hydrolysis of racemic para-nitrostyrene oxide using an epoxide hydrolase preparation from *Aspergillus niger*. 20(6), 446-452.
- Munnecke, D. M., & Hsieh, D. P. (1974). Microbial decontamination of parathion and p-nitrophenol in aqueous media. *Appl Microbiol*, 28(2), 212-217. doi:10.1128/am.28.2.212-217.1974
- Muraki, T., Taki, M., Hasegawa, Y., Iwaki, H., & Lau, P. C. (2003). Prokaryotic homologs of the eukaryotic 3-hydroxyanthranilate 3,4-dioxygenase and 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase in the 2-nitrobenzoate degradation pathway of *Pseudomonas fluorescens* strain KU-7. *Appl Environ Microbiol*, 69(3), 1564-1572. doi:10.1128/AEM.69.3.1564-1572.2003
- Nagata, Y., Endo, R., Ito, M., Ohtsubo, Y., & Tsuda, M. (2007). Aerobic degradation of lindane (gamma-hexachlorocyclohexane) in bacteria and its biochemical and molecular basis. *Appl Microbiol Biotechnol*, 76(4), 741-752. doi:10.1007/s00253-007-1066-x
- Nagata, Y., Futamura, A., Miyauchi, K., & Takagi, M. (1999). Two different types of dehalogenases, LinA and LinB, involved in gamma-hexachlorocyclohexane degradation in *Sphingomonas paucimobilis* UT26 are localized in the periplasmic space without molecular processing. *J Bacteriol*, 181(17), 5409-5413. doi:10.1128/JB.181.17.5409-5413.1999
- Nagata, Y., Miyauchi, K., & Takagi, M. (1999). Complete analysis of genes and enzymes for gamma-hexachlorocyclohexane degradation in *Sphingomonas paucimobilis* UT26. *J Ind Microbiol Biotechnol*, 23(4-5), 380-390. doi:10.1038/sj.jim.2900736
- Nagata, Y., Mori, K., Takagi, M., Murzin, A. G., & Damborsky, J. (2001). Identification of protein fold and catalytic residues of gamma-hexachlorocyclohexane dehydrochlorinase LinA. *Proteins*, 45(4), 471-477. doi:10.1002/prot.10007
- Nagata, Y., Ohtsubo, Y., & Tsuda, M. (2015). Properties and biotechnological applications of natural and engineered haloalkane dehalogenases. *Appl Microbiol Biotechnol*, 99(23), 9865-9881. doi:10.1007/s00253-015-6954-x
- Nam, H., Lewis, N. E., Lerman, J. A., Lee, D. H., Chang, R. L., Kim, D., & Palsson, B. O. (2012). Network context and selection in the evolution to enzyme specificity. *Science*, 337(6098), 1101-1104. doi:10.1126/science.1216861
- Nardini, M., & Dijkstra, B. W. (1999). Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol*, 9(6), 732-737. doi:10.1016/s0959-440x(99)00037-8
- Nardini, M., Ridder, I. S., Rozeboom, H. J., Kalk, K. H., Rink, R., Janssen, D. B., & Dijkstra, B. W. (1999). The x-ray structure of epoxide hydrolase from *Agrobacterium radiobacter* AD1. An enzyme to detoxify harmful epoxides. *J Biol Chem*, 274(21), 14579-14586.
- Naughton, S. X., & Terry, A. V., Jr. (2018). Neurotoxicity in acute and repeated organophosphate exposure. *Toxicology*, 408, 101-112. doi:10.1016/j.tox.2018.08.011
- Nesvera, J., Rucka, L., & Patek, M. (2015). Catabolism of Phenol and Its Derivatives in Bacteria: Genes, Their Regulation, and Use in the Biodegradation of Toxic Pollutants. *Adv Appl Microbiol*, 93, 107-160. doi:10.1016/bs.aambs.2015.06.002

- Ney, B., Ahmed, F. H., Carere, C. R., Biswas, A., Warden, A. C., Morales, S. E., . . . Greening, C. (2017). The methanogenic redox cofactor F420 is widely synthesized by aerobic soil bacteria. *ISME J*, 11(1), 125-137. doi:10.1038/ismej.2016.100
- Nies, D. H. (2003). Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol Rev*, 27(2-3), 313-339. doi:10.1016/S0168-6445(03)00048-2
- Noor, S., Taylor, M. C., Russell, R. J., Jermin, L. S., Jackson, C. J., Oakeshott, J. G., & Scott, C. (2012). Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS One*, 7(6), e39822. doi:10.1371/journal.pone.0039822
- Oelschlagel, M., Ruckert, C., Kalinowski, J., Schmidt, G., Schlomann, M., & Tischler, D. (2015). *Sphingopyxis fribergensis* sp. nov., a soil bacterium with the ability to degrade styrene and phenylacetic acid. *Int J Syst Evol Microbiol*, 65(9), 3008-3015. doi:10.1099/ij.s.0.000371
- Oelschlagel, M., Zimmerling, J., Schlomann, M., & Tischler, D. (2014). Styrene oxide isomerase of *Sphingopyxis* sp. Kp5.2. *Microbiology (Reading)*, 160(Pt 11), 2481-2491. doi:10.1099/mic.0.080259-0
- Okai, M., Kubota, K., Fukuda, M., Nagata, Y., Nagata, K., & Tanokura, M. (2010). Crystal structure of gamma-hexachlorocyclohexane Dehydrochlorinase LinA from *Sphingobium japonicum* UT26. *J Mol Biol*, 403(2), 260-269. doi:10.1016/j.jmb.2010.08.043
- Olsen, R. H., Kukor, J. J., & Kaphammer, B. (1994). A novel toluene-3-monooxygenase pathway cloned from *Pseudomonas pickettii* PKO1. *J Bacteriol*, 176(12), 3749-3756. doi:10.1128/jb.176.12.3749-3756.1994
- Orru, R. V., & Faber, K. (1999). Stereoselectivities of microbial epoxide hydrolases. *Curr Opin Chem Biol*, 3(1), 16-21. doi:10.1016/s1367-5931(99)80004-0
- Padhi, S. K., Fujii, R., Legatt, G. A., Fossum, S. L., Berchtold, R., & Kazlauskas, R. J. (2010). Switching from an esterase to a hydroxynitrile lyase mechanism requires only two amino acid substitutions. *Chem Biol*, 17(8), 863-871. doi:10.1016/j.chembiol.2010.06.013
- Pandeeti, E. V., Chakka, D., Pandey, J. P., & Siddavattam, D. (2011). Indigenous organophosphate-degrading (opd) plasmid pCMS1 of *Brevundimonas diminuta* is self-transmissible and plays a key role in horizontal mobility of the opd gene. *Plasmid*, 65(3), 226-231. doi:10.1016/j.plasmid.2011.02.003
- Pansegrau, W., Schoumacher, F., Hohn, B., & Lanka, E. (1993). Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *Proc Natl Acad Sci U S A*, 90(24), 11538-11542. doi:10.1073/pnas.90.24.11538
- Parapatla, H., Gudla, R., Konduru, G. V., Devadasu, E. R., Nagarajaram, H. A., Sritharan, M., . . . Siddavattam, D. (2020). Organophosphate hydrolase interacts with ferric-enterobactin and promotes iron uptake in association with TonB-dependent transport system. *Biochem J*, 477(15), 2821-2840. doi:10.1042/BCJ20200299
- Parthasarathy, S., Azam, S., Lakshman Sagar, A., Narasimha Rao, V., Gudla, R., Parapatla, H., . . . Siddavattam, D. (2017). Genome-Guided Insights Reveal Organophosphate-Degrading *Brevundimonas diminuta* as *Sphingopyxis wildii* and Define Its Versatile Metabolic Capabilities and Environmental Adaptations. *Genome Biol Evol*, 9(1), 77-81. doi:10.1093/gbe/evw275
- Parthasarathy, S., Parapatla, H., Nandavaram, A., Palmer, T., & Siddavattam, D. (2016). Organophosphate Hydrolase Is a Lipoprotein and Interacts with Pi-specific Transport System to Facilitate Growth of *Brevundimonas diminuta* Using OP Insecticide as Source of Phosphate. *J Biol Chem*, 291(14), 7774-7785. doi:10.1074/jbc.M116.715110
- Parthasarathy, S., Parapatla, H., & Siddavattam, D. (2017). Topological analysis of the lipoprotein organophosphate hydrolase from *Sphingopyxis wildii* reveals a periplasmic localisation. *FEMS Microbiol Lett*, 364(19). doi:10.1093/femsle/fnx187
- Pati, A., Ivanova, N. N., Mikhailova, N., Ovchinnikova, G., Hooper, S. D., Lykidis, A., & Kyrpides, N. C. (2010). GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods*, 7(6), 455-457. doi:10.1038/nmeth.1457
- Pearce, S. L., Oakeshott, J. G., & Pandey, G. (2015). Insights into Ongoing Evolution of the Hexachlorocyclohexane Catabolic Pathway from Comparative Genomics of Ten *Sphingomonadaceae* Strains. *G3 (Bethesda)*, 5(6), 1081-1094. doi:10.1534/g3.114.015933
- Pearson, W. R. (2013). An introduction to sequence similarity ("homology") searching. *Curr Protoc Bioinformatics, Chapter 3*, Unit3 1. doi:10.1002/0471250953.bi0301s42
- Pedragosa-Moreau, S., Archelas, A., & Furstoss, R. (1996). Microbiological transformations. 31: Synthesis of enantiopure epoxides and vicinal diols using fungal epoxide hydrolase mediated hydrolysis. *Tetrahedron Letters*, 37(19), 3319-3322. doi:https://doi.org/10.1016/0040-4039(96)00540-0

- Perez-Pantoja, D., De la Iglesia, R., Pieper, D. H., & Gonzalez, B. (2008). Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium *Cupriavidus necator* JMP134. *FEMS Microbiol Rev*, 32(5), 736-794. doi:10.1111/j.1574-6976.2008.00122.x
- Perez-Pantoja, D., Donoso, R., Agullo, L., Cordova, M., Seeger, M., Pieper, D. H., & Gonzalez, B. (2012). Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales. *Environ Microbiol*, 14(5), 1091-1117. doi:10.1111/j.1462-2920.2011.02613.x
- Pikkemaat, M. G., & Janssen, D. B. (2002). Generating segmental mutations in haloalkane dehalogenase: a novel part in the directed evolution toolbox. *Nucleic Acids Res*, 30(8), E35-35. doi:10.1093/nar/30.8.e35
- Prokop, Z., Oplustil, F., DeFrank, J., & Damborsky, J. (2006). Enzymes fight chemical weapons. *Biotechnol J*, 1(12), 1370-1380. doi:10.1002/biot.200600166
- Raju, N. J. (2022). Arsenic in the geo-environment: A review of sources, geochemical processes, toxicity and removal technologies. *Environ Res*, 203, 111782. doi:10.1016/j.envres.2021.111782
- Rui, L., Cao, L., Chen, W., Reardon, K. F., & Wood, T. K. (2005). Protein engineering of epoxide hydrolase from *Agrobacterium radiobacter* AD1 for enhanced activity and enantioselective production of (R)-1-phenylethane-1,2-diol. *Appl Environ Microbiol*, 71(7), 3995-4003. doi:10.1128/AEM.71.7.3995-4003.2005
- Russell, R. J., Scott, C., Jackson, C. J., Pandey, R., Pandey, G., Taylor, M. C., . . . Oakeshott, J. G. (2011). The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. *Evol Appl*, 4(2), 225-248. doi:10.1111/j.1752-4571.2010.00175.x
- Sahasrabudhe, S. R., & Modi, V. V. (1985). Hydroxylation of benzoate and its chlorinated derivatives in *Aspergillus niger*. *Biochem Int*, 10(4), 525-529.
- Saier, M. H., Jr., & Paulsen, I. T. (2001). Phylogeny of multidrug transporters. *Semin Cell Dev Biol*, 12(3), 205-213. doi:10.1006/scdb.2000.0246
- Salzberg, S. L., Delcher, A. L., Kasif, S., & White, O. (1998). Microbial gene identification using interpolated Markov models. *Nucleic Acids Res*, 26(2), 544-548. doi:10.1093/nar/26.2.544
- Samantarrai, D., Lakshman Sagar, A., Gudla, R., & Siddavattam, D. (2020). TonB-Dependent Transporters in Sphingomonads: Unraveling Their Distribution and Function in Environmental Adaptation. *Microorganisms*, 8(3). doi:10.3390/microorganisms8030359
- Sanchez-Santed, F., Colomina, M. T., & Herrero Hernandez, E. (2016). Organophosphate pesticide exposure and neurodegeneration. *Cortex*, 74, 417-426. doi:10.1016/j.cortex.2015.10.003
- Sawyer, S. (1989). Statistical tests for detecting gene conversion. *Mol Biol Evol*, 6(5), 526-538. doi:10.1093/oxfordjournals.molbev.a040567
- Schmitz, L. M., Rosenthal, K., & Lutz, S. (2019). Recent advances in heme biocatalysis engineering. *Biotechnol Bioeng*, 116(12), 3469-3475. doi:10.1002/bit.27156
- Schober, M., & Faber, K. (2013). Inverting hydrolases and their use in enantioconvergent biotransformations. *Trends Biotechnol*, 31(8), 468-478. doi:10.1016/j.tibtech.2013.05.005
- Schuiten, E. D., Badenhorst, C. P. S., Palm, G. J., Berndt, L., Lammers, M., Mican, J., . . . Bornscheuer, U. T. (2021). Promiscuous Dehalogenase Activity of the Epoxide Hydrolase CorEH from *Corynebacterium* sp. C12. *ACS Catalysis*, 11(10), 6113-6120. doi:10.1021/acscatal.1c00851
- Schwarz, J., Rosenthal, K., Snajdrova, R., Kittelmann, M., & Lutz, S. (2020). The Development of Biocatalysis as a Tool for Drug Discovery. *Chimia (Aarau)*, 74(5), 368-377. doi:10.2533/chimia.2020.368
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068-2069. doi:10.1093/bioinformatics/btu153
- Seffernick, J. L., de Souza, M. L., Sadowsky, M. J., & Wackett, L. P. (2001). Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *J Bacteriol*, 183(8), 2405-2410. doi:10.1128/JB.183.8.2405-2410.2001
- Seffernick, J. L., & Wackett, L. P. (2001). Rapid evolution of bacterial catabolic enzymes: a case study with atrazine chlorohydrolase. *Biochemistry*, 40(43), 12747-12753. doi:10.1021/bi011293r
- Segers, P., Vancanneyt, M., Pot, B., Torck, U., Hoste, B., Dewettinck, D., . . . De Vos, P. (1994). Classification of *Pseudomonas diminuta* Leifson and Hugh 1954 and *Pseudomonas vesicularis* Busing, Doll, and Freytag 1953 in *Brevundimonas* gen. nov. as *Brevundimonas diminuta* comb. nov. and *Brevundimonas vesicularis* comb. nov., respectively. *Int J Syst Bacteriol*, 44(3), 499-510. doi:10.1099/00207713-44-3-499

- Serdar, C. M., Gibson, D. T., Munnecke, D. M., & Lancaster, J. H. (1982). Plasmid Involvement in Parathion Hydrolysis by *Pseudomonas diminuta*. *Appl Environ Microbiol*, 44(1), 246-249. doi:10.1128/aem.44.1.246-249.1982
- Shapir, N., Mongodin, E. F., Sadowsky, M. J., Daugherty, S. C., Nelson, K. E., & Wackett, L. P. (2007). Evolution of catabolic pathways: Genomic insights into microbial s-triazine metabolism. *J Bacteriol*, 189(3), 674-682. doi:10.1128/JB.01257-06
- Shaw, J. P., & Harayama, S. (1992). Purification and characterisation of the NADH:acceptor reductase component of xylene monooxygenase encoded by the TOL plasmid pWW0 of *Pseudomonas putida* mt-2. *Eur J Biochem*, 209(1), 51-61. doi:10.1111/j.1432-1033.1992.tb17260.x
- Shields, M. S., Reagin, M. J., Gerger, R. R., Campbell, R., & Somerville, C. (1995). TOM, a new aromatic degradative plasmid from *Burkholderia* (*Pseudomonas*) *cepacia* G4. *Appl Environ Microbiol*, 61(4), 1352-1356. doi:10.1128/aem.61.4.1352-1356.1995
- Siegel, D., Hui, H. C., Doerffler, E., Clarke, M. O., Chun, K., Zhang, L., . . . Mackman, R. L. (2017). Discovery and Synthesis of a Phosphoramidate Prodrug of a Pyrrolo[2,1-f][triazin-4-amino] Adenine C-Nucleoside (GS-5734) for the Treatment of Ebola and Emerging Viruses. *J Med Chem*, 60(5), 1648-1661. doi:10.1021/acs.jmedchem.6b01594
- Stover, C. K., Warrenner, P., VanDevanter, D. R., Sherman, D. R., Arain, T. M., Langhorne, M. H., . . . Baker, W. R. (2000). A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature*, 405(6789), 962-966. doi:10.1038/35016103
- Stucki, G., & Thueer, M. (1995). Experiences of a large-scale application of 1,2-dichloroethane degrading microorganisms for groundwater treatment. *Environ Sci Technol*, 29(9), 2339-2345. doi:10.1021/es00009a028
- Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*, 189(1), 113-130. doi:10.1016/0022-2836(86)90385-2
- Sun, J., Deng, Z., & Yan, A. (2014). Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. *Biochem Biophys Res Commun*, 453(2), 254-267. doi:10.1016/j.bbrc.2014.05.090
- Supreeth, M., & Raju, N. S. (2017). Biotransformation of chlorpyrifos and endosulfan by bacteria and fungi. *Appl Microbiol Biotechnol*, 101(15), 5961-5971. doi:10.1007/s00253-017-8401-7
- Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*, 34(Web Server issue), W609-612. doi:10.1093/nar/gkl315
- Svedendahl, M., Carlqvist, P., Branneby, C., Allner, O., Frise, A., Hult, K., . . . Brinck, T. (2008). Direct epoxidation in *Candida antarctica* lipase B studied by experiment and theory. *ChemBiochem*, 9(15), 2443-2451. doi:10.1002/cbic.200800318
- Szymanski, W., Westerbeek, A., Janssen, D. B., & Feringa, B. L. (2011). A simple enantioconvergent and chemoenzymatic synthesis of optically active α -substituted amides. *Angew Chem Int Ed Engl*, 50(45), 10712-10715. doi:10.1002/anie.201105164
- Takeuchi, M., Hamana, K., & Hiraishi, A. (2001). Proposal of the genus *Sphingomonas* sensu stricto and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analyses. *Int J Syst Evol Microbiol*, 51(Pt 4), 1405-1417. doi:10.1099/00207713-51-4-1405
- Taylor, M. C., Jackson, C. J., Tattersall, D. B., French, N., Peat, T. S., Newman, J., . . . Oakeshott, J. G. (2010). Identification and characterization of two families of F420 H₂-dependent reductases from *Mycobacteria* that catalyse aflatoxin degradation. *Mol Microbiol*, 78(3), 561-575. doi:10.1111/j.1365-2958.2010.07356.x
- Teelucksingh, T., Thompson, L. K., & Cox, G. (2020). The Evolutionary Conservation of *Escherichia coli* Drug Efflux Pumps Supports Physiological Functions. *J Bacteriol*, 202(22). doi:10.1128/JB.00367-20
- Topp, E., Xun, L. Y., & Orser, C. S. (1992). Biodegradation of the herbicide bromoxynil (3,5-dibromo-4-hydroxybenzonitrile) by purified pentachlorophenol hydroxylase and whole cells of *Flavobacterium* sp. strain ATCC 39723 is accompanied by cyanogenesis. *Appl Environ Microbiol*, 58(2), 502-506. doi:10.1128/aem.58.2.502-506.1992
- Tsai, P. C., Bigley, A., Li, Y., Ghanem, E., Cadieux, C. L., Kasten, S. A., . . . Raushel, F. M. (2010). Stereoselective hydrolysis of organophosphate nerve agents by the bacterial phosphotriesterase. *Biochemistry*, 49(37), 7978-7987. doi:10.1021/bi101056m

- Tsai, P. C., Fox, N., Bigley, A. N., Harvey, S. P., Barondeau, D. P., & Raushel, F. M. (2012). Enzymes for the homeland defense: optimizing phosphotriesterase for the hydrolysis of organophosphate nerve agents. *Biochemistry*, 51(32), 6463-6475. doi:10.1021/bi300811t
- Udikovic-Kolic, N., Scott, C., & Martin-Laurent, F. (2012). Evolution of atrazine-degrading capabilities in the environment. *Appl Microbiol Biotechnol*, 96(5), 1175-1189. doi:10.1007/s00253-012-4495-0
- Ueberbacher, B., Osprian, I., Mayer, S., & Faber, K. (2005). A Chemoenzymatic, Enantioconvergent, Asymmetric Total Synthesis of(R)-Fridamycin E. *European Journal of Organic Chemistry*, 2005, 1266-1270. doi:10.1002/ejoc.200400720
- Vargas, D. F., Larghi, E. L., & Kaufman, T. S. (2021). Evolution of the Synthesis of Remdesivir. Classical Approaches and Most Recent Advances. *ACS Omega*, 6(30), 19356-19363. doi:10.1021/acsomega.1c03082
- Vasu, K., & Nagaraja, V. (2013). Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol Mol Biol Rev*, 77(1), 53-72. doi:10.1128/MMBR.00044-12
- Verma, H., Dhingra, G. G., Sharma, M., Gupta, V., Negi, R. K., Singh, Y., & Lal, R. (2020). Comparative genomics of *Sphingopyxis* spp. unravelled functional attributes. *Genomics*, 112(2), 1956-1969. doi:10.1016/j.ygeno.2019.11.008
- Visser, H., de Bont, J. A., & Verdoes, J. C. (1999). Isolation and characterization of the epoxide hydrolase-encoding gene from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol*, 65(12), 5459-5463. doi:10.1128/AEM.65.12.5459-5463.1999
- Vyas, N. K., Nickitenko, A., Rastogi, V. K., Shah, S. S., & Quijcho, F. A. (2010). Structural insights into the dual activities of the nerve agent degrading organophosphate anhydrolase/prolidase. *Biochemistry*, 49(3), 547-559. doi:10.1021/bi9011989
- Wackett, L. P. (2009). Questioning our perceptions about evolution of biodegradative enzymes. *Curr Opin Microbiol*, 12(3), 244-251. doi:10.1016/j.mib.2009.05.001
- Wackett, L. P., Kwart, L. D., & Gibson, D. T. (1988). Benzylic monooxygenation catalyzed by toluene dioxygenase from *Pseudomonas putida*. *Biochemistry*, 27(4), 1360-1367. doi:10.1021/bi00404a041
- Wackett, L. P., & Robinson, S. L. (2020). The ever-expanding limits of enzyme catalysis and biodegradation: polyaromatic, polychlorinated, polyfluorinated, and polymeric compounds. *Biochem J*, 477(15), 2875-2891. doi:10.1042/BCJ20190720
- Wahler, D., & Reymond, J. L. (2002). The adrenaline test for enzymes. *Angew Chem Int Ed Engl*, 41(7), 1229-1232. doi:10.1002/1521-3773(20020402)41:7<1229::aid-anie1229>3.0.co;2-5
- Walker, C. H. (1993). The classification of esterases which hydrolyse organophosphates: recent developments. *Chem Biol Interact*, 87(1-3), 17-24. doi:10.1016/0009-2797(93)90021-p
- Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., . . . Sobral, B. W. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res*, 42(Database issue), D581-591. doi:10.1093/nar/gkt1099
- Wozniak, R. A., & Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol*, 8(8), 552-563. doi:10.1038/nrmicro2382
- Wu, S., Li, A., Chin, Y. S., & Li, Z. J. A. C. (2013). Enantioselective hydrolysis of racemic and meso-epoxides with recombinant *Escherichia coli* expressing epoxide hydrolase from *Sphingomonas* sp. HXN-200: preparation of epoxides and vicinal diols in high ee and high concentration. 3(4), 752-759.
- Yang, F., Feng, H., Massey, I. Y., Huang, F., Guo, J., & Zhang, X. (2020). Genome-Wide Analysis Reveals Genetic Potential for Aromatic Compounds Biodegradation of *Sphingopyxis*. *Biomed Res Int*, 2020, 5849123. doi:10.1155/2020/5849123
- Yen, K. M., Karl, M. R., Blatt, L. M., Simon, M. J., Winter, R. B., Fausset, P. R., . . . Chen, K. K. (1991). Cloning and characterization of a *Pseudomonas mendocina* KR1 gene cluster encoding toluene-4-monooxygenase. *J Bacteriol*, 173(17), 5315-5327. doi:10.1128/jb.173.17.5315-5327.1991
- Zhang, D. C., Liu, H. C., Xin, Y. H., Zhou, Y. G., Schinner, F., & Margesin, R. (2010). *Sphingopyxis bauzanensis* sp. nov., a psychrophilic bacterium isolated from soil. *Int J Syst Evol Microbiol*, 60(Pt 11), 2618-2622. doi:10.1099/ijs.0.018218-0
- Zhang, W. M., Zhang, J. J., Jiang, X., Chao, H., & Zhou, N. Y. (2015). Transcriptional activation of multiple operons involved in para-nitrophenol degradation by *Pseudomonas* sp. Strain WBC-3. *Appl Environ Microbiol*, 81(1), 220-230. doi:10.1128/AEM.02720-14
- Zuckerkindl, E., & Pauling, L. (1965). Evolutionary Divergence and Convergence in Proteins. In V. Bryson & H. J. Vogel (Eds.), *Evolving Genes and Proteins* (pp. 97-166): Academic Press.

Aparna Nandavaram, Annapoorni Lakshman Sagar, Ashok Kumar Madikonda and Dayananda Siddavattam*

Proteomics of *Sphingobium indicum* B90A for a deeper understanding of hexachlorocyclohexane (HCH) bioremediation

DOI 10.1515/reveh-2015-0042

Received October 13, 2015; accepted October 13, 2015; previously published online March 8, 2016

Abstract: Genome wide expression profiling of *Sphingobium indicum* B90A revealed induction of *lin* genes, *linA* and *linB*, involved in dechlorination of hexachlorocyclohexane (HCH), in the presence of all four isomers of HCH. Supporting proteomics data, the qPCR and promoter assay showed upregulation of *linA* transcription in the presence of HCH isomers. Analysis of the upstream region of the *linA* gene revealed the existence of the GntR binding site overlapping the -10 hexamer of the putative promoter motif. As GntR is a known transcription repressor its dissociation from the *linA* promoter is expected to induce *lin* genes in the presence of HCH isomers. Comparison of *in situ* and in-culture proteomics indicated expression *lin* genes at the dumpsite, an indication for the *in situ* HCH degradation.

Keywords: biodegradation; bioremediation; hexachlorocyclohexane (HCH); *lin* genes; proteomics.

Introduction

Hexachlorocyclohexane (HCH), an organochlorine compound, has extensively been used as an insecticide to control important agricultural pests. The commercial formulation of HCH primarily consists of α , β , γ , δ isomers of which the γ isomer (lindane) has potent insecticidal activity. One ton purification of lindane from technical HCH produces 10 tons of HCH muck consisting of α , β , γ and δ HCH. During the past 60 years around 60,000 tons of

lindane has been used worldwide, leading to the formation of 4–7 million tons of HCH muck that is scattered in different places around the globe. Regardless of the stereochemistry, stability and persistence of HCH isomers they have been shown to cause adverse effects to the environment and human health (1). Considering its influence on environment and human health it has been declared as one of the key global pollutants.

The persistence of HCH isomers in the environment is chiefly due to the absence of microbes that can degrade/use them as sole source of carbon. A number of attempts have been made to evaluate microbial degradation of HCH. However, very few reports are available on microbial degradation of HCH both under aerobic and anaerobic conditions. Detailed investigations revealed the existence of upper and lower degradation pathway enzymes in HCH degrading *Sphingobium japonicum* UT26S (2) and *Sphingobium indicum* B90A (3). The upper pathway enzymes, primarily the dehalogenases, contribute to the dehalogenation of HCH. The lower pathway enzymes, the dioxygenases, convert the dehalogenated HCH into the tricarboxylic acid cycle intermediates. HCH dehydrochlorinase (LinA), haloalkane dehalogenase (LinB) and dehydrogenase (LinC/LinX) are upper pathway enzymes. The reductive dechlorinase (LinD), ring cleavage oxygenase (LinE), maleylacetate reductase (LinF), an acyl-CoA transferase (LinG, H), a thiolase (LinJ) contribute to the lower degradation pathway. The HCH degrading (*lin*) genes are organized as clusters. Along with *linI* and *linR*, the genes that code for regulatory proteins, the *linK*, *linL*, *linM* and *linN* coding for a putative ABC-type transporter form a cluster that resembles a genomic island (3). The existence of transposase and recombinase coding genes flanking the *lin* cluster suggests extensive recombination and horizontal transfer of *lin* genes among soil bacteria.

The meta-genome sequence extracted from the soil samples collected from the dumpsite revealed the coexistence of several HCH degraders and non-degraders (4). Interestingly, most of the *Sphingomonads* found enriched at the HCH dumpsites, have shown existence

*Corresponding author: Dayananda Siddavattam, Department of Animal Biology, School of Life Sciences, Room No- S40, University of Hyderabad, Hyderabad, Telangana 500046, India, E-mail: siddavattam@gmail.com

Aparna Nandavaram, Annapoorni Lakshman Sagar and Ashok Kumar Madikonda: Department of Animal Sciences, School of Life Sciences, University of Hyderabad, Hyderabad, India

Genome-Guided Insights Reveal Organophosphate-Degrading *Brevundimonas diminuta* as *Sphingopyxis wildii* and Define Its Versatile Metabolic Capabilities and Environmental Adaptations

Sunil Parthasarathy^{1,†}, Sarwar Azam^{2,†}, Annapoorni Lakshman Sagar¹, Veera Narasimha Rao², Ramurthy Gudla¹, Hari Parapatla¹, Harshita Yakkala¹, Sujana Ghanta Vemuri¹, and Dayananda Siddavattam^{1,*}

¹Department of Animal Biology, School of Life Sciences, University of Hyderabad, Hyderabad, India

²National Institute of Animal Biotechnology, Hyderabad, India

†These authors contributed equally to this work.

*Corresponding authors: E-mails: sds1@uohyd.ernet.in; siddavattam@gmail.com.

Accepted: November 5, 2016

Abstract

The complete genome sequence of *Brevundimonas diminuta* represented a chromosome (~4.15 Mb) and two plasmids (pCMS1 and pCMS2) with sizes of 65,908 and 30,654 bp, respectively. The sequence of the genome showed no significant similarity with the known bacterial genome sequences, instead showed weak similarity with the members of different genera of family, Sphingomonadaceae. Contradicting existing taxonomic position, the core genome-guided phylogenetic tree placed *B. diminuta* in the genus *Sphingopyxis* and showed sufficient genome-to-genome distance warranting a new species name. Reflecting the strains ability to grow in harsh environments, the genome-contained genetic repertoire required for mineralization of several recalcitrant man-made aromatic compounds.

Key words: biodegradation, aromatic compound degradation, biotransformation, Sphingomonadales.

Introduction

The safe disposal of neurotoxic organophosphate (OP) residues has attracted the attention of several microbiologists. Bacterial strains possessing organophosphate hydrolase (OPH) activity have been isolated from sewage and soil samples (Munnecke and Hsieh 1974). Using conventional taxonomic tools, the isolated OP-degrading bacterial strains have been placed in the genus, *Pseudomonas* (Munnecke and Hsieh 1974). However, when the genus *Pseudomonas* was reclassified, bacterial strains that were previously named as *Pseudomonas diminuta* and *Pseudomonas vesicularis* were moved to a separate genus known as *Brevundimonas*. Accordingly, the *P. diminuta* MG was renamed *Brevundimonas diminuta* MG (Segers et al. 1994). In this study, we generated the complete genome sequence of *B. diminuta* MG and report several interesting features pertaining to its taxonomy, evolution, and degradation potential.

Results and Discussion

A total of 6.8 Gbp data was generated to assemble the complete genome of *B. diminuta*. The scaffold level assembly contained 28 scaffolds with N50 of 1,786,567 bp. These scaffolds were further merged to obtain three super scaffolds using data generated from a mate-pair library. The super scaffolds were circularized by manually generating sequence for the DNA, amplified using primers specific to the right and left flanks of the assembled scaffold sequences. The largest super scaffold with a length of 4,147,822 bp is regarded as the chromosome sequence of *B. diminuta* (fig. 1A). The other two scaffolds gave circular DNA molecules with a size of 65,908 bp (fig. 1B) and 30,654 bp (fig. 1C), respectively. The 65,908 bp circular sequence matched perfectly with the physical map and partial sequence of pCMS1, a previously reported indigenous plasmid of *B. diminuta* (Pandeeti et al. 2011). However, the second circular DNA showed no complete similarity to any other



Review

TonB-Dependent Transporters in Sphingomonads: Unraveling Their Distribution and Function in Environmental Adaptation

Devyani Samantarrai [†], Annapoorni Lakshman Sagar [†], Ramurthy Gudla and Dayananda Siddavattam ^{*}

Department of Animal Biology, School of Life Sciences, University of Hyderabad, Hyderabad 500 046, India; devyani66@gmail.com (D.S.); annapoornilakshmansagar@gmail.com (A.L.S.); grmurthygp120@gmail.com (R.G.)

^{*} Correspondence: siddavattam@gmail.com; Tel.: +91-040-66794578

[†] These authors contributed equally to this work.

Received: 30 November 2019; Accepted: 7 January 2020; Published: 3 March 2020



Abstract: TonB-dependent transport system plays a critical role in the transport of nutrients across the energy-deprived outer membrane of Gram-negative bacteria. It contains a specialized outer membrane TonB-dependent transporter (TBDT) and energy generating (ExbB/ExbD) and transducing (TonB) inner membrane multi-protein complex, called TonB complex. Very few TonB complex protein-coding sequences exist in the genomes of Gram-negative bacteria. Interestingly, the TBDT coding alleles are phenomenally high, especially in the genomes of bacteria surviving in complex and stressful environments. Sphingomonads are known to survive in highly polluted environments using rare, recalcitrant, and toxic substances as their sole source of carbon. Naturally, they also contain a huge number of TBDTs in the outer membrane. Out of them, only a few align with the well-characterized TBDTs. The functions of the remaining TBDTs are not known. Predictions made based on genome context and expression pattern suggest their involvement in the transport of xenobiotic compounds across the outer membrane.

Keywords: TonB-dependent transporter (TBDT); sphingomonads; xenobiotics

1. Introduction

The outer membrane of Gram-negative bacteria performs several important functions. It acts as a barrier to prevent the entry of antibiotics and other toxic chemicals and protects the cell wall by denying access to cell wall degrading enzymes. However, existence of an energy-deprived outer membrane is a hurdle for the uptake of nutrients in Gram-negative bacteria [1]. A majority of nutrients gains entry into periplasmic space by diffusing through the outer membrane via a pore-like structure formed in outer membrane-associated β -barrel containing proteins, otherwise known as porins [2]. However, certain scarcely available nutrients depend on active transport to cross the energy-deprived outer membrane. The active transport mechanism of the outer membrane is known as TonB-dependent transport system. The system contains two components, the inner membrane-associated TonB complex and an outer membrane-associated TonB-dependent transporter (TBDT). The TonB complex contains proton motive force (PMF) components, ExbB/ExbD, and energy transducer TonB in a ratio of 7:2:1 [3]. The TonB complex has unique role in outer membrane transport. The PMF components ExbB/ExbD generate energy by pumping protons across the inner membrane, while TonB transduces this energy to the outer membrane-localized TBDT. The TonB protein contains three domains: the N-terminal transmembrane helix, C-terminal domain, and a proline-rich rigid central domain. The N-terminal region is embedded in the inner membrane and is associated with one of the transmembrane domains of ExbB [4].



RAJIV GANDHI CENTRE FOR BIOTECHNOLOGY

Computational Biology Group Bio Innovation Centre

Thiruvananthapuram



Certificate

Annapoorni Lakshman Sagar

This is to certify that.....

participated in the workshop on "Exploring Gene Expression data using Transcriptome and Microarrays" held at the Computational Biology Facility, Bio-Innovation Centre, Rajiv Gandhi Centre for Biotechnology (Government of India), April 18 - 20, 2016.

Professor M. Radhakrishna Pillai
Director, Rajiv Gandhi Centre for Biotechnology
Thiruvananthapuram

Mining of *Brevundimonas diminuta* MG genome to identify haloalkane dehalogenase with promiscuous epoxide hydrolase activity and its significance in green synthesis of bioactive molecules

Submission date: 26-Apr-2022 05:26PM (UTC+0530)
by Annapoorni Lakshman Sagar

Submission ID: 1820808178

File name: Annapoorni_Lakshman_Sagar.pdf (924.35K)

Word count: 31984

Character count: 173075

Mining of Brevundimonas diminuta MG genome to identify haloalkane dehalogenase with promiscuous epoxide hydrolase activity and its significance in green synthesis of bioactive molecules

ORIGINALITY REPORT

14%

SIMILARITY INDEX

8%

INTERNET SOURCES

7%

PUBLICATIONS

9%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to University of Hyderabad,
Hyderabad

Student Paper

7%

S. S. Suresh

2

www.ncbi.nlm.nih.gov

Internet Source

Prof. S. DAYANAND
Dept. of Animal Biology
School of Life Sciences
University of Hyderabad
Hyderabad-500 046, T.

1%

3

tohoku.repo.nii.ac.jp

Internet Source

<1%

4

link.springer.com

Internet Source

<1%

5

Submitted to Mahidol University

Student Paper

<1%

6

curis.ku.dk

Internet Source

<1%

7

go.gale.com

Internet Source

<1%

8

hdl.handle.net

Internet Source

<1%

9

Deviprasanna Chakka, Ramurthy Gudla, Ashok Kumar Madikonda, Emmanuel Vijay Paul Pandeeti et al. " The Organophosphate Degradation () Island-borne Esterase-induced Metabolic Diversion in and Its Influence on - Nitrophenol Degradation ", Journal of Biological Chemistry, 2015

Publication

<1 %

10

baadalsg.inflibnet.ac.in

Internet Source

<1 %

11

edepot.wur.nl

Internet Source

<1 %

12

Radka Chaloupkova, Veronika Liskova, Martin Toul, Klara Markova et al. "Light-Emitting Dehalogenases: Reconstruction of Multifunctional Biocatalysts", ACS Catalysis, 2019

Publication

<1 %

13

"Vaccine Design", Springer Science and Business Media LLC, 2016

Publication

<1 %

14

J. Koca. "Analysis of the reaction mechanism and substrate specificity of haloalkane dehalogenases by sequential and structural comparisons", Protein Engineering Design and Selection, 11/01/1999

Publication

<1 %

- | | | |
|-------|--|--------|
| 15 | Yuji Nagata, Katsuki Mori, Masamichi Takagi, Alexey G. Murzin, Jiří Damborský.
"Identification of protein fold and catalytic residues of γ -hexachlorocyclohexane dehydrochlorinase LinA", Proteins: Structure, Function, and Bioinformatics, 2001
<small>Publication</small> | $<1\%$ |
| <hr/> | | |
| 16 | Eva D. Schuiten, Christoffel P. S. Badenhorst, Gottfried J. Palm, Leona Berndt et al. "
Promiscuous Dehalogenase Activity of the Epoxide Hydrolase CorEH from sp. C12 ", ACS Catalysis, 2021
<small>Publication</small> | $<1\%$ |
| <hr/> | | |
| 17 | Submitted to Universiti Sains Malaysia
<small>Student Paper</small> | $<1\%$ |
| <hr/> | | |
| 18 | Ghader Bashiri, Edward N Baker. "Convergent pathways to biosynthesis of the versatile cofactor F420", Current Opinion in Structural Biology, 2020
<small>Publication</small> | $<1\%$ |
| <hr/> | | |
| 19 | www.rhea-db.org
<small>Internet Source</small> | $<1\%$ |
| <hr/> | | |
| 20 | Submitted to VIT University
<small>Student Paper</small> | $<1\%$ |
| <hr/> | | |
| 21 | Eri Nishiyama. "Identification of <i>Burkholderia multivorans</i> ATCC 17616 genes induced in soil environment by <i>in vivo</i> expression | $<1\%$ |

technology", Environmental Microbiology,
05/2010

Publication

22

cansarblack.icr.ac.uk

Internet Source

<1 %

23

en.wikipedia.org

Internet Source

<1 %

24

Sunil Parthasarathy, Annapoorni Lakshman
Sagar, Dayananda Siddavattam.

"Biodegradation of organophosphates:
biology and biotechnology", Elsevier BV, 2022

Publication

<1 %

25

idoc.pub

Internet Source

<1 %

26

"Aerobic Utilization of Hydrocarbons, Oils, and
Lipids", Springer Science and Business Media
LLC, 2019

Publication

<1 %

27

Antonin Kunka, Jiri Damborsky, Zbynek
Prokop. "Haloalkane Dehalogenases From
Marine Organisms", Elsevier BV, 2018

Publication

<1 %

28

Submitted to Aston University

Student Paper

<1 %

29

Handbook of Hydrocarbon and Lipid
Microbiology, 2010.

Publication

<1 %

30	Helianthous Verma, Gauri Garg Dhingra, Monika Sharma, Vipin Gupta, Ram Krishan Negi, Yogendra Singh, Rup Lal. "Comparative genomics of Sphingopyxis spp. unravelled functional attributes", Genomics, 2020 Publication	<1 %
31	Submitted to Kenyatta University Student Paper	<1 %
32	epdf.tips Internet Source	<1 %
33	Submitted to University of Bedfordshire Student Paper	<1 %
34	elifesciences.org Internet Source	<1 %
35	Katsutoshi Hori, Seira Takahashi, Fumiko Taguchi. "Non-enzymatic degradation of carbon nanotubes in the presence of bacterial enzymes", Cambridge University Press (CUP), 2021 Publication	<1 %
36	mic.sgmjournals.org Internet Source	<1 %
37	othes.univie.ac.at Internet Source	<1 %
38	settlement.org Internet Source	<1 %

39

www.ijdra.com

Internet Source

<1 %

40

Submitted to Jawaharlal Nehru University
(JNU)

Student Paper

<1 %

41

Lian Hongzhen, Danni Li, Weihang Wang.
"Determination of p - Cresidine Sulfonic Acid
and p - Cresidine in p - Cresidine Sulfonic Acid
for Industrial Use by Reversed - Phase Ion -
Pair HPLC", Journal of Liquid Chromatography
& Related Technologies, 2007

Publication

<1 %

42

research-information.bris.ac.uk

Internet Source

<1 %

43

Jennifer L. Seffernick, Lawrence P. Wackett. "
Rapid Evolution of Bacterial Catabolic
Enzymes: A Case Study with Atrazine
Chlorohydrolase ", Biochemistry, 2001

Publication

<1 %

44

Li Cao. "Enantioconvergent production of
(R)-1-phenyl-1,2-ethanediol from styrene oxide
by combining the *Solanum tuberosum* and an
evolved *Agrobacterium radiobacter* AD1
epoxide hydrolases", Biotechnology and
Bioengineering, 06/20/2006

Publication

<1 %

45	Oosterkamp, Margreet J., Teun Veuskens, Flávia Talarico Saia, Sander A. B. Weelink, Lynne A. Goodwin, Hajnalka E. Daligault, David C. Bruce, John C. Detter, Roxanne Tapia, Cliff S. Han, Miriam L. Land, Loren J. Hauser, Alette A. M. Langenhoff, Jan Gerritse, Willem J. H. van Berkel, Dietmar H. Pieper, Howard Junca, Hauke Smidt, Gosse Schraa, Mark Davids, Peter J. Schaap, Caroline M. Plugge, and Alfons J. M. Stams. "Genome Analysis and Physiological Comparison of Alicyclophilus denitrificans Strains BC and K601T", PLoS ONE, 2013. Publication	<1 %
----	--	------

46	bmcbiochem.biomedcentral.com Internet Source	<1 %
----	---	------

47	eprints.nottingham.ac.uk Internet Source	<1 %
----	---	------

48	scholar.lib.vt.edu Internet Source	<1 %
----	---	------

49	archiv.ub.uni-heidelberg.de Internet Source	<1 %
----	---	------

50	manualzz.com Internet Source	<1 %
----	---	------

51	openaccess.iyte.edu.tr Internet Source	<1 %
----	---	------

52

www.bgb-info.com

Internet Source

<1 %

53

Submitted to Higher Education Commission
Pakistan

Student Paper

<1 %

54

Lygie Esquirol, Thomas S. Peat, Elena Sugrue,
Sahil Balotra et al. "Bacterial catabolism of s-
triazine herbicides: biochemistry, evolution
and application", Elsevier BV, 2020

Publication

<1 %

55

Supriya S. Pai, Andrew D. Ellington. "Chapter
21 Using RNA Aptamers and the Proximity
Ligation Assay for the Detection of Cell
Surface Antigens", Springer Science and
Business Media LLC, 2008

Publication

<1 %

56

biblio.ugent.be

Internet Source

<1 %

57

researchonline.rvc.ac.uk

Internet Source

<1 %

58

www.beilstein-journals.org

Internet Source

<1 %

59

"Bacterial Cell Surfaces", Springer Science and
Business Media LLC, 2013

Publication

<1 %

60

Andrea Fortova, Eva Sebestova, Veronika Stepankova, Tana Koudelakova, Lenka Palkova, Jiri Damborsky, Radka Chaloupkova. "DspA from Strongylocentrotus purpuratus: The first biochemically characterized haloalkane dehalogenase of non-microbial origin", Biochimie, 2013

Publication

<1 %

61

B. M. Campbell, T. Lynam, J. C. Hatton. "Small-scale patterning in the recruitment of forest species during succession in tropical dry forest, Mozambique", Vegetatio, 1990

Publication

<1 %

62

Submitted to Taylor's Education Group

Student Paper

<1 %

63

Submitted to University of Lancaster

Student Paper

<1 %

64

Vincent Burrus. "Mechanisms of stabilization of integrative and conjugative elements", Current Opinion in Microbiology, 2017

Publication

<1 %

65

epdf.pub

Internet Source

<1 %

66

Submitted to Queen's University of Belfast

Student Paper

<1 %

67

d-nb.info

Internet Source

<1 %

Exclude quotes On

Exclude matches

< 14 words

Exclude bibliography On