

**DIVERSITY AND CHARACTERIZATION OF
POLYHYDROXYALKANOATE SYNTHASE
(PhaC) IN SEAWATER AND MANGROVE
METAGENOMES**

FOONG CHOON PIN

UNIVERSITI SAINS MALAYSIA

2016

**DIVERSITY AND CHARACTERIZATION OF
POLYHYDROXYALKANOATE SYNTHASE
(PhaC) IN SEAWATER AND MANGROVE
METAGENOMES**

by

FOONG CHOON PIN

**Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy**

FEBRUARY 2016

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude and appreciation to my supervisor Prof. Dr. K Sudesh Kumar for his patient, encouragement and guidance during my PhD study as well as his kind arrangement for my short-term attachment at RIKEN Yokohama campus, Japan. I would also like to thank my lab members and friends in Ecobiomaterial Research Laboratory for their valuable supports and cherish moments that we had spent together.

I am grateful to my academic mentors from RIKEN, Professor Dr. Minami Matsui and Dr. Todd Taylor as well as their postdoctoral fellows and staffs for their advises and supports during my one year attachment in their laboratories. I am also grateful to Dr. Foong Swee Yoke and Dr. Shinji Kondo for their technical advises in mangrove sampling and bioinformatics analysis, respectively.

I would also like to thank Assoc. Prof. Dr. Yutaka Suzuki and his NGS-team members at the Department of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo for providing the sequencing service as well as Assoc. Prof. Dr. Shigeru Deguchi and Dr. Takashi Toyofuku from Japan Agency for Marine-Earth Science and Technology (JAMSTEC) for providing the seawater samples.

I would like to give a special thank you to my beloved family. They are always encouraging and supporting me throughout my postgraduate study. I would like to acknowledge MyBrain15 scholarship from the Ministry of Higher Education Malaysia and short-term International Program Associate (IPA) from RIKEN for their financial support in my PhD study.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	x
LIST OF FIGURES	xii
LIST OF SYMBOLS AND ABBREVIATIONS	xvi
ABSTRAK	xxi
ABSTRACT	xxiii
CHAPTER 1 – INTRODUCTION	
1.0 Introduction	1
1.1 Objectives	4
CHAPTER 2 – LITERATURE REVIEW	
2.1 Biobased plastics from microorganisms	6
2.1.1 Polyhydroxyalkanoate (PHA)	7
2.1.2 Properties of PHA	9
2.1.3 Applications of PHA	10
2.2 PHA producers	11
2.3 PHA biosynthesis pathways and PHA synthase (PhaC)	16
2.4 Culture-independent or metagenomics approaches	27
2.5 Metagenomic studies in mangrove and seawater biomes	31
2.6 Review on PHA synthase discovered from metagenomics resources	33
2.7 DNA sequencing technologies	34
2.7.1 First generation sequencer	34

2.7.2	Second generation sequencer	35
	(a) Roche 454	35
	(b) Illumina/ Solexa	36
	(c) Life Technologies SOLiD	39
	(d) Life Technologies Ion Torrent and Ion Proton	40
2.7.3	Third generation sequencer	40
	(a) Pacific Bioscience (PacBio)	41
2.8	Bioinformatics analysis in metagenomics studies	41
2.9	Genome walking	44
2.9.1	Restriction digestion-independent GW methods	45
2.9.2	Restriction digestion-dependent GW methods	45
2.9.3	GW application in metagenomics DNA	47

CHAPTER 3 – MATERIALS AND METHODS

3.1	General techniques	48
3.1.1	Weighing of chemicals and materials	48
3.1.2	Sterilization	48
3.1.3	Measurement of optical density (OD) and pH	48
3.2	Media preparation	49
3.2.1	Lysogeny broth (LB)	49
3.2.2	Nutrient rich (NR) medium	49
3.2.3	Mineral salts medium (MM)	49
3.2.4	Preparation of antibiotic stock solutions	50
3.2.5	Preparation of structurally related carbon sources	50
	(a) Sodium valerate	50

	(b) Sodium 4-hydroxybutyrate	51
	(c) Sodium hexanoate	51
	(d) Sodium heptanoate	51
3.3	General molecular biology techniques	52
3.3.1	Agarose gel electrophoresis	52
3.3.2	DNA quantification	52
3.3.3	PCR amplification of 16S ribosomal RNA (rRNA) gene	52
3.3.4	PCR and gel purification	53
3.3.5	Cloning of PCR product	54
3.3.6	Preparation of chemically competent cells	54
3.3.7	Plasmid DNA extraction	55
3.3.8	DNA sequencing	56
3.4	General bioinformatics analyses	56
3.5	Plasmids and bacterial strains	57
3.6	Primers for PCR amplification	58
3.7	Sampling sites	60
3.7.1	Mangrove soil (Penang, Malaysia)	60
3.7.2	Seawater (Japan)	62
3.8	Total DNA extraction from mangrove soil samples	63
3.8.1	Conventional cetyl trimethyl ammonium bromide (CTAB)/ sodium dodecyl sulfate (SDS)-based method (Zhou <i>et al.</i> , 1996)	63
3.8.2	MO BIO PowerClean ® DNA Clean-Up Kit	64
3.8.3	MO BIO PowerSoil ® DNA Isolation Kit	65
3.9	Total DNA extraction from seawater samples	67
3.10	Whole genome amplification (WGA)	69

3.11	Methodology for objective (a)	69
3.11.1	Whole genome shotgun sequencing	69
3.11.2	Sequence annotation and analyses via MG-RAST portal	70
	(a) Annotation	70
	(b) Taxonomic profile analyses	73
3.12	Methodology for objective (b)	74
3.12.1	Identification of partial PHA synthase gene	74
3.13	Methodology for objective (c)	78
3.13.1	<i>De novo</i> assembly of full length PHA synthase gene	78
3.13.2	Sequence verification of the <i>de novo</i> assembled PHA synthase genes	79
3.13.3	PCR amplification of partial Class I and Class II PHA synthase genes	80
3.13.4	Sequence analyses for partial PHA synthase gene	81
3.13.5	Comparison of the PHA synthases from various environmental samples	81
3.14	Methodology for objective (d)	82
3.14.1	Genome walking for PHA synthase gene from seawater metagenomes	82
	(a) Restriction digestion and self-ligation of DNA fragments	82
	(b) Inverse PCR	82
	(c) Affinity purification	83
	(d) Nested PCR	83
3.14.2	Construction of <i>Cupriavidus necator</i> PHB ⁻ 4 transformants	84
	(a) Ribosome binding site (RBS) prediction	84

(b) Recombinant plasmids preparation	84
(c) Transformation of recombinant plasmid into <i>Escherichia coli</i> S17-1	85
(d) Bacterial transconjugation	86
3.14.3 PHA biosynthesis	87
3.14.4 PHA content quantification	88
(a) Preparation of methanolysis solution	88
(b) Preparation of caprylic methyl ester (CME) solution	88
(c) Methanolysis	88
(d) Gas chromatography (GC)	89
(e) Calculation of PHA content and monomer composition	90
(f) Statistical analysis	92
3.14.5 Fluorescence microscopic imaging	92
(a) Preparation of Nile blue A [1% (w/v)] and acetic acid [8% (v/v)] solutions	92
(b) Fluorescence microscopic observation	91
3.14.6 <i>In vitro</i> PHA synthase activity assay	93
(a) Preparation of reagents	93
(b) Sonication	93
(c) Bradford assay	93
(d) Measurement of PHA synthase activity	94

CHAPTER 4 – RESULTS AND DISCUSSION

4.1 Mangrove soil texture and metal content analyses	96
4.2 Challenges in extracting mangrove soil metagenomic DNA	99

4.2.1	Conventional CTAB/SDS-based DNA extraction with purification	100
4.2.2	DNA extraction using MOBIO PowerSoil ® DNA Isolation Kit	102
4.3	PCR inhibitory assay and removal of exogenous DNA in PCR preparation	106
4.4	Metagenomic microbial diversity of Penang Island mangrove soils	110
4.4.1	Relative read abundance of bacterial population in Penang Island mangrove soils	116
4.4.2	Prokaryotic diversity in mangrove soils compared to known diversity in the public database	120
4.4.3	Comparative study with other biome	127
4.5	PHA synthase (PhaC) from the Penang Island mangroves soil metagenomes	135
4.5.1	Relative read abundance of putative PhaC DNA fragments	134
4.5.2	Genus diversity of putative PhaC DNA fragments	141
4.6	Japan seawater metagenomics DNA for novel PHA synthase exploration	146
4.7	Whole genome amplification (WGA)	148
4.8	Partial PHA synthase gene fragments from Japan seawater metagenomic DNA	150
4.9	Phylogenetic comparison of PHA synthase from culture-independent studies	158
4.9.1	<i>De novo</i> sequence assembly of <i>phaC</i> gene from Penang Island mangrove soil metagenomes	158

4.9.2	Comparative analysis of PHA synthase from culture-independent studies	160
4.10	Characterization of full-length novel PHA synthase genes from Japan seawater metagenomes	174
4.10.1	Genome walking for novel PHA synthases from seawater metagenomes	174
4.10.2	Polyhydroxyalkanoate (PHA) production of seawater derived PHA synthases in <i>Cupriavidus necator</i> PHB ⁻ 4 transformants	179
4.10.3	PHA synthase activity	186
CHAPTER 5 – CONCLUSION		188
REFERENCES		192
APPENDIX		

LIST OF TABLES

	PAGE
Table 2.1	Summary of PHA-producing genera from the domain Bacteria 13
Table 2.2	Summary of PHA-producing genera from the domain Archaea 15
Table 2.3	Major enzymes involved in the PHA biosynthesis and biodegradation pathways 19
Table 2.4	Primers targeting on various classes of PHA synthase 23
Table 2.5	Overview of the specifications for the latest next-generation sequencing platforms 38
Table 3.1	List of plasmids and bacterial strains used in this study 57
Table 3.2	List of primers used in this study 59
Table 3.3	GPS coordinates and description of the seawater sampling sites 63
Table 4.1	Soil physicochemical and soil texture of the Batu Maung (BM), Balik Pulau (BP) 96
Table 4.2	Metal contents in mangrove soil from Batu Maung (BM) and Balik Pulau (BP) and also selected contaminated mangrove studies 98
Table 4.3	Concentration, purity and yield of the total soil DNA from the Batu Maung (BM) and Balik Pulau (BP) mangrove using MO BIO PowerSoil® DNA Isolation Kit with modifications 104
Table 4.4	Summary of DNA sequencing statistics and annotations from MG-RAST 111
Table 4.5	Read abundance of orders from the class <i>Deltaproteobacteria</i> in BM and BP mangrove soil metagenomes 119
Table 4.6	Top 10 most abundant known microbial genus in BM and BP mangrove soil metagenomes 120
Table 4.7	Richness of genus diversity in Penang Island mangrove soils compared to known genus diversity in the domain Bacteria 126
Table 4.8	Richness of genus diversity in Penang Island mangrove soils compared to known genus diversity in the phylum <i>Proteobacteria</i> 127
Table 4.9	Richness of genus diversity in Penang Island mangrove soils compared to known genus diversity in the domain Archaea 127

Table 4.10	Metagenome data sets from different biomes that are publicly available	131
Table 4.11	Number of putative PHA synthase DNA fragments using different calculations (C1-ALL and C2-PLB)	137
Table 4.12	Genus diversity and relative read abundance of PHA synthase from the class <i>Deltaproteobacteria</i> in proportion to the total PhaC calculated using C2-PLB	141
Table 4.13	Summary of PHA producing genera and PHA synthase genes	143
Table 4.14	DNA concentration and yield for seawater metagenomic DNA	147
Table 4.15	Genetic group classification and closest organism matches for the partial putative PHA synthase genes from Japan seawater metagenomes	151
Table 4.16	Statistics of the output from the <i>de novo</i> sequence assembly of PHA synthase DNA fragments in Penang Island mangrove soil metagenomes	158
Table 4.17	Closest organism matches for the putative full-length CDS of <i>phaC</i> genes from Penang Island mangrove soil metagenomes	159
Table 4.18	PHA synthase from culture-independent studies (uncultured bacteria)	160
Table 4.19	Checklist of the eight highly conserved amino acid residues and proposed catalytic triad of PHA synthase for the undefined (UD) clusters	165
Table 4.20	Closest organism matches for the putative full-length CDS of <i>phaC</i> genes from Japan seawater metagenomes	177
Table 4.21	PHA production by different <i>C. necator</i> PHB ⁻ 4 transformants that contained putative PHA synthase genes obtained from Japan seawater metagenomes	180
Table 4.22	PHA production of different strains/transformant using fructose (10 g/L) or CPKO (5 g/L) as sole carbon source	182
Table 4.23	PHA production of transformant I-GG18 using fructose (5 g/L) and added with different precursor carbon sources	184
Table 4.24	<i>In vitro</i> PHA synthase activity for different bacterial strain/transformant using different carbons sources	187

LIST OF FIGURES

		PAGE
Figure 1.1	The flow of ideas, aims and major workflow in this study	5
Figure 2.1	Classification of bioplastics and conventional petrochemical-based plastics according to their raw materials and biodegradability	7
Figure 2.2	The general chemical structure of different PHAs	8
Figure 2.3	Major PHA biosynthesis and biodegradation pathways in bacteria	18
Figure 2.4	Classification of PHA synthases	21
Figure 2.5	Position of the primers targeting on different classes of PHA synthase	25
Figure 2.6	A summary of molecular biological methods to study microorganisms using both sequence- and function-based approaches at the DNA level	29
Figure 2.7	General library preparation workflow and sequencing chemistry of the 1st-, 2nd- and 3rd-generation sequencers	37
Figure 3.1	Sampling locations for mangrove soil samples	61
Figure 3.2	Sampling locations for seawater samples	62
Figure 3.3	Schematic diagram of MG-RAST version 3 analysis pipeline	71
Figure 3.4	Bioinformatics analyses workflow	74
Figure 3.5	Construction of recombinant plasmids pBBR1MCS-2 with insertion of <i>phaC1</i> promoter from <i>C. necator</i> H16 and putative PHA synthase gene from seawater metagenomes: (a) I-GG18, (b) I-GG1 and (c) I-GG12	85
Figure 4.1	Total soil DNA (metagenomic DNA) extracted from Batu Maung (BM) and Balik Pulau (BP) mangroves using CTAB/SDS-based DNA extraction and DNA purification with MO BIO PowerClean Clean-Up Kit	101
Figure 4.2	Total soil DNA extracted from Batu Maung (BM) and Balik Pulau (BP) mangroves using MO BIO PowerSoil® DNA Isolation Kit without modification	103

Figure 4.3	Total soil DNA extracted from Batu Maung (BM) and Balik Pulau (BP) mangroves using MO BIO PowerSoil® DNA Isolation Kit with modifications	104
Figure 4.4	PCR amplification of the 16S rRNA gene using various dilutions of DNA template from the conventional CTAB/SDS-based DNA extraction method	107
Figure 4.5	PCR amplification of the 16S rRNA gene using DNA template purified with (a) MO BIO PowerClean Clean-Up Kit; (b) MO BIO PowerSoil DNA Isolation Kit	108
Figure 4.6	16S rRNA gene PCR amplification using DNA template extracted with MO BIO PowerSoil® DNA Isolation Kit	109
Figure 4.7	Known and unknown species read abundance based on LCA classification approach using cutoffs of e-value $\leq 1e-5$, identity $\geq 60\%$ and alignment length ≥ 15	113
Figure 4.8	Rarefaction and alpha diversity analyses from the MG-RAST portal	114
Figure 4.9	Phylogenetic tree showing the comparison of microbial diversity and abundance at the phylum level between BM and BP mangrove soil metagenomes	115
Figure 4.10	Read abundance of the major bacterial phylum in BM and BP mangrove soil metagenomes (in proportion to total bacterial read count)	117
Figure 4.11	Read abundance of classes from the phylum <i>Proteobacteria</i> in BM and BP mangrove soil metagenomes (in proportion to total bacterial read count)	118
Figure 4.12	Known bacterial diversity at the genus level (categorized according to their respective phylum) based on the NCBI RefSeq 16S rRNA gene sequences.	122
Figure 4.13	Bacterial genus diversity of the Penang Island mangrove soils (categorized according to their respective phylum) based on the LCA classification approach against M5NR database	123
Figure 4.14	Known bacterial diversity at the genus level for the phylum <i>Proteobacteria</i> (categorized according to their respective	124

	class) based on the NCBI RefSeq 16S rRNA gene sequences	
Figure 4.15	<i>Proteobacteria</i> genus diversity of the Penang Island mangrove soils (categorized according to their respective class) based on the LCA classification approach against M5NR database	124
Figure 4.16	Known archaeal diversity at the genus level (categorized according to their respective phylum) based on the NCBI RefSeq 16S rRNA gene sequences	125
Figure 4.17	Archaeal genus diversity of the Penang Island mangrove soils (categorized according to their respective phylum) based on the LCA classification approach against M5NR database	125
Figure 4.18	Rarefaction and alpha diversity analyses from the MG-RAST portal	128
Figure 4.19	Phylogenetic tree showing the comparison of the diversity and abundance of microbial phylum (Archaea, Bacteria and Eukaryota domains) between the Penang Island and Brazilian mangrove soil biomes	130
Figure 4.20	Comparative analysis of various metagenome biomes using principal component analysis based on LCA taxonomic classification in the MG-RAST portal	134
Figure 4.21	Relative read abundance of DNA fragments annotated as putative PHA synthase in the Penang Island mangrove soil metagenomes	138
Figure 4.22	Relative read abundance of DNA fragments annotated as putative PHA synthase and containing putative lipase box-like motif “[G/A/S]-X-C-X-G-[G/A/S]” in the Penang Island mangrove soil metagenomes	139
Figure 4.23	Genus diversity of the putative PhaC DNA fragments in the Penang Island mangroves soil metagenomes	142
Figure 4.24	Distribution of amino acid sequence identity of the annotated putative PHA synthase based on RefSeq database using the C2-PLB calculation	144
Figure 4.25	Total seawater DNA extraction using FastDNA™ 2 mL SPIN Kit for Soil (MP Biomedicals, USA) with modified protocol	147

Figure 4.26	Whole genome amplification of seawater metagenomic DNA	149
Figure 4.27	PCR amplification of partial Class I and Class II PHA synthase gene from WGA seawater metagenomic DNA. Annealing temperature (T_a) = 54°C	150
Figure 4.28	Neighbor-joining phylogenetic tree showing the seawater putative PHA synthase clones closely clustered according to their designated genetic groups respectively based on a cut off of 90 % nucleotide sequence similarity	154
Figure 4.29	MUSCLE multiple sequence alignment of putative partial PHA synthases from Japan seawater metagenomes	155
Figure 4.30	Neighbor-Joining phylogenetic tree of all the classes of PHA synthase (amino acid sequence)	162
Figure 4.31	Undefined classes or clusters of PHA synthase	164
Figure 4.32	Neighbor-joining phylogenetic tree of Class I PHA synthase (protein sequence) from culture-independent studies	168
Figure 4.33	Neighbor-joining phylogenetic tree of Class II PHA synthase (protein sequence) from culture-independent studies	170
Figure 4.34	Neighbor-joining phylogenetic tree of Class III PHA synthase (protein sequence) from culture-independent studies	172
Figure 4.35	Neighbor-joining phylogenetic tree of Class IV PHA synthase (protein sequence) from culture-independent studies	173
Figure 4.36	Schematic diagram of 3 genome walking DNA fragments from Japan seawater metagenomes that contained putative PHA synthase genes	175
Figure 4.37	Multiple sequence alignment of 3 full-length putative CDS of PHA synthases obtained from Japan seawater metagenomes with <i>Cupriavidus necator</i> H16 PhaC1 protein (WP_011615085)	178
Figure 4.38	Observation of different <i>Cupriavidus necator</i> strains under fluorescence microscope (1000 ×) after 48 h cultivation in nitrogen-limiting conditions	183

LIST OF SYMBOLS AND ABBREVIATIONS

–	Minus
%	Percentage
&	And
(<i>R</i>)	<i>Rectus</i> -isomer
(<i>S</i>)	<i>Sinister</i> -isomer
~	Approximately
±	Plus-minus
×	Times
≥	Greater-than or equal
°C	Degree Celsius
α	Alpha
β	Beta
γ	Gamma
μ	Micrometers
μg	Microgram
μm	micrometer
μL	Microliter
μM	Micromolar
3HB	3-hydroxybutyrate
3HHp	3-hydroxyheptanoate
3HHx	3-hydroxyhexanoate
4HB	4-hydroxybutyrate
am	Ante meridem
ATP	Adenosine triphosphate
BLASTn	Basic Local Alignment Search Tool for nucleotide
BLASTp	Basic Local Alignment Search Tool for protein

BLASTx	Basic Local Alignment Search Tool for protein using translated nucleotide query
BM	Batu Maung mangrove forest
bp	Base pair
BP	Balik Pulau mangrove forest
CCD	Charge-coupled device
CDS	Coding sequence
CoA	Coenzyme-A
CPKO	Crude palm kernel oil
ddNTPs	Dideoxynucleic acids
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphates
eDNA	Environmental DNA
FISH	Fluorescence <i>in situ</i> hybridization
g	Gram
<i>g</i>	gravity
GB	Gigabyte
GC	Gas chromatography
GPS	Global positioning system
GW	Genome walking
h	Hour
HA	Hydroxyacyl
I-PCR	Inverse PCR
kb	Kilo-base
kDa	Kilo dalton
kPa	Kilopascal

L	Liter
LB	Lysogeny broth
LCA	Lowest common ancestor
LCL-PHAs	Long chain length polyhydroxyalkanoates
M	Molar
MCL-PHAs	Medium chain length polyhydroxyalkanoates
MDA	Multiple displacement amplification
MFS	Major facilitator superfamily
mg	Milligram
Min	Minute
mL	Milliliter
mm	Millimeter
mM	Milimolar
MM	Mineral salts medium
mol%	Mole percent
MPa	Megapascal
NCBI	National center for Biotechnology Information
ng	Nanogram
NGS	Next-generation sequencing
nm	Nanometer
No.	Number
NR	Nutrient rich
NTC	No-template control
OD	Optical density
OD ₆₀₀	Optical density at wavelength 600 nm
P3HB	Poly-3-hydroxybutyrate
P(3HB- <i>co</i> -4HB)	Poly(3-hydroxybutyrate- <i>co</i> -4-hydroxybutyrate)

P(3HB- <i>co</i> -3HV)	Poly(3-hydroxybutyrate- <i>co</i> -3-hydroxyvalerate)
P(3HHx)	Poly(3-hydroxyhexanoate)
PA	Polyamide
PBAT	Poly(butylene adipate- <i>co</i> -terephthalate)
PBS	Polybutylene succinate
PCA	Principal component analysis
PCL	Polycaprolactone
PCR	Polymerase chain reaction
PE	Polyethylene
PE	Polyethylene
PET	Polyethylene terephthalate
pH	Potential hydrogen
PHA	Polyhydroxyalkanoate
PhaA	beta-ketothiolase
PhaB	NADPH-dependent acetoacetyl-CoA reductase
PhaC	PHA synthase
PhaE	Polyhydroxyalkanoate granule associated protein
PhaR	Repressor protein
phaZ	PHA depolymerase
PLA	Poly(lactic acid)
pm	Post meridiem
PP	Polypropylene
PPP	Poly(para-phenylene)
PS	Polystyrene
Psi	Pounds per square inch
PTT	Polytrimethylene terephthalate
PVC	Polyvinyl chloride

RBS	Ribosome binding sequence
RNA	Ribonucleic acid
rpm	Revolutions per minute
rRNA	ribosomal ribonucleic acid
s	Second
SCL-PHAs	Short chain length polyhydroxyalkanoates
SSCP	Single-strand confirmation polymorphism
SSU	Small subunit
TCA	Tricarboxylic acid
TGGE	Temperature gradient gel electrophoresis
T-RFLP	Terminal restriction fragment length polymorphism
U	Unit
UV	Ultraviolet
V	Voltage
v/v	Volume per volume
WGA	Whole genome amplification
wt%	Weight percent
ZMW	Zero-mode waveguide

**KEPELBAGAIAN DAN PENCIRIAN SINTASE
POLIHIDROKSIALKANOAT (PhaC) DALAM METAGENOM AIR LAUT
DAN PAYA BAKAU**

ABSTRAK

Komuniti mikrob bagi dua tanah paya bakau Pulau Pinang (Batu Maung dan Balik Pulau) yang dipengaruhi oleh aktiviti antropogenik telah dikaji dengan menggunakan pendekatan penjujukan metagenomik “shotgun” tanpa-kultur. Dua set data metagenomik (~250 GB) dihasilkan melalui platform “Next-generation Sequencing (NGS)” Illumina HiSeq dan disimpan dalam pelayan awam “Metagenomic-Rapid Annotations using Subsystems Technology (MG-RAST)”. Analisis taksonomi mikrob menunjukkan bahawa kedua-dua tanah paya bakau Pulau Pinang didominasi oleh Bakteria (97 %), *Proteobacteria* (43 %) dan *Deltaproteobacteria* (15 %) pada peringkat domain, filum dan kelas masing-masing. Pada peringkat genus, kebanyakan bakteria anaerobik diperhatikan terdiri daripada *Deltaproteobacteria*. Sebahagian besar daripada jujukan adalah milik spesis mikrob (70 %) dan filum (32 %) yang belum dikenalpasti atau belum dikultur. Kajian kepelbagaian sintase PHA (PhaC) menunjukkan bahawa lebih kurang 21-23% daripada jumlah genera mikrob yang dikesan (Bakteria and Arkea) dalam tanah paya bakau Pulau Pinang mengandungi PhaCs dengan motif putatif “lipase-box-like” “(G/A/S)-X-C-X-G-(G/A/S)” berdasarkan keputusan BLASTx terhadap pangkalan data Jujukan Rujukan (RefSeq) dalam Pusat Kebangsaan untuk Maklumat Bioteknologi (NCBI). Jangkaan PhaC separa ini secara keseluruhannya (>80 %) dimiliki oleh filum *Proteobacteria* (*Alphabacteria*, *Betabacteria*, *Deltabacteria* dan *Gammabacteria*). Lebih kurang 27-37 % daripada PhaC berpotensi kepunyaan genus

mikrob baru sekiranya purata 70 % kadar takat identiti asid amino (AAI) digunakan. Pada masa yang sama, pendekatan pemeriksaan yang berbeza berasaskan PCR genotip telah digunakan untuk menyiasat PhaC Kelas I and II dari metagenom air laut cetek dan laut dalam (24 m hingga 5373 m) yang diperolehi dari Palung Nankai dan Jurang Jepun. Sebanyak 20 kumpulan genetik (KG) separa PhaC telah ditentukan. Kesemua KG PhaC mempunyai organisma yang terdekat, iaitu *Proteobacteria* dan didominasi oleh *Alphaproteobacteria*. Lima KG PhaC mempunyai AAI <70% dan berkemungkinan tinggi dimiliki oleh genus mikrob baru dari *Alphaproteobacteria*. Tambahan itu, analisis filogenetik dengan menggunakan semua PhaCs yang diperolehi daripada sumber-sumber metagenomik menunjukkan tiga kelompok baru atau kluster PhaC yang belum dikenali sebagai tambahan kepada empat kelompok PhaC (Kelas I hingga IV) yang sedia ada. Pengesahan fungsi PhaC juga dikaji dan tiga jujukan lengkap kod DNA telah berjaya diperolehi daripada metagenom air laut Jepun melalui kaedah “genome walking”. Hanya PhaC I-GG18 berfungsi aktif dan mampu menghasilkan PHA dalam transforman *Cupriavidus necator* PHB⁻4 (mutan PHB-negatif). PhaC I-GG18 mempunyai identiti jujukan protein yang tinggi (97 %) kepada PhaC dari genus penghasil PHA baru *Marinobacter*. PhaC I GG18 ini mempunyai substrat khusus terhadap monomer PHA berantai pendek (SCL-PHA) seperti 3-hydroxybutyryl-CoA dan 4-hydroxybutyryl-CoA. Aktiviti sintase PhaC I-GG18 dalam transformant *C. necator* PHB⁻4 adalah 10 kali ganda lebih rendah daripada *C. necator* H16 jenis liar pada 24 jam penggeraman di dalam medium terhad nitrogen.

DIVERSITY AND CHARACTERIZATION OF POLYHYDROXYALKANOATE SYNTHASE (PhaC) IN SEAWATER AND MANGROVE METAGENOMES

ABSTRACT

The microbial communities of two local Penang mangrove soils (Batu Maung and Balik Pulau) which are under anthropogenic influences were investigated using culture-independent shotgun metagenome sequencing approach. Two metagenome data sets (~250 GB) were generated from the Illumina HiSeq next-generation sequencing (NGS) platform and then deposited in Metagenomic-Rapid Annotations using Subsystems Technology (MG-RAST) public server. Microbial taxonomic analysis showed that both Penang mangrove soils were dominated by Bacteria (97 %), *Proteobacteria* (43 %) and *Deltaproteobacteria* (15 %) at the domain, phylum and class levels, respectively. At the genus level, predominance of anaerobic bacteria was observed and mostly belonged to *Deltaproteobacteria*. A large portion of the reads belonged to unknown or yet uncultured microbial species (70 %) and microbial phyla (32 %). Investigation on the PHA synthase (PhaC) diversity shown that about 21-23 % of the total detected microbial (bacteria and archaea) genera in the Penang mangrove soils contained PhaCs with putative lipase-box-like motif “(G/A/S)-X-C-X-G-(G/A/S)” based on the BLASTx results against National Center for Biotechnology Information Reference Sequence (NCBI RefSeq) database. These partial putative PhaCs predominantly (>80 %) belonged to the phylum *Proteobacteria* (*Alphaproteobacteria*, *Betaproteobacteria*, *Deltaproteobacteria*, and *Gammaproteobacteria*). About 27-37 % of the PhaCs potentially belonged to new

microbial genus if a 70 % average amino acid identity (AAI) cutoff was applied. At the same time, a different PCR-based genotypic screening approach was employed in this study to investigate Class I and II PhaCs from shallow and deep-sea seawater metagenomes (24 m to 5373 m) which were collected from Nankai Trough and Japan Trench. A total of 20 partial PhaC genetic groups (GGs) were determined. All the GGs had closest organism matches to *Proteobacteria* and predominated by *Alphaproteobacteria*. Five PhaC GGs had AAI < 70 % and most probably belonged to new microbial genus from *Alphaproteobacteria*. Furthermore, phylogenetic analysis using all the PhaCs derived from metagenomic resources showed three new or undefined clusters of PhaC in addition to four existing known clusters of PhaC (Class I to IV). For functional verification, three complete DNA coding sequences were successfully obtained from Japan seawater metagenomes by genome walking approach. Only I-GG18 PhaC was functionally active and able to produce PHA in transformant *Cupriavidus necator* PHB⁻4 (PHB-negative mutant). I-GG18 PhaC had very high protein sequence identity (97 %) to the PhaCs of new PHA producing genus *Marinobacter*. This I-GG18 PhaC had substrate specificity towards short-chain-length PHA (SCL-PHA) monomers such as 3-hydroxybutyryl-CoA and 4-hydroxybutyryl-CoA. The synthase activity of I-GG18 PhaC in transformant *C. necator* PHB⁻4 was 10 folds lower than the wild-type *C. necator* H16 at 24th hour of incubation in nitrogen-limiting medium.

CHAPTER 1

1.0 INTRODUCTION

Plastic products have been widely integrated into our lifestyle due to their flexible and durable features. However, non-biodegradable nature of conventional petrochemical- or fossil-based plastics has made them a serious threat to our environment and also other living organisms. Scientists and public are now becoming aware about global energy crisis, waste and pollution issues due to increasing human population. Therefore, sustainable and eco-friendly materials such as polyhydroxyalkanoates (PHAs) as well as other biobased and biodegradable polymers [polylactic acid (PLA) and polybutylene succinate (PBS)] are promising alternative plastic materials to protect our planet from plastic waste accumulation. Commercial productions and applications of PHAs are ongoing in a few countries, while some countries have also started to ban the usage of fossil-based plastic products especially the single-use items.

PHAs are carbon and energy reserve biopolymers which are produced from microorganisms (bacteria and archaea) under unfavorable growth and stress conditions. There are three major factors that determine the types of PHA polymer that can be produced in a microorganism: (1) substrate specificity of the PHA synthase (PhaC), (2) metabolic pathways in the microbial host, and (3) types of carbon source provided. Carbon sources and microbial metabolic pathways would influence the types of PHA monomers or substrates supplied to the PHA synthase. The key enzyme in PHA biosynthesis pathway is the PhaC, which has the “absolute power” to select what types of PHA monomer to be incorporated into the PHA polymer chain depending on its substrate specificity. Various types of PhaC have been reported. Together, they have very broad substrate specificity with more than

150 different PHA constituents that can be polymerized. One of the possible reasons could be their low protein sequence similarity (8 to 96 %). Thus, it is impossible to detect all the four classes of PhaC using a single universal primer set. The current evidences for a PHA synthase at the primary structural level are composed of eight highly conserved amino acid residues, a putative lipase-box-like motif “G-X-C-X-G” in the α/β domain and a catalytic triad (Steinbüchel and Valentin, 1995; Madison and Huisman, 1999; Rehm, 2003).

To date, the diversity of PHA, PHA producer and PhaC are mostly being studied through pure isolates using culture-dependent approaches. A total of four classes of PhaC and 167 PHA producers have been reported from the existing cultivable microbial collections which are believed to constitute not more than 15 % of the total microorganisms (Rehm, 2003; Koller *et al.*, 2013). Microbiologists generally accept that at least 85 % of the microorganisms have not been cultured due to unsuitable *in vitro* conditions in the laboratory (Amann *et al.*, 1995; Lok *et al.*, 2015). Therefore, there is a huge knowledge gap in PhaC diversity from the under-discovered microbial world. Culture-independent or metagenomic approaches are the only tools that can directly access this untapped and huge microbial genomic information.

Previous high-throughput shotgun metagenome sequencing studies have shown highly complex microbial diversity (> 700 species) in mangrove soils (Andreote *et al.* 2012; Thompson *et al.* 2013). Sequencing output has become the only limitation to uncover the complete or total microbial diversity in the mangrove soil biome. This is especially important for the detection of rare or low abundance unculturable microbial species. Microbial communities of two local Penang mangrove soils from Batu Maung and Balik Pulau that are under the influence of

anthropogenic activities were investigated in this study by using the state-of-the-art next-generation sequencing (NGS) platform. The Illumina HiSeq platform can generate a much higher sequencing output (> 500 folds) compared to the two previous studies which had used the Roche 454 FLX+ platform. In addition to descriptive analysis on the taxonomic information (microbial diversity and relative abundance), these shotgun metagenome data sets can also provide functional information. Mangrove soil biome contains high microbial diversity and is continuously exposed to various abiotic stresses such as saline and anoxic conditions. No study on PhaC from mangrove soil metagenome has been reported. Therefore, there will be a high chance to discover large numbers of novel PhaCs from new microbial genera in the mangrove soil metagenome particularly from the anaerobic microorganisms.

In addition, precious seawater samples from shallow to deep-sea (24 m to 5373 m) were collected from Nankai Trough and Japan Trench by Japan Agency for Marine-Earth Science and Technology (JAMSTEC). There is currently only one published study on the finding of PhaCs from Northern Baltic Sea metagenomes (Pärnänen *et al.*, 2015), while no report was found on the PhaC from deep-sea environments. Deep-sea biome is considered as an extreme and stressed environment with low availability of sunlight, low temperature and high hydrostatic pressure. Besides, it is also difficult to access deep-sea environment due to technical challenges and high cost of conducting deep-sea research. A previous study showed that deep-sea contains high diversity of unknown low abundance or rare microbial species (Sogin *et al.*, 2006). Thus, it will be interesting to discover new PhaC from these Japan deep-sea metagenomes.

Overall, two different sequence-based culture-independent approaches were applied in this study to explore PhaC from mangrove soil and seawater metagenomes. The first approach was high-throughput shotgun metagenome sequencing, which could provide both microbial taxonomic information and diversity of PhaC from the Penang mangrove soils. The second approach was PCR-based genotypic screening to detect Class I and II PhaC from the Japan seawater metagenomes. Phylogenetic analysis of PhaCs was also performed in this study by using all the PhaC sequences obtained from various metagenomic resources in order to identify new cluster of PhaC. In addition, an interesting genome walking approach was applied on the Japan seawater metagenomes to determine the complete coding sequences of PhaCs without having any prior knowledge on the genomic content of the uncultured microorganisms. Finally, examination of these full-length PhaCs through PHA biosynthesis was carried out to verify their functionality *in vivo* (Figure 1.1).

1.1 Objectives

- a) To study the microbial diversity and their relative abundance in Batu Maung and Balik Pulau mangrove soils in Penang Island using culture-independent shotgun metagenome sequencing approach.
- b) To investigate the prevalence of PHA synthase diversity and abundance in the Penang mangrove soils.
- c) To identify novel cluster of PHA synthase from the Penang mangrove soils, Japan seawaters (Japan Trench and Nankai Trough) and other metagenomic resources through phylogenetic comparison.

d) To examine novel PHA synthases for PHA production in heterologous host.

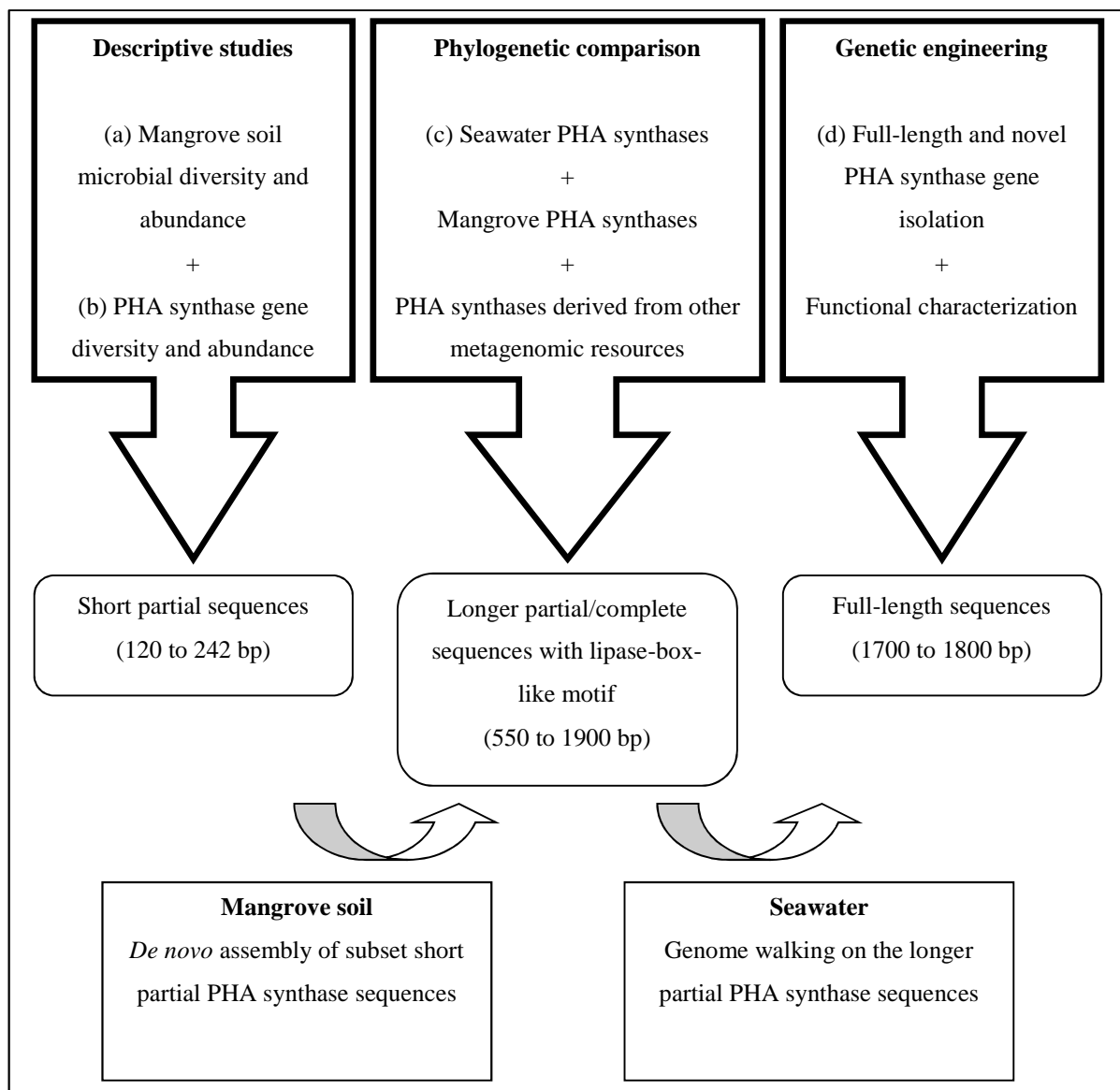


Figure 1.1: The flow of ideas, aims and major workflow in this study.

CHAPTER 2

2.0 LITERATURE REVIEW

2.1 Biobased plastics from microorganisms

Biodegradability and sustainability are two major concerns in the search for “green” materials to replace petrochemical-based (oil and natural gas) plastics such as polyethylene terephthalate (PET), polyvinyl chloride (PVC), polyethylene (PE), polypropylene (PP), polystyrene (PS) and polyamide (PA). These petrochemical-based plastics are very durable and tend to end up in landfill or unfavorably in the oceans as floating marine plastics such as the Great Pacific Garbage Patch (Kaiser, 2010). Plastics are found in about 90 % of seabirds as well as contributed to the deaths of 1 million seabirds and 100,000 sea mammals every year (Saikia and de Brito, 2012; Wilcox *et al.*, 2015).

Generally, biobased plastics include plant-derived plastics (starch, protein and cellulose) and microbial-derived plastics. Partially biobased plastics are produced through the blending of biobased materials with petrochemical-based plastics and they are eventually only partially biodegraded. Microorganisms are able to synthesize six types of monomers of biobased plastics such as hydroxyalkanoic acids for polyhydroxyalkanoates (PHAs), D- & L-lactic acids for polylactic acid (PLA), succinic acid for polybutylene succinate (PBS), bioethylene for biopolyethylene (PE), 1,3-propanediol for polytrimethylene terephthalate (PTT) and cis-3,5-cyclohexadiene-1,2-diols for poly(para-phenylene) (PPP). However, only the first three polymers are fully biodegradable (Figure 2.1). Among them, hydroxyalkanoic acids have a large number of structural variations. These microbial biobased plastics have very similar properties to the petrochemical-based plastics (Steinbüchel and Fächtenbusch, 1998; Chen, 2009).

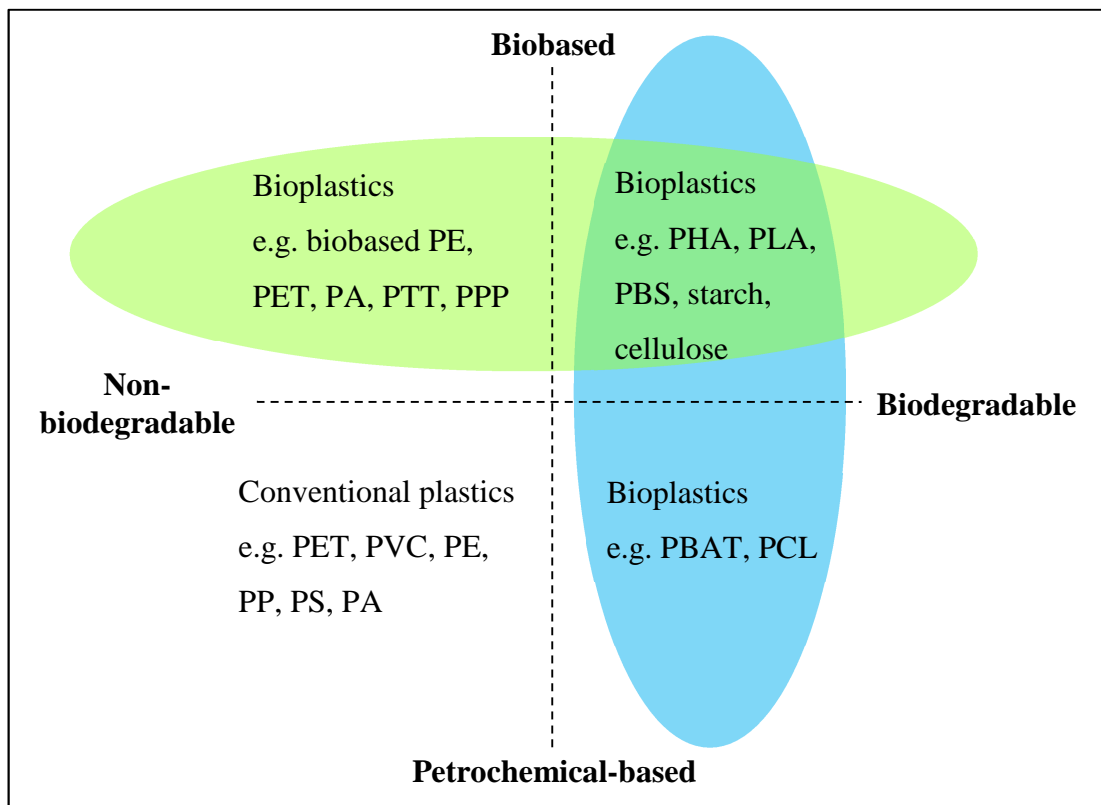


Figure 2.1: Classification of bioplastics and conventional petrochemical-based plastics according to their raw materials and biodegradability.

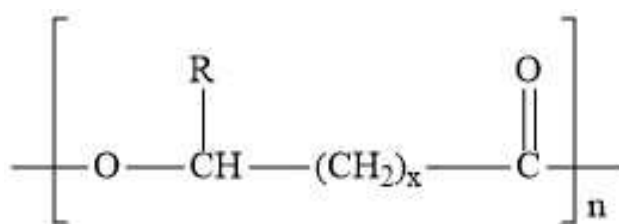
Polyethylene (PE); Polyethylene terephthalate (PET); polyamide (PA); Polytrimethylene terephthalate (PTT); Poly(para-phenylene) (PPP); Polyhydroxyalkanoate (PHA); Polylactic acid (PLA); Polybutylene succinate (PBS); polyvinyl chloride (PVC); polypropylene (PP); polystyrene (PS); poly(butylene adipate-co-terephthalate) (PBAT); polycaprolactone (PCL).

(Source: modified from Fact Sheet European Bioplastics, 2015)

2.1.1 Polyhydroxyalkanoate (PHA)

Polyhydroxyalkanoates (PHAs) are naturally produced by many bacteria and archaea under unbalanced growth conditions but with excess supply of carbon. The unbalanced growth conditions are such as limitations of nitrogen, phosphorus,

sulphur, magnesium or oxygen. PHAs are stored as carbon and energy reserves intracellularly (cytoplasm) in the form of water insoluble inclusions or granules (Anderson and Dawes, 1990). Maurice Lemoigne was the first to discover poly(3-hydroxybutyrate) (PHB) in *Bacillus megaterium* in 1926 (Lemoigne, 1926; Doi, 1990). PHB is the most common type of PHA produced by microorganisms. PHA other than PHB was first discovered in 1974 as a poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) [P(3HB-*co*-3HV)] copolymer (Wallen and Rohwedder, 1974; Sudesh *et al.*, 2000). Since then, more than 150 different PHA monomers have been identified (Steinbüchel and Valentin, 1995; Madison and Huisman, 1999). The general chemical structure of PHAs is shown in Figure 2.2.



Number of repeating units, x	Alkyl group, R	Polymer type
1	Hydrogen	Poly(3-hydroxypropionate)
	Methyl	Poly(3-hydroxybutyrate)
	Ethyl	Poly(3-hydroxyvalerate)
	Propyl	Poly(3-hydroxyhexanoate)
	Pentyl	Poly(3-hydroxyoctanoate)
	Nonyl	Poly(3-hydroxydodecanoate)
2	Hydrogen	Poly(4-hydroxybutyrate)
	Methyl	Poly(4-hydroxyvalerate)
3	Hydrogen	Poly(5-hydroxyvalerate)
	Methyl	Poly(5-hydroxyhexanoate)

n refers to number of repeating unit (100 – 30000)

Figure 2.2: The general chemical structure of different PHAs.

Source: Lee (1996a)

2.1.2 Properties of PHA

The major advantages of PHA compared to petrochemical-based plastics are biodegradability (via microbial enzymatic reactions), biocompatibility (natural and non-toxic) and sustainability (synthesized from renewable resources) (Zinn *et al.*, 2001; Jendrossek and Handrick, 2002; Sudesh and Iwata, 2008). PHA is completely biodegraded into carbon dioxide and water under aerobic condition, while under anaerobic condition it is biodegraded into methane and carbon dioxide by microorganisms (Lee, 1996b; Abou-Zeid *et al.*, 2001). The physical and thermal properties of PHAs are dependent on the monomer type, monomer composition and molecular weight of the polymer.

In general, PHA can be categorized into three major groups based on the carbon chain length of the monomers. Short chain length PHAs (SCL-PHAs) consists of monomers with 3 to 5 carbon atoms, medium chain length PHAs (MCL-PHAs) consists of monomers with 6 to 14 carbon atoms and long chain length PHAs (LCL-PHAs) consists of monomers with more than 14 carbon atoms (Lee, 1996b; Lu *et al.*, 2009). SCL-PHAs have thermoplastic properties (stiff and brittle material) such as high crystallinity, high tensile modulus and low elongation at break. MCL-PHAs have elastomeric properties (rubber-like material) such as low crystallinity, low melting temperature and high elongation at break (Sudesh *et al.*, 2000; Yu, 2007). PHAs with high mol % of SCL monomers and low mol % of MCL monomers have properties similar to polypropylene (PP). In contrast, PHAs with low mol % of SCL monomers and high mol % of MCL monomers have properties similar to low-density polyethylene (LDPE) (Abe and Doi, 2002, Sudesh *et al.*, 2007; Yu, 2007).

The molecular weights of microbial PHAs are in the range of 2×10^5 to 3×10^6 Da (Lee, 1996a). *Escherichia coli* transformant (a non-native PHA producer that

is lacking in PHA depolymerase activity) harboring PHA synthase gene from *Cupriavidus necator* could produce ultra-high molecular weight P(3HB) ranging from 3×10^6 to 1×10^7 Da (Kusaka *et al.*, 1998). The elongation at break and tensile strength are higher or better than low molecular weight P(3HB).

2.1.3 Applications of PHA

PHA have been commercialized by many companies since 1982 in several countries such as UK (ICI), USA (Metabolix, MHG, P&G and Newlight Technologies), Japan (Kaneka), Canada (Biomatera), Germany (Biomer), Italy (Bio-On), Brazil (PHB Industrial Brasil), Malaysia (SIRIM) and China (Tianjin GreenBio Materials and TianAn Biopolymer) (website: <http://bioplasticsinfo.com/polyhydroxy-alkonates/companies-concerned/>). PHA can be used as coating and packaging materials, disposable items, bio-implants, drug carriers, precursors for fine chemicals and biofuel productions (Amara, 2008; Chen 2009; Gao *et al.*, 2011). Packaging and disposable items are the most common applications of PHA and these include bottles, cups, razors, utensils, mulch films, diapers and feminine hygiene products. PHA can also be used as oil-blotting film in cosmetics and skin care industry (Sudesh *et al.*, 2007). In biomedical field, the biocompatibility and biodegradability features of PHA make it suitable for osteosynthetic materials, bone plates, surgical sutures, cardiovascular patches, wound dressings and tissue engineering scaffolds (Steinbüchel and Fächtenbusch, 1998; Zinn *et al.*, 2001; Chen and Wu, 2005; Jain *et al.*, 2010).

PHA could also be used as biodegradable carriers for long-term dosage of drugs, medicines, hormones, insecticides, herbicides and fertilizers under controlled release formulations (Pouton and Akhtar, 1996; Khanna and Srivastava, 2005; Jain *et*

al., 2010). Besides, PHAs have uniform chirality and are excellent starting chemicals (precursors) for the synthesis of other optically active compounds such as drugs vitamins and pheromones (Lee *et al.*, 1999; Reddy *et al.*, 2003; Jain *et al.*, 2010). The most recent discovery of PHA application is as a biofuel precursor which is first reported in 2009. PHA could be esterified with methanol to generate *R*-3-hydroxyalkanoate methyl ester (3HAME) via acid-catalyzed hydrolysis, which could be further used to generate combustion heat (Zhang *et al.*, 2009).

2.2 PHA producers

The first known PHA producer is *Bacillus megaterium* (Lemoigne, 1926). However, the study on PHA was relatively slow until the first crude oil crisis occurred in mid-1970s, which has triggered the efforts to look for alternative resources for petrochemical-based plastics. During the 1980s until 2010s, a large number of findings on new PHA producers were reported, for instance from the genus *Aeromonas*, *Azotobacter*, *Burkholderia*, *Chromobacterium*, *Cupriavidus*, *Delftia*, *Nocardia*, *Pseudomonas*, *Rhizobium*, *Rhodococcus* and *Streptomyces* (Valappil *et al.*, 2007; Chen, 2009).

Cupriavidus necator (previously known as *Wautersia eutropha*, *Ralstonia eutropha*, *Alcaligenes eutrophus* or *Hydrogenomonas eutrophus*) especially strain H16 (Schlegel and Lafferty, 1965) is the most extensively studied PHA producer and is a well-known model organism for PHA study (Reinecke and Steinbüchel, 2008). It can accumulate PHA up to 90 wt% of the dry cell weight using simple carbon sources and plant oil (Chen, 2009; Lee *et al.*, 2008). Whole bacterial genome sequencing of *C. necator* H16 has been completed and it contains two chromosomes and one megaplasmid (Pohlman *et al.*, 2006). Genome-wide transcriptome analyses

of *C. necator* H16 has also been performed using microarray to detect genes that are differentially transcribed during PHB biosynthesis by comparing it with PHB-negative mutant strains (PHB⁻ 4 and Δ *phaC1*) (Peplinski *et al.*, 2010). Besides, the first industrial scale production of PHA (Biopol®, PHBV copolymer) was achieved using *C. necator* in 1982 by Imperial Chemical Industries (ICI) (Luengo *et al.*, 2003; Verlinden *et al.*, 2007).

Pseudomonads (belonging to rRNA homology-group I) are also widely studied due to their unique ability to produce MCL-PHAs. The 3-hydroxyacyl-CoA substrates (C6 to C14) for the production of MCL-PHAs are derived from fatty acid β -oxidation and *de novo* fatty acid biosynthesis pathways (Huisman *et al.*, 1989; Anderson and Dawes, 1990; Witholt and Kessler, 1999; Sudesh *et al.*, 2000). Photosynthetic bacteria such as *Rhodospirillum rubrum* (Brandl *et al.*, 1989) and *Cyanobacteria* (*Synechocystis* sp., *Aulosira fertilissima* and *Spirulina subsalsa*) (Panda and Mallick, 2007; Shrivastav *et al.*, 2010; Samantaray and Mallick, 2014) are also interesting PHA producers because they are able to utilize sunlight and carbon dioxide to synthesize PHAs (photoautotrophic) without addition of extra carbon sources.

Besides, PHA producers have also been isolated from extreme environments such as hot springs, salt lakes and polar-regions. Extremophiles such as *Halobacteriaceae*, *Thermus thermophiles*, thermophilic *Streptomyces* sp. and psychrophilic *Pseudomonas* sp. possess the ability to synthesize PHAs (Fernandez-Castillo *et al.*, 1986; Pantazaki *et al.*, 2003; Phithakrotchanakoon *et al.*, 2009; Ayub *et al.*, 2009; Legat *et al.*, 2010).

To date, there are about 167 microbial PHA producing genera (150 bacteria and 17 archaea) (Reddy *et al.*, 2003; Zinn *et al.*, 2001; Koller *et al.*, 2010; Koller *et*

al., 2013) (Table 2.1 and 2.2). Majority of them belong to the phylum *Proteobacteria* (*Alpha-*, *Beta-*, *Delta* and *Gamma-proteobacteria*), followed by *Cyanobacteria*, *Euryarchaeota*, *Actinobacteria*, *Firmicutes*, *Thaumarchaeota*, *Chloroflexi* and *Deinococcus-Thermus*. The presence of PHA in eukaryote has been reported in human (blood and tissue) and fungi (*Aureobasidium*, *Penicillium*, *Physarum*) in the form of (*R*)-3-hydroxybutyrate oligomers (low molecular weight PHA) and poly- β -malic acid (similar chemical composition as natural PHA), respectively (Steinbüchel and Hein, 2001; Zinn *et al.*, 2001; Koller *et al.*, 2010).

Table 2.1: Summary of PHA-producing genera from the domain Bacteria

<i>Actinobacteria</i> (7)		
<i>Actinomycetes</i>	<i>Corynebacterium</i>	<i>Micrococcus</i>
<i>Microlunatus</i>	<i>Nocardia</i>	<i>Rhodococcus</i>
<i>Streptomyces</i>		
<i>Chloroflexi</i> (1)		
<i>Chloroflexus</i>		
<i>Cyanobacteria</i> (27)		
<i>Anabaena</i>	<i>Aphanocapsa</i>	<i>Aphanothece</i>
<i>Aulosira</i>	<i>Calothrix</i>	<i>Chlorogloea</i>
<i>Chroococcus</i>	<i>Cyanobacterium</i>	<i>Cyanothece</i>
<i>Fischerella</i>	<i>Gloeocapsaa</i>	<i>Gloeothece</i>
<i>Gomphosphaeria</i>	<i>Microcoleus (Microvoleus)</i>	<i>Microcystis</i>
<i>Nodularia</i>	<i>Nostoc</i>	<i>Oscillatoria</i>
<i>Pleurocapsa</i>	<i>Pseudoanabaen</i>	<i>Rivularia</i>
<i>Scytonema</i>	<i>Spirulina</i>	<i>Synechococcus (Anacystis)</i>
<i>Synechocystis</i>	<i>Tolypothrix</i>	<i>Westiellopsis</i>
<i>Deinococcus-Thermus</i> (1)		
<i>Thermus</i>		

<i>Firmicutes</i> (5)		
<i>Bacillus</i>	<i>Caryophanon</i>	<i>Clostridium</i>
<i>Staphylococcus</i>	<i>Syntrophomonas</i>	
<i>Alphaproteobacteria</i> (32)		
<i>Asticcaulus</i>	<i>Azospirillum</i>	<i>Beijerinckia</i>
<i>Bradyrhizobium</i>	<i>Brevundimonas</i>	<i>Caulobacter</i>
<i>Chelatococcus</i>	<i>Defluviicoccus</i>	<i>Hyphomicrobium</i>
<i>Labrenzia</i>	<i>Magnetospirillum</i>	<i>Mesorhizobium</i>
<i>Methylococcus</i>	<i>Methylobacterium</i>	<i>Methylocystis</i>
	(<i>Protomonas</i>)	
<i>Methylosinus</i>	<i>Mycoplana</i>	<i>Nitrobacter</i>
<i>Novosphingobium</i>	<i>Oligotropha</i>	<i>Paracoccus</i>
<i>Pedomicrobium</i>	<i>Rhizobium</i>	<i>Rhodobacter</i>
<i>Rhodopseudomonas</i> ,	<i>Rhodospirillum</i>	<i>Ruegeria</i>
<i>Sinorhizobium</i> (<i>Ensifer</i>)	<i>Sphingomonas</i>	<i>Sphingopyxis</i>
<i>Stella</i>	<i>Xanthobacter</i>	
<i>Betaproteobacteria</i> (31)		
<i>Accumulibacter</i>	<i>Acidovorax</i>	<i>Alcaligenes</i>
		(<i>Azohydromonas</i>)
<i>Aquaspirillum</i>	<i>Aromatoleum</i>	<i>Brachymonas</i>
<i>Burkholderia</i>	<i>Caldimonas</i>	<i>Chromobacterium</i>
<i>Comamonas</i>	<i>Cupriavidus</i> (<i>Ralstonia</i>)	<i>Dechloromonas</i>
<i>Delftia</i>	<i>Derxia</i>	<i>Herbaspirillum</i>
<i>Hydrogenophaga</i> ,	<i>Ideonella</i>	<i>Janthinobacterium</i>
<i>Lampropedia</i>	<i>Leptothrix</i>	<i>Methylibium</i>
<i>Pelomonas</i> ,	<i>Roseateles</i>	<i>Rubrivivax</i>
<i>Schlegelella</i>	<i>Sphaerotilus</i>	<i>Spirillum</i>
(<i>Caenibacterium</i>)		
<i>Thauera</i>	<i>Thiobacillus</i>	<i>Variovorax</i>
<i>Zoogloea</i>		
<i>Deltaproteobacteria</i> (2)		
<i>Desulfobacterium</i>	<i>Desulfococcus</i>	

Gammaproteobacteria (44)

<i>Acidithiobacillus</i> (<i>Ferrobacillus</i>)	<i>Acinetobacter</i>	<i>Actinobacillus</i>
<i>Aeromonas</i>	<i>Alcanivorax</i> (<i>Fundibacter</i>)	<i>Alkalilimnicola</i>
<i>Allochromatium</i>	<i>Amphritea</i>	<i>Azomonas</i>
<i>Azotobacter</i> (<i>Axobacter</i>)	<i>Beggiatoa</i>	<i>Chromatium</i>
<i>Chromohalobacter</i>	<i>Cobetia</i>	<i>Competibacter</i>
<i>Ectothiorhodospira</i>	<i>Erwinia</i>	<i>Escherichia</i> (recombinant)
<i>Haemophilus</i>	<i>Hahella</i>	<i>Halomonas</i>
<i>Halorhodospira</i>	<i>Klebsiella</i> (recombinant)	<i>Kushneria</i>
<i>Lamprocystis</i>	<i>Legionella</i>	<i>Marinobacter</i>
<i>Marinospirillum</i>	<i>Methylomonas</i> (<i>Methanomonas</i>)	<i>Moraxella</i>
<i>Neptunomonas</i>	<i>Nitrococcus</i>	<i>Oceanospirillum</i>
<i>Photobacterium</i>	<i>Plasticicumulans</i>	<i>Pseudomonas</i>
<i>Saccharophagus</i>	<i>Thiocapsa</i>	<i>Thiococcus</i>
<i>Thiocystis</i> (<i>Thiosphaera</i>)	<i>Thiodictyon</i>	<i>Thiopedia</i>
<i>Vibrio</i> (<i>Beneckeia</i>)	<i>Zobellella</i>	

(Source: Koller *et al.*, 2013)

Table 2.2: Summary of PHA-producing genera from the domain Archaea

Euryarchaeota (15)

<i>Haloarcula</i>	<i>Halobacterium</i>	<i>Halobiforma</i>
<i>Halococcus</i>	<i>Haloferax</i>	<i>Halogeometricum</i>
<i>Halopiger</i>	<i>Haloquadratum</i>	<i>Halorhabdus</i>
<i>Halorubrum</i>	<i>Haloterrigena</i>	<i>Natrialba</i>
<i>Natrinema</i>	<i>Natronobacterium</i>	<i>Natronococcus</i>

Thaumarchaeota (2)

<i>Cenarchaeum</i>	<i>Nitrosopumilus</i>
--------------------	-----------------------

(Source: Koller *et al.*, 2013)

PHA producers are commonly identified via simple and rapid phenotypic screening using viable colony staining method. Lipophilic dyes such as Sudan Black B (Schlegel *et al.*, 1970), Nile Blue A (Ostle and Holt 1982) and Nile Red (Gorenflo *et al.*, 1999; Spiekermann *et al.*, 1999) can bind to the PHA granules. However, these dyes could also bind to lipids and fatty materials (Burdon, 1946; Spiekermann *et al.*, 1999). The presence of PHA granules inside the cells could also be observed using phase contrast microscope (Dawes and Senior, 1972; Sudesh *et al.*, 2000).

2.3 PHA biosynthesis pathways and PHA synthase (PhaC)

The central PHA biosynthesis pathway consists of three basic enzymatic steps which will convert acetyl coenzyme A (acetyl-CoA) intermediate to PHB. In the first step, condensation of two molecules of acetyl-CoA to acetoacetyl-CoA is catalyzed by β -ketothiolase (PhaA). This is followed by the reduction of acetoacetyl-CoA to *R*-3-hydroxybutyryl-CoA by NADPH-dependent acetoacetyl-CoA reductase (PhaB). Finally, the polymerization of the *R*-3-hydroxybutyryl-CoAs into PHB is catalyzed by PHA synthase (PhaC) (Anderson and Dawes, 1990). The genes for these three important enzymes were successfully cloned during the late 1980s (Schubert *et al.*, 1988; Slater *et al.*, 1988; Peoples and Sinskey, 1989).

In microorganisms, substrates or monomers for the PHA synthase could be supplied from various metabolic pathways such as fatty acid β -oxidation, fatty acid *de novo* biosynthesis and citrate acid cycle (Madison and Huisman, 1999; Steinbüchel, 2001; Taguchi *et al.*, 2002) (Figure 2.3 and Table 2.3). Monomers of MCL-PHA such as 3-hydroxyhexanoate (3HHx) and 3-hydroxyheptanoate (3HHp) can be channeled from the fatty acid β -oxidation pathway to PHA synthase via the catalysis reaction of *R*-specific enoyl-CoA hydratase (PhaJ), which convert enoyl-

CoA intermediates to (*R*)-3-hydroxyacyl-CoA. In the same pathway, epimerase and 3-ketoacyl-CoA reductase (FabG) can convert (*S*)-3-hydroxyacyl-CoA and 3-ketoacyl-CoA intermediates to (*R*)-3-hydroxyacyl-CoA, respectively (Eggink *et al.*, 1992; Madison and Huisman, 1999; Taguchi *et al.*, 1999).

Besides, MCL-PHA monomers could also be supplied from the fatty acid *de novo* biosynthesis pathway, in which 3-hydroxyacyl-ACP-CoA transferase (PhaG) can convert (*R*)-3-hydroxyacyl-ACP intermediates to (*R*)-3-hydroxyacyl-CoA (Eggink *et al.*, 1992; Madison and Huisman, 1999). Meanwhile, 4HB monomer can be supplied from the citric acid or tricarboxylic acid (TCA) cycle through the conversion of succinyl-CoA to succinic semialdehyde and then 4-hydroxybutyrate. This 4-hydroxybutyrate intermediate can be converted to 4-hydroxybutyrate-CoA via the catalysis reaction of 4-hydroxybutyrate-CoA:CoA transferase (OrfZ) (Valentin and Dennis, 1997; Zhou *et al.*, 2012).

In some cases, supplementation of precursors or structurally related substrates as exogenous carbon sources to the microorganisms could produce PHAs with unusual copolymers but this is also dependent on the substrate specificity of the PHA synthase (Sudesh and Doi, 2005). For instance, (i) sodium propionate or sodium valerate could be added as precursors for the synthesis of poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) (Lee *et al.*, 2008); (ii) γ -butyrolactone, 1,4-butanediol or sodium 4-hydroxybutyrate could be added as precursors for the synthesis of poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate) (Lee *et al.*, 2004); (iii) isocaproic acid could be added as precursors for the synthesis of poly(3-hydroxybutyrate-*co*-3-hydroxy-4-methylvalerate) (Lau *et al.*, 2010); and (iv) 3-mercaptopropionic acid or 3,3-thiodipropionic acid could be added as precursors for the synthesis of poly(3-hydroxybutyrate-*co*-3-mercaptopropionate) (Lütke-Eversloh *et al.*, 2002).

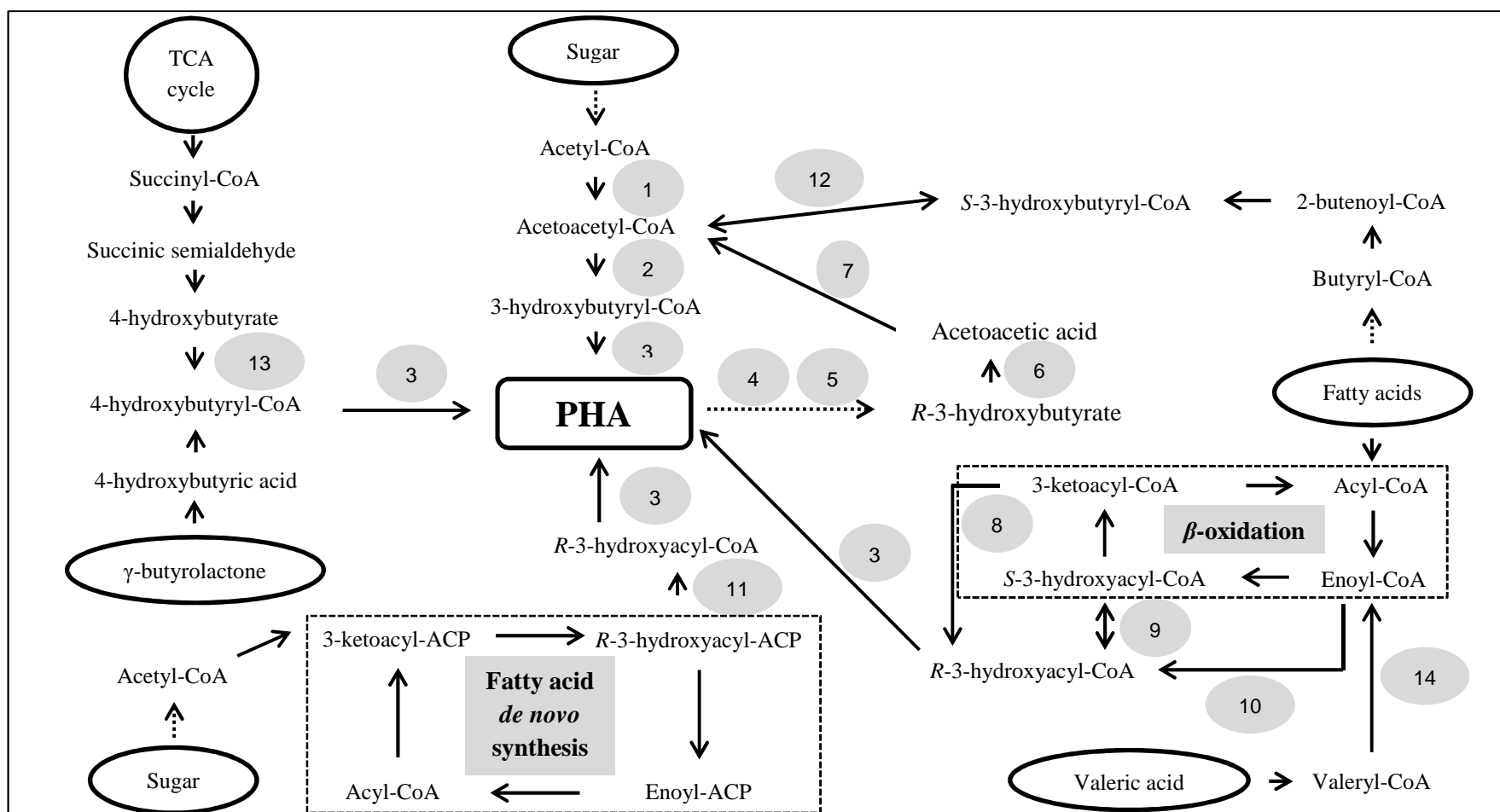


Figure 2.3: Major PHA biosynthesis and biodegradation pathways in bacteria. Major enzymes are indicated by the numbering in grey circles and descriptions are shown in Table 2.3. (Modified from Chen, 2009)

Table 2.3: Major enzymes involved in the PHA biosynthesis and biodegradation pathways

No.	Abbreviation	Enzymes
1	PhaA	β -ketothiolase
2	PhaB	NADPH dependent acetoacetyl-CoA reductase
3	PhaC	PHA synthase
4	PhaZ	PHA depolymerase
5	-	Dimer hydrolase
6	-	(<i>R</i>)-3-hydroxybutyrate dehydrogenase
7	-	Acetoacetyl-CoA synthetase
8	FabG	3-ketoacyl-CoA reductase
9	-	Epimerase
10	-	(<i>R</i>)-enoyl-CoA hydratase
11	PhaG	3-hydroxyacyl-ACP-CoA transferase
12	-	NADH-dependent acetoacetyl-CoA reductase
13	OrfZ	4-hydroxybutyrate-CoA:CoA transferase
14	-	Acyl-CoA dehydrogenase

Among the PHA biosynthesis and biodegradation genes, PHA synthase has received the most attention because it is the key enzyme in the PHA biosynthesis process. It has a partial Enzyme Commission number [EC: 2.3.1.-], in which PhaC belongs to Transferases (main class EC 2), Acyl transferases (subclass EC 2.3) and other than amino-acyl groups (sub-subclass EC 2.3.1). The unknown serial number “-” of PhaC is because of the catalytic activity of the protein is not exactly known or the protein catalyzes a reaction that is known but not yet included in the International Union of Biochemistry and Molecular Biology (IUBMB) EC list (UniProt Consortium, 2010). A recent study demonstrated that PHA synthase of *Bacillus megaterium* confer depolymerase activity via alcoholic cleavage of PHA chains (Hyakutake *et al.*, 2015).

In general, PhaC catalyzes the polymerization reaction of the hydroxyacyl (HA) moiety in HA-CoA to PHA, with the concomitant release of CoA (Sudesh *et al.*, 2000; Stubbe and Tian 2003; Rehm, 2003). Initially, three classes of PHA synthase (Class I to III) were proposed by Rehm and Steinbüchel (1999) based on the amino acid sequence, *in vivo* substrate specificity and subunit composition. This classification is later revised with the addition of Class IV PHA synthase by Rehm (2003) (Figure 2.4). Class IV PHA synthase was discovered from the *Bacillus megaterium* in 1999 (McCool and Cannon, 1999).

Class I and II PHA synthases contain only one type of subunit (PhaC). Class I PHA synthase comprises of a single PhaC subunit which has molecular mass around 61 to 73 kDa. Class I PHA synthase is represented by *Cupriavidus necator* and can produce short chain length PHA. Class II PHA synthase comprise of two PhaC subunits which have molecular masses around 60 to 65 kDa. Class II PHA synthase is represented by *Pseudomonas aeruginosa* and can produce medium chain length PHA. Meanwhile, Class III and IV PHA synthases contain two different types of subunits. Class III PHA synthase comprises of one PhaC subunit (~ 40 kDa) and one PhaE subunit (~ 40 kDa). Class III PHA synthase is represented by *Allochromatium vinosum* and can produce short chain length PHA. Class IV PHA synthase comprises of one PhaC subunit (~ 40 kDa) and one PhaR subunit (~ 22 kDa). Class IV PHA synthase is represented by *B. megaterium* and can produce short chain length PHA.

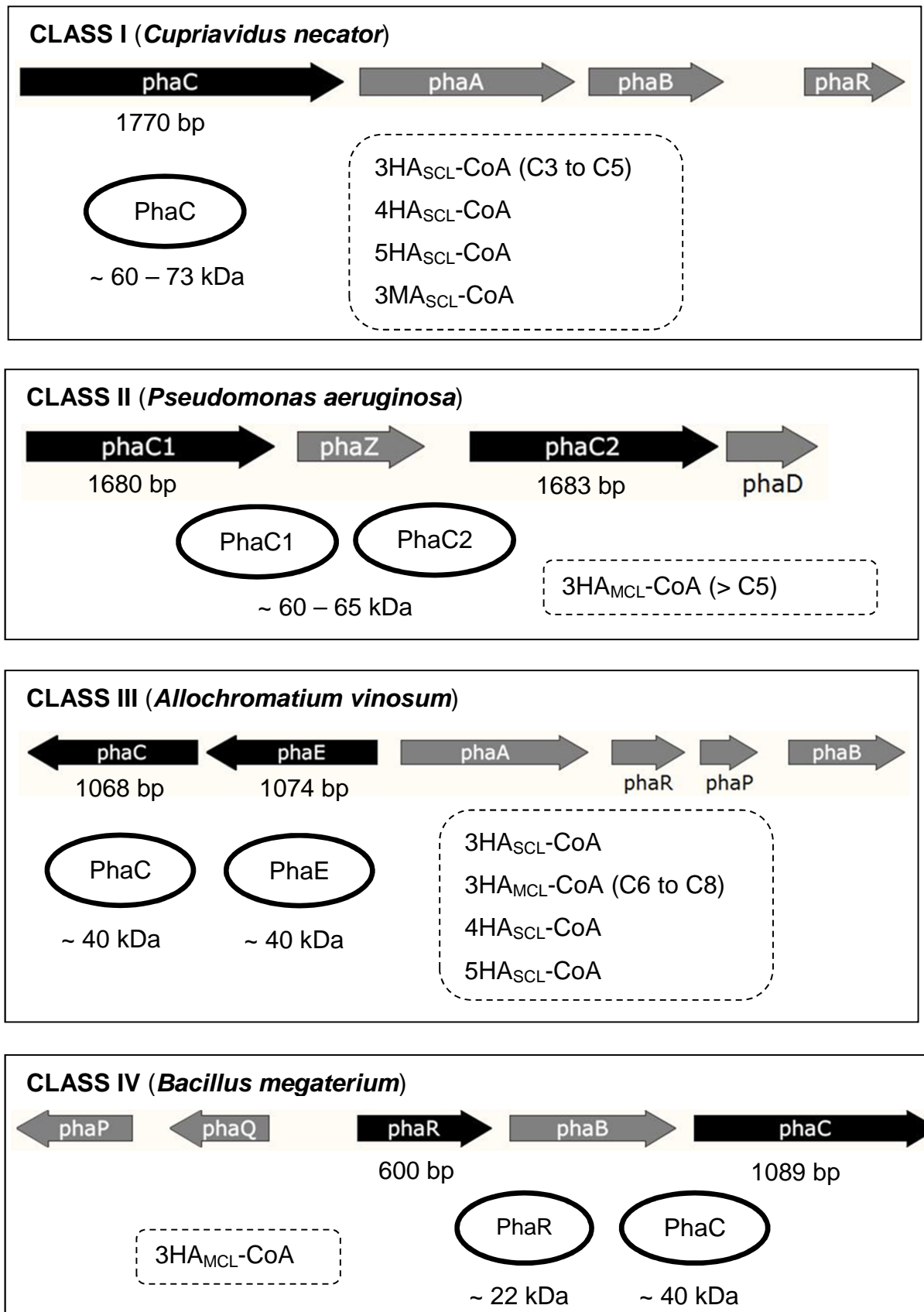


Figure 2.4: Classification of PHA synthases (modified from Rehm *et al.*, 2003).

Interestingly, PhaE and PhaR subunits have no similarity to PhaC subunit. Multiple sequence alignment of the PHA synthase protein sequences of the PhaC subunit show the presence of six conserved blocks and eight highly conserved amino acid residues. Besides, a lipase-box-like motif “G-X-[S/C]-X-G” is present in all the PHA synthases, where the serine residue in lipase is replaced with cysteine residue in PHA synthase. A catalytic triad comprising of Cys-319, His-508 and Asp-480 (positions are based on *C. necator* PhaC1) is required for catalytic activity (Rehm and Steinbüchel, 1999; Qi and Rehm, 2001; Rehm, 2003).

The first genotypic detection method of PHA synthase was developed by Sheu *et al.* (2000) using degenerate primer sets (phaCF1, phaCF2 and phaCF4) to amplify partial Class I and II PHA synthase genes from six Gram-negative bacterial genera (*Alcaligenes*, *Comamonas*, *Hydrogenophaga*, *Pseudomonas*, *Ralstonia* and *Sphaerotilus*). Romo *et al.* (2007) improved these previous primer sets and their newly designed primers (G-D, G-1R and G-2R) are able to amplify partial Class I and II PHA synthase genes from nine Gram-negative bacterial genera (*Aeromonas*, *Acinetobacter*, *Azospirillum*, *Azotobacter*, *Burkholderia*, *Rhizobium*, *Pseudomonas*, *Ralstonia* and a member of *Enterobacteriaceae* family) (Table 2.4 and Figure 2.5).

Primers for Class II PHA synthase was developed by Solaiman *et al* (2000) and they are able to amplify both *phaC1* and *phaC2* genes (partial) from *Pseudomonas* MCL-PHA producers. The complete open reading frame (ORF) of *phaC1* and *phaC2* genes could be amplified from most *Pseudomonas* strains belonging to γ subdivision *Proteobacteria* (rRNA group I) using primer sets designed by Zhang *et al.* (2001). A degenerate primer set developed by Kung *et al.* (2007) is able to amplify partial *phaC1* gene from MCL-PHA producers such as genus *Acinetobacter*, *Aeromonas*, *Exiguobacterium* and *Pseudomonas*.

Table 2.4: Primers targeting on various classes of PHA synthase

Class	Name and sequence (5' to 3')	References
I	phaCI730F (731-750, <i>phaCI</i>^a) CGCCCTGCATCAACAAGTTC phaCI1218R (1198-1218, <i>phaCI</i>^a) GTAGTTCCAGACCAGGTCGTT	Pärnänen <i>et al.</i> 2015
I and II	phaCF1 (739-764, <i>phaCI</i>^a) ATCAACAA(GGG/A)T(TT/A)CTAC(AA/G)TC(CC/T)T(CC/G)GACCT phaCF2 (814-839, <i>phaCI</i>^a) GT(CCC/GG)TTC(GGG/AA)T(GGG/CC)(AAA/GG)T(C C/G)(TT/A)(CCC/GG)CTGGCGCAACCC phaCF4 (1210-1235, <i>phaCI</i>^a) AGGTAGTTGT(TT/C)GAC(CCC/GG)(AAA/CC)(AAA/ CC)(GGG/A)TAG(TTT/G)TCCA	Sheu <i>et al.</i> , 2000
I and II	G-D (727-748, <i>phaCI</i>^a) GTGCCGCC(GC)(CT)(AG)(GC)ATCAACAAGT G-1R (1258-1278, <i>phaCI</i>^a) GTTCCAG(AT)ACAG(GC)A(GT)(AG)TCGAA G-2R (1198-1218, <i>phaCI</i>^a) GTAGTTCCA(GC)A(CT)CAGGTCGTT	Romo <i>et al.</i> , 2007
II	I-179L (669-696, <i>phaCI</i>^b) ACAGATCAACAAGTTCTACATCTTCGAC I-179R (1177-1206, <i>phaC2</i>^b) GGTGTGTCGTTGTTCCAGTAGAGGATGTC	Solaiman <i>et al.</i> , 2000
II	ORF1 forward (228-251, ORF2^b) CCA(C/T)GACAGCGGCCTGTTACCTG <i>phaZ</i> reverse primer (640-663^b) GTCGTCGTC(A/G)CCGGCCAGCACCAG <i>phaZ</i> forward primer (640-663^b) CTGGTGCTGGCCGG(C/T)GACGACGAC <i>phaD</i> reverse primer (236-260^b) TCGACGATCAGGTGCAGGAACAGCC	Zhang <i>et al.</i> , 2001

II	phal-1 (forward) (283-308, <i>phaCI</i>^b) CARACNTAYYTNGCNTGGMGNAARGA phal-2(reverse) (1123-1148, <i>phaCI</i>^b) TARTTRTTNACCCARTARTTCCADAT	Kung <i>et al.</i> , 2007
II	phaCII36F (38-57, <i>phaCI</i>^b) GAGCGAAAAACAGTACGCCA phaCII1056R (1139-1158, <i>phaCI</i>^b) CATCGGTGGGTAGTTCTGGT	Pärnänen <i>et al.</i> 2015
III	P1 (313-330, <i>phaC</i>^c) ATNGA(CT)TGGGGNTA(CT) CCN P2 (733-750, <i>phaC</i>^c) (AG)AA(AGT)ATCCA(CT)TT(CT)TCCAT	Hai <i>et al.</i> , 2004
III	codehopEF (266-290, <i>phaE</i>^d) CGACCGAGTTCCGCGAYATHTGGYT codehopER (475-497, <i>phaE</i>^d) GCGTGCTGGCGGCKYTCNAVYTC codehopCF (133-161, <i>phaC</i>^d) ACCGACGTCGTCTACAAGGARAAYAARYT codehopCR (388-412, <i>phaC</i>^d) GGTCGCGGACGACGTCNACRCARTT	Han <i>et al.</i> , 2010
IV	B1F (333-352, <i>phaC</i>^e) AACTCCTGGGCTTGAAGACA B1R (912-931, <i>phaC</i>^e) TCGCAATATGATCACGGCTA B2R (692-711, <i>phaC</i>^e) ACGGTCCACCAACGTTACAT	Shamala <i>et al.</i> , 2003
IV	phaCIV9F (9-28, <i>phaC</i>^e) TCCTTACGTGCAAGAGTGGG phaCIV921R (902-921, <i>phaC</i>^e) ATCACGGCTAGCAGCAATGT phaCIII110F (110-129, <i>phaC</i>^f) CAGAGCCGCAAGTCGGATTA	Pärnänen <i>et al.</i> 2015