CHARACTERIZATION OF TWO POLYHYDROXYALKANOATE SYNTHASE GENES (*phaC*) OF *Cupriavidus malaysiensis* USMAA1020^{T} ISOLATED FROM KULIM LAKE, KEDAH, MALAYSIA

NUR ASILLA HANI BINTI SHAFIE

UNIVERSITI SAINS MALAYSIA

2021

CHARACTERIZATION OF TWO POLYHYDROXYALKANOATE SYNTHASE GENES (*phaC*) OF *Cupriavidus malaysiensis* USMAA1020^{T} ISOLATED FROM KULIM LAKE, KEDAH, MALAYSIA

by

NUR ASILLA HANI BINTI SHAFIE

Thesis submitted in fulfillment of the requirement
for the degree of
Master of Science

September 2021

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to my supervisor, Prof. Dr. Amirul Al- Ashraf for his guidance and support throughout my study. I am forever grateful for the opportunity given to me and giving me space to learn and improve myself. I would also like to thank Dr. Lau Nyok Sean, Dr. Huong Kai Hee and Dr. Amira for their guidance throughout this project.

Thirdly, I would like to express my appreciation to all my seniors and friends, Dr. Ira, Dr. Afifah, Azura, Yasmin, Hafizah, Aidda, Fatin, and Hanim. Also I would like to extend this appreciation to my labmates at CCB for all their support and guidance, and to all those supportive discussions and ideas whenever I encountered a set back on the project.

Fourthly, I would like to thank everyone who had helped directly or indirectly for this project including the staffs of CCB. Besides that, I would like to extend my gratefulness to my loved ones which are my family Mr. Shafie, Mrs. Yusnani, Mrs. Paridah, Mr. Marek, Nani Ilyana, Muhd Shafeeq, Nur Aliaa, Muhd Aliff, Hayati, Fahmi, Rasyidag, Zulhelmi and also Laila who are always there when I need them, giving me support and encouragements throughout my study. Last but not least, I would like to express my gratitude to MyBrain15 scholarship by the Ministry of Higher Education for the financial aid to ease my study.

TABLE OF CONTENT

CHAPTER 3 MATERIALS AND METHODS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

%	Percentage
x g	Times gravity
β	Beta
γ	Gamma
ºC	Degree celcius
µg	Microgram
µL	Microliter
µM	Micromole
3HV	3-hydroxyvalerate
4HB	4-hydroxybutyrate
ANI	Average nucleotide identity
BDPs	Biodegradable polymers
BLAST	Basic local alignment search tool
BLASTn	Basic local alignment search tool for nucleotide
bp	Base pair
CDS	Protein coding sequence
CLR	Continuous long read
CME	Caprylic methyl ester
CoA	Coenzyme A
COG	Cluster of Orthologous Groups

CRISPR	Clustered regularly interspaced short palindromic repeats
DCW	Dry cell weight
DDH	DNA-DNA hybridization
DNA	Deoxyribonucleic acid
dNTPs	Dideoxy-nucleotides
dsDNA	Double-stranded deoxyribonucleic acid
FDA	Food and Drug Administration
g	Gram
g/L	Gram per liter
Gb	Gigabase
GC	Gas chromatography
gDNA	Genomic deoxyribonucleic acid
h	Hour
HA	Hydroxyapatite
HCL	Hydrocloric acid
HGAP	Hierarchical genome assembly process
ICI	Imperial Chemical Industry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobase
kDa	Kilodalton
Kg	Kilogram
KOH	Potassium hydroxide
LB	Luria bertani medium

LCB	Locally collinear block
M	Molar
Mb	Megabase
mcl-PHA	Medium-chain length polyhydroxyalkanoate
mg	Miligram
min	Minute
mL	mililiter
mm	Milimeter
MM	Mineral salts medium
mol%	Mol percent
MPa	Megapascal
NCBI	National centre for biotechnology information
ng	Nanogram
NGS	Next generation sequencing
nm	Nanometer
NR	Nutrient rich medium
OD	Optical density
ORF	Open reading frame
oriT	Origin of transfer
P(3HB)	Poly(3-hydroxybutyrate)
P(3HB-*co*-3HV)	Poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate)
P(3HB-*co*-3HHx)	Poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate)
P(3HB-*co*-4HB)	Poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate)

PHA	Polyhydroxyalkanoate
PHB	Polyhydroxybutyrate
PhaA	*β*-ketothiolase
PhaB	Acetoacetyl-CoA reductase
PhaC	Polyhydroxyalkanoate synthase
$phaC_{RE}$	Polyhydroxyalkanoate synthase *Ralstonia eutropha*
PhaF	Polyhydoxyalkanoate granule-associated protein
PhaG	3-hydroxyacyl-acyl carrier protein-coenzyme A transferase
PhaJ	Enoyl-CoA hydratase
PhaZ	PHA depolymerase
PBS	Polybutylene succinate
PCR	Polymerase chain reaction
PLA	Poly(lactic acid)
psi	Pounds per square inch
RAST	Rapid annotation using subsystem technology
RBC	Real Biotech Corporation
RNA	Ribonucleic acid
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
s	Second
scl-PHA	Short-chain length polyhydroxyalkanoate
SMRT	Single molecule real time sequencing
TBE	Tris-borate EDTA

TMHMM	Transmembrane helices based on hidden Markov Model
tra	Transfer gene
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume
WGS	Whole genome sequencing
wt%	Dry weight percent
ZMW	Zero-mode waveguides

PENCIRIAN DUA GEN SINTASE POLIHIDROKSIALKANOAT (*phaC*) *Cupriavidus malaysiensis* USMAA1020^T YANG DIPENCILKAN DARIPADA TASIK KULIM, KEDAH, MALAYSIA

ABSTRAK

Strains *Cupriavidus malaysensis* USMAA1020^T, USMAA2-4 dan USMAHM13 telah dipencilkan daripada pelbagai persekitaran semulajadi di Semenanjung Malaysia. USMAA1020^T telah dipencilkan daripada enap cemar di Tasik Kulim, Kedah. Manakala, USMAA2-4 telah dipencilkan daripada tanah di Sungai Pinang, Pulau Pinang sedangkan USMAHM13 telah dipencilkan daripada enap cemar sawah padi di Sungai Manik, Perak. Kesamaan jujukan gen 16S rRNA di antara USMAA1020^T, USMAA2-4, dan USMAHM13 adalah 99%. Berdasarkan analisis filogenetik, ketiga-tiga strain telah membentuk satu kluster dalam pokok filogenetik yang dibina apabila menggunakan dua kaedah iaitu kebolehjadian maksimum dan penyambungan bersebelahan. Sehingga kini, terdapat banyak strain *Cupriavidus* yang telah dicirikan dan dikenalpasti dilaporkan tetapi maklumat tentang penjujukan seluruh genom untuk strain *Cupriavidus* adalah terhad terutama sekali strain yang mempunyai dua gen sintase polihidroksialkanoat (*phaC*) yang berfungsi. Dalam kajian ini, analisis penjujukan keseluruhan genom untuk ketiga- tiga strain yang berkait rapat antara satu sama lain tidak kira perbezaan persekitaran yang dipencilkan. Penjujukan keseluruhan genom telah dilakukan dengan menggunakan platform PacBio RS II. Genom yang dijujuk mempunyai saiz dalam lingkungan 7.82 hingga 8.70 Mb secara relatifnya lebih besar daripada strain *C. necator* H16^T (7.42 Mb).

Dua gen polihidroksialkanoat sintase (phaC) telah dijumpai di dalam annotasi genom. Oleh itu, strain USMAA1020^{T} dicirikan dengan lebih lanjut dengan memasukkan kedua-dua gen phaC secara individu ke dalam mutan negatif *Cupriavidus necator* PHB^{-}4 untuk mengetahui fungsinya dalam penghasilan PHA. Strain USMAA1020^{T} dipilih kerana strain ini yang pertama dipencilkan antara ketiga-tiga strain dan juga telah dijadikan sebagai rujukan untuk dua strain yang lain. Menariknya, polihidroksialkanoat granul protein yang berkaitan hanya dijumpai pada strain USMAA1020^{T} tetapi tidak pada strain USMAA2-4, USMAHM13 dan juga H16^{T}. Bagi menggalakkan akumulasi PHA yang mengandungi monomer 4HB, γ-buyrolactone telah digunakan sebagai sumber karbon tunggal, asid oleik digunakan untuk menghasilkan monomer 3HB dan pentanol digunakan sebagai ko-karbon untuk menghasilkan monomer 3HV. Akumulasi PHA menggunakan sumber karbon yang berbeza menunjukkan bahawa kedua-dua gen phaC adalah gen aktif yang berfungsi. Penemuan ini menunjukkan bahawa transforman *C. necator* PHB^{-}4 yang mengandungi phaC1 dapat menghasilkan poli(3-hidroksibutirat-*co*-4-hidroksibutirat) dengan mmonomer 4-hidroksibutirat (4HB) sehingga 83% dan PHA dengan 26.5 wt% apabila dibekalkan dengan γ-buyrolactone sebagai sumber karbon. Manakala transforman *C. necator* PHB^{-}4 yang mengandungi phaC2 tidak dapat menghasilkan 4HB monomer dan hanya dapat menghasilkan 2.3 wt% PHA apabila dibekalkan dengan γ-buyrolactone. Disamping itu, transforman phaC2 dapat menghasilkan PHA sehingga 100 mol% 3HB monomer dan 41.9 wt% apabila asid oleik dibekalkan sebagai sumber karbon tunggal manakala transforman phaC1 hasilkan PHA 25.8 wt%. Penemuan ini menunjukkan bahawa kedua-dua gen phaC adalah gen aktif yang berfungsi bergantung dengan sumber karbon yang dibekalkan.

CHARACTERIZATION OF TWO POLYHYDROXYALKANOATE SYNTHASE GENES (*phaC*) OF *Cupriavidus malaysiensis* USMAA1020^T ISOLATED FROM KULIM LAKE, KEDAH, MALAYSIA

ABSTRACT

Cupriavidus malaysiensis strains USMAA1020^T, USMAA2-4 and USMAHM13 were isolated from various natural environments of Peninsular Malaysia. USMAA1020^T was isolated from sludge in Kulim Lake, Kedah. Meanwhile, USMAA2-4 was isolated from soil of Sungai Pinang, Penang whereas USMAHM13 was isolated from sludge of a paddy field in Sg. Manik, Perak. The 16S rRNA gene sequence similarity between USMAA1020^T, USMAA2-4, and USMAHM13 was 99%. In phylogenetic analysis, these strains formed a cluster in phylogenetic tree constructed when using both maximum-likelihood and neighbour joining method. Till date, there were many reported *Cupriavidus* strains which have been characterized and identified but only limited information available on the whole genome sequence of *Cupriavidus* strains especially strains that contain two functional polyhydroxyalkanoate synthase genes (*phaC*). In this study, we analyzed the whole genome sequence of these strains which are closely related to each other regardless the differences of isolating environments. The whole genome sequences were done using the PacBio RS II platform. The genomes sequenced (7.82 to 8.70 Mb) were comparatively larger than the type strain *Cupriavidus necator* H16^T (7.42 Mb). Two polyhydroxyalkanoates synthase genes (*phaC*) were found in the annotated genomes of all the three strains. Hence, strain USMAA1020^T was further

characterized by inserting both *phaC* gene individually in a negative mutant strain *Cupriavidus necator* PHB^{-}4 for its functionality in producing PHA. This strain was chosen as this was the first strain isolated amongst the three strain and also was made as a reference for thee other two strains. Interestingly, polyhydroxyalkanoate granule-associated protein (PhaF) was found only in strain USMAA1020^{T} but not in strains USMAA2-4, USMAHM13 and H16^{T}. To promote the accumulation of PHA containing 4HB monomer, γ-buyrolactone was supplied as the sole carbon source, oleic acid was provided to produce 3HB monomer and pentanol was supplied as co-carbon source to produce 3HV monomer. Accumulation of PHA using different carbon sources showed that both transformant consisting *phaC1* and *phaC2* of USMAA1020 were functionally active. These finding demonstrated that the transformant *C. necator* PHB^{-}4 harbouring *phaC1* was capable of accumulating poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate) with 4-hydroxybutyrate (4HB) monomer of up to 83 mol% and 26.5 wt% of PHA content when using γ-butyrolactone as sole carbon source. Meanwhile, the transformant *C. necator* PHB^{-}4 harbouring *phaC2* was not able to produced 4HB monomer and produce only 2.3 wt% of PHA content when γ-butyrolactone was used as sole carbon source. In contrast, *phaC2* transformant can produce up to 100 mol% of 3HB monomer and 41.9 wt% of PHA content when oleic acid was provided as sole carbon source whereas transformant *phaC1* produces only 25.8 wt% of PHA. These findings shows that both *phaC* and *phaC2* were functionally active depending on the type of carbon sources being used.

CHAPTER 1

INTRODUCTION

1.1 Background of Research

Southeast Asia is known as one of the major hot spots that are rich in biodiversity (Hughes, 2017). Biodiversity a term introduced by Wilson in 1992 and is a blend of the phrase biological diversity. It is complex where at times believe as a measure of what we want to retain, but it is also irregularly believes as a tool: an evaluation of an instrumentally significant dimension of biological systems (Maclaurin & Sterelny, 2008). In the past, agricultural practices have failed to promote healthy populations of microorganisms, limiting production yields and threatening sustainability (Johns, 2017). These massive activities have impacted the below-ground diversity in the tropics of Southeast Asia.

Bacteria play a major role in soil processes and they constitute a major portion of diversity in soils (Miyashita *et al.*, 2013). Microbial diversity in soil is deemed to be important as wide variety of microorganisms is involved in maintenance of soil health and quality (Garbeva *et al.*, 2004; Meliani *et al.*, 2012). The factors which strongly correlate to the bacterial community and diversity are the soil properties that include soil pH, total carbon and C/N ratio. As reported by Tripathi et al. in year 2012, the soil community variation was dominated by Alphaproteobacteria, Beta/Gammaproteobacteria, Acidobacteria, and Actinobacteria which significantly correlate with soil pH.

Cupriavidus is a bacterial genus that belongs to the Betaproteobacteria class and is characterised as Gram-negative, motile and rod-shaped bacteria. A few *Cupriavidus* species were reported to have been isolated from various surroundings.

Cupriavidus necator, which was isolated from the soil of University Park, PA, USA, is highly resistant to copper (Makkar & Casida, 1987). *Cupriavidus gilardii* was isolated from a whirl pool (Coenye *et al.*, 1999). To date other reported species were *Cupriavidus alkaliphilus* (Estrada-De Los Santos *et al.*, 2012), *Cupriavidus basilensis* (Vandamme & Coenye, 2004), *Cupriavidus campinensis* (Goris *et al.*, 2001), *Cupriavidus laharis* (Sato *et al.*, 2006), *Cupriavidus metallidurans* (Goris *et al.*, 2001), *Cupriavidus nantongensis* (Sun *et al.*, 2016), *Cupriavidus numazuensis* (Kageyama *et al.*, 2005), *Cupriavidus oxalaticus* (Şahin *et al.*, 2000), *Cupriavidus pampae* (Cuadrado *et al.*, 2010), *Cupriavidus pauculus* (Vandamme & Coenye, 2004), *Cupriavidus pinatubonensis* (Sato *et al.*, 2006), *Cupriavidus plantarum* (Estrada-de los Santos *et al.*, 2014), *Cupriavidus respiraculi* (Vandamme & Coenye, 2004), *Cupriavidus taiwanensis* (Chen *et al.*, 2001), and *Cupriavidus yeoncheonensis* (Singh *et al.*, 2015).

Three *Cupriavidus* strains were isolated at different locations of Kulim lake in Kedah, Sungai Pinang in Penang and Sungai Manik in Perak respectively (Amirul *et al.*, 2004). These strains were different morphologically where *Cupriavidus malaysiensis* USMAA1020^T had a circular, convex with an entire edge, opaque and beige in colour meanwhile USMAA2-4 had a circular but with a flat with a wrinkled edges, opaque and beige in colour. Whereas USMAHM13 had a glistening, smooth, convex with an entire edge and yellow-pigmented. But their 16S rRNA sequence showed 99% similarity. These strains were reported to have the capability to produce P(3HB-*co*-4HB) copolymer with high 4HB composition of 53 mol% for wild-type USMAA1020^T (Amirul *et al.*, 2008) and up to 95% for transformant USMAA1020^T (Norhafini *et al.*, 2017), 84 mol% for USMAA2-4 (Vigneswari *et al.*, 2009), 43 mol% for USMAHM13 (Ramachandran & Amirul, 2013). To prove that these strains

were different, the whole genomes of these strains were sequenced and compared. Whole genome sequencing is a procedure where the complete DNA sequence of an organism was determined (Ng & Kirkness, 2010). In the early days of this technology, the cost was very high, but over time there has been a rapid decline on the cost for generating genome information (van El *et al.*, 2013). .

In this research, PacBio platform was chosen for microbial whole genome sequencing since longer reads are required in facilitating *de novo* genome assembly and genome finishing (Chin *et al.*, 2013). With the aid of hierarchical genome assembly process (HGAP) which is a tool for de novo genome assembly using PacBio reads, it can generate finished, high-quality genome assemblies in automated workflow. This research focused on the USMAA1020^{T} strain and its two polyhydroxyalkanoate synthase (*phaC*) genes *phaC1* and *phaC2*. Polyhydroxyalkanoates (PHA) are biopolyesters synthesised by microorganisms as intracellular carbon and energy storage compound (Khanna & Srivastava, 2005). The PHA is synthesised as water- insoluble inclusions when microorganisms are exposed to excess carbon sources and limited amounts of nutrients (Chek *et al.*, 2017). The PHA has a broad range of applications in various industries, such as the biomedical sector that includes tissue engineering, tissue bio-implantation, drug delivery, surgery, and wound dressing due to its biocompatible and biodegradable properties (Raza *et al.*, 2018). Polyhydroxyalkanoate synthase (PhaC) is the key enzyme involved in PHA biosynthesis and functions by polymerising monomeric hydroxyalkanoates substrates (Bhubalan *et al.*, 2011). It can affect the properties of PHA in terms of the monomer composition, molecular weights and also material properties (Zou *et al.*, 2017).

Pseudomonas mendocina NK-O1 has been reported that it consists of two *phaC* subunits which are *phaC1* and *phaC2* where strains in this study also possessed two phaC subunits *phaC1* and *phaC2*. The research reported that both subunits have similar substrate specificities but *phaC1* displays higher catalytic rate and produces PHA with lower molecular weight than *phaC2*. The PHA synthases are usually homologs and belong to the same class, but there are a few reported cases of bacterial strains carrying multiple functional PHA synthase homologs that come from different PHA synthase classes (Hein *et al.*, 2002).

Although many studies were published on the heterologous transformation studies of polyhydroxyalkanoate synthase genes (*phaC1* and *phaC2*), but mostly common were reported on the *Pseudomonas* strains such as *Pseudomonas mendocina*, *Pseudomonas putida* and many more. To date, there is only one study on *Cupriavidus* strains that were reported to contain two functional PhaC. Consequently, this study was conducted to determine the functional activity of both *phaC* genes, *phaC1* and *phaC2* in a mutant strain *Cupriavidus necator* PHB^{-}4 by transconjugation. Herein, the study on bacterial strains that possess more than one *phaC* will give more insight on the bacteria genetically and biochemical diversity of the PHA synthases.

1.2 Research objectives

i. To sequence, assemble and annotate whole genomes of three *Cupriavidus malaysiensis* strains USMAA1020^{T}, USMAA2-4 and USMAHM13.

ii. To determine the functional activity of PhaC enzymes from *Cupriavidus malaysiensis* strain USMAA1020^{T} towards PHA.

CHAPTER 2

LITERATURE REVIEW

2.1 Petroleum-based plastics vs Biodegradable plastics

Nowadays the usage of packaging materials such as paper, glass, wood, metals and plastics have increased drastically due to the growth of population. The increased usage of plastics which are widely applied especially in household and also in industries (Prasteen *et al.*, 2018). Excessive usage of plastics from petroleum based has prompted for renewable substitution. The term bioplastics literally alludes to either: 1) plastics that is generated from fossil fuel, such as polybutylene succinate (PBS) or aliphatic plastics that can be utilised as substrate by microorganisms, or 2) bio-based plastic suchlike poly(lactic acid) (PLA) and polyhydroxyalkanoate (PHA) that is synthesised from biomass and renewable resource (Emadian *et al.*, 2017). Environmentally significant drawbacks of petroleum-based plastics are carbon dioxide emissions and extended accumulation period due to its non-biodegradability. Petroleum-based plastics withstands the degradation processes with a very slow kinetics in landfills. On the contrary, the degradation of bioplastics in landfills is at much higher rates. The production of bioplastics has been about 750,000 tons/year, whereas petroleum-derived plastics has been about 200 million tons/year (Gironi & Piemonte, 2011). According to Widdecke et al. (2009), the production growth of bioplastics forthcoming will be exponentially reaching up to 1,000,000 tons/year (Widdecke, 2009).

Bioplastics industry has been expanding over the last few years. In 2014, whereby annually 1.7 million tons were produced globally. This figure is forecasted to rise to 7.8 million tons of production in 2019 (Gironi & Piemonte, 2011). Despite

being an environmentally friendly material, bioplastics also possesses some restrictions, for instance significantly high cost in production and poor mechanical properties. The high production cost could be mitigated by utilising cost efficient renewable resources such as agricultural wastes (Emadian et al., 2017). Biodegradable polymers (BDPs) or biodegradable plastics are polymeric materials which the microorganisms enzyme action plays the predominant mechanism that decomposes into carbon dioxide, methane, water, inorganic compounds or biomass in . This process could be assessed by standardised tests in a specified time frame and considering the availability of disposal condition.

2.2 Polyhydroxyalkanoates (PHA)

2.2.1 PHA properties

Polyhydroxyalkanoates or PHA are microbial polymers that were synthesised when there is excess carbon source and at least one other growth-essential nutrient that is depleted (Sharma *et al.*, 2017). PHA is the only polymer with similar properties to various synthetic thermoplastics, such as polypropylene which were entirely biodegradable and biocompatible (Kung *et al.*, 2007).

PHA differs in properties and chemical compositions as homo- or copolymers depending on the structural variation of PHA monomers (Noda *et al.*, 2005). PHA is comparable to polypropylene, showing great resistance to moisture and harbouring excellent barrier properties to gases where it has the ability to reduce or slow down the passage of gases. Major advantage of PHA is its biodegradability that can be beneficial towards its respective functions and applications. Up to a certain extent this beneficial traits allows PHA to compete against conventional plastics but due to its high in cost, can be a major drawbacks. Besides that, PHA is are insoluble in

water, relatively resistant to hydrolytic attack, resistant to UV, and sinks in water whereas polypropylene floats thus facilitating anaerobic biodegradation in sediments (Kalia *et al.*, 2000).

2.2.2 Types and structures of PHA

There were about 150 different types of PHA had been reported. PHA could be categorized into two classes: short-chain length (scl-PHA) and medium-chain length (mcl-PHA) according to the number of carbons in the side chains. The scl-PHA has less than five carbon atoms while mcl-PHA has 5 to 14 carbon atoms (Li *et al.*, 2016). The monomers of 3-hydroxyvalerate and 3- hydroxybutyrate are examples of scl-PHA, while 3-hydroxydecanoate, 3-octanoate and 3-hydroxyhexanoate are of mcl-PHA (Sharma *et al.*, 2017).

P(3-hydroxybutyrate) is the most extensive, well studied and best described member of PHA (Bugnicourt *et al.*, 2014). The scl-PHA was initially isolated and characterized by Lemoigne in 1926 at the Pasteur Institute in Paris whom bring to light the production of P(3HB) by *Bacillus megaterium* (Shah, 2014). When poly 3-hydroxybutyrate or P(3HB) was first discovered it was considered as the only PHA produced by bacteria as source of energy supply. Up until 1974, Wallen and Rohwedder discovered and identified new PHA materials in chloroform extracts of activated sewage sludge namely 3-hydroxyvalerate and 3-hydroxyhexanoate (Doi, 1990). Polyhydroxybutyrate (PHB) can be characterised as a highly crystalline polymer with relatively poor physical properties as it turns out to be stiffer, more brittle and thermally unstable than polypropylene (Chan *et al.*, 2011). Meanwhile, PHA copolymers are more flexible and less crystalline. Therefore, much efforts had been done for the enhancement of mechanical and thermal properties of P(3HB) by

integrating monomers such as 3-hydroxyvalerate (3HV) (Khosravi-Darani & Bucci, 2015). In industrial, a process drafted by Imperial Chemical Industries (ICI) had used glucose and propionic acid as the precursor carbon source through fed-batch culture of *C. necator* to synthesize poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate), P(3HB-*co*-3HV) (Holmes, 1988; Jain *et al.*, 2010). When compared with P(3HB), P(3HB-*co*-3HV) is less brittle than P(3HB) hence potentially more usable. It is not tough nor stiff which makes it applicable in packaging. It also has lower meting points. In fact, the copolymers become tougher (increased impact strength) and more flexible (reduced Young's Modulus) with the increment of tensile strength when increased the fraction of 3HV. Young's modulus is a measure of the ability of a material to withstand changes in length when under lengthwise tension or compression. Although, copolymer P(3HB-*co*-3HV) possesses better properties than P(3HB), it was discovered that the copolymer is difficult to processed due to low elongation at break and slow crystallisation rate (Khosravi-Darani & Bucci, 2015).

The propyl group in poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate) or P(3HB-*co*-3HHx) is more effective than the ethyl group in P(3HB-co-3HV) for the improvement of the polymer's elongation at break. According to Doi *et al*. (1995), poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate) P(3HB-*co*-3HHx), a random copolymer is produced by *Aeromonas caviae* when fed with alkanoic acids of even carbon numbers ranging from C_{12} to C_{18} or plant oils (Doi *et al.*, 1995). It is also a flexible material which is suited to be used as a film (Wong *et al.*, 2012).

Besides that, poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate), P(3HB-*co*-4HB) is another type of PHA copolymer that shows competent physical properties which were found in brain tissue extracts of rats, pigeons and humans. The first synthetic 4HB in the form of sodium salt was available in the early 1960s (Domb *et al.*, 1998).

Over the years, ongoing studies resulted in the production of P(3HB-*co*-4HB) which acquires a wide range of 4HB content. Carbon sources that also resulted in the incorporation of 4HB units were 4-chlorobutyric acid, 1,4-butanediol, 1,6 hexanediol, 1,8 octanediol, 1,10-decanediol, and 1,12-dodecanediol (Jain *et al.*, 2010). Meanwhile, a study had been described where γ-butyrolactone was fed on *Alcaligenes latus* that obtained 4HB fractions ranging between 4 to 45 mol% subject to the concentration of γ-butyrolactone (Hiramitsu *et al.*, 1993). By varied the monomer composition of 4HB, this copolymer can also be customized to attained a broad range of polymeric materials, from hard crystalline plastics to very elastic rubber. This has gained attention as it could be applicable in the medical and pharmaceutical field since this copolymer displays desirable mechanical properties (Vigneswari *et al.*, 2009).

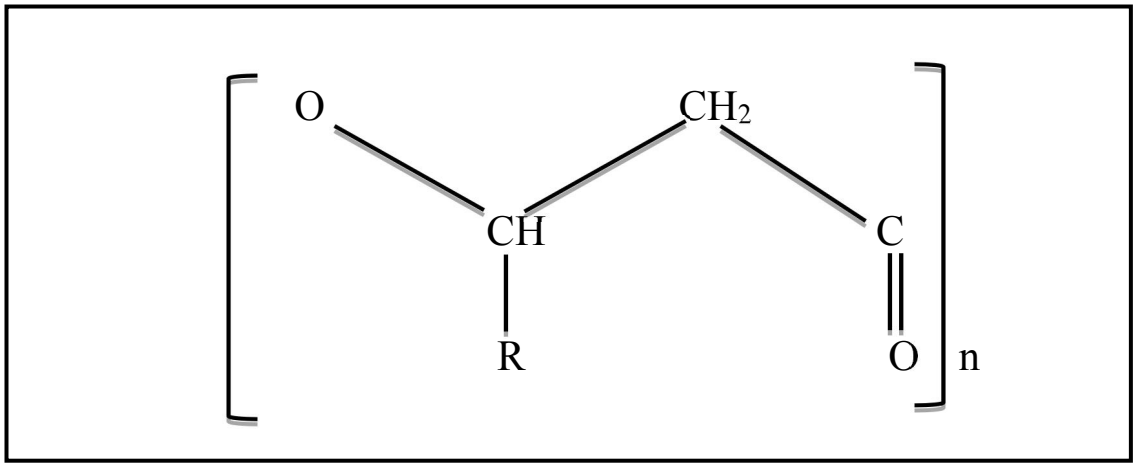

Figure 2.1 General structure of polyhydroxyalkanoate (PHA) (Lee, 1996; Tan *et al.*, 2014)

n varies from 100 to 30000

n = 1	R= hydrogen	Poly (3- hydroxypropionate)
	R= methyl	Poly (3- hydroxybutyrate)
	R= ethyl	Poly (3- hydroxyvalerate)
	R= propyl	Poly (3- hydroxyhexanoate)
	R= pentyl	Poly (3- hydroxyoctanoate)
	R= nonyl	Poly (3- hydroxydodecanoate)
n = 2	R= hydrogen	Poly (4- hydroxybutyrate)
	R= methyl	Poly (3- hydroxyvalerate)
n = 3	R= hydrogen	Poly (5- hydroxyvalerate)
	R= methyl	Poly (5- hydroxyhexanoate)
n = 4	R= hexyl	Poly (6- hydroxydodecanoate)

The monomer units were in (*R*)-configuration that denotes the residual group. The residual group can consist of different carbon chain length. In determining the identity of a monomer unit, the composition of side chain or atom R should be taken into consideration (Kootstra *et al.*, 2017).

2.3 Bacterial PHA Synthesis

There are two groups of bacteria used for the production of PHAs that can be classified based on the culture conditions required for PHA synthesis. The first group belongs to bacteria such as *Cupriavidus necator*, *Protomonas extorquens*, and *Protomonas oleovorans,* that require limitation of an essential nutrient, for instance phosphorous, nitrogen, sulphur or magnesium, to accumulate PHA from excess carbon source but not during the growth phase (Amadu *et al.*, 2021). Next, the second group belongs to some bacteria, such as *Alcaligenes latus*, a mutant strain of *Azotobacter vinelandii*, and recombinant *Escherichia coli,* that accumulate PHA during the growth phase but does not need nutrient limitation for PHA biosynthesis (Muhammadi *et al.*, 2015).

Meanwhile, two ways can be used in obtaining high PHA productivity, which are by fed-batch or continuous fermentation. The first group belonging bacteria mostly utilized the two-step cultivation which are more employed for fed-batch culture. During the first step, the required concentration of biomass is obtained with the absence of nutrient limitation. Afterwards, essential nutrient is kept limited in the second step for efficient PHA synthesis. However, PHA was produced more efficiently when a nutrient is limited but not completely depleted by the bacteria belonging to the first group. A mixture of carbon source and a limited amount of nutrient should be fed at the optimal ratio for bacterial cultivation and high PHA productivity (Khanna & Srivastava, 2005).

The nutrient feeding strategy has been employed for the fed-batch bacterial culture of the second group to obtain high PHA yield. Since the second group of PHA-synthesising bacteria does not depend on nutrient constraint, they can be supplemented with complex nitrogen sources, such as corn steep liquor, yeast extract,

or fish peptone to increase the growth of cell accumulation of polymer. Balanced cell growth and PHA accumulation are needed to avoid incomplete accumulation of PHA or premature termination of fermentation (Khanna & Srivastava, 2005).

2.4 Carbon sources

Over the years, in order to decrease the usage of synthetic plastics, continuing research had been done to increased the PHA production capacity for betterment of the future (Jiang *et al.*, 2016). The PHA production economy depends on various factors, such as cost of substrate and yield of PHA. In this context, the entire expense of PHA production where carbon source plays the main contributor among varying nutrients in the fermentation medium or substrate (Madison & Huisman, 1999). Many research has been focused on producing PHA efficiently using inexpensive carbon source as the production cost is high in comparison to the production of synthetic plastics (Kahar *et al.*, 2004).

Most PHA is produced mainly from food produce, sugar cane and vegetable oils which has raised concerns on food supply production. There are reports of domestic waste utilized as a substitute for cheaper carbon source towards mass production of sustainable biopolymers. The utilisation of agriculture waste, such as whey, olive mills of wastewater, treacle, corn steep liquor, starchy wastewater, and palm oil mill effluent, helps to reduce the PHA production cost by 40 to 50% (Kamilah *et al.*, 2013). Table 2.1 shows the summary on the effect of substrate cost and P(3HB) yield on the production cost of P(3HB).

Table 2.1 Effect of substrate cost and P(3HB) yield on the production cost of P(3HB) (Madison & Huisman, 1999)

Substrate	**Substrate price (US$ kg-1)**	**P(3HB) yield [g P(3HB) (g substrate)$^{-1}$]**	**Substrate cost {US$ [kg P(3HB)]$^{-1}$}**
Glucose	0.493	0.380	1.300
Sucrose	0.290	0.400	0.720
Methanol	0.180	0.430	0.420
Acetic acid	0.595	0.380	1.560
Ethanol	0.502	0.500	1.000
Cane molasses	0.220	0.420	0.520
Cheese whey	0.071	0.330	0.220
Hemicellulose hydrolysate	0.069	0.200	0.340

Besides that, wood hydrolysate is also a potentially inexpensive and renewable carbon source that can be produced through enzymatic or dilute acid hydrolysis of cellulose or hemicellulose to fermentable sugars, such as glucose, galactose, xylose, and mannose (Kumar, 2017). Moreover, the development of non-food based substrates, such as methanol, acetate, and syngas, is quickly becoming one of the most important research areas of industrial biotechnology. Acetate is the second simplest carboxylic acid, and many microorganisms can utilize acetate as alternative carbon source for cell growth. The acetate assimilation species include *Escherichia coli* , *Cryptococcus curvatus*, *Clostridium* sp., *Halomonas boliviensis*, and many more (Chen *et al.*, 2018).

2.5 PHA Synthase

PHA synthases or PhaC are the key enzymes in PHA biosynthesis. The natural biosynthetic pathways involve multiple steps, which begins with the PHA synthase catalysing the polymerisation of coenzyme A (CoA) thioesters of hydroxyalkanoic acids (HA) into PHA polymers (Rehm & Steinbüchel, 2005). Other genes involving in PHA biosynthesis, such as *phaA* (*β*-ketothiolase), *phaB* (acetoacetyl-CoA reductase), *phaG* (3-hydroxyacyl-acyl carrier protein-coenzyme A transferase), *phaJ* (enoyl-CoA hydratase), and other PHA genes were typically pooled with the *phaC* genes. PHA synthases are categorized into four classes which are class I, II, III, and IV (Zou *et al.*, 2017). Class I PHA synthase consists of a single subunit (PhaC) with molecular mass ranging from 60-70 kDa. It consists of one type of PhaC which forms a homodimer. Class I synthase tends to favour scl-HA of C3-C5 carbon chain length with a hydroxyl group at C3 or C4 (Rehm, 2003). The earliest and firstly discovered PHA synthase was the class I PHA synthase $PhaC_{RE}$ from *Ralstonia eutropha* and the PHA synthase is the most widely studied. Nevertheless, class I synthase may as well recognise mcl-HA under specific conditions. For instance, carbon sources from even-chain fatty acids can accumulate P(3HB-co-3HHx); whereas odd-chain fatty acids can accumulate P(3HB-*co*-3HV) when provided as a carbon source.

Meanwhile, class II PHA synthase primarily polymerises CoA thioesters of mcl monomers comprising of 6 to 12 carbon atoms, which are provided via the *β*-oxidation of fatty acids or the *de novo* biosynthesis of fatty acids (Zou *et al.*, 2017). This can be shown from the study of *Pseudomonas mendocina* NK-01 for the PhaC1 and PhaC2, where both subunits had the same substrate specificity. However,

PhaC1 shows a higher catalytic rate and produces PHA of lower molecular weight compared to that of PhaC2 (Guo *et al.*, 2013).

Class III PHA synthase is a heterodimer that needs two subunits for complete activity, where PhaC subunit is one of it while PhaE subunit is the other with a molecular weight of approximately 40 kDa, that is different compared to class I and II. Class III PHA synthases prefer CoA thioesters of (*R*)-3-hydroxy fatty acids which comprised of 3 to 5 carbon atoms.

Class IV PHA synthase resembles class III PHA synthase, whereby it comprises of two different subunits of PhaC and PhaR with a molecular mass of approximately 20 kDa (Zou *et al.*, 2017).In general, class IV PHA synthases is believed to prefer SCL monomers such as 3-hydroxybutyrate (C4) and 3-hydroxyvalerate (C5) for polymerisation, but it is also possible to polymerized another monomer as minor components (Tsuge *et al.*, 2015).

2.6 PHA Applications

Polyhydroxyalkanoates were considered attractive as biomaterials for applications in different fields as they are known to be highly biodegradable and thermoprocessable. PHA can be applied in food industry as packaging films for food packets, baggage, container and coat for paper. It can also be applied for disposals such as shavers blade, utensils, diapers, products containers for female sanitary, beauty containers, shampoo bottles, cups and more. PHA can as well be the starting material for chiral compound, served for the chemical synthesis of optically active compounds. These compounds were applied as biodegradable carriers in the long run controlled dosage of medicinal products, insecticides and herbicides. Besides that, PHA can be implemented in medical applications, such as surgical pins, sutures,

staples, swabs, bandages, bone replacements, plates and blood vessel replacements, stimulation of bone growths by piezoelectric properties, and tissue engineering (Poltronieri & Kumar, 2017).

An interesting yet auspicious application of PHA is its usage in medical and surgical applications due to its biodegradability and biocompatibility. Scaffolding material in tissue engineering is an obvious application in medical. Blends of PHA and hydroxyapatite (HA) were used as scaffolds to treat damage on the bone as reported in earlier studies (Misra *et al.*, 2006). Furthermore, a copolymer of polyglycolic acid (PGA) and PHA were used to produce pulmonary leaflets and pulmonary artery scaffolds in sheep. A continuity of this research was done by constructing a PHA-based heart valve scaffold that was surgically introduced into sheep. It is illustrated that based on these studies, tissue engineering using biopolymer scaffold was possible (Brigham & Sinskey, 2012). Over the years, PHA has successfully been employed as a graft matrix for neuronal generation after spinal cord injury in rats. A factor had showed for the interactions of PHA with cartilage chondrocytes was polymer crystallinity. Scaffolds produced from unblended P(HB-*co*-HHx) were also shown to be effective in cartilage repair. Besides that, Brigham & Sinskey (2012) further explained that matrices fabricated from P(HB-*co*-HV) displayed better healing response than scaffolds fabricated from collagen impregnated with calcium phosphate when embedded into cartilage defects in rabbits. Meanwhile, mild tissue response was shown when scaffolds produced from PHA copolymer were being implanted in rats.

In order for usage on surgical appliances such as sutures, a polymeric material of exceptional tensile strength must be applied to ensure the effectiveness in wound closure. With that, PHB and P(HB-*co*-HV) sutures were shown to be able to promote

recovery of muscle-fascial wounds (Dhingra, 2020). Poly(4-hydroxybutyrate) (P4HB) is a common type of PHA used for fabrication of surgical materials. Furthermore, suture material from P(4HB) fibre (545 MPa) was stronger than polypropylene sutures (410-460 MPa). Also, the Young's modulus of P4HB sutures is substantially lower than other monofilament sutures, produced from other substances, that are on the market. A manufacturer from Cambridge, MA, USA called Tepha Inc. were found to manufactured a few medical devices from PHA. The most widely known product, TephaFLEX® suture fabricated from P4HB was the first to be approved by the US Food and Drug Administration (FDA). Apart from that, they also manufactures surgical meshes and films fabricated from PHA (Brigham & Sinskey, 2012).

2.7 Genus *Cupriavidus*

The species of genus *Cupriavidus*, *Cupriavidus necator*, is first described by Makkar and Casida (1987), as a non-obligate bacterial predator of various soil microorganisms. This species type strain is strain N-1^{T}. It is a Gram-negative, aerobic, mesophilic, short rod that multiplies by binary fission. It is also motile and has peritrichous flagella (Makkar & Casida, 1987). The strain was isolated from soil in the surrounding area of University Park, Pennsylvania, USA (Vandamme & Coenye, 2004). According to Makkar and Casida (1987), *Cupriavidus* strains share several characteristics with groups in the genus *Alcaligenes*, that comprises of multiple species at that time that included *Alcaligenes faecalis*, *Alcaligenes xylosoxidans* and allied species (currently all categorized in the genus *Achromobacter*) and *Alcaligenes eutrophus* (reclassified in the genus *Ralstonia*) and recently classified to the novel genus *Wautersia*. However, Makkar and Casida (1987)

had classified strains into novel genus and species based on few exceptional biochemical characteristics and astonishing predatory activity. After long-term study, the sequences of *Cupriavidus* strains were very similar to *Wautersia eutropha* based on their DNA-DNA hybridization, phenotypic characteristics, DNA base ratios, and 16S rRNA gene sequences. Subsequently, the entire species of genus *Wautersia* were recategorized into *Cupriavidus*, abiding the rules 15, 17, 23a and 37a(1) of the International Code of Nomenclature of Bacteria, the genus name priority goes to *Cupriavidus* over *Wautersia* (Makkar & Casida, 1987).

2.7.1 Taxonomic status

Cupriavidus (L. n. *cuprum,* copper; L. adj. *avidus,* eager for, loving; M. L. neut. n. *Cupriavidus,* favourable of copper). It belongs to the class Betaproteobacteria that includes infamous species of bacteria such as *Burkholderia*, *Bordetella,* and *Cupriavidus* which were established in 1995 where mostly found in water and soil. Most species from genus *Cupriavidus* were isolated from soil. Colonies of bacteria grown on nutrient agar plates appear to be beige but some were yellow-pigmented (Ramachandran *et al.*, 2018). It also appears as glistening, mucoid, smooth, and convex with an entire edge; 2 to 4 mm diameter size on agar plate (Makkar & Casida, 1987).

2.7.2 *Cupriavidus malaysiensis*

The three *Cupriavidus* strains were isolated from natural vicinity in Malaysia. *Cupriavidus malaysiensis* USMAA1020^T was isolated from sludge in Kulim Lake, Kedah, Malaysia after screening of 663 isolates. *Cupravidus malaysiensis* USMAA2-4 was isolated from soil collected in Sungai Pinang, Penang, Malaysia. Meanwhile, *Cupriavidus malaysiensis* USMAHM13 was isolated from sludge of a paddy field in Sungai Manik, Perak, Malaysia (Ramachandran *et al.*, 2018).

2.7.3 Polyhydroxyalkanoate Production

Polyhydroxyalkanoate refers to biopolyesters synthesised by various bacterial species. *Cupriavidus necator*, previously known as *Ralstonia eutropha* or *Alkaligenes eutrophus* is the species that had been widely studied (Vandamme & Coenye, 2004). It had been extensively produced in large scale especially industrially. This particular strain had been used to produce PHBV or poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) copolymer with the market name Biopol by an industrial company called Imperial Chemical Industries and its patents were recently acquired by Metabolix Inc. (USA). In 2007, they produced 100 tonnes per year and has been planning to increased their capacity to 50,000 tonnes per year in 2008 (Verlinden *et al.*, 2007).

Cupriavidus pinatubonensis JMP134 (previously known as *Cupriavidus necator* JMP134) had been reported to contained two putative PHA synthase genes, *phaC1* (Reut_A1347) and *phaC2* (Reut_A2138). This *C. necator* is a versatile aromatic compounds degrader isolated from polluted environments. Both *phaC* genes were categorized into two classes, class I and class II respectively on the basis of their protein sequences but both genes displayed activities towards substrate from SCL

and not MCL *in vitro*. In the study, it is shown that *phaC1* transcriptional level was approximately 70-fold higher (70±9.17) than that of *phaC2* while culturing in 0.5 mM meta-nitrophenol (MNP). Meanwhile, the transcriptional level of *phaC2* was approximately 40-fold higher (244±45.21) than *phaC1* when grown in 0.2% octanoate. This indicated that once induced with dissimilar substrates or metabolites, both genes were differently transcribed (Jiang *et al.*, 2015).

Cupriavidus taiwanensis has also been reported to accumulate PHA. This was expected due to its remarkable genome similarity to *C. necator* H16^T. According to Chien et al. (2010), PHA was synthesized by cloning the PHA synthesis genes (*phaCAB*) of *C. taiwanensis* 184 into *Escherichia coli*. They were able to synthesize approximately 66 to 70% PHA of total cell material from the recombinant strains (Chien *et al.*, 2010).

Last but not least, *Cupriavidus malaysiensis* strains USMAA1020^T, USMAA2-4 and USMAHM13 have been reported to produce PHA when supplemented with the appropriate carbon sources. In strain USMAA1020^T, utilization of γ-butyrolactone as sole carbon source, lead to the a revelation that the concentration of γ-butyrolactone in the culture medium increased from 2.5 g/L to 20.0 g/L, the 4HB fraction in copolyester also increased from 25 to 60 mol% (Amirul *et al.*, 2008). Beside that strain USMAA1020^T was able to accumulate high PHA of 76 wt% and high dry cell weight (9.4 g/L) with the combination of substrate oleic acid and 1-pentanol to synthesize P(3HB-co-3HV) (Huong *et al.*, 2017). Meanwhile in strain USMAA2-4, utilization of oleic acid and 1-pentanol lead to a substantially higher P(3HB-*co*-3HV) concentration of 25.7 g/L and PHA content of 66 wt% (Shantini *et al.*, 2015). Strain USMAA2-4 had also been reported to be able to synthesize terpolymer P(3HB-co-3HV-co-4HB) through the combination of oleic acid with

γ-butyrolactone at different concentrations and 1-pentanol at a fixed concentration with the composition of 9 to 35 mol% 3HV and 4 to 24 mol% 4HB monomers by combining (Aziz *et al.*, 2017). As for USMAHM13, it managed to produce the highest 4HB composition of 43 mol% with cell dry weight and PHA content of 6.0 g L^{-1} and 49 wt% respectively by combining glycerine pitch (5 g L^{-1}) and 1,4-butanediol (5 g L^{-1}) through one-stage cultivation (Ramachandran & Amirul, 2013).

2.7.4 Genome of *Cupriavidus*

Up to date there were eleven reported and published complete genome of *Cupriavidus* strains. Among them are *Cupriavidus pinatubonensis* JMP134 which has two circular chromosomes and two plasmids (CP000090- CP000093) with a total size of 7.23 Mb. It encodes 6,631 predicted protein coding sequences (CDSs), where 4,898 (73.8%) can be assigned a putative function, and 87 RNA genes (Lykidis *et al.*, 2010). *Cupriavidus necator* $H16^T$ which comprised of two chromosomes and one megaplasmid (AM260479, AM260480 and AY305378) with a total size of 7.42 Mb which encodes 6,116 predicted protein coding sequences (Pohlmann *et al.*, 2007).

Meanwhile, *Cupriavidus necator* $N\text{-}1^T$ consists of two chromosomes and two circular plasmids (CP002877-CP002880) with a total size of 8.48 Mb that encodes 7,947 predicted protein coding sequences (Poehlein *et al.*, 2011). *Cupriavidus metallidurans* CH34 isolated from a metal processing factory comprises of two circular chromosomes and two megaplasmids (NC_007971-NC_007974) with a total size of 6.91 Mb which encodes 4,518 predicted protein coding sequences (Janssen *et al.*, 2010). *Cupriavidus taiwanensis* LMG19424 comprises two chromosomes and one megaplasmid (CU633749, CU633750, and CU633751) with a total size of 6.48

Mb that encodes 3,145 predicted protein coding sequences (Amadou *et al.*, 2008). *Cupriavidus basilensis* 4G11 consists of two chromosomes (CP010536-CP010537) with a total size of 8.42 Mb, which encodes 7,661 predicted protein coding sequences (Ray *et al.*, 2015). *Cupriavidus gilardii* CR3 consists of two large chromosomes with a total size of 5.58 Mb, which encodes 4,502 predicted protein coding sequences (Wang *et al.*, 2015).

Other than that *Cupriavidus nantongensis* X1 comprises of two circular chromosomes and one circular plasmid (CP014844-CP014846) with a total size of 7.14 Mb, that encodes 6,524 predicted protein coding sequences (Fang *et al.*, 2016). Additionally, the other genomes were reported as draft genomes.