DIRECTED EVOLUTION OF PHA SYNTHASE (*phaC1*$_{7\text{-}7}$) FROM *Pseudomonas* sp. USM 7-7 ISOLATED FROM ANTARCTICA

EMMANUEL JAIRAJ MOSES

UNIVERSITI SAINS MALAYSIA

2009

DIRECTED EVOLUTION OF PHA SYNTHASE (*phaC1*$_{7\text{-}7}$) FROM *Pseudomonas* sp. USM 7-7 ISOLATED FROM ANTARCTICA

By

EMMANUEL JAIRAJ MOSES

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

JULY 2009

EVOLUSI TERARAH PHA SINTASE (*phaC1*$_{7\text{-}7}$)

DARIPADA *Pseudomonas* sp. USM 7-7 PENCILAN ANTARTIKA

Oleh

EMMANUEL JAIRAJ MOSES

Tesis yang diserahkan untuk memenuhi keperluan bagi Ijazah Sarjana Sains

JULAI 2009

ACKNOWLEDGEMENTS

I would like to take this wonderful opportunity to thank the many people who have helped me in completing this thesis.

Firstly, I would like to thank my supervisor, Prof Madya Dr Mohd Razip Samian for all his guidance, valuable advice, constructive criticisms and scientific discussions. I would also like to thank Prof Mohd Nazalan for introducing me to the concept of "one science" and also for the many intellectual and interesting conversations that we had together.

I also would like to thank the members of LAB 414 for being a second family to me throughout this endeavor. Thank you for all the fun, laughter, tears, sadness, joy, frustrations, hearty meals and wonderful moments that we shared together.

I also like to thank the members of Lab 409, 406,418,318 and 218 for their cooperation in allowing me to use their equipments for the experiments that was carried out.

My family was a great help to me throughout this tough period. Thank you for all the support, love, encouragement and care.

Above all, I would like to thank The Almighty God for His Grace and Strength in enabling me to complete this thesis.

TABLE OF CONTENTS

CHAPTER 4 – RESULTS

LIST OF TABLES

PAGE

LIST OF FIGURES

LIST OF PLATES

PAGE

LIST OF SYMBOLS /ABBREVIATIONS

α	alpha
A	Alanine (Ala)
β	beta
BLAST	Basic Local Alignment Search Tool
C	Cysteine (Cys)
°C	degree celcius
DNA	deoxyribonucleic acid
DMSO	dimethyl sulfoxide
D	Aspartate (Asp)
E	Glutamate (Glu)
F	Phenylalanine (Phe)
G	Glycine (Gly)
g	gram
H	Histidine (His)
I	Isoleucine (Ile)
K	Lysine
kPa	kilo pascal
kb	kilo base
L	Leucine (Leu)
L	litre
LA	luria agar
M	Methionine (Met)
$MgSO_4$	magnesium sulphate
mg	miligram
ml	mililitre
N	Asparagine (Asp)
N_2	Nitrogen

nm	nanometer
P	Proline (Pro)
PEG	polyethelene glycol
psi	pound per square inch
Q	Glutamine (Glu)
R	Arginine
rpm	revolution per minute
S	Serine
s	seconds
T	Threonine (Thr)
TAE	Tris-acetate EDTA
TSS	Transformation and Storage Buffer
U	units
µg	microgram
µl	microlitre
V	Valine (Val)
W	Tryptophan
wt%	percentage of dried cell weight
w/v	weight/volume
Y	Tyrosine (Tyr)

EVOLUSI TERARAH PHA SINTASE (*phaC1*$_{7\text{-}7}$) DARIPADA *Pseudomonas* sp. USM 7-7 PENCILAN ANTARTIKA

ABSTRAK

Polihidroksialkanoat (PHA) sintase ialah enzim utama dalam biosintesis polihidroksialkanoat (PHA), sekumpulan biopoliester linear yang berfungsi sebagai bahan simpanan karbon dan tenaga di dalam sel prokariot. Biopoliester ini banyak diberi tumpuan sebagai pengganti kepada plastik yang diperbuat daripada petroleum kerana ia memiliki ciri yang agak sama dengan polipropelin, boleh dibiodegradasi, mempunyai ciri kebioserasian dan diperoleh daripada sumber yang boleh diperbaharui. Walaubagaimanapun, kos penghasilan yang tinggi menghindarkan penggunaan polimer ini berleluasa. Salah satu cara untuk menghasilkan PHA sintase berkualiti tinggi ialah melalui evolusi terarah yang sebenarnya adalah proses evolusi semulajadi yang dilakukan dalam makmal. Satu kitaran evolusi tearah melalui 'Staggered Extension Process' (StEP) telah dilakukan ke atas PHA sintase (*phaC1*$_{7\text{-}7}$) dan proses pemilihan telah menunjukkan bahawa 400 (80%) daripada 500 klon dalam perpustakaan mutan berfungsi. Dua puluh klon kemudiannya dipilih untuk analisis kromatografi gas (GC) dan penjujukkan DNA. Analisa GC menujukkan bahawa *phaC1*$_{7\text{-}7}$ jenis liar mengumpul sebanyak 0.4 peratus daripada berat sel kering (wt%) poli-3-hidroksibutirat P(3HB) dan empat mutan (MC1 1-3, MC1 4-43, MC15-13 dan MC15-16) mengumpul lebih banyak P(3HB). MC1 1-3 dan MC1 5-16 mengumpul 25% lebih banyak P(3HB) manakala MC1 4-43 dan MC1 5-13 mengumpul 50% lebih banyak P(3HB). Beberapa perubahan asid amino PhaC1 yang mungkin meningkatkan kebolehan mutan-mutan tersebut untuk

mengumpul P(3HB) telah dikenalpasti apabila mutan-mutan ini dianalisa dengan lebih mendalam. Perubahan asid amino yang mungkin terlibat dalam peningkatan pengumpulan P(3HB) adalah R101P, F148L, D157N, P223L, D328V, H350Y dan Y444H.

DIRECTED EVOLUTION OF PHA SYNTHASE (*phaC1*$_{7\text{-}7}$) FROM *Pseudomonas* sp. USM 7-7 ISOLATED FROM ANTARCTICA

ABSTRACT

Polyhydroxyalkanoate (PHA) synthase is the key enzyme in the biosynthesis of polyhydroxyalkanoates (PHAs), which is a family of linear biopolyesters that function as carbon and energy reserves in prokaryotic cells. These biopolyesters have been drawing much attention as alternatives for petrochemically based plastics since they have physical properties similar to that of polypropylene, are biodegradable, biocompatible and can be obtained from renewable resources. However the main drawback of the utilization of theses polymers on a large scale is the high cost of production. Improvement of PHA synthase properties seems to be an attractive solution to alleviate this problem. One of the ways to enhance the properties of this unique enzyme is by subjecting it to directed evolution which is actually the natural evolutionary process mimicked in laboratory conditions. The PHA synthase (*phaC1*$_{7\text{-}7}$) in this study was subjected to one cycle of directed evolution via the Staggered Extension Process (StEP) and the selection process revealed that 400 (80%) out of the 500 clones in the mutant library was functional. Twenty clones were then selected and subjected to DNA sequence and gas chromatography (GC) analyses. Analysis showed that wild type *phaC1*$_{7\text{-}7}$ accumulated 0.4 percent of dried cell weight (wt%) P(3HB) and four mutants (MC1 1-3, MC1 4-43, MC1 5-13 and MC1 5-16) accumulated higher amounts of P(3HB) as compared to the wild type. MC1 1-3 and MC1 5-16 accumulated

25% higher amounts of P(3HB) while MC1 4-43 and MC1 5-13 accumulated 50% higher amounts of P(3HB). Several beneficial amino acid substitutions that likely increased the P(3HB) accumulation ability of *phaC1*$_{7\text{-}7}$ were identified when the mutants were analyzed further. The possible beneficial amino acid substitutions were R101P, F148L, D157N, P223L, D328V, H350Y and Y444H.

CHAPTER 1: INTRODUCTION

Synthetic polymers otherwise known as plastics have been widely used as an alternative to glass, wood and other construction materials in various industrial, domestic and environmental applications for the past 50 years (Lee *et al*., 1991; Poirier *et al*., 1995; Lee, 1996). The widespread use of plastic is due to the fact that plastics are durable and stable. Furthermore, it possesses several mechanical and thermal properties that make it suitable for various industrial purposes (Rivard *et al*., 1995)

However, disposal of plastics poses major environmental hazards. This is the due to its recalcitrant nature which makes it persistent in the environment. In addition, incineration of plastics releases dangerous fumes such as dioxin (Ojumu *et al*., 2004). The solution to this grave situation seems to lie on the utilization of a group of unique and interesting biopolymers known as polyhydroxyalkanoates (PHAs).

Polyhydroxyalkanoates (PHAs) are biopolyesters composed of 3-hydroxy fatty acid monomers which function as carbon and energy reserves in prokaryotic cells (Anderson and Dawes, 1990). Polyhydroxyalkanoic acids (PHAs) are biotechnologically relevant products currently being investigated in many laboratories in academia and industry because these biological polyesters are biodegradable thermoplastics and elastomers that exhibit interesting material properties (Hocking and Marchessault, 1996; Jendrossek *et al*., 1996; Steinbuchel, 1996). Furthermore, its production is based on renewable sources and the polymers are biocompatible (Sudesh *et al.,* 2000; Volova *et al.,* 2003). Therefore, they are considered for several applications in the packaging, pharmaceutical, agriculture and food industry (Anderson and Dawes, 1990; Walle *et al*., 2001), or even as

bioimplant materials for tissue engineering (Deng *et al*., 2002). Polyhydroxyalkanoates (PHAs) seem to be suitable candidates to replace conventional plastics.

The main deterrent of utilizing polyhydroxyalkanotes on a larger scale to fully replace petroleum based plastic is the high cost of production (Lenz and Marchessault, 2005; Verlinden *et al.,* 2007). One of the initiatives currently undertaken by many researchers both in industries and academics to alleviate this problem is to improve the properties of PHA synthase which is a key enzyme in PHA production via directed evolution (Rehm, 2003; Rehm, 2007; Nomura and Taguchi, 2007). The current study that was carried out revolved around this theme.

An interesting bacterium known as *Pseudomonas* sp USM 7-7 was previously isolated from Antarctic sea ice sample. This bacterium can produce up to 9 wt% of PHA at 30°C when supplied with glucose as a carbon source. Two isogenic genes (*phaC1*$_{7\text{-}7}$ and *phaC2*$_{7\text{-}7}$) that encoded for two different PHA synthases (PhaC1and PhaC2) were identified, cloned and characterized from this microorganism. Further analysis showed these PHA synthases produced only medium chain length (mcl) PHA and was unable to produce poly-3-hydroxybutyrate P(3HB) homopolyesters naturally (Yit, 2006).

A closely related PHA synthase (PhaC1), was isolated from *Pseudomonas* sp 61-3, a bacterial strain of mesophilic origin (Matsusaki, 1998) and reported to accumulate poly-3-hydroxybutyrate P(3HB), a type of short chain length (scl) PHA albeit at extremely low levels (Matsumoto *et al*., 2002). Recent attempts were made to improve the P(3HB) accumulation ability of this PHA synthase by directed evolution (Takase *et al*., 2003).

Blast 2 Seq. alignments (Tatiana et al., 1999) provided by NCBI showed that this PHA synthase shared 84% similarity with the PHA synthase C1 (PhaC1) isolated from *Pseudomonas* sp. USM 7-7. Therefore, this study was conducted to investigate the P(3HB) accumulation ability of PhaC1 from *Pseudomonas* sp. USM 7-7 and further improve it via directed evolution.

This study was divided into two major parts. The first part involved the construction of a *phaC1*$_{7\text{-}7}$ directed evolution mutant library using random gene recombination techniques which included DNA shuffling and Staggered Extension Process (StEP). The second part of this study involved the selection and characterization of *phaC1*$_{7\text{-}7}$ mutants with the desired properties.

The main objectives of this study were;

1. To construct a *phaC1*$_{7\text{-}7}$ directed evolution mutant library via gene recombination techniques (Staggered Extension Process and DNA Shuffling)
2. To determine poly-3-hydroxybutyrate P(3HB) accumulation levels in *E. coli* cells habouring *phaC1*$_{7\text{-}7}$ mutant genes
3. To identify possible amino acid substitutions that likely contributed to the increase of poly-3-hydroxybutyrate P(3HB) production

2.0 LITERATURE REVIEW

2.1 Engineering Industrial Enzymes: An Overview

Enzymes initiate and govern a myriad of biochemical reactions in living organisms. Virtually all known enzymes are proteins and have the capability to sustain life itself (Bott, 2004). This is because enzymes are able to work under ambient temperatures and pH. Furthermore, these natural catalysts are highly specific (Cherry and Fidantsef, 2003).

The unique properties mentioned make enzymes an attractive alternative to chemical catalysts for industrial applications. Enzymes are considered superior to their counterparts; the chemical catalysts, because they can accelerate reaction rates by enormous factors (Benkovic and Hammes-Schiffer, 2003), are chemospecific, regiospecific as well as stereospecific (Marrs *et al.,* 1999). All these properties are exhibited at relatively benign conditions as compared to chemical catalysts (Cippola, 2004). Further advantages of using enzymes as industrial catalysts include reduced energy consumption, lower material consumption and simpler process design (Zagrebelny, 2005). Enzymes used in industries are generally known as industrial enzymes (Marrs *et al*., 1999).

Several industries have replaced chemical catalysts with enzymes. This can clearly be seen in laundry detergent industries where phosphates have been replaced by proteases and cellulases. Similarly, textile industries extensively utilize amylases and pectinases as catalyst instead of sodium hydroxide. The bread making industry on the other hand has substituted chemical emulsifiers with lipases. The number of

enzymes used in industries has increased exponentially over the years (Schmidt-Dannert, 2001).

Enzymes have been used as catalysts for over 500 products in 50 different industrial applications (Cherry and Fidantsef, 2003). Based on the statistics above, it can be said that the number of enzymes used in industries will definitely continue to rise. However, using enzymes for industrial applications poses a major obstacle to enzyme technologists because these exquisite catalysts of nature are not created to withstand harsh industrial conditions such as high temperature, organic solvents, pH, detergents and oxidants (Schmidt-Dannert and Arnold, 1999). Therefore, harnessing enzymes for industrial purposes usually requires a certain degree of engineering to improve their activity or stability (Bloom *et al.,* 2005). Besides improvement of enzyme properties, enzyme engineering can also force enzymatic reactions to proceed in the desired direction (Cipolla, 2004).

Many methods of enzyme engineering such as chemical modification, phage display, rational design and directed evolution have been employed in attempts to produce enzymes with improved functions and properties (Marrs *et al*., 1999). Of all the methods mentioned above, rational design and directed evolution have been the two principle strategies used in enzyme engineering over this past few years (Cipolla, 2004).

Rational design uses advance technologies such as molecular modeling, computational chemistry and X-ray crystallography (Schmidt-Dannert and Arnold, 1999) to determine the structure of enzymes and elucidate the relationships between sequence, structure, function and mechanism in order to produce enzymes with enhanced properties and functions. This approach requires a great deal of knowledge regarding the enzyme (Bornscheuer and Pohl, 2001). Enzymes that have not been

crystallized cannot be engineered using this method. Interactions between enzymes and the surrounding environment cannot be determined as well. Furthermore, the effect of mutations beyond enzyme structure cannot be predicted using this approach (Powell *et al*., 2001).

In contrast, directed evolution techniques which mimics natural evolution, enable the engineering of enzymes without a deep understanding in relation to the structure or function of the enzyme (Dalby, 2003). The latter will be the focus of this review.

2.2 Directed Evolution: A Brief Introduction

Organisms have adapted to change over millions of years. This is clearly evident as life can be found in various environments. Some microbes and other life forms have adapted their physiology by changing enzymes, proteins and other characteristics to enable them to thrive in extremely harsh conditions such as the depth of oceans, volcanic areas, polar region and highly saline lakes. This astounding feat is accomplished by a process known as natural evolution. Natural evolution selects for specific traits by applying environmental pressure (Eijsink *et al.*, 2005). Creation of functionalities or new traits is primarily achieved through mutagenesis, recombination and survival of the fittest (Otten and Quax, 2005).

Biologists have developed a strategy to mimic the natural evolutionary process mentioned above in laboratory conditions. This process is known as directed evolution and is also referred to as molecular breeding occasionally (Powell *et al*., 2001). It is based on Darwinian principles of mutation and selection (Chatterjee and Yuan, 2006). However, it differs from natural evolution in two major aspects. Firstly, natural evolution occurs under multiple and variable selection pressure

whereas directed evolution is accomplished under controlled selection pressure for predetermined functions. The other key aspect that makes directed evolution different is that functions of practical use can be obtained through the design of appropriate selection schemes while natural evolution is biased towards functions that are crucial to the survival of the organism (Schmidt-Dannert, 2001, Williams, 2004).

The first directed evolution experiment was carried out by Mills and co-workers about four decades ago in 1967. Ironically, the study was not conducted on proteins or enzymes but DNA (Chatterjee and Yuan, 2006). The first directed evolution experiment involving enzymes was carried out by Chen and Arnold only in 1993. The enzyme used in this study was subtilisin E (Chen and Arnold, 1993). In 1994, Pim Stemmer introduced the concept of DNA shuffling which was based on the natural homologous DNA recombination process (Otten and Quax, 2005). These two works marked a new advent in the development of directed enzyme evolution. This field has grown in leaps and bounds ever since.

All directed evolution experiments consist of two very important steps. The first step involves the generation of a directed evolution mutant library. The second step involves the screening and selection of these mutants for improved properties or function (Cipolla, 2004; Williams *et al*., 2004, Wang *et al*., 2006 and Dalby, 2007).

2.3 Construction of Directed Evolution Mutant Libraries

All directed evolution experiments are initiated by the construction of a mutant library. The main criterion to be considered in directed evolution mutant library construction is the creation of molecular diversity. This has been achieved in

laboratory conditions through the mimicry of two essential evolutionary processes namely random mutagenesis and gene recombination (Zhao and Zha, 2004). The principles revolving around these two key evolutionary have been applied in mutant library construction.

2.3.1 Mutant Library Construction via Random Mutagenesis.

Random mutagenesis occurs when there is error during DNA replication. This includes nucleotide substitutions, insertions, deletions and inversions. Random mutagenesis techniques utilized in directed evolution mutant library construction are based on the first three phenomenon (Zhao and Zha, 2004).

The most simple and common random mutagenesis technique is the introduction of point mutations over the entire length of the target gene. This can be achieved by using chemical mutagens (Myers *et al.,* 1985) and ultraviolet (uv) radiation (Botstein and Shortle, 1985). Mutator strains can be used to generate point mutations as well (Botstein and Shortle, 1985). An example of a commonly used mutator strain is *E. coli* XL-1 Red which is commercially available from Stratagene (Bornscheuer *et al.,* 1999; Alexeeva *et al.,* 2002).This specially engineered strain is deficient in three primary DNA repair pathways and has a mutation rate that is 5000 times higher than that of wild type *E. coli*. However, the use of mutator strains is quite limited because the genome is not stable and the doubling time is slower than wild type *E. coli* (Wang *et al.,* 2006). Error prone polymerase chain reaction (PCR) is another approach that is routinely used to introduce point mutations. In error prone PCR, magnesium (Mg^{2+}) ion is usually substituted with manganese (Mn^{2+}) as the cofactor. Other parameters that are manipulated in the experiment include

dinucleoside triphosphate (dNTP) concentrations and the number of amplification cycles (Leung *et al.,* 1989; Cadwell and Joyce 1992; Cirino *et al.,* 2003).

The second technique involves saturation mutagenesis. Saturation mutagenesis refers to the creation of all possible amino acids at a particular residue or region of a protein (Zhao and Zha, 2004). The target residues or regions are usually predicted through structure-function relationship knowledge (Olson and Sauer, 1988) or point mutation experiments (Miyazaki and Arnold, 1999). Combinatorial cassette mutagenesis (Wells *et al.,* 1988; Olson and Sauer, 1988), recursive ensemble mutagenesis (Delagrave *et al.,* 1993), scanning saturation mutagenesis (Chen *et al.,* 1999) and codon cassette (Kegler-Ebo *et al.,* 1994) are some of the common saturation mutagenesis technique that are usually employed.

Another approach commonly used is random mutagenesis through insertions and deletions. A well published method using this approach is termed RID (Random Insertion/Deletion) mutagenesis. In this method, up to 16 bases can be inserted or deleted. A major advantage of this method is that insertions and deletions can be performed concurrently (Murakami *et al.,* 2002). However, this approach is less popular because it is laborious, time consuming and requires large amounts of DNA templates (Zhao and Zha, 2004; Neylon, 2004).

All the random mutagenesis techniques mentioned thus far have their own pros and cons. However, error prone PCR remains the most popular and is an almost universal approach to create directed evolution mutant libraries through random mutagenesis. This is due to its robustness, efficiency and simplicity (Neylon, 2004).

2.3.2 Recombination Techniques in Mutant Library Construction

Almost all random mutagenesis techniques mentioned earlier suffer some drawbacks. The major drawbacks include bias in the type of nucleotide mutations (error bias) and bias in the types of amino acid substitutions (codon bias). This problem could be resolved by employing gene recombination techniques to construct mutant libraries (Neylon, 2004; Zhao and Zha, 2004).

Gene recombination plays a crucial role in evolution as it can repair damaged genes and combine different variants to increase the diversity of a population. The various gene recombination techniques are modeled based on this events and include homologous and non homologous recombination. Recombination techniques offer a major advantage over random mutagenesis as it can accumulate beneficial mutations and remove deleterious ones (Zhao and Zha, 2004). Recombination techniques in mutant library construction can be divided into six broad categories namely shuffling, full length parent shuffling, single crossover, domain swapping, *in vivo* recombination and synthetic shuffling (Otten and Quax, 2005).

DNA shuffling is by far the most common recombination technique that is used in the generation of mutant libraries. A general workflow of DNA shuffling involves the digestion of the source DNA with DNAse. The fragments are then purified, mixed together and subjected to repeated cycles of melting, annealing and extension. The assembled fragments are then produced in substantial amounts using a final PCR amplification step (Stemmer, 1994; Zhao, 1997). Sometimes, restriction enzymes (Kikuchi *et al.,* 1999) or endonuclease V (Miyazaki, 2002) is used to fragment the genes instead of DNAse.

The Staggered Extension Process (StEP) is another similar *in vitro* homologous gene recombination technique that can be used. In this method, full length genes are used as templates to synthesize chimeric gene products via multiple cycles of denaturation and extremely short annealing/extension periods as opposed to DNA shuffling (Zhao *et al.,* 1998).

Random Chimeragenesis on Transient Template (RACHITT) is another *in vitro* homologous gene recombination technique that is quite popular. This approach is different from the first two that were mentioned in that no thermocycling, overlap extension or staggered extension is involved. The method involves ordering, trimming and assembly of randomly cleaved single stranded parental gene fragments annealed onto a transient single stranded template which is prepared from one of the parent genes and contains uracil. A major advantage of this technique is that large numbers of crossovers can be achieved (up to 14). However, this method is technically difficult because it involves additional steps in generating the single stranded DNA (Coco *et al.,* 2001, Neylon, 2004, Zhao and Zha 2004).

All the methods mentioned above have one common weakness that is a high dependence on the homology of the DNA sequences that need to be recombined. Therefore, these methods are not applicable for DNA sequences with little or no homology. A number of alternatives have been developed to address this issue. These methods are collectively known as non homologous gene recombination methods (Zhao and Zha, 2004).

Incremental truncation for creation of hybrid enzymes (ITCHY) is one example of a non homologous recombination technique that can be used. This method entails the use of exonuclease digestion to incrementally truncate the parental genes. The truncated genes are then ligated using blunt end ligation to create functional hybrid

enzymes (Ostermeier *et al*., 1999). The major disadvantage of this method is that the exonuclease digestion is very difficult to control and optimize (Neylon, 2004). Another example of a non homologous gene recombination method is the Sequence Homology Independent Protein Recombination (SHIPREC). This method involves truncation of parental genes with DNAse I, fragment selection and blunt end ligation (Sieber *et al.,* 2001). Other examples of non homologous recombination methods include degenerate oligonucleotide gene shuffling (DOGS), *in vitro* exon shuffling and random multirecombinant PCR (Gibbs *et al.,* 2001; Kolkman and Stemmer, 2001; Tsuji, 2001).

Besides the *in vitro* recombination methods described above, there have been instances in which *in vivo* recombination techniques have been used (Zhao and Zha, 2004). A famous example would be the combinatorial libraries enhanced by recombination in yeast (CLERY) method. This approach combines both *in vitro* DNA shuffling and *in vivo* homologous recombination in yeast (Cherry *et al.,* 1999).

Another *in vivo* recombination technique that has been reported in literature is the Random Chimeragenesis by Heteroduplex Recombination method which relies on the DNA repair system to rectify regions of non identity in the heteroduplex formed among different parental genes (Volkov *et al.,* 1999).

2.4 Screening and Selection

The second step in directed evolution experiments involves the screening and selection of mutants or variants with improved functions or properties. Various screening and selection strategies have been developed for this purpose. Linkages between the gene, the enzyme it encodes and the enzyme product form the basis of these strategies (Boersma *et al.,* 2007).

All screening and selection assay systems need to meet some requirements. Firstly, the substrate used should be identical or at least similar to the substrate of the target enzyme. Secondly, the assay should be applicable over the desired dynamic range. Thirdly, the assay should be fast and available in a high throughput format (Bornscheuer, 2004; Boersma *et al.,* 2007)

Screening is usually done using fluorogenic or chromogenic substrates which are readily converted into spectroscopically different products (Wahler and Raymond, 2001; Aharoni *et al.,* 2005; Otten and Quax, 2005). A major advantage of screening is that the difference between the substrate and the product can be readily determined indiscriminately in most cases. A major disadvantage of using this approach is that each mutant in the library must be tested for activity or functionality (Boersma *et al.,* 2007).

Selection assays are usually performed on the entire mutant library simultaneously (Aharoni *et al.,* 2005; Leemhuis *et al.,* 2005). The selection assays are usually done either *in vitro* or *in vivo*. *In vitro* selection does not require viable cells. Besides that, no membrane barrier exists between the substrate and products. On the other hand, in vivo selection usually links cell survival or function to enzyme activity and involves growth selection (Boersma *et al.,* 2007). In general, selection methods are preferred over screening as it is less tedious and allows analysis of more variants simultaneously (Otten and Quax, 2005; Boesrma *et al.,* 2007).

A summary of some screening and selection techniques as well as their advantages and disadvantages is presented in Table 2.1

Table 2.1 Selection and screening methods.

Method	Requirement	Strength	Weakness	Example	Reference
Selection					
Growth on appropriate antibiotic	Mutant cells release enzymes that will destroy the antibiotic	Highly sensitive. High throughput screening possible.	Not generally feasible. Can generate highly resistant strains.	Increased moxalactame resistance using Cephalosporinase genes	Crameri *et al*. (1998)
Complementation	Enzymatic product occurs in metabolism	Very specific. High throughput screening possible	Restricted to product of metabolism	Identification of tryptophan producing Mutants	Juergens *et al.*(2000)
Display methods Coupled with detection e.g fluorescence-activated cell sorting (FACS), bio-panning or suicide substrate	Proteins must be displayed	High throughput screening. Ultra sensitive	Difficult to detect improved variants for existing activity	Identification of proteases. Epitope mapping	Jose *et al*. (2001) Christmann *et al.* (2001)
Screening					
Microtiter plate (using chromogenic/ fluorogenic substrates)	Synthesis/design of assay substrates Photometer/fluorimeter	Very sensitive Low background signal. High throughput screening applicable. Detection of improved properties	Synthesis required Surrogate substrates used	Identification of more stereoselective lipase/esterase variants Fingerprinting of various enzymatic activities	Reetz *et al*.(1997). Badalassi *et al*. (2000). Reymond and Wahler (2002).
Microtiter plate (using true substrates)	Photometer/fluorimeter Electrospray ionization mass spectroscopy (ESI-MS) Nuclear Magnetic Resonance (NMR)	Direct detection of true activity	Less sensitive. Strong background	Determination of lipase/esterase activity	Reetz *et al*.(1999) Baumann et al. (2001).

(Adapted from Bornscheuer, 2004)

2.5 Directed Evolution of Various Industrial Enzymes

Developing enzymes for industrial purposes remain a great challenge to many scientists by and large. However, in recent years, this endeavor has been greatly eased by the utilization of directed evolution. Various enzyme properties such as specificity, activity, stability and solubility have been improved through directed evolution (Johannes and Zhao, 2006). There have been numerous publications regarding this subject and some of these examples will be described shortly.

2.5.1 Improvement of Enzyme Specificity via Directed Evolution

Specificities of enzymes have been improved through directed evolution. This includes substrate specificity and enantioselectivity.

The specificity of the enzyme organophosphorous hydrolase towards Chlorpyrifos (a commercially available pesticide) was increased after two rounds of DNA shuffling (Chen *et al.,* 2004). The substrate specificity of cytochrome c peroxide from *Saccharomyces cerevisiae* was altered from the protein cytochrome c to small molecules such as guiacol using DNA shuffling and saturation mutagenesis (Iffland *et al.,* 2000; Iffland *et al.,* 2001). Similarly, the substrate specificity of Cytochrome P450 oxygenases has been broadened to accept short chain alkanes as substrates using directed evolution techniques. This enzyme is particularly important in pharmaceutical industries because it is the key enzyme in hepatic drug metabolism (Julsing *et al.,* 2008)

The enantioselectivity of *Pseudomonas aeruginosa* lipase towards 2-methyldecanoate was considerably increased using error prone polymerase chain

reaction (ePCR) and site saturation mutagenesis (Reetz *et al*., 2000, Reetz *et al.,* 2001). A similar approach was employed to invert the enantioselectivity of hydantoinase from D- selectivity to L-selectivity (Arnold *et al.,* 2000). Another example is the improvement of *Agrobacterium radiobacter* epoxide hydrolase enantioselectivity by 13 fold using error prone PCR and DNA shuffling (Spelberg *et al.,* 2004).

2.5.2 Improvement of Enzyme Stability via Directed Evolution

The stability of an enzyme is affected by many factors such as temperature, pH, oxidative stress, the solvent used, binding of metal ions or cofactors and the presence of surfactants (Eijinsk *et al.,* 2005).

The stability of subtilisins was improved using a family shuffling method that involved 26 closely related subtilisins. The improved variants exhibited greater thermostability, solvent stability and greater tolerance towards higher pH (Ness *et al.,* 1999).

The pH tolerance of a thermophilic xylanase isolated from *Thermobifida fusca* was recently broadened to withstand alkaline conditions using two rounds of DNA shuffling. This property is of extreme importance as xylanase catalyzes the hydrolysis of xylan which is a major component in hemicelluloses and is commonly used in paper and pulp industries which require high temperatures and alkaline conditions (Wang and Xia, 2008).

The stability of phosholipase A_1 in organic solvents was increased after being subjected to random mutagenesis and recombination (Song and Rhee, 2001). Variants of horse radish peroxidase that exhibited greater stability in the presence of

hydrogen peroxide (H_2O_2), sodium dodecyl sulfate (SDS) and salts were also obtained using the approaches mentioned above (Arnold *et al.,* 2001).

2.5.3 Improvement of Enzyme Catalytic Activity via Directed Evolution

The catalytic activities of natural occurring enzymes are usually too low to be of industrial importance. Therefore, the catalytic activities of these enzymes need to be enhanced. This has been routinely done through directed evolution lately (Johannes and Zhao, 2006).

The catalytic activity of glyphosphate-N-actetyltransferase (GAT) was improved a thousand fold after eleven rounds of DNA shuffling. This enzyme is important as it confers resistance to glyphosphate (Castle *et al.,* 2004; Johannes and Zhao*,* 2006). Error prone PCR and DNA shuffling was also used to increase the total activity of barley α amylase to about 1000 times (Wong *et al.,* 2004). This enzyme is important in the production of corn sweeteners and bioethanol (Textor *et al.,* 1998).

Recently, the catalytic activity of a thermostable xylanase was improved 12 fold using directed evolution (Wong and Xia, 2008).

2.5.4 Improvement of Enzyme Solubility via Directed Evolution

Enzyme solubility is usually desired for a high level of expression in heterologous hosts. This feat has been routinely achieved using directed evolution approaches (Johannes and Zhao, 2006).

The solubility of phosphotriesterase from *Pseudomonas diminuta* was increased twenty fold using error prone PCR and DNA shuffling (Roodvelt and Tawfik, 2005). Phosphotriesterases are important because it catalyzes the detoxification of organophosphate compounds (Johannes and Zhao, 2006). The tobacco etch virus protease was also expressed in a more soluble form by using directed evolution techniques mentioned above (Berglund *et al.,* 2005).

2.6 Directed Evolution of Psychrophilic Enzymes

The examples described thus far focus on enzymes isolated from mesophilic or thermophilic organisms. However, in recent years, enzymes isolated from psychrophilic organisms have become more attractive and important for industrial applications. This group of enzymes is collectively known as psychrophilic or cold-adapted enzymes. These enzymes have higher specific activity at lower temperature, better flexibility and enhanced thermolability as compared to their mesophilic and thermophilic counterparts (van den Burg, 2003; Watanabe, 2004). Psychrophilic enzymes are particularly important in laundry, paper and pulp as well as food processing industries which favour enzymes that can function at relatively moderate temperatures. This is to decrease energy consumption (van den Burg, 2003).

Psychrophilic enzymes have also been subjected to directed evolution to make it feasible for industrial purposes. For example, the enzyme activity of β-galactosidase (Family 42) was increased 1.5 fold at 18°C using directed evolution (Panasik, 2002). Another β-galactosidase (Family 2) produced variants that exhibited 2.5 fold increase in activity at 15°C after being subjected to saturation mutagenesis (Coker and Brenchely, 2006).

A psychrophilic subtilisin was subjected to random mutagenesis, saturation mutagenesis and shuffling and produced variants that showed improvement in catalytic activity by three fold at 10°C and the stability increased 500 fold at 60°C (Miyazaki *et al*., 2000).

2.7 Polyhydroxyalkanoate (PHA) Synthases are Important Industrial Enzymes

The preceding portion of this review has made it clearly evident that directed evolution is the epitome of enzyme engineering in this decade. A group of important industrial enzymes known as polyhyroxyalkanoate (PHA) synthases will be used as a paradigm to further reinstate this fact in the subsequent portion of this review.

Polyhydroxyalkanaoate (PHA) synthases are important industrial enzymes because it catalyzes the polymerization of R-3-hydroxyacyl-coenzyme-A (3HA-CoA) to produce a group of important biopolyesters known as polyhydroxyalkanoates (PHAs). These biodegradable polymers have attracted much attention in both industry and academia due to its similar properties to polypropylene and other thermoplastics thus being considered an attractive alternative to conventional petrochemical-based plastics (Jendrossek *et al*., 1996; Steinbuchel, 1996; Kichise *et al*., 2002).

The properties of this interesting biopolymer and the intriguing enzyme that catalyzes its polymerization will be elaborated shortly. Various attempts to improve the properties of the polymerase via directed evolution will be described as well.

2.8 Polyhydroxyalkanoates (PHA): An Introduction

Polyhydroxyalkanoates (PHAs) are a group of biological polyesters synthesized by a wide consortium of prokaryotic microorganisms from kingdoms eubacteria and archaea. These biopolymers are accumulated as water insoluble inclusions when there is an excess of carbon source and other nutrients such as nitrogen, phosphorus or oxygen is growth limiting. The microorganisms will then utilize these storage polymers as carbon and energy source during conditions of starvation (Anderson and Dawes, 1990; Steinbuchel, 1991; Rehm, 2003; Shang *et al.,* 2003; Rehm, 2007; Verlinden *et al.,* 2007). In addition to being carbon and energy source, these biopolymers can also function as electron sinks and have an impact on bio-film formation (Pham *et al.,* 2004; Rehm, 2006).

Most PHAs are aliphatic polyesters composed of carbon, oxygen and hydrogen with polyhydroxybutyrate (PHB) being the most abundant and widely studied (Braunegg *et al.,* 1998; Khanna and Srivastava, 2005). Different PHAs vary at the structures of the pendant groups in C-3 or β-position (Lu *et al*., 2005). The general chemical structure of PHA is depicted in Figure 2.1

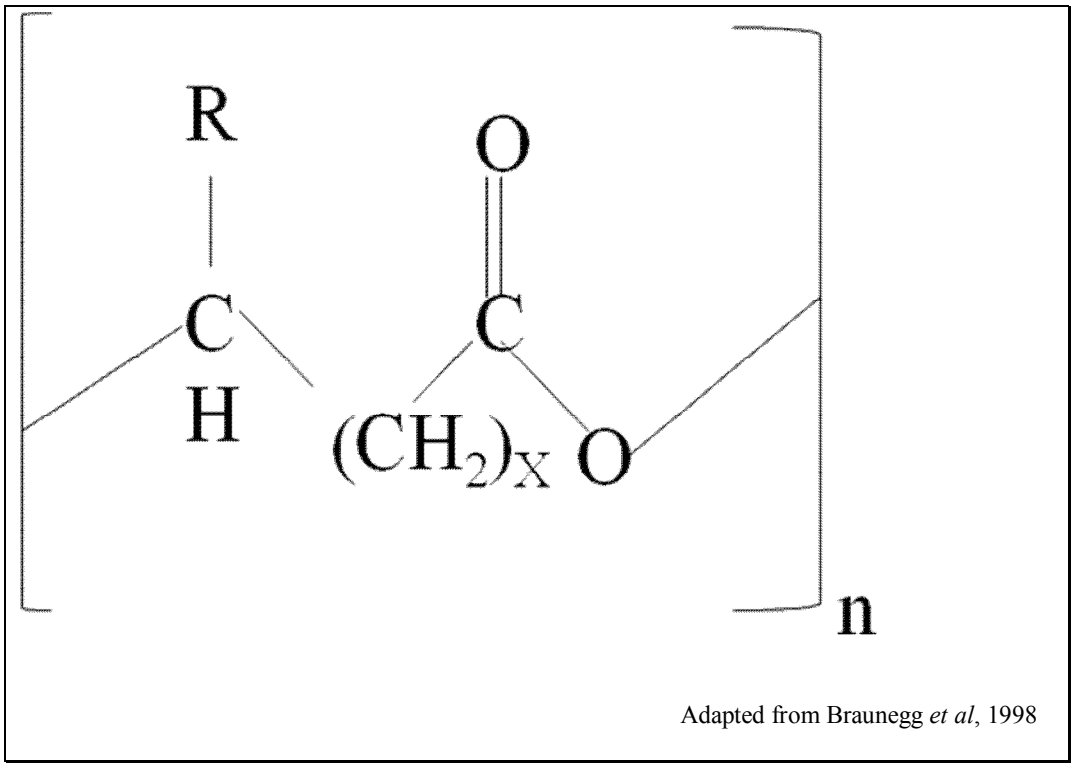

Figure 2.1 General chemical structure of PHA. The composition of the side chain or atom R and the value of X determine the identity of the monomer unit.

2.8.1 Types of Polyhydroxyalkanoates (PHA)

PHAs are abundant in nature and generally can be classified into three different groups based on the size of comprising monomers. PHAs containing three carbons (C3) to five carbon (C5) monomers are classified as short chain length PHA (scl-PHA) while PHAs with C6-C14 monomers are classified as medium chain length (mcl-PHA). PHA with more than C14 monomers are classified as long chain length PHA (lcl-PHA). Copolymers of PHAs which contain more than one type of monomer in a single chain have been widely reported as well (Lee, 1996; Madison and Huisman, 1999; Suriyamongkol *et al.,* 2007; Nomura and Taguchi, 2007). These biopolymers exhibit rather interesting physical and material properties which make it desirable for industrial applications (Rehm and Steinbuchel, 1999).

Short chain length (scl) PHAs are highly crystalline, stiff and brittle (Padermshoke *et al.,* 2005; Verlinden *et al.,* 2007). It exhibits material properties and tensile strength that are close to polypropelene although it has a markedly lower extension to break (Khanna and Srivastava, 2005). This homopolymer has a helical structure and behaves as an elastic material when spun into fibers (Padermshoke *et al.,* 2005; Antipov *et al.,* 2006). It also exhibits other interesting properties such as moisture resistance, water insolubility, optical purity and good oxygen impermeability (Holmes, 1988; Lindsay, 1992; Ojumu *et al.,* 2004).

Medium chain length PHAs (mcl-PHAs) generally are elastomeric semi crystalline polymers with a low melting point, low tensile strength and high elongation to break. It can be used as a biodegradable rubber after cross linking (Preusting *et al.,* 1990; Khanna and Srivastava, 2005; Nomura and Taguchi, 2007).

Copolymers of PHA generally have the same degree of crystalinity as homopolymers are tougher and flexible. Furthermore, the properties of these copolymers can be controlled by adjusting the mole fractions of the co-monomers. This feature alone makes it highly attractive for various industrial commercial applications (Ojumu *et al.,* 2004; Khanna and Srivastava, 2005; Nomura and Taguchi, 2007).

There is another group of PHB which has no important industrial application known to date and is usually ignored due to its obscurity. This group of low molecular weight PHB molecules are usually present at low concentrations in many bacterial and eukaryotic cells such yeast, spinach, sheep intestine and cat's muscle. It is usually found in association with calcium and polyphosphate ions. This molecule has been known to play a role in forming ion channels and also has been postulated to be involved in *E. coli* competence acquisition. However, the synthesis and genetics of these molecules remain a mystery till today (Reusch, 1995; Das and Reusch, 1999; Reusch, 2000; Zin *et al.,* 2001; Rehm, 2003; Addison *et al.,* 2004; Rehm, 2007).

2.8.2. Unique Features of Polyhydroxyalkanoates (PHA)

PHAs possess several unique features besides the interesting physical and material properties mentioned earlier that make it stand out among all biopolymers. Firstly, these biopolyesters are readily degraded by various microorganisms in the environment that produce PHA hydrolases and PHA depolymerases thus making it completely biodegradable (Jendrossek and Hondrick, 2002; Choi *et al.,* 2004; Verliden *et al.,* 2007). Secondly, the production of these biopolymers is based on

renewable sources such as sugars, plant oils and CO_2 (Sudesh *et al.,* 2000) instead of fossil fuels (Braunegg *et al.,* 2004; Gavrilescu and Chisti, 2005). Thirdly, these biopolymers are biocompatible meaning that it does not have toxic effects in living organisms (Volova *et al.,* 2003). This comes as no surprise because lower molecular weight PHB have been found to be a normal constituent of many prokaryotic and eukaryotic cells (Seebach and Fritz, 1999).

2.8.3 Industrial Applications of Polyhydroxyalkanoates (PHA)

PHAs are used in a wide range of applications which include medical, agricultural, pharmaceutical, food and manufacturing industries (Rehm and Steinbuchel, 1999). These biopolymers are particularly important in packaging industries where it is used to manufacture containers and films (Bucci and Tavares, 2005). It has also been used as raw materials to manufacture biodegradable personal hygiene articles such as diapers (Noda, 2001). PHAs have also been processed into toners for printing and adhesives for coating (Madison and Huisman, 1999). It has also been reported that composites of bioplasticss are used in electronic products such as mobile phones (Verlinden *et al.,* 2007).

There are huge potential applications of PHAs in the field of agriculture. Examples of potential applications include encapsulation of seeds, encapsulation of fertilizers for slow release, biodegradable plastic films for crop protection and biodegradable containers for hothouse facilities (Verlinden *et al.,* 2007).

Recently, PHAs have been widely utilized as sutures, repair patches, orthopaedic pins, adhesion barriers, stents and nerve guides in medical applications (Verlinden *et al.,* 2007). This biopolymer has also been used to construct scaffolds in tissue