Isolation and characterisation of genes involved in polyhydroxyalkanoate production in a *Pseudomonas* sp. isolated from Adelaide soil

Shuguang Zhang, M. Sc.

A Thesis Submitted for the Degree of Doctor of Philosophy

Discipline of Microbiology and Immunology
Department of Molecular BioSciences
Adelaide University
Adelaide, SA, 5005
Australia

April, 2001

I dedicate this thesis to my mother Zhenglan Wang,
father Shilin Zhang and
my wife Jing Wang.

Acknowledgments

Firstly, I am indebted to my supervisors, Dr Connor Thomas, Dr. Anton P. J. Middelberg (now at Cambridge University, England) and Dr. David R. G. Williams (Department of Chemical Engineering). As a result of their guidance, encouragement and patience, I have finally completed this project. I also greatly appreciate the support from Connor's computer expertise, from which I learned a lot and which makes me feel my computer knowledge is not enough and there is so much to learn. I also thank Professor Barry Scott for his help.

I would like to thank the Department of Molecular BioSciences, for providing me the excellent facilities for my work as a Ph.D. student. I gratefully acknowledge the support of the Adelaide University for providing me an Overseas Postgraduate Research Scholarship (OPRS) and an Adelaide University scholarship (AUS).

The following people were also generous in assisting me with various aspects of this work:

Ailong Huang for helpful discussions during his stay in this department as a visiting scholar. Mark Webster and Garry Penny for assistance with some techniques, especially at the beginning of my project. Yao Ling and Richard van Wegen (Department of Chemical Engineering) for assistance with gas chromatography analysis of PHA and various other help. Lyn Waterhouse of CEMMSA, for her excellent work with preparation of samples for TEM analysis. The past and present members of Lab 523, Lene, Heather, Rebecca, Fan, Jeng and Kim, a huge thanks for your help and friendship.

A very special thankyou to the family of Dr Fazu Chen for friendship, and the food we enjoyed at your place. Thank you Xi, for your phone calls to Jing, which made her happy when I was not at home.

During the years in Chengdu and in Adelaide, I got so much encouragement from Professor Faqun Zhang and Professor Jiafa Guan. Thank you all very much.

Family is always the sources of strength and love to me. Thanks to my parents, who have worked so many years to support a family with three sons both financially and morally, for suffering so much to bring the sons up and giving them good education with a so long hard time in their lives. I have realised some of your dreams, which you had when you were young but could not realise for a sort of reasons. When I was far away from home, brothers Xiaoguang and Xinguang and sister in-law Li Yang gave my parents so much love and care. Thank you so much. I love you all dearly.

Finally, I think my wife, Jing, for her love, support and encouragement during this project.

This work contains no material which has been accepted for the award of any other degree or diploma in any university or any other tertiary institution. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for photocopying and loan.

20/04/2001

Shuguang Zhang

Errata and Corrections:

Reference to "PHA granules" in various places throughout this thesis, should be replaced by the term "PHA inclusions".

Page 26, Paragraph 3, line 7: replace the words "poor qualities", with "poor physical properties"

Page 30, Paragraph 2, line 1: replace the sentence "Some good strains have been selected for PHA production." by "Isolates able to express high levels of PHA have been described."

Page 86: Insert the sentence "All strains isolated in 1997." as a footnote to Table 3.1.

Page 168, Paragraph 2, Line 7: Add the sentences "Future studies of expression of PHA, should consider using a PHA negative strain of *R. eutropha* as an alternative expression host. This could be achieved through the use of any of a number of broad host range plasmid vectors or by introduction of a PHA expression cassette located within a synthetic transposable element located on a suicide plasmid."

Shuguang Zhang
Tuesday, 24 July 2001

NOTE

Wherever possible, this thesis refers to *Alcaligenes eutrophus* strains as *Ralstonia eutropha* as proposed by Yabuuchi *et al.* (1995).

References:

Yabuuchi, E.; Kosako, Y.; Yano, I.; Hotta, H. and Nishiuchi, Y. (1995). Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. Nov.: Proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff, 1973) comb. Nov., *Ralstonia solanacearum* (Smith, 1896) comb. Nov. and *Ralstonia eutropha* (Davis,1969) comb. Nov. *Microbiology and immunology.* **39**(11), 897-904.

Abbreviations

µg	microgram
µl	microlitre
× g	relative centrifugal force
A	adenine
aa	amino acid(s)
AP	alkaline phosphatase
Ap	ampicillin
APS	ammonium persulphate
ATP	adenosine 5′-triphosphate
bp	base pair
BSA	bovine serum albumin
C	cytosine
cm	centimetre
CTP	cytosine 5′-triphosphate
Cyc	Cycloheximide
ddNTP	dideoxyribonucleotide triphosphate
DEPC	diethyl pyrocarbonate
Dig	digoxigenin
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
DSC	differential scanning calorimetry
DTT	dithiothreitol
dsDNA	Double stranded DNA
E. coli	*Escherichia coli*
EDTA	ethylene-diamine-tetra-acetic-acid disodium salt
EM	electron microscopy
EtBr	ethidium bromide
F	forward
ft	feet

G	guanine
GC	gas chromatography
GTP	guanosine 5′-triphosphate
h	hour
HA	hydroxyalkanoic acid
HGT	high gelling temperature
Hly	haemolysis
IMVS	institute of medical and veterinary science
IPTG	isopropyl-β-D-thio-galactopyranoside
Kan	kanamycin
kbp	kilobase pair
kDa	kilodalton
Klenow	Klenow fragment of *E. coli* DNA polymerase I
kv	kilovolt
L	litre
M	molar
MCS	multiple cloning site
mg	milligram
min	minute(s)
ml	millilitre
mM	millimolar
MOPS	3-[N-Morpholino]propane-sulfonic acid
mRNA	messenger ribonucleic acid
MS	Mass spectrometry
mv	millivolt
NA	Nutrient agar
NB	Nutrient broth
NBT	4-Nitroblue tetrazolium chloride
ng	nanogram
nm	nanometre
NMR	nuclear magnetic resonance
O/N	overnight
OD_{260}	optical density at 260 nm

OD_{600}	optical density at 600 nm
oligo	oligodeoxyribonucleotides
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PHA	polyhydroxyalkanoate
$phaA_{int}$	an internal region of *phaA*
$phaB_{int}$	an internal region of *phaB*
phage	bacteriophage
P(3HB)	poly-β-hydroxybutyrate
pmol	picomoles
R	reverse
RBS	ribosome binding site
RNA	ribonucleic acid
RNase	ribonuclease
Rp	rifampicin
rRNA	ribosomal ribonucleic acid
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulphate
sec	second
SEM	scanning electron micrsocopy
Sm	streptomycin
SM	suspension medium
SSC	standard saline citrate
ssDNA	single stranded DNA
T	thymidine
TAE	tris-acetate EDTA buffer
TBE	tris-boric acid EDTA buffer
Tc	tetracycline
TE	tris-EDTA buffer
TEM	transmission electron microscopy
TEMED	N,N,N',N'-Tetramethyl-ethylenediamine
Tris	tris[hydroxymethyl]amino-methane

Triton X-100	α-[4-(1,1,3,3-tetramethylbutyl)phenyl]-ω-hydroxypoly(oxy-1,2-ethanediyl)
TTP	thymine 5′-triphosphate
UTP	uridine 5′ triphosphates
UV	ultraviolet light
v/v	volume per volume
w/v	weight per volume
X-gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X-pho	5-Bromo-4-chloro-3-indolyl-phosphate

Abstract

Poly-β-hydroxyalkanoate (PHA), as an alternative to the present petroleum-based plastics, is synthesised and accumulated in bacterial cells as a carbon storage through a PHA synthetic pathway, which involves different PHA biosynthetic enzymes. In this study, 19 PHA producing bacterial isolates were isolated from 15 soil samples, which were collected from different sites around Adelaide. PHA production was confirmed by microscopy of Nile Blue A stained cell smears and gas chromatography (GC). GC and differential scanning calorimetry (DSC) analysis showed that one isolate, *Pseudomonas* 10c-1-3 produced PHA heteropolymers at a high level (58.5% w/w) of its dry cell mass and this isolate was selected for further analysis..

The presence of *pha* genes in *Pseudomonas* strain 10c-1-3 was identified by PCR amplification of an internal *phaA* fragment using degenerate primers designed from the conservative regions of *phaA* genes from other bacterial sources. To clone the *pha* genes from *Pseudomonas* strain 10c-1-3, a cosmid library containing 4,556 clones was constructed and screened using PCR to detect the presence of an internal *phaA* fragment. None of the 17 positive clones identified were able to synthesised and accumulate PHA in *E. coli*. One of these cosmids, pCT400 was identified to encode an entire *pha* gene cluster by PCR, Southern hybridisation and DNA sequence analysis. The *pha* genes in pCT400 were subcloned and the *pha* cluster was recovered in plasmid pCT411. PHA production was identified in *E. coli* DH5α harbouring pCT411 and a total of 6.6 kbp of DNA encoding all the genes necessary for biosynthesis of PHA was sequenced and then characterised.

Analysis of the sequence data has identified three open reading frames with potential to encode for polypeptides of 729, 393 and 246 amino acids. The DNA and polypeptide sequences encoded by these ORFs show significant similarity to the *phaC*, *phaA* and *phaB* genes, and protein products, described for *Burkholderia* sp., *Alcaligenes latus*, and *Alcaligenes* sp. SH-69. Furthermore, these open reading frames are of similar size and are arranged in the same order (*phaC*, *phaA* and *phaB*) as that found in *Alcaligenes* sp. SH-69. On the basis of similarity to other genes *phaC* is proposed to

encode a polyhydroxyalkanoate synthase with a molecular mass of 81.3 kDa. Similarly, *phaA* is proposed to encode an acetoacetyl-CoA thiolase (40.6 kDa) and *phaB* an acetoacetyl-CoA reductase (26.5 kDa). DNA sequence analysis showed that there are *phaC*, *phaA*, *phaB* and a fourth open reading frame (ORF4) in this fragment in the above order to form an operon.

PHA production plasmids pCT411, pCT415, pCT416 pCT417 and pCT418 were constructed and used to transform *E. coli* strains. PHA production by *E. coli* DH5α harbouring pCT411, provided evidence that a native promoter in the *pha* operon from *Pseudomonas* strain 10c-1-3 enabled transcription of the *pha* genes in *E. coli*. pCT415, pCT416 and pCT417 were constructed with the *pha* genes under the *LacZ* promoter from the vector pBluescript II KS(+) as IPTG inducible PHA production plasmids and identified by IPTG induction. Plasmids pCT415 and pCT417 contain *phaCAB*ORF4 and pCT416 contains *phaCAB* and a partial ORF4, and PHA accumulation was identified in all the *E. coli* strains harbouring these plasmids respectively, indicating that the ORF4 does not participate in PHA biosynthesis.

Small scale PHA production with different bacterial systems resulted in a low level PHA production. To increase the plasmid stability so as to increase the PHA production, an IPTG inducible PHA production plasmid pCT418 was constructed by the introduction of *parB* locus into pCT415. Plasmid stability examination showed that pCT418 was much more stable than pCT415. Although *pha* clones are able to produce PHA granules in *E. coli*, results which confirmed expression of the *pha* genes in *E. coli* at transcriptional level and at the translational level were not successfully obtained.

The function of ORF4 was investigated by construction of an isogenic ORF4 mutant *Pseudomonas*, PS002. GC and TEM examination indicated that this ORF4 is associated with PHA synthesis and accumulation and involved in PHA granule formation.

Table of Contents

Acknowledgments i
Abbreviations iii
Abstract vii

Chapter 1 Introduction

1.1 General information 2
1.1.1 Petroleum-based plastics and their problems to the environment 2
1.1.2 Solutions for the environmental problem 3
1.1.2.1 Photo-degradable polymers and copolymers 4
1.1.2.2 Biodegradable plastics 4
1.1.2.2.1 Definition of biodegradable polymers 4
1.1.2.2.2 Microbial biodegradable plastics 5
1.1.2.2.3 Chemically synthetic biodegradable plastics 6
1.2 Terminology 6
1.3 Methods for analysis of PHA 8
1.3.1 Light and electron microscopy 8
1.3.2 PHA determination: spectrophotometry and gas chromatography 9
1.3.3 PHA extraction and purification 9
1.3.4 PHA composition and physical properties analysis: MS, NMR, DSC and GPC 10
1.3.5 Biodegradation of PHA 11
1.4 Biosynthesis of PHA 11
1.4.1 PHA producing microorganisms 11
1.4.2 PHA biosynthesis pathways 11
1.4.3 Regulation of PHA biosynthesis 14
1.4.4 Diversity of bacterial polyhydroxyalkanoate 15
1.4.5 PHA granules 15
1.4.6 PHA biodegradation 16
1.5 Genetics of PHA synthesis pathway 17
1.5.1 Screening strategies for cloning *pha* genes 18

1.5.2 Genetics of PHA synthesis pathway 20
1.5.3 Classification of PHA synthases 24
1.6 Physical properties and application of PHA 26
1.6.1 Physical properties of P(3HB) and strategies for property improvement 26
1.6.2 Application of PHA 29
1.7 Industrial scale production by bacteria 29
1.7.1 Strain selection criteria for large scale PHA production 29
1.7.2 PHA production from naturally occurring bacteria 30
1.7.3 PHA production with recombinant *E. coli* 32
1.7.3.1 Development of recombinant *E. coli* strains 32
1.7.3.2 Development of PHA expression systems 34
1.7.4 Media for bacterial PHA production 39
1.7.5 Industrial scale PHA production methods and technologies 41
1.7.6 A systematic strategy for PHA production system construction on an industrial scale 41
1.8 Production and disposal 45
1.9 Specific aims of this thesis 46

Chapter 2 Materials and Methods

2.1 Chemicals and reagents 48
2.2 Enzymes and antibodies 59
2.3 Culture Media 49
2.3.1 General growth media 49
2.3.2 Media for isolation of PHA producing bacteria 49
2.3.3 PHA production media 50
2.4 Antibiotics 50
2.5 Maintenance and propagation of bacterial strains 50
2.6 Characterisation of PHA producing isolates 51
2.7 PHA analysis 51
2.7.1 Staining of bacterial cells containing PHA granules 51
2.7.2 Transmission electron microscopy examination of PHA granules 54
2.7.3 Lyophilisation of PHA-containing bacterial cells 54
2.7.4 Dry cell mass determination 54

2.7.5 Gas chromatography (GC) analysis 55
2.7.5.1 Sample preparation for GC analysis 55
2.7.5.2 PHA standard sample preparation 55
2.7.5.3 Gas chromatography 55
2.7.6 Differential scanning calorimetry of P(3HB) 56
2.8 Bacterial strains and vectors 58
2.9 Solutions and buffers 58
2.10 DNA extraction 58
2.10.1 Preparation of genomic DNA 58
2.10.2 Small scale plasmid and cosmid DNA preparation 65
2.10.3 Large scale plasmid (or cosmid) DNA isolation 66
2.10.4 DNA preparation for library screening and subclones identification 67
2.11 DNA analysis and manipulation 68
2.11.1 DNA quantitation 68
2.11.2 Restriction endonuclease digestion of DNA 68
2.11.3 Agarose gel electrophoresis of DNA 68
2.11.4 Calculation of restriction fragments 69
2.11.5 Purification of DNA fragments from agarose gels 69
2.11.6 Dephosphorylation of DNA fragments 70
2.11.7 Synthesis of oligonucleotides 70
2.11.8 Polymerase chain reaction (PCR) 70
2.11.9 *In vitro* cloning 73
2.11.9.1 Cloning of restriction digest fragments 73
2.11.9.2 Cloning of purified PCR product 74
2.11.10 DNA sequence analysis 74
2.11.10.1 DNA sequencing 74
2.11.10.2 Computer analysis of DNA and protein sequences 75
2.12 Transformation 75
2.12.1 Electroporation 75
2.12.2 Chemical transformation 76
2.13 Construction of a cosmid library 76
2.13.1 DNA preparation and *in vitro* packaging 76
2.13.2 Preparation of *E. coli* culture for infection 77

2.13.3 Transduction of *E. coli* 77
2.14 Bacterial conjugation 77
2.15 Southern hybridisation 78
2.15.1 DNA probe preparation 78
2.15.1.1 Random primed Digoxigenin labelling of DNA 78
2.15.1.2 Digoxigenin labelling of PCR product 78
2.15.2 Southern hybridisation analysis 79
2.15.2.1 Southern transfer 79
2.15.2.2 Hybridisation 79
2.15.2.3 Detection 79
2.16 RNA isolation from bacteria 79
2.17 RNA analysis 80
2.17.1 RNA quantitation 80
2.17.2 RNA slot blot and hybridisation 80
2.18 Expression of *pha* genes in *E. coli* 81
2.18.1 Construction of *E. coli* clones for over-expression of *pha* genes 81
2.18.2 Over-expression of *pha* genes 82
2.18.3 SDS polyacrylamide gel electrophoresis (SDS-PAGE) 82

Chapter 3 Isolation of PHA producing bacteria

3.1 Introduction and strategy 84
3.2 Results 85
3.2.1 Isolation and identification of PHA producing bacteria 85
3.2.1.1 Isolation of PHA producing bacteria from soil 85
3.2.1.2 Identification of PHA producing isolates 85
3.2.2 Gas chromatography (GC) analysis of PHA 85
3.2.3 DSC of PHA produced by the isolate 10c-1-3 94
3.2.4 Transmission electron microscopy analysis of PHA granules 94
3.3 Discussion 97

Chapter 4 construction of a *Pseudomonas* 10c-1-3 genomic library and selection of *pha* clones

4.1 Strategies 100

4.1.1 Strategy for the detection of *pha* genes 100

4.1.2 Strategy for genomic library construction 100

4.1.3 Strategy for genomic library screening 101

4.2 Results 101

4.2.1 Detection of *pha* genes 101

4.2.1.1 Design of degenerate primers 101

4.2.1.2 Detection of *phaA* genes by PCR 104

4.2.1.3 Cloning of the PCR products 104

4.2.1.4 Sequence analysis of the cloned PCR product in pSGZ1 104

4.2.1.5 Probe preparation from pSGZ1 by PCR 109

4.2.2 Construction of a cosmid genomic library 109

4.2.2.1 Preparation and evaluation of the library 109

4.2.2.2 The screening of the cosmid library 109

4.2.2.3 Selection of a hybrid cosmid for subcloning of *pha* genes 111

4.3 Discussion 113

Chapter 5 Subcloning, DNA sequence analysis of *pha* genes and construction of plasmids for expression of PHA

5.1 Introduction and Strategy 120

5.2 Results 120

5.2.1 Construction of plasmids pCT401, pCT402 and pCT403, and the detection of the presence of the *pha* genes in these plasmids 120

5.2.2 Nucleotide sequence analysis 127

5.2.3 Analysis of DNA translation products 135

5.2.4 Construction of the PHA expression plasmid pCT411 148

5.2.5 Construction of IPTG inducible PHA production plasmids pCT415, pCT416, and pCT417 151

5.2.6 Detection of the expression of *pha* genes in *E. coli* 158

5.2.7 Improvement for PHA production in *E. coli* 160

5.3 Discussion 161

Chapter 6 Investigation of the function of ORF4

6.1 Strategy 170

6.2 Results 171

6.2.1 Preparation of the mutant *Pseudomonas* strain 10c-1-3 PS 002 171

6.2.1.1 Cloning of ORF4 into pGEM-T 171

6.2.1.2 Cloning of a Kan^r cartridge from pBSL15 171

6.2.1.3 Cloning of the Kan^r cartridge into pCT421 174

6.2.1.4 Construction of a mobilisable vector for the ORF4::*Kan* mutation 174

6.2.1.5 Conjugation and preparation of a mutant of wild type *Pseudomonas* strain 10c-1-3 174

6.2.2 PHA production by wild type *Pseudomonas* strain 10c-1-3 and the mutant *Pseudomonas* strain 10c-1-3 PS002 178

6.3 Discussion 182

Chapter 7 Summary, conclusion and general discussion

7.1 Introduction 187

7.2 Isolation of PHA producing bacteria 188

7.3 Cloning of the *pha* genes 190

7.4 Construction of PHA production plasmids 193

7.5 Future work 194

7.6 Perspective in PHA production plasmid construction 195

Bibliography 198

Appendix A Buffers and solutions 214

Appendix B Media 219

Appendix C DNA sequence alignment analysis of *pha* genes 222

Appendix D Kanamycin resistant gene from Genbank 240

List of figures

Figure	Title	Page
1.1.	General structural formula of polyhydroxyalkanoate.	7
1.2.	Pathway of P(3HB) synthesis and related reaction steps in *R. eutropha.*	13
1.3.	Schematic drawing showing the molecular organisation of relevant genes for PHA biosynthesis.	21
1.4.	Proposed pathway of P(3HB/3HHx) biosynthesis by *A. caviae* from alkanoic acids or oils.	25
1.5.	Production of P(3HB/3HV) by *E. coli fadR atoC* mutant, MD9101 harbouring PHA synthetic genes.	35
1.6.	Molecular model describing the postsegregational killing mediated by the *parB* (*hok-sok*) locus.	37
2.1.	A scheme for the identification of bacteria isolated from soil samples.	52
2.2.	A typical GC calibration curve for P(3HB) concentration (mg/ml) vs peak-area.	57
2.3.	A typical DSC thermogram of P(3HB).	59
3.1.	Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.	88
3.2.	Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.	89
3.3.	Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.	90
3.4.	Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.	91
3.5.	Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.	92
3.6.	Gas chromatogram of PHA extracts prepared from isolate M14c-1.	93
3.7.	DSC thermogram of PHA produced by isolate 10c-1-3.	96
3.8.	Transmission electron micrographs of PHA granules in *Pseudomonas* strain 10c-1-3.	98
4.1.	Alignment of 11 *phaA* DNA sequences and location of degenerate primers used to amplify DNA from *Pseudomonas* strain 10c-1-3 and isolate 2-3-2.	102

4.2.	Alignment of 6 *phaB* sequences and location of degenerate primers used to amplify DNA from *Pseudomonas* strain 10c-1-3 and isolate 2-3-2.	103
4.3.	PCR products from plasmid p4A, and chromosomal DNAs of isolate 2-3-2 and *Pseudomonas* strain 10c-1-3 using primer pairs #4033 and #4034; #4035 and #4036; and #4033, #4036.	105
4.4.	Strategy used for cloning $phaA_{int}$, $phaB_{int}$ PCR products into pGEM-T using degenerate primer pairs #4033 and #4034, #4035 and #4036.	106
4.5.	Restriction analysis of plasmid pSGZ1.	107
4.6.	DNA sequence analysis of the PCR product $phaA_{int}$ cloned in pSGZ1.	108
4.7.	Digestion of hybrid cosmids with *Eco* RI/*Sal* I and *Eco* RI respectively.	110
4.8.	A schematic illustration of the microtitre plate subpooling procedure used to identify PCR positive cosmid clones.	112
4.9.	Cosmid DNA digested with *Apa* I, *Bgl* II, *Eco* RI and *Eco* RV respectively and Southern hybridisation analysis using Digoxigenin labelled $phaA_{int}$ as a probe.	114
4.10.	Digestion of cosmid (50)D6 with *Apa* I.	115
4.11.	DNA sequence analysis of the PCR products from cosmids (36)H9 and (50)D6.	116
5.1.	Subcloning of plasmid pCT400 and construction of plasmids pCT401, pCT402 and pCT403.	121
5.2.	Confirmation of the plasmid constructs, pCT401, pCT402 and pCT403 by restriction endonuclease enzyme digestion.	123
5.3.	PCR amplification of a *phaA* fragment from plasmids pCT401, pCT402 and pCT403.	124
5.4.	Comparison of the insert in pCT401 with the *R. eutropha pha* operon.	125
5.5.	Preparation of Digoxigenin labelled *phaC* and *phaAB* DNA probes from *R. eutropha pha* operon.	126
5.6.	Detection of *pha* genes in pCT401, pCT402, and pCT403 by Southern hybridisation.	128
5.7.	(a) Complete annotated DNA sequence encoding the entire *pha* biosynthetic region of isolate 10c-1-3 and (b) a schematic showing the arrangement of the putative *pha* genes and their promoter of isolate 10c-1-3.	134
5.8.	Unrooted phylogenetic tree derived from a comparison of PHA synthases peptide sequences.	137

5.9.	Alignments of the key functional domains of the predicted PhaA from *Pseudomonas* 10c-1-3 with related proteins.	139
5.10.	Aligned hydropathy profiles of the predicted Pseudomonas 10c-1-3 polypeptide PhaA and selected β-ketothiolases involved in PHA biosynthesis.	140
5.11.	Unrooted phylogenetic tree derived from a comparison of β-ketothiolases.	142
5.12.	Alignment of a key functional domain of PhaB from *Pseudomonas* 10c-1-3 with related proteins.	144
5.13.	Aligned hydropathy profiles of PhaB and selected acetoacetyl-CoA reductases involved in PHA biosynthesis.	145
5.14.	Unrooted phylogenetic tree derived from a comparison of acetoacetyl-CoA reductases.	147
5.15.	Construction of pCT408.	149
5.16.	Construction of an IPTG non-inducible PHA production plasmid pCT411.	150
5.17.	PCR products from pCT400 and pCT411 using primer pair #5577 and #5578.	152
5.18.	GC analysis of PHA extracts from *E. coli* DH5α harbouring plasmid pCT411.	153
5.19.	Construction of the IPTG inducible PHA production plasmid pCT415.	155
5.20.	Effects of IPTG concentration on induction of PHA expression from *E. coli* DH5α harbouring pCT415, pCT416 or pCT417.	157
5.21.	Detection of the expression of the *pha* genes in *E. coli* strains and *Pseudomonas* 10c-1-3 by Northern hybridisation using the Digoxigenin labelled *phaAB* probe.	159
5.22.	Construction of plasmid pCT418.	162
5.23.	Stability of plasmids pCT415 and pCT418 in an *E. coli* DH5α host during culture in NB with or without addition of antibiotic selection.	163
6.1.	Isolation of ORF4 on a *Sph* I-*Sal* I fragment and construction of pCT419 and pCT420.	172
6.2.	Construction of pCT421.	173
6.3.	Construction of a *kan* cartridge located on a *Pst* I fragment in pCT422.	175
6.4.	Construction of ORF4 insertion mutations encoded on plasmids pCT423 and pCT424.	176
6.5.	Construction of pCT425.	177

6.6.	DNA sequence analysis of a PCR product from *Pseudomonas* strain 10c-1-3 PS002 which spans the chromosomal ORF4*::kan* insertion mutation.	179
6.7.	GC profiles of the PHA extracted from the wild type of *Pseudomonas* sp. 10c-1-3 and the ORF4 mutant, PS002.	180
6.8.	TEM micrographs of PHA granules in the cells of wild type and the ORF4 mutant strains of *Pseudomonas* strain 10c-1-3.	181
6.9.	Distribution of cells of *Pseudomonas* strain 10c-1-3 wild type and the ORF4 mutant which contain different numbers of PHA intracellular granules.	183
7.1.	Strategy for the construction of a PHA production plasmid	197
7.2.	Schematic of the designed artificial *pha* operons	197

List of tables

Table	Title	Page
1.1.	Comparison of the promoter regions of *pha* operons and *E. coli* σ^{70}	22
1.2.	Typical properties of PHA	27
1.3.	Summary of PHA production by various microorganisms and culture methods	42
2.1.	Bacterial strains used in this thesis.	60
2.2.	Vectors and plasmids used in this thesis.	62
2.3.	Oligonucleotides used for sequencing and PCR in this study.	71
3.1.	PHA producing isolates from soil samples.	86
3.2.	Characterisation of 19 soil isolates.	87
3.3.	PHA content in dry cell mass of each isolate.	95
5.1.	Comparison of the amino acid composition of *Pseudomonas* 10c-1-3 PHA synthase (PhaC) with other closely related PHA synthase sequences.	136
5.2.	Comparison of the amino acid composition of *Pseudomonas* 10c-1-3 β-ketothiolase (PhaA) with other closely related β-ketothiolases sequences.	141
5.3.	Comparison of the amino acid composition of *Pseudomonas* 10c-1-3 acetoacetyl-CoA reductase (PhaB) with other closely related acetoacetyl-CoA reductase sequences.	146
A.	Description of Genbank sequence entries used in the multiple alignment of *phaA* genes.	222
B.	Description of Genbank sequence entries used in the multiple alignment of *phaB* genes.	227
C.	Description of Genbank sequence entries used in the multiple alignment of *phaC* genes.	230

Chapter 1

Introduction

1.1 General information

1.1.1 Petroleum-based plastics and their problems to the environment

Water-soluble and water-insoluble petroleum-based plastic materials are widely used and are an integral part of contemporary life. Water-soluble plastics are in general speciality polymers with chemical groups that effect water-solubility (eg. carboxyl, hydroxyl, amido, etc.), whereas the water-insoluble plastics are usually referred to as commodity plastics.

Water-soluble polymers are generally low volume products with a wide range of applications. They are used in cosmetics, and in water treatment systems as dispersants, thickeners, detergents, and super absorbents including poly(acrylic acid), poly(acrylamide), poly(vinyl alcohol), and poly(ethylene glycol). Some enter the environment as landfill, along with the insoluble commodity plastics eg. super absorbents in diapers [poly(acrylic acid) and poly(alkylene oxides)], in industrial detergent effluents [eg. poly(vinyl alcohol)] from paper and textile mills. By contrast, water-soluble polymers entering the environment do not cause a serious waste disposal problem, because of the small volume in use, even though most of them are non-biodegradable (Kawai, 1995).

Water insoluble commodity plastics are used to manufacture an extraordinary array of products, including packaging, disposable diaper backing, fishing nets, and agricultural films. These plastics comprise polymers such as polyethylene, polypropylene, polystyrene, poly(vinyl chloride, poly(ethylene terephthalate), and Nylon. Of particular concern is the fact that most of these commodity plastics are non-biodegradable. Consequently these types of plastics pose a major environmental threat.

Plastics compose over 10% (by weight) and 30% (by volume) of solid waste in the USA and Japan (Kawai, 1995). In Korea, 1.3 million tons of waste plastics are generated annually, with this figure increasing by more than 30% per year. In the global environment, these non-biodegradable plastics accumulate at a rate of 25 million tons per year (Lee *et al.*, 1991). Lack of biodegradability and subsequent persistence, impact significantly on the rate of use of landfill sites. Furthermore they constitute a major land surface litter problems. These issues have raised growing concern about future use of

non-degradable polymers and promoted research activity worldwide to either modify current petroleum-based products to promote degradability or to develop new alternatives that are degradable by any or all of the following mechanisms: biodegradation, photo-degradation, environmental erosion and chemical degradation.

1.1.2 Solutions for the environmental problem

Solid wastes can be disposed of in several different ways: landfill, recycling, and incineration.

In most countries landfill appears to be the principal disposal method. However because of the low bulk density of plastics, available landfill is rapidly used (Swift, 1993). Clearly the landfill option will soon be exhausted and this is exacerbated by the long half lives of commodity plastics and lack of physical space. For example, a soft-drink bottle, which is made from poly(ethylene terephthate) is estimated to take greater than 100 years to degrade, and a polystyrene foam cup more than 500 years (Chang, 1994).

Recycling is considered to be a good alternative to disposal of plastic waste. However efficient recycling, as with some other materials, is not possible with plastics (Wegner and Wagemann, 1994). Estimates suggest that only about 25% of plastic waste can be recycled (Chang, 1994), primarily due to contamination with dyes or food wastes. Cascades of reuse for plastics are feasible, and are in part, already employed. For example, plastic scrap can be used to produce products of less demanding quality, because each processing cycle reduces the inherent purity and quality, with concomitant reductions in mechanical properties as a general rule. At the end of their useful lives, plastics should be incinerated to recover the inherent thermal energy, rather than disposal in landfill. A key problem in the recycling of plastics is their degree of dispersion and the purity in which they can be reclaimed. If the energy input and costs of labour to transport, separate and purify, dry and reprocess the plastic scrap exceed the energy value that can be recovered by incineration in a modern environmentally safe power plant, then material recycling becomes unreasonable.

These solutions cannot satisfactorily solve the environmental problem mentioned previously. Another option is the development of degradable plastics, which are either photo- or bio- degradable.

1.1.2.1 Photo-degradable polymers and copolymers

Incorporation of new materials into conventional plastics can make the final products degradable by light and/or by microorganisms. This approach has been used to develop photodegradable plastics and copolymers based on cellulose or starch.

Photodegradable polymers contain a photo-stabilising agent that is active for a predetermined length of time (Gilead, 1985). When this agent becomes inactive, UV induced photo oxidation of the polymer begins. This process results in formation of carbonyl ends on the polymer chains, which in turn supports microbial degradation. Another approach which combines biodegradability with photo-degradability, is to incorporate azo-aromatic units and UV-sensitive keto-groups into the polymer monomers (Daniell and Guda, 1997). The azo-groups protect the UV-sensitive keto-groups but when the plastics are exposed to bacteria, the azo-groups are attacked by bacteria. Ketones are released and the polymer is then open to photo-degradation. However, several technical problems remain to be solved: (1) how to control the start of photo-degradation and (2) how to reduce the degradation time once polymers are buried in landfill disposal systems.

Another approach is based on the fact that copolymers produced with natural polymers like cellulose or starch improve biodegradability (Smock, 1987; and Narayan, 1988). To achieve significant biodegradation of these copolymers, the content of natural polymers should be about 30-50%. However, this results in a product with inferior physical properties that in turn causes problems during processing. When these copolymers are exposed to environment, only the natural polymer content is biologically degraded, while the conventional plastic content remains non-degradable (Swift, 1993).

Consequently, the two types of degradable polymers described are not completely biodegradable and clearly are not ideal alternatives to current commodity plastics.

1.1.2.2 Biodegradable plastics

1.1.2.2.1 Definition of biodegradable polymers

Biologists, biochemists, polymer chemists, engineers, lawyers, legislator, environmentalists, manufacturers, and lay people are involved in biodegradable polymer research and development, each with their own agenda, perspective and opinion on

requirements. This makes it difficult to coin a universally accepted definition for biodegradable polymer. Swift (1993) proposed two equations to describe the biodegradability of a given polymer exposed to aerobic and anaerobic environments.

Aerobic environment:

$$C_t=CO_2+H_2O+C_r+C_b \quad (1)$$

Anaerobic environment:

$$C_t=CO_2+CH_4+H_2O+C_r+C_b \quad (2)$$

where C_t: total carbon content of the polymer;

C_r: carbon residue in the environment;

C_b: carbon converted into biomass.

Complete biodegradation, partial biodegradation, and zero biodegradation are defined by $C_r=0$, $0<C_r<C_t$, and $C_r=C_t$, respectively.

Thus the difference between the definition of a biodegradable polymer and an environmentally acceptable biodegradable polymer is: A biodegradable polymer may be completely or partially degraded by enzymes, while an environmentally acceptable biodegradable polymer must be mineralised or produce no environmentally harmful residues, if it is partially biodegraded. This difference is an issue to be considered when developing environmentally biodegradable polymers.

1.1.2.2.2 Microbial biodegradable plastics

Many microorganisms can produce biopolymers as energy and carbon storage sources. For example, lactic acid polymers can be derived from lactic acid bacteria (Athanasiou *et al.*, 1995). Over the last few decades, considerable interest has been shown in polymers produced by those microorganisms, eg. aliphatic polyesters and polysaccharides, which can be completely degraded into carbon dioxide and water in environment (Lee, 1995). The most useful of all microbially derived biodegradable plastics are aliphatic polyester polymers known as polyhydroxyalkanoates (PHA's). These polymers possess physical properties similar to conventional plastics but can be completely

biodegraded to carbon dioxide and water. The general structural formula of PHA is shown in Figure 1.1.

Most interest has focused on poly-β-hydroxybutyrate [P(3HB)], a homopolymer of PHA, in which the residue is β-hydroxybutyrate. The initial motivation for this interest was a desire to find a cheap alternative to oil-based synthetic plastics especially in light of the oil shortage in the late 70's. At that time, oil prices were predicted to be greater than $100 per barrel in the 1980's, and the chemical company, ICI, began research on P(3HB) (See Figure 1.1). However oil prices dropped, and ICI shifted its interests in P(3HB) to the biodegradability and biocompatibility of the polymer (Byrom, 1990).

1.1.2.2.3 Chemically synthetic biodegradable plastics

In Japan, the Showa Kobunshi company synthesised high molecular weight biodegradable aliphatic polyesters, which seem to have processing advantages over microbial PHA. Polymer companies can use current production facilities to produce these aliphatic polyesters and it has been estimated that the price of these aliphatic polyesters is cheaper than microbial PHA (Chang, 1994).

Research on PHA has been carried out for several decades on many aspects including the metabolism of PHA in bacteria, genetic analysis of *pha* genes involved in PHA biosynthesis and PHA production in different expression systems. In this chapter, these aspects will be discussed.

1.2 Terminology

PHA is a family of biodegradable polymers consisting of different hydroxyalkanoic acids as monomeric units. In this thesis, the following definitions of the terms related to PHA have been adopted (Steinbuchel *et al.*, 1992).

1. Microorganisms produce PHA of different composition, depending on source substrates and their biosynthetic pathways. Where the composition of the polymer is known, the following notation will be used eg. P(3HB) – a homopolymer, poly-β-hydroxybutyrate; P(3HB/3HV) – a copolymer, poly(β-hydroxybutyrate-co-β-hydroxyvalerate); or P(3HB)/P(3HO) – a blend of two different polymers, poly-β-hydroxybutyrate and

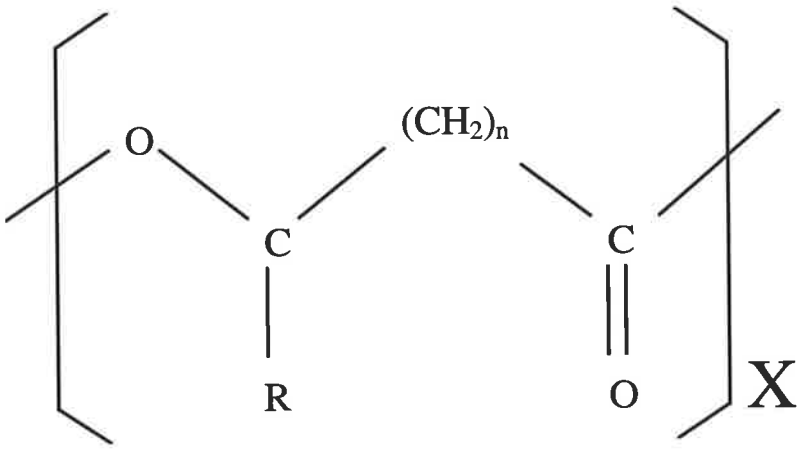


Figure 1.1. General structural formula of polyhydroxyalkanoate (Lee and Chang, 1995).

X=100-30,000

n=1,	R= $-CH_3$	→ poly-β-hydroxybutyrate [P(3HB)]
	R= $-CH_2-CH_3$	→ poly-β-hydroxyvalerate [P(3HV)]
	R= $-(CH_2)_4-CH_3$	→ poly-β-hydroxyoctanoate [P(3HO)]
n=2	R= H	→ poly-4-hydroxybutyrate [P(4HB)]
	R= $-CH_3$	→ poly-4-hydroxyvalerate [P(4HV)]
n=3	R= H	→ poly-5-hydroxyvalerate [P(5HV)]

poly-β-hydroxyoctanoate. If the composition has not been known, the polymer is referred to as PHA.

2. SCL refers to short-chain-length hydroxy-alkanoic acids (HAs), which consist of 3-5 carbon atoms, MCL to medium-chain-length hydroxy-alkanoic acids, which consist of 6-14 carbon atoms, and LCL to long-chain length hydroxy-alkanoic acids, which consist of more than 14 carbon atoms.

3. Each of the enzymes involved in the steps of PHA synthesis does not have their substrate specificity. They are referred generally to as, eg. PHA β-ketothiolase, PHA reductase and PHA synthase which catalyse synthesis of PHA (described in Section 1.4.2). Subscripts (eg. PHA_{SCL} synthase) are used to indicate substrate specificity of PHA synthase, if PHA synthase exhibits a preference towards HA_{SCL}-CoA thioesters, HA_{MCL}-CoA thioesters or HA_{LCL}-CoA thioesters.

4. To identify the structural genes encoding the enzymes involved in PHA synthesis from different microorganisms, *phaA*, *phaB* and *phaC* refer to the genes encoding β-ketothiolase, reductase and PHA synthase respectively. If there are more than one same genes in a genome, these genes are indicated by an Arabic number which follows the last letter (eg. *phaA1*, *phaA2*,...; *phaB1*, *phaB2*, ...; *phaC1*, *phaC2*,...). Other genes coding for proteins involved in the PHA synthesis or associated with the PHA granules are referred to consecutively as *phaD*, *phaE*, etc. Open reading frames (ORFs) coding for proteins with unknown functions are referred to and indicated by an Arabic number which follows the last letter as ORF1, ORF2, etc.

1.3 Methods for analysis of PHA

1.3.1 Light and electron microscopy

Presence of PHA granules in bacterial cells has been traditionally confirmed by staining with Sudan Black B. Sudan Black B mainly stains lipids and is less reliable when for PHA. Ostle and Holt (1982) used Nile Blue A, a water-soluble basic oxazine dye that has a greater affinity and higher specificity than Sudan Black B for P(3HB). Nile Blue A gives a bright orange fluorescence at a wavelength of 460 nm, when PHA granules are

stained and has the advantage that it does not stain other inclusion bodies eg. glycogen, and polyphosphate. This is useful for identification of PHA granules within bacterial cells.

Electron microscopy is usually employed for the analysis of morphological alteration of intracellular PHA granules and the extracted PHA as a material. Molitoris *et al.* (1996) studied degradation of PHA by bacteria using SEM and found that PHA hydrolysis started at the surface and at physical lesions in the polymer granule. The structure of crystallised PHA has also been studied with SEM (Alper, and Lundgren, 1963; and Barham *et al.*, 1984).

1.3.2 PHA determination: spectrophotometry and gas chromatography

Spectrophotometric methods for quantitative analysis and determination of PHA have been established (Law and Slepecky, 1961). When heated with concentrated sulfuric acid, PHA converts to crotonic acid. Crotonic acid is then assayed at A_{235nm}. This method is relatively inaccurate and time-consuming. Braunegg *et al.* (1978) established a direct and rapid gas chromatographic method for P(3HB) determination in microbial biomass without P(3HB) extraction. This method is characterised by high precision (detectable as low as 10-5 g PHA/L) and a short time (4 h) for P(3HB) determination.

1.3.3 PHA extraction and purification

B. megaterium (Merrick and Doudoroff, 1964) and *Azotobacter beijerinckii* (Ritchie and Dawes, 1969) were first employed for the isolation of native PHA granules. DNase-treated cell extracts were repeatedly centrifuged with glycerol and PHA granules were collected from the glycerol water interface. Fukui *et al.* (1976) isolated crude PHA granules of *Zogloea ramigera* from sonically disrupted cells by centrifugation at $10^5 \times g$ for 60 min. The crude granules were then resuspended into 50 mM potassium phosphate buffer (pH7.0), and fractionated by sucrose density gradient centrifugation.

PHA is usually extracted from microorganisms using chlorinated hydrocarbons. Dry biomass is refluxed with chloroform and the resulting solution is filtered to remove cellular debris. The PHA is then concentrated by evaporation of solvent. The polymer is precipitated with methanol, ethanol or diethyl ether, leaving the low molecular-weight lipids in solution (Slater *et al.*, 1992).

1.3.4 PHA composition and physical properties analysis: MS, NMR, DSC and GPC

Mass spectrometry (MS) and nuclear magnetic resonance (NMR) techniques have been successfully used to determine polymer composition and the distribution of monomer units. Gas chromatography (GC) and mass spectrometry (MS) are often used as tools for determination of PHA and identification of structure.

Lee and Choi (1997) analysed the methanolysis products of PHA containing 4HB (4-hydroxybytyric acid), 4HV (4-hydroxyvaleric acid), and 4HHx (4-hydroxyhexanoic acid) with GC/MS and detected two major chromatographic peaks with characteristic retention times of each methyl ester of 4HA and the corresponding λ-lactone (λ-butyrolactone, λ-valerolactone, λ-caprolactone, respectively). They suggested that this method could be incorporated into an efficient screening procedure for isolation of bacterial strains which could accumulate PHA containing 4HAs as the principal monomer from structurally related carbon substrates.

Doi *et al.* (1986) analysed PHA in chloroform solution by 500-MHz ^{1}H NMR spectroscopy, and determined the sequence distribution of the monomeric units in P(3HB/3HV) produced by *R. eutropha* by analysing 125-MHz ^{13}C spectra.

PHA is a partially crystalline polymer. In its amorphous state, there is a glass-to-rubber transition temperature (T_g) and in the crystalline phase, it has a defined melting temperature (T_m). T_g and T_m are terms used to express the thermal and mechanical properties of PHA. Differential Scanning Calorimetry (DSC) is used for the determination of T_g and T_m. Microbial P(3HB) is in a highly crystalline state after extraction due to its exceptional native stereochemical regularity with a degree of crystallinity in the range of 60-80% and a Tm of 174 °C, which decreases by 30 to 40% (Tm of 143 °C), if the 3HV content in P(3HB/3HV) increases to 30 mol% (Mitomo *et al.*, 1987). T_g of P(3HB) is usually in the range -5 to 5 °C (Barham *et al.*, 1984) and 0 to 20 °C (Cesaro and Scandola, 1989; and Scandola *et al.*, 1988) and it decreases if the 3HV content in P(3HB/3HV) is increased.

Barham *et al.* (1984) used gel permeation chromatography (GPC) in chloroform at 30 °C with a set of five micro-styragel columns to determine the viscosity-average

molecular weight of P(3HB) from *R. eutropha*. The standard molecules used were of polystyrene and the molecular weight of P(3HB) was calculated by applying the Mark-Houwink relationship equation $[\eta]=KM^{\alpha}$, where $[\eta]$ is the extrinsic viscosity, M is the molecular weight and K and α are constants for the particular solute-solvent-temperature combination respectively.

1.3.5 Biodegradation of PHA

Biodegradability is a very important property to be considered in development and evaluation of polymers. These properties can be characterised by enzymatic, microbial and field burial methods. Enzymatic methods are rapid, but require use of specific enzymes which limit their application. Microbial methods are very similar to the enzymatic methods but they take a long time. Burial methods demonstrate the ultimate polymer degradability, but require a long incubation time and results obtained vary with the burial sites. However, the test conditions for burial methods are representative of those found in landfills and consequently are probably the most desirable (Chang, 1994).

1.4 Biosynthesis of PHA

1.4.1 PHA producing microorganisms

P(3HB) is the most prominent PHA and was first described in 1925 by Lemoigne in *Bacillus megaterium* (Muller and Seebach, 1993). By 1991, approximately 300 different bacteria have been reported to synthesise various PHA (Steinbuchel, 1991), including the *Pseudomonas* species, *Rhodospirillum rubrum, Halobacterium mediterrana, Rastonia eutropha, Rhizobium melioti, Zoogloea ramigera* and *Corynebacterium*. Enterobacteria are exceptions and do not normally synthesise P(3HB) as a storage compound.

1.4.2 PHA biosynthesis pathways

In the organisms investigated by the early 90's, four different pathways are used for the synthesis of PHA under unbalanced growth conditions (Steinbuchel and Schlegel, 1991; and Anderson and Dawes, 1990).

The first type of PHA biosynthetic pathway was found in *R. eutropha*, which comprises a three step pathway for P(3HB) synthesis from acetyl coenzyme A (acetyl-CoA) and is probably present in the majority of PHA-accumulating bacteria (Steinbuchel and Schlegel, 1991). Two acetyl-CoA molecules are firstly condensed to acetoacetyl-CoA by β-ketothiolase (EC 2.3.1.9). Secondly, an NADPH-dependent acetoacetyl-CoA reductase (EC 1.1.1.36) catalyses the reduction of acetoacetyl-CoA to D(-)-β-hydroxybutyryl-CoA which is finally catalysed by P(3HB) synthase and linked to the growing chain of P(3HB) (Muller and Seebach, 1993; and Anderson *et al.*, 1990) (Figure 1.2). β-ketothiolase and NADPH-dependent acetoacetyl-CoA reductase are soluble proteins. PHA synthase is soluble only as long as no P(3HB) accumulation occurs and it becomes granule-associated under storage conditions. A second β-ketothiolase (EC 2.3.1.16) and an NADH-dependent acetoacetyl-CoA reductase (EC 1.1.1.35) have also been detected, and these may play a role in the β-oxidation of fatty acids (Haywood *et al.*, 1988a,b; and Steinbuchel and Schlegel, 1991).

A similar type of PHA biosynthetic pathway has been found in *Rhodospirillum rubrum*, but comprises five steps. Two acetyl-CoA molecules are condensed to acetoacetyl-CoA by β-ketothiolase (EC 2.3.1.9). Secondly, an NADH-dependent acetoacetyl-CoA reductase catalyses the reduction of acetoacetyl-CoA to L(+)-β-hydroxybutyryl-CoA which is then converted to D(-)-β-hydroxybutyryl-CoA by two enoyl-CoA hydratases. Finally D(-)-β-hydroxybutyryl-CoA is catalysed by P(3HB) synthase and linked to the growing chain of P(3HB).

The third type of PHA biosynthetic pathway is active in most of pseudomonads belonging to the rRNA homology group I. For instance, *Pseudomonas oleovorans* synthesises PHA_{MCL} with MCL-β-hydroxyalkanoates that are derived from intermediates of the β-oxidation of fatty acids during cultivation on various MCL-alkanes, alkanols, or alkanoates.

Most pseudomonads belonging to the rRNA homology group I, except *P. oleovorans,* possess an additional pathway for the synthesis of copolymers consisting of MCL-β-hydroxybutyrate from acetyl-CoA, which is known as the fourth PHA biosynthetic pathway. *P. aeruginosa* as an example, synthesises a copolymer consisting of

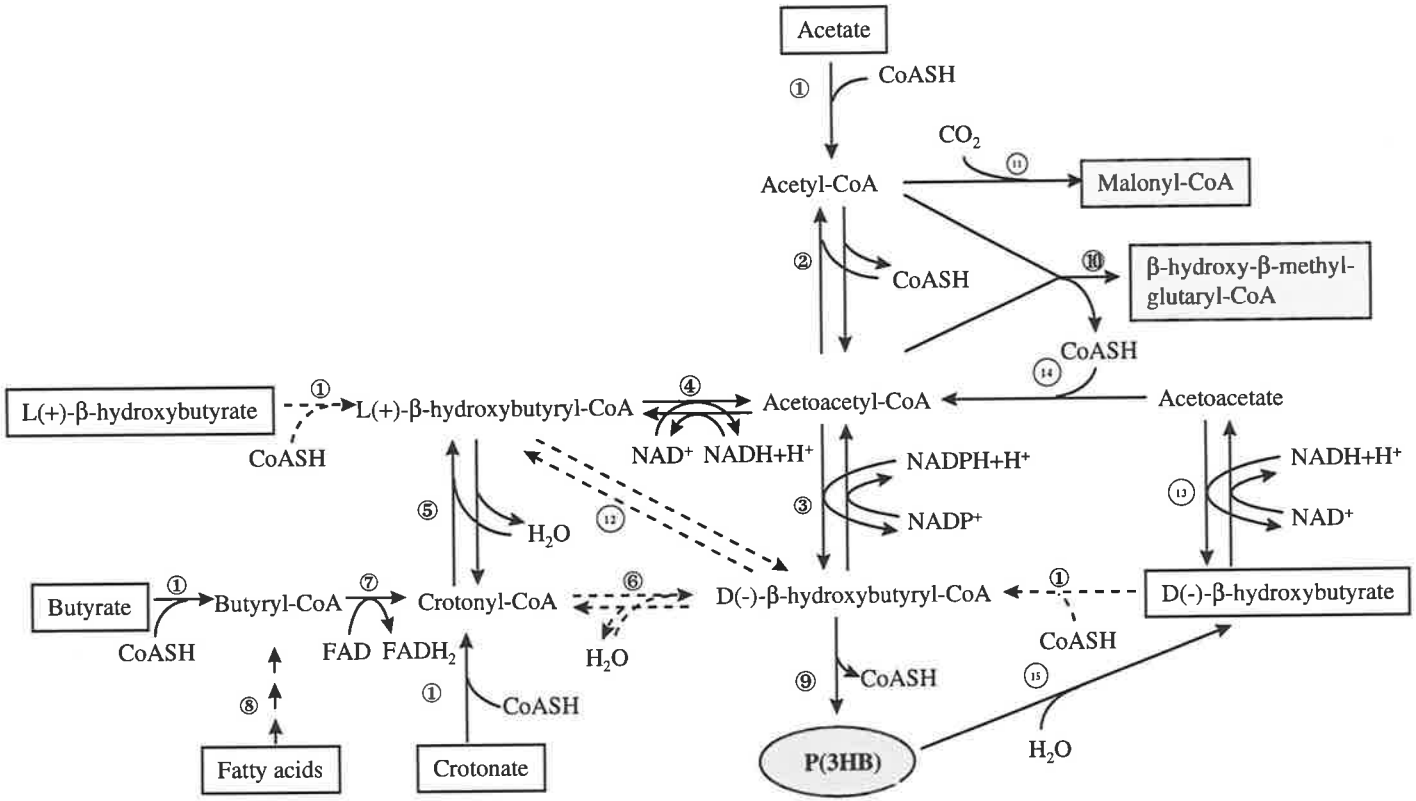


Figure 1.2. Pathway of P(3HB) synthesis and related reaction steps in *R. eutropha* (Adapted from Steinbuchel and Schlegel, 1991).

1. Acyl-CoA synthetases (it is uncertain whether specific synthetases exist for all possible substrates); 2. β-ketothiolase; 3. acetoacetyl-CoA reductase (NADPH-dependent); 4. acetoacetyl-CoA reductase (NADH-dependent); 5. enoyl-CoA hydratase [forming L-(+)-β-hydroxybutyryl-CoA]; 6. enoyl-CoA hydratase [forming D-(-)-β-hydroxybutyryl-CoA]; 7. butyryl-CoA dehydrogenase; 8. enzymes involved in the β-oxidation pathway; 9. P(3HB) synthase; 10. hydroxymethylglutaryl-CoA synthase; 11. acetyl-CoA carboxylase; 12. β-hydroxybutyryl-CoA epimerase; 13. D(-)-β-hydroxybutyrate dehydrogenase (NAD-dependent); 14. CoA transferase; 15. P(3HB) depolymerase; Uncertain reactions in *R. eutropha* are indicated by broken lines.

β-hydroxydecanoate as the main constituent, with β-hydroxydodecanoate and β-hydroxyoctanoate being minor constituents as a result of cultivation on gluconate.

1.4.3 Regulation of PHA biosynthesis

All the three P(3HB) biosynthetic enzymes, β-ketothiolase, NADPH-dependent acetoacetyl-CoA reductase, and P(3HB) synthase are constitutively synthesised in *R. eutropha,* even under conditions that do not restrict growth. It is not known whether expression of P(3HB) synthetic genes is regulated at the transcriptional or translational level. However, the regulation of P(3HB) accumulation does occur at the enzyme level (Steinbuchel and Schlegel, 1991). For example, in the P(3HB) biosynthetic pathway used by *R. eutropha*, β-ketothiolase initiates the first step of P(3HB) biosynthesis. Evidence indicates that β-ketothiolase plays a dominant role in the regulation of P(3HB) synthesis. Oeding and Schlegel (1973) showed that free Coenzyme A inhibits the activity of β-ketothiolase. Condensation of two acetyl-CoA molecules to acetoacetyl-CoA by β-ketothiolase is competitively inhibited by Coenzyme A, whereas the activity of the β-ketothiolase in the degradative pathway is inhibited by acetoacetyl-CoA (Muller and Seebach, 1993). Under unlimited growth conditions, cellular concentration of free coenzyme A is high, and this inhibits synthesis of β-ketothiolase and prevents P(3HB). However, under N, P, S, K or Fe limiting growth conditions, intracellular protein synthesis ceases, resulting in an increase in the intracellular NADPH concentration (Schlegel *et al.*, 1961; Repaske and Repaske, 1976; and Steinbuchel and Schlegel, 1989). Elevated NADPH concentration reduces the efficiency of acetyl CoA degradation by the tricarboxylic acid cycle, resulting an increased concentration of acetyl CoA available for diversion into the P(3HB) biosynthetic pathway (Anderson and Dawes, 1990; and Muller and Seebach, 1993). Under nitrogen-limited conditions, a 2.2-fold increase in the level of acetyl-CoA was detected in cells (Steinbuchel and Schlegel, 1991). Under these limiting growth conditions, the concentration of free Coenzyme A is presumably low, and the activity of biosynthetic β-ketothiolase is released to condense acetyl-CoA molecules to acetoacetyl-CoA molecules. Excretion of pyruvate by P(3HB)-negative mutants cultivated in the presence of fructose, gluconate, lactate, or CO_2 plus H_2 under limiting conditions, under which the wild type produced P(3HB), indicated that the regulation of β-ketothiolase

by free coenzyme A is a main regulatory valve controlling the flow of metabolites (Steinbuchel and Schlegel, 1991).

The biosynthesis of P(3HB) is also a sink for reducing equivalents. When *R. eutropha* was cultivated under conditions without any nutrient depletion, but limited oxygen supply, P(3HB) synthesis begins immediately (Steinbuchel and Schlegel, 1991). Under limited growth conditions, increased concentrations of NADPH assist in the action of the NADPH dependent acetoacetyl-CoA reductase of the P(3HB) biosynthetic pathway. The P(3HB) synthase remains soluble and inactive until polymer synthesis is underway. In the presence of P(3HB) synthesis, the P(3HB) polymerase becomes active and granule-associated (Anderson and Dawes, 1990).

1.4.4 Diversity of bacterial polyhydroxyalkanoate

By 1995, about 91 different monomers had been identified as constituents of PHA synthesised by various bacteria (Steinbuchel and Valentin, 1995). Theoretically, these 91 different constituents could arbitrarily combine to produce several million possible polymers. However, only few monomers are synthesised by bacteria under any one set of growth conditions. The most common polymer found is P(3HB).

Production of novel PHA seems to be limited by the availability of carbon sources supplied to PHA producing bacteria. Conversion of carbon in growth media to PHA involves two key steps: conversion of substrates to certain intermediates (precursor substrates) (which varies with different bacteria due to differences in their metabolism), and recognition of these intermediates by the PHA synthesis enzymes. Therefore, one way to obtain novel PHA is to isolate bacteria with novel metabolic pathways capable of converting carbon sources to different precursor substrates followed by synthesis to PHA.

1.4.5 PHA granules

PHA granules are typically 0.2 to 0.5 μm in diameter and possess a membranous coat about 2 nm thick, composed of lipid and protein, representing about 0.5 to 2%, respectively, of the granule weight (Lundgren *et al.*, 1964). Both PHA synthase and depolymerase are known to be associated with this coat (Anderson and Dawes, 1990).

Granule-associated PHA synthase has been described in *R. Rubrum* (Merrick and Doudoroff, 1961), *B. Megaterium* (Griebel *et al.*, 1968), *Azotobacter beijerinckii* (Ritchie and Dawes, 1969; and Ritchie *et al.*, 1971), *and Z. Ramigera* (Fukui *et al.*, 1976; and Tomita *et al.*, 1983). Wieczorek *et al.* (1996) investigated the PHA granule-associated proteins of fifty different PHA-accumulating bacterial strains and found that the PHA granules of many of these strains exhibited a similar protein pattern. However, PHA synthases constitute either a minor protein at the granule surface or they are readily dissociated from the surface during the preparation of the granules. Other proteins, which are very tightly bound to the granules, represent the predominant protein components in the PHA granules isolated from most species. In *R. eutropha* strain H16 and *R. ruber*, two proteins (24-kDa and 14-kDa) predominate the surface of PHA granules. These proteins apparently determine the size, number of granules and the efficiency of PHA accumulation in cells. McCool and Cannon (1999) investigated the PHA granule-associated proteins and cloned the respective genes from *Bacillus megaterium*. One of these, PhaP, is an extremely hydrophilic protein. The presence of these proteins suggested that PHA granules are not only a source of carbon, energy, and reducing equivalents, but also a source of amino acids.

Ballard *et al.* (1987) investigated the PHA granules in *R. eutropha* and found that when the PHA content was about 80% of cell volume, PHA production ceased, even though PHA synthase activity was still high. This observation indicates that physical constraints limit PHA production within a cell wall.

PHA granules synthesised in microorganisms such as *R. eutropha* and *Bacillus megaterium* are typically spherical with a diameter of 0.1-0.8 μm (Ellar *et al.*, 1968; and Lafferty *et al.*, 1988). Investigation of the sizes of PHA granules in *E. coli* showed that there was little variation in the mean diameter of 1.13-1.25 μm of PHA granules synthesised by different strains under different culture conditions (Middelberg *et al.*, 1995).

1.4.6 PHA biodegradation

PHA can be degraded by intracellular depolymerases, expressed by the PHA accumulating strains, in the absence of a suitable exogenous carbon source, or extracellularly by extracellular depolymerases. The latter are secreted by many bacteria as

a means of utilising PHA left in the environment after PHA producing bacteria have died. Intracellular PHA depolymerases are unable to hydrolyse extracellular PHA, and extracellular depolymerases cannot hydrolyse intracellular granules. This is apparently due to differences in the physical structures of intracellular "native" granules and extracellular "denatured" PHA. The latter are highly crystalline polymers. However, intracellular PHA granules are completely amorphous and covered by a surface layer consisting of proteins and phospholipids, and this may explain resistance to extracellular depolymerases. The surface layer is lost or damaged during rough isolation of the granules, or by treatment with solvents, or by other chemical or physical stresses (Jendrossek *et al.*, 1996).

PHA-degrading bacteria have been found in nearly all terrestrial and aquatic ecosystems and all of them are able to accumulate PHA if they are cultivated under appropriate conditions. Many PHA-degrading fungi (including 91 genera) have also been identified (Jendrossek *et al.*, 1996). Intracellularly, PHA is enzymatically degraded at the granule surface, to low molecular weight fragments, which are rapidly metabolised by the microbial cells. PHA can also be degraded to methane and carbon dioxide in an anaerobic environment. Budwill *et al.* (1992) investigated the degradation of P(3HB) and P(3HB/3HV) by an anaerobic sewage sludge consortium and found that the polymers were completely degraded within 16 days under laboratory conditions.

The extracellular PHA depolymerases of many bacteria and the fungi, *Penicillium funiculosum* and *Fusarium solani* have been purified and characterised. The molecular biology of some PHA depolymerases has also been studied (Jendrossek *et al.*, 1996).

1.5 Genetics of PHA synthesis pathway

In the late 1980's, three research groups (Slater *et al.*, 1988; Peoples and Sinskey,1989b,c; and Schubert *et al.*, 1988) used different strategies to clone *pha* genes from *R. eutropha* and express these genes in *E. coli*. Their work represented a breakthrough in PHA research. Subsequently, *pha* genes from a variety of different bacteria have been identified, characterised and expressed in recombinant hosts.

1.5.1 Screening strategies for cloning *pha* genes

Strategy A: Enzyme activity detection

Slater *et al.* (1988) constructed a cosmid library of *R. eutropha* and screened this library by detection of clones expressing β-ketothiolase activity. Some clones, which were β-ketothiolase activity positive, also harboured the structural genes for acetoacetyl-CoA reductase and PHA synthase.

Strategy B: Screening for a target gene with a marker introduced by insertion mutagenesis

In the absence of DNA sequence data that might facilitate cloning PHA associated genes, these target genes can be identified by tagging with genetic markers such as transposons. Loss of PHA expression or associated enzyme activity can then be used to identify *pha* mutants. A transposon associated marker gene can then be used to clone flanking DNA, which in turn can be used to identify clones from a gene library. This strategy was successfully used to clone the PHA-biosynthetic genes of *R. eutropha* (Schubert *et al.*, 1988). Tombolini *et al.* (1995) also successfully used this strategy for cloning a *phaC* gene from a mutant of *Rhizobium meliloti* 41, in which *phaC* was inactivated by the insertion of transposon Tn*5*.

Strategy C: Southern hybridisation with a homologous DNA fragment

Alignment of PHA-biosynthetic genes indicates a high degree of similarity at the DNA level. Thus heterologous DNA probes should allow identification of PHA-biosynthetic genes using standard Southern hybridisation methods. For example, the gene encoding the biosynthetic β-ketothiolase from *Z. ramigera* was used as a heterologous probe to identify a genomic *Eco*RI restriction fragment harbouring the PHA-biosynthetic genes of *R. eutropha* (Peoples and Sinskey, 1989b). Similarly, *R. eutropha pha* genes have been successfully used to identify genomic DNA fragments of *Chromatium vinosum* D (Liebergesell and Steinbuchel, 1992) *Thiocapsa pfennigii* (Steinbuchel *et al.*, 1992), *Thiocystis violacea* (Liebergesell and Steinbuchel, 1993), and *Ectothiorhodospira shaposhnikovii* (Steinbuchel *et al.*, 1992).

As an alternative to DNA probes, oligonucleotides homologous to highly conserved regions of *pha* genes can be used to probe the presence of uncharacterised *pha* genes. A ^{32}P-labelled 30-mer oligonucleotide was designed from the a highly conserved region of the PHA synthases of *R. eutropha* and *P. oleovorans* and synthesised to probe the *pha* genes of *P. aeruginosa* (Timm and Steinbuchel, 1992), *P. citronellolis*, *P. mendocina*, *Pseudomonas* sp. DSM 1650 and *Pseudomonas* sp. GP4BH1 by hybridisation of cosmid libraries (Timm *et al.*, 1994).

Strategy D: Phenotypic complementation

PHA-biosynthetic genes can be also identified by screening a genomic library for phenotypic complementation of PHA-negative mutants, or for conferring the ability to synthesise and accumulate PHA in a PHA-negative wild type background. The PHA-negative mutant P(3HB)-4 of *R. eutropha* was employed to identify the PHA synthase genes from *Lamprocystis roseopersicina* (Steinbuchel *et al.*, 1992), *Rhodospirillum rubrum* (Hustede *et al.*, 1992), *Rhodobacter sphaeroides* (Hustede *et al.*, 1992), *Methylobacterium extorquens* (Valentin and Steinbuchel, 1993) and *Rhodococcus ruber* (Pieper and Steinbuchel, 1992). A PHA-negative mutant of *P. putida* was employed to identify the *pha* loci of *P. oleovorans* and *P. putida* (Huisman *et al.*, 1991). PHA-synthetic genes of *Syntrophomonas wolfei* (McInerney *et al.*, 1992) and *Alcaligenes latus* (Choi *et al.*, 1998) were identified by screening PHA producing *E. coli* clones.

However, this strategy is not always successful when *E. coli* is used as the PHA-negative wild type background. Tombolini *et al.* (1995) subcloned three *pha* genes from a cosmid library prepared from *Rhizobium meliloti* 41 and constructed a plasmid for expression of PHA. When *E. coli* was used as a host, high levels of PHA were expressed. However, the *E. coli* strain harbouring the cosmid, from which the *pha* genes were subcloned, and the PHA production plasmid constructed, did not produce PHA even though all the necessary *pha* genes were present. The reason for this phenomenon is unknown.

Strategy E: Application of PCR

As discussed in strategy C, oligonucleotides can be designed based on highly conserved regions of *pha* genes. These can be used to prime PCR amplification of target

regions within previously uncharacterised *pha* genes. If needed, the amplified PCR products can be further sequenced. This approach has been exploited by Lopez *et al.* (1997) to amplify a partial *phaC* fragment from different bacteria.

1.5.2 Genetics of PHA synthesis pathway

The known PHA biosynthetic genes coding for the three enzymes, β-ketothiolase, NADPH dependent acetoacetyl-CoA reductase and PHA synthase, and the genes coding for other proteins involved in PHA metabolism, are often clustered in operon structures within bacterial genomes. Based on genetic analysis, five types of clusters can be identified (Figure 1.3).

For type I *pha* genes clusters, the three PHA synthetic genes are in an operon. Analysis of *pha* gene arrangements in *R. eutropha* has shown that three structural genes required for the synthesis of PHA in *R. eutropha* are in an operon in the order of *phaC, phaA,* and *phaB* and are located on a 4 kbp segment of the chromosome (Slater *et al.*, 1988; Peoples and Sinskey, 1989b,c; and Schubert *et al.*, 1988). In *E. coli*, expression of *pha* genes is constitutive and nutrient limitation is not required for PHA synthesis. Choi *et al.* (1998) cloned *pha* genes from *Alcaligenes latus* and found these *pha* genes are also arranged in an operon with the same gene order as that for *R. eutropha*. These three genes share a single promoter located upstream of *phaC*. Another cryptic ORF (ORF4) with its own promoter is located upstream of *phaC*. These *pha* genes are expressed constitutively in *E. coli* from their natural promoter. When expressed in *E. coli,* the product of ORF4 seems to be involved in maintenance of the stability and PHA productivity of PHA production plasmids. *Pseudomonas* sp. strain 61-3 also possesses a *pha* operon with the order of *phaB*, *phaA* and *phaC* (Matsusaki *et al.*, 1998). An additional ORF (*phaR*) is located upstream of *phaB*, and is transcribed in the opposite direction to that of the *pha* operon. It is thought that *phaR* encodes a positive regulatory protein that controls transcription of *phaBAC*. The shared promoters of *pha* operons are similar to the *E. coli* σ^{70} consensus promoter sequence (Table 1.1).

For type II *pha* genes clusters, only *phaA* and *phaB* are arranged in an operon; *phaC* is transcribed separately. In *P. denitrificans* (Ueda *et al.*, 1996), *R. meliloti* (Tombolini *et al.*, 1995), and *Z. ramigera* (Peoples and Sinskey, 1989a), the genes for β-ketothiolase

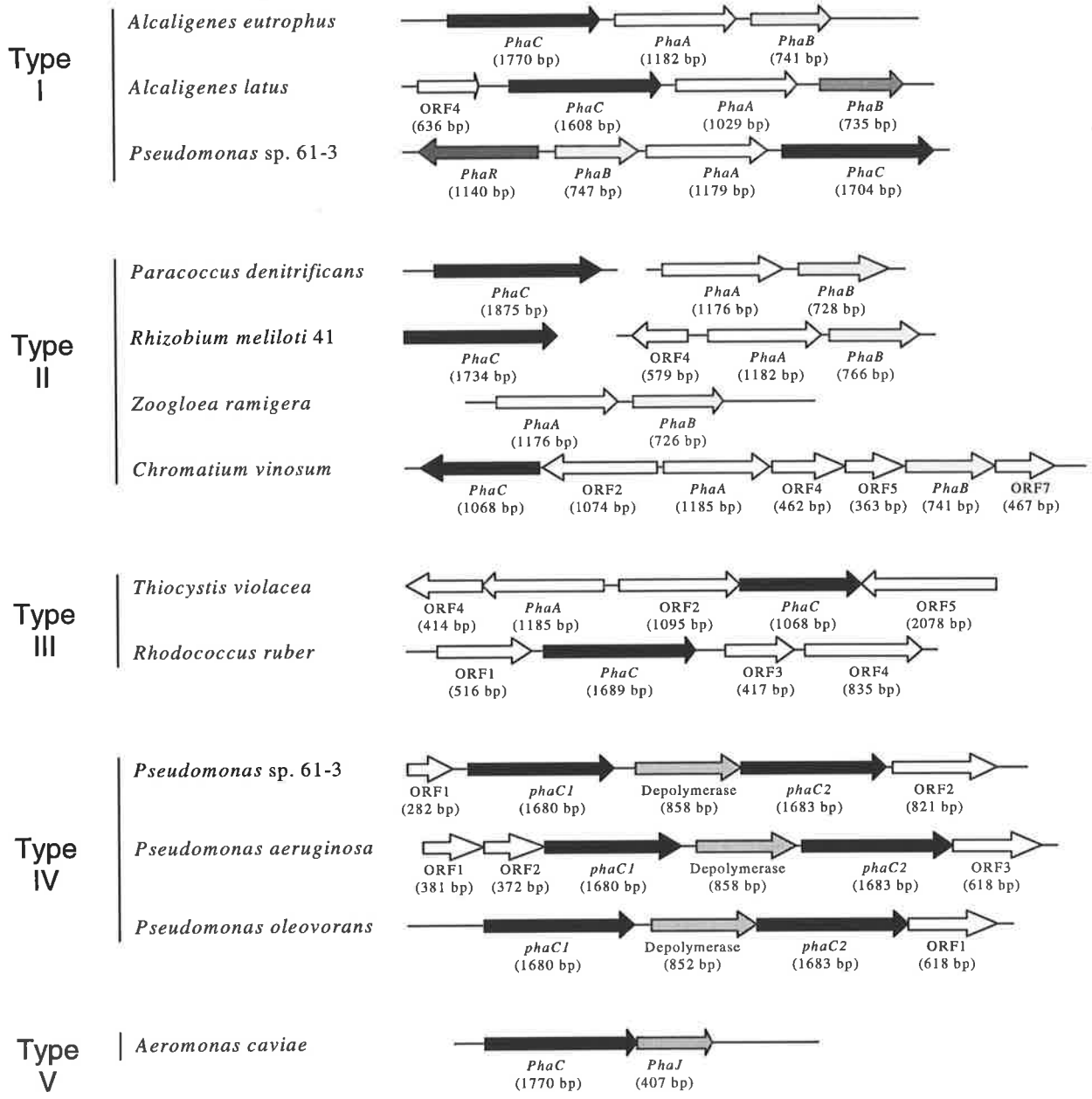


Figure 1.3. Schematic drawing showing the molecular organisation of relevant genes for PHA biosynthesis.

The organisation of *pha* genes from *R. caviae* (Fukui *et al.* 1998), *R. eutropha* H16 (Slater *et al.*, 1988; Peoples and Sinskey, 1989b,c; and Schubert *et al.*, 1988), *A. latus* (Choi *et al.*, 1998), *C. vinosum* D (Liebergesell and Steinbuchel, 1992), *P. denitrificans* (Ueda *et al.*, 1996), *P. aeruginosa* PAO1 (Timm and Steinbuchel, 1992), *P. Oleovorans* ATCC 29347 (Huisman *et al.*, 1991), *Pseudomonas* sp. strain 61-3 (Matsusaki *et al.*, 1998), *R. meliloti* 41 (Tombolini *et al.*, 1995), *R. ruber* (Pieper and Steinbuchel, 1992), *T. violacea* 2311 (Liebergesell and Steinbuchel, 1993), and *Z. ramigera* I-16M (Peoples and Sinskey, 1989a) are shown. The ORFs refer to open reading frames with unknown function.

Table 1.1. Comparison of the promoter regions of *pha* operons and *E. coli* σ^{70}.

Source	Promoter region	Ref
R. eutropha	CCGCATTGACAGCGCGTGCGTTGCAAGGCAACAATGG-ACTCAAATGTCTC----	1
A. latus	----GTAGAAT-GAAAAGAGTTGTCATGATGCGGTAAGACACGAAGCCTACAACG	2
P. sp. 61-3	--ATATCGACC-TGTGTAGAACATATTTCAACTATGTTGCACAAATATTAACT--	3
E. coli σ^{70}	* ** * * * * ** | || || TTGACA TATAAT -35 -10	

1. Schubert *et al.*, 1991. Genbank accession No. M64341
2. Genser *et al.*, 1998. Genbank accession No. U47026
3. Matsusaki *et al.*, 1998. Genbank accession No. AB014757

(*phaA*) and for NADPH-dependent acetoacetyl-CoA reductase (*phaB*) are adjacent to each other and constitute one single operon. However, in *Chromatium vinosum* strain D, genes coding for β-ketothiolase (*phaA*) and NADH-dependent acetoacetyl-CoA reductase (*phaB*) are probably in an operon, but separated by two ORFs (ORF4 and ORF5) of unknown function. The structural gene coding PHA synthase (*phaC*) in *C. vinosum* D is preceded by a second ORF (ORF2). Both probably constitute an operon and a σ^{70}-dependent promoter has been identified upstream of ORF2. The function of the product of ORF2 is unknown. Mutations within ORF2 result in strains which have decreased PHA synthase activity. Furthermore, the protein product is associated with PHA granules (Liebergesell and Steinbuchel, 1992).

Type III *pha* genes clusters are characterised by separation of *phaC* from other *pha* genes. In *R. ruber*, the gene for PHA synthase (*phaC*) is not adjacent to other genes coding for β-ketothiolase (*phaA*) and NADPH-dependent acetoacetyl-CoA reductase (*phaB*) (Pieper and Steinbuchel, 1992). Liebergesell and Steinbuchel (1993) cloned a 5361 bp *Eco*R I chromosomal DNA fragment from *Thiocystis violaceae* strain 2311 which encoded five ORFs: *phaA*, *phaC* and three other ORFs (ORF2, ORF4 and ORF5) of unknown function. *phaA* and *phaC* are in opposite orientation and separated by ORF2. It is not known that whether the *phaA* and *phaB* in *R. ruber* and *T. violaceae* are arranged in an operon.

Type IV *pha* genes clusters are characterised by the presence of two separated *phaC* genes. *Pseudomonas oleovorans* (Huisman *et al.*, 1991), *Pseudomonas aeuginosa* (Timm and Steinbuchel, 1992) and *Pseudomonas* sp. strain 61-3 (Matsusaki *et al.*, 1998) possess two genes coding for PHA synthases which are separated by a gene coding for PHA depolymerase. A cryptic ORF was identified downstream of *phaC2* in the genome of each bacterium.

Type V *pha* genes clusters are characterised by presence of novel *pha* genes in an operon. In *Aeromonas caviae*, *phaJ* is located downstream of *phaC*, but is transcriptionally coupled to the latter (Fukui *et al.* 1998). *phaJ* encodes an enoyl-CoA hydratase which participates in PHA biosynthesis from alkanoic acids. This enzyme exhibits (*R*)-specific hydration activity toward trans-2-enoyl-CoA with four to six carbon atoms, demonstrating that (*R*)-specific hydration of 2-enoyl-CoA catalysed by the product of *phaJ* is a

channelling pathway for supply of (*R*)-3-hydroxyacyl-CoA monomer units from fatty acid β-oxidation to poly(3HB/3HHx) biosynthesis in *A. caviae* (Figure 1.4). For the biosynthetic route from β-oxidation intermediates to (*R*)-3HA-CoA for PHA biosynthesis, three candidates have been proposed: (*R*)-specific hydration of 2-enoyl-CoA, (*R*)-specific reduction of 3-ketoacyl-CoA, and epimerisation of (*S*)-3HA-CoA (de Waard *et al.*, 1993; Eggink *et al.*, 1992; and Steinbuchel, 1996). This was the first study to prove that (*R*)-specific hydration of 2-enoyl-CoA a channelling pathway for supplying (*R*)-3-hydroxyacyl-CoA monomer units for PHA synthesis through the fatty acid β-oxidation pathway. The discovery of *phaJ* coding for (*R*)-hydratase activity within the *pha* locus was also novel.

1.5.3 Classification of PHA synthases

PHA synthases catalyse the last step for PHA synthesis and play an important role in both PHA biosynthesis and determination of PHA composition. Steinbuchel *et al.* (1992) divided PHA synthases into three types with respect to the primary structures deduced from their DNA sequences and to the substrate specificity of the PHA synthases.

Type I PHA synthases are PHA_{SCL} synthases. The primary structures exhibit 36.8-39.0% amino acids similarity, and include PHA synthases from *R. eutropha* (M_r 63,940), *R. ruber* (M_r 61,371), *R. sphaeroides* (M_r 66,740), and *M. extorquens* (M_r 66,742).

Type II PHA synthases are PHA_{MCL} synthases. The primary structures exhibit 53.7-79.6% amino acids similarity, and include PHA synthases from *P. oleovorans* (M_r 62,400 and M_r 62,600), and *P. aeruginosa* (M_r 62,363 and M_r 62,630). Type II PHA synthases differ from type I synthases by the lack of 26 amino acids N-terminal truncation, and by their substrate specificity. Type I and type II PHA synthases are 34.4-39.9% similar.

Type III PHA synthases are PHA_{SCL} synthases and exhibit 87.3% amino acids similarity. They include PHA synthases from *C. vinosum* (M_r 39,730), and *T. violacea* (M_r 39,550). These PHA synthases are smaller than type I and type II PHA synthases and contain a 175 amino acid N-terminal truncation and a 55 amino acid C-terminal truncation. Type III synthases also exhibit low amino acid sequence similarity (21.0-27.7%) when compared with type I and type II synthases. In *C. vinosum* and *T. violacea*, the ORF2 for a

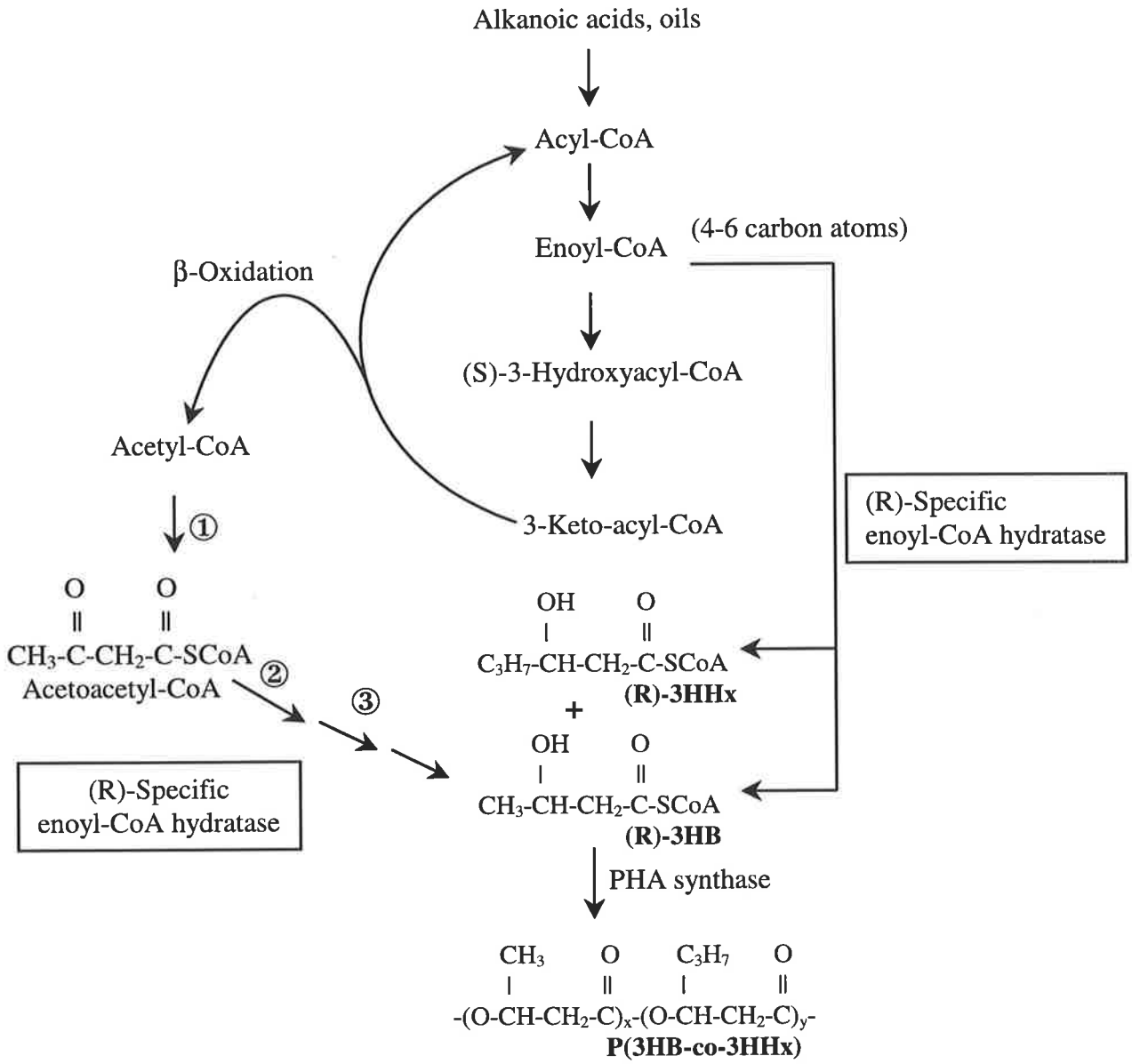


Figure 1.4. Proposed pathway of P(3HB/3HHx) biosynthesis by *A. caviae* from alkanoic acids or oils (Adapted from Fukui *et al.*, 1998).

1. β-ketothiolase
2. NADH-acetoacetyl-CoA dehydrogenase,
3. crotonase [(*S*)-specific enoyl-CoA hydratase]

product of unknown function is located upstream of *phaC*, and both ORF2 and *phaC* most probably constitute a single operon. It has been found that the expression of PHA synthase activities requires both the *phaC* and the ORF2 products. Steinbuchel *et al.* (1992) suggested that ORF2 product may fulfil a particular function for the polymerisation step which the larger type I PHA synthases fulfil themselves but the smaller type III PHA synthases can not.

1.6 Physical properties and application of PHA

1.6.1 Physical properties of P(3HB) and strategies for property improvement

As a polymer, PHA compares well with polypropylene in terms of molecular mass, melting point, crystallinity and tensile strength (Evans and Sikdar, 1990) (Table 1.2). However, compared to the common petroleum based commodity plastic products, P(3HB) barely fulfils any requirements for toughness, impact-resistance, flexibility, resilience, and fat-, water-, and heat-resistance.

Brittleness

P(3HB) can be prepared as an extremely pure product with no residual catalysts or other contaminants as the case for other plastics. During moulding, extrusion, or blowing, P(3HB) behaves like an ultra-pure fluid: it stays fluid (ie. in a glassy state), when temperatures drop below the melting point (De Koning and Lemstra, 1992). In the absence of contaminants, few crystallisation nuclei develop and the melt (glass) hardens in large crystallites (Barham, 1984). The space between the crystallites is almost void of material and cracks easily. This results in a very brittle solid with poor qualities and inherent difficulties for processing.

One way to avoid brittleness is to alter the properties of the polymer chain by incorporation of other HA units such as HV so that crystallisation is disturbed. The irregularity of the chains 'loosens' the tightness of the crystals. While this in itself does not solve the problem of the formation of large crystallites, it helps by producing a softer material.

Table 1.2. Typical properties of PHA (Luzier, 1992).

	HV content (mol %)		
Property	0	10	20
Melting point (°C)	177	140	130
Crystallinity (%)	80	60	35
Tensile strength (Mpa)	40	25	20
Flexural modulus (Gpa)	3.5	1.2	0.8
Extension at break (%)	8	20	50
Notched Izod impact strength (J/m)	60	110	350

A second way to avoid brittleness is to hasten uniform crystal formation. This can be achieved by adding nucleating agents to the melt. Standard nucleation agents like talc also work with P(3HB) and do not interfere with biodegradability. A most intriguing nucleating agent is saccharin. Saccharin crystals apparently have lattices that perfectly fit the repeat units of P(3HB) and thus serve as a neat nucleating surface (Black *et al.*, 1990). Blends with saccharin or other nucleating agents crystallise much faster than materials without these agents and yield smaller and many more crystals. This in turn drastically decreases the number of 'empty' spaces. Goods produced from such blend have superior strength and flexibility (Hanggi, 1995).

A third way to avoid brittleness is to add 'softening agents' or plasticisers. Plasticisers act like solvents and if added in small quantities to P(3HB), prevent the polymer from forming crystallites. A number of plasticising agents used in plastics manufacture are available at present, eg. esters of citric acid. When blended with P(3HB), they reduce brittleness and enhance flexibility. By using different combinations of PHA and plasticisers, it is possible to produce goods with properties equal to those made of polystyrene or polypropylene (Hanggi, 1995).

Impact strength

The major drawback of P(3HB) is its low impact strength. This can also be overcome by inducing the bacteria to produce a copolymer, such as P(3HB/3HV). This copolymer has a reduced melting temperature, reduced crystallinity and increased flexibility and toughness (Table 1.2). In addition, P(3HB/3HV) is degraded to carbon dioxide and water at a faster rate than the homopolymer [P(3HB)] (Daniell and Guda, 1997) by a wide variety of soil microorganisms, including Gram-negative bacteria, Gram-positive Bacilli, streptomyces and moulds (Mergaert *et al.*, 1992).

Other requirements like resistance to water, fat and heat

For wide spread application of PHA, PHA based plastics must be resistant to water, fat and heat. PHA is inert toward water and products made of it are 100% waterproof. Homopolymers and some of the heteropolymers with low hydroxyvalerate content can also be heated to 132 °C in the autoclave. Resistance toward fats and oils is limited since PHA

is lipophilic. However, resistance appears to be adequate for articles with a storage time of a few days to a few weeks (Hanggi, 1995).

1.6.2 Application of PHA

From the perspective of biodegradability, biocompatibility and manufacture from renewable resources, the primary application areas of PHA based plastics include: (1) disposable personal hygiene: PHA could be used as the sole structural material or as part of a degradable composite; (2) packaging: films, blow-moulded bottles, as a coating on paper; (3) Medical: biocompatibility coupled with its slow hydrolytic degradation lead to potential in reconstructive surgery and controlled release fields based on the property that PHA is immunologically inert and it is slowly degraded in animal or human tissue. This property can be used to produce devices for release of drugs or other pharmaceutical compounds in the body (Luzier, 1992; Holmes, 1985; Koosha *et al.*, 1989; Knowles *et al.*, 1992; and Kassab *et al.*, 1997). PHA can also be used as a substrate and matrix for a *Pseudomonas* strain capable of denitrifying drinking water (Biedermann *et al.*,1997).

1.7 Industrial scale production by bacteria

Production of PHA is a sophisticated process and has been carried out using natural PHA producing bacteria and recombinant *E. coli*.

1.7.1 Strain selection criteria for large scale PHA production

As of 1994, about 300 PHA producing bacteria had been isolated. However not all of them are useful for PHA production on a large scale. The following criteria are important factors determining selection of PHA producing microorganisms (Chang, 1994): (1) a high growth rate; (2) the ability to utilise inexpensive substrates; (3) a high polymer synthesis rate; (4) the maximum extent of polymer accumulation and yield; and (5) the ease of polymer recovery.

For example, *Azotobacter* was first employed for P(3HB) production by ICI. However, this strain diverted carbon to polysaccharide synthesis which made polymer recovery difficult. *Methylobacterium* can also utilise the cheap substrate methanol to produce PHA, but the PHA productivity, yield and molecular weight are low and there are

also difficulties for PHA extraction. For these reasons, *Azotobacter* and *Methylobacterium* were finally rejected by ICI as candidate organisms for large scale production (Byrom, 1992).

Some good strains have been selected for PHA production. *R. eutropha* H16 expresses high polymer content with a good molecular mass and the PHA produced can be relatively easily extracted. However, this strain does not utilise glucose and mutants able to utilise glucose have since been selected for industrial production of P(3HB) (Schlegel, 1990). *A. latus* is also a very good PHA producer which was isolated and employed by an Austrian company Chemie Linz AG in 1980s. This bacterium has some advantages over *R. eutropha*, eg. More rapid growth, ability to use cheap sucrose (beat and cane molasses) as a carbon source and can accumulate P(3HB) during normal growth. A single fermentation process only is required (Lee and Chang, 1995; and Hrabak, 1992).

Some *Pseudomonas* spp. have advantages compared to *R. eutropha* and *A. latus*. One of them is that *R. eutropha* and *A. latus* mainly produce PHA_{SCL}, while most of the investigated *Pseudomonas* strains mainly produce PHA_{MCL}. Among those *Pseudomonas* strains, *P. oleovorans* is widely employed. Other bacteria employed for PHA production on a large scale are listed in Table 1.3.

1.7.2 PHA production from naturally occurring bacteria

PHA synthesis and accumulation by naturally occurring bacteria usually happens under nutrient limiting conditions, necessitating a two-stage fermentation process for large scale production. During the first stage, there is no nutrient limitation, bacteria multiply and grow, but no significant accumulation of PHA occurs. When cell numbers reach a high density, nutrient-limiting growth conditions are used to trigger PHA synthesis.

Doi *et al.* (1990) applied this strategy to produce three different PHA copolymers from a strain of *R. eutropha*. A random copolymer P(3HB/3HP) was produced from a nitrogen-free medium containing 3-hydroxypropionic acid, 1,5-pentanediol, 1,7-heptanediol, or 1,9-nonanediol; P(3HB/4HB) was produced from carbon sources such as 4-hydroxybutyric acid, γ-butyrolactone, 1,4-butanediol, 1,6-hexanediol, 1,8-octanediol, 1,10-decanediol, and 1,12-dodecanediol; and P(3HB/3HV) from propionic acid or pentanoic acid. During the first of the two-stage cultivation, cells were grown under

aeration at 30 °C in a nutrient-rich medium. After the growth phase, PHA expression was induced by transfer to a nitrogen-free mineral medium containing different carbon sources. Temperature and pH were automatically controlled at 30 °C and 7.5 respectively. ^{1}H and ^{13}C NMR spectra analysis showed that in P(3HB/3HP), the contents of 3HB and 3HP were 93-96% and 4-7% respectively, in P(3HB/4HB), the contents of 3HB and 4HB were 66-91% and 9-36% respectively, in P(3HB/3HV), the contents of 3HB and 3HV were 25-69% and 31-75% respectively.

Kim *et al.* (1994a,b) used a fed batch culture technique to produce P(3HB) and P(3HB/3HV) by *R. eutropha.* Ryu *et al.* (1997) used high cell density fed-batch fermentation of *R. eutropha* for the production of P(3HB) in a 60 L fermenter with phosphate limitation instead of nitrogen limitation. This strategy resulted in a final cell concentration of 281 g/L, a P(3HB) concentration of 232 g/L, and a P(3HB) productivity of 3.14 g/L/h.

Asenjo *et al.* (1995) studied the effect of magnesium and phosphate limitation on the molecular weight distribution of P(3HB) in *R. eutropha* in continuous culture. Under N-limitation and glucose excess, PHA M_w decreased when the magnesium content was decreased below 50% (19.7 mg/L) of the basal medium content. This resulted in a broadening of the polymer molecular weight distribution (M_w/M_n) from 2 to 5 and a decrease in M_w from 2×10^6 to 0.9×10^6. Below 20% of the basal content of magnesium (7.9 mg/L) these two trends were reversed. Under N-limitation and glucose excess, phosphate had virtually no effect on P(3HB) M_w or its distribution. However, with no (or little) glucose excess, the M_w of the P(3HB) produced decreased with phosphate concentrations below 50% of the basal level (0.705 g/L). Their investigation suggested that in continuous or fed-batch cultures, where nitrogen limitation is used to trigger P(3HB) accumulation, it is necessary to control both the addition of glucose (no excess) and also to maintain magnesium limitation [ca. 25% (9.9 mg/L) of basal medium level] and phosphate above 50% of the basal level (0.705 g/L).

Kim *et al.* (1997) studied production of PHA_{MCL} by a strain of *Pseudomonas putida* in a two-stage fed-batch cultivation using a combination of glucose and octanoate. Octanoate supports efficient PHA_{MCL} production, but PHA yield is low because this substrate is also used for bacterial growth. To overcome this problem, a two-step fed-batch

cultivation with glucose and octanoate as the main carbon sources for bacterial growth and PHA accumulation was used. During the first step, glucose was utilised for bacterial growth. Octanoate was then supplied as the sole carbon source and converted to PHA under the nitrogen- and oxygen-limiting conditions. This approach resulted in a PHA_{MCL} yield of 18.6 g/L with a carbon conversion rate of about 40% (g PHA/g octanoate).

Thus the PHA content in cells may be raised to unusually high levels under certain growth conditions, eg. to 74 % of the cell dry weight in *R. eutropha* under conditions of nitrogen limitation. However since the synthesis and the degradation of PHA are often coupled, there are difficulties in isolation of an undegraded PHA polymer. Even though some natural bacteria have been successfully employed for large-scale production of PHA, they have several major drawbacks such as slow growth rate, narrow spectrum of useable substrates, intracellular PHA degradation as mentioned above and usually a requirement for complex fermentation (Lee *et al.*, 1994c).

1.7.3 PHA production with recombinant *E. coli*

Use of *E. coli* as a production host for PHA production has several advantages over natural PHA producing bacteria. Its genetics, molecular biology, biochemistry, physiology, and the range of utilisable carbon sources have been well documented. *E. coli* cells are fragile and easily disrupted to facilitate recovery of PHA granules. Furthermore, the lack of intracellular depolymerases in *E. coli* means that degradation of the PHA is unlikely to occur (Fidler and Dennis, 1992; and Lee, 1997).

An ideal PHA production system involves two aspects, an optimum *E. coli* strain as the PHA production background, and the genetic expression system designed accordingly. Genetic manipulation of *E. coli* strains and development of mechanisms which allow control of expression of *pha* genes have allowed significant technological advances.

1.7.3.1 Development of recombinant *E. coli* strains

E. coli strains which allow cheap production of PHA must meet the following criteria: the true cell mass of the culture must be high, and most of its cells must accumulate PHA. If both of these criteria are met, then a large amount of PHA will be produced. However, metabolic fluxes differ widely among *E. coli* strains. Introduction of

foreign PHA biosynthetic enzymes into different *E. coli* strains results in establishment of a new metabolic pathway that competes with existing pathways for substrates. Hence differences in PHA production are noted for different strains (Lee *et al.*, 1994b).

Production of P(3HB) is readily achieved in *E. coli*, although the production of copolymer or heteropolymer has been less than successful. When *E. coli* harbouring the *R. eutropha* PHA biosynthetic genes is grown under conditions which would stimulate copolymer production in *R. eutropha*, no hydroxyvalerate can be detected in the produced polymer (Slater *et al.*, 1992). Slater *et al.* (1992) hypothesised, and later proved that the lack of copolymer production in *E. coli* is due to the fact that *E. coli* does not possess an efficient system for conversion of propionate to propionyl-CoA. When *E. coli* is grown in a rich medium, the pathways of fatty acid uptake and fatty acid utilisation are not expressed because of the regulation by *atoC* gene (activator) and *fadR* gene (repressor) products respectively. Thus substrates such as propionate can not be efficiently converted to propionyl-CoA.

The gene *atoC*, encodes an activator of fatty acid uptake. A mutation in this gene [*atoC*(Con)] results in constitutive levels of the ATO (acetoacetate) enzymes (Jenkins and Nunn, 1987) which are involved in short-chain fatty acid degradation in *E. coli*. The *fadR* gene encodes a multifunctional regulator of fatty acid and acetate metabolism (DiRusso, 1988) which acts as a negative regulator. For production of PHA by *E. coli* from substrates such as acetate or fatty acids, the enzymes for fatty acid degradation (*fad* operon) and acetate metabolism (*ace* operon) must be derepressed (DiRusso, 1988, Maloy and Nunn, 1982). Under these conditions cellular macromolecules must be derived from acetyl-CoA which requires the operation of the glyoxylate shunt. The glyoxylate shunt allows the net assimilation of carbon from acetyl-CoA as it bypasses the two carbon dioxide evolving steps of the TCA (DiRusso, 1988, Maloy and Nunn, 1982). The two unique enzymes of the glyoxylate shunt, isocitrate lyase and malate synthase A, are encoded by *aceA* and *aceB* respectively (Maloy and Nunn, 1982). Derepression of these genes takes time when the bacteria are moved from a rich medium to a medium containing acetate or fatty acids as the sole carbon source. However, mutations in *fadR* result in constitutive expression of the glyoxylate shunt enzymes under non-inducing conditions.

Slater *et al.* (1992) developed a *fadR atoC* mutant *E. coli* strain, MD9101 from *E. coli* strain, LS5218 by transduction with a transducing lysate of *E. coli* RM1981 (*recA56 srl*::Tn*10*) and used this strain for the production of copolymers. Development of this mutant overcame the problem of inefficient conversion of propionate to propionyl-CoA. For production of copolymer P(3HB/3HV) by *E. coli* MD9101 (p4A), incorporation of 3HV is absolutely dependent on the presence of both glucose and propionate and thus production is essentially a two-step process. When *E. coli* was cultivated in M9 minimal medium, with acetate as the sole carbon source, production of copolymer was subsequently induced by the addition of glucose and propionic acid. Under these culture conditions, 3HB-3HV ratios in the copolymer were successfully manipulated by altering the propionate concentration and/or the glucose concentration in the culture.

As shown in Figure 1.5, a required step in production of P(3HB/3HV) is the condensation of acetyl-CoA and propionyl-CoA to form β-acetopropionyl-CoA. This activity has generally been attributed to the β-ketothiolase (PhaA) encoded by a *phaA* gene in a *pha* operon. Slater *et al.* (1998) cloned a *bktB* gene from *R. eutropha*, and found that the β-ketothiolase encoded by this gene is also capable of forming β-acetopropionyl-CoA.

Different *E. coli* strains have been evaluated for their growth rate, utilisation of glucose and ability to accumulate P(3HB), when harbouring the *R. eutropha* PHA biosynthetic genes. Strains tested include wild-type strains, K-12, B and W; derivatives of K-12, which are often used in fermentations to produce recombinant proteins, DH5α, JM109 and XL1-Blue; and a hybrid of K-12 and B, HB101. *E. coli* strains XL1-Blue and B have been found to be the best P(3HB) producing strains on the basis that they supported expression of high yields of P(3HB) as well as a high true cell mass (Lee *et al.*, 1994a,b).

1.7.3.2 Development of PHA expression systems

The need to control PHA expression in *E. coli* during large scale production, has lead to construction of expression systems which enable the combination of high levels of cell growth, expression of PHA biosynthetic genes, and high PHA yields. Based on the mechanism of regulation of expression, three types of expression systems have been described.

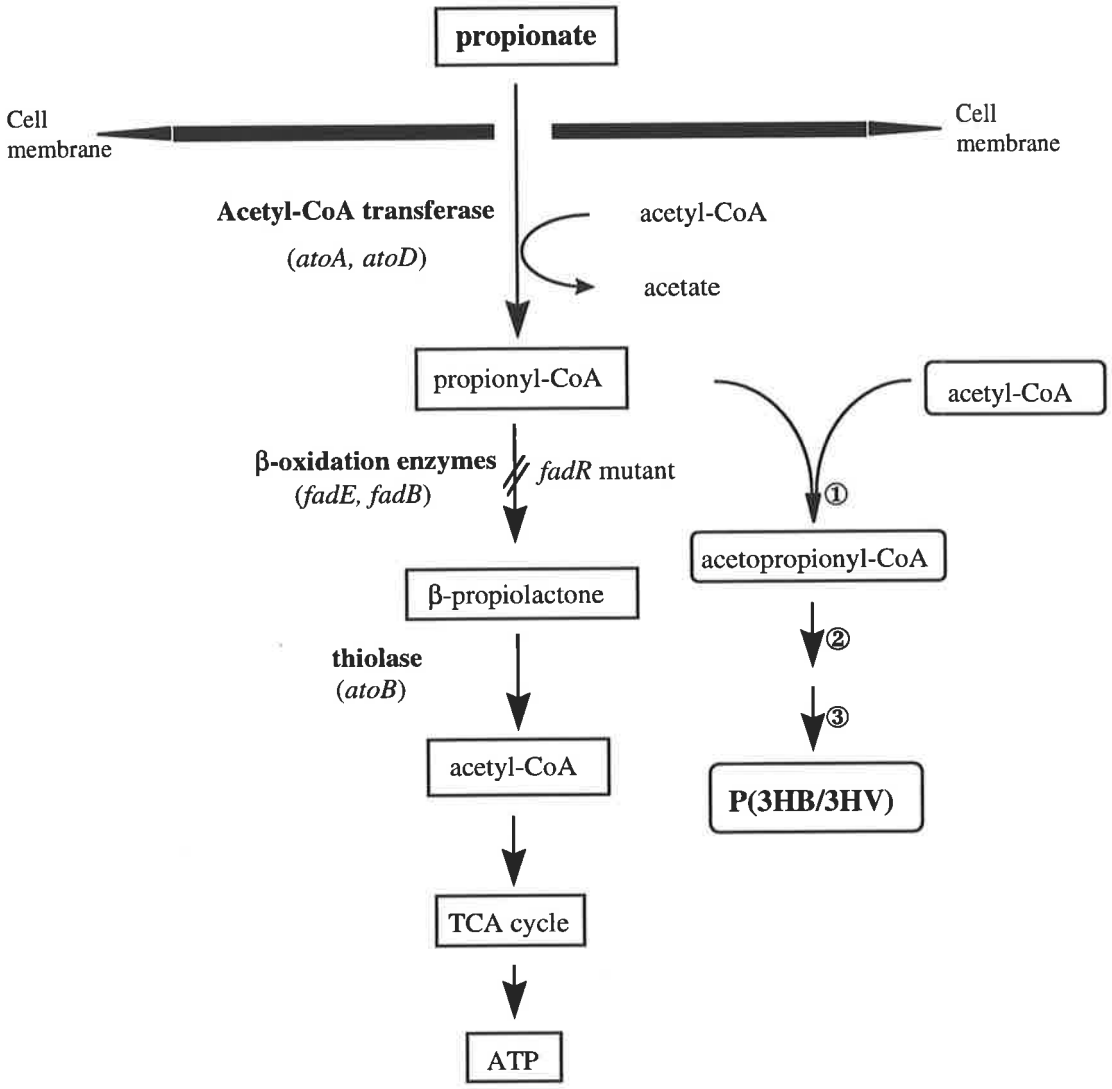


Figure 1.5. Production of P(3HB/3HV) by an *E. coli fadR atoC* mutant, MD9101 harbouring PHA synthetic genes (Adapted from Jenkins and Nunn, 1987).

The *atoB, atoA,* and *atoD* genes are constitutive because of the *atoC* (Con) mutation, and *fadE* and *fadB* are not expressed because of the *fadR* mutation. Therefore propionyl-CoA is not degraded but is shunted into the PHA pathway to form copolymers. 1. β-ketothiolase, 2. NADPH-dependent acetoacetyl-CoA reductase, and 3. PHA synthase.

"First generation" plasmid expression systems are represented by plasmids P4A (pTZ18U-PHB) and pSYL101. Constructs which simply harbour *R. eutropha* PHA biosynthesis genes allow PHA production to less than 50% of dry cell weight (Fidler and Dennis, 1992). To eliminate gene dosage effects on PHA accumulation, a plasmid with a high copy number, designated p4A, was created, which enabled P(3HB) production levels to 70% of total cell weight (Janes *et al.*, 1990). Lee *et al.* (1994a,c) subsequently demonstrated that the level of PHA expression correlated with high expression vector copy number. These first generation plasmids relied on: (1) constitutive expression of *pha* genes from a native promoter, resulting in uncontrollable PHA production; and (2) use of antibiotic resistant genes to ensure stable maintenance of the plasmid. Plasmid instability is the most common problem encountered in cultures involving the recombinant organisms. However, growth in the presence of antibiotics is often not practical on an industrial scale, and alternative selection methods must be used to ensure plasmid stability.

"Second generation" plasmid expression systems differ in that they incorporate mechanisms designed to improve plasmid stability. Stability can be obtained by exploitation of the *parB* locus of plasmid R1, which mediates stabilisation via postsegregational killing of plasmid-free cells (Figure 1.6). This strategy was employed in construction of a number of expression plasmids including pJM9123, pSYL104, pSYL105 and pSYL107 (Slater *et al.*, 1992; and Lee *et al.* 1994a,b,c).

Plasmid pJM9123 was created by introducing a copy of the *parB* plasmid stabilisation locus into p4A plasmid (Slater *et al.*, 1992). A similar approach was employed by Lee *et al.* (1994a,b,c). They introduced the *parB* locus into a proven high copy number expression plasmid (pSYL101). This strategy was also used to manipulate the high copy number vector, pSK2665 containing the *R. eutropha pha* operon, to create stable PHA production plasmids, pSYL104 and pSYL105 respectively. *E. coli* MD9101 (harbouring pJM9123) and *E. coli* XL1-blue (harbouring pSYL104, and pSY105 respectively) show high level PHA production.

One of the problems associated with over-expression of PHA in recombinant *E. coli* is filamentation. *E. coli* XL1-Blue cells expressing the *R. eutropha pha* genes are

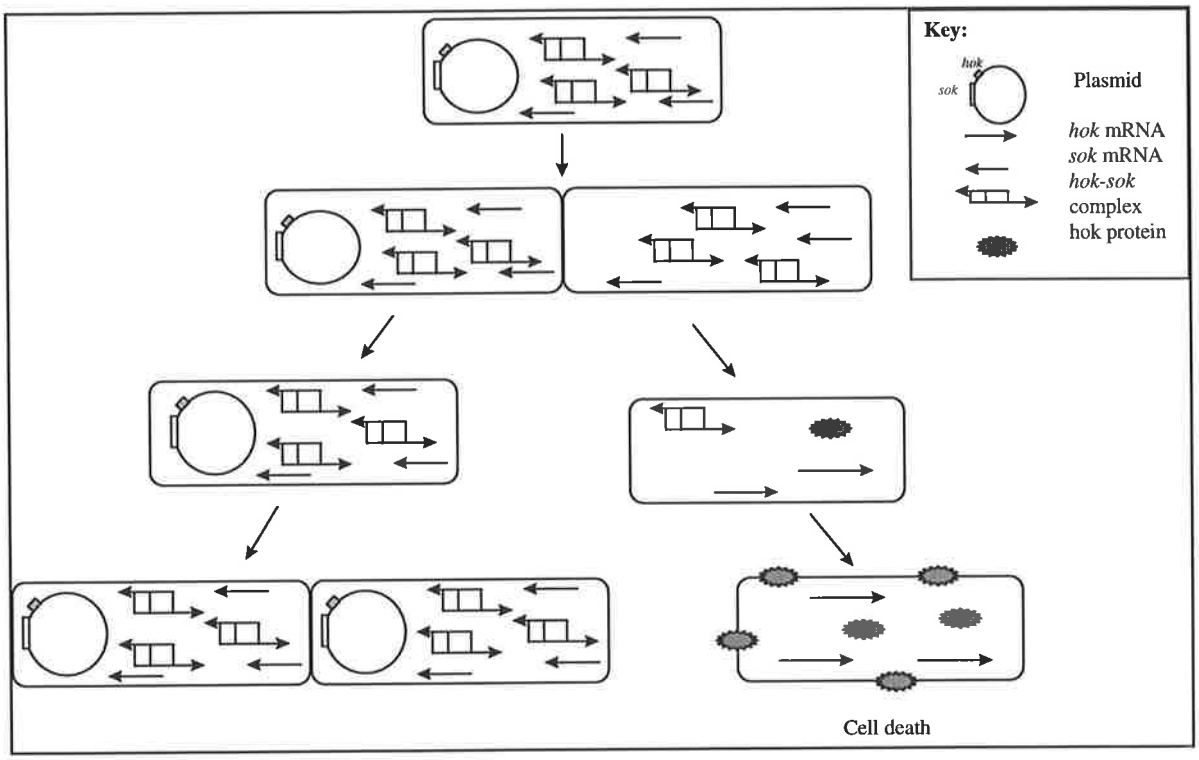


Figure 1.6. Molecular model describing the postsegregational killing mediated by the *parB* (*hok-sok*) locus (Adapted from Gerdes, 1988).

The *parB* locus consists of two genes, *hok* (host killing) and *sok* (suppression of killing). This locus ensures plasmid stabilisation by postsegregational killing of plasmid free cells. The expression of the Hok protein results in the rapid collapse potential of the cell membrane, arrest of respiration and cell death. The *sok* gene product is a small antisense RNA which inhibits the translation of the *hok* mRNA. When a cell contains the plasmid with *parB*, it survives since the *sok*-RNA prevents the translation of the *hok* mRNA. However, the *sok*-RNA is rapidly degraded whilst the *hok* mRNA is extraordinarily stable. Therefore if a cell loses the plasmid, *sok*-RNA is no longer synthesised and the Hok protein will be produced, resulting in cell death. Hence plasmids containing the *parB* locus are very stable.

very long (some greater than 100μm in length) (Lee *et al.*, 1994b). This phenomenon results from a blockage of cell division, which causes a reduction in growth rate and hence reduces the productivity of recombinant protein(s) as well as cellular levels of PHA. The molecular basis for this phenomenon involves the heat-shock response to foreign proteins and inhibition of the activity of the FtsZ protein. FtsZ is an important protein implicated in circumferential "pinching" of the cell division septum (Lutkenhaus, 1993). FtsZ forms a ring at the division site of the cell. By incorporating copies of *ftsZ* in PHA expression systems, filamentation can be suppressed and twice as much PHA can accumulate (Lee, 1994). Choi *et al.* (1997) employed *E. coli* XL1-Blue harbouring pSYL107 (including *ftsZ*) for P(3HB) production by fed-batch culture and achieved a P(3HB) concentration of 104g/L, and a P(3HB) content of 70% with a productivity of 2 g P(3HB)/L/h.

Constitutive expression of *pha* genes places a large metabolic burden on the cell and retards cell growth, resulting low PHA production. Thus efficient expression of PHA in *E. coli* requires not only a high gene dosage, but also a controlled expression of the *pha* genes via a heterologous promoter. Third generation expression systems incorporate mechanisms designed to control expression of PHA. Kidwell *et al.* (1995) constructed a high copy number and temperature inducible PHA production plasmid, pJM9238 (based on pRA90), in which the native promoter in *R. eutropha pha* operon was replaced by a *tac* promoter, forming a *tac::pha* operon fusion. *parB* was also introduced for plasmid stability. The copy number of this plasmid is temperature dependent. At non-inducing temperatures (eg. 30 °C), the plasmid copy number is depressed to between 1 and 10 and the copy number of LacI protein is sufficient to repress *tac::pha* transcription. However, at inducing temperatures (eg. 41 °C), the copy number of the plasmid may increase to as high as 1,000 and insufficient LacI repressor protein is available to prevent rapid induction of transcription from the *tac* promoter and subsequent synthesis of PHA-specific proteins, and polymer synthesis. Kidwell *et al.* (1995) was able to show that temperature shift from 30 °C to 41 °C during fed-batch fermentation allowed high level production of PHA by *E. coli* HMS174 (pJM9238) (1.07 g P(3HB)/L/h). This work indicated the importance of tight transcriptional control and high gene dosage on PHA production in *E. coli*. More recently, Sim *et al.* (1997) found that an increased induction of PHA synthase (PhaC) alone can result in a decrease in PHA molecular weight in *E. coli*, suggesting PHA synthase plays a key role in control of molecular weight and the polydispersity of PHA. Thus in

construction of expression systems, the control of expression of *phaC* should be carefully considered.

1.7.4 Media for bacterial PHA production

Production cost estimates of $5-15 per kilogram mean that PHA cannot currently compete with petroleum based synthetic plastics (about $1 a kilogram) (Daniell and Guda, 1997). The main reasons for the high cost is the cost of carbon sources and co-substrates, maintenance costs associated with a large fermentation facility and associated purification costs (Rhee *et al.*, 1992). Consequently, selection of appropriate growth media for the recombinant *E. coli* strains is as important as the selection of host strains in which to express the recombinant protein(s) and the improvement of the expression systems. The chosen media must support good bacterial growth, and must also be cost effective.

Various media have been examined for their effects on the production of PHA in recombinant *E. coli*. These include M9, Harrisons and Riesenberg's defined media (R-medium). Lee *et al.* (1994c) found that of the defined media, Riesenberg's was the medium of choice, since it supported the best *E. coli* growth and highest PHA accumulation. However, PHA accumulation is highest when the host microorganisms are grown on a complex medium. A possible reason for this is that when in a complex medium, intermediary precursors such as vitamins, amino acids and cofactors, which are required for macromolecule biosynthesis are readily available. Thus bacterial cells do not have to expend extra energy on synthesising these nutrients, particularly when under pressure to overproduce a recombinant protein. This assists the growth of the bacteria and more acetyl-CoA is available for the synthesis of PHA. In a defined medium, acetyl-CoA must be used as a biosynthetic precursor in other pathways including citrate and acetate formation, and in membrane synthesis. There is therefore, less acetyl-CoA available for the production of PHA. Considering the cost of complex media, they are not economically available on a large scale. Thus it is necessary to develop cheap defined medium for bacterial growth and PHA production.

Lee and Chang (1994) supplemented R-medium with yeast extract and achieved greater yields of PHA in *E. coli*. This suggested that a complex nitrogen source may be required for efficient synthesis of PHA. Based on this hypothesis, Lee and Chang (1994)

tested ten complex nitrogen sources with a defined medium. Of these, tryptone (2g/L) supported up to a fivefold increase in PHA concentration when added to R-medium. Similarly, when cysteine, isoleucine, methionine and proline were added to flask cultures of recombinant *E. coli*, PHA synthesis was increased significantly (Lee *et al.*, 1995). Clearly added amino acids decrease the need for amino acids synthesis and this probably results in more NADPH available for the activity of the NADPH-dependent acetoacetyl-CoA reductase. Using a similar reasoning, Lee *et al.* (1995) found that the addition of oleic acid enhanced PHA synthesis. Apparently, addition of oleic acid into *E. coli* culture provides precursors for fatty acid synthesis and releases more acetyl-CoA available for PHA synthesis.

However, a disadvantage to use of additives in defined media is increased cost of production. PHA production cost can best be reduced by using unpurified, low cost organic wastes from agriculture and food processing plants. These wastes, such as cheese whey, molasses, tallow, and pig waste liquor represent potential economical nutrient sources for PHA production on an industrial scale. For example, van Wegen *et al.* (1997) employed *E. coli* Topp 1 (pJM9123) for P(3HB) production from whey and production cost of as low as US1995$2.67 kg^{-1} was estimated (van Wegen *et al.*, 1998). Li *et al.* (1997) carried out PHA production by fed-batch culture with *R. eutropha* using beet and sugarcane molasses as the sole carbon source respectively and achieved PHA yield and content of 15.7 g/L, 56% (of dry cell weight) and 10.7 g/L, 50%, respectively after 70 h fermentation. Zhang *et al.* (1994) employed *K. aerogenes* for the production of P(3HB) from sugarcane molasses as the sole carbon source and achieved P(3HB) production at the rate of approximately 1 g P(3HB)/L/h. Cromwick *et al.* (1996) produced PHA from unhydrolysed tallow by employing *P. resinovorans* with a PHA content of 15% of its dry cell weight. Cho *et al.* (1997) employed *Azotobacter vinelandii* UWD, and produced 2.0 g/L of dry cells which contained 34.0% (w/w) of P(3HB/3HV) with 7.9% (mol/mol) of 3HV from two-fold diluted pig waste liquor, which was supplemented with other components. Supplementation of 30 g glucose/L increased cell dry weight to 9.4 g/L with 58.3% (w/w) of P(3HB/3HV) and 4.3% (mol/mol) of 3HV at a production rate of 0.11 g P(3HB/3HV)/L/h.

Although waste products may provide a cost effective source of nutrients for PHA synthesis, it is clear from the discussion in section 1.7.2 that nutrient composition can

adversely affect yield and molecular weight distribution of the resulting polymer. Clearly optimisation of the producer bacterium and nutrient source for maximum PHA expression must be considered. PHA quality should not be compromised by cost of production.

1.7.5 Industrial scale PHA production methods and technologies

PHA production by different bacteria, including recombinant *E. coli* with different methods and carbon sources have been carried out in recent years. Lee and Chang (1995) have summarised those methods used for PHA production as shown in Table 1.3. At that time, P(3HB/3HV) was produced (ca. 600 tonnes per year) by ZENECA Bio Products (UK), and was sold at $16/kg under the tradename BIOPOL. There were also several other companies involved in PHA research and development. Berlin Packaging Corp. (USA) marketed and distributed BIOPOL to a number of companies worldwide. Bio Ventures Alberta Inc. (Canada) carried out simulation studies for the production of PHA by recombinant *E. coli*. Metabolix Inc. (USA) looked for joint ventures and licensing technology. Metabolix, Monsanto (USA), and ZENECA Seeds (UK) investigated strategies for PHA production by transgenic plants. Polyferm, Inc. (Canada) was developing a system for PHA production from cheap hemicellulose by employing *Pseudomonas cepacia* (Lee, 1995).

1.7.6 A systematic strategy for PHA production system construction on an industrial scale

PHA producing bacterial cells can be considered to consist of two key components: the cytoplasm, which supplies the substrates and an environment like a bioreactor for PHA synthesis and accumulation, and the *pha* genes, which encode the enzymes for PHA synthesis. For the naturally occurring PHA producing bacteria, both of the above components are unified. Production of PHAs with different composition can simply be achieved by optimisation of the media or by employing different types of bacteria. However, there are limits to benefits obtained by strain selection and media optimisation. Genetic engineering has partially solved this problem by employing recombinant *E. coli* strains, in which the "bioreactors", and *pha* genes from different sources can be flexibly combined. With this strategy, several PHA production plasmids have been constructed and used for PHA production by *E. coli* (p4A, pSYL104 and others), and some *E. coli* strains have been selected and improved. Nevertheless, the process is still 'hit and miss'. At

Table 1.3. Summary of PHA production by various microorganisms and culture methods (Adapted from Lee and Chang, 1995) (to be continued).

Bacteria	PHA	Culture method	Major substrate	Culture time (h)	Cell Conc. (g/L)	PHA conc. (g/L)	PHA conc. (%)	Productivity (g/L/h)	Reference
Ralstonia eutropha	P(3HB)	Glucose control fed batch	Glucose	49	124	92	74	1.87	Kim *et al.* (1994a)
Ralstonia eutropha	P(3HB)	Glucose control fed batch	Glucose	50	164	121	76	2.42	Kim *et al.* (1994a)
Ralstonia eutropha	P(3HB)	Recycled gas culture system	CO_2/H_2	40	85	61.5	72	1.54	Lee and Chang, (1995)
Ralstonia eutropha	P(3HB)	Two-stage continuous	CO_2/H_2	-	-	20	-	0.9	Lee and Chang, (1995)
Ralstonia eutropha	P(3HB)	Fed batch	Ethanol	50	63.5	47	74	0.94	Alderete *et al.* (1993)
Ralstonia eutropha	P(3HB/3HV)	Glucose control fed batch	Glucose + propionic acid	46	158	117	74	2.55	Kim *et al.* (1994b)
Ralstonia eutropha	P(3HB/3HV)	Glucose control fed batch	Glucose + propionic acid	39	113	64	56.5	1.64	Kim *et al.* (1994b)
Ralstonia eutropha	P(3HB/3HV)	Fed batch	Glucose + pentanoic acid	48	9.8	6.4	65	0.13	Ramsay *et al.* (1990)
Ralstonia latus	P(3HB)	Continuous	Sucrose	D=0.16/h	-	16.2	-	2.6	Lee and Chang (1995)
Ralstonia latus	P(3HB/3HV)	One-stage continuous	Sucrose + propionic acid	D=0.15/h	4.65	2	43	0.3	Ramsay *et al.* (1990)
Azotobacter vinelandii	P(3HB)	Glucose control fed batch	Glucose + fish peptone	47	40.1	32	79.8	0.68	Page and Cornish, (1993).
Azotobacter vinelandii	P(3HB/3HV)	Fed batch	Beet molasses + pentanoic acid	-	-	19-22	59-71	-	Page *et al.* (1992).
Haloferax mediterrani	P(3HB/3HV)	Continuous	Starch	D=0.02/h	-	1.5	-	0.03	Lillo and Rodriguez-Valera, (1990)
Klebsiella aerogenes	P(3HB)	Fed batch	Molasses	32	37	24	65	0.75	Zhang *et al.* (1994)
Paracoccus denitrificans	P(3HB/3HV)	Fed batch	Methanol + n-amyl alcohol	120	9	2.34	26	0.02	Ueda *et al.* (1992)

Table 1.3. Summary of PHA production by various microorganisms and culture methods (Adapted from Lee and Chang, 1995) (continued).

Bacteria	PHA	Culture method	Major substrate	Culture time (h)	Cell Conc. (g/L)	PHA conc. (g/L)	PHA conc. (%)	Productivity (g/L/h)	Reference
Protomonas extorquens	P(3HB)	Fully automatic fed batch	Methanol	121	223	136	61	1.12	Suzuki *et al.* (1986a)
Protomonas extorquens	P(3HB)	Fully automatic fed batch	Methanol	170	233	149	64	0.88	Suzuki *et al.* (1986a)
Pseudomonas oleovorans	P(3HH/3HO)	Continuous	n-Octane	D=0.09/h	2.25	1.05	46.7	0.09	Preusting *et al.* (1991)
Pseudomonas oleovorans	P(3HH/3HO)	Continuous	n-Octane	D=0.2/h	11.6	2.9	25	0.58	Preusting *et al.* (1993a)
Pseudomonas oleovorans	P(3HH/3HO)	Fed batch	n-Octane	38	37.1	12.1	33	0.32	Preusting *et al.* (1993b)
Pseudomonas oleovorans	P(3HH/3HO)	Fed batch	Octanoic acid	45	41.8	15.5	37.1	0.34	Lee and Chang, (1995)
Recombinant *E. coli*	P(3HB)	pH-stat fed batch	Glucose + LB medium	42	117	89	76	2.11	Kim *et al.* (1992)
Recombinant *E. coli*	P(3HB)	pH-stat fed batch	Glucose	35	71.4	16.3	22.8	0.46	Lee *et al.* (1994a)
Recombinant *E. coli*	P(3HB)	pH-stat fed batch	Glucose + tryptone + thiamine	44	104.5	66.7	63.8	1.52	Lee and Chang, (1994)
Recombinant *E. coli*	P(3HB)	pH-stat fed batch	Glucose + yeast extract +corn steep liquor	41.5	116	72.2	62.2	1.74	Lee and Chang, (1994)
Recombinant *E. coli*	P(3HB)	pH-stat fed batch	Glucose + yeast extract + corn steep liquor + casein hydrolysate	41	112	71	72.3	1.98	Lee and Chang, (1995)

present the highest productivity obtained with recombinant *E. coli* was 3.4 g of PHA/L/h (Wang and Lee, 1997b), which is considerably less than that obtained from the best naturally occurring bacterial isolates of *A. latus* (4.9 g of PHA/L/h) and the copolymer is mainly P(3HB/3HV) (Wang and Lee, 1997a). PHA heteropolymers and copolymers have superior physical properties than P(3HB), but the recombinant systems cannot reliably produce these heteropolymers.

Despite this problem, it should be possible through careful systematic design of expression systems, strain selection and manipulation as well as choice of growth substrates to develop systems that can produce useful PHAs. Induction of expression through the use of temperature shifts, or use of chemical inducers such as IPTG, are not economically feasible in an industrial setting. Ideally, control systems that can be automatically switched on/off by naturally occurring components in the growth media should be investigated. For example, the *trp* promoter is negatively regulated by tryptophan. When there is tryptophan, *trp* promoter is switched off, and *pha* genes do not express, while tryptophan is used up, this promoter is switched on, resulting expression of *pha* genes and PHA synthesis.

Other approaches could involve mixing and matching *pha* genes on an expression system. Although *phaA* and *phaB* genes are highly conserved, *phaC* genes are quite variable. Thus expression of PHA_{SCL} or PHA_{MCL} could use $phaC_{SCL}$ genes from bacteria such as *R. eutropha*, *A. latus*, or $phaC_{MCL}$ genes from bacteria such as *P. putida*, and *P. resinovoran*. Another advantage of employing *phaC* genes from different sources is that the feedback regulation of the end product PHA to β-ketothiolase and/or acetoacetyl-CoA reductase in PHA synthetic pathway maybe decreased, resulting a high level of PHA production. A recent and exciting finding is a new *phaJ* gene isolated from *Aeromonas caviae* by Fukui *et al.* (1998). This gene encodes an enoyl-CoA hydratase and participates in PHA biosynthesis by hydration trans-2-enoyl-CoA to (*R*)-3-hydroxyacyl-CoA monomer units from fatty acid β-oxidation. Fukui *et al.* (1999) employed *E. coli* LS5218 [*fadR*, *atoC* (Con)], harbouring the A. *cavia phaC* and *phaJ* genes and produced P(3HB/3HHx) copolymer from a medium containing octanoate and dodecanoate. Thus *phaJ* could be combined with the *phaCAB* genes for PHA heteropolymer production. As well, the *parB* locus will also be introduced into PHA production plasmid to keep its stability in culture.

If tallow is used as a supply for fatty acids, another gene encoding lipase can also be introduced in to the PHA production plasmid.

From the above discussion, it is clear that if commercial production of PHA is to be successful, different countries or regions will develop production facilities that are best catered for by available raw materials and resources.

1.8 Production and disposal

Biodegradable plastics such as PHA are expected to be a solution to the current disposal problem of commodity plastic waste. They are attractive for their biodegradability as well as savings of fossil energy resources. However, they also have problems in terms of production as discussed previously, and disposal. Wegner and Wagemann (1994) proposed the following points to be considered in biodegradable polymer research.

1. Biodegradable polymers must be strictly separated from those which are not biodegradable. Recycling of the latter would be strongly hampered by the presence of even small amounts of biodegradable polymers, since these would compromise the stability and tense.

2. The complete degradation of "biodegradable" polymers in waste dumps can not be guaranteed. Some "natural" materials, eg. wood, paper, cellulose and protein-based textiles have been demonstrated their survivals in dumps for many decades by excavations.

3. Degradation of biodegradable polymers in a large amount may make the CO_2 balance even worse and the methane converted from these polymers may contribute significantly to the greenhouse effect.

4. The production of biodegradable plastics from agricultural wastes is probably problematic with regard to the need for large amounts of water, which in turn cause water treatment problems.

1.9 Specific aims of this thesis

Although PHA production by naturally occurring bacteria offers more opportunities to produce PHA heteropolymers, this option is likely to be replaced by recombinant *E. coli*. P(3HB) is the dominant product of current recombinant *E. coli* strains, but is a less than a perfect alternative to commodity plastics, due to its inferior physical properties. Clearly PHA heteropolymers have superior physical properties than P(3HB), and solutions which facilitate large scale production of heteropolymers must be developed if PHAs are to survive as an economical alternative source of plastics. As discussed in Section 1.4.4, for PHA heteropolymer production, the bacterial metabolic system should provide different precursors for the PHA synthetic enzymes and the PHA synthetic enzymes should have a wide range of substrate specificity to these precursors. If the genes coding for these PHA synthetic enzymes with a wide substrate specificity are cloned and expressed in an appropriate *E. coli* strain, hopefully, PHA heteropolymers can be produced. With this strategy, the specific aims of this thesis were to clone novel PHA synthetic genes and then express these genes in *E. coli* for PHA production, including:

1. Isolation and identification of novel PHA producing bacteria, from which novel PHA synthetic genes would be cloned;

2. Cloning and genetic characterisation of the PHA synthetic genes from the selected bacterial isolate;

3. Evaluation of the use of the cloned genes for PHA production using different *E. coli* host strains.

The work describing isolation and identification of novel PHA producing bacteria is described in Chapter 3. Detection of *pha* genes, genomic library construction and screening is described in Chapter 4. Subcloning, DNA sequence analysis of *pha* genes and construction of plasmids for PHA expression are described in Chapter 5. Finally, investigation of a novel *pha* gene is described in Chapter 6.

Chapter 2

Materials and Methods

2.1 Chemicals and reagents

Chemicals were Analar grade. Unless otherwise stated, all chemicals used in this study were purchased from either Ajax Chemicals (Auburn, New South Wales, Australia), BDH Laboratory Supplies (Poole, Dorset, England) or Sigma Chemical Company (St Louis, Missouri, USA). Acetic acid, HCl, CsCl, phenol, EDTA, SDS, sodium chloride, sodium acetate, formamide, and sucrose were purchased from BDH Laboratory Supplies. Ethanol, methanol, iso-propanol, iso-amyl alcohol, Triton X-100, chloroform, formaldehyde, calcium chloride, cobalt chloride, magnesium chloride, magnesium sulphate, potassium chloride, potassium di-hydrogen orthophosphate, di-potassium hydrogen orthophosphate, acetone, and sodium hydrogen carbonate were obtained from Ajax Chemicals. Coomassie brilliant blue R250, TEMED, EGTA, X-phosphate, were obtained from Sigma Chemical Company. X-gal was purchased from Progen Industries Ltd. (Darra, Queensland, Australia). Digoxigenin DNA labelling and detection kits, Tris base, IPTG, NBT, and glycine were purchased from Boehringer-Mannheim (GmbH, Mannheim, Germany). Amphotericin B was purchased from ICN, Biomedicals Inc. (Aurora, Ohio, USA). chloramphenicol, cycloheximide, and Tetracycline were purchased from Sigma Chemical Company (St. Louis, Missouri, USA). Ampicillin and Kanamycin were purchased from Progen (Darra, Queensland, Australia). Ultrapure dATP, dCTP, dGTP and dTTP were purchased from Pharmacia (Uppsala, Sweden). Phenol (Special grade) for isolation of bacterial RNA, was obtained from Wako Pure Chemical Industries Ltd. (Osaka, Japan).

The following electrophoresis grade reagents were obtained from the sources as indicated: DNA grade agarose (Progen Industries Ltd. Queensland, Australia), acrylamide and APS (Bio-Rad, Richmond, California, USA), and ultra pure N,N-methylene bis acrylamide and urea (Bethesda Research Laboratories, Grand Island, New York, USA).

Milli Q grade, organic free reagent water (Millipore Corporation, Bedford, Massachusetts, USA.) was used to prepare all buffers and reagents for DNA and RNA manipulation.

2.2 Enzymes and antibodies

Restriction endonucleases were purchased from Boehringer Mannheim (GmbH, Mannheim, Germany), MBI Fermentas (Buffalo, New York, USA), New England Biolabs (Beverly, Massachusetts, USA), Progen Industries Ltd. (Queensland, Australia), or Promega Corporation (Madison, Wisconsin, USA). Lysozyme, pronase, T4 DNA ligase, and shrimp alkaline phosphatase were obtained from Boehringer Mannheim. RNase A was acquired from Sigma Chemical Company, and DNase I (RNase free) was purchased from Promega. Stock solutions of RNase A (10 mg/ml) were heated at 95 °C for 10 min prior to use, to inactivate contaminating DNase. *Taq* DNA polymerase (Taqbead^{TM} hot start polymerase) and T4 DNA ligase were obtained from Promega. Dye primer/terminator sequencing kits were purchased from Applied Biosystems (Perkin-Elmer Corporation, Norwalk, Connecticut, USA).

Anti-Digoxigenin-Alkaline Phosphatase conjugate (Fab fragments) was obtained from Boehringer Mannheim.

2.3 Culture Media

2.3.1 General growth media

Nutrient broth (NB) and nutrient agar (NA) were Oxoid (Oxoid Ltd, London, England). Luria broth was prepared as described by Miller (1972). Terrific broth, SOC and 2×YT media were prepared as described by Sambrook *et al.* (1989). All media were sterilised in an autoclave before use. The composition of each medium is listed in Appendix B.

2.3.2 Media for isolation of PHA producing bacteria

PHA producing bacteria were cultivated on and isolated from PHA medium. This medium is a modification of poly-β-hydroxybutyrate medium agar as described by Atlas (1993). In this medium, hydroxybutyrate was replaced by glucose as the carbon source (Appendix B). Amphotericin B and cycloheximide were added to this medium to inhibit the growth of soil-borne fungi.

Antibiotics, where indicated, were added to broth and solid media at the final concentrations as indicated in Section 2.4.

PHA producing isolates were incubated at 30 °C and strains of *E. coli* were incubated at 37 °C unless otherwise indicated.

2.3.3 PHA production media

For broth cultures, NB (supplemented with 10 or 20 g/L glucose) was used as a complex growth medium and R-medium (supplemented with 10 or 20 g/L glucose) (Kim *et al.*, 1994a) was used as a defined growth medium. Broths were dispensed in 100 ml volumes in 500 ml flasks. Flasks of culture were incubated in a rotary shaker (New Brunswick Co. Inc., Edson, New Jersey) set at 114 rpm and 37 °C for *E. coli* and at 114 rpm and 30 °C for PHA producing isolates. The composition of each medium is listed in Appendix B. Where required, antibiotics were added to broth and solid media at the final concentrations as indicated in Section 2.4.

2.4 Antibiotics

Filter (Millipore, Millex-GP pore size 0.22 μm) sterilised solutions of antibiotics were added to broth and solid media as required at the following final concentrations: Amphotericin B, 15 μg/ml; Ampicillin (Ap), 100 μg/ml; Cycloheximide (Cyc) (actidione), 70 μg/ml; Kanamycin (Kan), 50 μg/ml; Streptomycin (Sm), 100 μg/ml; and Tetracycline (Tc), 15 μg/ml.

2.5 Maintenance and propagation of bacterial strains

Suspensions of all strains were stored at –70 °C in Wheaton vials (Millville, New Jersey, USA) containing 1 ml solution of 1% (w/v) Bacto peptone (Difco) containing 15% (v/v) glycerol. Single colonies of PHA producing isolates and *E. coli* strains were prepared by streaking a loopful of frozen glycerol stock onto the appropriate media and incubated O/N at 30 °C for PHA producing isolates and 37 °C for *E. coli* strains. For routine use, cultures were maintained on agar plates at 4 °C.

2.6 Characterisation of PHA producing isolates

PHA producing bacterial strains isolated from soil samples were identified using a scheme proposed by Sirockin and Cullimore (1969) (Figure 2.1). O/N NA cultures were used to assess Gram reaction (Pelczar and Chan, 1977), and presence of endospores (Sirockin and Cullimore, 1969). The cultures were also used to test the utilisation of ethanol and catalase reaction (MacFaddin, 1976). Motility in NA, oxidation-fermentation of sugars, and oxidase tests were carried out as described by MacFaddin (1976).

2.7 PHA analysis

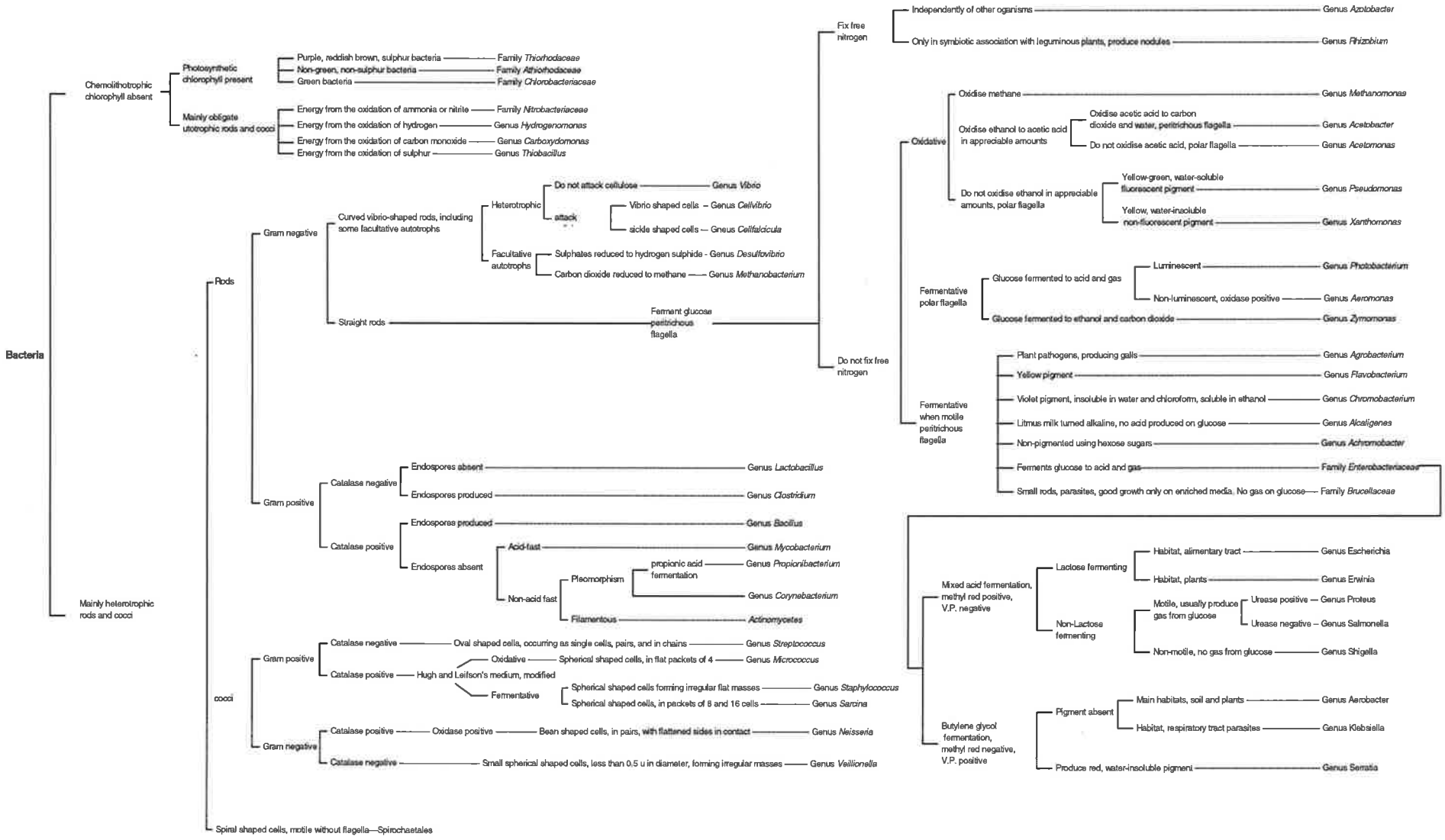
2.7.1 Staining of bacterial cells containing PHA granules

To visualise PHA granules in bacterial cells, cells were grown in NB (supplemented with 10 or 20 g/L glucose) O/N for PHA production, taken, washed 3 times in distilled water and heat-fixed onto a clean microscope slide. Smears were then stained with crystal violet or Nile Blue A as follows.

Smears were stained with 1% (w/v) crystal violet solution for 1 min, washed with tap water for 1 min to remove any excess stain and finally blotted dry with tissue paper. The preparations were covered with a glass cover slip (mounting medium-Movoil with anti-bleaching agent) and examined by transmitted light microscopy. PHA granules are not stained and appear as clear zones, while the other parts of the cells are stained purple.

Smears were stained with filtered 1% (w/v) aqueous solution of Nile Blue A at 55°C for 10 min in a Coplin jar (Ostle and Holt, 1982). Slides were washed with tap water to remove excess stain, then soaked in 8% (v/v) aqueous acetic acid for 1 min. The stained smear was then washed, blotted dry with tissue paper and covered with a glass cover slip (mounting medium-Movoil with anti-bleaching agent). Stained smears were viewed with an Olympus BHS Microscope fitted with a BH2-RFC UV fluorescent attachment. PHA granules exhibit a strong orange fluorescence when observed at an excitation wavelength of 460 nm.

Figure 2.1. A scheme for the identification of bacteria isolated from soil samples [Adapted from Sirockin and Cullimore (1969)].


Bacteria
Chemolithotrophic chlorophyll absent
Photosynthetic chlorophyll present
Purple, reddish brown, sulphur bacteria — Family Thiorhodaceae
Non-green, non-sulphur bacteria — Family Athiorhodaceae
Green bacteria — Family Chlorobacteriaceae
Mainly obligate autotrophic rods and cocci
Energy from the oxidation of ammonia or nitrite — Family Nitrobacteriaceae
Energy from the oxidation of hydrogen — Genus Hydrogenomonas
Energy from the oxidation of carbon monoxide — Genus Carboxydomonas
Energy from the oxidation of sulphur — Genus Thiobacillus
Mainly heterotrophic rods and cocci
Rods
Gram negative
Curved vibrio-shaped rods, including some facultative autotrophs
Heterotrophic
Do not attack cellulose — Genus Vibrio
attack
Vibrio shaped cells – Genus Cellvibrio
sickle shaped cells — Gneus Cellfalcicula
Facultative autotrophs
Sulphates reduced to hydrogen sulphide - Genus Desulfovibrio
Carbon dioxide reduced to methane — Genus Methanobacterium
Straight rods
Ferment glucose peritrichous flagella
Fix free nitrogen
Independently of other organisms — Genus Azotobacter
Only in symbiotic association with leguminous plants, produce nodules — Genus Rhizobium
Do not fix free nitrogen
Oxidative
Oxidise methane — Genus Methanomonas
Oxidise ethanol to acetic acid in appreciable amounts
Oxidise acetic acid to carbon dioxide and water, peritrichous flagella — Genus Acetobacter
Do not oxidise acetic acid, polar flagella — Genus Acetomonas
Do not oxidise ethanol in appreciable amounts, polar flagella
Yellow-green, water-soluble fluorescent pigment — Genus Pseudomonas
Yellow, water-insoluble non-fluorescent pigment — Genus Xanthomonas
Fermentative polar flagella
Glucose fermented to acid and gas
Luminescent — Genus Photobacterium
Non-luminescent, oxidase positive — Genus Aeromonas
Glucose fermented to ethanol and carbon dioxide — Genus Zymomonas
Fermentative when motile peritrichous flagella
Plant pathogens, producing galls — Genus Agrobacterium
Yellow pigment — Genus Flavobacterium
Violet pigment, insoluble in water and chloroform, soluble in ethanol — Genus Chromobacterium
Litmus milk turned alkaline, no acid produced on glucose — Genus Alcaligenes
Non-pigmented using hexose sugars — Genus Achromobacter
Ferments glucose to acid and gas — Family Enterobacteriaceae
Small rods, parasites, good growth only on enriched media. No gas on glucose — Family Brucellaceae
Gram positive
Catalase negative
Endospores absent — Genus Lactobacillus
Endospores produced — Genus Clostridium
Catalase positive
Endospores produced — Genus Bacillus
Endospores absent
Acid-fast — Genus Mycobacterium
Non-acid fast
Pleomorphism
propionic acid fermentation — Genus Propionibacterium
Genus Corynebacterium
Filamentous — Actinomycetes
cocci
Gram positive
Catalase negative — Oval shaped cells, occurring as single cells, pairs, and in chains — Genus Streptococcus
Catalase positive — Hugh and Leifson's medium, modified
Oxidative — Spherical shaped cells, in flat packets of 4 — Genus Micrococcus
Fermentative
Spherical shaped cells forming irregular flat masses — Genus Staphylococcus
Spherical shaped cells, in packets of 8 and 16 cells — Genus Sarcina
Gram negative
Catalase positive — Oxidase positive — Bean shaped cells, in pairs, with flattened sides in contact — Genus Neisseria
Catalase negative — Small spherical shaped cells, less than 0.5 u in diameter, forming irregular masses — Genus Veillonella
Spiral shaped cells, motile without flagella—Spirochaetales
Mixed acid fermentation, methyl red positive, V.P. negative
Lactose fermenting
Habitat, alimentary tract — Genus Escherichia
Habitat, plants — Genus Erwinia
Non-Lactose fermenting
Motile, usually produce gas from glucose
Urease positive — Genus Proteus
Urease negative – Genus Salmonella
Non-motile, no gas from glucose — Genus Shigella
Butylene glycol fermentation, methyl red negative, V.P. positive
Pigment absent
Main habitats, soil and plants — Genus Aerobacter
Habitat, respiratory tract parasites — Genus Klebsiella
Produce red, water-insoluble pigment — Genus Serratia

2.7.2 Transmission electron microscopy examination of PHA granules

Washed cultures were pelleted by centrifugation, then fixed by suspension in a solution of 1.25% (v/v) glutaraldehyde, 4% (v/v) paraformaldehyde, and 4% (w/v) sucrose in 0.1 M phosphate buffer (pH7.2) for about 2 h at 4 °C. Cells were gently pelleted and washed with 0.1 M phosphate buffer [containing 4% (w/v) sucrose] then postfixed in 0.1 M phosphate buffer containing 2% (w/v) osmium tetroxide (OsO_4). After washing in 0.1 M phosphate buffer, the cell pellet was dehydrated in graded solutions of ethanol [50%, 60%, 70%, 80%, 90%, 95%, and 100% (v/v), 10-15 min for each step] followed by three washes in 100% anhydrous ethanol before infiltration with ethanol:resin (1:1, v/v) (Procure-Araldite embedding kit, ProSciTech) mixture O/N. The pellet was further infiltrated with fresh resin O/N before polymerisation at 70 °C. Thin sections were prepared with Reichert Ultracut ultramicrotome, placed on 300 mesh copper EM grids and stained with 5% (w/v) uranyl acetate and Reynolds' lead citrate (Reynolds, 1963). Stained sections were examined in a Philips CM100 transmission electron microscope.

2.7.3 Lyophilisation of PHA-containing bacterial cells

Culture samples were pelleted in a Beckman J2-20M at 15,000 rpm for 5 min and washed 3 × with distilled water. Cells were lyophilised in 10 ml glass tubes for approximately 24h in an Edwards Modulyo Freeze Drier. Dry cell preparations were then used for GC analysis.

2.7.4 Dry cell mass determination

Eppendorf tubes (1.5 ml) were placed in an oven at 105 °C for at least 24 h. A 1.5 ml culture sample was placed in a preweighed tube and pelleted at 15,000 rpm for 2 min in a Biofuge 15 benchtop centrifuge. An additional 1.5 ml culture sample was then added to the tube and pelleted. The cells were washed twice in 1 ml of distilled water and the pellet lyophilised for 24 h. The dry cell mass was determined gravimetrically.

2.7.5 Gas chromatography (GC) analysis

GC assay for PHA in culture samples was based on a method described by Braunegg *et al.* (1978). The amount of PHA was determined by methanolysis of PHA monomers followed by GC analysis.

2.7.5.1 Sample preparation for GC analysis

Lyophilised cell samples were crushed into a fine powder, placed in a screw-cap test tube and mixed with 2 ml methanol [containing 3% concentrated H_2SO_4 (v/v)], 2 ml chloroform, and 200 µl of 0.4% (w/v) benzoic acid (dissolved in methanol) as an internal standard. The tubes were closed tightly and incubated at 100°C for 3.5 h during which the mixture was shaken two or three times. After the mixture was cooled to room temperature, 1 ml of Milli Q water was added into the mixture and the sample shaken vigorously for 2 min or more. Three phases were allowed to separate, during which cell debris gathered at the interphase. The top aqueous phase was removed and discarded and the remaining organic phase was analysed for presence of PHA by GC analysis.

2.7.5.2 PHA standard sample preparation

20-30 mg of pure P(3HB) (CarboMer Inc., MA, USA) was dissolved into 10 ml of chloroform using a volumetric flask in a water bath. 100 µl, 300 µl, 500 µl, and 700 µl volumes of this P(3HB) solution were pipetted into four GC tubes and 1.9 ml, 1.7 ml, 1.5 ml, and 1.3 ml of chloroform were added into these four tubes respectively. 2 ml of Methanol and 200 µl of 0.4% (w/v, g/ml) benzoic acid dissolved in methanol as internal standard were also added into each tube. Methanolysis was carried out as described in Section 2.7.5.1.

2.7.5.3 Gas chromatography

An Hewlett-Packard gas chromatograph (Model HP 5830) equipped with a double FID was used for GC analysis. A 5 µl Hamilton syringe (AdeLab) was used to inject all samples into a stainless steel 6 ft capillary column (internal diameter of 0.125 inch) filled with 10% Carbowax 20 M on a Chromosorb WAW 80 100 mesh support (Supelco). The initial temperature was 100 °C and the final temperature was 180 °C. The holding time for

the initial temperature was 1 min and for the final temperature 10 min. The temperature of the column was increased at a rate of 8 °C/min. The nitrogen carrier gas flow rate was 30 ml/min.

The gas chromatograph was calibrated as follows. 2 µl of P(3HB) standard solution (see Section 2.7.5.2) was injected into the GC and then the thermographs containing peaks representing P(3HB) and benzoic acid as an internal standard were obtained. The peak-area and the pure P(3HB) amount in each standard sample were calculated and then the curve of P(3HB) (mg) vs peak-area as a calibration curve was plotted (Figure 2.2). As assay response varied between experiments, calibration was repeated once every day.

Unknown samples were injected as above. Peak areas and PHA associated peak were determined. The PHA amount in each sample was then determined using the calibration data and the PHA concentration of each sample was calculated against its dry cell weight.

2.7.6 Differential scanning calorimetry of P(3HB)

Differential scanning calorimetry (DSC) is a technique for detecting thermal effects accompanying physical or chemical change in a sample by means of programmed heating or cooling. A polymer will undergo major physical change in its morphological state at transition points [eg. the glass transition temperature (Tg) for amorphous polymers and the crystallisation temperature (Tcr) for crystalline polymers]. A change in enthalpy occurs and this can be followed by calorimetry. Microbial-sourced P(3HB) is in a highly crystalline state after extraction due to its native and exceptional stereochemical regularity, and an endotherm in the DSC thermogram appears around the melting point (Tm). The crystallinity of P(3HB) can be estimated by comparison of the enthalpy of fusion with that of pure crystalline P(3HB) 146 J/g (Barham *et al.*, 1984). P(3HB) glass transition can also be monitored using DSC. The amorphous state of P(3HB) is obtained after quenching the sample from the fusion point to a low temperature (about -100 °C).

A DuPont 2200 Thermal Analyser was used for this analysis. Standard indium was used for temperature calibration and nitrogen as the carrier gas. A lyophilised sample containing about 10 mg of PHA was weighed, encapsulated in

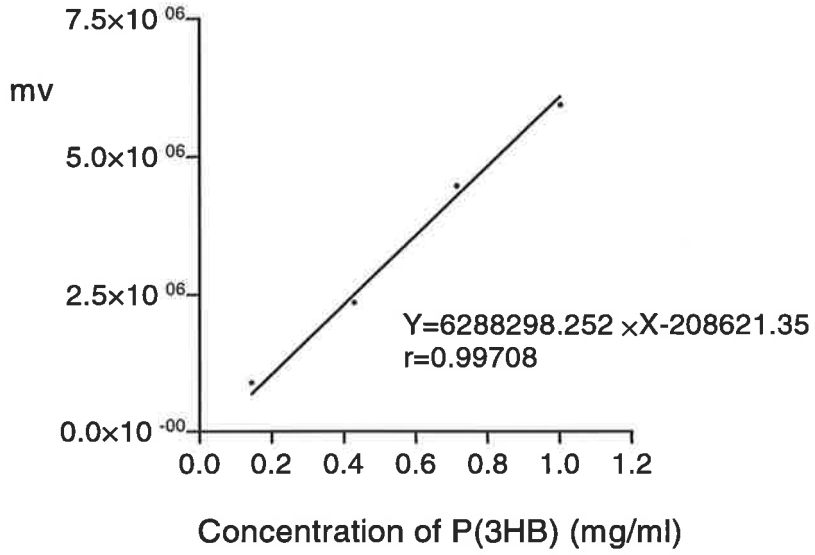


Figure 2.2. A typical GC calibration curve for P(3HB) concentration (mg/ml) vs peak-area.

r: coefficient of linear regression; Y: Peak area (mv); and X: P(3HB) concentration (mg/ml).

a standard aluminium pan and placed into the sample chamber. A thermal cycle was programmed as listed below:

1. Equilibrate at 25 °C;
2. Ramp to 200 °C at 10 °C/min;
3. Isothermal for 5.0 min;
4. Ramp 50.0 °C/min to -100 °C;
5. Iso-track for 5.0 min;
6. Ramp 10.0 °C/min to 200.0 °C.

Figure 2.3 illustrates a typical DSC thermogram of P(3HB) under the above stated operating conditions. Morphological transition of polymer is temperature-dependent. At a heating rate of 10 °C/min, the peak point of the fusion endotherm was taken as the Tm (~174 °C). The inflection point of the specific heat increment at the glass transition range in the second scan was taken as the glass transition temperature (~0 °C).

2.8 Bacterial strains and vectors

Bacterial strains and vectors used in this study are listed in Table 2.1 and Table 2.2 respectively. All the strains were stored as suspensions in a glycerol-based medium and maintained at -70 °C.

2.9 Solutions and buffers

Compositions of all solutions and buffers used in this study are listed in Appendix A.

2.10 DNA extraction

2.10.1 Preparation of genomic DNA

Chromosomal DNA was isolated by a modification of the method described by Manning *et al.* (1986). Briefly, bacterial cells were incubated O/N in NB, pelleted in 20 ml McCartney bottles in an MSE Minor S centrifuge (Crawley, West Sussex, England) at

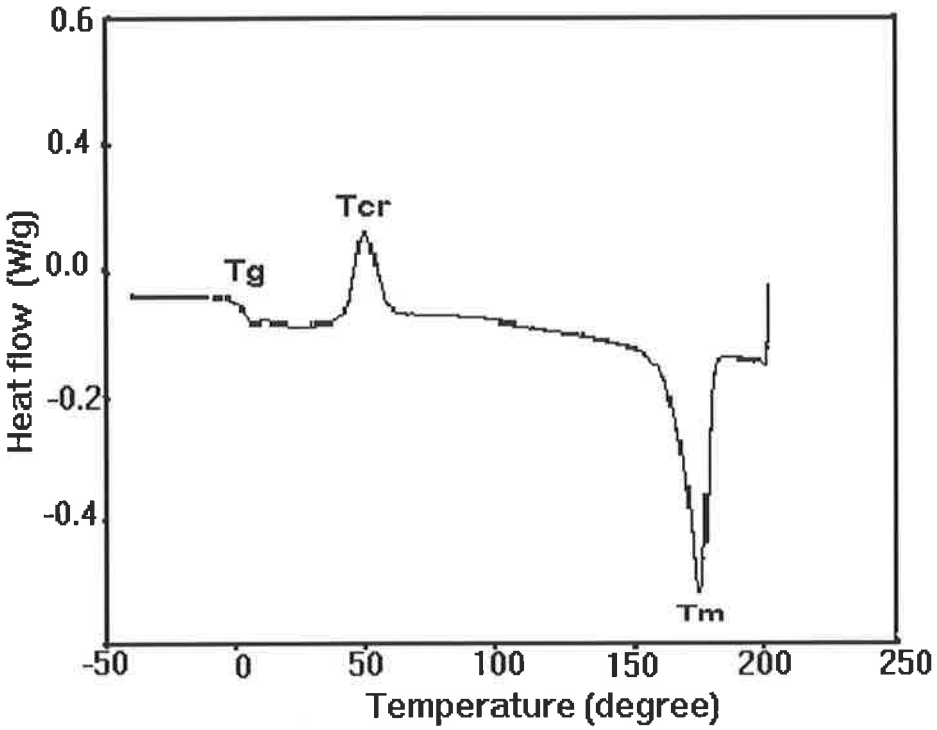


Figure 2.3. A typical DSC thermogram of P(3HB).

Tg: P(3HB) glass transition temperature (~0 °C); Tcr: P(3HB) crystallisation temperature (~50 °C); and Tm: P(3HB) melting temperature (~174 °C).

Table 2.1. Bacterial strains used in this thesis.

Species/Isolate No.	Genotype/Phenotype	Source
R. eutropha H16	producing PHA, wild type	S.Y., Lee[1]
Isolate 1c-1	producing PHA, wild type	This study
Isolate 1c-1-1	producing PHA, wild type	This study
Isolate 2-3-2	producing PHA, wild type	This study
Isolate 3	producing PHA, wild type	This study
Isolate 6c-2	producing PHA, wild type	This study
Isolate 10c-1-1	producing PHA, wild type	This study
Isolate 10c-1-3	producing PHA, wild type	This study
Isolate 14b-1	producing PHA, wild type	This study
Isolate 14c-1-2	producing PHA, wild type	This study
Isolate 14c-1-3	producing PHA, wild type	This study
Isolate K1b-1	producing PHA, wild type	This study
Isolate K1c-1	producing PHA, wild type	This study
Isolate M1c-1(1)	producing PHA, wild type	This study
Isolate M2	producing PHA, wild type	This study
Isolate M4-1	producing PHA, wild type	This study
Isolate M4-2	producing PHA, wild type	This study
Isolate M10c-1	producing PHA, wild type	This study
Isolate M10c-2	producing PHA, wild type	This study
Isolate M14c-1	producing PHA, wild type	This study
Pseudomonas strain 10c-1-3 PS001	producing PHA, Sm^r	This study
Pseudomonas strain 10c-1-3 PS002	producing PHA, ORF4::*kan*, Sm^r, Kan^r	This study
E. coli DH5α	F⁻ ϕ80d *recA1 endA1 hsdR17*(r_k^-, m_k^+) *supE44*λ^- *thi1 gyrA96 relA1 deoR lacZ*ΔM15Δ(*lacZYA-argF*) U169	GibcoBRL
E. coli SM10 λpir	*thi-1 thr leu tonA lacY supE recA*::RP4-2-Tc::Mu, Kan^r, λpir	Miller and Mekalanos, 1988
E. coli SY327 λpir	Δ(*lac pro*) *argE*(Am) *rif nalA recA56*	Donnenberg and Kaper, 1991
E. coli MD9101	*fadR601*, *atoC*2(Con), *recA56*, *srl*::Tn10	Slater *et al.*, 1992

Species/Isolate No.	Genotype/Phenotype	Source
E. coli XL1-Blue	*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac*, [F' *proAB*, *laq1^q*ZΔ*M15*, Tn10(Tc^r)]	Stratagene
E. coli Topp1	Rif^r[F', *proAB*, *lac1^q*, ZΔ*M15*, Tn10, (Tc^r)]	R. Morona[2]
E. coli Topp2	Rif^r[F', *proAB*, *lac1^q*, ZΔ*M15*, Tn10, (Tc^r)]	R. Morona[2]
E. coli Topp3	Rif^r[F', *proAB*, *lac1^q*, ZΔ*M15*, Tn10, (Tc^r, Kan^r)]	R. Morona[2]
E. coli Topp4	Rif^r[F', *proAB*, *lac1^q*, ZΔ*M15*, Tn10, (Tc^r)]	R. Morona[2]
E. coli Topp5	Rif^r[F', *proAB*, *lac1^q*, ZΔ*M15*, Tn10, (Tc^r)]	R. Morona[2]
E. coli Topp6	Rif^r[F', *proAB*, *lac1^q*, ZΔ*M15*, Tn10, (Tc^r)]	R. Morona[2]
E. coli SGZ001	*E. coli* DH5α (pCT411)	This study
E. coli SGZ002	*E. coli* DH5α (pCT415)	This study
E. coli SGZ003	*E. coli* DH5α (pCT416)	This study
E. coli SGZ004	*E. coli* DH5α (pCT417)	This study
E. coli SGZ005	*E. coli* Topp1 (pCT415)	This study
E. coli SGZ006	*E. coli* Topp1 (pCT416)	This study
E. coli SGZ007	*E. coli* Topp1 (pCT417)	This study
E. coli SGZ008	*E. coli* Topp2 (pCT415)	This study
E. coli SGZ009	*E. coli* Topp2 (pCT416)	This study
E. coli SGZ010	*E. coli* Topp2 (pCT417)	This study
E. coli SGZ011	*E. coli* MD9101 (pCT415)	This study
E. coli SGZ012	*E. coli* MD9101 (pCT416)	This study
E. coli SGZ013	*E. coli* MD9101 (pCT417)	This study
E. coli SGZ014	*E. coli* XL1-Blue (pCT415)	This study
E. coli SGZ015	*E. coli* XL1-Blue (pCT416)	This study
E. coli SGZ016	*E. coli* XL1-Blue (pCT417)	This study
E. coli SGZ017	*E. coli* DH5α (pCT418)	This study
E. coli SGZ018	*E. coli* MD9101 (pCT418)	This study

1. Department of Chemical Engineering, Advanced Institute of Science and Technology (KAIST), Taejon, Korea.
2. Department of Molecular BioSciences, Adelaide University.

Table 2.2. Vectors and plasmids used in this thesis.

Vector/Plasmid	Description	Source
pBluescript II SK(+)	Cloning plasmid vector (Ap^r, *lacZ*)	Stratagene
pBluescript II KS(+)	Cloning plasmid vector (Ap^r, *lacZ*)	Stratagene
pGEM7Zf(+)	Cloning plasmid vector (Ap^r, *lacZ*)	Promega
pGEM5Zf(+)	Cloning plasmid vector (Ap^r, *lacZ*)	Promega
pGEM-T	Cloning plasmid vector. pGEM5Zf(+) linearised with T overhangs for direct cloning of PCR products (Ap^r, *lacZ*)	Promega
pHC79	Cloning cosmid vector (Ap^r, Tc^r)	Hohn and Collins, 1980
pQE-50	Expression vector (Ap^r)	Qiagen
pCVD442	Suicide vector (Ap^r, *sacB*, *mob*RP4)	Donnenberg and Kaper, 1991
pBSL15	Ap^r, Kan^r gene cassette with flanking polylinkers	Alexeyev, 1995
pGP1-2	T7 RNA Polymerase gene and λ repressor gene in pKC30 (Kan^r)	Tabor and Richardson, 1985
pHB1	2.9 kbp *Eco* RI/*Pst* I *hlyB* fragment from pPM431 in pBluescript II SK(+)	P. Ahtonen[1]
p4A	*phaCAB* of *R. eutropha* in pTZ18U (Ap^r)	Slater *et al.*, 1992
pJM9123	*phaCAB* of *R. eutropha* in pTZ18U, (Kan^r, *parB*)	Slater *et al.*, 1992
pSGZ1	0.2 kbp $phaA_{int}$ PCR product of *Pseudomonas* strain 10C-1-3 in pGEM-T (Ap^r)	This study
pSGZ2	1.6 kbp *Xba* I DNA fragment (Kan^r, *parB*) from pJM9123 in pBluescript II KS(+) (Ap^r)	This study
pSGZ3	1.6 kbp *Xba* I DNA fragment (Kan^r, *parB*) from pJM9123 in pQE-50 (Ap^r)	This study

Vector/Plasmid	Description	Source
pCT400	~34 kbp chromosomal DNA (including *phaCAB*ORF4) of *Pseudomonas* strain 10c-1-3 in pHC79 (Ap^r, Tc^s)	This study
pCT401	~4.1 kbp chromosomal DNA (including *phaC*, and partial *phaA*) of *Pseudomonas* strain 10c-1-3 in pGEM7Zf(+) (Ap^r)	This study
pCT402	~28 kbp chromosomal DNA (including partial *phaA* and *phaB*ORF4) of *Pseudomonas* strain 10c-1-3 in pGEM7Zf(+) (Ap^r)	This study
pCT403	~1.1 kbp truncated chromosomal DNA of *Pseudomonas* strain 10c-1-3 in pHC79 (Ap^r, Tc^s)	This study
pCT404	~0.7 kbp truncated chromosomal DNA; derivative of pCT401 (Ap^r)	This study
pCT405	~2.2 kbp truncated chromosomal DNA; derivative of pCT401 (Ap^r)	This study
pCT406	~9.0 kbp truncated chromosomal DNA; derivative of pCT402 (Ap^r)	This study
pCT407	~6.0 kbp truncated chromosomal DNA; derivative of pCT402 (Ap^r)	This study
pCT408	~4.8 kbp chromosomal DNA (including partial *phaA*, *phaB*, and ORF4) in pBluescript II SK(+) (Ap^r)	This study
pCT409	~1.5 kbp chromosomal DNA (including partial *phaA*, *phaB*, and partial ORF4) in pBluescript II SK(+) (Ap^r)	This study
pCT410	~3.2 kbp chromosomal DNA (including partial *phaA*, *phaB*, and ORF4) in pBluescript II SK(+) (Ap^r)	This study
pCT411	~8.9 kbp chromosomal DNA (including *phaCAB*ORF4) in pBluescript II SK(+) (Ap^r)	This study
pCT412	~4.8 kbp chromosomal DNA (including partial *phaA*, *phaB*, and ORF4) in pBluescript II KS(+) (Ap^r)	This study

Vector/Plasmid	Description	Source
pCT413	~1.5 kbp chromosomal DNA (including partial *phaA*, *phaB*, and partial ORF4) in pBluescript II KS(+) (Ap^r)	This study
pCT414	~3.2 kbp chromosomal DNA (including partial *phaA*, *phaB*, and ORF4) in pBluescript II KS(+) (Ap^r)	This study
pCT415	~8.9 kbp chromosomal DNA (including *phaCAB*ORF4) in pBluescript II KS(+) (Ap^r)	This study
pCT416	~5.6 kbp chromosomal DNA (including *phaCAB* and partial ORF4) in pBluescript II KS(+) (Ap^r)	This study
pCT417	~7.3 kbp chromosomal DNA (including *phaCAB*ORF4) in pBluescript II KS(+) (Ap^r)	This study
pCT418	1.6 kbp *Xba* I DNA fragment (Kan^r, *parB*) from pJM9123 in pCT415 (Ap^r)	This study
pCT419	0.7 kbp ORF4 PCR product of *Pseudomonas* strain 10C-1-3 in pGEM-T (Ap^r)	This study
pCT420	0.7 kbp ORF4 PCR product of *Pseudomonas* strain 10C-1-3 in pGEM-T (Ap^r)	This study
pCT421	0.7 kbp ORF4 PCR product of *Pseudomonas* strain 10C-1-3 in pGEM-T (Ap^r)	This study
pCT422	1.1 kbp PCR product of Kan^r cartridge from pBSL15 in pGEM-T (Kan^r, Ap^r)	This study
pCT423	1.8 kbp ORF4::*Kan* gene in pGEM-T (Kan^r, Ap^r)	This study
pCT424	1.8 kbp ORF4::*Kan* gene in pGEM-T (Kan^r, Ap^r)	This study
pCT425	1.8 kbp ORF4::*Kan* gene in pCVD442 (Kan^r, Ap^r, *sacB*, *mob*RP4)	This study

1. Department of Molecular BioSciences, Adelaide University.

3,100 rpm for 10 min, and resuspended in 2 ml 25% (w/v) sucrose in 0.05 M Tris-HCl (pH 8.0). Lysozyme (10 mg/ml) in 1 ml of 0.25 M EDTA (pH 8.0) was added and then incubated on ice for 20 min, followed by addition of 0.75 ml TE buffer, 0.25 ml lysis solution [5% (w/v) sarkosyl, 50 mM Tris-HCl (pH 8.0), 62.5 mM EDTA], and 1 mg solid pronase. After incubating at 56 °C for 1 h, DNA was gently extracted three times with phenol and twice with diethyl ether, and then dialysed overnight at 4 °C against 5 L of 1 × TE buffer.

2.10.2 Small scale plasmid and cosmid DNA preparation

Method 1: Small scale quantities of plasmid (or cosmid) DNA (3 to 5 µg per ml) was purified from *E. coli* by a modification of the three-step alkali lysis method of Sambrook *et al.* (1989). An O/N culture (1.5 ml), was collected in a 1.5 ml reaction tube (Sarstedt, Newton, North Carolina, USA) by centrifugation for 30 sec at 14,000 rpm in an Eppendorf tube (Eppendorf, Germany), and the pellet resuspended in 100 µl of solution 1 (Appendix A) with the addition of 5 µl of 10 mg/ml RNase A. After the addition of 200 µl solution 2 (Appendix A), the sample was gently mixed and then incubated on ice for 5 min. A further 5 min incubation on ice was preceded after the addition of 150 µl of solution 3 (Appendix A). The supernatant was collected by centrifugation at 14,000 rpm for 5 min, mixed with 250 µl of solution 4 (Appendix A), incubated on ice for 5 min and then centrifuged at 14,000 rpm for 10 min. The supernatant was transferred to a fresh 1.5 ml reaction tube and plasmid DNA was precipitated by the addition of 0.7 volume of iso-propanol, collected by centrifugation (10 min at 14,000 rpm), washed with 70% (v/v) ethanol and dried *in vacuo* (Speedivac, Savant Instruments, Farmingdale, New York, USA). The DNA pellet was resuspended in 40 µl H_2O.

Method 2: Plasmid (or cosmid) DNA for quick clone identification was extracted by a modification of the three-step alkali lysis method of Sambrook *et al.* (1989) by Cormack and Somssich (1997). Colonies were used to inoculate 1-2 ml of LB broth containing the appropriate antibiotic and grown overnight at 37 °C with agitation. Then 0.2 ml of each culture was transferred to microcentrifuge tubes, 0.2 ml of Solution 2 (Appendix A) was added and the tubes were closed and inverted four times to mix. Immediately, 0.2 ml of Solution 3 (Appendix A) was added and the samples were gently

mixed by inversion several times. The tubes were then centrifuged at 14,000 rpm for 1-2 min to pellet cell debris. The supernatants were transferred to fresh microfuge tubes containing 0.5 ml 100% iso-propanol and then mixed by inversion. After centrifugation for 1 min, the supernatants were discarded by decanting and blotting. The tubes were briefly spun again to collect the residual supernatants which were then removed with a pipette. Alternatively, the supernatants were removed by gentle aspiration. The wet pellets were then resuspended in 50 µl of 1 × TE buffer containing 10 mg/ml of RNase A.

Method 3: Sequencing grade plasmid (or cosmid) DNA was prepared using the reagents and protocols provided in the Bresaspin Plasmid Mini Kit (Bresatec), or in the Qiagen Plasmid Mini Kit (Qiagen GmbH, Hilden, Germany).

2.10.3 Large scale plasmid (or cosmid) DNA isolation

Plasmid (or cosmid) DNA was isolated by either of the following procedures.

Method 1: Large scale quantities of plasmid (or cosmid) DNA (2 to 5 µg) was prepared from 300 ml O/N cultures by the three step alkali lysis method and CsCl gradient centrifugation (Garger *et al.,* 1983). The cell pellet, collected by centrifugation in a JA10 rotor using a Beckman J2-21M ultracentrifuge (Beckman Instruments Inc., Palo Alto, California, USA) at 12,000 rpm for 10 min, was resuspended in 2.4 ml of solution 1 (Appendix A) and transferred to a SS-34 tube (Nalgene Labware, Rochester, New York, USA). The suspension was incubated at room temperature for 10 min after the addition of 0.6 ml lysozyme (20 mg/ml in solution 1). Two volumes of solution 2 (Appendix A) was added prior to a 5 min incubation on ice. A further 15 min incubation on ice proceeded the addition of 2.8 ml of solution 3 (Appendix A). Cell debris was removed by centrifugation at 12,000 rpm for 10 min at 4 °C in a JA20 rotor using a Beckman J2-21M ultracentrifuge. DNA was extracted at least once with phenol:chloroform:isoamyl alcohol (25:24:1) in a 20 ml McCartney bottle, and precipitated with 0.7 volume of iso-propanol. DNA was recovered by centrifugation (12,000 rpm for 20 min, Beckman J2-21M ultracentrifuge), washed with 70% (v/v) ethanol, and air dried.

A gradient was prepared after resuspension of the DNA pellet in 2.8 ml of 1 × TE buffer. To the DNA solution, 2.9 g CsCl was added, prior to the addition of 0.3 ml

(10 mg/ml) of EtBr. The solution was adjusted to a refractive index of 1.386 and placed into a 3.9 ml Beckman Quick-Seal ultracentrifuge polyallomer tube (Beckman Instruments Inc.). The tube was sealed and centrifuged in a Beckman Optima^{TM} TLX Ultracentrifuge (Beckman Instruments Inc.) using a TLN-100 rotor at 391,000 rpm for 4 h at 20 °C. The plasmid (or cosmid) band was recovered with a 19 gauge needle attached to a 1 ml syringe. EtBr was extracted 3 to 4 times with an equal volume of iso-amyl alcohol. CsCl was removed by dialysis in 2 L of 1 × TE buffer at 4 °C with at least one change of buffer.

Method 2: Large scale quantities of plasmid (or cosmid) DNA (2 to 5 mg) isolated from 500 ml O/N cultures was prepared by a modified three step alkali lysis procedure. The cell pellet collected by centrifugation in a JA10 rotor using a Beckman J2-21M ultracentrifuge at 12,000 rpm for 10 min, was resuspended in 10 ml of solution 1 containing 1 mg/ml RNase A, and transferred to a SS-34 tube (Nalgene Labware). The suspension was incubated at room temperature for 5 min after the addition of 10 ml solution 2. A further 15 min incubation on ice proceeded the addition of 10 ml of solution 3. Cell debris was carefully removed by centrifugation at 12,000 rpm for 10 min at 4 °C in a JA-20 rotor using a Beckman J2-21M ultracentrifuge. The supernatant was applied to a Qiagen-tip 500 column (Qiagen GmbH, Hilden, Germany) equilibrated with buffer QBT, and allowed to pass through the column by gravity flow. The column was washed twice in 30 ml QC buffer, and the DNA eluted with 15 ml QF buffer. DNA was precipitated with 0.7 volume of iso-propanol, recovered by centrifugation (12,000 rpm for 20 min, Beckman J2-21M ultracentrifuge), washed with 70% (v/v) ethanol, and air dried. The DNA pellet was resuspended in an appropriate volume of sterile H_2O.

2.10.4 DNA preparation for library screening and subclones identification

When PCR was used to screen a cosmid library and subclones, template DNA was prepared using a rapid technique as described by Gussow and Clackson (1989) and Holmes and Quigley (1981). Collected *E. coli* cells were washed 2-3 times in 1 ml Milli Q water, resuspended in 100 µl Milli Q water and followed by incubation in boiling water bath for 7 min. After centrifugation for 10 min at 14,000 rpm, 2 µl of the supernatant was used as a template for PCR.

2.11 DNA analysis and manipulation

2.11.1 DNA quantitation

Double stranded DNA (dsDNA) and oligonucleotides as single stranded DNA (ssDNA) concentrations were determined by measuring their absorbance at 260 nm (using Pharmacia LKB Ultrospec Plus Spectrophotometer) and assuming one OD_{260} unit is equal to 50 µg/ml for dsDNA and 33 µg/ml for ssDNA (Miller, 1972) or by computing densitometry as follows: DNA samples were run on 1.0% TAE agarose gel. The gel was stained in 1 × TAE buffer containing 2 µg/ml EtBr for 20 min, destained in water for 20 min, and photographed using a polaroid 665 film. The amount of DNA was quantified by comparison with a Qiagen purified standard [plasmid pBluescript II SK(+)].

2.11.2 Restriction endonuclease digestion of DNA

Restriction enzyme digestion of DNA was performed using restriction buffers and reaction conditions specified by the manufacturer. Typically, a standard digestion reaction involved an appropriate amount of target DNA (0.1 to 0.5 µg), restriction buffer supplied by the manufacturer, Milli Q water and 2 µl of restriction enzyme (2 units/µl). The reaction mixture was incubated at an appropriate temperature for at least 2 h. The reactions were terminated by heating at the recommended temperature (65 °C for 15 min, or at 85 °C for 20 min), or by phenol:chloroform (1:1) extraction. Prior to loading onto an agarose gel, 0.1 volume of 10 × tracking dye was added.

2.11.3 Agarose gel electrophoresis of DNA

Electrophoresis was carried out at room temperature in horizontal 0.5% to 2.0% (w/v) agarose gels (Saekam HGT). The concentration of agarose depended on the expected size of the DNA fragments to be analysed. Gels were electrophoresed in a Minicell EC370M electrophoresis system (E-C Apparatus Corporation, St. Petersberg, Florida, USA) or a Easy-Cast™ electrophoresis system, Model #B1A (Owl Scientific Inc., Portsmouth, New Hampshire, USA) at a maximum of 120 V for 1 h to 3 h in 1 × TAE buffer, followed by staining in 1 × TAE buffer containing 2 µg/ml EtBr for 10 min, and destained in distilled water for 10 min. DNA fragments were visualised using a UV

transilluminator (λ=254 nm) (UVP Inc., Upland, California USA) and documented with a Tractel Gel Documentation System (Vision Systems, Salisbury, South Australia, Australia) coupled with a Mitsubishi video copy processor, or photographed on Polaroid 665 negative or 667 positive film (Polaroid Corporation, Cambridge, Massachusetts, USA).

2.11.4 Calculation of restriction fragment size

The size of restriction enzyme fragments were determined by comparing their relative mobility on agarose gel with that of *Eco*RI digested *B. subtilis* bacteriophage SPP1 or *Hind* III digested λ phage c1857S7 DNA. The sizes of the SPP1 *Eco*RI digested fragments in kbp are: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.49 and 0.36. The sizes of the λ *Hind* III digested fragments in kbp are: 23.13, 9.416, 6.557, 4.361, 2.322, 2.027, 0.564 and 0.125.

2.11.5 Purification of DNA fragments from agarose gels

DNA fragments were isolated by one of the following procedures. Where required, glycogen was added to assist in the precipitation of DNA by acting as a carrier molecule, especially if the DNA was in low abundance or linearised DNA fragments were less than 1.0 kbp.

Method I 1% (w/v) low melt preparative grade agarose (Bio-Rad) gels were prepared with 1 × TAE buffer and run at 70 V or 45 V in 1 × TAE buffer at room temperature. The part of the gel containing DNA marker and a small amount of the target fragment was cut out and stained in 1 × TAE buffer containing 2 μg/ml EtBr. The appropriate restriction fragment in the left gel which was not EtBr stained was excised aseptically after comparison with its adjacent stained target fragment under UV light. DNA was recovered from the excised agarose gel using Bresaclean DNA extraction kit (Bresatec, Adelaide, Australia) according to manufacturer's recommendations.

Method II 1.0% (w/v) agarose gels (DNA grade, Progen) were prepared with 1 × TAE buffer and used for separation of restriction fragments. Gels were run at 100V in 1 × TAE buffer at room temperature. The part of the gel containing DNA marker and a small amount of the target fragment was cut out and stained in 1 × TAE buffer containing

2 µg/ml EtBr. The appropriate restriction fragment in the remaining unstained gel was excised and then cut into small pieces after comparison with the stained gel counterpart. An equal volume of Tris-HCl (pH 8.0) saturated phenol was added to the agarose gel pieces collected into a 1.5 ml reaction tube. After thorough mixing, the tube was placed at -70 °C for at least 1 h. The sample was then thawed at 37 °C, vortexed and centrifuged in a Heraeus Biofuge 15 benchtop centrifuge at 15,000 rpm for 5 min. The upper aqueous phase was transferred to a new tube and then extracted with equal volume of chloroform. After extraction with chloroform, DNA was precipitated by addition of 0.1 volumes of 3 M NaAc (pH 4.6), and 2.5 volumes of absolute ethanol followed by incubation at -70 °C for at least 30 min. Following centrifugation in a Beckman J2-21M at 18,000 rpm (4 °C) for 30 min, the supernatant was removed. Pelleted DNA was then washed with 70% ethanol, dried *in vacuo* and resuspended in 20 µl Milli Q.

2.11.6 Dephosphorylation of DNA fragments

DNA fragments were dephosphorylated using shrimp alkaline phosphatase (Boehringer Mannheim). Two units of shrimp alkaline phosphatase was added directly to digested DNA together with dephosphorylation buffer and incubated at 37 °C for 10 min. The reaction was terminated by heating at 65 °C for 10 min. The linearised and dephosphorylated DNA fragment was then gel purified.

2.11.7 Synthesis of oligonucleotides

The sequences of all oligonucleotides used in this study are shown in Table 2.3. All of these oligonucleotides were prepared in the non tritylated form. Purification of these oligonucleotides from 30% ammonium hydroxide solution was performed as described by Sawadogo and Van Dyke (1991). The purified oligonucleotides were stored at -20 °C.

2.11.8 Polymerase chain reaction (PCR)

DNA amplification using PCR was performed using a modification of the procedures of Li *et al.* (1988) and Sambrook *et al.* (1989).

PCR products to be used for cloning and sequencing were prepared in a total reaction volume of 50 µl [5 µl 10 × PCR buffer ($MgCl_2$ 2.5 mM), 0.2 mM of each dNTP,

Table 2.3. Oligonucleotides used for sequencing and PCR in this study.

Oligo	Sequence	Locations in the *pha* operon of *Pseudomonas* strain 10c-1-3*
M13 (Forward)	5' GTA AAA CGA CGG CCA GT 3'	
M13 (Reverse)	5' GAC CAT GAT TAC GCC AAG 3'	
T7	5' TAA TAC GAC TCA CTA TAG GG 3'	
#3512	5' CGT GTC GCC AGG CTA TAT CG 3'	5055
#3513	5' TTC TCG ACC GGT GCG GAC TT 3'	5206
#3578	5' GCATGC TCT CGA CCG GTG CGG ACT 3'	5207**
#3579 (Reverse)	5' GTCGAC GAG ACG CCT GAG AAC GAT 3'	5981***
#3623	5' CTGCAG CCG GAT GAA TGT CAG CTA 3'	in Kan^r cartridge****
#3624 (Reverse)	5' CTGCAG CCG CTC AGA AGA ACT CGT 3'	in Kan^r cartridge****
#3722	5' GAT ATT GCT GAA GAG CTT GG 3'	in Kan^r cartridge****
#4033	5' AG)(GCT)C AT(GC) AAC (AC)AG GC)T(GCT) TGC GG 3'	3494
#4034	5'(GC)(AT)(GC)AG(GC)CCGTC(GC)(AT) (GCT)GA(AT)CAT 3'	3693
#4035	5' (AG)TC AAC (AT)(AG)(CT)GCCGG(CT) ATC AC 3'	4771
#4036	5' AT(GC) (GT)C(GC) GT(GC) (GCT)CG A(CT)A TAG CC 3'	5065
#5017	5' GGC ATC AAC CAG GTG 3'	3494
#5018	5' GCC GTC GTG GAT CAT 3'	3688
#5384 (Reverse)	5' GCT AAC TTG TCC TGA TCC TC 3'	3782
#5404	5' GAT GAT CTG GCT CGC CGA CT 3'	526
#5405	5' GGT CGA ACG TCC GTA AGA AC 3'	4414
#5426(Reverse)	5' CGA TAC GAT CAC TAC GTC AG 3'	3271
#5453	5' CAG ATA CCG CCG ACA GCA CA 3'	1190
#5454	5' GTT ATA CAT CCT CGA CCT GC 3'	2130
#5455	5' GTT CGG CCA GAC CAA CTA CT 3'	4956
#5456 (Reverse)	5' CAG ATA CGT GTT GCG CAG GT 3'	2721
#5457 (Reverse)	5' TGC TTG AGC AGG TCG ATC AC 3'	1640
#5486	5' GAC GCG AAG TCC AAC CAC AA 3'	5539

Oligo	Sequence	Locations in the *pha* operon of *Pseudomonas* strain 10c-1-3*
#5576	5' CGT GCT GGA TCA GGA AGA CT 3'	5508
#5577	5' TCG AAG GAA GGC ACG GTG AC 3'	3956
#5578 (Reverse)	5' ATC TGC TTG TGC ACC GCG AG 3'	4239
#5608	5' GAA GTC TGG TCG CAG TTC AT 3'	5782
#5671 (Reverse)	5' CGC AGT GTG AGA CAA GAG AT 3'	6459
#5672a (Reverse)	5' TTC GCG AGA CGC CTG AGA AC 3'	5986
#5672b (Reverse)	5' GCG TGA TGT AGG TGC TCG TT 3'	5488
#5673 (Reverse)	5' TGT CCG TGC CGA TAT AGC CT 3'	5083
#5674 (Reverse)	5' ACC GTA TAG CCG TCC TTA TG 3'	4604
#5675 (Reverse)	5' CCA TCA CCT TCG GAT CGA CG 3'	4115
#5676 (Reverse)	5' CTT CGG CCT TGT TCT GCG AG 3'	3809
#5677 (Reverse)	5' AGG ATC ACT TCG CTC ACC TG 3'	3405
#5678 (Reverse)	5' ATC CAG AAG CTG CGC TTG TT 3'	2957
#5679 (Reverse)	5' GAT CGT CTG CTC GCG CAT CT 3'	2505
#5680 (Reverse)	5' CAG CAG GTC GTT CTC GAA CA 3'	2052
#5681 (Reverse)	5' TGC TTG AGC AGG TCG ATC AC 3'	1640
#5682 (Reverse)	5' AGC CAT TGC TCG AAC ACT TG 3'	1267
#5683 (Reverse)	5' GGA TTC CGG ATG CGA TCA CC 3'	770
#5684 (Reverse)	5' CGC ATA GAT GTC CGA GAA GT 3'	394

* See Figure 5.7a;
** *Sph* I restriction site was introduced;
*** *Sal* I restriction site was introduced;
**** See Appendix D, a *Pst* I restriction site was introduced into #3623 and #3624.

0.2-1.0 μM of each primer, approximately 100 ng experimental DNA template and 1.25 units of TaqBead™ (Promega)].

When PCR was used as a method to screen a cosmid library, a total PCR reaction volume of 25 μl contained 2.5 μl 10 × PCR buffer, 4% (v/v) of formamide (Sarkar *et al.*, 1990), 0.2 mM of each dNTP, 0.2-1.0 μM of each primer, #4033 and #4034, 2 μl cell lysate (see following description) as DNA template and 1 unit *Taq* polymerase.

The reaction mixture was overlaid with mineral oil then subjected to thermal cycling in either a Corbett Research FTS-1 or a Corbett Research FTS-320 thermal cycler (Corbett Research, North Ryde, New South Wales, Australia).

The amplification cycle comprised of an initial 3 min denaturation at 95 °C, followed by 30 or 25 cycles of 95 °C for 30 s, an appropriate annealing temperature [usually around (Tm-5 °C) of each primer] for 1 min, and 72 °C for an appropriate period of time determined by the length of the amplified DNA fragment at the assumed extension rate of 0.5 kbp/min, followed by a final extension at 72 °C for 4 min. When primer pair #4033 and #4034 were used for cosmid library screening and subclones identification, all the conditions were the same except that the annealing temperature was 56 °C and the extension time at 72 °C was 30 s.

2.11.9 *In vitro* cloning

2.11.9.1 Cloning of restriction digest fragments

DNA fragments to be cloned and the vector DNA were similarly cleaved in either single or double restriction enzyme digests (see Section 2.11.2). Where dephosphorylated vector and insert DNA had compatible cohesive termini, a standard ligation reaction was performed as follows. Vector (50 ng) and insert were mixed in a reaction tube at a molar ratio of 1:3. The tube was heated at 45 °C for 5 min to melt their cohesive termini, cooled on ice for 1 min, combined together with 2 μl 10 × ligation buffer, 2 units of T4 DNA ligase (Boehringer Mannheim), and made up to a final volume of 20 μl with Milli Q water. The ligation reaction was incubated O/N at 4 °C. For blunt end ligations, 4 units of T4 DNA ligase was used in a reaction. After ligation, T4 DNA ligase was inactivated by

heating at 65 °C for 10 min or by phenol:chloroform extraction. The ligated DNA was precipitated by addition of absolute ethanol, washed with 70% ethanol, dried *in vacuo* and resuspended in 10 µl Milli Q water.

For a simple and quick cleanup, the ligation reaction was *n*-butanol purified using the method of Thomas (1994) as follows. The reaction was added with sterile Milli Q water to a final volume of 50 µl, mixed with 500 µl of *n*-butanol by vortexing for 5 s, and centrifuged at 14,000 rpm for 10 min. The DNA pellet was dried *in vacuo* and resuspended in 10 µl Milli Q water.

When chromosomal DNA fragments were used for construction of a cosmid library, the molar ratio of vector, pHC79 (0.5 µg) to insert DNA (1 µg) was changed to 3:1 and the total amount of DNA used in ligation reactions ranged from 0.5 µg to 5 µg, as specified by the manufacturer of the cosmid packaging system (Promega, USA).

2.11.9.2 Cloning of purified PCR product

PCR product was purified as described in Section 2.11.5 and added to 50 ng of pGEM-T (Promega) vector at a 3:1 molar ratio of PCR product:vector. The ligation reaction was carried out using the reagents and protocol as described in *the pGEM-T Vector System Technical Manual* (Promega, 1997). Before electrotransformation of *E. coli*, the ligation reaction was terminated and DNA was precipitated, dried *in vacuo* and resuspended in 10 µl Milli Q water. Transformants were selected on NA containing ampicillin and X-gal. While, *lacZ* colonies were selected for further examination.

2.11.10 DNA sequence analysis

2.11.10.1 DNA sequencing

DNA sequencing using dye-labelled primers and/ or terminators was carried out using the PRISM Ready Reaction Cycle Sequencing kit (Applied Biosystems).

2.11.10.2 Computer analysis of DNA and protein sequences

DNA and protein sequences were analysed using various WWW based molecular analysis software programs available via the Department of Microbiology and Immunology WWW page (*http://www.microbiology.adelaide.edu.au/*).

pha related DNA and protein sequences were retrieved from Genbank, Genbank Updates, SWISS-PROT and PIR data banks. Alignment of DNA and amino acid sequences was carried using CLUSTLW (Thompson *et al.*, 1994).

Primers for PCR and DNA sequencing were designed using Primer Designer (Version 2.0, Scientific and Educational Software). DNA sequencing data were corrected with the aid of Chromas (version 1.44, *http://www.technelysium.com.au/chromas.html*). Translation of DNA sequences were carried out by using DNAsis (Hitachi Software Engineering Co., Ltd., 1990).

Searches for homology to known DNA and protein sequences contained in Genbank, Genbank Updates, SWISS-PROT and PIR data banks were via BLAST2 search service at EMBL (*http://dove.embl-heidelberg.de/Blast2/*) (Altschul *et al.*, 1990). Searches for open reading frames and the identification of a promoter in a DNA sequence were carried out using the programs at *http://www.ncbi.nlm.nih.gov/gorf/gorf.html* and *http://www.fruitfly.org/seq_tools/promoter.html* respectively. Restriction endonuclease sites within a DNA sequence were determined using the WWW based program Webcutter 2.0 (*http://www.ccsi.com/firstmarket/cutter/cut2.html*) (Heiman, 1997).

2.12 Transformation

2.12.1 Electroporation

E. coli DH5α cells were made competent for transformation using the following procedures. 300 ml NB was inoculated with 5 ml of an overnight stationary phase culture and incubated at 37 °C for 2.5 h in a shaken 5 litre conical flask to an OD_{600} of approximately 0.4. The cells were chilled on ice for 30 min, and centrifuged at 7,400 rpm in a JA14 rotor using a Beckman J2-21M ultracentrifuge at 4 °C for 10 min. The pellet was consequently washed in 100 ml and 50 ml ice cold Milli Q water, followed by a 20 ml

ice cold 10% (v/v) glycerol wash. The cells were resuspended in 2 ml ice cold 10% (v/v) glycerol and aliquots of 100 µl were snap frozen in a dry ice/ethanol bath and stored at -70 °C.

Electrocompetent *E. coli* was thawed on ice prior to the addition of plasmid DNA contained in a maximum volume of 10 µl Milli Q water. A BioRad *E. coli* Pulser and 0.2 cm electrode gap cuvettes (BioRad) were used as described in the BioRad Electroporation Manual to transform *E. coli* electrocompetent cells with plasmid DNA. Electrocompetent *E. coli* was electroporated using a field strength of 2.0 kilovolts (kv). Immediately following electroporation, 1 ml SOC (Appendix B) medium was added to the cells. The cells were then incubated at 37 °C for 1 h to allow for plasmid expression and then plated onto selective media.

2.12.2 Chemical transformation

Suspensions of competent cells were prepared by a modification of the method described by Hanahan (1983). Briefly, 100 µl of an O/N culture of *E. coli* DH5α was subcultured into 2.0 ml of LB medium, and grown with agitation at 37 °C for 2-4 h. The cells were collected by centrifugation in Eppendorf 5417R microcentrifuge at 10,000 rpm for 30 sec, and then resuspended in 200 µl of 75 mM $CaCl_2$.

After cells were chilled on ice for 2 h, DNA was added in a maximum volume of 20 µl and incubated on ice for 30 min. The mixture was shocked at 42 °C for 2 min, chilled on ice again for 2 min and followed by addition of 1 ml LB medium. Then the cell suspension was incubated at 37 °C for 45 min. Transformants were recovered by plating 200 µl culture onto NA plates (containing appropriate antibiotics).

2.13 Construction of a cosmid library

2.13.1 DNA preparation and *in vitro* packaging

Chromosomal DNA of the PHA producing isolate, *Pseudomonas* strain 10c-1-3 was prepared as described in Section 2.10.1 with a size of larger than 100 kbp as required for cosmid library construction. Genomic DNA was partially digested with *Mbo* I (in a final volume of 750 µl) using the method described by Frischauf (1987). *Mbo* I was

inactivated by addition of 15 µl of 0.5 M EDTA (pH 8.0) and DNA was extracted with phenol:chloroform followed by precipitation with ethanol. Purified DNA was pelleted by centrifugation at 15,000 rpm for 15 min at 4 °C, dried *in vacuo* and dissolved in 100 µl of Milli Q water O/N at 4 °C. Cosmid vector pHC79 was completely digested with *Bam* HI and dephosphorylated as described in Sections 2.11.2 and 2.11.6. The vector and insert were ligated as described in Section 2.11.9.1.

The *Packagene® Lambda DNA Packaging System* was used to package DNA in the ligation reaction mixture according to instructions supplied by the manufacturer (Promega, 1997).

2.13.2 Preparation of *E. coli* culture for infection

10 ml of Terrific Broth (TB) medium supplemented with 0.2% maltose and 10 mM $MgCl_2$ was inoculated with a single colony of *E. coli* DH5α and incubated overnight at 37 °C with agitation. 100 µl of the O/N culture was subcultured into 10 ml of LB medium supplemented with 0.2% maltose and 10 mM $MgCl_2$ and incubated at 37 °C for approximately 2.5 h, or until the OD_{600} reached 0.6-0.8.

2.13.3 Transduction of *E. coli*

10 µl of the bacteriophage suspension prepared in Section 2.13.1 was mixed with 0.1 ml of SM buffer and 0.2 ml of the fresh *E. coli* DH5α culture prepared in Section 2.13.2. The bacteriophage particles were allowed to adsorb for 20 min at 37 °C. Then 1 ml of TB was added and the incubation *continued* for a further 45 min at 37 °C for bacteriophage expression. After incubation, 0.2 ml of the bacterial culture was plated onto NA plates (Ap, 100 µg/ml) followed by incubation O/N at 37 °C.

2.14 Bacterial conjugation

Broth cultures were grown to early exponential phase with slow agitation to avoid damage to sex pili of donor. The cells were washed in NB and resuspended in 10 ml NB. Donor and recipient bacteria were mixed at a ratio of 1:10 and the cells pelleted by centrifugation at 5,000 rpm for 5 min. The pellet was gently resuspended in 300 µl of NB

and spread onto a nitrocellulose membrane filter (0.45 μm, type HA, Millipore Corporation) on a NA plate. This plate was incubated O/N at 30 °C. The cells were then resuspended in 1.0 ml of NB and samples plated onto selective growth medium and incubated overnight at 30 °C.

2.15 Southern hybridisation

Southern hybridisation was generally carried out following the procedures described in *The Dig System User's Guide for Filter Hybridisation* (Boehringer Mannheim).

2.15.1 DNA probe preparation

2.15.1.1 Random primed Digoxigenin labelling of DNA

Denatured *Pst* I digested plasmid p4A fragments and SPP1 fragments were labelled by random primed incorporation of Digoxigenin-11-dUTP using the components of the Digoxigenin DNA labelling kit and the method described in *The Dig System User's Guide for Filter Hybridisation* (Boehringer Mannheim). The DNA probe was then added to prehybridisation solution to form the hybridisation solution.

2.15.1.2 Digoxigenin labelling of PCR product

PCR product was labelled by PCR mediated incorporation of Dig-11-dUTP using the component of the PCR Dig probe synthesis kit and following the procedure described in *The Dig System User's Guide for Filter Hybridisation* (Boehringer Mannheim). The DNA template used was plasmid pSGZ1, and the primers used were #4033 and #4034. Cycling conditions were as same as that described in Section 2.11.8. After the concentration optimisation of the labelled PCR product with different dilutions, the DNA probe was then accordingly added to prehybridisation solution with an optimal concentration to form hybridisation solution.

2.15.2 Southern hybridisation analysis

2.15.2.1 Southern transfer

Unidirectional transfers of DNA from 1.0% agarose gels to nylon (Hybond N+, Amersham International) were performed as described by Southern (1975) and modified by Sambrook *et al.* (1989), using capillary transfer. DNA was irreversibly bound to the filter after UV-crosslinking (254 nm, 2 min).

2.15.2.2 Hybridisation

UV-treated nylon membranes were incubated in a prehybridisation solution for a minimum of 2 h at 42 °C. The DNA probes diluted in prehybridisation solution were denatured by heat at 68 °C for 10 min in water bath and then incubated with membranes O/N at 42 °C in a Extron HI 2001 hybridisation oven (Bartelt Instruments Pty. Ltd., Heildelberg West, Victoria, Australia). Membranes were washed 2 × 10 min in 2 × SSC (60 mM NaCl, 60 mM sodium citrate, pH 7.0) and 0.1% (w/v) SDS at room temperature, followed by 2 × 20 min washes with 0.1 × SSC and 0.1% (w/v) SDS at 68 °C. Nylon membranes were directly detected or dried in air for later detection.

2.15.2.3 Detection

Digoxigenin-labelled probe DNA was detected colorimetrically as recommended by the manufacturer (Boehringer Mannheim)

2.16 RNA isolation from bacteria

Total cellular RNA from *R. eutropha*, *Pseudomonas* strain 10c-1-3 and *E. coli* was isolated using the hot-phenol method of Aiba *et al.* (1981). Logarithmic phase cultures (10 ml) were centrifuged in 20 ml McCartney bottles in an MSE Minor S centrifuge at 3,100 rpm for 10 min, stored as a pellet at -70 °C. Cells were resuspended in 700 µl lysis buffer (20 mM sodium acetate, 1 mM EDTA, 0.5% SDS, pH 5.5), and immediately transferred to a 1.5 ml Eppendorf tube containing 500 µl phenol [equilibrated with 20 mM sodium acetate, 1 mM EDTA (pH 5.5)] at 65 °C. After vortexing to mix the phases and incubating at 65 °C for 5 min in a dry block heater (Ratek Instruments, Boronia, Victoria,

Australia), the samples were centrifuged for 3 min at 14,000 rpm in Eppendorf centrifuge 5417R. The upper aqueous layer was removed and transferred into a fresh Eppendorf tube containing 500 µl of hot phenol and 200 µl of lysis buffer. This extraction was repeated tree times. RNA was precipitated by addition of 0.1 volumes of 3 M sodium acetate (pH 5.2), and 2.5 volumes of absolute ethanol followed by incubation at -70 °C for at least 2 h. RNA was collected by centrifugation (14,000 rpm, 20 min, at 4 °C). Samples were resuspended in 50 µl of RNA suspension buffer [20 mM Tris-HCl (pH 7.6), 5 mM $MgCl_2$] and contaminating DNA in the samples was removed by digestion with 1 µl RNase-free DNase I at 37 °C for 30 min. Then samples were purified with phenol:chloroform (1:1). RNA was re-precipitated as above and resuspended in 200 µl DEPC-treated Milli Q water. The integrity of extracted RNA and efficiency of DNA removal were examined by gel electrophoresis.

2.17 RNA analysis

2.17.1 RNA quantitation

The concentration of total RNA in solution was determined by spectrophotometric measurement of absorption at 260 nm, assuming an OD_{260} of 1.0 is equal to 40 µg RNA/ml (Miller, 1972). Samples with OD_{260}/OD_{280} ratios between 1.8 and 2.0 were accepted for use in hybridisation experiments. Samples with ratios below 1.8 were re-extracted with phenol:chloroform (1:1).

2.17.2 RNA slot blot and hybridisation

A suitably sized Hybond N^+ nylon membrane, pre-soaked in 10 × SSC, was assembled in a slot blot apparatus (Hoefer Scientific Instruments, San Francisco, California, USA). RNA samples in a final volume of 100 µl, were denatured by addition of 300 µl of a solution containing 6.15 M formaldehyde in 10 × SSC, incubated for 10 min at 65 °C and stored on ice until loading. Samples (200 µl per well) were applied to the membrane under vacuum. Wells were washed once by addition of 100 µl of 10 × SSC, and RNA was fixed onto the membrane at 80 °C for 2 h in a vacuum oven (Thermoline Scientific Equipment Pty. Ltd., Wetherill Park, New South Wales, Australia) prior to O/N

hybridisation (see Section 2.15.2.2). Target RNA was detected using the method described in Section 2.15.2.3.

2.18 Expression of *pha* genes in *E. coli*

2.18.1 Construction of *E. coli* clones for over-expression of *pha* genes

Plasmid pCT401, harbouring a 4.3 kbp *Apa* I *Pseudomonas* strain 10c-1-3 chromosomal DNA fragment (including *phaC,* and *partial phaA*) cloned into pGEM7Zf(+) (see Section 5.2.1), plasmid pCT408, harbouring a 4.8 kbp *Apa* I/*Eco* RV *Pseudomonas* strain 10c-1-3 chromosomal DNA fragment (including partial *phaA*, *phaB*, and ORF4) cloned into pBluescript II SK(+) (see Section 5.2.4), plasmid pCT409 harbouring a 1.5 kbp *Apa* I/*Pst* I *Pseudomonas* strain 10c-1-3 chromosomal DNA fragment (including partial *phaA*, *phaB*, and partial ORF4) cloned into pBluescript II SK(+), plasmid pCT410 harbouring a 3.6 kbp *Apa* I/*Sac* I *Pseudomonas* strain 10c-1-3 chromosomal DNA fragment (including partial *phaA*, *phaB*, and ORF4) cloned into pBluescript II SK(+), plasmid pCT411 harbouring a 6.1 kbp *Pseudomonas* strain 10c-1-3 chromosomal DNA fragment (including *phaCAB*ORF4) cloned into pBluescript II SK(+) (see Section 5.2.4), and pCT419 harbouring ORF4 (see Section 6.2.1.1) were used to overexpress the PHA biosynthetic genes and ORF4 in *E. coli*. The expression of the *pha* genes in these plasmids is under the control of a vector-borne T7 promoter. Plasmid pGP1-2 contains the gene for T7 RNA polymerase under the control of λ P_L promoter that is repressed by a temperature sensitive repressor, and p15A origin of replication which is compatible with the ColEl origin of the replication located on the expression vectors, pGEM5Zf(+), pGEM7Zf(+) and pBluescript II SK(+) (Tabor and Richardson, 1985). Plasmid pGP1-2 and each of above plasmids were sequentially transformed into *E. coli* DH5α to produce *E. coli* strains, harbouring (pCT401 and pGP1-2), (pCT408 and pGP1-2), (pCT409 and pGP1-2), (pCT410 and pGP1-2), (pCT411 and pGP1-2), and (pCT419 and pGP1-2) respectively. *E. coli* DH5α harbouring (pHB1 and pGP1-2) was used as a positive control and *E. coli* DH5α harbouring [pBluescript II SK(+) and pGP1-2] a negative control.

2.18.2 Over-expression of *pha* genes

The over-expression of *pha* genes under the control of the temperature inducible T7 RNA polymerase promoter was performed using a modification of the method described by Tabor and Richardson (1985).

E. coli DH5α strains harbouring (pCT401 and pGP1-2), (pCT408 and pGP1-2), (pCT409 and pGP1-2), (pCT410 and pGP1-2), (pCT411 and pGP1-2), (pCT419 and pGP1-2), (pHB1 and pGP1-2), and [pBluescript II SK(+) and pGP1-2] were streaked onto NA plates containing Kan (50 μg/ml) and Ap (100 μg/ml) and incubated at 30 °C, O/N. A single colony was used to inoculate 5 ml of 2×YT broth containing Kan (50 μg/ml) and Ap (200 μg/ml), and incubated at 30 °C with agitation to a density of OD_{600} 0.8. T7 RNA polymerase was then induced by raising the temperature to 42 °C for 20 min with agitation. Rp (200 μg/ml) was added to inhibit the transcription by *E. coli* RNA polymerase. The cultures were further cultivated at 37 °C O/N and samples were taken at 0, 3, and 24 h for SDS-PAGE analysis. An O/N culture of *Pseudomonas* strain 10c-1-3, which was incubated in NB (containing 2.0% glucose) and had accumulated PHA was also prepared for SDS-PAGE analysis.

2.18.3 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS PAGE was a modification of the method of Lugtenberg *et al.* (1975), and was performed as described by Achtman *et al.* (1978). Culture samples were suspended in SDS lysing buffer and heated to 100 °C for 3 min prior to loading. 3 μl of each sample was loaded on an acrylamide gel slabbed on a mini gel apparatus (Hoefer Scientific). Proteins were separated by electrophoresis at 60 V through the stacking gel (5% acrylamide gel) and 150 V through the separating gel (12% acrylamide gel). Proteins were stained by a O/N incubation at room temperature in a Coomassie stain solution with agitation and the gel was then destained in destain solution with several changes. Low molecular weight marker proteins (Pharmacia) used were Phosphorylase B (94 kDa), Bovine serum albumin (67 kDa), Ovalbumin (43 kDa), Carbonic anhydrase (30 kDa), Soybean trypsin inhibitor (20.1 kDa), and a-Lactalbumin (14.4 kDa).

Chapter 3

Isolation of PHA producing bacteria

3.1 Introduction and strategy

By the end of the 1980's, 300 different bacteria had been reported to synthesise various PHAs (Steinbuchel, 1991), and about 91 different monomers had been identified as the constituents of PHA in various bacteria by 1995 (Steinbuchel and Valentin, 1995). Due to the superior physical properties of P(3HB) copolymers or heteropolymers, these are of greater interest to industry. Of the polymers identified, the most prominent was P(3HB). Analysis of PHA biosynthesis pathways involved indicated that production of novel PHA was limited by availability of chemicals supplied as precursor substrates to the bacteria. From substrate to PHA, two key steps are involved: conversion of substrates to intermediary precursors, and recognition of these precursors by PHA synthetic enzymes, the latter depending on the substrate specificity of the PHA enzymes concerned. Consequently one method to obtain novel PHA is to isolate novel bacteria with metabolic pathways suitable for novel PHA production. Novel *pha* genes could then be cloned using standard DNA recombination technology. With this in mind, the primary aim of this chapter was the isolation and identification of some PHA producing bacteria from Adelaide soils. In later chapters, the cloning and characterisation of the PHA biosynthetic genes from one PHA producer is described.

Since PHA is a carbon and energy storage for bacterial cells, biosynthesis and composition depend on supplies of carbon source within micro-environments of the soil habitat. Thus different soil habitats may support growth of bacteria capable of producing different PHAs. For these reasons, a variety of soils were sampled from the Adelaide district.

Although fungi are common in soil, so far no fungi have been identified for PHA production. Nevertheless their growth in a culture with bacteria can result in difficulties for the isolation for PHA producing bacteria. To circumvent this problem, anti-fungal antibiotics, Amphotericin B, and Cycloheximide (actidione) were added to media and used to isolate PHA producing bacteria (see Section 2.4). Crystal violet and Nile Blue A stains were primarily used to identify presence of PHA granules in different isolates. The presence and relative yield of PHA was confirmed by gas chromatography (GC) while differential scanning calorimetry (DSC) was used to evaluate some physical properties of

the polymer. Transmission electron microscopy was used to determine the distribution of PHA granules within producer cells.

3.2 Results

3.2.1 Isolation and identification of PHA producing bacteria

3.2.1.1 Isolation of PHA producing bacteria from soil

Soil samples (1 g each) were suspended and diluted in saline. Each dilution (200 µl) was spread onto P(3HB) medium [containing glucose (20 g/L), Amphotericin B (15 µg/ml), and Cycloheximide (actidione, 70 µg/ml)] agar. After incubation at 30 °C for 2 days, single colonies were isolated. Cells from each colony were used to prepare smears on microscope slides. Simple stains of smears were prepared by treatment with a Crystal Violet solution and cells examined for presence of unstained intracellular granules. The presence of PHA granules was confirmed by staining smears with Nile Blue A. Nineteen PHA producing isolates were identified from fifteen soil samples (Table 3.1).

3.2.1.2 Identification of PHA producing isolates

All 19 isolates were characterised where possible to genus level using standard taxonomic tests (Figure 2.1) (Table 3.2). Using the identification key outlined by Sirockin and Cullimore (1969), isolate 10c-1-3 was tentatively identified as *Pseudomonas* strain 10c-1-3 and the other isolates were tentatively identified as strains of Genera *Methanomonas*, *Acetobacter*, *Pseudomonas*, *Xanthomonas*, or *Aeromonas*.

3.2.2 Gas chromatography (GC) analysis of PHA

To confirm PHA production, the 19 isolates described above were cultured in 500 ml flasks of R-medium (+2.0% Glucose) at 30 °C for 2 days, and PHA extracted from cells using procedures outlined in Section 2.7.5.1. These extracts were then examined by gas chromatography. *E. coli* MD9101 (p4A) was used as a positive control. The GC results are shown in Figures 3.1 - 3.6. Extracts from all 19 isolates contained components with a retention time similar to that of pure P(3HB), or the P(3HB) produced by *E. coli*

Table 3.1. PHA producing isolates from soil samples.

No.	Site of collection	Soil description	Isolate
1	Happy Valley	straw compost	1c-1; 1c-1-1; M1c-1(1); K1b-1; K1c-1
2	Happy Valley	aged mushroom compost	M2; 2-3-2
3	Adelaide University campus	tree litter	3
4	Adelaide University campus	park land soil	M4-1; M4-2
5	Adelaide University campus	park land soil	Nil
6	Adelaide University campus	tree litter	6c-2
7	Adelaide University campus	tree litter	Nil
8	Torrens River	river bank soil	Nil
9	Torrens River	river bank soil	Nil
10	Adelaide University campus	compost	M10c-1; M10c-2; 10c-1-1; 10c-1-3
11	Torrens River	river bank soil	Nil
12	Adelaide University campus	compost	Nil
13	Adelaide University campus	tree litter	Nil
14	Adelaide University campus	park land soil	14b-1; M14c-1; 14c-1-2; 14c-1-3
15	Adelaide University campus	tree litter	Nil

Table 3.2. Characterisations of 19 soil isolates.

Isolate	Morphology	colony pattern	Pigmentation	Gram Stain	O/F*	Moti -lity	Oxidase	Cata-lase**	Ethanol utility (growth/acid/gas)	Endo-spore	Tentative Genus
1c-1	rod, filamentous	irregular, convex-concave	light yellow	-	OF	+	+	+	N/A***	-	*Methanomonas* *Acetobacter* *Pseudomonas* *Xanthomonas* *Aeromonas*
1c-1-1	rod, filamentous	irregular, convex-concave	light yellow	-	OF	+	-	+	+/-/-	-	
2-3-2	rod	circular, convex	light yellow	-	O	+	+	-	+/-/-	-	
3	rod	circular, convex	water soluble yellow-green fluorescent pigment	-	O	+	+	+	N/A***	-	
6c-2	rod, filamentous	irregular, convex	light yellow	-	OF	+	-	+	N/A***	-	
10c-1-1	rod, filamentous	irregular, convex	light yellow	-	OF	+	-	+	+/-/-	-	
10c-1-3	rod	circular, convex	water soluble yellow-green fluorescent pigment	-	O	+	+	+	+/-/-	-	*Pseudomonas*
14b-1	rod	irregular, flat	light yellow	+/-	OF	+	-	+	+/-/-	-	*Methanomonas* *Acetobacter* *Pseudomonas* *Xanthomonas* *Aeromonas*
14c-1-2	rod, filamentous	irregular, convex	light yellow	-	OF	+	-	+	+/-/-	-	
14c-1-3	rod	irregular, convex	light yellow	-	O	+	+	+	NnA***	-	
K1b-1	rod, filamentous	irregular, flat	light yellow	-	OF	+	-	+	N/A***	-	
K1c-1	rod, filamentous	irregular, convex	light yellow	-	OF	+	-	+	N/A***	-	
M1c-1(1)	rod	irregular, convex-concave	light yellow	-	OF	+	-	+		-	
M2	rod	irregular, convex	light yellow	-	OF	+	-	+	+/-/-	-	
M4-1	rod	circular, convex	light yellow	-	O	+	+	+	N/A***	-	
M4-2	rod, filamentous	irregular, convex	light yellow	-	OF	+	-	+	N/A***	-	
M10c-2	rod	irregular, convex	light yellow	+/-	OF	+	-	+	N/A***	-	
M10c-1	rod	circular, convex	light yellow	-	O	+	+	+	N/A***	-	
M14c-1	rod, filamentous	irregular, convex	light yellow	-	OF	+	-	+	N/A***		

* O/F: Oxidation/Fermentation of glucose;
** +: immediately bubbling;
*** N/A: not applicable.

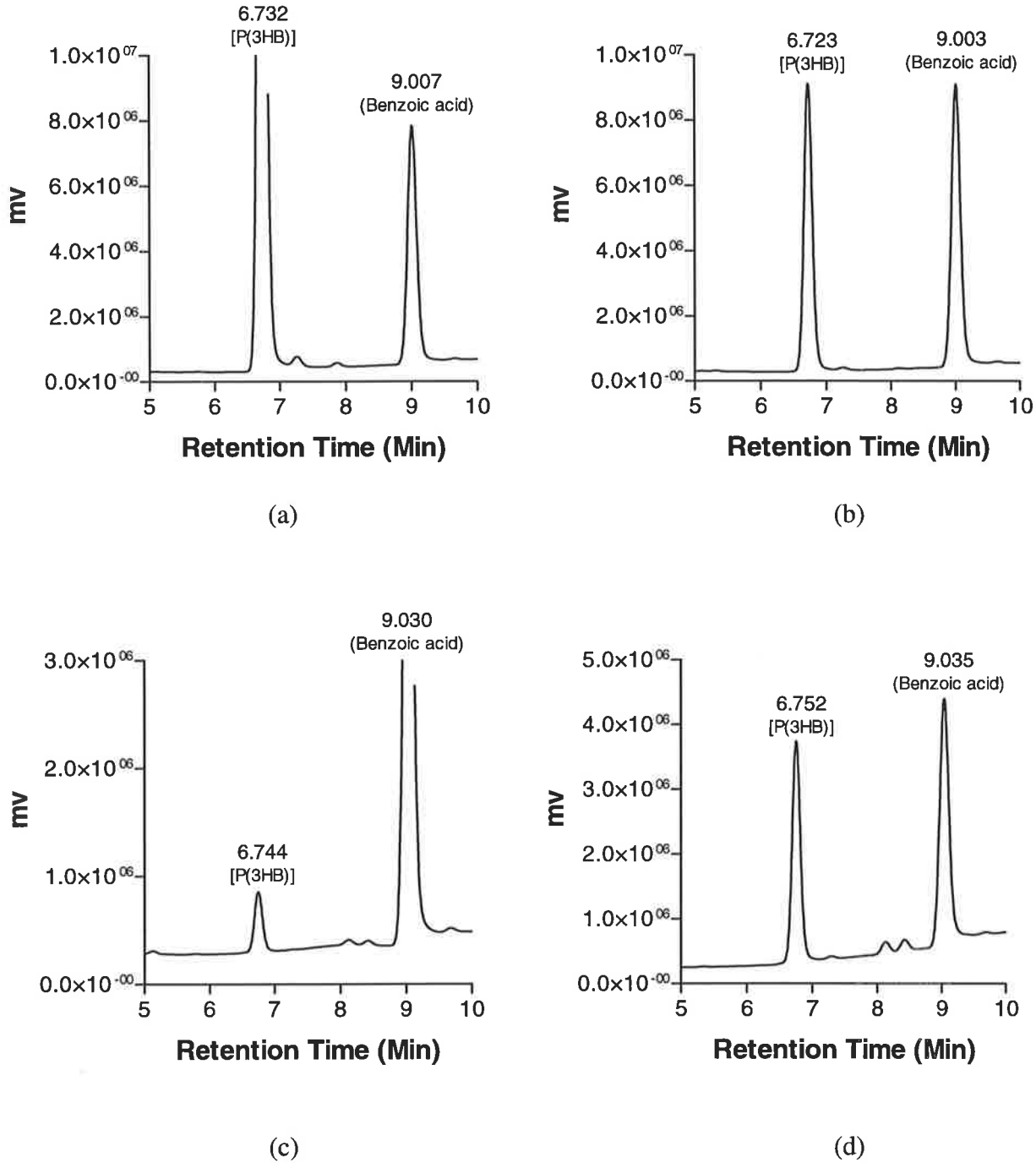


Figure 3.1. Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.

a. pure P(3HB); b. P(3HB) produced by *E. coli* MD9101 (p4A); c. isolate 1C-1; d. isolate 1C-1-1. Note that chromatograms prepared for isolates 1C-1 and 1C-1-1 show peaks with a retention time similar to that of pure P(3HB) sample or P(3HB) produced by *E. coli* MD9101 (p4A), indicating P(3HB) production by these two isolates. Benzoic acid was used as an internal control.

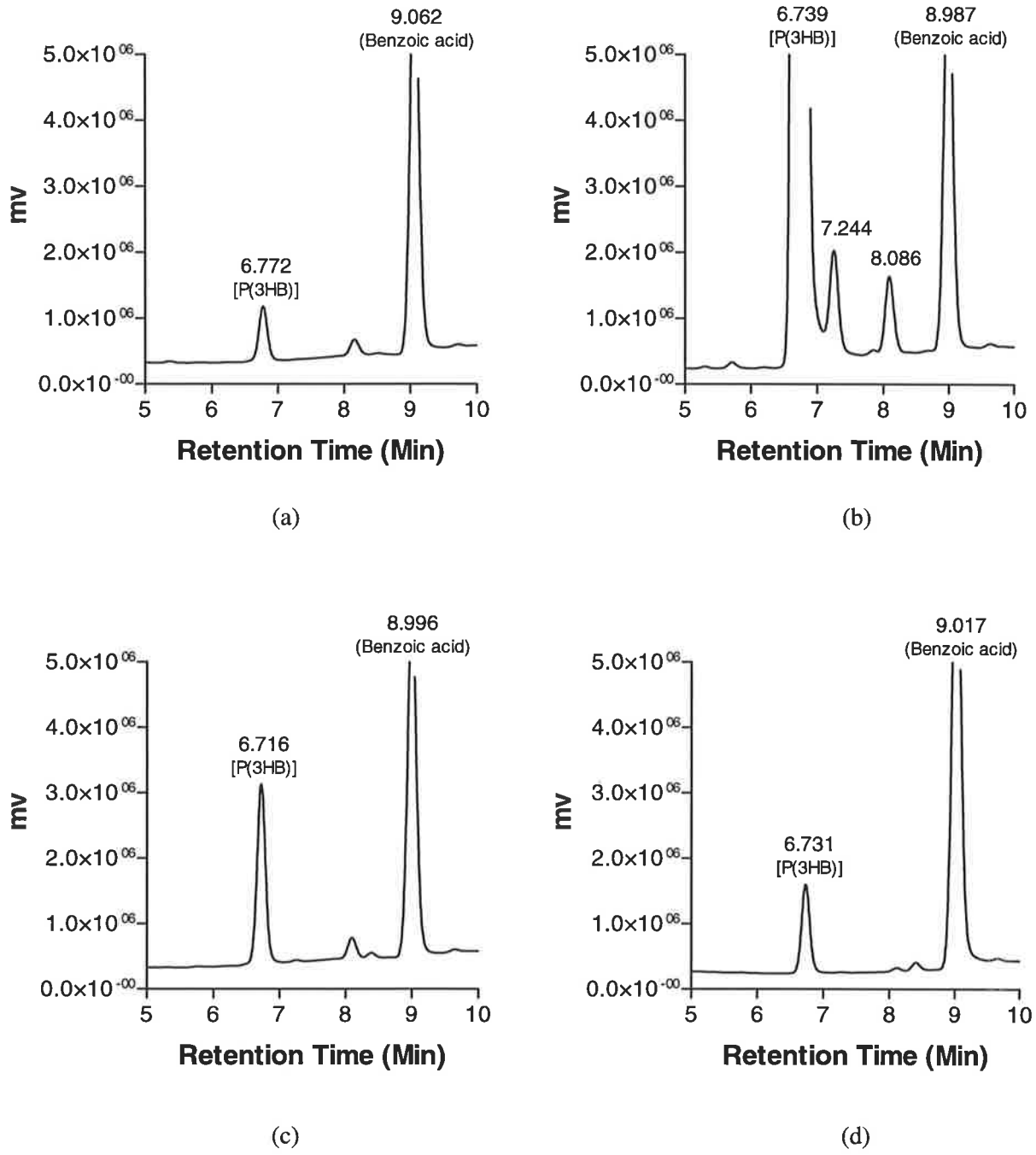


Figure 3.2. Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.

a. isolate 2-3-2; b. isolate 3; c. isolate 6C-2; d. isolate 10C-1-1. Chromatograms of samples prepared from isolates 2-3-2, 6C-1, and 10C-1-1 have retention times similar to that of the pure P(3HB) sample or P(3HB) produced by *E. coli* MD9101 (p4A), indicating P(3HB) production by these three isolates. For isolate 3, GC analysis suggests that it produced PHA with 3HB and other monomers, of which the retention times are similar to but different from that of P(3HB), indicating production of novel PHA. Benzoic acid was used as an internal control.

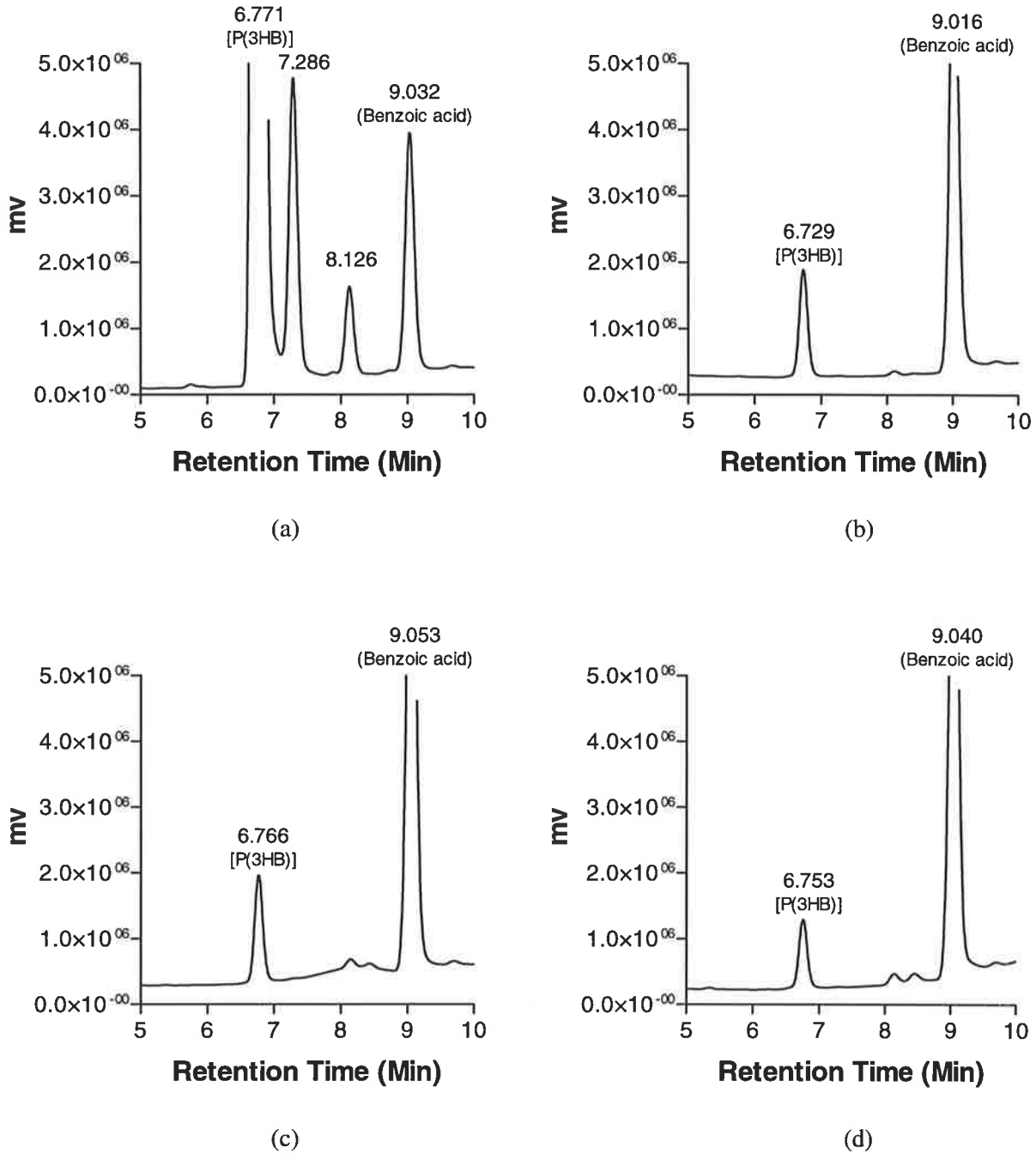


Figure 3.3. Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.

a. isolate 10C-1-3; b. isolate 14b-1; c. isolate 14C-1-2; d. isolate 14C-1-3. Chromatograms of samples prepared from isolates 14b-1, 14C-1-2, and 14C-1-3 indicate presence of products with a retention time similar to that of pure P(3HB) sample or P(3HB) produced by *E. coli* MD9101 (p4A), indicating P(3HB) production by these three isolates. For isolate 10C-1-3, GC analysis suggests that it produced PHA with 3HB and other monomers, of which the retention times are similar to but different from that of P(3HB), suggesting production of novel PHA. Benzoic acid was used as an internal control.

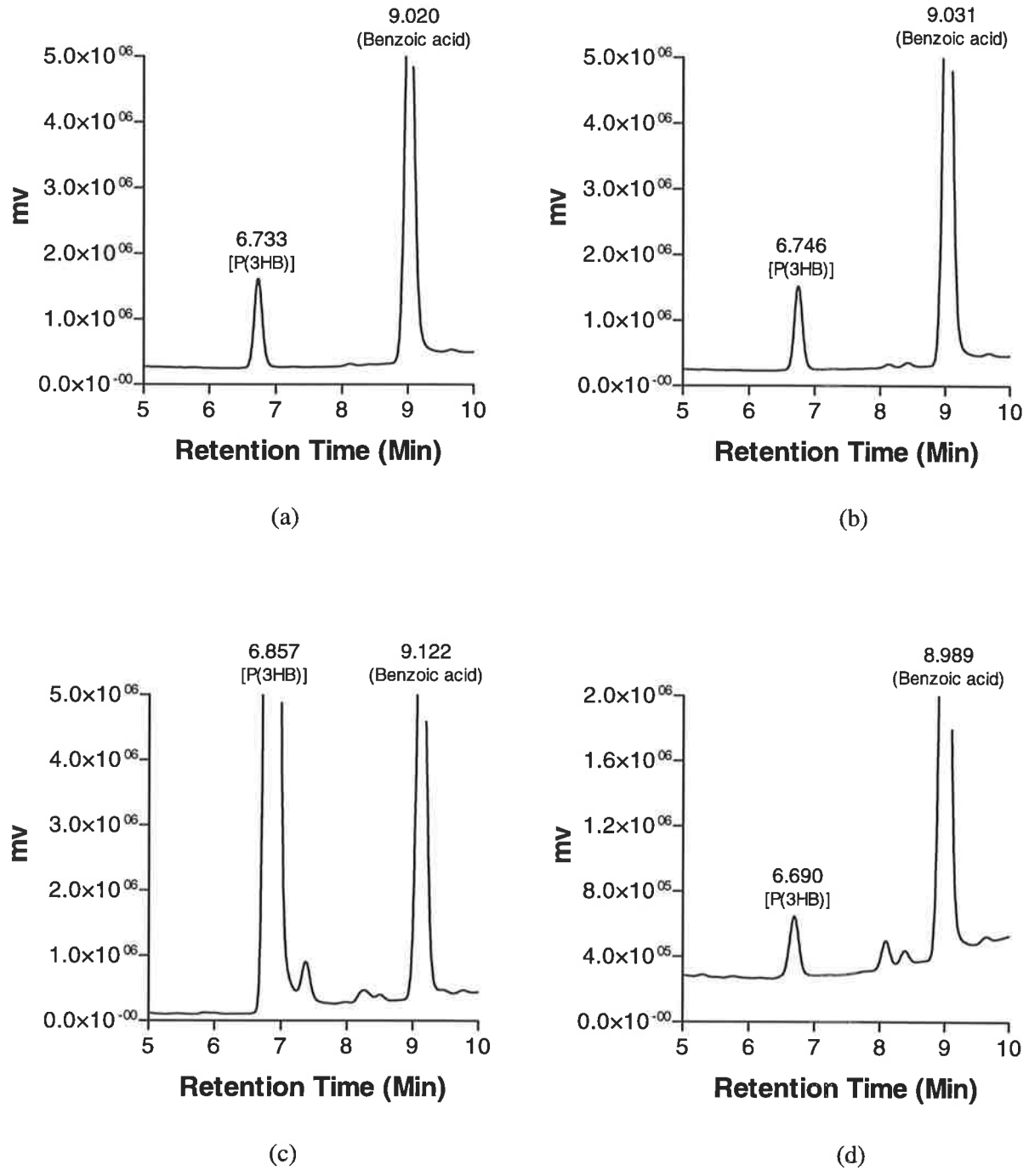


Figure 3.4. Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.

a. isolate K1B-1; b. isolate K1C-1; c. isolate M1c-1(1); d. isolate M2. Chromatograms of samples prepared from all four isolates indicated they produced products with a retention time similar to that of pure P(3HB) sample or P(3HB) produced by *E. coli* MD9101 (p4A), suggesting P(3HB) production by these isolates. Benzoic acid was used as an internal control.

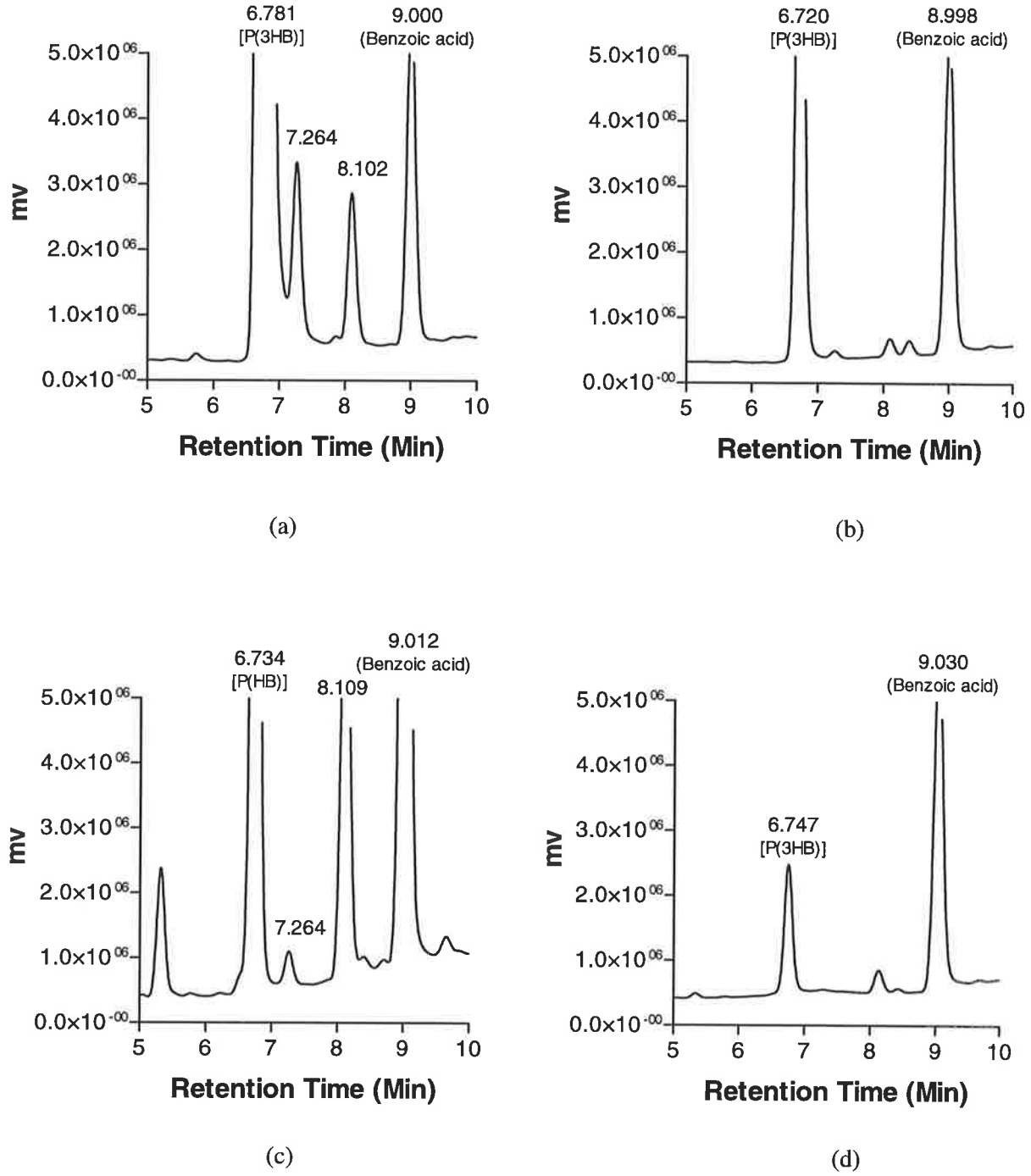


Figure 3.5. Gas chromatograms of PHA extracts prepared from soil isolates capable of producing PHA.

a. isolate M4-1; b. isolate M4-2; c. isolate M10c-1; d. isolate M10c-2. These chromatograms showed that isolates M4-2, and M10c-2 produced products with a retention time similar to that of pure P(3HB) sample, or P(3HB) produced by *E. coli* MD9101 (p4A), suggesting P(3HB) production by these three isolates. Isolates M4-1, and M10c-1, also produced compounds with retention times similar to but different from that of P(3HB), suggesting production of novel PHA from 3HB and other monomers..

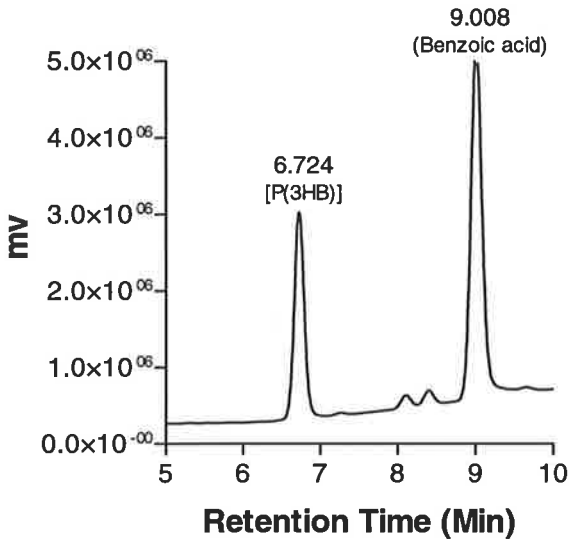


Figure 3.6. Gas chromatogram of PHA extracts prepared from isolate M14c-1.

GC analysis showed isolate M14c-1 produced a product with a retention time similar to that of pure P(3HB) sample or P(3HB) produced by *E. coli* MD9101 (p4A), indicating P(3HB) production by this isolate. Benzoic acid was used as an internal control.

MD9101 (p4A) (with a retention time of ~6.7 min). This indicated that all isolates tested were able to produce PHA. In addition, extracts from 3, 10c-1-3, M4-1 and M10c-1 contained components with retention times of ca. 7.24-7.26 min and 8.08-8.12 min respectively. These data suggest the producer bacteria may synthesise novel monomers with molecular weights larger than that of 3HB.

The yield of PHA expressed as percentage of dry cell weight is shown in Table 3.3. Of the nineteen isolates, 10c-1-3 produced the highest yield of PHA. On the basis of PHA yield and potential to express novel polymers, isolate 10c-1-3 was selected for further study.

3.2.3 DSC of PHA produced by the isolate 10c-1-3

Physical and chemical properties of PHA produced by 10c-1-3 were determined by differential scanning calorimetry. Lyophilised cells (about 10 mg) prepared from cultures grown in R-medium (+2.0% glucose) at 30 °C for 2 days were encapsulated in a standard aluminium pan and placed into the sample chamber of DuPont 2200 Thermal Analyser and the sampled subjected to analysis using the program described in Section 2.7.6. Figure 3.7 shows a typical DSC thermogram of the PHA produced by the isolate 10c-1-3. These results show that the polymer has a glass transition temperature of about 0 °C, a crystallisation temperature of 47.5 °C and a melting temperature of 171.2 °C. The crystallisation temperature and the melting temperature of the polymer are lower than that of pure P(3HB) (50 °C and 174 °C respectively). This data indicates the PHA produced by 10c-1-3 is likely to be a heteropolymer and consequently that incorporation of monomers other than 3HB into this PHA resulted in decrease of melting temperature. This result supports the GC analysis shown previously.

3.2.4 Transmission electron microscopy of PHA granules

Strain 10c-1-3 was cultured in R-medium (+2.0% glucose) at 30 °C for 2 days and cells were prepared for examination by TEM as outlined in Section 2.7.2. When stained, ultra thin sections were examined by transmission electron microscopy, individual cells (0.59-0.69 μm × 1.49-1.94 μm in size) were found to contain 1 to 9 PHA granules (with a

Table 3.3. PHA content in dry cell mass of each isolate.

Isolate	P(3HB)	PHAs other than P(3HB)	
		A*	B*
1c-1	1.70	-	-
1c-1-1	3.20	-	-
2-3-2	0.80	-	-
3	30.10	0.70	0.50
6c-2	6.00	-	-
10c-1-1	3.00	-	-
10c-1-3	58.50	4.10	1.40
14b-1	3.90	-	-
14c-1-2	3.20	-	-
14c-1-3	2.10	-	-
K1b-1	4.50	-	-
K1c-1	4.10	-	-
M1c-1(1)	42.20	-	-
M2	1.70	-	-
M4-1	27.60	0.70	0.40
M4-2	16.80	-	-
M10c-1	9.00	0.30	4.90
M10c-2	4.10	-	-
M14c-1	5.60	-	-

*: A and B are different PHAs from P(3HB).

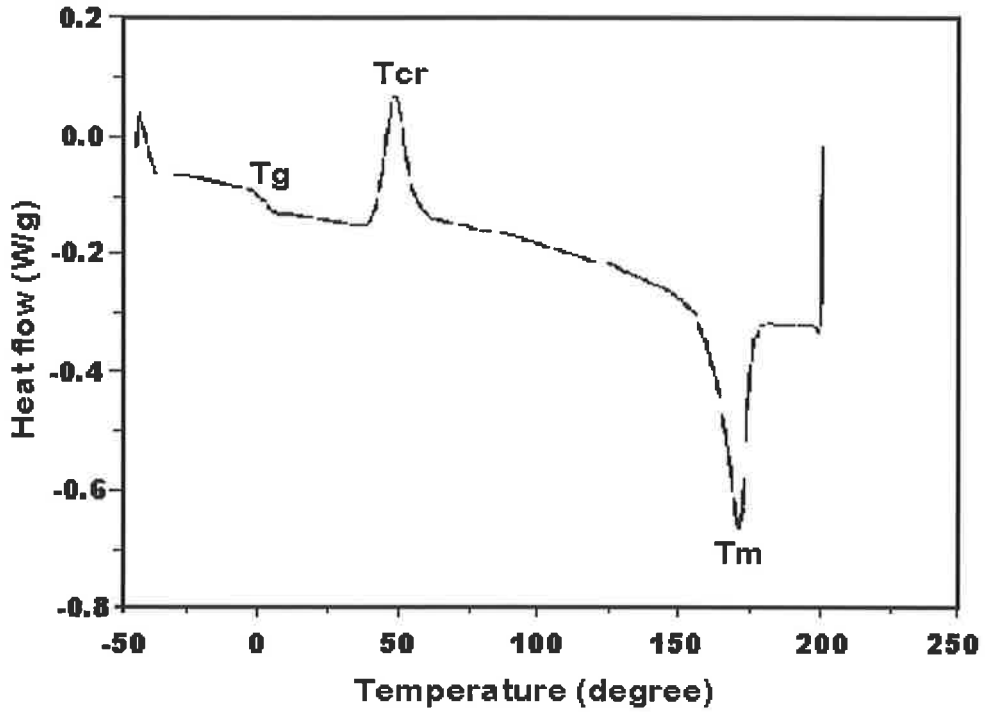


Figure 3.7. DSC thermogram of PHA produced by isolate 10c-1-3.

Tg: glass transition temperature (about 0 °C);
Tcr: crystallisation temperature (47.5 °C);
Tm: melting temperature (171.2 °C).

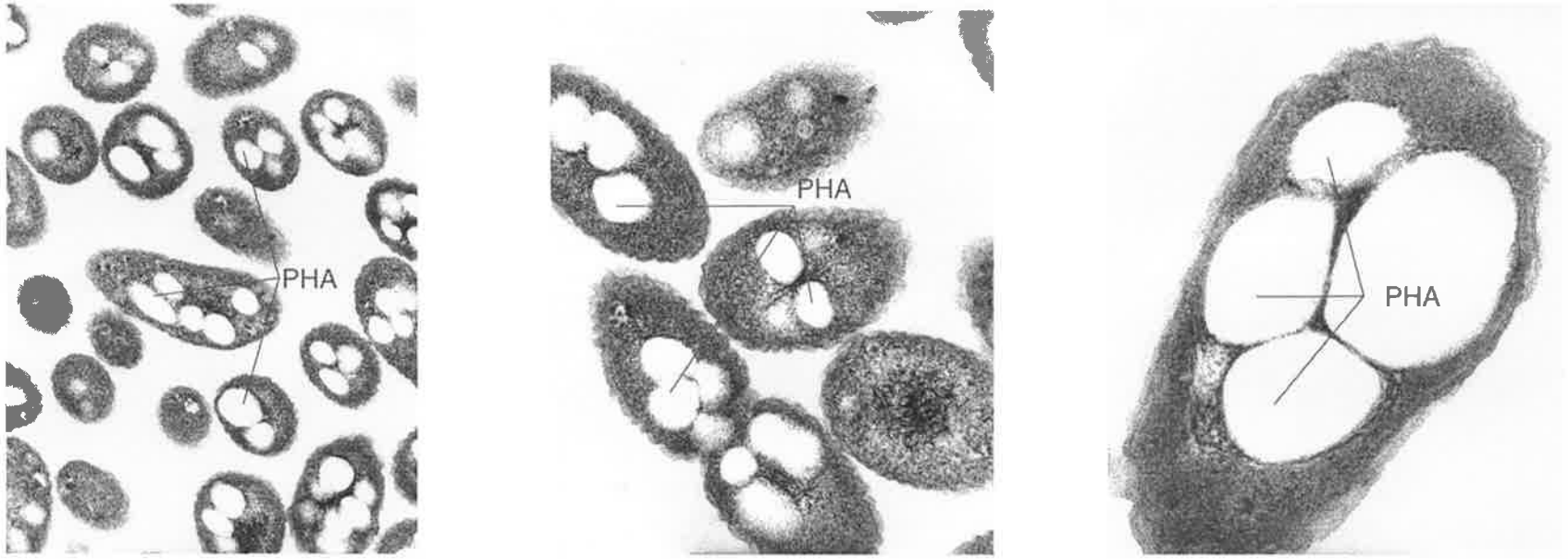
size of approximately 0.18-0.33 μm × 0.32-0.44 μm) (Figure 3.8). These granules were usually aggregated together within the cytoplasm of individual cells.

3.3 Discussion

Nineteen PHA producing isolates were isolated from 15 soil samples collected from different sites around Adelaide. Presence of intracellular granules and PHA production was confirmed microscopically using the Nile Blue A staining method and by comparative gas chromatography analysis. All the 19 isolates were characterised to genus level where possible. Isolates belonging to the genera *Methanomonas*, *Acetobacter*, *Pseudomonas*, *Xanthomonas*, or *Aeromonas* were identified. This variety of genera capable of expressing PHA is consistent with previously published observation (Tezuka *et al.*, 1980; Foellner *et al.* 1993; Suzuki *et al.*, 1986b; Ramsay *et al.* 1994; and Fukui and Doi, 1997).

Gas chromatography analysis showed that isolates 3, 10c-1-3, M1-1(1), M4-1, and M4-2 produced PHA at high levels. Furthermore, this data indicated isolates 3, 10c-1-3, M4-1 and M10c-1 apparently produced PHA comprising monomers other than 3HB. In particular, one isolate, *Pseudomonas* strain 10c-1-3, produced large amounts of PHA when grown on R-medium [+2.0% (w/v) glucose]. Cells from this culture contain between 1 and 9 intracellular granules. However, yields of PHA expressed [58.5% (w/w, of dry cell weight] are considerably less than that previously reported from *R. eutropha* (Kim *et al.*, 1994a,b; and Ryu *et al.*, 1997) for example, as well as some other PHA expressing bacteria (Page *et al.* 1992; and Zhang *et al.* 1994). Nevertheless, the lower melting temperature (171 °C) of this polymer compared with that of P(3HB) (174 °C) (Holmes, 1985), as well as the unusual GC profile, provide evidence that monomers other than 3HB are incorporated (Slater *et al.*, 1992).

Together, this information indicates that the PHA biosynthetic capacity and spectrum of substrates of *Pseudomonas* 10c-1-3 are likely to be greater than that of the other isolates obtained in this study. On this basis it was hypothesised that 10c-1-3 may possess novel PHA biosynthetic genes and was consequently selected for further study at molecular level.



(a) (b) (c)

Figure 3.8. Transmission electron micrographs of PHA granules in *Pseudomonas* strain 10c-1-3.

(a) magnification 16,000 ×; (b) magnification 32,000 × and (c) magnification 64,000 ×

Chapter 4

Construction of a *Pseudomonas* 10c-1-3 genomic library and selection of *pha* clones

4.1 Strategies

4.1.1 Strategy for the detection of *pha* genes

Inspection of aligned DNA sequences for the *pha* genes published in the Genbank database shows that *phaA* and *phaB* genes have significant segments with conserved sequence homology and *phaC* genes have little similarity in both size and sequence (Appendix C). Thus studies designed to examine distribution of *pha* genes by PCR amplification techniques may be limited to amplification of *phaA* and *phaB*. The lack of conserved blocks of sequence in *phaC* would restrict the design of oligonucleotide primers to ones with very high degeneracy. Nevertheless, Lopez *et al.* (1997) designed primers [5'-GTTGCAGTACAAGCCGCT-3' (forward); and 5'-CACGGCACGATATGGTCT-3' (reverse), see Appendix C for their binding sites] based on the published sequence of the *phaC* cloned from *R. eutropha* to study homologues in some other bacteria (*Bacillus* sp., and some Gram positive and negative river bacterial isolates). However, for reasons already stated, amplification of segments of *phaA* and *phaB* genes represents a more reliable strategy. Using *phaA* and *phaB* based primers and PCR techniques, DNA fragments of the *pha* genes could be amplified, sequenced and used as a probe for detection of full-length *pha* clones using Southern hybridisation techniques. The use of this strategy to identify putative *pha* biosynthesis genes in PHA producing soil isolates and recombinant DNA clones is described in this chapter.

4.1.2 Strategy for genomic library construction

A wealth of published data indicates that all known *pha* biosynthetic genes are clustered in operons in single discrete regions of the bacterial genome (Figure 1.3). This suggests that the *pha* genes in *Pseudomonas* strain 10c-1-3 are likely to be clustered in an operon. Relevant literature suggests these genes from this isolate would be located on a ~5 kbp region. Consequently cosmid cloning techniques were selected as an appropriate method of obtaining full-length clones. Since the *Pseudomonas* genome is unlikely to be in excess of 1×10^7 bp, and the average size of a DNA insert in a cosmid clone is ~40 kbp, a library of about 250 clones would be required in order to be sure of obtaining a single clone harbouring the PHA biosynthetic genes.

4.1.3 Strategy for genomic library screening

Two possible methods for screening a cosmid library for *pha* positive clones were considered in this study.

Positive clones producing PHA as granules should be able to be detected by microscopic examination of Nile Blue A stained preparations (see Section 2.7.1), or by assay of the enzyme activities of the PHA pathway (see Section 1.5.1). Screening a large library by these methods is tedious and based on the assumption that PHA is synthesised by cosmid clones. These technical problems limit their use.

An alternative approach, which makes no assumptions about expression of recombinant genes, relies on use of labelled DNA probes and PCR or colony hybridisation methods to select positive clones. While colony hybridisation is sensitive, it sometimes gives false positive results. PCR amplification is more stringent and reliable, but assumes that suitable oligonucleotide primers are available to amplify a target from the desired gene(s). The availability of techniques which allow rapid small scale isolation of crude DNA preparations from colonies, make PCR methods potentially useful for screening large numbers of samples. To further simplify the PCR screening process, a genomic library can be divided into pools and the positive pools were subdivided into subpools of clones in microtitre plates. Cross plate sampling procedures can then be used to allow detection of positive clones in separate plates, and at the same time substantially reduce the number of PCR amplification reactions required to about 20% of the number of clones to be screened.

4.2 Results

4.2.1 Detection of *pha* genes

4.2.1.1 Design of degenerate primers

Multiple alignment of sequence data for published *phaA* and *phaB* sequences allowed identification of conserved regions useful for construction of PCR primers (Figures 4.1, 4.2 and Appendix C). Oligonucleotide primer pair #4033 and #4034 flank a 200 bp region of *phaA* of *R. eutropha pha* operon. Similarly oligonucleotide primer pair #4035 and #4036 flank a 314 bp fragment of *phaB* of *R. eutropha pha* operon.

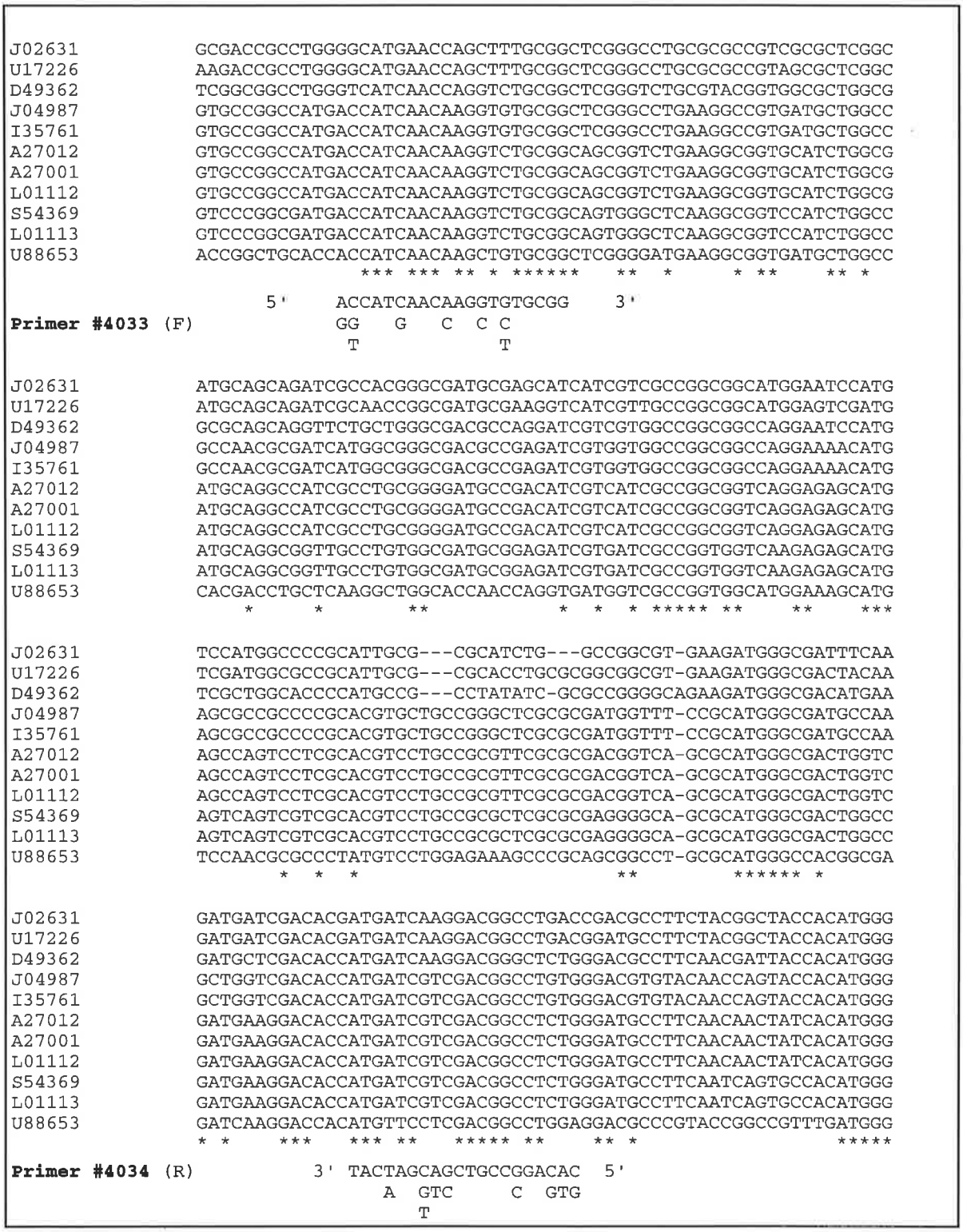


Figure 4.1. Alignment of 11 *phaA* DNA sequences and location of degenerate primers used to amplify DNA from *Pseudomonas* strain 10c-1-3 and isolate 2-3-2.

Genbank accession numbers for each of the source sequences are shown. For each block of sequence, identities are indicated by asterisks, and gaps are shown as hyphens.

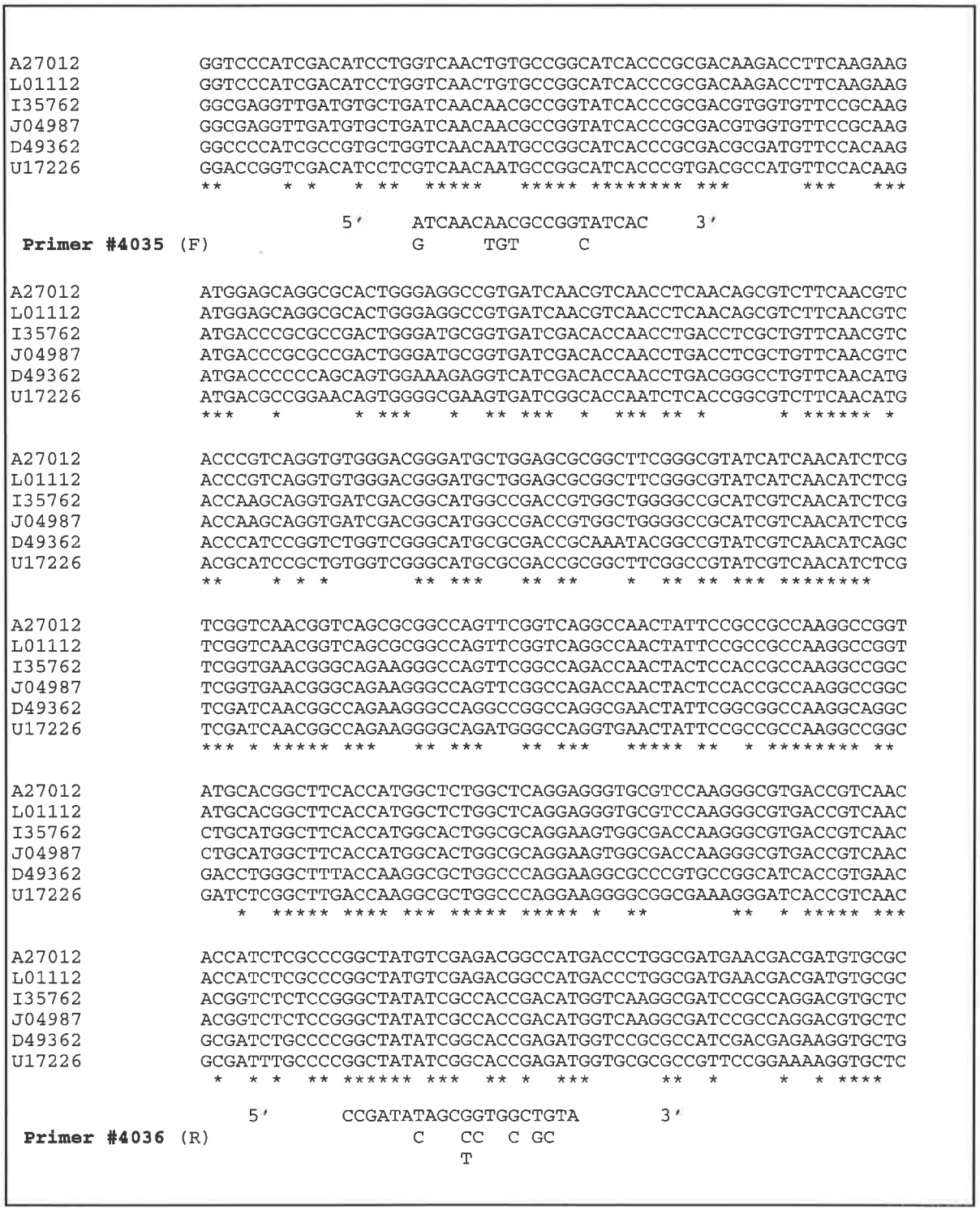


Figure 4.2. Alignment of 6 *phaB* sequences and location of degenerate primers used to amplify DNA from *Pseudomonas* strain 10c-1-3 and isolate 2-3-2.

Genbank accession numbers for each of the source sequences are shown. For each block of sequence, identities are indicated by asterisks, and gaps are shown as hyphens.

4.2.1.2 Detection of *phaA* genes by PCR

Chromosomal DNAs extracted from isolate 2-3-2 and *Pseudomonas* strain 10c-1-3 were used as DNA templates for PCR. PCR was carried out as described in Section 2.11.8 with the primer pairs #4033 and # 4034; #4035 and # 4036 respectively. Plasmid p4A was used as a positive control. Milli Q water was substituted for DNA in negative controls.

Amplification of isolate 2-3-2 and *Pseudomonas* strain 10c-1-3 chromosomal DNA and p4A plasmid (positive control) all gave expected *phaA* and *phaB* based PCR products of approximately 0.2 kbp and 0.3 kbp respectively (Figure 4.3). Negative control samples did not produce any PCR products.

4.2.1.3 Cloning of the PCR products

Agarose gel purified 0.2, 0.3 kbp PCR products from isolate 2-3-2 and *Pseudomonas* strain 10c-1-3 were cloned separately into the pGEM-T vector (Figure 4.4). One clone harbouring plasmid pSGZ1 was selected from the clones transformed with the ligation reaction of the *phaA* specific 0.2 kbp product from *Pseudomonas* strain 10c-1-3 and pGEM-T. Clones containing the *phaB* specific 0.3 kbp fragment were not obtained.

Restriction enzyme digestion analysis of plasmid pSGZ1 using *Pvu* II produced fragments of ~2.67 kbp and ~0.64 kbp when separated by agarose gel electrophoresis (Figure 4.5). This is consistent with the insertion of a ~0.2 kbp DNA fragment into the 0.4 kbp *Pvu* II fragment of pGEM-T. As a control, *Pvu* II digested pGEM5Zf(+) (from which pGEM-T was created) produced DNA fragments of ~2.67 kbp and ~0.44 kbp.

4.2.1.4 Sequence analysis of the cloned PCR product in pSGZ1

To confirm that the ~0.2 kbp insert within pSGZ1 represents an internal region of *phaA* ($phaA_{int}$) from *Pseudomonas* strain 10c-1-3, the ~0.2 kbp insert was sequenced using M13 forward and reverse Dye Primers (Figure 4.6a). BLASTN DNA sequence analysis showed that the ~0.2 kbp insert in pSGZ1 was similar to the *phaA* sequences of other bacteria with the highest homology to the *R. eutropha phaA* gene (65.0% similarity) (Figures 4.6b and 4.6c).

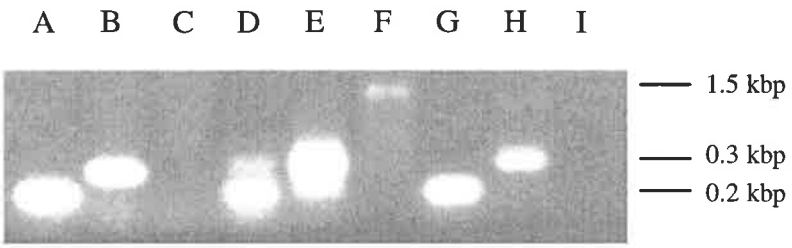


Figure 4.3. PCR products from plasmid p4A, and chromosomal DNAs of isolate 2-3-2 and *Pseudomonas* strain 10c-1-3 using primer pairs #4033 and #4034; #4035 and #4036; and #4033, #4036.

Lane A	Plasmid p4A, primer pair #4033, #4034
Lane B	Plasmid p4A, primer pair #4035, #4036
Lane C	Plasmid p4A, primer pair #4033, #4036
Lane D	Isolate 2-3-2 chromosomal DNA, primer pair #4033, #4034
Lane E	Isolate 2-3-2 chromosomal DNA, primer pair #4035, #4036
Lane F	Isolate 2-3-2 chromosomal DNA, primer pair #4033, #4036
Lane G	*Pseudomonas* strain 10c-1-3 chromosomal DNA, primer pair #4033, #4034
Lane H	*Pseudomonas* strain 10c-1-3 chromosomal DNA, primer pair #4035, #4036
Lane I	*Pseudomonas* strain 10c-1-3 chromosomal DNA, primer pair #4033, #4036

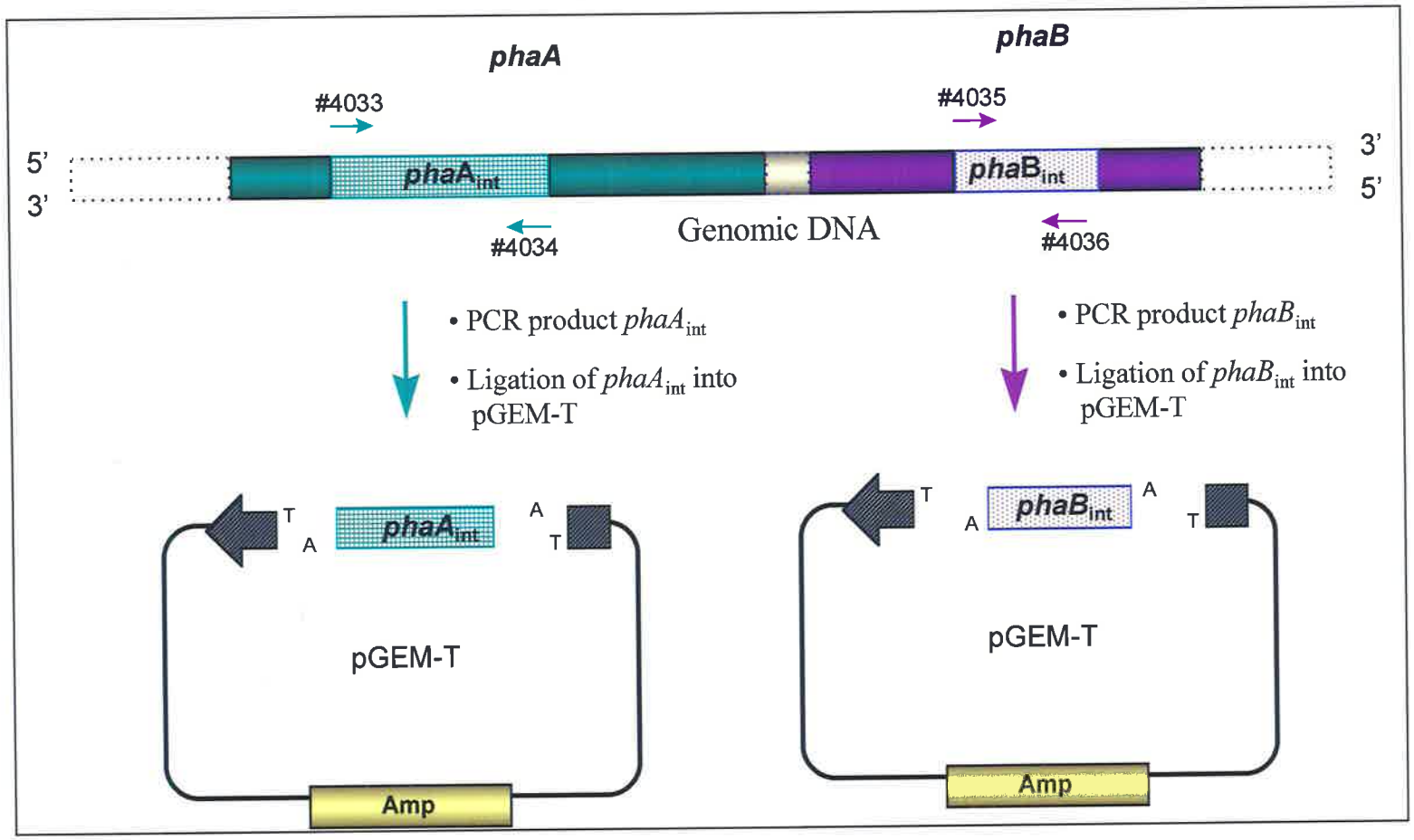


Figure 4.4. Strategy used for cloning *phaA*$_{int}$, *phaB*$_{int}$ PCR products into pGEM-T using degenerate primer pairs #4033 and #4034, #4035 and #4036.

Genomic DNA prepared from *Pseudomonas* 10c-1-3 was used as a template for PCR. The PCR products were agarose gel purified prior to ligation with pGEM-T.

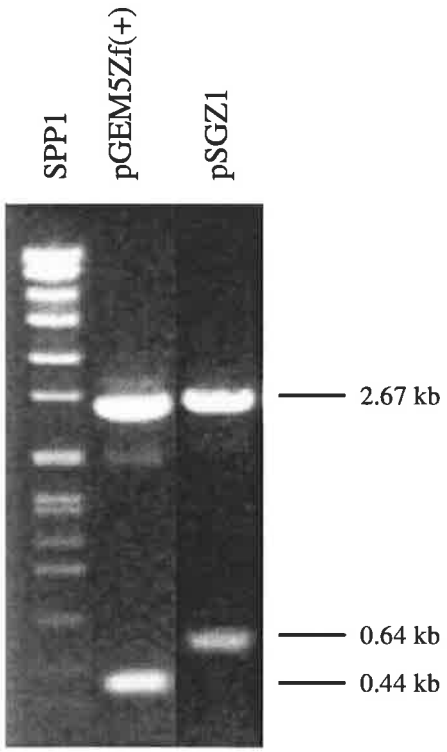


Figure 4.5. Restriction analysis of plasmid pSGZ1.

Vector pGEM-T [pGEM5Zf(+)] and plasmid pSGZ1 containing *phaA*$_{int}$ (0.2 kbp) from *Pseudomonas* strain 10c-1-3 were digested to completion with *Pvu* II and fragments separated by agarose gel electrophoresis. The presence of the 0.64 kbp fragment in pSGZ1 is consistent with the insertion of a 0.2 kbp fragment into pGEM-T.

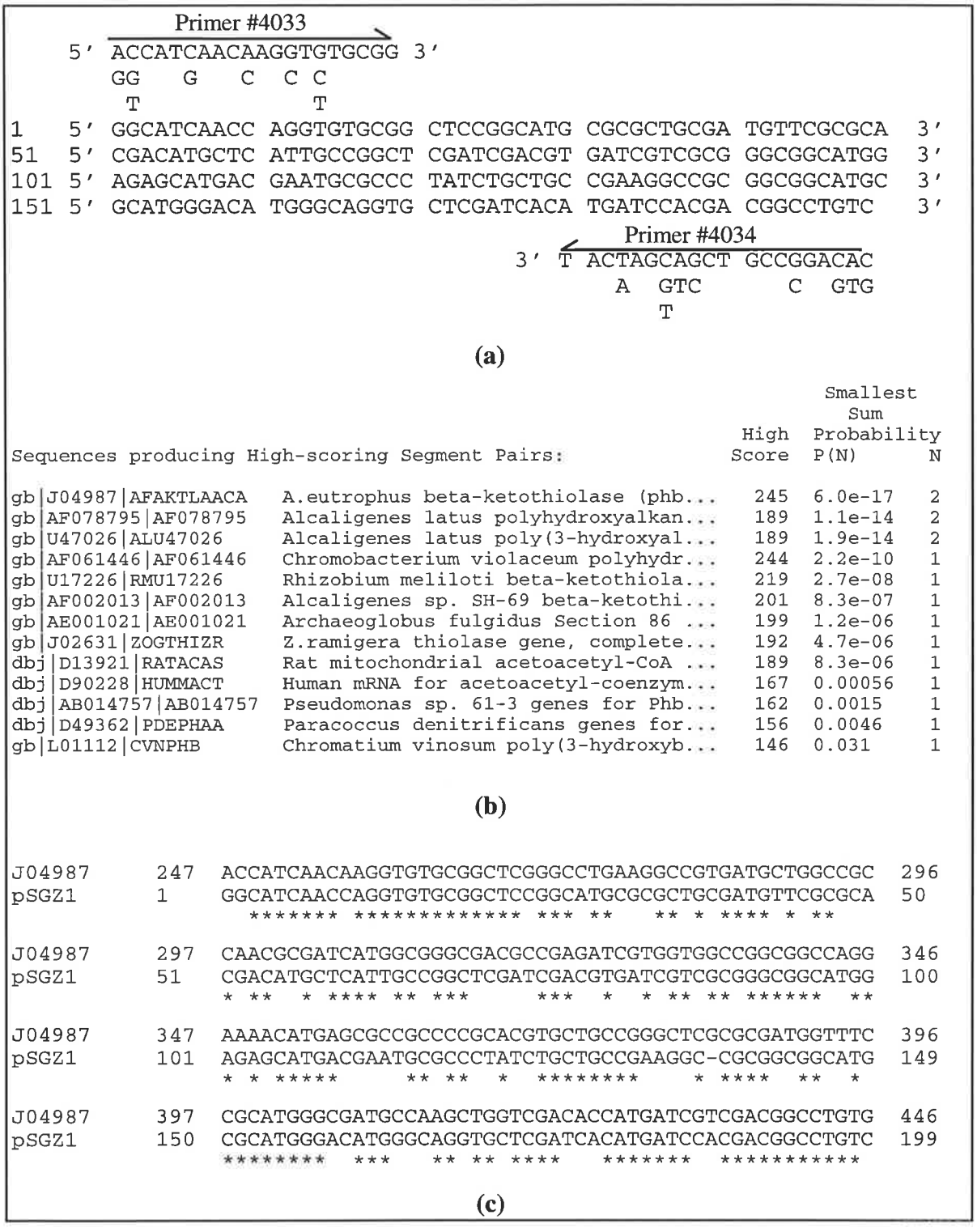


Figure 4.6. DNA sequence analysis of the PCR product *phaA*$_{int}$ cloned in pSGZ1.

(a) DNA sequence of the *phaA*$_{int}$ in pSGZ1; **(b)** Result of a BLASTN analysis of the *phaA*$_{int}$ sequence from pSGZ1, showing that the DNA sequence of this insert was similar to the *phaA* genes of other bacteria. The *phaA*$_{int}$ sequence was most similar to the *phaA* gene of *R. eutropha* (Genbank accession number J04987); **(c)** DNA sequence alignment of the *phaA*$_{int}$ insert in pSGZ1 and J04987. These sequences are 65% similar.

4.2.1.5 Probe preparation from pSGZ1 by PCR

To facilitate the use of the *phaA*$_{int}$ as a probe in colony and Southern hybridisation methods, this insert was Digoxigenin labelled by incorporation of Digoxigenin-11-dUTP using a PCR amplification method described in the Materials and Methods (Section 2.15.1.2). Purified pSGZ1 was used as DNA template and oligonucleotides #4033 and #4034 were used as primers.

4.2.2 Construction of a cosmid genomic library

4.2.2.1 Preparation and evaluation of the library

Unsheared genomic DNA prepared from *Pseudomonas* 10c-1-3 was partially digested with *Mbo* I and used to construct a cosmid library as described in the Materials and Methods (see Section 2.13). Clones containing cosmid DNA were selected by plating on NA containing ampicillin. As a result, a library comprising some 4,464 clones was prepared. About 2% of these were able to grow on NA containing tetracycline. Since tetracycline resistant clones do not contain DNA inserted into the *Bam* HI site of the tetracycline resistance gene of cosmid vector pHC79, this indicates that ~98% of clones contain *Pseudomonas* DNA. The size of this library was estimated to be sufficient to contain the entire *Pseudomonas* genome approximately 15 fold.

To estimate the average size of the DNA inserts, cosmid DNA was prepared from 7 randomly selected clones. Restriction enzyme digestions of these 7 cosmids and vector pHC79 (as a negative control) were performed with *Eco* RI and *Sal* I. Restriction sites for these two enzymes flank the *Bam* HI site in pHC79. Analysis of the sizes of DNA fragments produced by digestion indicated each of the 7 clones contained the correct vector fragment and insert fragments (Figure 4.7a). *Eco* RI digestion showed that the average size of the inserts was about 40 kbp (Figure 4.7b).

4.2.2.2 The screening of the cosmid library

The library was screened for cosmid which contained DNA homologous to *phaA*, using PCR screening techniques, based on those described by Holmes and Quigley (1981),

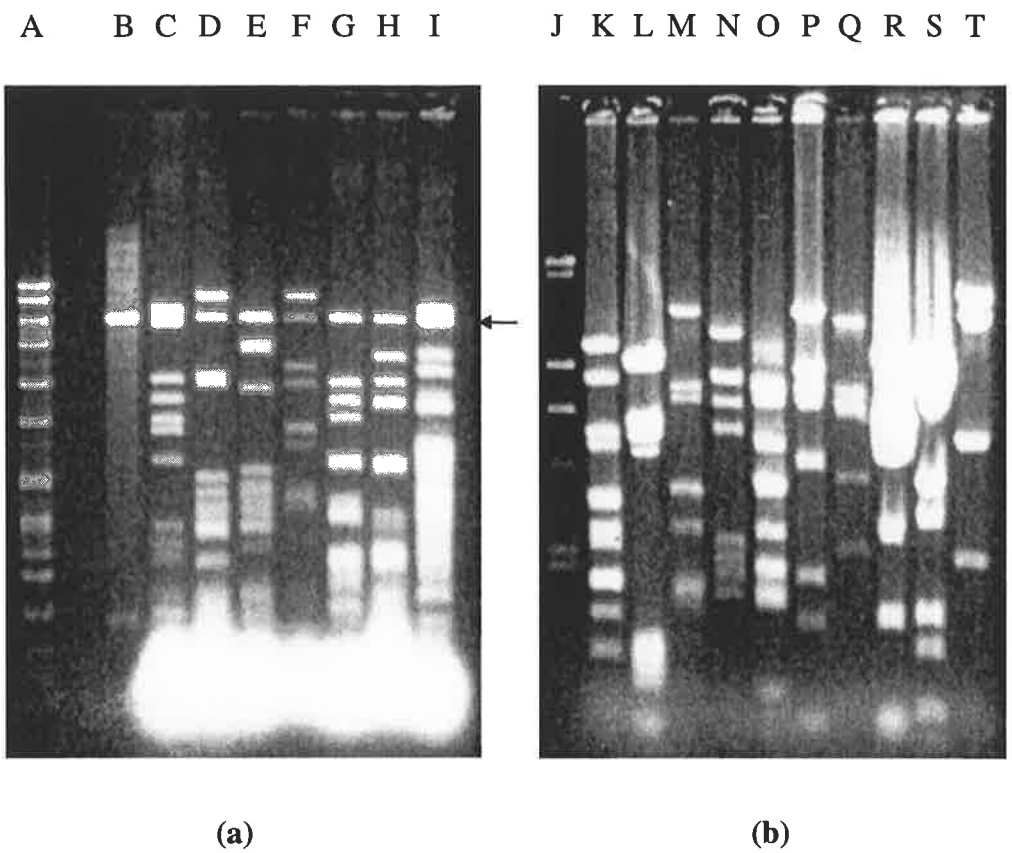


Figure 4.7. Digestion of hybrid cosmids with *Eco* RI/*Sal* I and *Eco* RI respectively.

A. DNA marker SPP1;
B. *Eco* RI/*Sal* I digested vector pHC79;
C-I. *Eco* RI/*Sal* I digested hybrid cosmids.

Compared to lane B, all the hybrid cosmids contained a 5.9 kbp DNA fragment (as indicated with the above arrow), indicating the presence of the vector in each hybrid cosmid.

J. DNA marker (*Hin*d III cut λ DNA);
K-T. *Eco* RI digested hybrid cosmids.

The size of each hybrid cosmid indicated that the average size of the inserts in hybrid cosmids was about 40.0 kbp.

Gussow and Clackson (1989), Rapley and Walker (1992), Gustafson *et al.* (1993), and Israel (1993). The screening procedure required 2 separate steps.

The first step was to select pools in which some clones contained the desired DNA fragment flanked by the primer pair #4033 and #4034 ($phaA_{int}$). To achieve this, the whole library was divided into 61 pools, each containing 96 clones. The cells of the 96 clones in a pool were mixed, DNA extracted and used as template for PCR. Among the 61 pools, 23 pools gave PCR products with the expected size of 200 bp. The second step involved using each set of 96 clones from these positive pools to individually inoculate the wells of microtitre plates. Each well contained 200 µl of NB plus Ap (100 µg/ml). The microtitre plates were then incubated at 37°C O/N. Cell culture (50 µl) from each microtitre plate well was then combined into 20 subpools as described in Figure 4.8. The cells from each subpool were pelleted, washed 2 times with Milli Q water, resuspended in 100 µl of Milli Q water, heat lysed, centrifuged and 2 µl of the supernatant was used as DNA template for PCR screening (Section 2.11.8). Clones defined by the intersection of PCR positive row and column subpools were recorded as containing DNA homologous to $phaA_{int}$. This was confirmed by PCR analysis for the individual candidate clones. As a result, 17 individual positive clones were selected and numbered according to their pool number and location on microtitre plates as follows: (2)D7, (3)F4, (6)H4, (20-1)C1, (20-2)E6, (20-2)G1, (32)A4, (32)C12, (32)G9, (34)F3, (36)H9, (44)C12, (44)H11, (50)D6, (52)H5, (55)G1, and (56)H4.

4.2.2.3 Selection of a hybrid cosmid for subcloning of *pha* genes

As a first step in the genetic analysis of PCR positive cosmid clones to identify subclones containing the desired *pha* genes, the cosmid clones were examined for potential to produce PHA granules. Production of PHA granules would indicate the clones encoded all genes necessary for biosynthesis of PHA. The individual 17 clones were cultivated in nutrient broth containing glucose (2.0%). After O/N incubation at 37°C with agitation, smears of each culture were prepared, stained with Nile Blue A and examined by incident light fluorescence microscopy. No PHA production was detected for any of the cells on smears prepared from the PCR positive clones. Thus in the absence of a phenotypic test to confirm the presence of PHA biosynthetic genes, an alternative molecular approach was

96 well microtitre plate

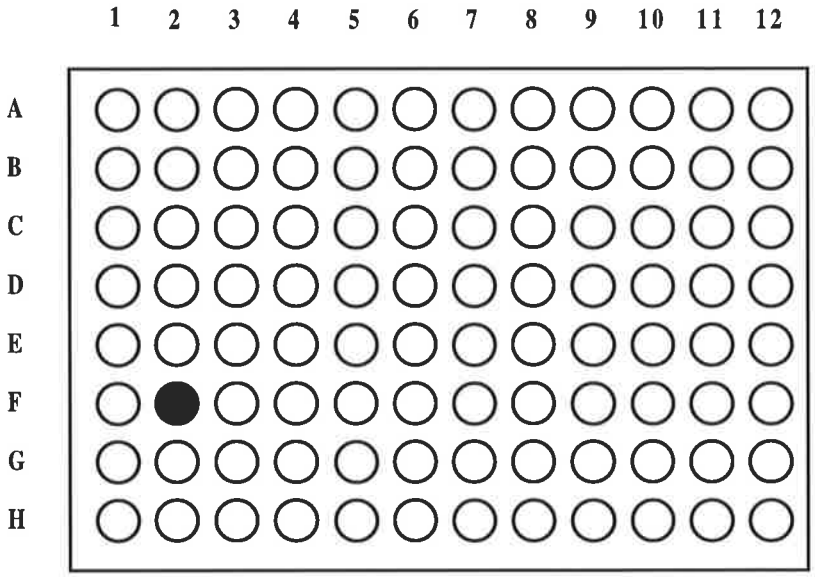


Figure 4.8. A schematic illustration of the microtitre plate subpooling procedure used to identify PCR positive cosmid clones.

Clones were cultured in separate wells and then pooled as described below. Crude DNA extracts for each subpool were obtained and used as the template for PCR. Amplification of 200 bp DNA fragments was taken to indicate a positive PCR result. A positive result obtained with subpool 2 and F for example, would mean that the clone cultured in well F2 contained cosmid DNA encoding $pha_{int.}$ Thus for every 96 clones, only 12 + 8 = 20 PCR reactions were required to identify clones containing the desired DNA.

Subpool details

Pools from Column

```
1=A1 + B1 + ..... +G1 + H1
2=A2 + B2 + ..... +G2 + H2
3=A3 + B3 + ..... +G3 + H3
4=A4 + B4 + ..... +G4 + H4
5=A5 + B5 + ..... +G5 + H5
6=A6 + B6 + ..... +G6 + H6
7=A7 + B7 + ..... +G7 + H7
8=A8 + B8 + ..... +G8 + H8
9=A9 + B9 + ..... +G9 + H9
10=A10 + B10 + ... +G10 + H10
11=A11 + B11 + ... +G11 + H11
12=A12 + B12 + ... +G12 + H12
```

Pools from Row

```
A=A1 + A2 + ..... + A11 + A12
B=B1 + B2 + ..... + B11 + B12
C=C1 + C2 + ..... + C11 + C12
D=D1 + D2 + ..... + D11 + D12
E=E1 + E2 + ..... + E11 + E12
F=F1 + F2 + ..... + F11 + F12
G=G1 + G2 + ..... + G11 + G12
H=H1 + H2 + ..... + H11 + H12
```

used.

Cosmid DNA was prepared from clones (36)H9, (44)H11, (50)D6, (52)H5, (55)G1 and this was subjected to restriction enzyme analysis. The DNA fragments obtained were then separated on agarose gel electrophoresis (Figure 4.9a, c) and blotted onto Nylon membranes. Southern analysis showed that DNA digested with *Apa* I, *Bgl* II, *Eco* RI, and *Eco* RV all gave DNA fragments which hybridised to digoxigenin labelled $phaA_{int}$ DNA probe under stringent conditions (Figure 4.9b, d). Of the four restriction enzymes, *Apa* I digestion gave the fewest DNA fragments, suggesting *Apa* I probably was the best enzyme of choice to begin subcloning DNA fragments from a cosmid. *Apa* I digestion of (50)D6 gave only three DNA fragments (4.1, 7 and ~28 kbp) (Figure 4.10). Furthermore, Southern analysis showed that when *Apa* I was used to digest cosmid DNA from (50)D6 and *Pseudomonas* 10c-1-3 chromosomal DNA, DNA fragments of identical size hybridised to labelled $phaA_{int}$ (Figure 4.9b). This suggested that (50)D6 could encode all the *pha* genes, if it is assumed that they are arranged in an operon with a size of about 5 kbp.

To confirm that the PCR positive clones (36)H9 and (50)D6 contained DNA homologous to $phaA_{int}$, primers #5017 and #5018 internal to $phaA_{int}$ were designed and used to sequence the PCR products amplified from cosmids prepared from these clones using primer pair #4033 and #4034. After sequencing, the PCR products were compared to the sequence obtained for $phaA_{int}$ (Section 4.2.1.4). The aligned sequence data is shown in Figure 4.11a. Sequence data derived from (36)H9 and (50)D6 are essentially identical, but differ significantly from that of $phaA_{int}$. However, BLASTN analysis showed the PCR based sequences from (36)H9 and (50)D6 were similar to *pha* related sequences identified in other organisms (Figure 4.11b). This suggests that *Pseudomonas* 10c-1-c may contain at least two separate *phaA* genes within the genome. On the basis of these results, cosmid isolated from (50)D6 was selected for genetic analysis and renamed pCT400.

4.3 Discussion

As a first step in cloning *pha* genes from the PHA producing soil isolate 10c-1-3, a 200 bp fragment internal to a putative *phaA* gene was amplified by PCR. The rationale behind this first step was to establish a mechanism by which the *pha* genes of *Pseudomonas* strain 10c-1-3 could be isolated by standard recombinant DNA techniques.

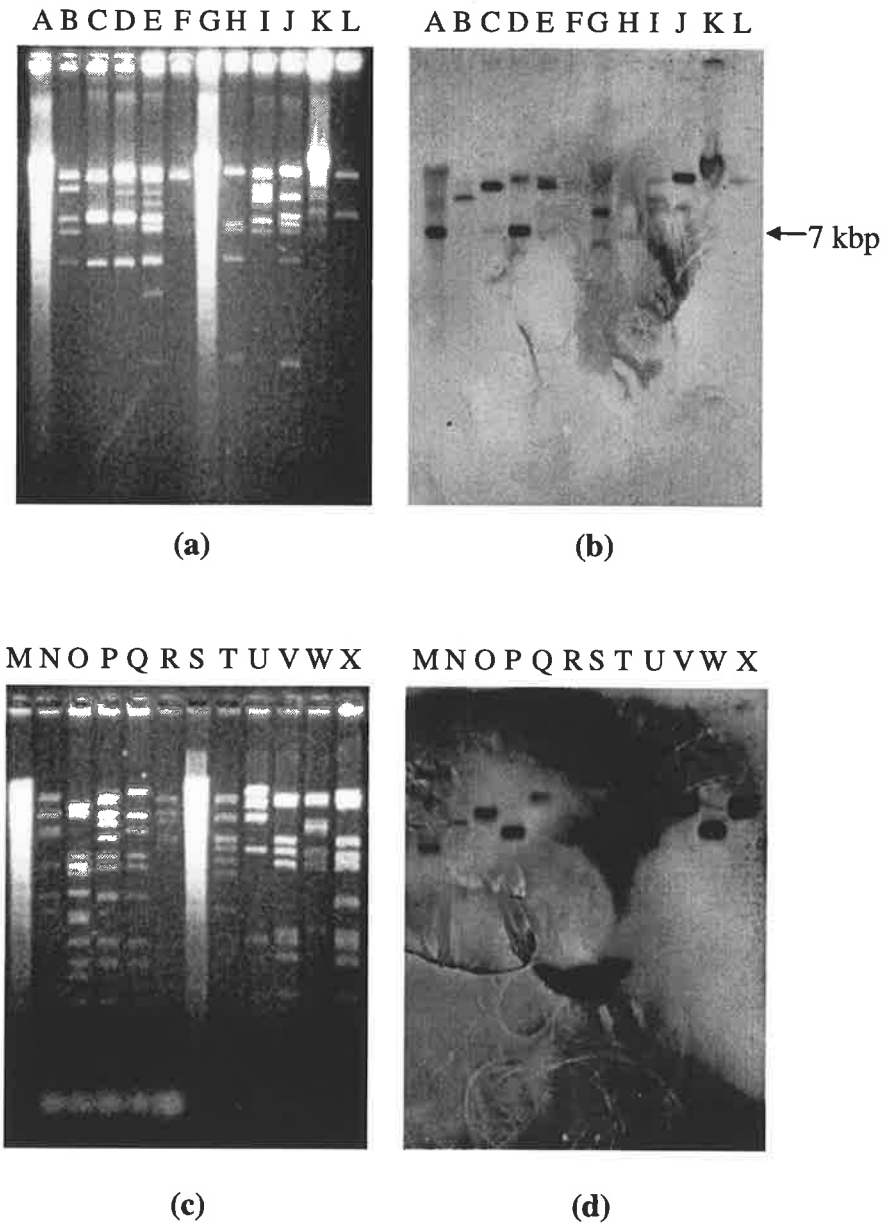


Figure 4.9. Cosmid DNA digested with *Apa* I, *Bgl* II, *Eco* RI and *Eco* RV respectively and Southern hybridisation analysis using Digoxigenin labelled $phaA_{int}$ as a probe.

Lanes **A**, **B**, **C**, **D**, **E**, **F** are *Apa* I digested *Pseudomonas* strain 10c-1-3 chromosomal DNA, cosmids (36)H9, (44)H11, (50)D6, (52)H5, and (55)G1. Lanes **G**, **H**, **I**, **J**, **K**, **L** are *Bgl* II digested *Pseudomonas* strain 10c-1-3 chromosomal DNA, cosmids (36)H9, (44)H11, (50)D6, (52)H5, and (55)G1. Lanes **M**, **N**, **O**, **P**, **Q**, **R** are *Eco* RI digested *Pseudomonas* strain 10c-1-3 chromosomal DNA, cosmids (36)H9, (44)H11, (50)D6, (52)H5, and (55)G1. Lanes **S**, **T**, **U**, **V**, **W**, **X** are *Eco* RV digested *Pseudomonas* strain 10c-1-3 chromosomal DNA, cosmids (36)H9, (44)H11, (50)D6, (52)H5, and (55)G1.

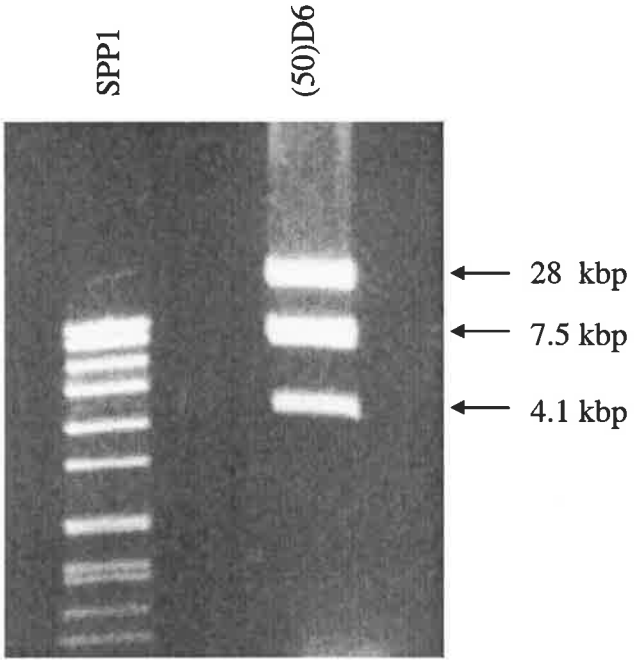


Figure 4.10. Digestion of cosmid (50)D6 with *Apa* I.

```
phaA_int        TGCGGCTCCGGCATGCGCGCTGCGATGTTCGCGCACGACATGCTCATTGCCGGCTCGATC
50D6                      AGTTCGCGTCGCCATG--------CGACAGCCGTCGCGCGGAGCCCATC
36H9                        TTCGCGTCGCCATG--------CGAAAGCCGTCGCGC-GAGCCCATC
consensus                   * ****  ** ***        *** *  *     ** *   * ***

phaA_int        GACGTGATCGTCGCGGGCGGCATGGAGAGCATGACGAATGCGCCCTATCTGCTGCCGAAG
50D6            AGCACGTGCGGCGCGGCGCTCATGTTTTCCTGGCCACCGGCGATCACGATGTCGGCGTCG
36H9            AGCACGTGCGGCGCGGCGCTCATGTTTTCCTGGCCACCGGCGATCACGATGTCGGCGTCG
consensus         *  *  ** *****    ****     *  * *    ***  *    **  * **  *

phaA_int        GCCGCGGCGGCATGCGCATGGGACATGGGCAGGTGCTCGATCA--CATGATCCACGACGG
50D6            CCTGCGATGAT---CGCGTTGGCCGCCAGCATCACCGCCTTCAGGCCCGAGCCGC-ACAC
36H9            CCTGCGATGAT---CGCGTTGGCCGCCAGCATCACCGCCTTCAGGCCCGAGCCGC-ACAC
consensus        * ***  *     *** * ** *    ***    * *  ***  *  ** ** * **

phaA_int        CCTGTC-----
50D6            CTTGTTCATGC
36H9            CTTGTTCATG-
consensus       * *** -

Identity: 71%
```

(a)

Sequences producing High-scoring Segment Pairs:		High Score	Probability P(N)	N
gb|J04987|AFAKTLAACA	A.eutrophus beta-ketothiolase (phb...	567	2.1e-37	1
dbj|AB009273|AB009273	Comamonas acidovorans *phaC*, *phaA* g...	396	2.7e-23	2
gb|AF002013|AF002013	Alcaligenes sp. SH-69 beta-ketothi...	374	6.2e-22	2
gb|AF061446|AF061446	Chromobacterium violaceum polyhydr...	368	7.4e-21	1
gb|L01112|CVNPHB	Chromatium vinosum poly(3-hydroxyb...	333	6.0e-18	1
gb|U88653|PAU88653	Pseudomonas aeruginosa thiolase (p...	320	7.2e-17	1
gb|L01113|TSSPHBAC	Thiocystis violacea beta-ketothiol...	306	1.1e-15	1
gb|S54369|S54369	phbCAB operon: phbA=beta-ketothiol...	306	1.1e-15	1
gb|U29581|ECU29581	Escherichia coli K-12 genome; appr...	293	1.3e-14	1
gb|AE000368|ECAE000368	Escherichia coli K-12 MG1655 secti...	293	1.3e-14	1
gb|AF026544|AF026544	Ralstonia eutropha beta-ketothiola...	260	7.1e-12	1
dbj|D49362|PDEPHAA	Paracoccus denitrificans genes for...	251	4.0e-11	1
gb|J02631|ZOGTHIZR	Z.ramigera thiolase gene, complete...	227	3.9e-09	1
gb|AE000236|ECAE000236	Escherichia coli K-12 MG1655 secti...	189	5.7e-06	1
dbj|D90777|D90777	E.coli genomic DNA, Kohara clone #...	189	5.7e-06	1
dbj|D90778|D90778	E.coli genomic DNA, Kohara clone #...	189	5.7e-06	1
gb|AE000582|HPAE000582	Helicobacter pylori Section 60 of ...	185	1.2e-05	1
gb|U17226|RMU17226	Rhizobium meliloti beta-ketothiola...	182	2.2e-05	1
gb|AF029365|AF029365	Rhodobacter capsulatus B10 nuo gen...	151	0.0082	1
dbj|D00512|RATACAL	Rattus sp. mRNA for mitochondrial ...	144	0.031	1
dbj|D13921|RATACAS	Rat mitochondrial acetoacetyl-CoA ...	144	0.031	1

(b)

Figure 4.11. DNA sequence analysis of the PCR products from cosmids (36)H9 and (50)D6.

(a) Multiple alignment of the DNA sequence data from *phaA*$_{int}$ and PCR products amplified from cosmids (36)H9 and (50)D6 using primer pair #4033 and #4034. Nucleotide bases identical for all three sequences are marked by asterisks. The consensus sequence is also shown; **(b)** BLASTN analysis of DNA sequence data obtained by PCR amplification from (50)D6 using primer pair #4033 and #4034.

The results are discussed in the following paragraphs.

Although the initial aim was to clone PCR fragments internal to both putative *phaA* and *phaB*, only clones containing the former were obtained. The degenerate primers designed to amplify both fragments were used successfully to amplify both target DNAs from isolates 10c-1-3 and 2-3-2. However, only an internal part of *phaA* gene, $phaA_{int}$, was successfully cloned from isolate 10c-1-3 as plasmid pSGZ1. When the cloned *phaA* internal fragment was sequenced and this DNA sequence data compared to other sequences in the Genbank database, it showed strong similarity to other previously identified *phaA* genes. In particular, the sequence showed strong similarity to that from *Ralstonia eutropha* (65% identity). This information provided strong evidence for the presence of PHA biosynthetic genes in the genome of isolate 10c-1-3.

Plasmid pSGZ1 also served another important purpose in this thesis. The cloned internal *phaA* fragment was labelled and used extensively as a specific DNA probe in Southern hybridisation analysis. Similarly, the degenerate primers used to amplify $phaA_{int}$ proved useful as a PCR detection technique in a variety of other recombinant techniques used in this thesis. Probes based on DNA amplified from known *phaA* genes could have been used, however, it is clear from the sequence analysis described that the significant differences in the DNA sequence may have compromised the utility of the probe in hybridisation analyses.

A *Pseudomonas* strain 10c-1-3 cosmid library containing 4,556 clones was successfully constructed and the PCR techniques described above used in a novel manner to identify clones harbouring the putative *phaA* gene. Using a pooling technique originally described by Gussow and Clackson (1989), 17 clones were putatively identified to contain DNA encoding *phaA*. This was achieved with just 521 separate PCR reactions. This approach was chosen because it offered specificity and freedom from problems associated with colony hybridisation techniques commonly used to screen libraries.

Restriction and Southern hybridisation analysis was used to confirm the presence of *phaA* DNA in 7 randomly selected cosmid clones. In particular this allowed confirmation that cosmid clone (50)D6 harbouring plasmid pCT400 was likely to contain the entire *pha* operon from *Pseudomonas* 10c-1-3. *E. coli* DH5α harbouring pCT400 did not however

facilitate production of PHA intracellular granules. There are several possible reasons for this observation. Firstly, the *E. coli* RNA polymerase may not be able to interact with the native promoter normally responsible for initiating transcription of the *pha* genes, or the promoter may not be present as a result of a truncation occurring during the cosmid cloning procedure. Secondly, the native promoter may not allow efficient transcription in *E. coli* either because *E. coli* DH5α lacked a necessary sigma factor required for activation of transcription or because the promoter sequence was not recognised by *E. coli* sigma factors. A third possible reason is that some cellular products inhibit the mRNA synthesis or the protein synthesis of these *pha* genes.

Intriguingly, when *phaA* DNA derived by PCR amplification from cosmid clones (36)H9 and (50)D6 was sequenced, the data obtained differed from that previously obtained for $phaA_{int}$ derived from pSGZ1. This difference suggested that *Pseudomonas* 10c-1-3 contains at least 2 loci with similarity to *phaA*. Nevertheless, that is sufficient similarity between $phaA_{int}$ derived from pSGZ1 and the *phaA* regions located on the cosmid clones to allow use of the former as a DNA probe in hybridisation analysis.

Finally, work described in this chapter showed that the restriction endonuclease *Apa* I would be useful as means of subcloning pCT400 and isolating the *pha* genes as small well defined recombinant clones. This work is described in the next chapter.

Chapter 5

Subcloning, DNA sequence analysis of *pha* genes and construction of plasmids for expression of PHA

5.1 Introduction and Strategy

Chapter 4 presented results of work designed to clone *pha* biosynthetic genes from the soil isolate 10c-1-3. This chapter describes the genetic and sequence analysis of *pha* related genes on cosmid (50)D6 (pCT400) and partial characterisation of the genes involved in PHA biosynthesis.

The basic subcloning strategy used was as follows. Digestion of pCT400 with *Apa* I allowed subcloning of three fragments that covered the whole DNA insert in this cosmid (Figure 4.10). These three DNA fragments were identified with Southern hybridisation using the digoxigenin labelled $phaA_{int}$, digoxigenin labelled *pha* genes from *R. eutropha* as probes, PCR amplification of the $phaA_{int}$, DNA sequencing and BLASTN analysis to determine which of the three DNA fragments contained *pha* genes. The *pha* genes were then subcloned and used to construct PHA production plasmids.

In addition the production of PHA by recombinant *E. coli* strains harbouring these genes was examined. High yield production of PHA by recombinant clones depends on high levels of expression of *pha* genes (Section 1.7.3.2). This in turn requires stable maintenance of the *pha* genes. In this chapter, the effect of recombinant host strain on PHA production and use of the *parB* locus as a mechanism for maintenance of plasmid stability in *E. coli* cell population was examined.

5.2 Results

5.2.1 Construction of plasmids pCT401, pCT402 and pCT403, and the detection of the presence of the *pha* genes in these plasmids

In Chapter 4, it was shown that the insert in plasmid pCT400 could be separated into three *Apa* I fragments. These fragments were isolated by agarose gel electrophoresis. The 4.1 and 28 kbp fragments were subcloned into *Apa* I digested, shrimp alkaline phosphatase (SAP) treated vector pGEM7Zf(+) and named pCT401, pCT402 respectively. pCT403 was constructed by religating the 7.5 kbp *Apa* I fragment. Details of the construction are shown in Figure 5.1. Plasmid DNA was used to transform competent *E. coli* DH5α and Ap^r and *LacZ* negative white clones were selected. Clones containing pCT401, pCT402 and pCT403 were confirmed by restriction endonuclease digestion of

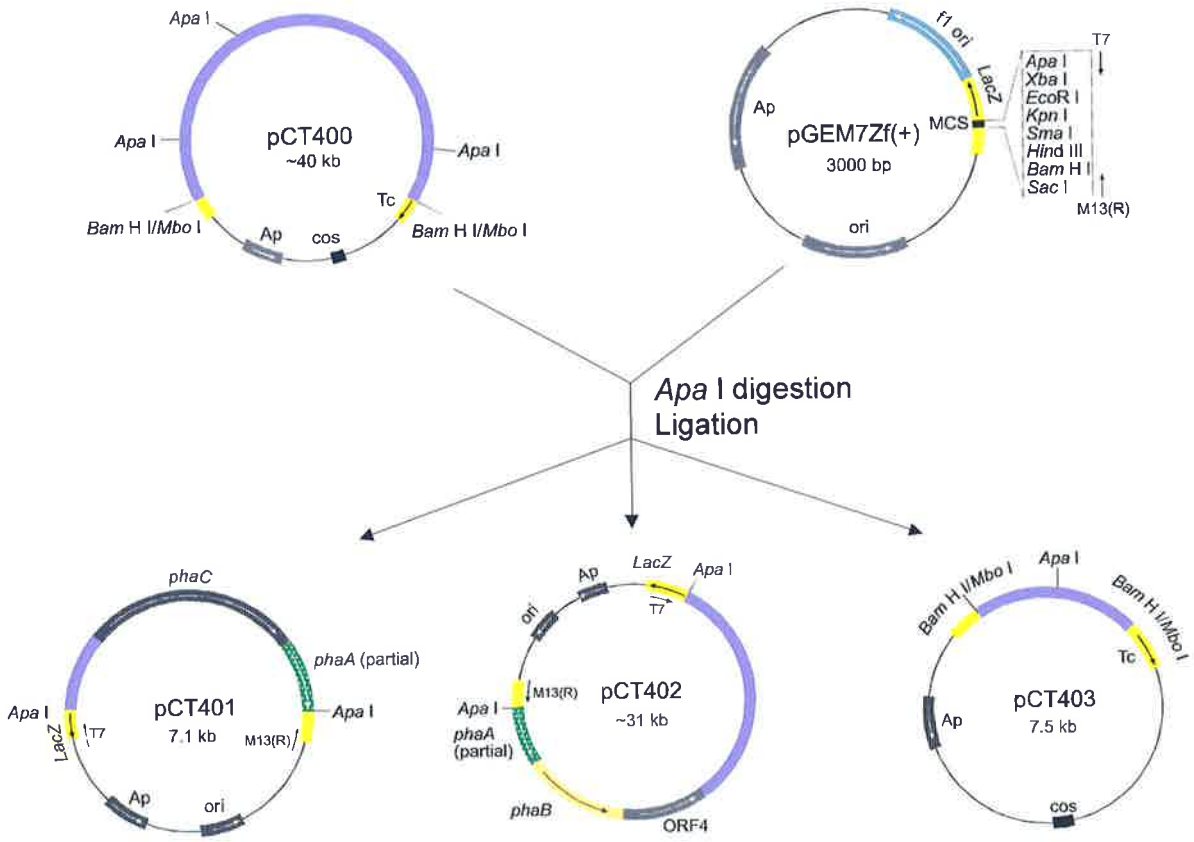


Figure 5.1. Subcloning of plasmid pCT400 and construction of plasmids pCT401, pCT402 and pCT403.

PCT400 was completely digested with *Apa* I, resulting in three *Apa* I fragments with sizes of ~4.1 kbp, ~7.5 kbp and ~28 kbp (See Figure 4.10). These three fragments were gel purified, the ~4.1 kbp and ~28 kbp fragments were ligated to SAP treated pGEM7Zf(+) respectively and the ~7.5 kbp fragment was religated. These constructs were used to transform *E. coli* DH5α. Transformants were plated on NA which contained Ap, IPTG and X-gal. White, Ap^r clones of plasmid constructs were confirmed by restriction endonuclease enzyme digestion. Clones harbouring the ~4.1 kbp fragment (pCT401), the ~28 kbp fragment (pCT402) and the self-ligated ~7.5 kbp fragment of pCT400 (pCT403) were selected and used for further work.

plasmid DNA preparations (Figure 5.2).

In order to determine which of the three clones encoded *pha* genes, plasmid DNA preparations were subjected to PCR amplification using primers (#4033 and #4034) designed to amplify a part of *phaA*. Plasmid pCT400 was used as a positive control (Figure 5.3). PCR analysis resulted in formation of 200 bp fragments from pCT401 and pCT400. No DNA fragments were produced from pCT402 or pCT403, indicating that plasmid pCT401 contained DNA encoding at least a part of *phaA*.

To determine the orientation of the *pha* genes in the 4.1 kbp DNA insert in pCT401, separate PCR amplifications were carried out using one primer which initiated DNA replication from the vector [M13 (R)] and one of the primers used to amplify $phaA_{int}$ (#4033 and #4034) to form primer pairs #4033 and M13 (R) and #4034 and M13 (R). Thermocycling was carried out as follows: one incubation at 95 °C for 3 min, followed by 30 cycles of 95 °C for 30 s, 55 °C for 60 s, and 72 °C for 5 min. A final extension was carried out by incubation at 72 °C for 4 min. A single amplicon of 820 bp was produced when primer pair M13 (R) and #4033 were used. This fragment was predicted to comprise the 200 bp $phaA_{int}$ plus another 500 bp fragment of this region. This indicated that the insertion orientation of DNA encoding *pha* genes was as described in Figure 5.1. If the *phaA* gene of the isolate 10c-1-3 is of similar size to the *phaA* gene from *R. eutropha* (~1.2 kbp), this 820 bp PCR product indicated that pCT401 contains insufficient DNA to encode a complete *phaA* gene. Furthermore, if the *pha* biosynthetic genes are arranged in an operon with the same order and size as that of *R. eutropha*, then plasmid pCT401 should contain sufficient DNA to encode a complete *phaC*. This also means a further 1.3 kbp DNA fragment encoding the rest of *phaA* and *phaB* should be located on either pCT402 or pCT403 (Figure 5.4).

The presence of *phaC* in pCT401, and 3' region of *phaA* and *phaB* in pCT402 or pCT403 was confirmed by Southern hybridisation analysis of *Apa* I digested plasmid DNA using Digoxigenin labelled *phaC* and *phaAB* DNA probes prepared from *Pst* I digested plasmid p4A (Figure 5.5). pCT401, pCT402, pCT403 and the positive control pCT400 were completely digested with *Apa* I, transferred to a Nylon membrane and then hybridised with the probes Digoxigenin labelled *phaC* and Digoxigenin labelled *phaAB* respectively

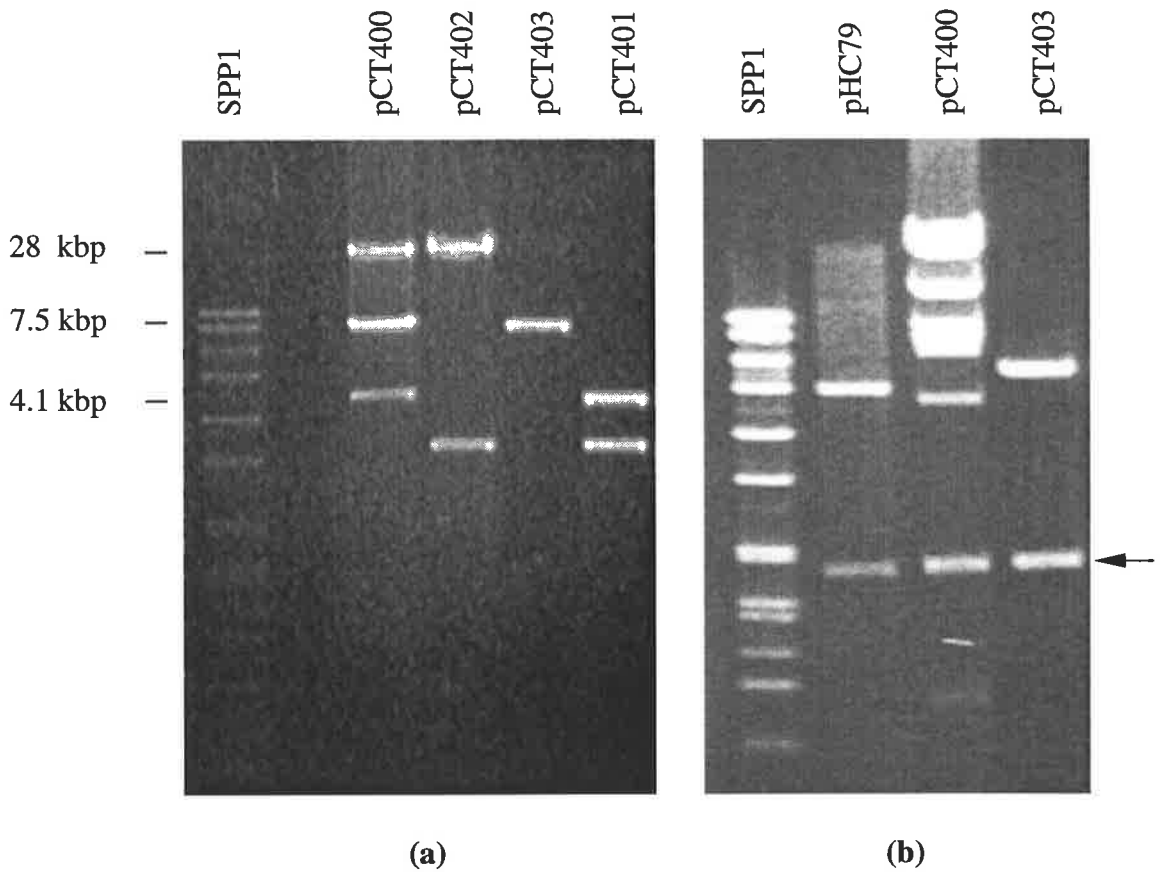


Figure 5.2. Confirmation of the plasmid constructs, pCT401, pCT402 and pCT403 by restriction endonuclease enzyme digestion.

(a) *Apa* I digestion of plasmids pCT400, pCT401, pCT402 and pCT403 showed these latter three constructs contained the ~4.1, ~7.5 and ~28 kbp DNA fragments purified from *Apa* I digest of pCT400. **(b)** *Bgl* II digestion of pHC79, pCT400 and pCT403 showed that they shared a 1.7 kbp pHC79 fragment (as indicated with the above arrow) and indicated that pCT403 contained the pHC79 vector and a truncated *Pseudomonas* strain 10c-1-3 chromosomal DNA insert of 1.1 kbp in size flanked by the *Mbo* I/*Bam* HI sites as shown in Figure 5.1.

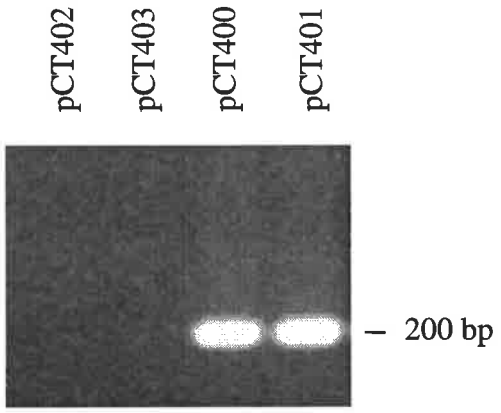


Figure 5.3. PCR amplification of a *phaA* fragment from plasmids pCT401, pCT402 and pCT403.

PCR amplifications were carried out using primer pair #4033 and #4034. The 200 bp fragments amplified from pCT400 and pCT401 indicated that the *Apa* I insert cloned in pCT401 contained $phaA_{int}$.

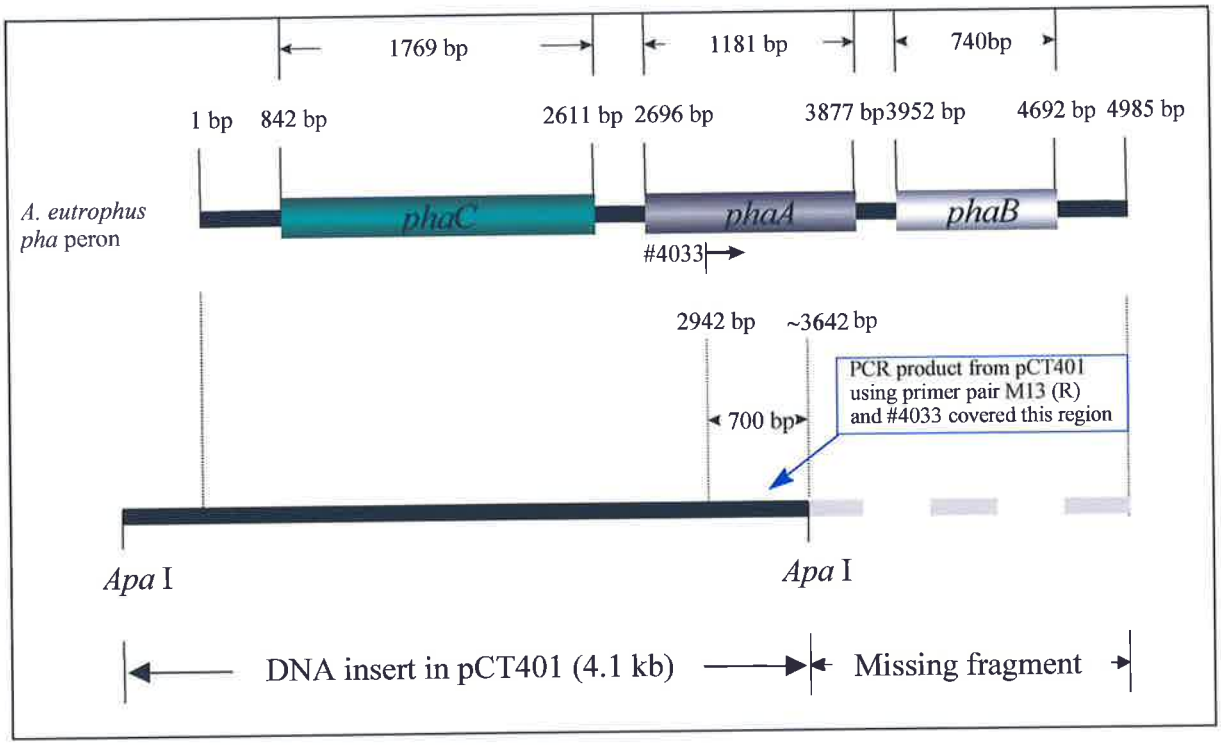


Figure 5.4. Comparison of the insert in pCT401 with the *R. eutropha pha* operon.

The 820 bp PCR product from pCT401 using primer pair M13(R) and #4033 covered a ~700 bp region of *phaA* and indicated the insertion orientation of the *pha* genes in pCT401 was as shown in Figure 5.1 and there was only a partial *phaA* in pCT401. If this PCR fragment was amplified from *R. eutropha pha* operon, it would cover the 2942 - ~3642 bp region. Given the *Pseudomonas* strain 10c-1-3 *pha* genes are also in an operon with a similar size to that of *R. eutropha*, the insert in pCT401 (~4.1 kbp) would contain the whole *phaC* and a surplus upstream fragment. This comparison also indicated the absence of the 3643 - 4985 bp region in *R. eutropha pha* operon in pCT401.

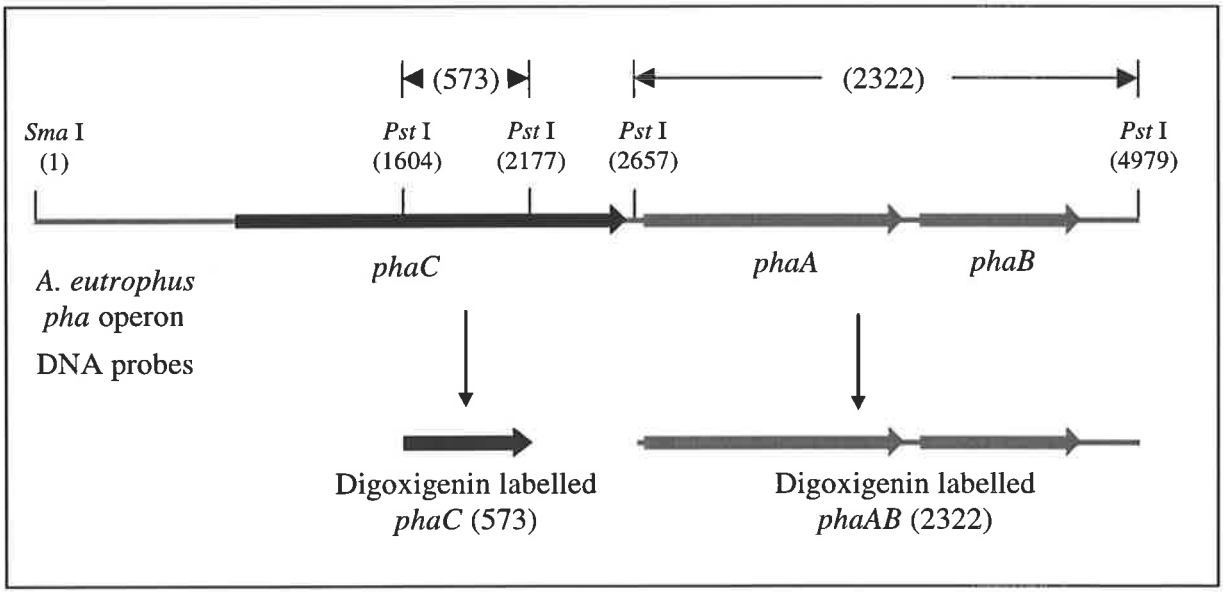


Figure 5.5. Preparation of Digoxigenin labelled *phaC* and *phaAB* DNA probes from *R. eutropha pha* operon.

Plasmid p4A was completely digested with *Pst* I, and the 573 bp fragment containing a part of *phaC* and the 2322 bp fragment containing a partial *phaA* and the whole of *phaB* were gel purified and Digoxigenin labelled. Labelled probes were then used for the identification of the presence of *phaC* in pCT401, and a part of *phaA* and the *phaB* in pCT402 or pCT403. Numbers in brackets refer to restriction sites or sizes in bp.

(Figure 5.6). Figure 5.6b shows that the 2.3 kbp *Pst* I Dig-*phaAB* hybridised with the 28 kbp *Apa* I fragment from pCT400, the 28 kbp *Apa* I insert of pCT402 and the 4.1 kbp *Apa* I insert of pCT401. This confirmed that the 5' and 3' *phaA* fragments are located on pCT401 and pCT402 respectively, and suggested that *phaB* is located on pCT402. The 0.5 kbp *Pst* I *phaC* probe hybridised to *Apa* I digested pCT400 and pCT401 (Figure 5.6c). Neither probe hybridised to pCT403 restriction fragments. Thus *phaC* is predicted on pCT401.

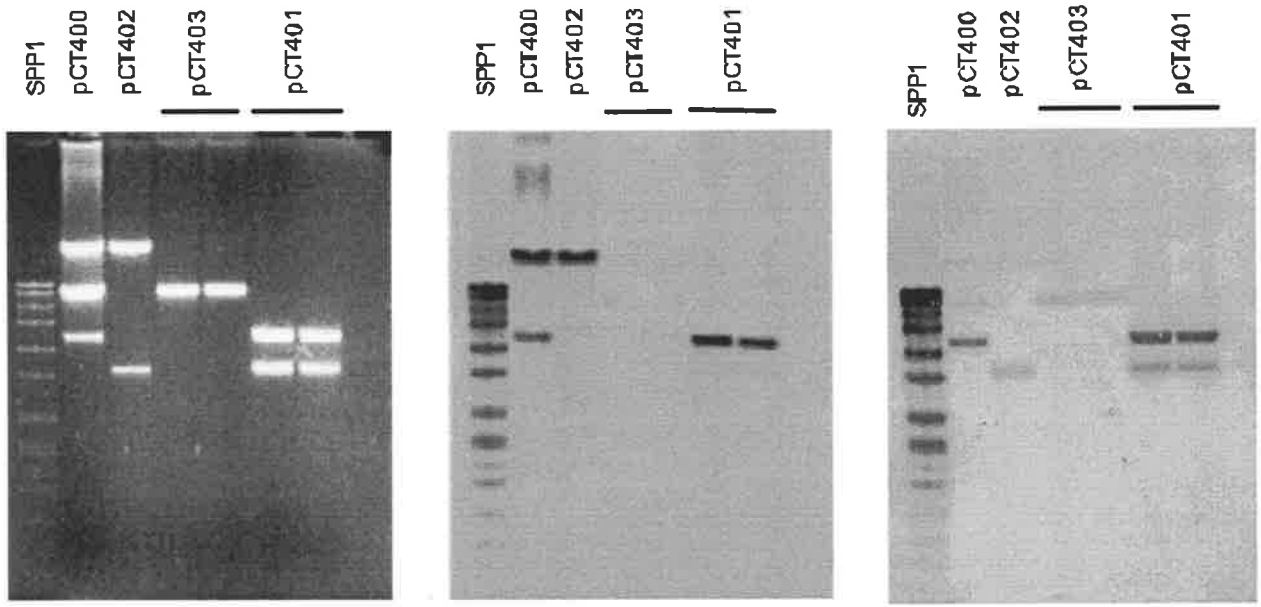
The presence of a part of *phaA* in pCT401 and in pCT402 was also shown by DNA sequence analysis. The inserts in pCT401 and pCT402 were sequenced from the T7 and M13(R) priming sites of the pGEM7Zf(+) vector. BLASTN analysis of DNA sequence data obtained from pCT401 and pCT402 using primer M13(R) confirmed similarity with other *phaA* DNA sequences in the Genbank databank. Similarly, DNA sequence data obtained from pCT401 using primer T7 was similar to *phaC*s from other bacterial sources. Sequence data produced from pCT402 using primer T7 showed no similarity with any *pha* genes from other bacterial sources (data not shown).

5.2.2 Nucleotide sequence analysis

DNA sequencing was achieved by 'sequence walking' in which successive sequence runs were used to design sequencing primers for the next sequencing reactions. Using this technique, the complete sequence of the 4.1 kbp insert in pCT401 and the sequence of a 2.5 kbp segment of the *Apa* I insert in pCT402 were obtained. The assembled DNA sequences were analysed for potential open reading frames and the annotated sequences are shown in Figure 5.7a.

The sequence data covering the *Apa* I site on pCT400, which joints the inserts cloned into pCT401 and pCT402 was confirmed by direct PCR product sequencing using primer set #5577 (F) and #5578 (R) and pCT400 as template.

Figure 5.7a shows the complete annotated sequence encoding the entire *pha* biosynthetic region of isolate 10c-1-3. Corresponding translated sequences are also shown. The sequence data was derived from plasmids pCT401 and part of the insert of pCT402. Four open reading frames are described (Figure 5.7b). Three of these (*phaCAB*) show high



(a) (b) (c)

Figure 5.6. Detection of *pha* genes in pCT401, pCT402, and pCT403 by Southern hybridisation.

A.: *Apa* I digested pCT400, pCT401, pCT402, and pCT403.
B: Southern hybridisation identified the presence of *phaA* and *phaB* in plasmids pCT400, pCT401 and pCT402, using the Digoxigenin labelled *phaAB* DNA probe (Figure 5.5).
C.: Southern hybridisation identified the presence of *phaC* in plasmids pCT400 and pCT401, using the Digoxigenin labelled *phaC* probe (Figure 5.5).

Apa I

```
1     GGG CCC GCC TTC GCA GGC TGC GCA GCC GGT CAA GGC CGA CGC GAG CGT GAC
52    CGG GAG CAC GGG CGC GGT GTG CGT CGT GAT GGT GGC GGA CTG CCT GCC GGT
103   GCT CTT GTG CGA CGA GTC CGG ACG CGC GGT CGG CGC GGC GCA TGC GGG CTG
154   GCG GGG CCT GGC CGC GGG TAT CGT GGA GAA GAC GGC CGA CCG GGT GGC CAC
205   GCT CGC CGG CGC CAG CAC GGC GAG CCT GCA TGC GTA CCT CGG GCC GGC GAT
256   CGG CCC GGC GGC GTT CGA AGT GGG CGA GGA CGT GCT CGA CGC GTT CGT CTC
307   GGC CGC CGA TGC CGC GCA GCG CGA CGC GAC CGC CGG CGC CTT CAC GCG GGC
358   GCA GGC GCC GCA CAA GTA CTT CTC GGA CAT CTA TGC GCT CGC GCG GCT GCG
409   GCT TGC AGT CGC GGG CGT CGA CCC TGC GCG CGT GCA GGG CGG CAC GCA CTG
460   CAC GGT CAC CGA GCG CGA ACG CTT CTA TTC GTA TCG GCG GGA CCG CGT GAC
511   GGG CCG CAT GGC CGC GAT GAT CTG GCT CGC CGA CTG AAA CCT GCC GCG CCC
562   GGG GGC CGC CCA CGC GAG AAG GGT TGC CCG GGC ATT GCG AAG CGG CCT TAT
613   TCG GCA TCG AGC GAT CTA AGG CGG CAT CAA GCG GCG GCA TCG AGC GAT CTC
664   AGG CGG CAT CAA GCG GCG TCA AGC GGA CTT CAG GCG GCC TTT TAT CGG CCG
715   CGA GCG GAT TTG TCA GAC ACG GTT TCG CGA ATC CCG GGT GAT CGC ATC CGG
766   AAT CCC GCT TCG ACC GGC ACG TCG TAC CGG CCG GGG GTT TTC TGC ACT GCT
817   CGA TTT CGG TTG ATA TTC CGC GTC CGC GCA GCG TCC GCG AGG CCC CGG CCC

                                                    promoter(pha operon)
868   GCC GAC GGG AAG GGC CGT GCA TGC GTG CTC ACG GAA GAT GAA ATG ATC GCA
                                                          -35
919   TGG GGA AAA CGA TAA TTA AAA AAA TAT CGC ACT GCG GCA AAA TCG GGG CAA
                    -10
        RBS(phaC)   start(? phaC)
970   AGC ATT GAC ATG CCT GAT CCT CGG GAG CAA GAA TGT TCC CGG AAT GTC AAA
                  Met Pro Asp Pro Arg Glu Gln Glu Cys Ser Arg Asn Val Lys

1021  TAT CGG GCA CGA CGT GGC AAG GTC ACC TTC CAC CGC GCA GTA TCG CGT CTC
      Tyr Arg Ala Arg Arg Gly Lys Val Thr Phe His Arg Ala Val Ser Arg Leu

1072  ATT GCG CGC TGG TCT CGG CGA GCC CTA CAC AGC GCC TGC AAC AGC AAT CAG
      Ile Ala Arg Trp Ser Arg Arg Ala Leu His Ser Ala Cys Asn Ser Asn Gln

                                              start(? phaC)
1123  TCT TTC GAG ACT CAC AGG TCA AAA GCG GGT ATG GCT GCC TCC AAA ACT TCT
      Ser Phe Glu Thr His Arg Ser Lys Ala Gly Met Ala Ala Ser Lys Thr Ser

1174  TCT TCA TCT TCC CGC CCA GAT ACC GCC GAC AGC ACA TCG GCC GGA ACG CGT
      Ser Ser Ser Ser Arg Pro Asp Thr Ala Asp Ser Thr Ser Ala Gly Thr Arg

1225  CCG CCG ACA CGG CCG GCG TGC AGC AAG TGT TCG AGC AAT GGC TCA ACG CAT
      Pro Pro Thr Arg Pro Ala Cys Ser Lys Cys Ser Ser Asn Gly Ser Thr His

1276  GGC GTT CGG TCG CCG ATC CGT CAC AGT GGA CGA CGC TCG CGC AGC AGG CGC
      Gly Val Arg Ser Pro Ile Arg His Ser Gly Arg Arg Ser Arg Ser Arg Arg

1327  AGC CGG GTC AGC CGA ATG GCC CCG ACG GCG GCA ATG CCG CAG CGT CGG CAG
      Ser Arg Val Ser Arg Met Ala Pro Thr Ala Ala Met Pro Gln Arg Arg Gln

1378  CGG CAT CGG CCA ATC CGT TCG CCG CTT TCG CAG CCG TTC CCG GAA TGC CGG
      Arg His Arg Pro Ile Arg Ser Pro Leu Ser Gln Pro Phe Pro Glu Cys Arg

1429  CGG CGC CGT TCG CGA TGC CGA CGG TGC CTG ATT TCG CGA AGC TCG GCA CGA
      Arg Arg Arg Ser Arg Cys Arg Arg Cys Leu Ile Ser Arg Ser Ser Ala Arg
```

```
1480 TGC CCG ACT TCT CGA AGC TCG CGG GCA GCA TGC CCG GCT TCG GCG CCG CGA
     Cys Pro Thr Ser Arg Ser Ser Arg Ala Ala Cys Pro Ala Ser Ala Pro Arg

1531 TGC CTG CCA TGC CGG CGA TGC CGC AGA TTC CGG GCG CGG CCA TCG CGC CGG
     Cys Leu Pro Cys Arg Arg Cys Arg Arg Phe Arg Ala Arg Pro Ser Arg Arg

1582 AGC GTC TTC AGC AGA TTG CAG GGC GAC TAT TCG CGC GAT GTG ATC GAC CTG
     Ser Val Phe Ser Arg Leu Gln Gly Asp Tyr Ser Arg Asp Val Ile Asp Leu

1633 CTC AAG CAG GCG AGC GCG CAG AGC ATC GAC CCC GCA GCG CTG AAG GAC CGG
     Leu Lys Gln Ala Ser Ala Gln Ser Ile Asp Pro Ala Ala Leu Lys Asp Arg

1684 CGC TTC AGC ACG ACG GCG TGG CAG TCC ACG CCG GCT AAC GCG TTC ACG GCC
     Arg Phe Ser Thr Thr Ala Trp Gln Ser Thr Pro Ala Asn Ala Phe Thr Ala

1735 GCG TGG TAT CTG CTC AAC GCG CGC TAT CTC CAG GAA CTC GCC GAC GCC GTC
     Ala Trp Tyr Leu Leu Asn Ala Arg Tyr Leu Gln Glu Leu Ala Asp Ala Val

1786 GAG GCC GAT CCC AAG ACG CGC GAG CGC ATC CGC TTC ACG GTG CAG CAG TGG
     Glu Ala Asp Pro Lys Thr Arg Glu Arg Ile Arg Phe Thr Val Gln Gln Trp

1837 ACG GCC GCG GCC TCG CCG AGC AAC TTT CTC GCG TTC AAT CCC GAA GCG CAG
     Thr Ala Ala Ala Ser Pro Ser Asn Phe Leu Ala Phe Asn Pro Glu Ala Gln

1888 CAG ACG CTC ATC GAG AGC AAG GGG GAG AGT CTG CGC CAG GGC ATG CTG AAT
     Gln Thr Leu Ile Glu Ser Lys Gly Glu Ser Leu Arg Gln Gly Met Leu Asn

1939 CTG CTG CAC GAC ATG CAG CGC GGC AAG ATC TCG CAA TCC GAC GAG TCG CGT
     Leu Leu His Asp Met Gln Arg Gly Lys Ile Ser Gln Ser Asp Glu Ser Arg

1990 TTC GTG GTC GGC AAG AAC ATC GCG ACG ACG GAA GGG TCG GTC GTG TTC GAG
     Phe Val Val Gly Lys Asn Ile Ala Thr Thr Glu Gly Ser Val Val Phe Glu

2041 AAC GAC CTG CTG CAA CTG ATC CAG TAC AAG CCG CAT ACG GAA GAA GTC TTC
     Asn Asp Leu Leu Gln Leu Ile Gln Tyr Lys Pro His Thr Glu Glu Val Phe

2092 GAG CGG CCG CTC CTG ATC GTG CCG CCG TGC ATC AAC AAG TTG TAC ATC CTC
     Glu Arg Pro Leu Leu Ile Val Pro Pro Cys Ile Asn Lys Leu Tyr Ile Leu

2143 GAC CTG CAG CCG CAG AGC TCG CTC GTC GCC CAT GCG CTC GAT GCG GGC CAT
     Asp Leu Gln Pro Gln Ser Ser Leu Val Ala His Ala Leu Asp Ala Gly His

2194 CAG GTG TTC ATC CTG TCC TGG CGC AAC GCG GAT CAG TCG ATC GCG CAC AAG
     Gln Val Phe Ile Leu Ser Trp Arg Asn Ala Asp Gln Ser Ile Ala His Lys

2245 ACC TGG GAC GAC TAC GTG CAG GAA GGC GTG CTC GAT CCG ATC GAA GCC GTC
     Thr Trp Asp Asp Tyr Val Gln Glu Gly Val Leu Asp Pro Ile Glu Ala Val

2296 AAG GCG ATC ACC GGG CGC GAG CAG ATC AAC ACG CTC GGC TTT TGC ATC GGC
     Lys Ala Ile Thr Gly Arg Glu Gln Ile Asn Thr Leu Gly Phe Cys Ile Gly

2347 GGA ACG ATC CTC GCT ACC GCG CTT TCG GTG GCG GCG GCG CGC GGC GAG CAC
     Gly Thr Ile Leu Ala Thr Ala Leu Ser Val Ala Ala Ala Arg Gly Glu His

2398 CCG GCG GCG TCG ATG ACG CTC CTC ACC GCG ATG CTC GAC TTC TCC GAC ACC
     Pro Ala Ala Ser Met Thr Leu Leu Thr Ala Met Leu Asp Phe Ser Asp Thr

2449 GGC GTG CTC GAC GTC TTC GTC GAC GAG GCG CAC GTG CAG ATG CGC GAG CAG
     Gly Val Leu Asp Val Phe Val Asp Glu Ala His Val Gln Met Arg Glu Gln

2500 ACG ATC GGC GGC AAG GGC GGC ACG CCG ACC GGG CTC ATG CGC GGC TTC GAG
     Thr Ile Gly Gly Lys Gly Gly Thr Pro Thr Gly Leu Met Arg Gly Phe Glu

2551 TTC GCG AAC ACG TTC TCG TAC CTG CGC CCG AAC GAT CTG GTG TGG AAC TAC
     Phe Ala Asn Thr Phe Ser Tyr Leu Arg Pro Asn Asp Leu Val Trp Asn Tyr
```

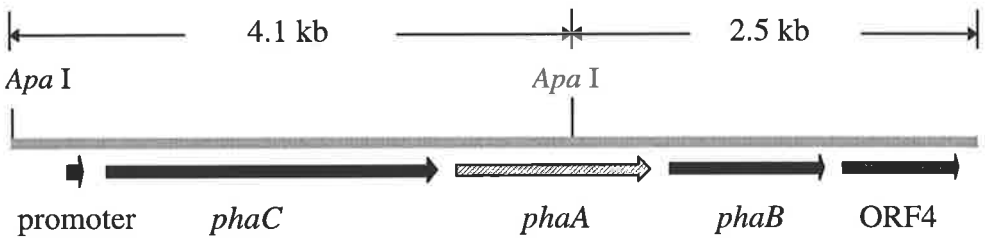
```
2602 GTC GTC GAC AAC TAC CTG AAG GGC GCA ACG CCG CAG GCG TTC GAT CTG CTC
     Val Val Asp Asn Tyr Leu Lys Gly Ala Thr Pro Gln Ala Phe Asp Leu Leu

2653 TTC TGG AAC AGC GAC TCG ACC AAT CTG CCG GGG CCG ATG TGC GTC TGG TAC
     Phe Trp Asn Ser Asp Ser Thr Asn Leu Pro Gly Pro Met Cys Val Trp Tyr

2704 CTG CGC AAC ACG TAT CTG GAA AAC AAG CTG CGC GAG CCG GGC AAG GTC ACG
     Leu Arg Asn Thr Tyr Leu Glu Asn Lys Leu Arg Glu Pro Gly Lys Val Thr

2755 ACG TGC GGC GAA CCG GTC GAT CTG TCG CGC ATG GAC GTG CCC ACT TTC ATT
     Thr Cys Gly Glu Pro Val Asp Leu Ser Arg Met Asp Val Pro Thr Phe Ile

2806 TAC GGT TCA CGG GAG GAC CAT ATC GTG CCC TGG ACG TCG GCG TAC GCG TCG
     Tyr Gly Ser Arg Glu Asp His Ile Val Pro Trp Thr Ser Ala Tyr Ala Ser

2857 GCG CCG CTC CTT GCC GGG CCG CAG AAA TTC GTG CTC GGC GCT TCG GGG CAC
     Ala Pro Leu Leu Ala Gly Pro Gln Lys Phe Val Leu Gly Ala Ser Gly His

2908 ATT GCG GGC GTC ATC AAC CCG CCG GCG AAG AAC AAG CGC AGC TTC TGG ATT
     Ile Ala Gly Val Ile Asn Pro Pro Ala Lys Asn Lys Arg Ser Phe Trp Ile

2959 TTC GAC ACC GAC GAC AAG GCG TTG CCG GCG GAC CCG GAC GCC TGG CTC GAA
     Phe Asp Thr Asp Asp Lys Ala Leu Pro Ala Asp Pro Asp Ala Trp Leu Glu

3010 GCC GCG ACG GAG CAT CCG GGA AGC TGG TGG CCC ACC TGG ACC GCT TGG CTC
     Ala Ala Thr Glu His Pro Gly Ser Trp Trp Pro Thr Trp Thr Ala Trp Leu

3061 GAC GGC TAC GCG GGC AAG AAG ACG AAG GCG CCG GCG GCG CCC GGG TCG AAG
     Asp Gly Tyr Ala Gly Lys Lys Thr Lys Ala Pro Ala Ala Pro Gly Ser Lys

3112 CAG TAC CCT GTC ATC GAG CCG GCC CCC GGC CGC TAT GTG CTG CAG CGT GAC
     Gln Tyr Pro Val Ile Glu Pro Ala Pro Gly Arg Tyr Val Leu Gln Arg Asp

          stop(phaC)
3163 CAG TAA CCG TTT CGT CGT CCG GCC GTG CGC GAC GCG CGC GGC ATG ACG ACC
     Gln ***

                                        RBS(phaA)      start(phaA)
3214 CCG CTT TAA CAT TCG TTG CCA GAG GAAA CTGAAA ATG ACT GAC GTA GTG ATC
                                                   Met Thr Asp Val Val Ile

3266 GTA TCG GCC GCA CGT ACC GCG GTC GGC AAA TTC GGC GGG TCG CTC GCG AAG
     Val Ser Ala Ala Arg Thr Ala Val Gly Lys Phe Gly Gly Ser Leu Ala Lys

3317 ATC GCA GCA CCC GAA TTG GGC GCA ACC GTG ATC CGC GCG GTA TTG GAG CGC
     Ile Ala Ala Pro Glu Leu Gly Ala Thr Val Ile Arg Ala Val Leu Glu Arg

3368 GCG AAC CTG AAA CCC GAG CAG GTG AGC GAA GTG ATC CTC GGC CAG GTG CTC
     Ala Asn Leu Lys Pro Glu Gln Val Ser Glu Val Ile Leu Gly Gln Val Leu

3419 GCC GCC GGC TCC GGA CAG AAT CCC GCA CGC CAG TCG CTC ATC AAG GCG GGG
     Ala Ala Gly Ser Gly Gln Asn Pro Ala Arg Gln Ser Leu Ile Lys Ala Gly

3470 CTG CCC TCG GCC GTT CCC GGC ATG ACG ATC AAC AAG GTC TGC GGC TCG GGC
     Leu Pro Ser Ala Val Pro Gly Met Thr Ile Asn Lys Val Cys Gly Ser Gly

3521 CTG AAG GCG GTG ATG CTG GCG GCC AAC GCG ATC ATC GCA GGC GAC GCC GAC
     Leu Lys Ala Val Met Leu Ala Ala Asn Ala Ile Ile Ala Gly Asp Ala Asp

3572 ATC GTG ATC GCC GGT GGC CAG GAA AAC ATG AGC GCC GCG CCG CAC GTG CTG
     Ile Val Ile Ala Gly Gly Gln Glu Asn Met Ser Ala Ala Pro His Val Leu

3623 ATG GGC TCG CGC GAC GGC TTT CGC ATG GGC GAC GCG AAG CTG ATC GAC TCG
     Met Gly Ser Arg Asp Gly Phe Arg Met Gly Asp Ala Lys Leu Ile Asp Ser

3674 ATG ATC GTC GAC GGC CTG TGG GAC GTG TTC AAC CAG TAC CAC ATG GGC ACG
     Met Ile Val Asp Gly Leu Trp Asp Val Phe Asn Gln Tyr His Met Gly Thr
```

```
3725 ACC GCC GAG AAC GTC GCG AAG GAA AAC GGC ATC TCG CGC GAG GAT CAG GAC
     Thr Ala Glu Asn Val Ala Lys Glu Asn Gly Ile Ser Arg Glu Asp Gln Asp

3776 AAG TTC GCG GCG CTC TCG CAG AAC AAG GCC GAA GCC GCG CAG AAG TCG GGC
     Lys Phe Ala Ala Leu Ser Gln Asn Lys Ala Glu Ala Ala Gln Lys Ser Gly

3827 CGC TTC AAC GAG GAG ATC GTG TCG GTC GAC ATT CCC CAG CGC AAG GGA GAC
     Arg Phe Asn Glu Glu Ile Val Ser Val Asp Ile Pro Gln Arg Lys Gly Asp

3878 CCC GTC AAG TTC GCC ACC GAC GAA TTC GTG CGC CAC GGC GTG ACG GCC GAA
     Pro Val Lys Phe Ala Thr Asp Glu Phe Val Arg His Gly Val Thr Ala Glu

3929 GCG CTC GCG AGC CTG AAG CCC GCC TTC TCG AAG GAA GGC ACG GTG ACG GCC
     Ala Leu Ala Ser Leu Lys Pro Ala Phe Ser Lys Glu Gly Thr Val Thr Ala

3980 GCC AAC GCA TCG GGC CTG AAC GAC GGC GCG GCC GCC GTG ATC GTG ATG TCG
     Ala Asn Ala Ser Gly Leu Asn Asp Gly Ala Ala Ala Val Ile Val Met Ser

4031 GCG AAG AAG GCC GAG GCG CTC GGC CTC ACG CCG CTC GCG CGC ATC AAG GCC
     Ala Lys Lys Ala Glu Ala Leu Gly Leu Thr Pro Leu Ala Arg Ile Lys Ala
                                                          Apa I
                                                            ↓
4082 TAC GCG AAC GCG GGC GTC GAT CCG AAG GTG ATG GGC ATG GGC CCG GTG CCG
     Tyr Ala Asn Ala Gly Val Asp Pro Lys Val Met Gly Met Gly Pro Val Pro

4133 GCA TCG AAG CGC TGT CTG GAG CGC GCG GGC TGG TCG GTG GGT GAT CTC GAT
4184 Ala Ser Lys Arg Cys Leu Glu Arg Ala Gly Trp Ser Val Gly Asp Leu Asp

4235 CTG ATG GAG ATC AAC GAA GCG TTC GCG GCG CAG GCG CTC GCG GTG CAC AAG
4286 Leu Met Glu Ile Asn Glu Ala Phe Ala Ala Gln Ala Leu Ala Val His Lys

4337 CAG ATG GGC TGG GAC ACG TCG AAG ATC AAC GTC AAC GGC GGC GCG ATC GCG
4388 Gln Met Gly Trp Asp Thr Ser Lys Ile Asn Val Asn Gly Gly Ala Ile Ala

4439 ATC GGG CAC CCG ATC GGC GCG TCC GGT TGC CGG ATC CTC GTC ACG CTG CTG
4490 Ile Gly His Pro Ile Gly Ala Ser Gly Cys Arg Ile Leu Val Thr Leu Leu

4541 CAC GAA ATG CAG AAG CGC GAT GCG AAA AAG GGC CTG GCG TCG CTG TGT ATC
     His Glu Met Gln Lys Arg Asp Ala Lys Lys Gly Leu Ala Ser Leu Cys Ile
                                                          stop(phaA)
4592 GGC GGC GGC ATG GGC GTG GCG CTG GCG GTC GAA CGT CCG TAA GAA CGA AGG
     Gly Gly Gly Met Gly Val Ala Leu Ala Val Glu Arg Pro ***

4643 CTA TAA CCG GCA CGG CCA TGC GGG AAG CGC ACG GGC AGC AGG CGG GCG CTT

                              RBS(phaB)      start(phaB)
4694 CCC GCA TAA CGA AAA TGG AGT GAG GA  ATG ACG AAA CGC ATA GCG TAC GTG
                                        Met Thr Lys Arg Ile Ala Tyr Val

4745 ACG GGC GGC ATG GGC GGC ATT GGC ACA AGC ATC TGC CAG CGT CTG CAT AAG
     Thr Gly Gly Met Gly Gly Ile Gly Thr Ser Ile Cys Gln Arg Leu His Lys

4796 GAC GGC TAT ACG GTC GTC GCG GGT TGC GGC CCG AAC TCT CCG CGC CGT GTC
     Asp Gly Tyr Thr Val Val Ala Gly Cys Gly Pro Asn Ser Pro Arg Arg Val

4847 AAA TGG CTC GAG GAA CAG AAG GCG AAC GGC TAT GAC TTC ATC GCG TCC GAG
     Lys Trp Leu Glu Glu Gln Lys Ala Asn Gly Tyr Asp Phe Ile Ala Ser Glu

4898 GGC AAC GTC GGC GAC TGG GAG TCC ACC AAG AAC GCC TTC GAC AAG GTG AAA
     Gly Asn Val Gly Asp Trp Glu Ser Thr Lys Asn Ala Phe Asp Lys Val Lys

4949 GCC GAA GTC GGC GAA GTC GAC ATC CTG GTG AAC AAC GCG GGC ATC ACG CGC
     Ala Glu Val Gly Glu Val Asp Ile Leu Val Asn Asn Ala Gly Ile Thr Arg
```

```
5000 GAC GTC GTG TTC CGC AAG ATG ACG CAC GAG GAC TGG ACG GCC GTC ATC CAC
     Asp Val Val Phe Arg Lys Met Thr His Glu Asp Trp Thr Ala Val Ile His

5051 ACC AAC CTG ACG AGC CTC TTC AAC GTG ACC AAG CAG GTG GTC CAC GGC ATG
     Thr Asn Leu Thr Ser Leu Phe Asn Val Thr Lys Gln Val Val His Gly Met

5102 GTG GAG CGC GGT TTC GGG CGG ATC ATC AAC ATT TCG TCG GTG AAC GGC CAC
     Val Glu Arg Gly Phe Gly Arg Ile Ile Asn Ile Ser Ser Val Asn Gly His

5153 AAA GGG CAG TTC GGC CAG ACC AAC TAC TCT ACG GCG AAA GCC GGC ATC CAC
     Lys Gly Gln Phe Gly Gln Thr Asn Tyr Ser Thr Ala Lys Ala Gly Ile His

5204 GGC TTC ACG ATG GCG CTC GCG CAG GAA GTG GCG ACC AAG GGC GTG ACG GTC
     Gly Phe Thr Met Ala Leu Ala Gln Glu Val Ala Thr Lys Gly Val Thr Val

5255 AAC ACC GTG TCG CCA GGC TAT ATC GGC ACG GAC ATG GTC AAG GCG ATC CGC
     Asn Thr Val Ser Pro Gly Tyr Ile Gly Thr Asp Met Val Lys Ala Ile Arg

5306 CCC GAG GTG CTG GAG AAG ATC GTC GCG ACG ATT CCG GTG CGC CGT CTC GGC
     Pro Glu Val Leu Glu Lys Ile Val Ala Thr Ile Pro Val Arg Arg Leu Gly

5357 CGC CCG GAC GAG ATC GGC TCG ATC GTG TCG TGG CTG GCA TCG GAG GAA TCG
     Arg Pro Asp Glu Ile Gly Ser Ile Val Ser Trp Leu Ala Ser Glu Glu Ser

5408 GGC TTC TCG ACC GGT GCG GAC TTC TCG CTC AAC GGC GGG CTG CAC ATG GGC
     Gly Phe Ser Thr Gly Ala Asp Phe Ser Leu Asn Gly Gly Leu His Met Gly

     stop(phaB)
5459 TGA CGG GCG GTG GCT CGC GCG GCC TGG AAC CGC GCG AGC CGT CAC TTG CCG
     ***

5510 GAA AAC GAA CCG GCG GCA ACG CTC TAT GCG TCA ACG CAG GAC ACG TTT GCT

                                          RBS (ORF4) start (ORF4)
5561 GCC ACC GCA GTA AGG TTT TGC GTT TAA AGG CGT TAA ATG ACC ACT ATT ACT
                                                 Met Thr Thr Ile Thr

5612 ACA AAG AAA CCC GCC GAA CGA CTC ATT AAA AAG TAT CCA AAC CGT CGG CTA
     Thr Lys Lys Pro Ala Glu Arg Leu Ile Lys Lys Tyr Pro Asn Arg Arg Leu

5663 TAC GAT ACG GAA ACG AGC ACC TAC ATC ACG CTT TCC GAC GTG AAG CAG CTC
     Tyr Asp Thr Glu Thr Ser Thr Tyr Ile Thr Leu Ser Asp Val Lys Gln Leu

5714 GTG CTG GAT CAG GAA GAC TTC AAG GTG ATG GAC GCG AAG TCC AAC GAC GAC
     Val Leu Asp Gln Glu Asp Phe Lys Val Met Asp Ala Lys Ser Asn Asp Asp
                                    Pst I
                                      ↓
5765 CTG ACG CGC AGC ATC CTG CTG CAG ATC ATC CTC GAA GAG GAG AGC GGC GGG
     Leu Thr Arg Ser Ile Leu Leu Gln Ile Ile Leu Glu Glu Glu Ser Gly Gly

5816 CTG CCG ATG TTC TCG TCG GTG ATG CTT TCG CAG ATC ATC CGT TTC TAC GGC
     Leu Pro Met Phe Ser Ser Val Met Leu Ser Gln Ile Ile Arg Phe Tyr Gly

5867 CAT GCG ATG CAG GGC ATG ATG GGC ACG TAT CTG GAA AAG AAC ATC CAG GCG
     His Ala Met Gln Gly Met Met Gly Thr Tyr Leu Glu Lys Asn Ile Gln Ala

5918 TTC ATC GAC ATC CAG CAG AAG CTC ACC GAT CAG AGC AAG GGC CTT TAC GAC
     Phe Ile Asp Ile Gln Gln Lys Leu Thr Asp Gln Ser Lys Gly Leu Tyr Asp

5969 GGC AAT GCG CTC AAC CCC GAA GTC TGG TCG CAG TTC ATG AAC ATG CAG GCG
     Gly Asn Ala Leu Asn Pro Glu Val Trp Ser Gln Phe Met Asn Met Gln Ala


6020 CCG ATG ATG CAG GGC ATG ATG ACG AGC TAC ATC GAG CAG TCG AAG AAC ATG
     Pro Met Met Gln Gly Met Met Thr Ser Tyr Ile Glu Gln Ser Lys Asn Met
```

```
 6071 TTC GTG CAG ATG CAG GAG CAG ATG CAG TCG CAG GCG AAG TCG ATG TTC AAC
      Phe Val Gln Met Gln Glu Gln Met Gln Ser Gln Ala Lys Ser Met Phe Asn

                                                                 stop(ORF4)
 6122 ACG TTT CCG TTT CCG CCG CCC GCC ACG CCG CCT TCG GAA AAG AAG TAA TCG
      Thr Phe Pro Phe Pro Pro Pro Ala Thr Pro Pro Ser Glu Lys Lys ***

 6173 TTC TCA GGC GTC TCG CGA AGG GTC CGC GGC GGG CGA ACT CGC TGA TTT GCA
 6224 TGC GCC GCG CCC GGC GCG GCC CGC GGC CGC GCA AAC CGC GGG CCG TCC ATG
 6275 GCG GTG CGG GCC GGG CGG CGC CGG TGC CGC CGT TGC GCG GCA CGG TAC AAT
 6326 ATG CCC TTT GCA GCG CGG CGC ACC CGG CGA ATC ATC GGC GGG AAG CAG GAA
 6377 ATG CGC GGC AGT GTC CGA ATC GTC CGC TGG TTC CGG CAC CGC GCG CCA CCG
 6428 CGC TCG TTA TTC GTG CCT TTT CCG GCC GCG CGC CTT CGC GCT CCA CGC CCG
 6479 GCC CGG GCA CTC ACT TCG CCA GAA CAT CTA TGT CGC ACA TTT CCT CTT CAG
 6530 CCG CGC CGA AAG TCG CTT CGT GTC CTC GGC TGC CCG AAA GCC CTC GTT GAT
 6581 TCC AAC AGA TCA TCA CGC AAC TGC GCG CGG AAG GCT ACA AAA TCT CNG GCA
 6632 CTA TGA CGG GCC GAT CTC TTG TCT CAC ACT GCG NTT CAT CAA CAA ACC TGC
 6683 AGG AAA NCT CAC CCA TCG GCA AGG GCT GAA CAA AAC GNA AGG GAT CTN ACN
 6734 GTT NCT CGC NCA AAA AAA CCC ACG NNC GGG TCT CAA GAA ATT TTC NAA GTT
 6785 TCC GTC CCG GCC CNC CCC ATN AGT ATA ANC GTC CCC CTT NCN AA
//
```

(a)



(b)

Figure 5.7. **(a)** Complete annotated DNA sequence encoding the entire *pha* biosynthetic region of isolate 10c-1-3 and **(b)** a schematic showing the arrangement of the putative *pha* genes and their promoter of isolate 10c-1-3.

homology to published sequences for *pha* biosynthetic genes encoded by other prokaryotic bacteria with the highest homology to the *pha* biosynthetic operon from *R. eutropha*.

The first ORF (*phaC*) is 2190 bp in length, spans nucleotides 979 to 3168 and encodes a polypeptide of 729 aa. An alternative in frame ATG codon for a truncated peptide is located at nucleotide 1153. This is followed by three other ORFs spanning nucleotides 3248 to 4429 (*phaA*), 4516 to 5256 (*phaB*) and 5392 to 5964 (ORF4) encoding predicted polypeptides of 393, 246 and 190 aa in length. The arrangement of these *pha* homologues is the same as that described for *R. eutropha* and other Type I PHA producing bacteria (see Section 1.5.2).

5.2.3 Analysis of DNA translation products

The predicted polypeptide sequences obtained were compared to related proteins in the Genpept, Swissprot and Pir protein databases using FASTA algorithms (Pearson and Lipman, 1988; and Pearson, 1990).

The polypeptide encoded by *phaC* has a predicted molecular mass of 81.3 kDa and lacks a recognisable signal peptide cleavage site and therefore is likely to be cytoplasmically located. PhaC has significant identity to bacterial PHA synthases (AAF23364, *Burkholderia* sp., 75% identity over a 533 aa overlap; P23608, *R. eutropha* H16, 65% identity over a 520 aa overlap; Q9ZB54, *Alcaligenes latus*, 61% identity over a 528 aa overlap). The polypeptide is about 100 amino acids longer than other published PhaC sequences and this may reflect the fact that the actual start of transcription has not been experimentally determined by primer extension analysis. In fact two potential in-frame start sites have been identified from the nucleotide sequence (as shown in Figure 5.7a). The second (internal to the ORF, *phaC*) would reduce the predicted length of the polypeptide by 58 amino acid residues. Broadly speaking however, the proteins have similar amino acid compositions (Table 5.1). Simple phylogenetic analysis (Figure 5.8) indicates a complex set of groupings which show only partial correlation to phylogenetic groups from which the predicted peptide sequences have been drawn. For example, although PhaCs from the β-Proteobacteria group well, other groups contain representatives from the γ-Proteobacteria and the low G+C Gram positive bacteria (*Methylobacterium* and

Table 5.1. Comparison of the amino acid composition of *Pseudomonas* 10c-1-3 PHA synthase (PhaC) with other closely related PHA synthase sequences.

Amino Acid	10c-1-3		AAF23364		P26496		AA99474		AA72004	
	N	%	N	%	N	%	N	%	N	%
Ala(A)	76	10.4	87	13.9	52	9.3	36	6.1	51	8.5
Cys(C)	16	2.2	6	1.0	3	0.5	4	0.7	4	0.7
Asp(D)	37	5.1	35	5.6	32	5.7	25	4.2	34	5.7
Glu(E)	30	4.1	27	4.3	20	3.6	37	6.3	35	5.8
Phe(F)	26	3.6	32	5.1	19	3.4	32	5.4	31	5.2
Gly(G)	40	5.5	42	6.7	38	6.8	35	5.9	47	7.8
His(H)	16	2.2	8	1.3	17	3.0	10	1.7	13	2.2
Ile(I)	28	3.8	19	3.0	19	3.4	30	5.1	24	4.0
Lys(K)	27	3.7	21	3.4	19	3.4	23	3.9	28	4.7
Leu(L)	56	7.7	56	9.0	77	13.8	59	10.0	54	9.0
Met(M)	12	1.6	15	2.4	8	1.4	13	2.2	14	2.3
Asn(N)	24	3.3	24	3.8	32	5.7	49	8.3	26	4.3
Pro(P)	52	7.1	42	6.7	41	7.3	28	4.7	40	6.7
Gln(Q)	32	4.4	39	6.2	23	4.1	36	6.1	29	4.8
Arg(R)	73	10.0	30	4.8	41	7.3	23	3.9	29	4.8
Ser(S)	69	9.5	37	5.9	38	6.8	36	6.1	36	6.0
Thr(T)	45	6.2	36	5.8	23	4.1	26	4.4	37	6.2
Val(V)	37	5.1	35	5.6	27	4.8	43	7.3	38	6.3
Trp(W)	16	2.2	17	2.7	15	2.7	16	2.7	11	1.8
Tyr(Y)	17	2.3	17	2.7	16	2.9	29	4.9	20	3.3
Glx(Z)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Acidic	67	9.2	62	9.9	52	9.3	62	10.5	69	11.5
Basic	116	15.9	59	9.4	77	13.8	56	9.5	70	11.6
Neutral Polar	327	44.9	328	52.5	281	50.2	276	46.8	299	49.8
Neutral Nonpolar	219	30.0	176	28.2	150	26.8	196	33.2	163	27.1
Total	729	100	625	100	560	100	590	100	601	100

All data are derived from the sequences shown in Figure 5.8 and from the Genbank database (Accession numbers shown). Compositional analysis of an amino acid is determined as a percentage of the total number of amino acids per sequence. Acidic (D, E); Basic (H,K R); Neutral Polar (A, F, G, I, L, M, P, V); Neutral Nonpolar (C, N, Q, S, T, W, Y).

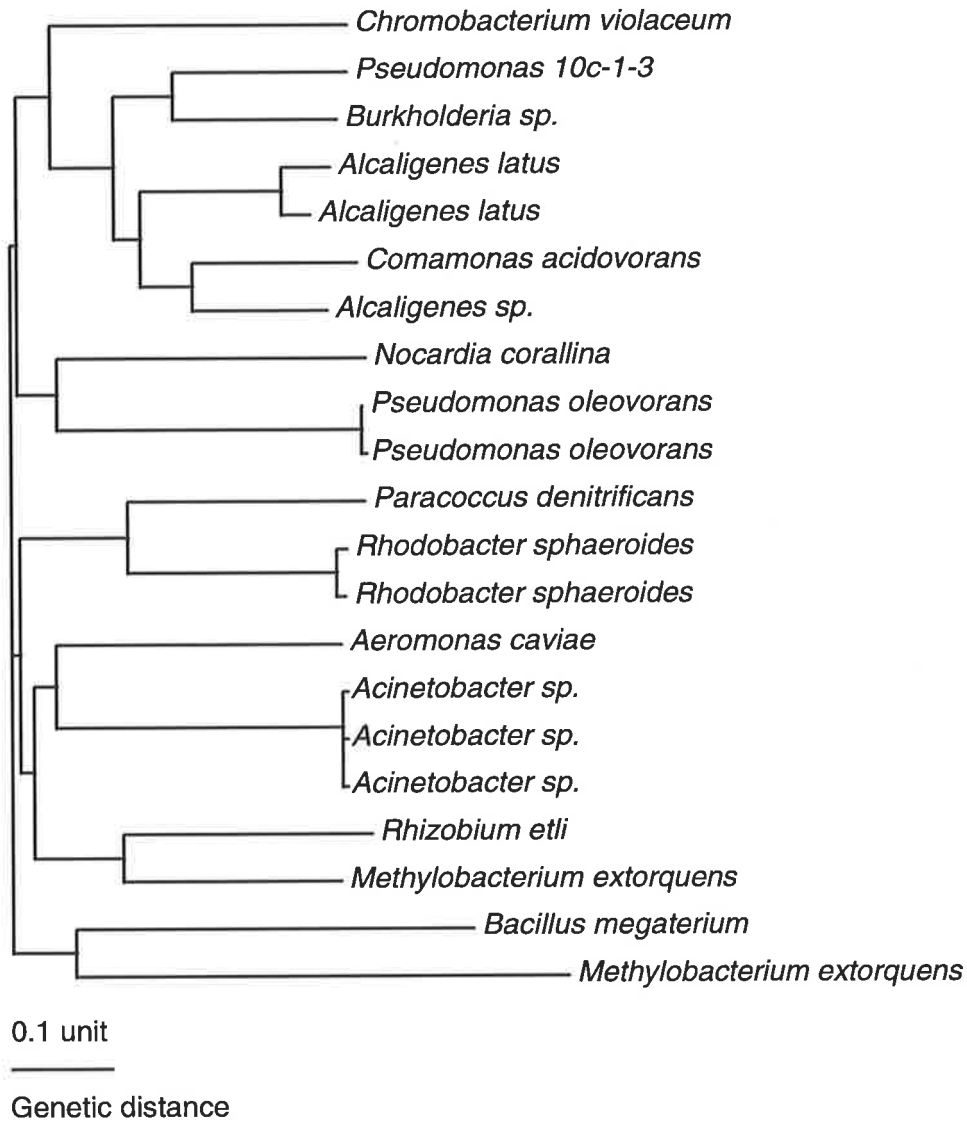


Figure 5.8. Unrooted phylogenetic tree derived from a comparison of PHA synthase peptide sequences.

The scale representing genetic distance is an arbitrary unit. The predicted PhaC identified in this study is labelled *Pseudomonas* 10c-1-3. Protein sequence database accession numbers sources for the other PhaC polypeptide sequences are as follows: AAF23364, *Burkholderia*; P26496, *Pseudomonas oleovorans*; AAC69615, *Chromobacterium violaceum*; BAA77257 *Paracoccus denitrificans*; AAD10274, *Alcaligenes latus*; AAD05260, *Bacillus megaterium*; AAC83658, *Alcaligenes latus*; BAA33155, *Comamonas acidovorans*; AAB94058, *Nocardia corallina*; BAA21815, *Aeromonas caviae*; CAA65833, *Rhodobacter sphaeroides*; AAB36944, *Alcaligenes* sp; 744165, *Acinetobacter* sp; AAB05020, *Rhizobium etli*; AAA99474, *Acinetobacter* sp.; AAA72330, *Methylobacterium extorquens*; AAA72004, *Rhodobacter sphaeroides*; AAA72003, *Rhodobacter sphaeroides*; AAA52191, *Acinetobacter* sp.; and AAA25934, *Pseudomonas oleovorans*.

Phylogenetic trees were calculated using the Clustal W alignment tool. This tool uses the Neighbor-Joining method on a matrix of "distances" between all sequences. These distances can be corrected for "multiple hits". Tree files were then displayed as a phylogram using Treeview (Page, R. D. M., 1998. *http//taxonomy.zoology.gla.ac.uk/rod/rod.html*).

Bacillus).

The second polypeptide (encoded by *phaA*) has a predicted molecular mass of 40.6 kDa and lacks an identifiable signal peptidase recognition site as a consequence is likely to be a cytoplasmic protein. PhaA shows significant identity to the family of β-ketothiolases, commonly associated with PHA biosynthesis (AAF23365, *Burkholderia* sp., 89% identity over a 393 aa overlap; P14611, *R. eutropha* H16, 84% identity over a 392 aa overlap; Q9ZG19, *A. latus*, 73% identity over a 392 aa overlap). The PhaA polypeptide contains three functional thiolase signature motifs characteristic of the thiolase family of proteins. These key functional domains were identified by using the predicted peptide sequence to search for thiolase motifs in the Prosite database (Bucher and Bairoch, 1994). Figure 5.9 shows alignments of these key functional domains for PhaA and related proteins and associated consensus sequences for the motifs. The cysteine residues associated with the acyl-enzyme intermediate signature in the N-terminal region of the protein and the β-ketothiolase active site located at the C-terminus are absolutely conserved in hydrophobic regions of all the aligned polypeptide sequences shown. In addition to the remarkable conservation of sequence and functional domains in this family of proteins, there is also considerable similarity in the hydropathy profiles (Figure 5.10) and amino acid composition. Selected representatives of the β-ketothiolases are rich in neutral polar amino acids such as alanine (15 - 17%) and glycine (10 - 11%) (Table 5.2). These similarities are further supported by simple phylogenetic analysis of a selection of published PhaA peptide sequences (Figure 5.11). The unrooted phylogram shows that none of the sequences examined vary by more than 0.2 Genetic units. As a general observation, peptide sequences inferred from genes broadly cluster according to phylogenetic grouping. For example, PhaA from *Paracoccus* (α-Proteobacteria) group separately from *Alcaligenes*, *Comamonas*, *Burkholderia* (β-Proteobacteria). The PhaA from 10c-1-3 is an exception given the tentative classification as a pseudomonad.

The third polypeptide, encoded by *phaB*, has a predicted mass of 26.5 kDa and shows extensive identity to acetoacetyl-CoA reductases involved in biosynthesis of PHA. Like PhaC and PhaB, the polypeptide lacks a signal peptide cleavage site and is therefore predicted to be located within the cytoplasm. FASTA analysis indicated the polypeptide was 90% identical over a 246 aa overlap to a *Burkholderia* sp. PhaB (AAF23366); 84%

Thiolase acyl-enzyme intermediate signature

```
Pseudomonas 10c-1-3    84 INKVCGSGLKAVMLAANAI 102
AAF23365               84 INKVCGSGLKAVMLAANAI 102
AAD10275               84 INAVCGSGLKAVMLAAQAI 102
AAC83659               84 INAVCGSGLKAVMLAAQAI 102
BAA33156               84 INAVCGSGLKAVMLAAQAV 102
AAB65779               84 INAVCGSGLKAVMLAAQAV 102
P54810                 84 INQVCGSGLRTVALAAQQV 102
BAA08357               84 INQVCGSGLRTVALAAQQV 102
1585881                84 INQVCGSGLRTVALAAQQV 102
AAA99475               84 LNVVCGSGLRAVHLAAQAI 102
AAB48515               84 INKLCGSGMKAVMLAHDLL 102
Consensus              84 .* .****...* **    .
```

Consensus pattern: [LIVM]-[NST]-x(2)-C-[SAGLI]-[ST]-[SAG]-[LIVMFYNS]-x-[STAG]-[LIVM]-x(6)-[LIVM]

Thiolase signature 2

```
Pseudomonas 10c-1-3  339 NVNGGAIAIGHPIGASG 355
AAF23365             339 NVNGGAIAIGHPIGASG 355
AAD10275             339 NVNGGAIAIGHPIGASG 355
AAC83659             339 NVNGGAIAIGHPIGASG 355
BAA33156             339 NVNGGAIAIGHPIGASG 355
AAB65779             339 NVNGGAIAIGHPIGASG 355
P54810               337 NVNGGAIAIGHPIGASG 353
BAA08357             337 NVNGGAIAIGHPIGASG 353
1585881              337 NVNGGAIAIGHPIGASG 353
AAA99475             338 NVNGGAIALGHPIGASG 354
AAB48515             337 NVYGGACAQGHPVGSTG 353
Consensus                ** *** * ***.*..*
```

Consensus pattern: N-x(2)-G-G-x-[LIVM]-[SA]-x-G-H-P-x-G-x-[ST]-G

Thiolase active site

```
Pseudomonas 10c-1-3  374 GLASLCIGGGMGVA 387
AAF23365             374 GLASLCIGGGMGVA 387
AAD10275             374 GVAALCIGGGMGVS 387
AAC83659             374 GLAALCIGGGMGVS 387
BAA33156             374 GLAGLCIGGGMGVA 387
AAB65779             374 GLAALCIGGGMGVS 387
P54810               372 GLATLCIGGGMGVA 385
BAA08357             372 GLATLCIGGGMGVA 385
1585881              372 GLATLCIGGGMGVA 385
AAA99475             373 GIATLCVGGGMGVA 386
AAB48515             372 GVASLCIGGGEATA 385
Consensus                *.* **.***    .
```

Consensus pattern: [AG]-[LIVMA]-[STAGLIVM]-[STAG]-[LIVMA]-C-x-[AG]-x-[AG]-x-[AG]-x-[SAG]

Figure 5.9. Alignments of the key functional domains of the predicted PhaA from *Pseudomonas* 10c-1-3 with related proteins.

Conserved amino acid sequences associated with 1, Thiolase acyl-enzyme intermediate signature; 2, Thiolase signature 2; and 3, the Thiolase active site. Asterisks indicate identical residues in all sequences represented. Periods indicate similar amino acid residues in all sequences represented. Sequences were aligned using CLUSTAL W (Thompson *et al*, 1994). Accession numbers and source organisms for all PhaA sequences shown are: AAF23365, *Burkholderia* sp. DSMZ 9242; P54810, *Paracoccus denitrificans*; AAD10275, *Alcaligenes latus*; AAC83659, *Alcaligenes latus*; BAA33156, *Comamonas acidovorans*; AAB65779, *Alcaligenes* sp. SH-69; BAA08357, *Paracoccus denitrificans*; AAB48515, *Pseudomonas aeruginosa*; 1585881, *Paracoccus denitrificans*; and AAA99475, *Acinetobacter* sp.

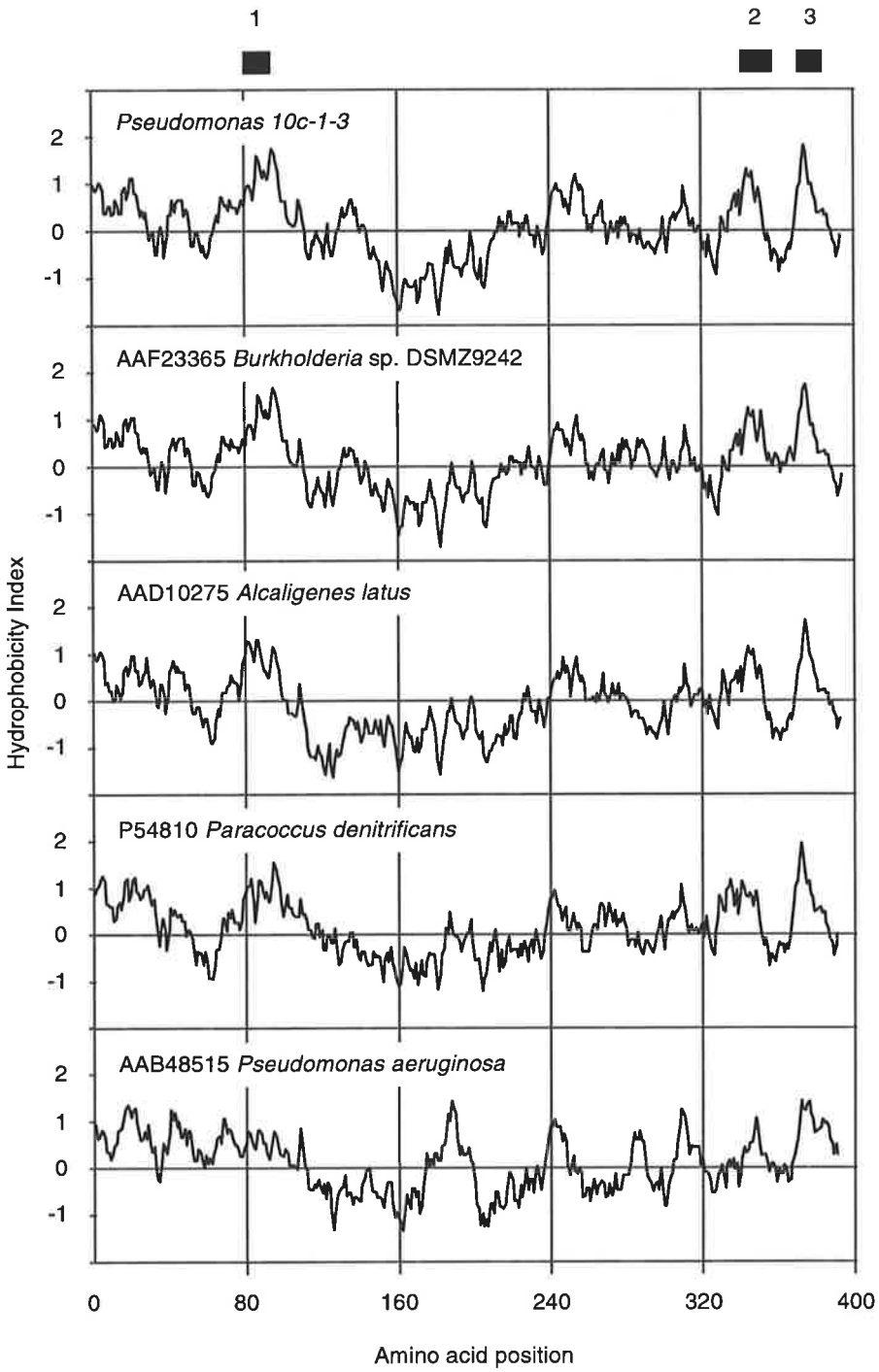


Figure 5.10. Aligned hydropathy profiles of the predicted *Pseudomonas* 10c-1-3 polypeptide PhaA and selected β-ketothiolases involved in PHA biosynthesis, calculated by the method of Kyte and Doolittle (1982) using a window span of 17 amino acid residues.

Hydrophobic domains are indicated by regions of the graphs above the horizontal axis. Also shown are the amino acid sequence segments associated with functional thiolase recognition sites. Region 1 is the β-ketothiolase acyl-enzyme intermediate signature involved in the formation of an acyl-enzyme intermediate; region 2, a β-ketothiolase signature 2; and region 3, the β-ketothiolase active site involved in deprotonation. The Genbank accession number and source organism are shown for each plot.

Table 5.2. Comparison of the amino acid composition of *Pseudomonas* 10c-1-3 β-ketothiolase (PhaA) with other closely related β-ketothiolases sequences.

Amino Acid	10c-1-3		AAF23365		AAD10275		P54810		AAB48515	
	N	%	N	%	N	%	N	%	N	%
Ala (A)	63	16.0	67	17.0	67	17.0	69	17.6	60	15.3
Cys (C)	4	1.0	5	1.3	5	1.3	4	1.0	5	1.3
Asp (D)	19	4.8	20	5.1	20	5.1	21	5.4	19	4.9
Glu (E)	20	5.1	18	4.6	18	4.6	19	4.8	23	5.9
Phe (F)	9	2.3	8	2.0	8	2.0	8	2.0	8	2.0
Gly(G)	43	10.9	44	11.2	44	11.2	46	11.7	44	11.3
His (H)	6	1.5	6	1.5	6	1.5	5	1.3	7	1.8
Ile (I)	24	6.1	24	6.1	24	6.1	23	5.9	20	5.1
Lys (K)	25	6.4	21	5.3	21	5.3	24	6.1	17	4.3
Leu (L)	30	7.6	31	7.9	31	7.9	27	6.9	36	9.2
Met (M)	15	3.8	14	3.6	14	3.6	17	4.3	17	4.3
Asn (N)	16	4.1	13	3.3	13	3.3	14	3.6	9	2.3
Pro (P)	15	3.8	18	4.6	18	4.6	15	3.8	13	3.3
Gln (Q)	13	3.3	13	3.3	13	3.3	18	4.6	14	3.6
Arg (R)	16	4.1	17	4.3	17	4.3	14	3.6	21	5.4
Ser (S)	23	5.9	17	4.3	17	4.3	14	3.6	23	5.9
Thr (T)	13	3.3	14	3.6	14	3.6	16	4.1	22	5.6
Val (V)	34	8.7	36	9.2	36	9.2	32	8.2	28	7.2
Trp (W)	3	0.8	3	0.8	3	0.8	4	1.0	1	0.3
Tyr (Y)	2	5.0	4	1.0	4	1.0	2	0.5	4	1.0
Glx (Z)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Acidic	39	9.9	38	9.7	38	9.7	40	10.2	42	10.7
Basic	47	12.0	44	11.2	44	11.2	43	11.0	45	11.5
Neutral Polar	233	59.3	242	61.6	242	61.6	237	60.5	226	57.8
Neutral Nonpolar	74	18.8	69	17.6	69	17.6	72	18.4	78	19.9
Total	393	100	393	100	393	100	392	100	391	100

All data are derived from the sequences shown in Figure 5.11 and from the Genbank database (Accession numbers shown). Compositional analysis of an amino acid is determined as a percentage of the total number of amino acids per sequence. Acidic (D, E); Basic (H,K R); Neutral Polar (A, F, G, I, L, M, P, V); Neutral Nonpolar (C, N, Q, S, T, W, Y)

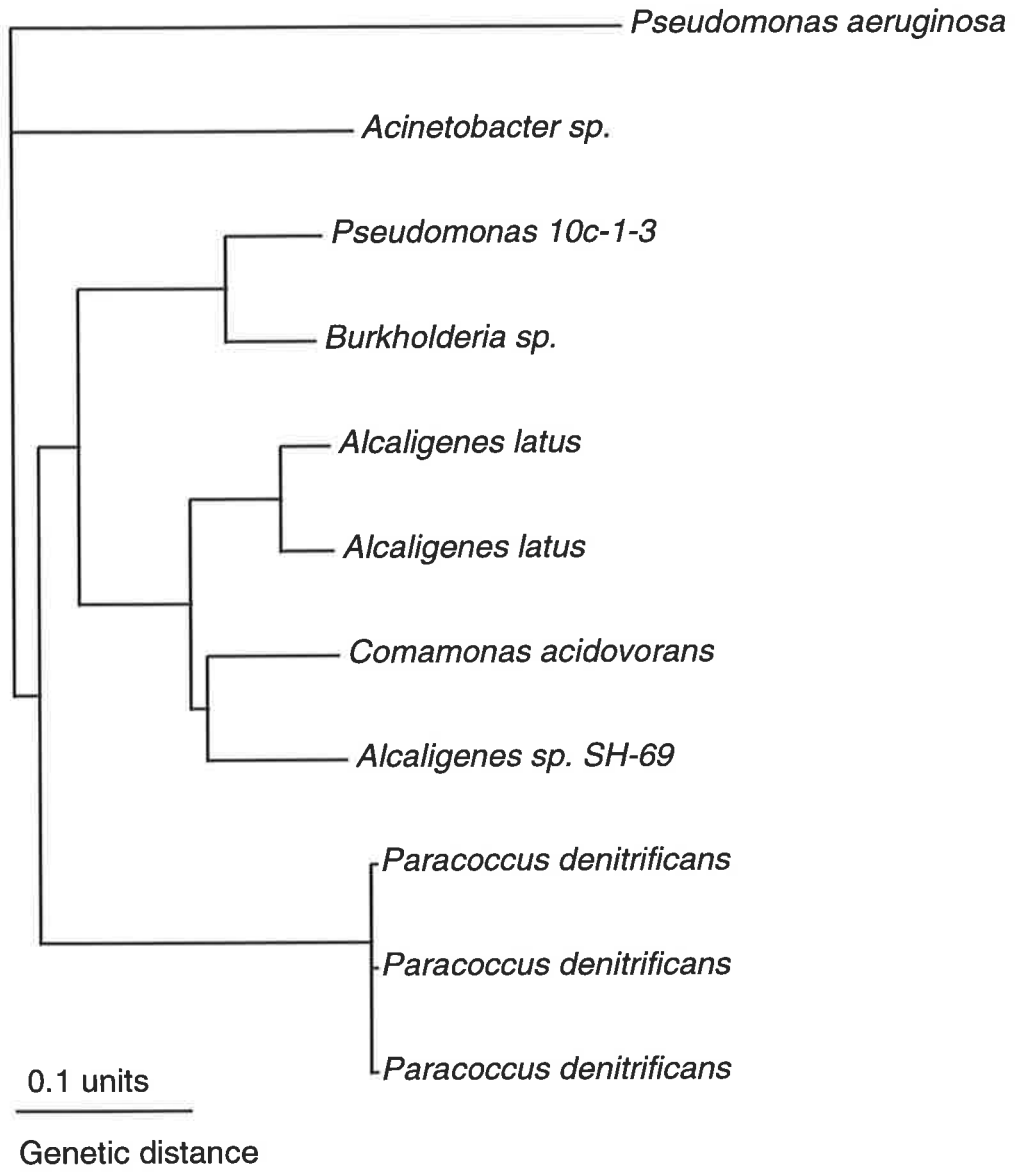


Figure 5.11. Unrooted phylogenetic tree derived from a comparison of β-ketothiolases.

The scale representing genetic distance is an arbitrary unit. The predicted polypeptide PhaA identified in this study is labelled *Pseudomonas* 10c-1-3. Protein sequence database accession numbers sources for the other PhaA polypeptide sequences are as follows: AAF23365, *Burkholderia* sp. DSMZ 9242; P54810, *Paracoccus denitrificans*; AAD10275, *Alcaligenes latus*; AAC83659, *Alcaligenes latus*; BAA33156, *Comamonas acidovorans*; AAB65779, *Alcaligenes* sp. SH-69; BAA08357, *Paracoccus denitrificans*; AAB48515, *Pseudomonas aeruginosa*; 1585881, *Paracoccus denitrificans*; and AAA99475, *Acinetobacter* sp.

Phylogenetic trees were calculated using the Clustal W alignment tool. This tool uses the Neighbor-Joining method on a matrix of "distances" between all sequences. These distances can be corrected for "multiple hits". Tree files were then displayed as a phylogram using Treeview (Page, R. D. M., 1998. *http://taxonomy.zoology.gla.ac.uk/rod/rod.html*).

identical to the *R. eutropha* H16 PhaB (P14697); and 75% identical over a 246 aa overlap to the *A. latus* PhaB (Q9ZB52). A dehydrogenase/reductase family signature motif was located between amino acid 140 and 168 of this peptide (Figure 5.12). This motif is highly conserved among representative PhaB polypeptides. Like PhaA, PhaB is characterised by similarity in the hydrophobicity profiles of acetoacetyl-CoA reductases (Figure 5.13) and amino acid composition (Table 5.3). These proteins share remarkable similarity in polypeptide length and amino acid composition. The structural and compositional similarity is also reflected by simple phylogenetic analysis (Figure 5.14). Although this simple analysis contains too few representative sequences to attach any conclusive phylogenetic inferences, it nevertheless clearly identifies four separate groups (or variants) of PhaB. Each group contains representative proteins from distinct phylogenetic groups. For example, the *Paracoccus* (α-Proteobacteria) are in a group distinct from *Acinetobacter* (γ-Proteobacteria). Although the proteins from *Pseudomonas*, *Burkholderia* and *Alcaligenes* cross two distinct phylogenetic boundaries (β- and γ-Proteobacteria) they are confined to a single group. This may be because there is insufficient representative sequences and sequence differences to develop any meaningful distinctions on phylogenetic grounds.

Finally, the 16 kDa polypeptide encoded by ORF4 shows strong similarity (82% identity) to the *phaF* gene of *R. eutropha* (Slater *et al.* 1998) as well as several hypothetical proteins thought to be associated with PHA biosynthetic enzymes. The peptide is predominantly hydrophilic (data not shown) and does not contain sequences representative of any motif currently characteristic of β-ketothiolases or those included in the Prosite database, nor does it contain an N-terminal signal peptide cleavage site and consequently is predicted to be a cytoplasmic protein.

Interestingly, the phylogenetic analysis indicated that the PhaC, PhaA and PhaB identified from isolate 10c-1-3 are all most similar to representative predicted polypeptides from the β-Proteobacteria organism, *Burkholderia* (Rodrigues *et al.*, 2000). BLASTN analysis also showed that the similarities between the *pha* genes from these two organisms are at high levels (83% at DNA level and 75% at peptide level for *phaC*s; 85% at DNA level and 81% at peptide level for *phaA*s; 87% at DNA level and 90% at peptide level for

Short chain dehydrogenase/reductase family signature

```
Pseudomonas 10c-1-3     140 SVNGHKGQFGQTNYSTAKAGIHGFTMALA 168
AAF23366                140 SVNGQKGQFGQTNYSTAKAGIHGFTMALA 168
AAD10276                139 SVNGEKGQFGQTNYSAAKAGMHGFTMALA 167
AAC83660                139 SVNGEKGQFGQTNYSAAKAGMHGFSMALA 167
AAB65780                139 SVNGEKGQAGQTNYSAAKAGMHGFSMALA 167
P50204                  134 SINGQKGQAGQANYSAAKAGDLGFTKALA 162
BAA08358                134 SINGQKGQAGQANYSAAKAGDLGFTKALA 162
1585882                 134 SINGQKGQAGQANYSAAKAGDLGFTKALA 162
P50203                  142 SVNGLKGQFGQANYSASKAGIIGFTKALA 170
AAA99472                142 SVNGLKGQFGQANYSASKAGIIGFTKALA 170
AAD05259                143 SIIGQAGGFGQTNYSAAKAGMLGFTKSLA 171
Consensus                   *. *  *  **.***..***  **. .**
```

Consensus pattern: [LIVSPADNK]-x(12)-Y-[PSTAGNCV]-[STAGNQCIVM]-[STAGC]-K-{PC}-[SAGFR]-[LIVMSTAGD]-x(2)-[LIVMFYW]-x(3)-[LIVMFYWGAPTHQ]-[GSACQRHM]

Figure 5.12. Alignment of a key functional domain of PhaB from *Pseudomonas* 10c-1-3 with related proteins.

Conserved amino acid sequences characteristic of the short chain dehydrogenase/reductase family of proteins. Asterisks indicate identical residues in all sequences represented. Periods indicate similar amino acid residues in all sequences represented.

The predicted PhaB polypeptide sequence from *Pseudomonas* strain 10c-1-3, was used to scan the Prosite database for motif patterns (*http://expasy.proteome.org.au/tools/scnpsit1.html*) (Hofmann *et al.*, 1999). Identified patterns, or motifs were then identified on other PhaB sequences aligned using CLUSTAL W alignment tool (Thompson *et al*, 1994). Accession numbers and source organisms for all PhaB sequences shown are: AAF23366, *Burkholderia* sp. DSMZ 9242; P50204, *Paracoccus denitrificans*; P50203, *Acinetobacter* sp. RA3849; AAD10276, *Alcaligenes latus*; AAD05259, *Bacillus megaterium*; AAC83660, *Alcaligenes latus*; AAB65780, *Alcaligenes* sp. SH-69; BAA08358, *Paracoccus denitrificans*; 1585882, *Paracoccus denitrificans*; and AAA99472, *Acinetobacter* sp.

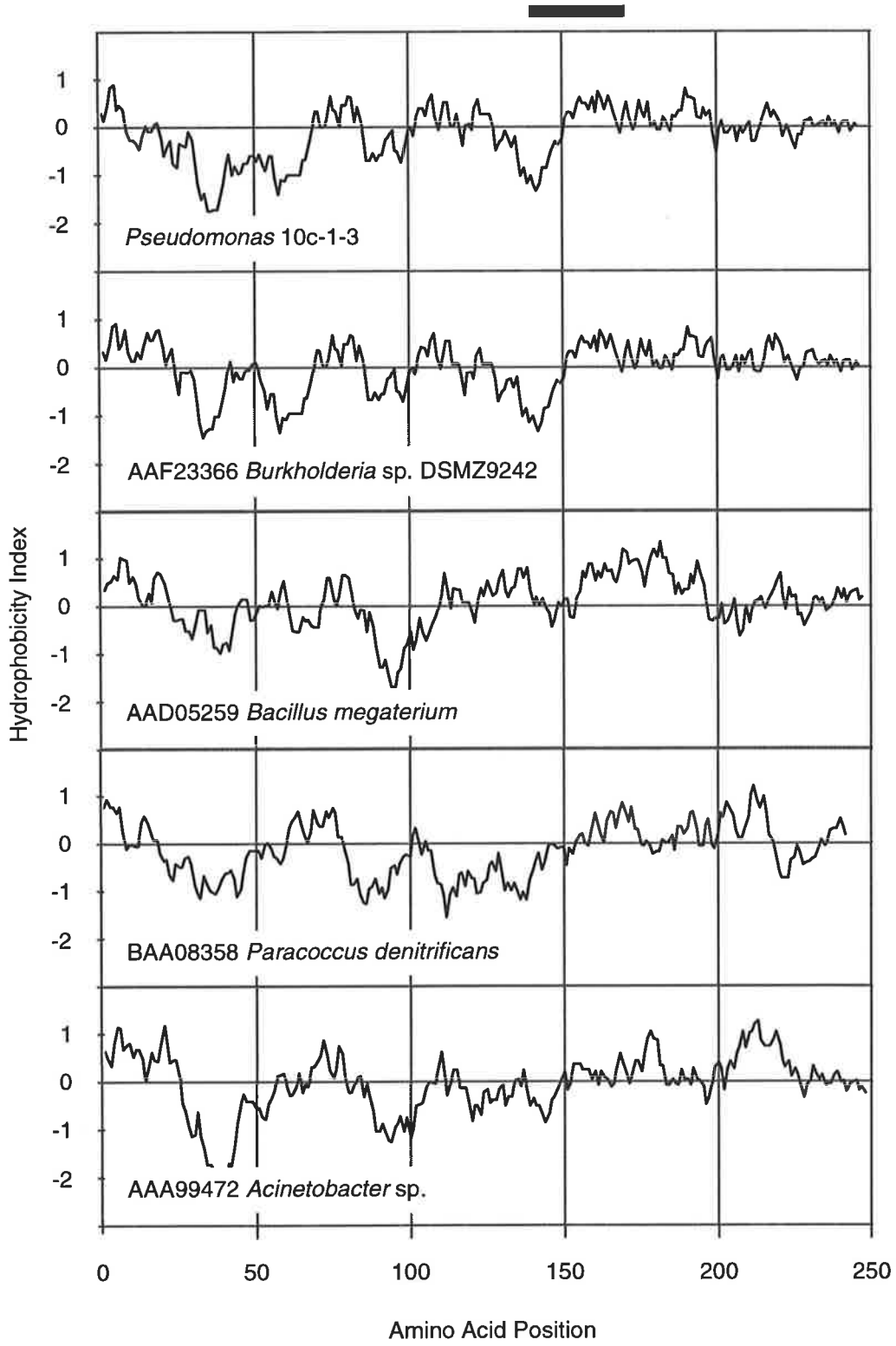


Figure 5.13. Aligned hydropathy profiles of PhaB and selected acetoacetyl-CoA reductases involved in PHA biosynthesis, calculated by the method of Kyte and Doolittle (1982) using a window span of 17 amino acid residues.

Hydrophobic domains are indicated by regions of the graphs above the horizontal axis. Also shown is the amino acid sequence segment characteristic of short chain dehydrogenase/reductase proteins. The Genbank accession number and source organism are shown for each plot.

Table 5.3. Comparison of the amino acid composition of *Pseudomonas* 10c-1-3 acetoacetyl-CoA reductase (PhaB) with other closely related acetoacetyl-CoA reductase sequences.

Amino Acid	10c-1-3		AAF23365		AAA99472		BAA08358		AAD05259	
AA	N	%	N	%	N	%	N	%	N	%
Ala (A)	17	6.9	19	7.7	30	12.1	37	15.3	30	12.1
Cys (C)	2	0.8	2	0.8	2	0.8	3	1.2	1	0.4
Asp (D)	10	4.1	16	6.5	8	3.2	10	4.1	9	3.6
Glu (E)	14	5.7	11	4.5	17	6.9	15	6.2	17	6.9
Phe (F)	9	3.7	10	4.1	7	2.8	8	3.3	5	2.0
Gly (G)	31	12.6	31	12.6	20	8.1	27	11.2	27	10.9
His (H)	7	2.8	3	1.2	3	1.2	3	1.2	4	1.6
Ile (I)	17	6.9	18	7.3	18	7.3	19	7.9	22	8.9
Lys (K)	14	5.7	13	5.3	12	4.8	13	5.4	18	7.3
Leu (L)	11	4.5	13	5.3	17	6.9	10	4.1	18	7.3
Met (M)	7	2.8	7	2.8	7	2.8	6	2.5	5	2.0
Asn (N)	13	5.3	12	4.9	12	4.8	12	5.0	13	5.3
Pro (P)	6	2.4	6	2.4	6	2.4	8	3.3	3	1.2
Gln (Q)	6	2.4	8	3.3	15	6.0	10	4.1	8	3.2
Arg (R)	12	4.9	11	4.5	10	4.0	9	3.7	8	3.2
Ser (S)	15	6.1	14	5.7	15	6.0	9	3.7	14	5.7
Thr (T)	20	8.1	19	7.7	17	6.9	15	6.2	16	6.5
Val (V)	26	10.6	25	10.2	23	9.3	17	7.0	21	8.5
Trp (W)	4	1.6	5	2.0	2	0.8	4	1.7	1	0.4
Tyr (Y)	5	2.0	3	1.2	7	2.8	7	2.9	7	2.8
Glx (Z)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Acidic	24	9.8	27	11.0	25	10.1	25	10.3	26	10.5
Basic	33	13.4	27	11.0	25	10.1	25	10.3	30	12.1
Neutral Polar	124	50.4	129	52.4	128	51.6	132	54.5	131	53.0
Neutral Nonpolar	65	26.4	63	25.6	70	28.2	60	24.8	60	24.3
Total	246	100	246	100	248	100	242	100	247	100

All data are derived from the sequences shown in Figure 5.14 and from the Genbank database (Accession numbers shown). Compositional analysis of an amino acid is determined as a percentage of the total number of amino acids per sequence. Acidic (D, E); Basic (H,K R); Neutral Polar (A, F, G, I, L, M, P, V); Neutral Nonpolar (C, N, Q, S, T, W, Y).

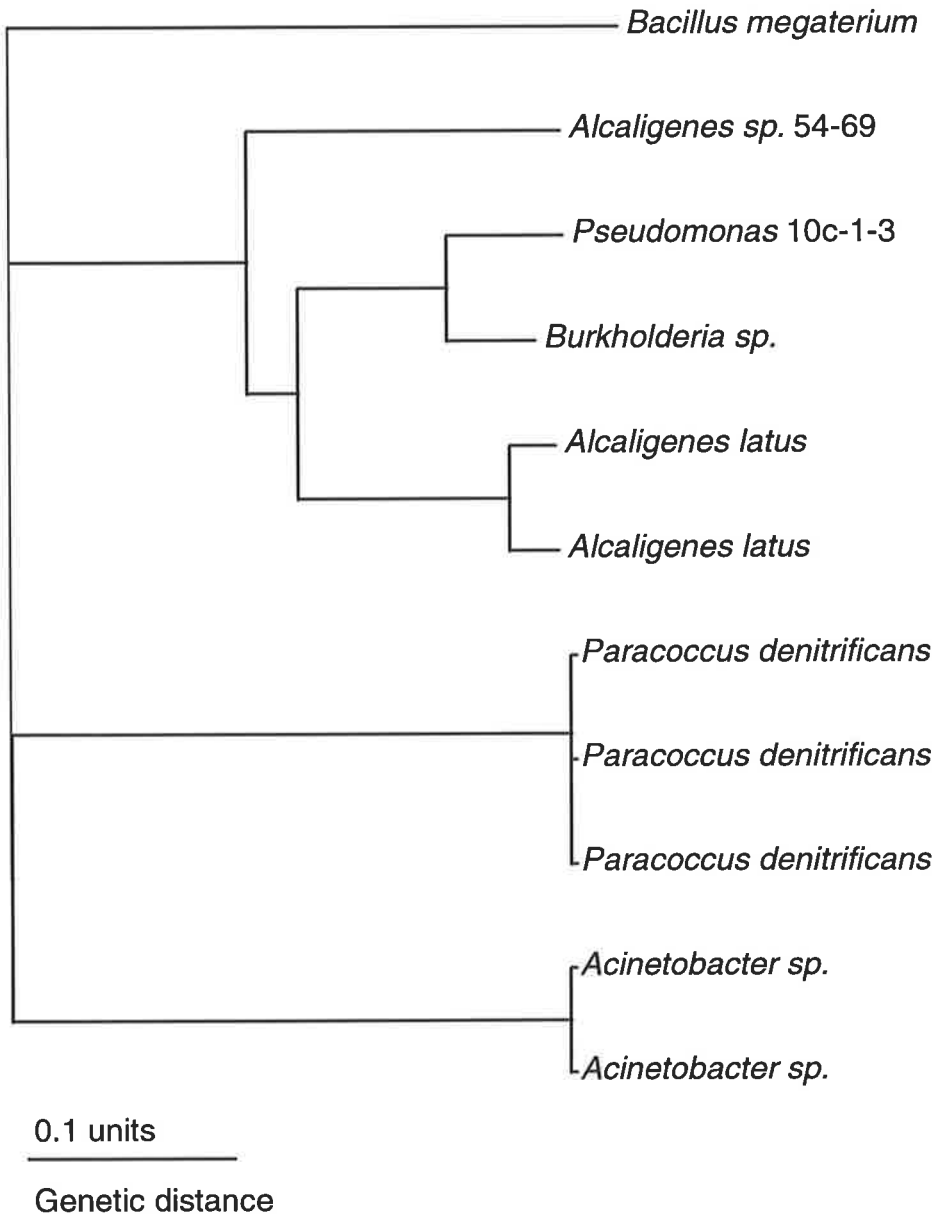


Figure 5.14. Unrooted phylogenetic tree derived from a comparison of acetoacetyl-CoA reductases.

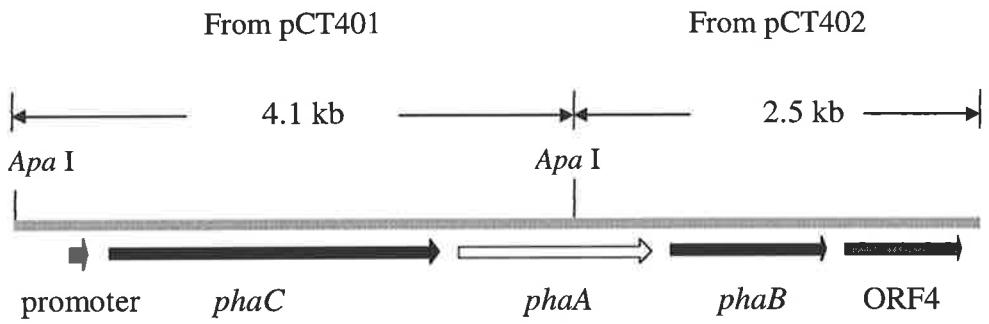
The scale representing genetic distance is an arbitrary unit. The predicted polypeptide PhaB identified in this study is labelled *Pseudomonas* 10c-1-3. Protein sequence database accession numbers sources for the other PhaA polypeptide sequences are as follows: AAF23366, *Burkholderia* sp. DSMZ 9242; P50204, *Paracoccus denitrificans*; P50203, *Acinetobacter* sp. RA3849; AAD10276, *Alcaligenes latus*; AAD05259, *Bacillus megaterium*; AAC83660, *Alcaligenes latus*; AAB65780, *Alcaligenes* sp. SH-69; BAA08358, *Paracoccus denitrificans*; 1585882, *Paracoccus denitrificans*; and AAA99472, *Acinetobacter* sp.

Phylogenetic trees were calculated using the Clustal W alignment tool. This tool uses the Neighbour-Joining method on a matrix of "distances" between all sequences. These distances can be corrected for "multiple hits". Tree files were then displayed as a phylogram using Treeview (Page, R. D. M., 1998. *http//taxonomy.zoology.gla.ac.uk/rod/rod.html*).

phaBs).

5.2.4 Construction of the PHA expression plasmid pCT411

Analysis of the distribution of restriction endonuclease restriction sites within the entire *pha* sequence indicated there were no suitable sites which might be used to subclone these genes from plasmid pCT400. To recover all the essential *pha* genes, *phaC* and the 5' partial *phaA* in pCT401 were ligated to the 3' partial *phaA*, *phaB* and ORF4 located on a piece of pCT402 as described below.



To achieve this, pCT406 and pCT408, which contain the 3' partial *phaA*, *phaB* and ORF4 were constructed as described in Figures 5.15. Plasmid pCT401 was then digested with *Apa* I and the 4.1 kbp fragment encoding *phaC* and the partial *phaA* was then cloned into *Apa* I digested pCT408 (Figure 5.16). These steps reconstructed the *pha* biosynthetic genes into their original order on plasmid pCT411. The construct was confirmed using a PCR method. PCR amplification was performed using the sequencing primer M13(F), which primes DNA replication from the vector at the site in front of the *Apa* I restriction site, and primer #5457 (for its binding site, see Table 2.3), which primes from the 4.1 kbp insert of pCT401. This PCR amplification also allowed selection of clones containing the 4.1 kbp insert in the correct orientation in pCT408. Techniques described in Section 4.2.2.2 were used to pool the transformants, prepare DNA templates for PCR and select positive clones. Plasmid pCT401 was used as a positive control which gave a ~1.7 kbp fragment after PCR amplification. Among 96 transformants, 3 clones were selected as positives. These were used for experiments designed to test whether *E. coli* strains harbouring plasmid pCT411 are able to produce PHA.

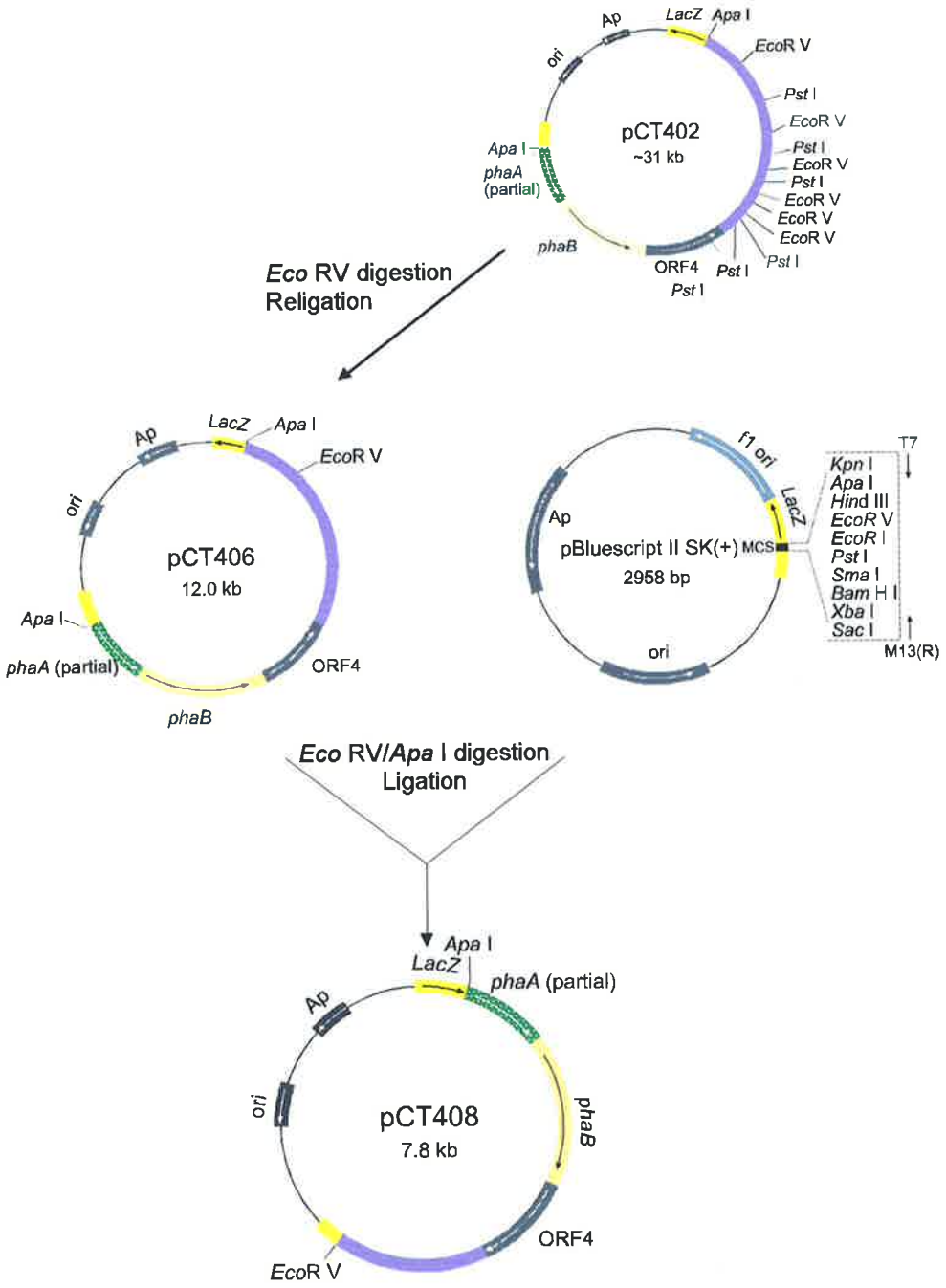


Figure 5.15. Construction of pCT408.

Plasmid pCT408 was constructed in two steps. Plasmid pCT402 was digested with *Eco* RV to remove ~19 kbp of DNA not associated with *phaA*, *phaB* and ORF4. The 12 kbp fragment containing vector plus *pha* genes, was religated to construct pCT406. Plasmid pCT406 was then digested with *Apa* I and *Eco* RV and the 4.8 kbp *Apa* I-*Eco* RV fragment containing *phaA* (partial), *phaB* and ORF4 was gel purified then subcloned into *Apa* I-*Eco* RV digested pBluescript II SK(+). The resulting plasmid, pCT408, was used to transform *E. coli* DH5α. The construct was then confirmed by restriction enzyme analysis (data not shown).

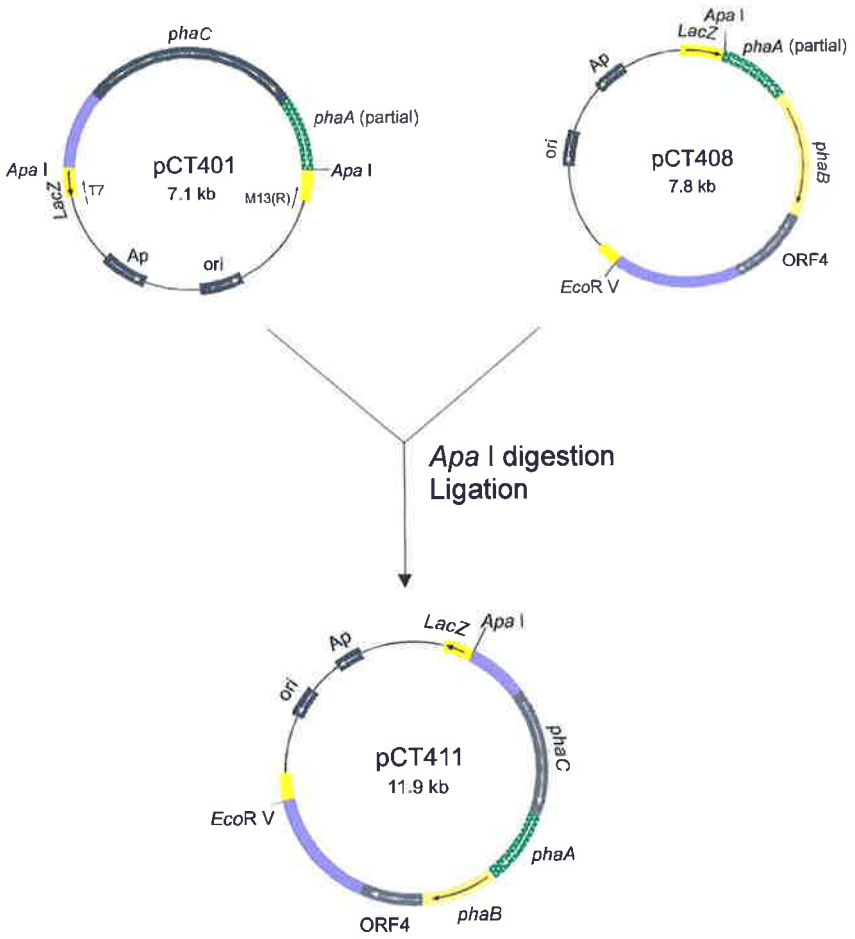


Figure 5.16. Construction of an IPTG non-inducible PHA production plasmid pCT411.

Plasmid pCT411 contains *phaC*, *phaA*, *phaB* and ORF4 under control of the native *phaC* promoter. These genes were inserted into pBluescript II SK(+) in the opposite orientation to *lacZ*.

This plasmid was constructed by digesting plasmid pCT401 with restriction endonuclease *Apa* I. The 4.1 kbp fragment harbouring *phaC* and the 5' partial *phaA* was gel purified and ligated to *Apa* I digested pCT408. The orientation of the 4.1 kbp *Apa* I insert in plasmid pCT411 was determined by PCR using the primer pair #5577 and #5578 to amplify DNA flanking the *Apa* I site in *phaA*.

To ensure that the *Apa* I restriction endonuclease recognition site within *phaA* of pCT411 was correctly reconstructed, PCR analysis using primers designed to amplify DNA flanking the *Apa* I site (#5577 5' TCG AAG GAA GGC ACG GTG AC 3', and #5578 5' ATC TGC TTG TGC ACC GCG AG 3', for their binding sites, see Table 2.3) was performed using pCT411 as the target DNA and pCT400 as a positive control. Both amplification reactions produced identical ~0.3 kbp products indicating the *Apa* I site internal to *phaA* was correctly reconstructed (Figure 5.17).

To determine whether the three *E. coli* DH5α strains harbouring pCT411 could produce PHA, they were used to inoculate NB (containing 2.0% glucose and Ap at final concentration of 100 μg/ml) and incubated at 37 °C O/N with agitation. The cells of each clone were collected, used to prepare smears on microscope slides and then stained with Nile Blue A. When examined by fluorescence microscopy, cells from all three strains were found to contain Nile Blue A stained granules. However, as a negative control, *E. coli* strain harbouring pBluescript II SK(+) vector did not produce Nile Blue A stained granules. This showed that pCT411 is a PHA production plasmid. The PHA production was also confirmed by GC analysis (Figure 5.18).

Plasmid pCT411 contains a DNA fragment encoding the *pha* operon cloned into the multiple cloning site within the *lacZ* gene of pBluescript II SK(+). The transcriptional direction of this *pha* operon was opposite to that of the *lacZ*. The detectable PHA accumulation in *E. coli* DH5α which harboured this plasmid indicated that expression of the *pha* genes, originally isolated from *Pseudomonas* strain 10c-1-3, occurred from the native promoter located at the 5' end to *phaC* (Figure 5.7a and 5.16), and not from the *lacZ* promoter.

5.2.5 Construction of IPTG inducible PHA production plasmids pCT415, pCT416, and pCT417

Plasmids capable of allowing controlled expression of *pha* genes were constructed by preparing a transcriptional fusion between plasmid DNA encoding the *pha* genes and the inducible *lac* promoter of pBluescript II KS(+). Three separate constructs were prepared: plasmids pCT415 and pCT417 contained *phaC*, *phaA*, *phaB* and ORF4 plus DNA from the 3' end of ORF4; plasmid pCT416 contained *phaC*, *phaA*, *phaB* and part of

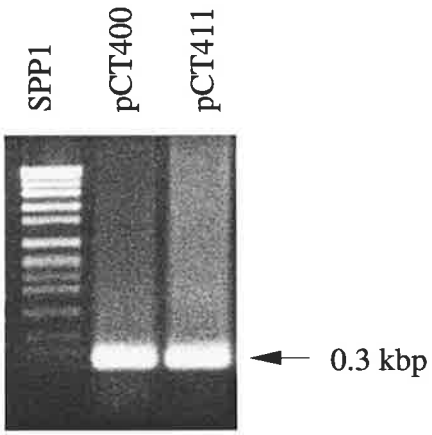


Figure 5.17. PCR products from pCT400 and pCT411 using primer pair #5577 and #5578.

Both PCR amplification reactions produced identical ~0.3 kbp products indicating the *Apa* I site internal to *phaA* was correctly reconstructed in pCT411.

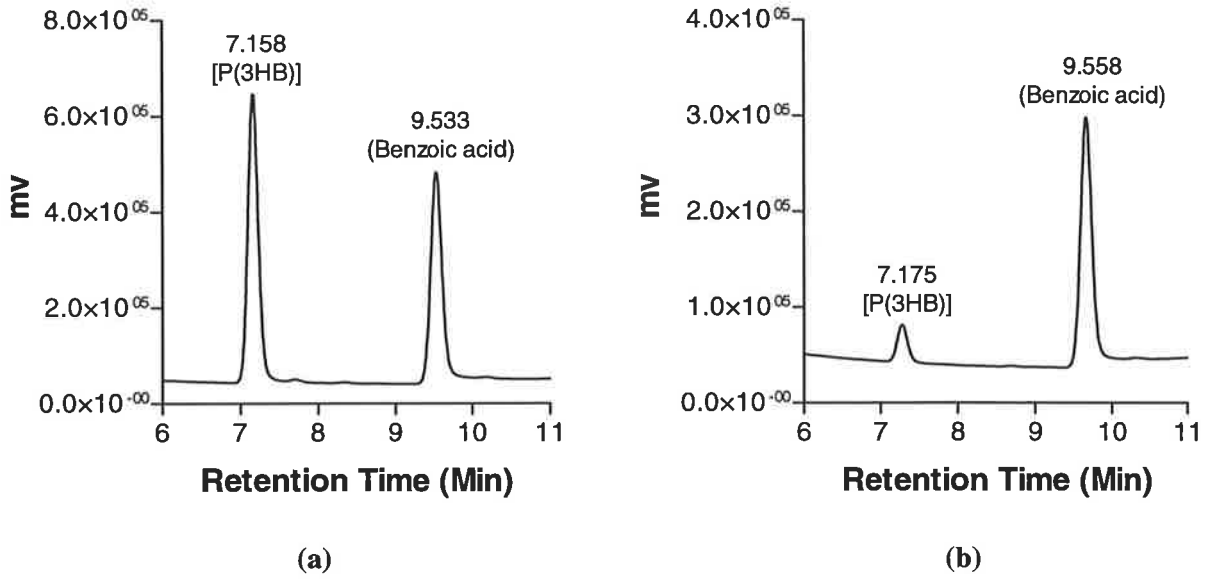


Figure 5.18. GC analysis of PHA extracts from *E. coli* DH5α harbouring plasmid pCT411.

(a) Pure P(3HB) as positive control; and **(b)** Extracts from *E. coli* DH5α harbouring pCT411. GC analysis showed that *E. coli* DH5α harbouring pCT411 produced PHA.

the 5' end of ORF4. These fusions were constructed in a two step manner analogous to that used to prepare pCT411. No attempt was made to prepare promoter fusions with optimised ribosome binding site - start codon spacings.

Plasmid pCT415 was constructed by first isolating an *Apa* I-*Eco* RV fragment carrying the 3' partial *phaA*, *phaB* and ORF4 from pCT402. This was subcloned into *Apa* I-*Eco* RV digested pBluescript II KS (+) to create pCT412. The *Apa* I fragment from pCT401, which carried *phaC* and the 5' partial *phaA* was then inserted into *Apa* I digested pCT412 to create pCT415 (Figure 5.19). This construct was confirmed by *Apa* I restriction enzyme analysis.

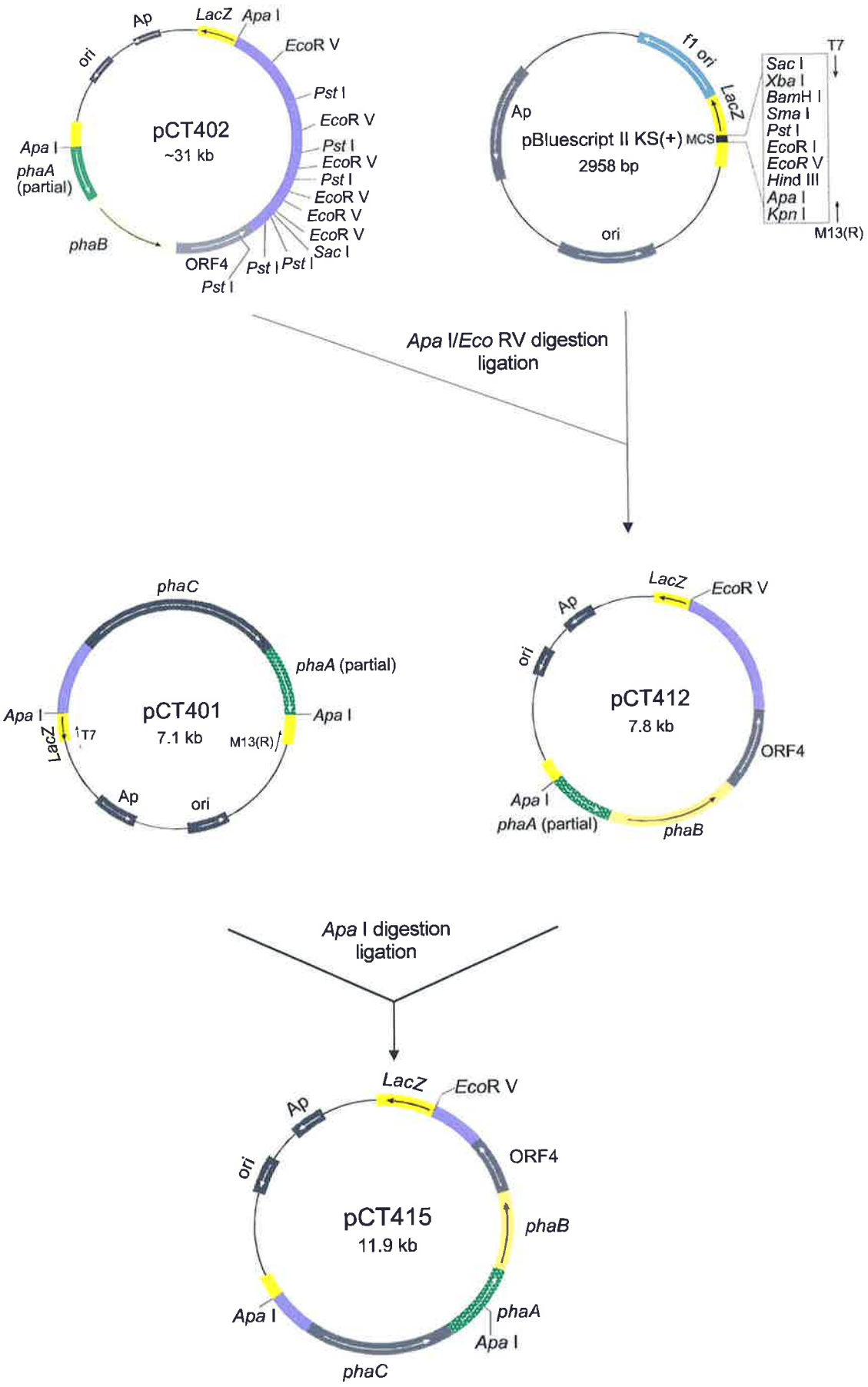
The related plasmids pCT416 and pCT417 were constructed in an identical manner, except that *Apa* I-*Pst* I and *Apa* I-*Sac* I fragments isolated from pCT402 were used as sources of 3' partial *phaA*, *phaB* and ORF4. Plasmid pCT416 differed from the others in that it contained a partial ORF4 in addition to *phaC*, *phaA* and *phaB*. Each construct was confirmed as described above.

To identify whether the *E. coli* DH5α strains harbouring pCT415, pCT416 and pCT417 respectively were capable of producing PHA, cultures of these constructs were used to inoculate 10 ml nutrient broths (containing 2.0% glucose and 100 µg/ml Ap) and incubated at 37 °C O/N with agitation in the absence of added inducer (IPTG). PHA production was determined by examination of Nile Blue A stained cell smears. All three *E. coli* plasmid combinations produced PHA, indicating that pCT415, pCT416 and pCT417 are all PHA production plasmids. No PHA was produced by pBluescript II KS(+) vector controls.

The effect of IPTG concentration on production of PHA by *E. coli* DH5α strains harbouring pCT415, pCT416 and pCT417 was then examined. The three constructs were incubated in 20 ml McCartney bottles containing 10 ml R-medium (1.5% glucose, 100 µg/ml Ap, and IPTG at a final concentration of 0.0 mM, 0.5 mM, 1.0 mM, 1.5 mM, or 2.0 mM). After incubation at 37 °C for 71 h with agitation, the PHA produced by each culture was measured using gas chromatographic analysis of cell extracts (Figure 5.20). These experiments demonstrated that the optimal PHA production was achieved by cultures in the presence of 1.0 mM IPTG. Furthermore, addition of IPTG increased the

Figure 5.19. Construction of the IPTG inducible PHA production plasmid, pCT415.

An identical strategy was used to construct the related plasmids pCT416 and pCT417 using the respective *Apa* I-*Pst* I and *Apa* I-*Sac* I fragments encoding the 3' partial *phaA*, *phaB* from plasmid pCT402. Plasmid pCT416 differs from pCT415 and pCT417 in that it contained only a partial copy of ORF4 in addition to *phaC*, *phaA* and *phaB*.


Ap
LacZ
Apa I
EcoR V
ori
Pst I
EcoR V
pCT402
~31 kb
Apa I
phaA
(partial)
Pst I
EcoR V
Pst I
EcoR V
EcoR V
EcoR V
Sac I
phaB
ORF4
Pst I
Pst I
Pst I
f1 ori
T7
Sac I
Xba I
BamH I
Sma I
Pst I
EcoR I
EcoR V
Hind III
Apa I
Kpn I
LacZ
Ap
pBluescript II KS(+) MCS
2958 bp
ori
M13(R)
Apa I/Eco RV digestion
ligation
phaC
phaA (partial)
Apa I
pCT401
7.1 kb
Apa I
LacZ
T7
M13(R)
Ap
ori
LacZ
EcoR V
Ap
ori
pCT412
7.8 kb
ORF4
Apa I
phaA (partial)
phaB
Apa I digestion
ligation
LacZ
EcoR V
Ap
ori
ORF4
pCT415
11.9 kb
phaB
Apa I
phaA
Apa I
phaC

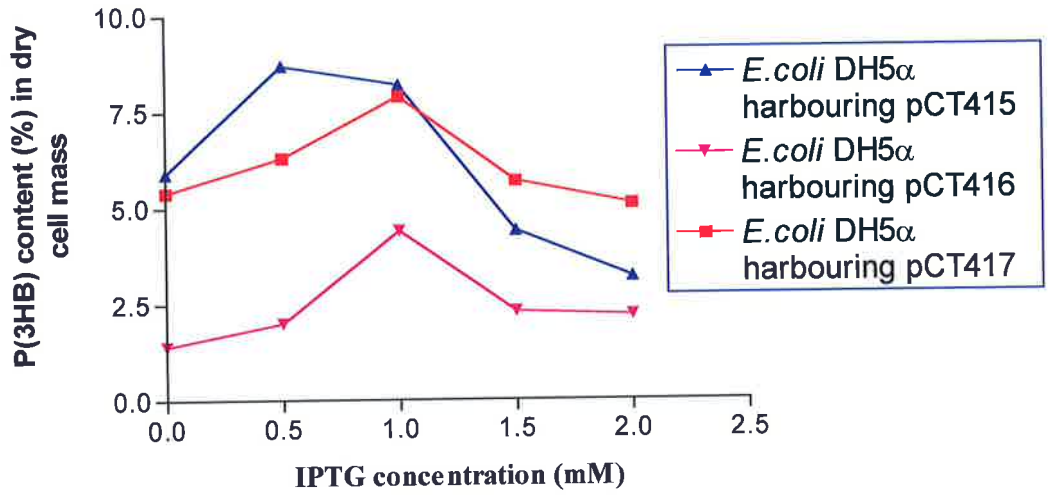


Figure 5.20. Effect of IPTG concentration on induction of PHA expression from *E. coli* DH5α harbouring pCT415, pCT416 or pCT417.

Each construct was cultured for 71 hours in R-medium (+1.5% glucose). Cells were then harvested and PHA expression was assessed as P(3HB) content (% dry cell mass) by gas chromatography.

quantity of produced PHA compared with control experiments with no added inducer. Interestingly, *E. coli* DH5α strains harbouring pCT415 and pCT417 produced similar amounts of PHA at all IPTG concentrations used, whereas *E. coli* DH5α harbouring pCT416 produced significantly less PHA.

5.2.6 Detection of the expression of *pha* genes in *E. coli*

Expression of the *pha* genes isolated from *Pseudomonas* 10c-1-3 in the recombinant host *E. coli* was tested at the transcription level and at the translational level.

Northern blot analysis was used to detect presence of *pha* mRNA. Total RNA was extracted from O/N cultures of *E. coli* DH5α harbouring pCT411, pCT415, pCT416 and pCT417 respectively, *E. coli* MD9101 harboring p4A, *Pseudomonas* 10c-1-3 and *R. eutropha*. RNA extracted from *E. coli* DH5α harboring pBluescript II SK(+) was used as a negative control.

RNA preparations were spot blotted on Nylon membrane and then hybridised with the 2.3 kbp Digoxigenin labelled *phaAB* probe isolated from plasmid p4A (see Figure 5.5) and developed as described in Section 2.17. The results are shown in Figure 5.21. The Digoxigenin labelled *phaAB* probe only hybridised with RNA extracted from *E. coli* MD9101 harbouring p4A.

To detect the product of each *pha* gene, plasmids pCT401, pCT408, pCT409, pCT410 and pCT411 were used to transform *E. coli* DH5α which harboured plasmid pGP1-2 (encodes T7 RNA polymerase under control of the temperature sensitive cI repressor). Plasmid pCT401 encode *phaC* and a partial *phaA*, pCT408 and pCT410 encode *phaB* and ORF4, pCT409 encode *phaB* and a partial ORF4, and pCT411 encode *phaC*, *phaA*, *phaB*, and ORF4. These constructs allowed high level expression from the T7 promoter located on the vector component of each plasmid (Section 2.18.2). *E. coli* DH5α harbouring pHB1 (encodes the *HlyB* gene from *Vibrio cholerae*) plus pGP1-2 was used as a positive control. *E. coli* DH5α harbouring pBluescript II SK(+) plus pGP1-2 was used as a negative control. An O/N culture of *Pseudomonas* 10c-1-3, incubated in NB (+ 2.0% glucose) was also analysed. Cell lysates were prepared for SDS-PAGE as described in Section 2.18.3.

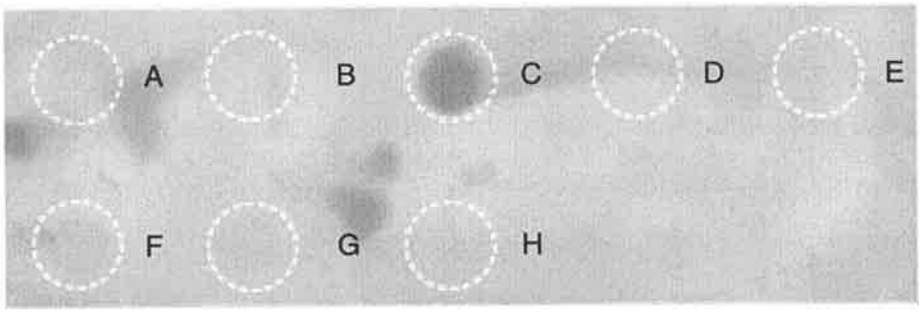


Figure 5.21. Detection of the expression of the *pha* genes in *E. coli* strains and *Pseudomonas* 10c-1-3 by Northern hybridisation using the Digoxigenin labelled *phaAB* probe (Figure 5.5).

A.: RNA extracted from *E. coli* DH5α harbouring pBluescript II SK(+).
B.: RNA extracted from *R. eutropha*.
C.: RNA extracted from *E. coli* MD9101 harbouring p4A.
D.: RNA extracted from *Pseudomonas* 10c-1-3.
E.: RNA extracted from *E. coli* DH5α harbouring pCT411.
F.: RNA extracted from *E. coli* DH5α harbouring pCT415.
G.: RNA extracted from *E. coli* DH5α harbouring pCT416.
H.: RNA extracted from *E. coli* DH5α harbouring pCT417.

Protein products equivalent in size to the predicted Pha peptides described earlier in this chapter were not obtained from either the recombinant constructs or from *Pseudomonas* 10c-1-3 (data not shown). However, a 65 kDa protein consistent with the predicted size of HlyB was obtained from cultures harbouring plasmid pHB1. This indicated that the experimental protocol used for over expression of proteins was not the reason for lack of expression of Pha peptides.

5.2.7 Improvement for PHA production in *E. coli*

Development of *E. coli* strains capable of yielding high levels of PHA requires expression of *pha* biosynthetic genes from high copy number plasmids and a cellular background which provides substrates for PHA synthesis and keeps the PHA synthesis at a high level.

In an attempt to optimise an *E. coli* host for PHA synthesis, pCT415, pCT416 and pCT417 were transformed into different *E. coli* strains including: Topp1, Topp2 (for Topp strains, see Melbourne, 1995), MD9101 (Slater *et al.*, 1992), and XL1-Blue (Lee *et al.*, 1994a,b). All of these host strains have shown to produce PHA at a high level when harbouring plasmid DNA encoding *pha* biosynthetic genes. These 12 *E. coli* strains were cultured separately in 100 ml NB (containing 2.0% glucose and 100 μg/ml Ap) in a shaking incubator at 37 °C O/N. The PHA content of each culture was examined by microscopic examination of Nile Blue A stained cell smears and gas chromatographic analysis of cell extracts. Microscopic analysis of cell smears showed that ~70% of the *E. coli* cells from each culture tested contained intracellular granules. These granules were, however quite small compared with those produced by the same host strains harbouring plasmids p4A or pJM9123. GC analysis showed that of the 12 strain plasmid combinations tested, *E. coli* Topp1 harbouring pCT417 produced the highest yield of PHA [12.7% (w/w), PHA/dry cell weight]. All other strain combinations had PHA contents of <10% (w/w, PHA/dry cell weight). PHA yields for these host/plasmid combinations are low compared with *E. coli* strains harbouring p4A or pJM9123 [50-70 % (w/w), PHA /dry cell mass] (Melbourne, 1995).

To overcome the problem of plasmid stability as described in Chapter 1 (Section 1.7.3.2), a 1.6 kbp fragment containing the *parB* locus and Kan^r gene was introduced into pCT415

(Figure 5.22). The resulting plasmid, pCT418 was then transformed into *E. coli* DH5α and tested for its stability. Plasmid stability was examined by incubation at 37 °C of *E. coli* DH5α strains harbouring pCT415 and pCT418 respectively in 5 L flasks containing 1 L NB (plus 2.0% glucose) in duplicate. One flask of each pair was supplemented with 100μg/ml Ap and the other was not. Each culture was sampled at a regular interval during incubation. Samples were diluted with saline and aliquots (0.1 ml) plated onto NA plates with no added antibiotic and on NA plates containing Ap. The percentage of the cells containing a plasmid in a culture at any time was calculated from the proportion of colonies growing on the NA plate containing Ap and the number of the colonies growing on the NA plate without added antibiotic. The results in Figure 5.23 showed that plasmid pCT418, which carried the *parB* locus, was stably maintained in *E. coli* DH5α even when cultured for 180 h in the absence of antibiotic. By contrast, pCT415, which does not contain *parB*, was not so stably maintained even when *E. coli* harbouring this plasmid was cultured in the presence of Ap.

5.3 Discussion

In this chapter, the sequencing and partial characterisation of the PHA biosynthetic genes from isolate 10c-1-3 has been described. A total of 6.6 kbp of DNA encoding all the genes necessary for biosynthesis of PHA was sequenced from plasmid pCT400 derived from the cosmid clone 50(D2) described in Chapter 4. Analysis of the sequence data has identified three open reading frames with potential to encode for polypeptides of 729, 393 and 246 amino acids. The DNA and polypeptide sequences encoded by these ORFs show significant similarity to the *phaC*, *phaA* and *phaB* genes, and protein products, described for *Burkholderia* sp., *Alcaligenes latus*, and *Alcaligenes* sp. SH-69 (DNA analysis data were not shown, for peptide analysis data see Figure 5.8, 5.11 and 5.14). Furthermore, these open reading frames are of similar size and are arranged in the same order (*phaC*, *phaA* and *phaB*) as that found in *Alcaligenes* sp. SH-69. Given the similarity of the open reading frames to previously described *pha* biosynthetic genes, a naming scheme consistent with that used for other *pha* related genes has been adopted.

On the basis of sequence similarity and the similar hydropathy profile of the inferred peptide, *phaC* is proposed to encode a polyhydroxyalkanoate synthase with a

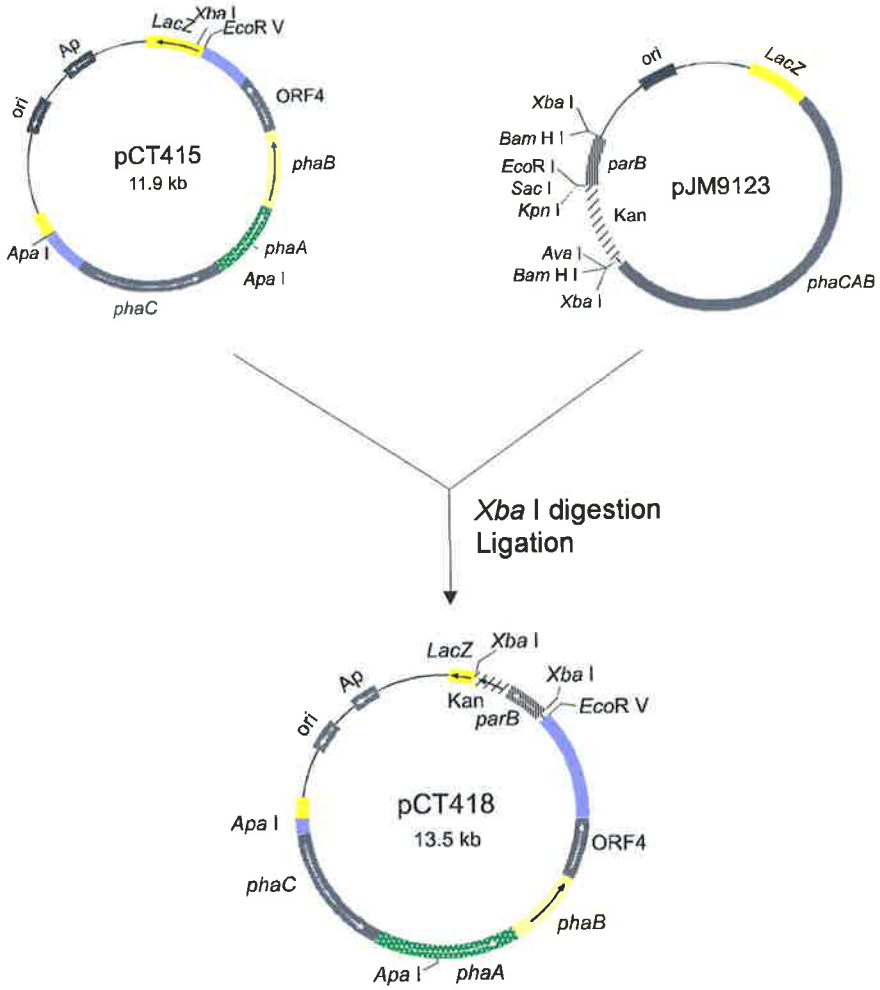


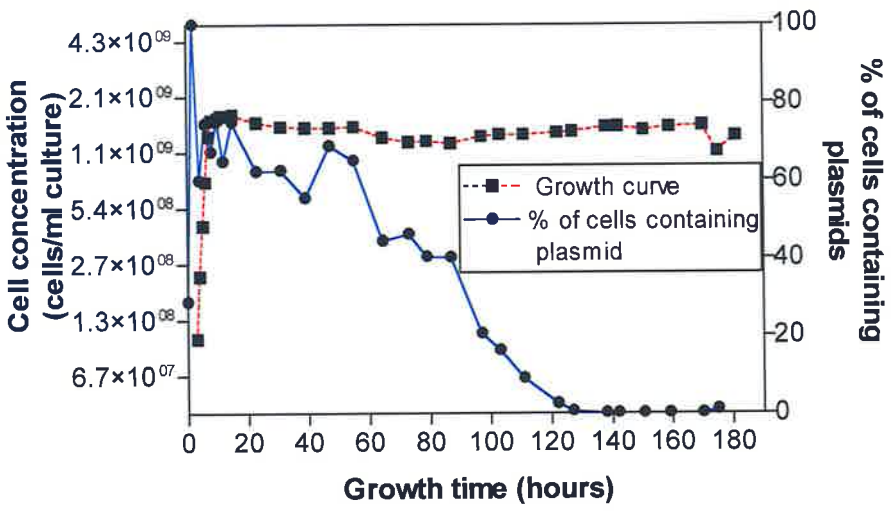
Figure 5.22. Construction of plasmid pCT418.

Plasmid pCT418 was constructed by cutting plasmid pJM9123 (Slater *et al.*, 1992) with restriction endonuclease *Xba* I and ligating the 1.6 kbp *Xba* I-*Xba* I fragment into *Xba* I digested pCT415.

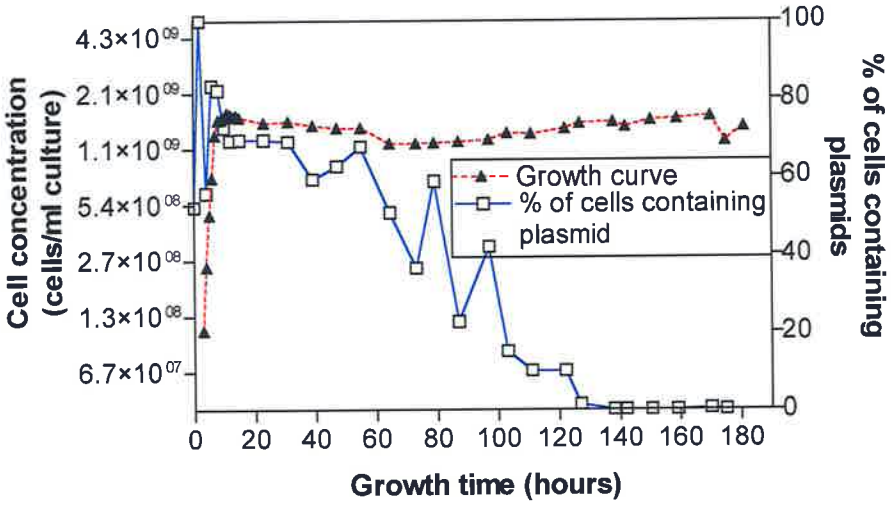
Figure 5.23. Stability of plasmids pCT415 and pCT418 in an *E. coli* DH5α host during culture in NB with or without addition of antibiotic selection. Plasmid stability was calculated by comparing the viable counts of cells obtained by plating these cells on NA with or without the same antibiotic selection.

A.: *E. coli* DH5α harbouring pCT415 cultured in NB.
B.: *E. coli* DH5α harbouring pCT415 cultured in (NB + ampicillin).
C.: *E. coli* DH5α harbouring pCT418 cultured in NB.
D.: *E. coli* DH5α harbouring pCT415 cultured in (NB + ampicillin).

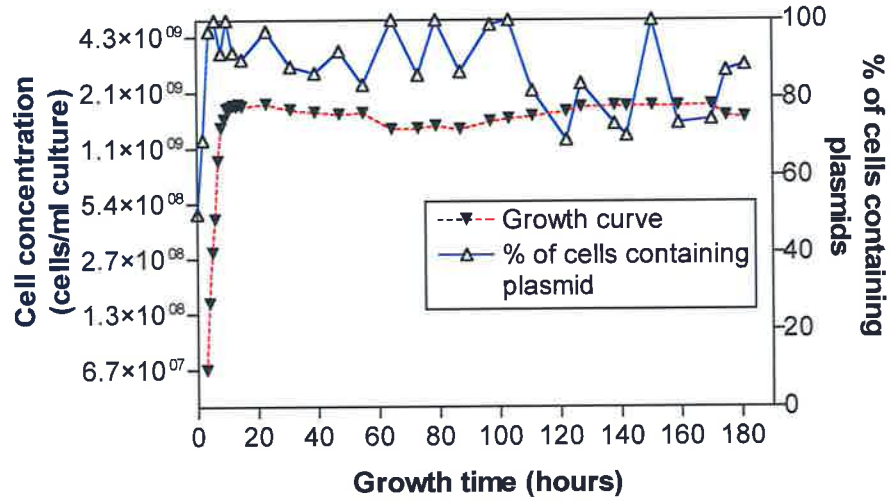
The growth curves shown represent the total viable counts obtained on NA without added antibiotic.


Cell concentration (cells/ml culture)
4.3×10^09
2.1×10^09
1.1×10^09
5.4×10^08
2.7×10^08
1.3×10^08
6.7×10^07
% of cells containing plasmids
100
80
60
40
20
0
0 20 40 60 80 100 120 140 160 180
Growth time (hours)
Growth curve
% of cells containing plasmid

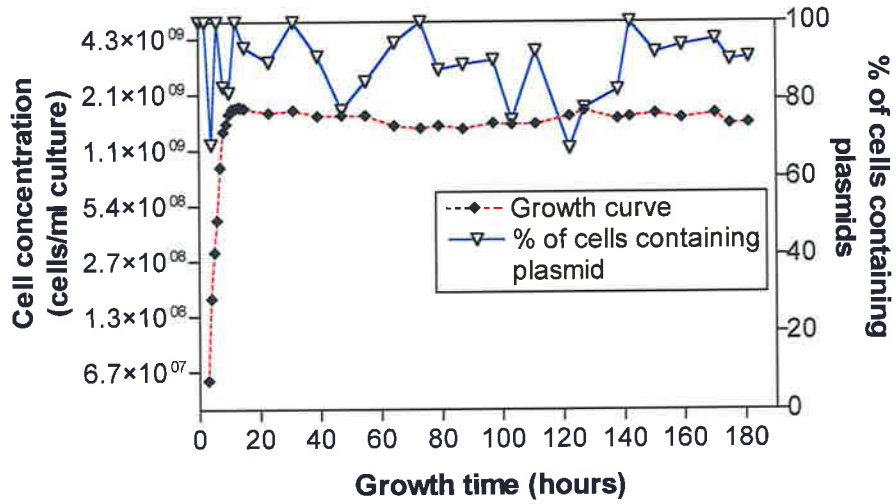
(A)


Cell concentration (cells/ml culture)
4.3×10^09
2.1×10^09
1.1×10^09
5.4×10^08
2.7×10^08
1.3×10^08
6.7×10^07
% of cells containing plasmids
100
80
60
40
20
0
0 20 40 60 80 100 120 140 160 180
Growth time (hours)
Growth curve
% of cells containing plasmid

(B)


Cell concentration (cells/ml culture)
4.3×10^09
2.1×10^09
1.1×10^09
5.4×10^08
2.7×10^08
1.3×10^08
6.7×10^07
% of cells containing plasmids
100
80
60
40
20
0
0 20 40 60 80 100 120 140 160 180
Growth time (hours)
Growth curve
% of cells containing plasmid

(C)


Cell concentration (cells/ml culture)
4.3×10^09
2.1×10^09
1.1×10^09
5.4×10^08
2.7×10^08
1.3×10^08
6.7×10^07
% of cells containing plasmids
100
80
60
40
20
0
0 20 40 60 80 100 120 140 160 180
Growth time (hours)
Growth curve
% of cells containing plasmid

(D)

molecular mass of 81.3 kDa.

The PhaA polypeptide contains three functional thiolase signature motifs characteristic of the thiolase family of proteins. The cysteine residues associated with the acyl-enzyme intermediate signature in the N-terminal region of the protein and the β-ketothiolase active site located at the C-terminus are absolutely conserved in hydrophobic regions of all the aligned polypeptide sequences shown. In addition to the remarkable conservation of sequence and functional domains in this family of proteins, there is also considerable similarity in the hydropathy profiles (Figure 5.10) and amino acid composition. Consequently, *phaA* is proposed to encode an acetoacetyl-CoA thiolase (40.6 kDa).

Using similar logic, *phaB* is proposed encode an acetoacetyl-CoA reductase (26.5 kDa). The inferred peptide sequence contains motifs characteristic of the short chain dehydrogenase/reductase family of proteins and PhaB has a conserved hydropathy profile. In addition to *phaC*, *phaA*, *phaB,* a fourth ORF of unknown function was identified. The sequence data indicated this ORF is probably transcriptionally linked to *phaCAB* to form a putative operon as shown in Figure 5.7b. The inferred peptide of this ORF shows strong similarity to the *phaF* gene of *R. eutropha* (Slater *et al.*, 1998). A more detailed discussion of the significance of ORF4 will be presented in Chapter 6.

From a phylogenetic point of view, PhaA and PhaB are very similar to other acetoacetyl-CoA thiolases and acetoacetyl-CoA reductases associated with PHA biosynthesis. Phylogenetic trees constructed from either representative PhaA or PhaB peptide sequences form clusters which generally match the broader taxonomic grouping of the host organisms. In particular, this analysis indicated that the PhaC, PhaA and PhaB identified from isolate 10c-1-3 are all most similar to representative predicted polypeptides from the β-Proteobacteria organism, *Burkholderia* (Rodrigues *et al.*, 2000).

Based on the nucleotide sequence analysis, several PHA production plasmids, pCT411, pCT415, pCT416 and pCT417 were constructed. Plasmid pCT411 in particular, harbours *phaC*, *phaA*, *phaB* and the ORF4 from *Pseudomonas* strain 10c-1-3. The *pha* genes in this plasmid were cloned into the vector in the opposite orientation to *lacZ*. In this orientation, it was expected that expression of the *pha* genes should be independent of the *lacZ* promoter and hence not subject to induction by IPTG. The fact that PHA is

constitutively expressed by *E. coli* cells harbouring this construct, suggested the presence of a functional promoter upstream of the *phaC* gene. This promoter could be vector derived, or it could be the putative *phaC* associated promoter shown in Figure 5.7. The -10 and –35 regions of this promoter are similar to that of the consensus σ^{70} *E. coli* promoter and hence it is possible that transcription of an mRNA encoding the identified *pha* genes could be initiated from this region. However, *E. coli* cells harbouring cosmid pCT400 (from which pCT411 was derived), are unable to produce PHA, although this construct contains all the necessary *pha* genes and the putative native promoter (see Section 4.2.2.3). This evidence indicated transcription of the *pha* genes located on pCT400 may occur only weakly, or not at all from the native promoter. However an alternative explanation is that transcription or translation of the *pha* genes was inhibited in this construct by a product of another gene(s) (eg. a repressor) located in the insert of pCT400. Nevertheless, irrespective of the potential transcriptional initiation sites mentioned, future work should employ primer extension analysis to confirm transcription of the *pha* genes from the putative promoter identified by analysis of the DNA sequence data.

Plasmids pCT415 and pCT417 encode *phaCAB* and ORF4. Plasmid pCT416 is identical except that it encodes only a partial ORF4. The transcriptional orientation of *pha* genes in all plasmids is identical to that of vector derived *lacZ*. Hence expression of the *pha* genes in these plasmids should be IPTG inducible. Unexpectedly, PHA was produced in the absence of inducer. However, in the presence of inducer (1.0 mM IPTG), PHA production by *E. coli* DH5α harbouring pCT415 and pCT416 when grown in R-medium, was significantly enhanced. This result is surprising given that glucose present in the growth medium would be expected to result in repression of transcription from the *lacZ* promoter (Magasanik and Neidhardt, 1987). It is possible that at the end of the culture period, glucose levels were sufficiently reduced to allow this catabolite repression to be overcome. Unfortunately in these preliminary experiments, glucose concentrations were not monitored.

The fact that PHA was produced by these constructs, even in the absence of IPTG, lends support to the hypothesis that transcription of the cloned *pha* genes can be initiated from a promoter located on cloned DNA upstream of *phaC*. Furthermore, since PHA is expressed by *E. coli* strains harbouring a plasmid lacking a complete ORF4 (pCT416), this suggests ORF4 is not essential for PHA biosynthesis.

Intriguingly, experimental evidence to demonstrate transcription and translation of the *pha* genes located on plasmids constructs prepared in this work (see Section 5.2.6) was not demonstrated. Northern blot analysis showed a 2.3 kbp DNA probe isolated from plasmid p4A only hybridised with the RNA extracted from *E. coli* MD9101 which harboured p4A, and not RNA prepared from constructs harbouring *pha* genes cloned from *Pseudomonas* 10c-1-3. These unexpected results could be due to a number of reasons: (1) yields of *pha* related mRNA in the *E. coli* strains harbouring the PHA production plasmids, pCT411, pCT415, pCT416 and pCT417 were low whereas the amounts of mRNA transcribed from the *pha* genes located on plasmids p4A was high. This could be the result of inefficient transcription from the putative promoter identified. This hypothesis correlates with low level PHA production by *E. coli* strains harbouring those PHA production plasmids, compared with high level PHA production by *E. coli* strains harbouring p4A; (2) the extracted RNAs degraded after extraction and during Northern blot, due to low stability and RNase contamination; (3) the DNA probe was less sensitive than a RNA probe for Northern blot.

Attempts to over express the PHA biosynthetic enzymes by transcription from vector derived T7 promoter also failed to demonstrate significant levels of protein of sizes consistent with the predicted molecular masses of PhaCAB. Over expression of a positive control (HlyB) was successful and this indicated that the methodology used was not flawed. Why the protein products of the *pha* genes were not detected in *E. coli* DH5α strains harbouring plasmids pCT401, pCT408, pCT409, pCT410 and pCT411 respectively is not known. However, these results could suggest that the levels of these proteins in PHA producing cells are low.

Nevertheless, even though the expression of the *pha* genes from *Pseudomonas* strain 10c-1-3 was not detected with SDS-PAGE and Northern blot, *E. coli* harbouring the PHA production plasmids, pCT411, pCT415, pCT416 and pCT417 express the *pha* genes in *E. coli* and this results in detectable quantities of PHA (in those plasmids encoding *phaCAB).* PHA was not detected in *E. coli* strains harbouring pCT401 which encodes a partial *Pseudomonas* strain 10c-1-3 *pha* operon, indicating *phaC*, *phaA* and *phaB* are essential for PHA production. Furthermore, as mentioned previously, PHA production was identified in *E. coli* strains, which harboured PHA production plasmids with a whole ORF4

or a partial ORF4, indicating the ORF4 is not essential for PHA biosynthesis. However, the function and relationship of ORF4 to PHA biosynthesis is unknown.

In this study, the effect of host strain on expression of PHA was evaluated. PHA production by different combinations of *E. coli* strains (Topp1, Topp2, MD9101 and XL1-blue) and PHA production plasmids (pCT415, pCT416 and pCT417) was examined. As with other related studies (Melbourne, 1995) host strain effects were noted. In this work, *E. coli* Topp1 harbouring pCT417 produced the highest yield of PHA [12.7% (w/w), PHA/dry cell weight]. Nevertheless, this PHA yield is significantly lower than that reported elsewhere (see Table 1.3).

One of the major causes of low PHA yields has been shown to be the stability of expression vectors in recombinant hosts. This study is no exception. Standard antibiotic selection procedures were inadequate and significant proportions of cells in culture lacked plasmid DNA. In attempts to overcome the effect of plasmid stability, the *parB* locus was introduced into pCT415 and pCT418 was constructed. Previous studies (Slater *et al.*, 1992; and Lee *et al.* 1994a,b,c) have shown inclusion of the *parB* locus on PHA production plasmids to significantly improve plasmid stability and consequently PHA yields. The presence of *parB* locus in pCT418 significantly increased the stability of this plasmid in an *E. coli* DH5α background. Even after extended culture, plasmid was retained in >80% of viable cells, whereas the proportion of cells able to retain plasmid lacking the *parB* locus (pCT415) was quite small.

Chapter 6

Investigation of the function of ORF4

6.1 Strategy

Chapter 5 described the identification of an open reading frame, ORF4 (or *phaD*) located immediately 3' of *phaB* in the *pha* operon cloned from *Pseudomonas* strain 10c-1-3. BLAST analysis showed that the inferred peptide of this ORF is homologous to the *phaF* gene from *R. eutropha* (Slater *et al.*, 1998) with an identity of 82% at the protein level. However, the peptide shares significant similarity to a number of other peptides thought to associated with PHA synthesis (Accession numbers: AF153086, U41265, P45368, P45373, S29377).

The lack of any transcriptionally important sequence motifs between PhaB and the putative ORF4 peptide indicates ORF4 is likely to be transcriptionally linked to the other *pha* genes and probably is transcribed from the promoter for *phaC*. Nevertheless, PHA production by *E. coli* strains harbouring PHA production plasmids with or without this ORF4 indicates the putative peptide is either not involved in PHA synthesis (Section 5.3) or not essential for production of core PHA. If indeed the ORF4 product is linked to PHA formation in cells, it is possible that the protein is either involved in post PHA synthesis processing such as modification of core PHA or in the initiation and formation of PHA granules within the cell. Consequently, if ORF4 plays a role in the latter process, then mutations within ORF4 may affect the appearance and number of PHA granules within the producer cell.

In this chapter, the construction of an insertion mutation within ORF4 (ORF4::*kan*) in wild type *Pseudomonas* strain 10c-1-3 is described. Briefly, DNA carrying the mutated ORF4 was prepared and cloned into a suicide vector, pCVD442. This construct was then transferred to *Pseudomonas* 10c-1-3 from *E. coli* by RP4 mediated conjugation. Transconjugants in which allelic replacement of ORF4::*kan* had occurred were isolated. PHA production by wild type and mutant strains was then examined by both GC analysis and transmission electron microscopy (TEM).

6.2 Results

6.2.1 Preparation of the mutant *Pseudomonas* strain 10c-1-3 PS 002

6.2.1.1 Cloning of ORF4 into pGEM-T

Examination of the DNA sequence data for ORF4 and flanking sequence indicated that there are no useful restriction endonuclease recognition sites which enable subcloning of ORF4 into the suicide vector, pCVD442. PCR amplification was therefore chosen as the method for introduction of DNA encoding ORF4 into the *Sph I*/*Sal* I digested suicide vector. PCR primers incorporating an *Sph* I (#3578) site and an *Sal* I (#3579) site into amplified ORF4 were constructed (see Table 2.3) and used for PCR amplification of a ~ 0.7 kbp ORF4 fragment from pCT408. The PCR product was gel purified and ligated to vector pGEM-T. The DNA was used to transform *E. coli* DH5α. Plasmid DNA was isolated from transformants and clones containing insert in both orientations were selected and designated pCT419 and pCT420 (Figure 6.1). Each construct was confirmed by restriction analysis with *Pst* I, *Sph* I and *Sal* I. The orientation of the insertion in pGEM-T was determined by PCR using primer #3578 and the sequencing primer M13(R).

The sequence analysis described in Figure 5.7 showed that the amplified ORF4 fragment present in pCT419 contained a unique *Pst* I restriction site. This site was chosen for construction of a Kan^r cartridge insertion mutation. To eliminate the *Pst* I recognition site from the pGEM-T vector component of pCT419, pCT419 was restricted with *Sal* I. This digest removed the *Pst* I site located between the *Sal* I site introduced on the ORF4 PCR amplicon and the *Sal* I site located in the multiple cloning site of pGEM-T. Digested plasmid pCT419 was then religated to form pCT421 (Figure 6.2).

Overexpression of the ORF4 encoded product was carried out with pCT419 and the inducible T7 RNA polymerase system (see Section 2.18). The ORF4 encoded product was not significantly detected with SDS-PAGE for unknown reasons.

6.2.1.2 Cloning of a Kan^r cartridge from pBSL15

The Kan^r cartridge encoded on pBSL15 (Alexeyev, 1995) was modified to allow isolation of the *kan* gene on a *Pst* I fragment. *Pst* I sites were introduced on either side of the *kan* gene by PCR using primers #3623, and #3624 (Table 2.3). PCR amplification of

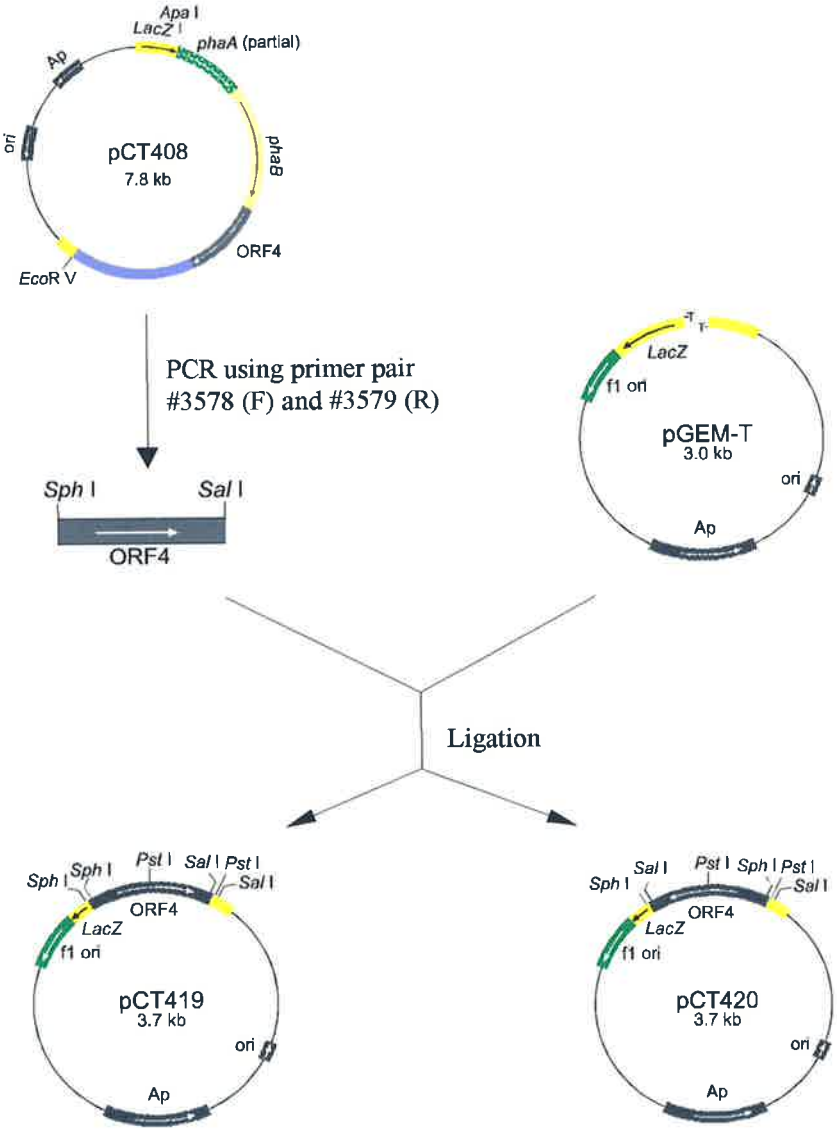


Figure 6.1. Isolation of ORF4 on a *Sph* I-*Sal* I fragment and construction of pCT419 and pCT420.

PCR using primer pair #3578 and #3579 with pCT408 as a DNA template, was used to amplify a ~0.7 kbp fragment encoding ORF4. The primers were used to introduce *Sph* I and *Sal* I restriction sites. The PCR product was gel purified, ligated into pGEM-T, transformed into *E. coli* DH5α and plated onto agar which contained Ap, IPTG and X-gal. White Ap^r clones were selected and clones containing the ORF4 amplicon confirmed by restriction analysis. PCR using primer pair #3578 and M13(R) was used to determine the orientation of the insert in pGEM-T. Plasmid DNA isolated from white Ap^r colonies were designated pCT419 and pCT420 respectively, as shown above.

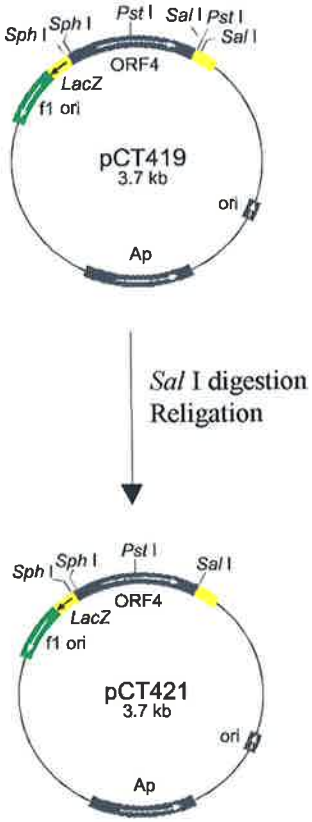


Figure 6.2. Construction of pCT421.

Plasmid pCT419 was digested with *Sal* I, religated, transformed into *E. coli* DH5α and plated onto agar which contained Ap. The clones were further identified by restriction analysis with *Pst* I. Plasmid DNA isolated from a single colony was designated pCT421 as shown above.

DNA from pBSL15 resulted in a product of ~1.1 kbp in size. This product was cloned into pGEM-T to produce plasmid pCT422 (Figure 6.3). White, Ap^r and *LacZ* negative *E. coli* DH5α clones were selected from the transformants of DNA ligation. Plasmid pCT422 was isolated from *E. coli* DH5α clones and the construct confirmed by restriction analysis with *Nco* I/*Sal* I and PCR using primer pair #3623 and #3624.

6.2.1.3 Cloning of Kan^r cartridge into pCT421

pCT423, pCT424 were constructed by ligating the *Pst* I fragment from pCT422 into *Pst* I digested pCT421. Details of the construction are shown in Figure 6.4. White, Ap^r, Kan^r and *LacZ* negative clones were selected from the transformants of ligation. Plasmids pCT423 and pCT424 isolated from *E. coli* DH5α were identified by restriction analysis with *Pst* I and *Nco* I and PCR using primer pair #3579 and #3623.

6.2.1.4 Construction of a mobilisable vector for the ORF4::*Kan* mutation

The 1.8 kbp *Sph* I/*Sal* I ORF4::*Kan* fragment located on pCT423 was subcloned into the mobilisable suicide vector pCVD442 as shown in Figure 6.5. Plasmid pCT425 was then used to transform *E. coli* SM10 λpir. This construct was then transferred to *Pseudomonas* 10c-1-3 by conjugation.

6.2.1.5 Conjugation and preparation of a mutant of wild type *Pseudomonas* strain 10c-1-3

A streptomycin (Sm) resistant mutant (*Pseudomonas* strain 10c-1-3 PS001) was obtained by incubation of the wild type culture in NB O/N at 37 °C with agitation followed by selection on NA containing Sm at 100 μg/ml. One Sm^r colony (PS001) was selected and used as a recipient in the conjugal mating experiment describe below.

Plasmid pCT425 carrying the ORF4::*kan* mutation was transferred from an *E. coli* SM10 λpir host to *Pseudomonas* strain 10c-1-3 PS001 by conjugation as described in Section 2.14. A *Pseudomonas* strain 10c-1-3 PS001 mutant was selected on NaCl free NA (containing sucrose of 5.0%, Kan at 50 μg/ml, and Sm at 100 μg/ml). pCVD442 (Ap^r, *sacB*, *mob*RP4) borne plasmid pCT425 is a suicide vector, and the host harbouring this plasmid could not survive with the presence of sucrose (5.0%). This characteristic of

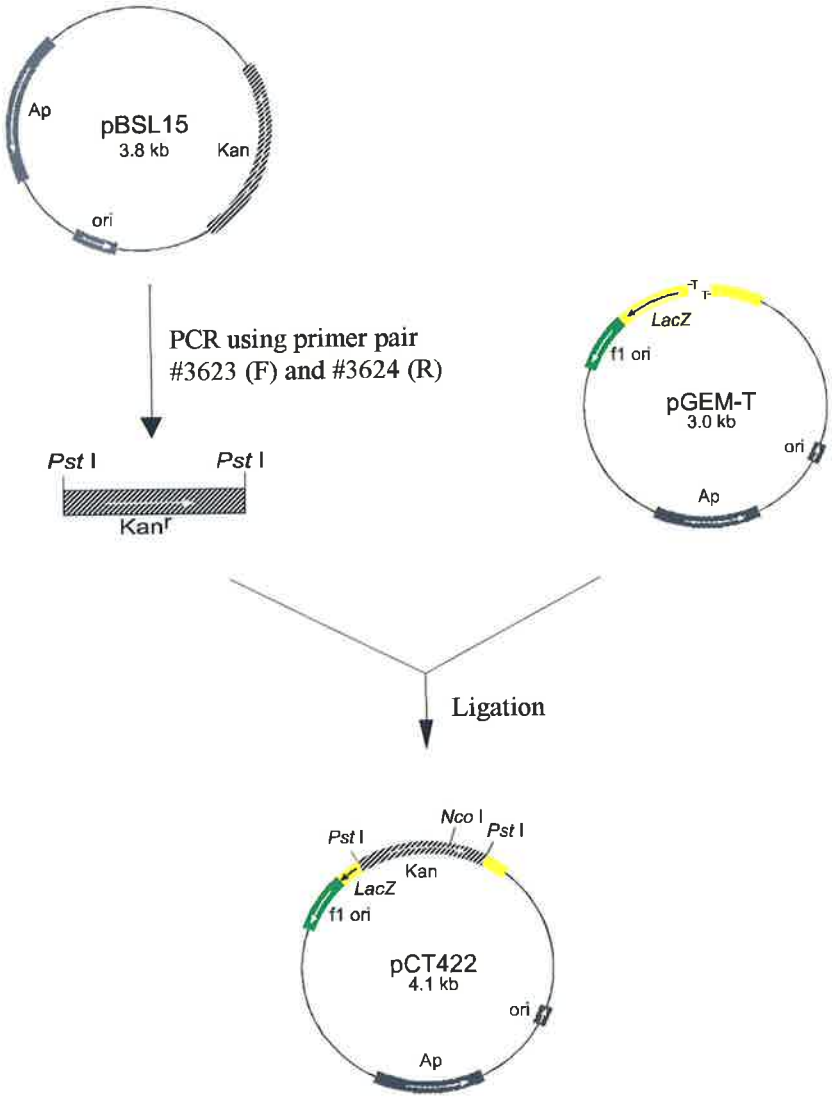


Figure 6.3. Construction of a *kan* cartridge located on a *Pst* I fragment in pCT422.

PCR, using primer pair #3623 and #3624, and DNA template pBSL15, was used to amplify an ~1.1 kbp fragment with a *Pst* I restriction site at each end. This PCR product was gel purified, ligated with pGEM-T. The ligation mixture was used to transform *E. coli* DH5α and transformants were plated onto agar which contained Ap, IPTG and X-gal. White and Ap^r clones were further identified by restriction analysis with *Nco* I and *Sal* I and PCR using primer pair #3623 and #3624 for insertion orientation identification. Plasmid DNA isolated from a single white Ap^r colony was designated pCT422 as shown above.

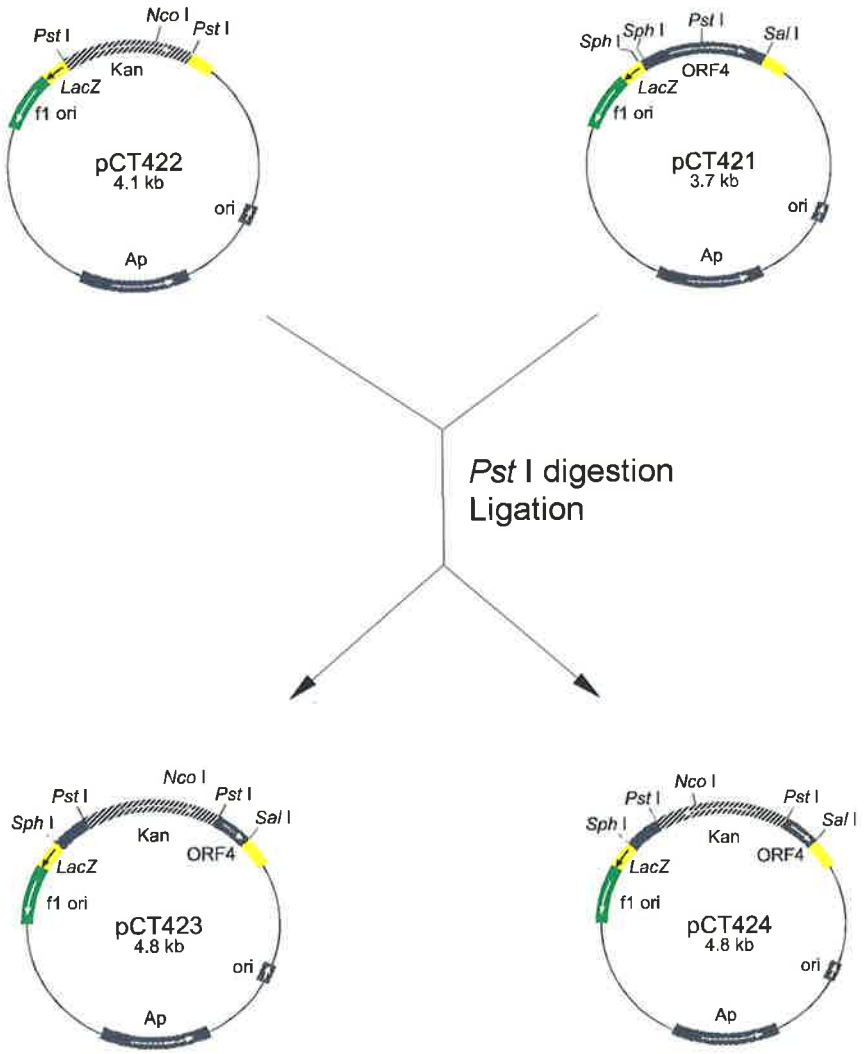


Figure 6.4. Construction of ORF4 insertion mutations encoded on plasmids pCT423 and pCT424.

PCT422 was completely digested with *Pst* I, and the 1.1 kbp *Pst* I fragment was gel purified and ligated with *Pst* I digested pCT421. The ligation mixture was used to transform *E. coli* DH5α and transformants were plated onto NA which contained Ap and Kan. White, Ap^r and Kan^r clones were further identified by restriction analysis with *Pst* I and *Nco* I and PCR using primer pair #3579 and #3623 to determine the orientation of the *kan* insertion. Plasmid DNAs isolated from a single Ap^r and Kan^r colony were designated pCT423 and pCT424 respectively, as shown above.

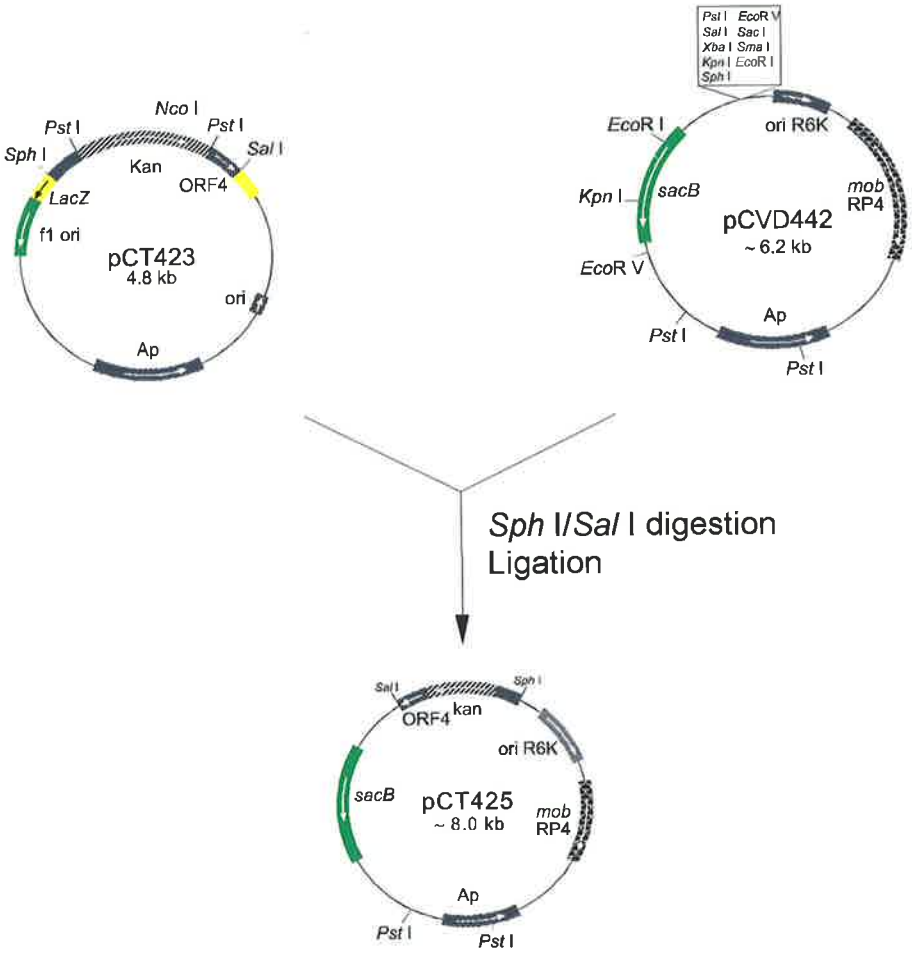


Figure 6.5. Construction of pCT425.

PCT423 was completely digested with *Sph* I and *Sal* I. The 1.8 kbp *Sph* I-*Sal* I fragment was gel purified, ligated with *Sph* I/*Sal* I digested pCVD422. The ligation mixture was used to transform *E. coli* SY327 λpir and transformants plated onto NA which contained Ap and Kan. Ap^r and Kan^r clones were further identified by restriction analysis with *Sph* I, *Sal* I and *Pst* I. Plasmid DNA isolated from a single Ap^r and Kan^r colony was designated pCT425, as shown above.

pCT425 indicated that the above mentioned mutant was plasmid free and the allelic replacement of ORF4::*kan* had occurred.

For the identification of the selected mutant, genomic DNA was prepared and PCR reactions were carried out. When primer pair #3512 (F), which binds to *phaB*, and #3624 (R), which binds to the Kan^r cartridge were used, a product of 1.8 kbp was obtained and When primer pair #3512 and #3579 were used, a 2.0 kbp product was obtained (for the binding sites of these primers, see Table 2.3). DNA sequencing of this 2.0 kbp PCR product using primer #3722 showed that the PCR product included both ORF4 (partial) and Kan^r cartridge (partial) (Figure 6.6), indicating that ORF4 was mutated by insertion of Kan^r cartridge in the selected colony of *Pseudomonas* strain 10c-1-3 PS001. This mutant was designated *Pseudomonas* strain 10c-1-3 PS002.

6.2.2 PHA production by wild type *Pseudomonas* strain 10c-1-3 and the mutant *Pseudomonas* strain 10c-1-3 PS002

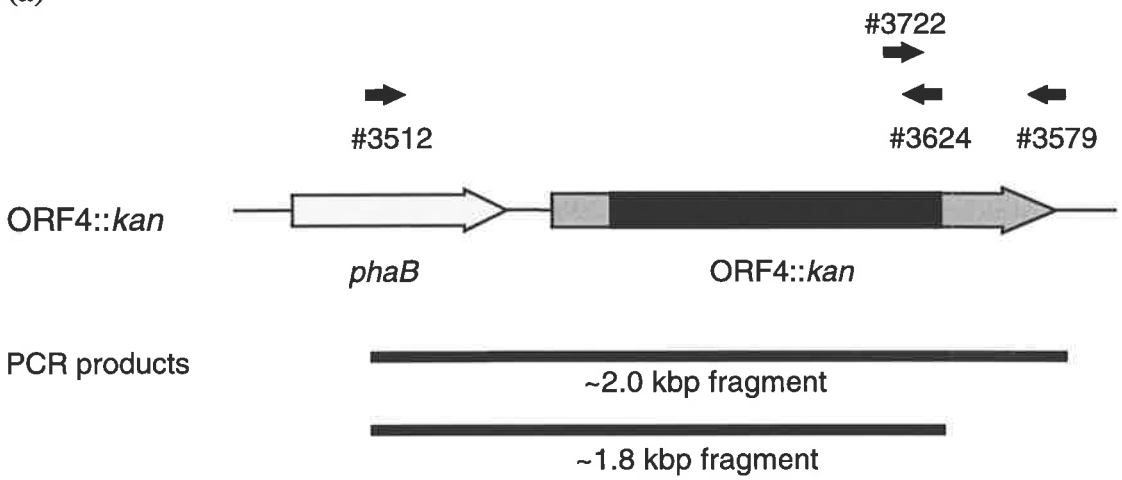
Wild type *Pseudomonas* strain 10c-1-3 and its mutant *Pseudomonas* strain 10c-1-3 PS002 were incubated in 1,000 ml flasks containing 200 ml R-medium (supplemented with 2% of glucose and kan at 50 μg/ml) for 2 days at 30 °C with agitation.

PHA production by the wild type and mutant was tested by examination of cells for the presence of PHA granules by microscopy of Nile Blue A stained smears. PHA granules were found in cells from both cultures. The presence of PHA was also tested by GC analysis and transmission electron microscopy of sections of resin embedded cells.

GC assay showed that wild type *Pseudomonas* strain 10c-1-3 produced P(3HB) at a level of 16.8% [w/w, P(3HB)/dry cell weight], and the mutant *Pseudomonas* strain 10c-1-3 PS002 produced P(3HB) at a level of 32.5% [w/w, P(3HB)/dry cell weight]. The GC profile of PHA extracted from the wild type (10c-1-3) was identical to that obtained for the PS002, the strain carrying the ORF4::*kan* mutation (Figure 6.7).

Transmission electron microscopy (TEM) analysis was used to examine stained sections of cells, prepared from each of the two strains. Sections of wild type *Pseudomonas* strain 10c-1-3 cells contained large PHA granules with distinct borders, arranged in clusters of up to six granules per cell (Figure 6.8). Counts of granules per cell

(a)



(b)

kan (partial) →

```
  1  CTGATCGCTT CCTCGCGCTT TACGGTATCG CCGCTNNCGA TTCGCAGCGC

                                          Stop(kan)
 51  ATCCCCTTCT ATCGCCTTCT TGACGAGTTC TTCTGAGCGG CTGCAGATCA
                                                 Pst I
```

ORF4 (partial) →

```
101  TCCTCGAAGA GGAGAGCGGC GGGCTGCCGA TGTTCTCGTC GGTGATGCTT
151  TCGCAGATCA TCCGTTTCTA CGGCCATGCG ATGCAGGGCA TGATGGGCAC
201  GTATCTGGAA AAGAACATCC AGGCGTTCAT CGACATCCAG CAGAAGCTCA
251  CCGATCAGAG CAAGGGCCTT TACGACGGCA ATGCGCTCAA CCCCGAAGTC
301  TGGTCGCAGT TCATGAACAT GCAGGCGCCG ATGATGCAGG GCATGATGAC
351  GAGCTACATC GAGCAGTCGA AGAACATGTT CGTGCAGATG CAGGAGCAGA
401  TGCAGTCGCA GGCGAAGTCG ATGTTCAACA CGTTTCCGTT TCCGCCGCCC

                              Stop(ORF4)
451  GCCACGCCGC CTTCGGAAAA GAAGTAATCG TTCTCAGGCG TCTCG
```

← #3579(R)

Figure 6.6. DNA sequence analysis of a PCR product from *Pseudomonas* strain 10c-1-3 PS002 which spans the chromosomal ORF4*::kan* insertion mutation.

(a) Diagramatic representation of primer binding sites and corresponding PCR products obtained from *Pseudomonas* 10c-1-3 strain PS002.

(b) DNA sequence data obtained from the 2.0 kbp PCR product amplified from *Pseudomonas* strain 10c-1-3 PS002 using primer pair #3512 (F) and #3579 (R). The data shown represents the 3' junction between the kan gene insert and ORF4. The *Pst* I site designating the 3' end junction is shown. This sequence data confirmed that the kan gene was inserted in the correct orientation and position within ORF4 of PS002.

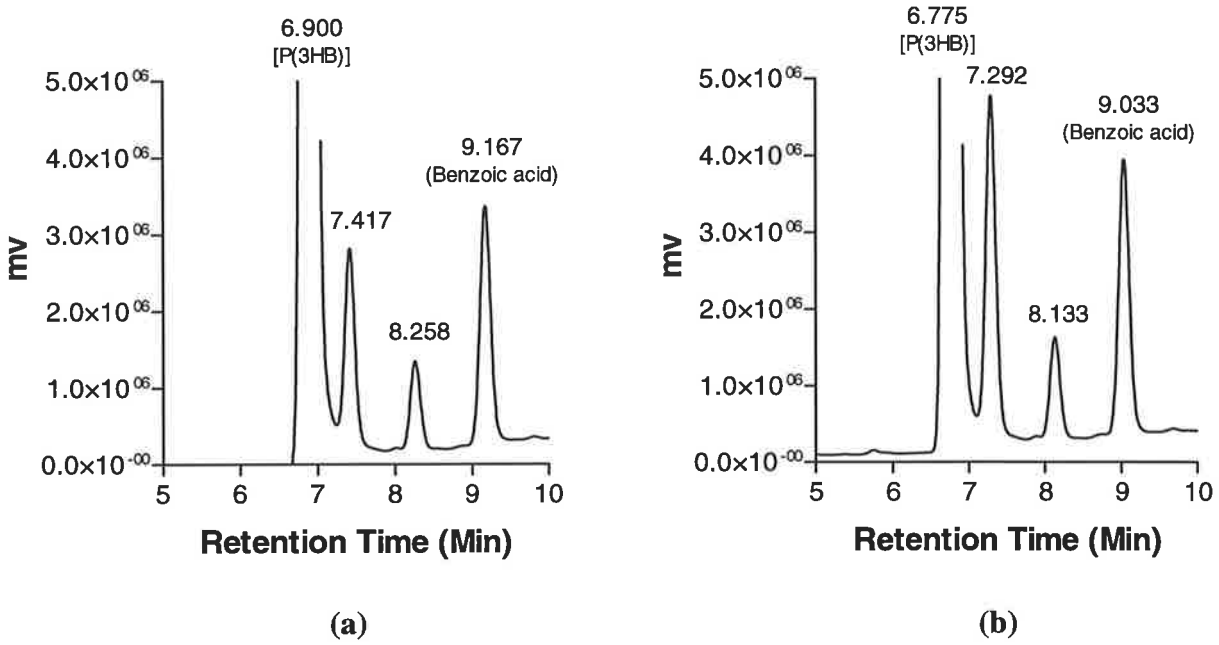
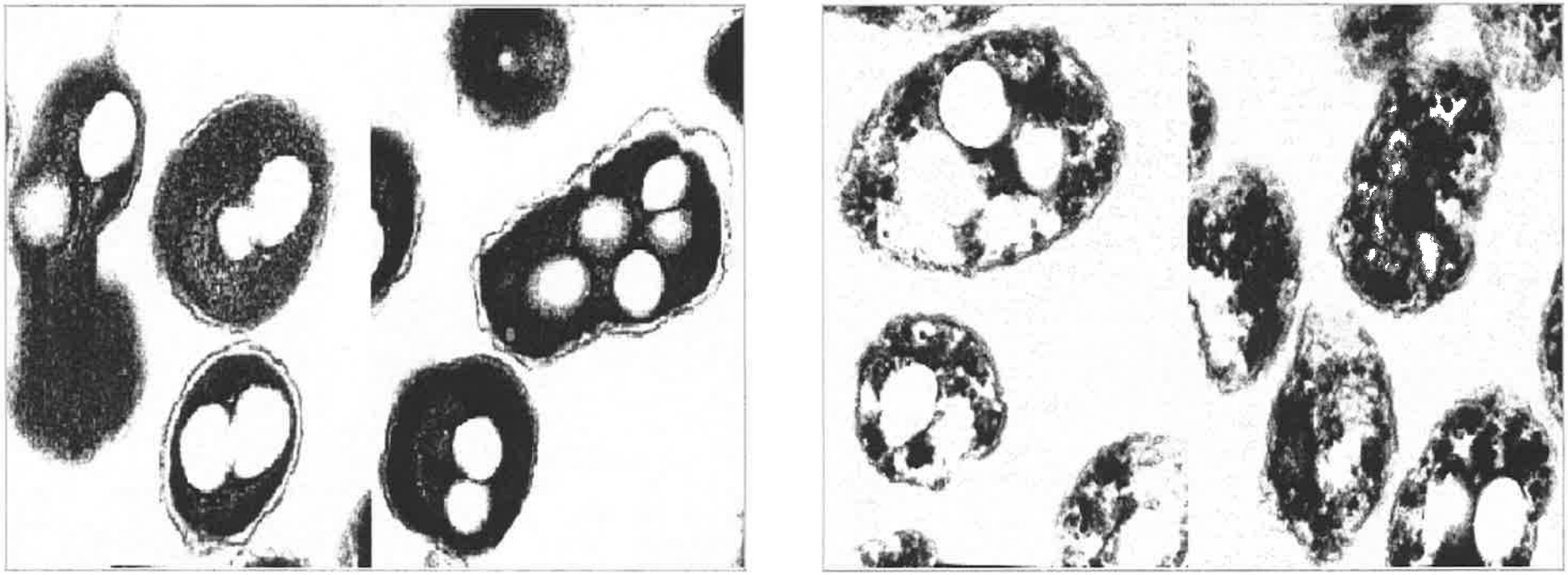


Figure 6.7. GC profiles of the PHA extracted from the wild type of *Pseudomonas* sp. 10c-1-3 and the ORF4 mutant, PS002.

(a) wild type *Pseudomonas* sp. 10c-1-3.
(b) mutant *Pseudomonas* strain 10c-1-3 PS002.



(a) (b)

Figure 6.8. TEM micrographs of PHA granules in the cells of wild type and the ORF4 mutant strains of *Pseudomonas* strain 10c-1-3.

(a) Wild type *Pseudomonas* strain 10c-1-3. Note the discrete and well defined borders of the intracellular PHA granules (magnification 41,000 ×).

(b) *Pseudomonas* strain 10c-1-3 PS002. Note that some of the intracellular PHA granules have borders less well defined than that of granules present in the wild type strain (magnification 41,000 ×).

section varied from 1 to 6 (mean =2.6 granules per cell section, n = 38 cell sections), with almost all cell sections examined containing at least one identifiable granule. By comparison, sections of cells prepared from the mutant, PS002, contained from 1 to 5 granules per cell (mean = 1.9 granules per cell section, n = 40 cell sections), and a significant proportion of cell sections contained no granules (20%). Moreover, the granule borders were less distinct. In addition, while the cytoplasm of wild type cell sections contained densely stained material characteristic of ribosomes, sections of cells prepared from the mutant strain contained only a lightly stained cytoplasmic region. To determine whether the number of granules per cell were significantly different for the wild type and mutant strain, a standard Student t-test was used to compare the mean number of granules per cell section. Approximately 40 cell sections of each strain were randomly selected and the granule numbers were determined. Statistical analysis showed that the means were likely to be different at the 95% confidence interval (Figure 6.9).

6.3 Discussion

In this study, an open reading frame (ORF4) downstream from *phaB* has been identified. Sequence analysis described in Chapter 5 suggests ORF4 is transcriptionally linked to the putative operon including the genes *phaCAB*. Although homologs of the putative polypeptide sequence encoded by ORF4 are available in public peptide databases and are in general thought to be associated with PHA biosynthesis, only a single entry (*phaF*, Slater *et al.* 1998) has been well described. In view of the possibility that ORF4 may encode a polypeptide involved in PHA biosynthesis/assembly, a mutation in ORF4 (ORF4::*Kan*) was constructed and this introduced into *Pseudomonas* 10c-1-3 using allelic replacement technology, as a first step in the functional analysis of this open reading frame.

PHA production by the wild type of *Pseudomonas* strain 10c-1-3 and the ORF4 mutant, *Pseudomonas* strain 10c-1-3 PS002 was tested by culture in R-medium. Cells from each culture were isolated and PHA expression tested using GC analysis and transmission electron microscopy of sections of embedded cells. Unexpectedly, GC analysis showed that PHA production level by the mutant strain was higher than that observed for the wild type culture. Morphologically, TEM analysis showed that the PHA

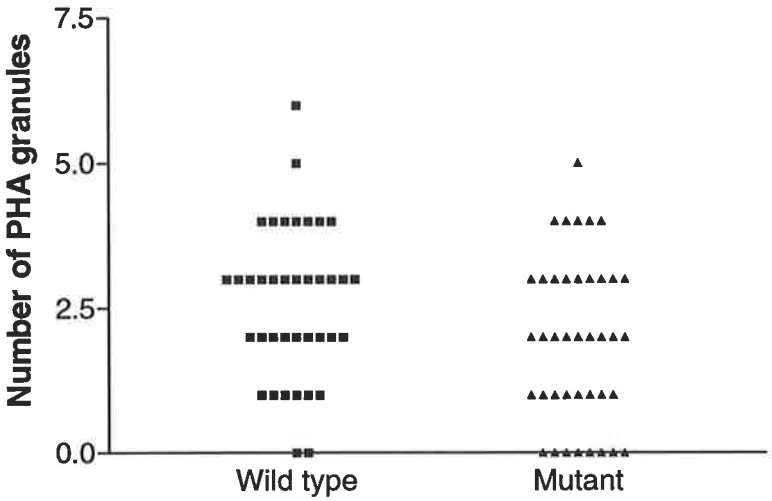


Figure 6.9. Distribution of cells of *Pseudomonas* strain 10c-1-3 wild type and the ORF4 mutant which contain different numbers of PHA intracellular granules.

The mean number of granules per wild type cell is 2.61 ± 1.326 and 1.95 ± 1.413 for the ORF4 mutant. A two tailed Student t test analysis showed the two means are significantly different at the 95% confidence level.

granules in wild type cells were arranged in clusters with distinct borders. However the PHAs in the mutant strain were distributed in cells with less distinct borders. Moreover, it was found the cytoplasm of wild type cell sections contained densely stained material characteristic of ribosomes, but the sections of cells prepared from the mutant strain contained only a lightly stained cytoplasmic region. To investigate the effect of ORF4 mutation on the quantity of PHA granules, the distribution of PHA granules per cell was tested. Wild type cells contained a significantly larger number of discrete PHA granules than the mutant. Together, this data suggests that ORF4 is likely to be associated with PHA synthesis and accumulation. In particular, an ORF4 encoded peptide is likely to play a role in PHA granule formation and may well be associated with the protein capsule of PHA granules. It is known that a large group of proteins are associated with PHA granules in other bacterial cells known to produce PHA (see Section 1.4.5). Investigation of these proteins mainly addressed patterns and characterisation (Anderson and Dawes, 1990; Wieczorek *et al.* 1996) but less on their encoding genes (McCool and Cannon, 1999).

Although the work presented here suggests a role for ORF4 in determining granule morphology, this work is preliminary. In view of the fact that the inferred ORF4 peptide shows high similarity to phaF from *R. eutropha*, it is also reasonable to suggest that ORF4 may encode an alternative PHA related biosynthetic enzyme. Under growth conditions that favour synthesis of co-polymer, it is possible that the PHA granules formed may have a morphology different from that produced from homopolymer. The growth conditions used in this work favour formation of homopolymer and no attempt was made to include substrates such as propionate which might allow formation of P(3HB-co-3HV). However, it would be of interest to compare the composition of polymers produced by *E. coli* clones containing *phaCAB* and *phaCAB* plus ORF4 when grown in a two step fermentation with glucose and propionate. If ORF4 encodes an alternative biosynthetic enzyme then the presence of ORF4 would correlate with a polymer enriched in hydroxyvalerate.

The homology between the ORF4 peptide and PhaF suggests that *Pseudomonas* 10c-1-3 encodes more than one set of core PHA biosynthetic enzymes. Chapter 4 also described amplification of a DNA fragment with a sequence similar to that of *phaA* and other *phaA* related sequences. This indicates that at least two distinct β-ketothiolase genes may be present in *Pseudomonas* 10c-1-3 as is the case for *R. eutropha* (Slater *et al.*, 1998).

Given a potential role for ORF4 in biosynthesis and assembly of PHA granules it is suggested that ORF4 be renamed as *phaD*.

Chapter 7

Summary, conclusions and general discussion

7.1 Introduction

Biodegradable bacterial polymers (PHA) are a proposed alternative to currently used petroleum-based plastics. There is an established need to solve environmental waste problems caused in part by the amount and non-degradability of the latter product. However, the key problem in PHA production by bacteria is the high cost, which can not presently compete with that of present petroleum-based plastics. This cost differential can only be reversed by research to develop bacteria which can produce a variety of PHAs from a range of cheaply available substrates.

PHA production was initially carried out using naturally occurring bacterial isolates, such as *R. eutropha*. Unfortunately, the growth rates of these bacteria are low (cf typical *E. coli* recombinant hosts) as are yields of PHA. In addition, there are difficulties in purifying the PHA from bacterial cells after production. Despite these problems, these bacteria have the advantage that PHA copolymers and heteropolymers can be readily produced which possess superior physical properties compared to the PHA homopolymer P(3HB) that is primarily produced by recombinant strains of *E. coli* expressing PHA biosynthetic genes.

E. coli is often the preferred host for commercial production of genetically engineered products. There are numerous reasons to support this choice including, extensive knowledge regarding the organism, rapid generation of biomass, the availability of low-cost culture media, and the large range of expression vectors available. To increase yields of PHA and decrease the cost of production, it has been widely accepted that there is a need to employ this recombinant host. *E. coli* strains harbouring PHA biosynthetic genes can accumulate PHA to a higher level than the natural PHA producing bacteria, from which the *pha* genes are isolated. The current recombinant PHA production *E. coli* strains mainly produce P(3HB), and a PHA copolymer P(3HB-co-3HV). However, the introduction of PHA biosynthetic genes into *E. coli* results in the establishment of a new biosynthetic pathway which competes with existing pathways and thus for existing metabolites within the cell (Lee *et al.*, 1994b). Furthermore, PHA synthesis enzymes do not have wide substrate specificities and therefore utilise a variety of different PHA precursors for PHA synthesis. For these reasons, it is not surprising that strains of *E. coli*

accumulate P(3HB) and co-polymer to different levels and at different rates. For example, competition for propionyl-CoA by PHA biosynthesis and fatty acid degradation means that little propionyl-CoA is incorporated into PHA. This problem can be partially overcome by introduction of mutations in the *fadR* and *atoC* genes that result in increased pools of propionyl-CoA available for PHA biosynthesis (eg. *E. coli* MD9101) (Slater *et al.*, 1992) in conjunction with optimisation of its growth media components and the use of two step growth methods. A broader range of co-polymers may also be expressed in *E. coli* if different PHA biosynthetic genes with wider substrate specificities are employed. A search for novel *pha* genes is therefore important for the future use and development of PHAs. The work described in this thesis therefore aimed to: (a) isolate novel PHA producing bacteria; (b) clone and describe the PHA synthetic genes from a selected isolate (*Pseudomonas* strain 10c-1-3) and (c) attempt to demonstrate production of PHA by introduction of the biosynthetic genes into *E. coli*.

7.2 Isolation of PHA producing bacteria

PHA is naturally synthesised and accumulated by many types of bacterial cells as a mechanism of carbon storage. Adelaide soil samples with high or low organic matter content were used to isolate 19 PHA producing bacterial isolates. PHA production by these isolates was confirmed by examination of cells for the presence of intracellular PHA granules and by gas chromatography. GC analysis showed that one isolate in particular, *Pseudomonas* 10c-1-3, produced PHA copolymer or heteropolymer at a high level of about 60% (w/w) of dry cell mass (see Table 3.3). On the basis of these properties, a molecular analysis of the PHA biosynthetic genes was undertaken.

In order to detect the presence of PHA biosynthesis genes in this isolate, a PCR method was devised to detect the presence of genes with similarity to known *phaA* and *phaB*. Multiple alignment of known DNA sequences showed these genes share remarkable similarity at the DNA level and this fact was used to design degenerate oligonucleotide primers (see Section 4.2.1.1) that could be used to amplify parts of these genes. One primer pair allowed amplification of an internal fragment of *phaA* cloned from *R. eutropha* and was subsequently used to amplify a similar fragment, $phaA_{int}$, from *Pseudomonas* strain 10c-1-3 chromosomal DNA. The use of PCR to screen isolates for *pha* genes is not entirely novel. Lopez *et al.* (1997) for example, used a primer pair (see Section 1.5.1),

designed primers from the conserved regions of *R. eutropha phaC* gene for the isolation of PHA producing bacteria. However, given that *phaC* genes from different bacteria share little homology, this approach is not likely to be as robust as PCR methods which amplify regions of *phaA* or *phaB*. The approach used in this thesis therefore should be applicable to detection of a wide range of bacterial types that encode genes involved in PHA biosynthesis.

A key issue in this work was isolation of bacteria encoding novel *pha* genes which allow the bacteria to capitalise on carbon sources present in their natural habitats and convert these to novel PHAs. A superior habitat selection approach was described by Hong *et al.* (1997), in which PHA producing bacteria were isolated from oil-contaminated soils, where there were a variety of different PHA precursors suitable for synthesis of PHA copolymers and heteropolymers. This study showed that it was possible to isolate bacteria from oil-contaminated soils and demonstrate presence of novel *pha* genes, especially *phaC* genes. This approach guided a novel direction for the isolation of novel PHA producing bacteria and novel *pha* genes. For this reason, a variety of soil types and organic matter content were selected for sampling.

Although the work described in this thesis concentrated on *Pseudomonas* strain 10c-1-3, the other PHA producing isolates should not be ignored. Even though GC analysis indicated they produced lower yields of PHA than 10c-1-3 and did not apparently produce PHA copolymers or heteropolymers from glucose, they may possess novel *pha* genes, which encode PHA biosynthetic enzymes with different properties from that of the known *pha* genes. Thus future work should examine these bacteria as sources of novel *pha* genes. In particular, these bacteria should be examined for *phaC* variants with novel substrate specificities. Justification for this approach is based on the fact that PHA synthase is probably the most important enzyme in the PHA biosynthetic pathway because of the potential to condense different substrates and consequently form a variety of PHA copolymers and heteropolymers. Since *phaC* encoded polymerases share the least conservation of all the known PHA biosynthesis genes, this task may be complicated by difficulties in identifying oligonucleotide primer pairs with low levels of degeneracy sufficient to ensure adequate specificity.

7.3 Cloning of the *pha* genes

In this work, a cosmid library of *Pseudomonas* strain 10c-1-3 DNA was constructed and screened for presence of *phaA* using the PCR approach mentioned previously. Cosmids (17) containing *phaA* related DNA were isolated from 4,464 clones. This was achieved through the use of strategic pooling of cosmids based on a technique originally described by Gussow and Clackson (1989). Compared to the other methods (discussed in Section 1.5.1), this screening method has the following advantages:

(a) It is very reliable and has a high stringency inherent in the PCR technique. The stringency depends on the thermocycling conditions used, eg. the annealing temperature for primers.

(b) It is very rapid and therefore time saving especially when combined with a rapid method for preparing DNA template for PCR.

(c) It is conservative in the requirement for consumables. For example, the use of a microplate for subpooling positive pools enables screening 96 clones with only 20 PCR reactions instead of 96 individual reactions.

As a result only a few hundred individual PCR reactions were required to screen the entire cosmid library. Consequently the entire cosmid bank was screened within 2-3 days.

Interestingly, none of the *E. coli* strains harbouring these cosmids synthesised and accumulated PHA. Reasons for this observation are not obvious given that later work which used subclones encoding *phaCAB* were shown to accumulate PHA as intracellular granules. Plasmid DNA from one cosmid clone, (50)D6, designated pCT400, was selected for further analysis. DNA sequence analysis showed pCT400 contained sufficient DNA to encode homologues of the *phaCAB* operon from *R. eutropha*. Sequence analysis showed that the *pha* genes from *Pseudomonas* strain 10c-1-3 are clustered as an operon in the order *phaCAB* and have the potential to encode for polypeptides of 729 (polymerase/synthase), 393 (β-ketothiolase) and 246 (NADH dependent acetoacetyl CoA reductase) amino acids respectively. An additional open reading frame, ORF4, which is apparently transcriptionally linked to these genes and with potential to encode for a polypeptide of 190 amino acids, was also identified downstream of the *phaB* gene. The peptide encoded by

this ORF is similar to PhaF from *R. eutropha* and other undescribed peptides thought to be associated with PHA biosynthesis/assembly. The putative *phaCAB* genes show significant similarity to the *phaC*, *phaA* and *phaB* genes, and protein products, described for *Burkholderia* sp., *Alcaligenes latus*, and *Alcaligenes* sp. SH-69. Furthermore, these open reading frames are of similar size to the *pha* genes found in *Alcaligenes* sp. SH-69. Nevertheless, although *phaAB* show remarkable conservation with other examples of these biosynthetic genes, significant differences in the size and sequence of *phaC* and its peptide from other *phaCs* were noted. The sequence conservation of *phaAB* is probably a reflection of the core function of the encoded enzymes, whereas the differences noted for *phaC* could indicate differences in substrate specificity of the synthase. As shown in other studies, *E. coli* DH5α cells harbouring these genes on plasmid pCT411 were able to express PHA.

Multiple alignment of the DNA sequences with equivalent sequence data from different bacterial sources confirmed that *phaA*s and *phaB*s are more conserved than *phaC*s. On the basis of sequence data alone, *phaC* can be regarded as novel and is predicted to encode a polymerase/synthase with different substrate specificity from that described by other workers.

Recently, multiple β-ketothiolase encoding genes have been located in *R. eutropha* strains (Slater *et al.*, 1998). In *R. eutropha*, the *phaA* gene encoded β-ketothiolase condenses two acetyl-CoA molecules to acetoacetyl-CoA but does not significantly catalyse the condensation reaction of acetyl-CoA and propionyl-CoA to β-ketovaleryl-CoA for the production of copolymer P(3HB-co-3HV). Slater *et al.* (1998) found the latter condensation reaction is catalysed by another β-ketothiolase, which is encoded by a *bktB* gene. This finding implies that *Pseudomonas* strain 10c-1-3 may possess multiple β-ketothiolase genes. Indirect experimental evidence to support this hypothesis is based on the finding that $phaA_{int}$ DNA originally amplified and cloned from *Pseudomonas* 10c-1-3 DNA (see Section 4.2.1) shares homology to a number of *phaA* related sequences, but has a significantly different sequence from that determined for *phaA* located on pCT400 (see Section 4.2.2.3). In fact the equivalent segments of DNA sequence share only 71% similarity (see Figure 4.11a). Thus *Pseudomonas* strain 10c-1-3 may encode at least two different β-ketothiolase genes.

Fukui *et al.* (1998) cloned a *phaJ* gene from *Aeromonas cavia*, which encodes enoyl-CoA hydratase. Enoyl-CoA hydratase participates in PHA biosynthesis from alkanoic acids in a different PHA biosynthetic pathway from that found *in R. eutropha* (see Section 1.5.2). Whether a *phaJ* gene is present in *Pseudomonas* strain 10c-1-3 is unknown. Other genes involved in formation of PHA polymer as granules may also be present in this isolate. However, given that *E. coli* cells carrying only *phaCAB* from *Pseudomonas* are able to form discrete PHA granules, this hypothesis seems unlikely. Gene(s) involved in PHA biodegradation however, should be encoded on the *Pseudomonas* 10c-1-3 genome and these remain to be identified. With appropriate screening methods, these unidentified genes should be able to be isolated from the existing cosmid bank.

The function of ORF4 and relationship to the PhaCAB gene products cannot be determined simply by peptide sequence comparisons. As mentioned previously, there are no characterised homologues to allow prediction of function. In this thesis, an insertion mutation within ORF4 was constructed as a first step to characterisation of the function of this open reading frame. Introduction of this mutation to *Pseudomonas* 10c-1-3 using allelic replacement technology, allowed direct comparison of PHA synthesis and accumulation of PHA granules by the wild type strain and the mutant. Interestingly, although no difference in the GC profile of PHA extracted from these strains was observed, the number of granules per cell and the morphology of PHA granules produced by the mutant was different from those produced by the wild type. In particular, the mean number of granules per cell was lower and the border of the granules was less distinct. Consequently, this preliminary evidence suggests that ORF4 may encode a polypeptide that in some way affects the formation and morphology of PHA granules.

When the mutant was grown in the presence of glucose, the PHA produced was indistinguishable from that produced by the wild type *Pseudomonas*. However, growth under conditions that favour formation of co-polymer was not tested. If ORF4 encodes an alternative PHA biosynthetic enzyme, growth of an *E. coli* strain encoding *phaCAB* plus ORF4 would in the presence of propionate, for example, may result in formation of copolymer, whereas the same strain lacking ORF4 would not. Similarly, the *Pseudomonas* ORF4 mutant would be expected to be unable to synthesise co-polymer under permissible growth conditions.

7.4 Construction of PHA production plasmids

Although the *pha* genes encoded on plasmid pCT400, originally isolated from cosmid clone 50(D2) were not apparently expressed in *E. coli*, subclones encoding *phaCAB* with or without ORF4, could direct biosynthesis and formation of PHA granules in *E. coli* background. Apart from demonstrating the *phaCAB* genes can direct PHA synthesis, these subclones allowed a demonstration that ORF4 was not required for PHA biosynthesis. Furthermore, production of PHA by these subclones occurred irrespective of the orientation of the *pha* genes in the multiple cloning site within *lacZ* of pBluescript or the presence of IPTG. This data suggests that transcription of the *pha* genes may occur from a putative promoter just upstream of *phaC*. However, the data can not preclude the possibility that transcription was initiated from a promoter sequence located on the vector plasmid. For this reason, primer extension analysis should be employed to resolve this issue. Certainly the putative promoter shares homology with the *E. coli* σ^{70} promoter consensus sequence and therefore may facilitate interaction with RNA polymerase via this sigma factor. In addition, *pha* genes from other bacteria have been successfully expressed from a native promoter in an *E. coli* background. Also a variety of *Pseudomonas* genes eg. genes involved in the RP4 mediated conjugal transfer system employed in this study to transfer the ORF4 mutation from *E. coli* to *Pseudomonas* 10c-1-3, can be normally expressed in *E. coli*.

The expression of the *pha* genes in *E. coli* was also examined at transcriptional level with Northern blot and at translational level with SDS-PAGE (see Sections 5.2.6 and 6.2.1.1). Unfortunately this work was unable to demonstrate *phaCAB* mRNA synthesis or over-expression of any of the enzymes involved in biosynthesis. The reason for this is unknown. However, one possible reason is that the genes are expressed only at low levels, especially once PHA granule accumulation becomes significant. The space required to maintain a number of granules may well preclude significant formation of mRNA. Certainly the cells used to isolate mRNA were grown under conditions that normally allowed accumulation of PHA granules.

The effect of host strain and plasmid stability on yield of PHA was examined. The *E. coli* strain Topp1 provided the best host background for expression of PHA using production plasmids described in this thesis. In addition, the presence of the *parB* plasmid

partitioning system was shown to significantly improve plasmid maintenance in host cells. This work supports previous studies by Slater *et al.* (1992) and Lee *et al.* (1994a,b,c). Nevertheless, PHA yields by all constructs used was significantly less than that demonstrated by other workers. Yields in excess of 90% on a cell dry weight basis have been recorded. In this study the maximum yield obtained by *E coli* strains harbouring *pha* genes from *Pseudomonas* 10c-1-3 was 12.7% [(w/w), PHA/dry cell weight]. Whether the yield of PHA can be improved by expressing these genes under control of an efficient *E. coli* promoter has yet to be demonstrated. Furthermore, the impact of the polymerase substrate specificity on yield also needs examination.

7.5 Future work

Pseudomonas strain 10c1-3 produces PHA copolymers or heteropolymers at a high level from glucose on a small scale without optimisation of growth conditions. Furthermore, cells of this organism are easily lysed. Consequently, this strain should be a good PHA producer for large scale production in terms of PHA accumulation and purification. However, further experiments are needed to improve PHA production by optimisation of the growth conditions, including optimising the incubation temperature, incubation time, media and the supplemented carbon sources. The PHA production by *Pseudomonas* strain 10c-1-3 was not optimised in this thesis as the main aim was to isolate novel *pha* genes.

GC analysis showed that *Pseudomonas* strain 10c-1-3 produced PHA copolymers or heteropolymer other than the homopolymer P(3HB) from glucose. It is possible that this isolate can produce other PHAs from other carbon sources. For a better understanding of the PHAs produced by this isolate, further information should be obtained by analysing the molecular structures of PHAs with MS and NMR.

One property of PHA synthase is its substrate specificity, eg. the recognition to HA_{SCL}, HA_{MCL} or HA_{LCL}, which is closely related to the physical properties of PHA. This property of the PHA synthase encoded by the *phaC* from *Pseudomonas* strain 10c-1-3 was not characterised in this thesis and further investigation with the addition of substrates in different chain lengths into media for PHA production is necessary for the characterisation.

The function of the ORF4 (*phaD*) from *Pseudomonas* strain 10c-1-3 is not quite clear. Several questions remain to be answered. First and foremost it would be important to demonstrate that ORF4 encoded a peptide. This could be achieved by cloning the open reading frame into a His-tag protein expression vector, for example. Any peptide expressed could be easily purified, characterised and used to prepare a specific antiserum. This antiserum could be used as an immuno-histochemical tool to demonstrate the physical location of the ORF4 peptide in PHA producing cells. Such a tool could help determine the spatial relationship between the peptide encoded by ORF4 and PHA granules located within the cytoplasm of producer cells. Secondly it is important to test the hypothesis that ORF4 encodes an alternative protein involved in PHA biosynthesis/assembly. This could be tested by compositional analysis of PHA produced by *Pseudomonas* 10c-1-3 and its isogenic ORF4 mutant, when grown under conditions which favour for example, incorporation of alternative substrates such as propionate, into PHA.

An additional problem that requires attention relates to reasons behind the inherently low yields of PHA production by *E. coli* that harboured the PHA production plasmids constructed in this study. As a starting point, the *pha* operon from *Pseudomonas* strain 10c-1-3 could be subcloned into an expression vector which preferably allows transcription from an inducible promoter. Secondly, the effect of host strain needs careful analysis since the host needs to be able to supply substrates for efficient conversion to PHA.

7.6 Perspective in PHA production plasmid construction

So far, among the PHA production plasmids, which were constructed by different research groups as discussed in Section 1.7.3.2, the *pha* genes in each plasmid were from a single source, eg. the *pha* genes in p4A, pJM9123 and pSYL plasmid series are from *R. eutropha*. After so many years of investigation on *pha* genes, many *pha* genes have been isolated from different bacteria and different PHA biosynthetic pathways with interesting properties, eg. the *phaC*s are classified into $phaC_{SCL}$ $phaC_{MCL}$ and $phaC_{LCL}$ due to the sizes of their preferred substrates. These achievements make it possible to construct artificial *pha* operons with the genes from different bacteria and different PHA biosynthetic pathways according to the needs to the PHA products with required physical properties. *phaC* is the most important gene in the PHA biosynthesis pathway and it will play a very

important role in this artificial operon construction. Fukui *et al.* (1999) constructed a plasmid which contained only the *phaC* and *phaJ* from *Aeromonas cavia* and PHA production was identified in *E. coli* harbouring this plasmid. *phaJ* encoded enzyme participates in PHA biosynthesis in a different pathway from that in *R. eutropha* and plays a very important role in the synthesis of substrates for the *phaC* encoded PHA synthase to produce PHA copolymers or heteropolymers. Therefore *phaJ* can be introduced into the artificial *pha* operons for the production of different PHAs at a high level. Another element in plasmid construction is an appropriate inducible promoter. Cheap agricultural by-products can be used for industrial scale PHA production to decrease production cost. However, different agricultural by-products contain different components, some of which may be used for the selection of an appreciate promoter. On this point of view, different regions or countries may construct different PHA production plasmids in accordance with their regional agricultural by-products. The strategy for the construction of a PHA production plasmid containing an artificial *pha* operon is illustrated in Figure 7.1 and the schematic of the artificial *pha* operon is illustrated in Figure 7.2.

PHA production plasmids optimised for PHA yield from specific low cost growth media, low cost of production in addition to the physical properties of PHA should allow development of ideal genetically engineered PHA production plasmids. However, the role of the recombinant host cannot be underestimated. The fact remains that *E. coli* does not tolerate a wide variety of potential growth substrates that encourage biosynthesis of heteropolymers and co-polymers. Significant work will need to be done to develop *E. coli* hosts that provide the necessary flow of biosynthetic substrates. To date only strain MD9101 has met even some of this fairly stringent requirement. It may be that alternative hosts will need to be developed. This obvious requirement is reflected in attempts to produce transgenic plants capable of expressing PHA.

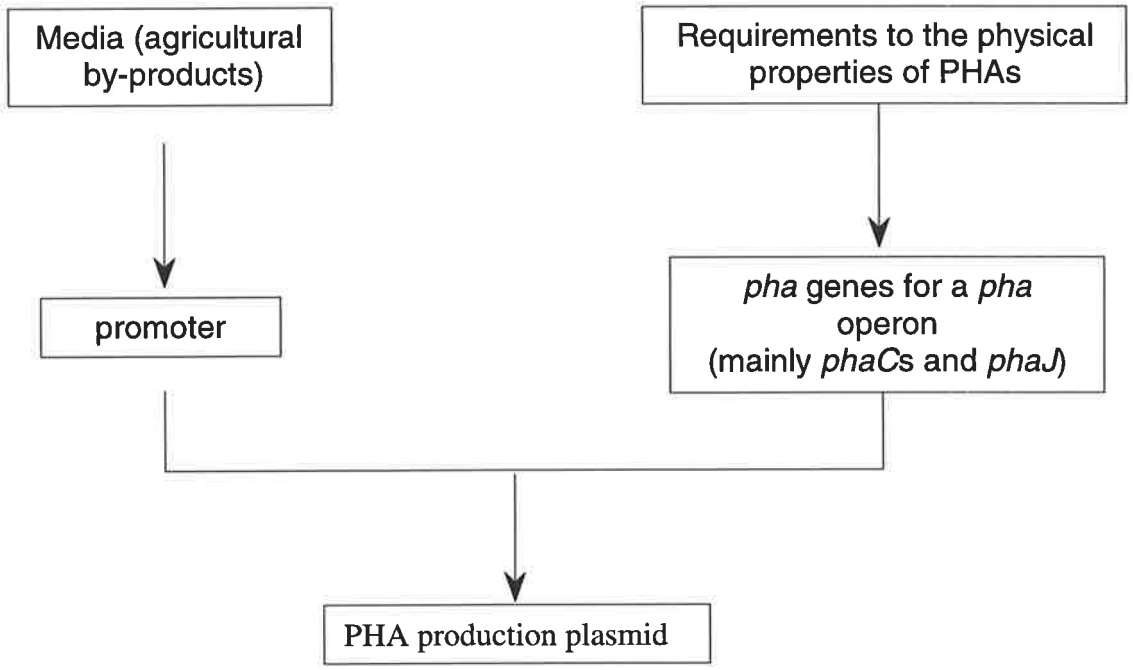


Figure 7.1. Strategy for the construction of a PHA production plasmid.

Figure 7.2. Schematic of the designed artificial *pha* operons.

Bibliography

Achtman, M.; Schwuchow, S.; Helmuth, R.; Morilli, G. and Manning, P. A. (1978). The cell-cell interactions in conjugating *Escherichia coli*: con-mutants and stabilisation mating aggregates. *Molecular and General Genetics*, **164**, 171-183.

Aiba, H.; Adhya, S. and de Crombrugghe, B. (1981). Evidence for two functional gal promoters in intact *Escherichia coli* cells. *Journal of Biological Chemistry*, **256**(22), 11905-11910.

Alderete, J. E.; Karl, D. W. and Park, C. H. (1993). Production of poly(hydroxybutyrate) homopolymer and copolymer from ethanol and propanol in a fed-batch culture. *Biotechnology Progress*, **9**(5), 520-525.

Alexeyev, M. F. (1995). Three kanamycin resistance gene cassettes with different polylinkers. *Biotechniques*, **18**(1), 52, 54, 56.

Alper, R. and Lundgren, D. G. (1963). Properties of poly-β-hydroxybutyrate. I. General considerations concerning the naturally occurring polymer. *Biopolymer*, **1**, 545-556.

Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, **215**(3), 403-410.

Anderson, A. J. and Dawes, E. A. (1990). Occurrence, metabolism, metabolic role and industrial uses of bacterial polyhydroxyalkanoates. *Microbiological Reviews*, **54**, 450-472.

Anderson, A. J.; Haywood, G. W. and Dawes, E. A. (1990). Biosynthesis and composition of bacterial poly(hydroxyalkanoates). *International Journal of Biological Macromolecules*, **12**, 102-105.

Asenjo, J. A.; Schmidt, A. S.; Andersen, P. R. and Andrews, B. A. (1995). Effect of single nutrient limitation on poly-β-hydroxybutyrate molecular weight distribution in *Alcaligenes eutrophus*. *Biotechnology and Bioengineering*, **46**, 497-502.

Athanasiou, K. A.; Niederauer, G. G.; Agrawal, C. M. and Landsman, A. S. (1995). Applications of biodegradable lactides and glycolides in podiatry. *Clinics in Podiatric Medicine and Surgery*, **12**(3), 475-495.

Atlas, R. M. (1993). *Handbook of Microbiological Media* (Ed. L. C. Parks). CRC press, 718-719.

Ballard, D. G. H.; Holmes, P. A. and Senior, P. J. (1987). Formation of polymers of β-hydroxybutyric acid in bacterial cells and a comparison of the morphology of growth with the formation of polyethylene in the solid state. In: *Recent Advances in Mechanistic and Synthetic Aspects of Polymerisation* (Eds. M. Fontanille and A. Guyot). Reidel (Kluwer) Publishing Co., Lancaster, U. K., **215**, 293-314.

Barham, P. J. (1984). Nucleation behaviour of poly-3-hydroxybutyrate. *Journal of Materials Science*, **19**, 3826-3834.

Barham, P. J.; Keller, A.; Otun, E. L. and Holmes, P. A. (1984). Crystallisation and morphology of a bacterial thermoplastic: poly-3-hydroxybutyrate. *Journal of Materials Science*, **19**, 2781-2794.

Biedermann, J.; Owen, A. J.; Schloe, K. T.; Gassner, F. and Süβmuth, R. (1997). Interaction between poly-3-hydroxybutyrate-co-3-hydroxyvalerate and a denitrifying *Pseudomonas* strain. *Canadian Journal of Microbiology*, **43**(6), 561-568.

Black, S. N.; Dobbs, B.; Dempsey, P. S. and Davey, R. J. (1990). Crystal chemistry and nucleating agents: saccharin and poly-3-hydroxybutyrate. *Journal of Materials Science Letters*, **9**, 51-52.

Braunegg, G.; Sonnleitner, B. and Lafferty, R. M. (1978). A rapid gas chromatographic method for the determination of poly-β-hydroxybutyric acid in microbial biomass. *European Journal of Applied Microbiology and Biotechnology*, **6**, 29-37.

Bucher, P. and Bairoch, A. (1994). A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. In: *Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology* (Eds. Altman, R.; Brutlag, D.; Karp, P.; Lathrop, R. and Searls, D.). ISMB-94, AAAIPress, Menlo Park, 53-61.

Budwill, K.; Fedorak, P. M. and Page, W. J. (1992). Methanogenic degradation of poly(3-hydroxyalkanoates). *Applied and Environmental Microbiology*, **58**, 1398-1401.

Byrom, D. (1990). Industrial production of copolymer from *Alcaligenes eutrophus*. In: *Novel Biodegradable Microbial Polymers: Proceedings of the NATO Advanced Research Workshop on New Biosynthetic, Biodegradable Polymers of Industrial Interest from Microorganisms* (Ed. E. A. Dawes). Kluwer Academic Publishers, 113-117.

Byrom, D. (1992). Production of poly-β-hydroxybutyrate: poly-β-hydroxyvalerate copolymers. *FEMS Microbiology Reviews*, **103**, 247-250.

Cesaro, A. and Scandola, M. (1989). Microbial polyesters. Scientific curiosity or industrial reality? *Chimicaoggi*, **7**, 81-85.

Cevallos, M. A.; Encarnacion, S.; Leija, A.; Mora, Y. and Mora, J. (1996). Genetic and physiological characterization of a *Rhizobium etli* mutant strain unable to synthesise poly-beta-hydroxybutyrate. *Journal of Bacteriology*, **178** (6), 1646-1654.

Chang, H. N. (1994). Biodegradable plastics and biotechnology. In: *Better living through innovative biochemical engineering: Proceedings of the third Asia-Pacific Biochemical Engineering Conference* (Eds. W. K. Teo; M.G. S. Yap and S. K. W. Oh). Singapore, 24-30.

Cho, K. S.; Ryu, H. W.; Park, C. H. and Goodrich, P. R. (1997). Poly(hydroxybutyrate-co-hydroxy-valerate). from swine waste liquor by *Azotobacter vinelandii* UWD. *Biotechnology Letters*, **19**(1), 7-10.

Choi, J.; Lee, S. Y. and Han, K. (1998). Cloning of the *Alcaligenes latus* polyhydroxyalkanoate biosynthesis genes and use of these genes for enhanced production of poly(3-hydroxybutyrate) in *Escherichia coli*. *Applied and Environmental Microbiology*, **64**(12), 4897-4903.

Choi, J.; Wang, F. and Lee, S. Y. (1997). Production of poly(3-hydroxybutyrate) by filamentation-suppressed recombinant *Escherichia coli* in defined medium. In: *Biochemical Engineering: Marching Toward the Century of Biotechnology: Proceedings of Asia-Pacific Biochemical Engineering Conference* (Eds. Z. Y. Shen; F. Ouyang; J. T. Yu and Z. A. Cao). Tsinghua University Press, Beijing, China, **1**, 630-633.

Cormack, R. S. and Somssich, I. E. (1997). Ultra-fast alkaline lysis plasmid extraction (UFX). *Elsevier Trens Journals Technical Tips Online*, No. 01210, at http://tto.trends.com/cgi-bin/tto/.

Cromwick, A. M.; Foglia, T. and Lenz, R. W. (1996). The microbial production of poly(hydroxyalkanoates) from tallow. *Applied Microbiology and Biotechnology*, **46**(5-6), 464-469.

Daniell, H. and Guda, C. (1997). Biopolymer production in microorganisms and plants. *Chemistry and Industry*, **14**, 555-558.

De Koning, G. J. G. and Lemstra, P. J. (1992). The amorphous state of bacterial poly[(*R*)-3-hydroxyalkanoate] *in vivo*. *Polymer Communications*, **33**, 3292-3294.

de Waard, P.; van der Wal, H.; Huijberts, G. N. and Eggink, G. (1993). Heteronuclear NMR analysis of unsaturated fatty acids in poly(3-hydroxyalkanoates): study of beta-oxidation in *Pseudomonas putida*. *Journal of Biological Chemistry*, **268**(1), 315-319.

DiRusso, C. C. (1988). Nucleotide sequence of the *fadR* gene, a multifunctional regulator of fatty acid metabolism in *Escherichia coli*. *Nucleic Acids Research*, **16**, 7995-8009.

Doi, Y.; Kunioka, M.; Nakamura, Y. and Soga, K. (1986). ^{1}H and ^{13}C NMR analysis of poly(β-hydroxybutyrate) isolated from *Bacillus megaterium*. *Macromolecules*, **19**, 1274-1276.

Doi, Y.; Segawa, A.; Nakamura, S. and Kunioka, M. (1990). Production of biodegradable copolyesters by *Alcaligenes eutrophus*. In: *Novel Biodegradable Microbial Polymers* (Ed. E. A. Dawes). Kluwer Academic Publishers, 37-47.

Donnenberg, M. S. and Kaper, J. B. (1991). Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infection and Immunity*, **59**(12), 4310-4317.

Eggink, G.; de Waard, P. and Huijberts, G. N. M. (1992). The role of fatty acid biosynthesis and degradation in the supply of substrates for poly(3-hydroxyalkanoate) formation in *Pseudomonas putida*. *FEMS Microbiology Reviews*, **103**, 159-164.

Ellar, D.; Lundgren, D. G.; Okamura, K. and Marchessault, R. H. (1968). Morphology of poly-beta-hydroxybutyrate granules. *Journal of Molecular Biology*, **35**(3), 489-502.

Evans, J. D. and Sikdar, S. K. (1990). Biodegradable plastics: an idea whose time has come? *Chemical Technology*, **20**, 38-42.

Fidler, S. and Dennis, D. (1992). Polyhydroxyalkanoate production in recombinant *Escherichia coli*. *FEMS Microbiological Letters*, **128**, 231-236.

Foellner, C. G.; Babel, W.; Valentin, H. E. and Steinbuechel, A. (1993). Expression of polyhydroxyalkanoic-acid-biosynthesis genes in methylotrophic bacteria relying on the ribulose monophosphate pathway. *Applied Microbiology and Biotechnology*, **40**(23), 284-291.

Frischauf, A. (1987). Digestion of DNA: size fractionation. In: *Methods in Enzymology* (Eds. S. L. Berger and A. R. Kimmel). Academic Press Inc., **152**, 183-189.

Fukui, T. and Doi, Y. (1997). Cloning and analysis of the poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) biosynthesis genes of *Aeromonas caviae*. *Journal of Bacteriology*, **179**(15), 4821-4830.

Fukui, T.; Shiomi, N. and Doi, Y. (1998). Expression and characterisation of (*R*)-specific enoyl coenzyme A hydratase involved in polyhydroxyalkanoate biosynthesis *by Aeromonas caviae*. *Journal of Bacteriology*, **180**(3), 667-673.

Fukui, T.; Yokomizo; S.; Kobayashi, G. and Doi, Y. (1999). Co-expression of polyhydroxyalkanoate synthase and (*R*)-enoyl-CoA hydratase genes of *Aeromonas caviae* establishes copolyester biosynthesis pathway in *Escherichia coli*. *FEMS Microbiology Letters,* **170**, 69-75.

Fukui, T.; Yoshimoto, A.; Matsumoto, M.; Hosokawa, S.; Saito, T.; Nishikawa, H. and Tomita, K. (1976). Enzymatic synthesis of poly-beta-hydroxybutyrate in *Zoogloea ramigera*. *Archives of Microbiology*, **110**(23), 149-156.

Garger, S. J.; Griffith, O. M. and Grill, L. K. (1983). Rapid purification of plasmid DNA by a single centrifugation in a two-step cesium chloride-ethidium bromide gradient. *Biochemical and Biophysical Research Communications,* **117**(3), 835-842.

Genser, K. F.; Renner, G. and Schwab, H. (1998). Molecular cloning, sequencing and expression in *Escherichia coli* of the poly(3-hydroxyalkanoate) synthesis genes from *Alcaligenes latus* DSM1124. *Journal of Biotechnology*, **64**(2-3), 125-135.

Gerdes, K. (1988). The *parB* (*hok/sok*) locus of plasmid R1: a general purpose plasmid stabilisation system. *Bio/Technology*, **6**, 1402-1405.

Gilead, D. (1985). Plastics that self destruct. *Chemtech*, **15**, 299-301.

Griebel, R.; Smith, Z. and Merrick, J. M. (1968). Metabolism of poly-β-hydroxybutyrate. I. Purification, composition and properties of native poly-β-hydroxybutyrate granules from *Bacillus megaterium*. *Biochemistry*, **7**(10), 3676-3681.

Gussow, D. and Clackson, T. (1989). Direct clone characterisation from plaques and colonies by the polymerase chain reaction. *Nucleic Acids Research*, **17**(10), 4000.

Gustafson, C. E.; Alm, R. A. and Trust, T. J. (1993). Effect of heat denaturation of target DNA on the PCR amplification. *Gene*, **123**, 241-244.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology,* **166**(4), 557-580.

Hanggi, U. J. (1995). Requirements on bacterial polyesters as future substitute for conventional plastics for consumer goods. *FEMS Microbiology Reviews*, **16**(2-3), 213-220.

Haywood, G. W.; Anderson, A. J.; Chu, L. and Dawes, E. A. (1988a). Characterisation of two 3-ketothiolases possessing differing substrate specificities in the polyhydroxyalkanoate synthesising organism *Alcaligenes eutrophus*. *FEMS Microbiology Letters*, **52**, 91-96.

Haywood, G. W.; Anderson, A. J.; Chu, L. and Dawes, E. A. (1988b). The role of NADH- and NADPH-linked acetoacetyl-CoA reductases in the poly-3-hydroxybutyrate synthesising organism *Alcaligenes eutrophus*. *FEMS Microbiology Letters*, **52**, 259-264.

Hofmann, K.; Bucher, P.; Falquet, L. and Bairoch, A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Research*, **27**, 215-219.

Hohn, B. and Collins, J. (1980). A small cosmid for efficient cloning of large DNA fragments. *Gene*, **11**, 291-298.

Holmes, P. A. (1985). Applications of PHB—A microbially produced biodegradable thermoplastic. *Physics in Technology*, **16**, 32-36.

Holmes, D. S. and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry*, **114**(1), 193-197.

Hong, K.; Chen, G.; Tian, W.; Huang, W. and Fan, Q. (1997). Production of novel polyhydroxyalkanoates by microorganisms isolated from oil-contaminated soils and water. In: *Biochemical Engineering: Marching Toward the Century of Biotechnology: Proceedings of Asia-Pacific Biochemical Engineering Conference* (Eds. Z. Y. Shen; F. Ouyang; J. T. Yu and Z. A. Cao). Tsinghua University Press, Beijing, China, **1**, 352-356.

Hrabak, O. (1992). Industrial production of poly-β-hydroxybutyrate. *FEMS Microbiology Reviews*, **103**, 251-256.

Huisman, G. W.; Wonink, E.; Meima, R.; Kazemier, B.; Terpstra, P. and Witholt, B. (1991). Metabolism of poly(3-hydroxyalkanoates) (PHAs) by *Pseudomonas oleovorans*:

identification and sequences of genes and function of the encoded proteins in the synthesis and degradation of PHA. *Journal of Biological Chemistry*, **266**(4), 2191-2198.

Hustede, E. and Steinbuechel, A. (1993). Characterization of the polyhydroxyalkanoate synthase gene locus of *Rhodobacter shpaeroides*. *Biotechnology Letters*, **15**, 709-714.

Hustede, E.; Steinbuchel, A. and Schlegel, H. G. (1992). Cloning of poly(3-hydroxybutyric acid) synthase genes of *Rhodobacter sphaeroides and Rhodospirillum rubrum* and heterologous expression in *Alcaligenes eutrophus*. *FEMS Microbiology Letters,* **72**(3), 285-90.

Israel, D. I. (1993). A PCR-based method for high stringency screening of DNA libraries. *Nucleic Acids Research*, **21**(11), 2627-2631.

Janes, B.; Hollar, J. and Dennis, D. (1990). Molecular characterisation of the poly-β-hydroxybutyrate biosynthetic pathway of *Alcaligenes eutrophus* H-16. In: *Novel Biodegradable Microbial Polymers* (Ed. E. A. Dawes). Kluwer Academic Publishers, 175-190.

Jendrossek, D.; Schirmer, A. and Schlegel, H. G. (1996). Biodegradation of polyhydroxyalkanoic acids. *Applied Microbiology and Biotechnology*, **46**(5-6), 451-463.

Jenkins, L. S. and Nunn, W. D. (1987). Genetic and molecular characterisation of the genes involved in short-chain fatty acid degradation in *Escherichia coli*: the *ato* system. *Journal of Bacteriology*, **169**, 42-52.

John, M. (1997). Transgenic cotton plants producing heterologous polyhydroxybutyrate bioplastic. *Patent*: US 5602321-A 23 11-FEB-1997.

Kassab, A. C.; Xu, K.; Denkbas, E. B.; Dou, Y.; Zhao, S. and Piskin, E. (1997). Rifampicin carrying polyhydroxybutyrate microspheres as a potential chemoembolisation agent. *Journal of Biomaterials Science Polymer Edition,* **8**(12), 947-961.

Kawai, F. (1995). Breakdown of plastics and polymers by microorganisms. *Advances in Biochemical Engineering/Biotechnology.* **52**, 151-194.

Kidwell, J.; Valentin, H. E. and Dennis, D. (1995). Regulated expression of the *Alcaligenes eutrophus pha* biosynthesis genes in *Escherichia coli*. *Applied and Environmental Microbiology*, **61**(4), 1391-1398.

Kim, B. S.; Lee, S. Y. and Chang, H. N. (1992). Production of poly-β-hydroxybutyrate by fed-batch culture of recombinant *Escherichia coli*. *Biotechnology Letters*, **14**, 811-818.

Kim, B. S.; Lee, S. C.; Lee, S. Y.; Chang, H. N.; Chang, Y. K. and Woo, S. I. (1994a). Production of poly(3-hydroxybutyric acid) by fed-batch culture of *Alcaligenes eutrophus* with glucose concentration control. *Biotechnology and Bioengineering*, **43**, 892-898.

Kim, B. S.; Lee, S. C.; Lee, S. Y.; Chang, H. N.; Chang, Y. K. and Woo, S. I. (1994b). Production of poly(3-hydroxybutyric-co-3-hydroxyvaleric acid) by fed-batch culture of

Alcaligenes eutrophus with substrate control using on-line glucose analyser. *Enzyme and Microbial Technology*, **16**, 556-561.

Kim, G. J.; Lee, I. Y.; Yoon, S. C.; Shin, Y. C. and Park, Y. H. (1997). Enhanced yield and a high production of medium-chain-length poly(3-hydroxyalkanoates) in a two-step fed-batch cultivation of *Pseudomonas putida* by combined use of glucose and octanoate. *Enzyme and Microbial Technology,* **20**(7), 500-505.

Knowles, J. C.; Hastings, G. W.; Ohta, H.; Niwa, S. and Boeree, N. (1992). Development of a degradable composite for orthopaedic use: *in vivo* biomechanical and histological evaluation of two bioactive degradable composites based on the polyhydroxybutyrate polymer. *Biomaterials*, **13**(8), 491-496.

Koosha, F.; Muller, R. H. and Davis, S. S. (1989). Polyhydroxybutyrate as a drug carrier. *Critical Reviews in Therapeutic Drug Carrier Systems,* **6**(2), 117-130.

Lafferty, R. M.; Korsatko, B. and Korsatko, W. (1988). Microbial production of poly-β-hydroxybutyric acid. In: *Biotechnology: a comprehensive treatise in 8 volumes* (Eds. H. J. Rehm and G. Reed). VCH Publishers, Weinheim, Germany, **6b**, 135-176.

Law, J. H. and Slepecky, R. A. (1961). Assay of poly-β-hydroxybutyric acid. *Journal of Bacteriology*, **82**, 33-36.

Lee, S. Y. (1994). Suppression of filamentation in recombinant *Escherichia coli* by amplified FtsZ activity. *Biotechnology Letters*, **16**, 1247-1252.

Lee, S. Y. (1995). Polyhydroxyalkanoates as biodegradable plastic material: academic and industrial effort. In: Chemeca'95, 23rd Australian Chemical Engineering Conference Proceedings (Ed. A. P. J. Middelberg). Adelaide, SA, Australia, **3**, 3-8.

Lee, S. Y. (1997). *E. coli* moves into the plastic age. *Nature Biotechnology*, **15**(1), 17-18.

Lee, S. Y. and Chang, H. N. (1994). Effect of complex nitrogen source on the synthesis and accumulation of poly(3-hydroxybutyric acid) by recombinant *Escherichia coli* in flask and fed-batch cultures. *Journal of Environmental Polymer Degradation*, **2**(3), 169-176.

Lee, S. Y. and Chang, H. N. (1995). Production of poly(hydroxyalkanoic acid). *Advances in Biochemical Engineering/Biotechnology,* **52**, 27-58.

Lee, S. Y.; Chang, H. N. and Chang, Y. K. (1994a). Production of poly(β-hydroxybutyric acid) by recombinant *Escherichia coli. Annals New York Academy of Sciences*, **721**, 43-53.

Lee, E. Y. and Choi, C. Y. (1997). Structural identification of polyhydroxyalkanoic acid (PHA) containing 4-hydroxyalkanoic acids by gas chromatography-mass spectrometry (GC-MS) and its application to bacteria screening. *Biotechnology Techniques*, **11**(3), 167-171.

Lee, S. Y.; Lee, Y. K. and Chang, H. N. (1995). Stipulatory effects of amino acids and oleic acid on poly(3-hydroxybutyric acid) synthesis by recombinant *Escherichia coli*. *Journal of Fermentation and Bioengineering*, **79**, 177-180.

Lee, S. Y.; Lee, K. M.; Chang, H. N. and Steinbuchel, A. (1994b). Comparison of Recombinant *Escherichia coli* strains for synthesis and accumulation of poly-(3-hydroxybutyric acid) and morphological changes. *Biotechnology and Bioengineering*, **44**, 1-11.

Lee, B.; Pometto, A. L. III.; Fratzke, A.; Bailey, T. B. Jr. (1991). Biodegradation of degradable plastic polyethylene by *Phanerochaete* spp. and *Streptomyces* spp. *Applied and Environmental Microbiology*, **57**, 678-685.

Lee, S. Y.; Yim, K. S.; Chang, H. N. and Chang, Y. K. (1994c). Construction of plasmids, estimation of plasmid stability and use of stable plasmids for the production of poly(3-hydroxybutyric acid) by recombinant *Escherichia coli*. *Journal of Biotechnology*, **32**, 203-211.

Li, H. H.; Gyllensten, U. B.; Cui, X. F.; Saiki, R. K.; Erlich, H. A. and Arnheim, N. (1988). Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature*, **335**(6189), 414-417.

Li, H. Q.; Sha L.; Wang, L. N. and Shen, Z. Y. (1997). Production of PHB by *A. eutrophus* with molasses as the sole carbon source. In: *Biochemical Engineering: Marching Toward the Century of Biotechnology: Proceedings of Asia-Pacific Biochemical Engineering Conference* (Eds. Z. Y. Shen; F. Ouyang; J. T. Yu and Z. A. Cao). Tsinghua University Press, Beijing, China, **1**, 348-351.

Liebergesell, M. and Steinbuchel, A. (1992). Cloning and nucleotide sequences of genes relevant for biosynthesis of poly(3-hydroxybutyric acid) in *Chromatium vinosum* strain D. *European Journal of Biochemistry*, **209**(1), 135-150.

Liebergesell, M. and Steinbuchel, A. (1993). Cloning and molecular analysis of the poly(3-hydroxybutyric acid) biosynthetic genes of *Thiocystis violacea*. *Applied Microbiology and Biotechnology*, **38**(4), 493-501.

Lillo, J. G. and Rodriguez-Valera, F. (1990). Effects of culture conditions on poly(beta-hydroxybutyric acid) production by *Haloferax mediterranei*. *Applied and Environmental Microbiology*, **56**(8), 2517-2521.

Lopez, N. I.; Pettinari, M. J. and Mendez, B. S. (1997). Detection of reserve polymer synthesis genes in natural bacterial populations. *FEMS Microbiology Ecology*, **22**, 129-136.

Lugtenberg, B.; Meijers, J.; Peters, R.; van der Hoek, P. and van Alphen, L. (1975). Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS Letters*, **58**(1), 254-258.

Lundgren, D. G.; Pfister, R. M. and Merrick, J. M. (1964). Structure of poly(β-hydroxybutyric acid) granules. *Journal of General Microbiology*, **34**, 441-446.

Lutkenhuas, J. (1993). *Escherichia coli* cell division. *Current Opinion in Genetics and Development*, **3**, 783-788.

Luzier, W. D. (1992). Materials derived from biomass/biodegradable materials. *Proceedings of the National Academy of Sciences of the USA*, **89**, 839-842.

MacFaddin, J. F. (1976). *Biochemical tests for identification of medical bacteria.* The Williams & Wilkins Company, Baltimore, MD, USA.

Magasanik, B. and Neidhardt, F. C. (1987). Regulation of carbon and nitrogen utilisation. In: *Escherichia coli and Salmonella typhimurium cellular and molecular biology* (Eds. J. L. Ingraham; K. B. Low; B. Magasanik; M. Schaechter and H. E. Umbarger). American Society for Microbiology Press, 1318-1325.

Maloy, S. R. and Nunn, W. D. (1982). Genetic regulation of the glyoxylate shunt in *Escherichia coli* K-12. *Journal of Bacteriology*, **149**, 173-180.

Manning, P. A.; Heuzenroeder, M. W.; Yeadon, J.; Leavesley, D. I.; Reeves, P. R. and Rowley, D. (1986). Molecular cloning and expression in *Escherichia coli* K-12 of the O antigens of the Inaba and Ogawa serotypes of the *Vibrio cholerae* O1 lipopolysaccharides and their potential for vaccine development. *Infection and Immunity*, **53**(2), 272-277.

Matsusaki, H.; Manji, S.; Taguchi, K.; Kato, M.; Fukui, T. and Doi,Y. (1998). Cloning and molecular analysis of the Poly(3-hydroxybutyrate) and Poly(3-hydroxybutyrate-co-3-hydroxyalkanoate) biosynthesis genes in *Pseudomonas* sp. strain 61-3. *Journal of Bacteriology*, **180**(24), 6459-6467.

McCool, G. J. and Cannon, M. C. (1999). Polyhydroxyalkanoate inclusion body-associated proteins and coding region in *Bacillus megaterium. Journal of Bacteriology*, **181,** 585-592.

McInerney, M. J.; Amos, D. A.; Kealy, K. S. and Palmer, J. (1992). Synthesis and function of polyhydroxyalkanoates in anaerobic syntrophic bacteria. *FEMS Microbiology Reviews*, **103**, 195-206.

Melbourne, C. (1995). Production of poly(hydroxyalkanoate) in recombinant *Escherichia coli.* Thesis for degree of bachelor of science with honours, The Department of Microbiology and Immunology, The University of Adelaide.

Mergaert, J.; Anderson, C.; Wouters, A.; Swings, J. and Kersters, K. (1992). Biodegradation of polyhydroxyalkanoates. *FEMS Microbiology Reviews*, **9**(2-4), 317-321.

Merrick, J. M. and Doudoroff, M. (1961). Enzymatic synthesis of poly-β-hydroxybutyric acid in bacteria. *Nature,* **189**, 890-892.

Merrick, J. M. and Doudoroff, M. (1964). Depolymerisation of poly-β-hydroxybutyrate by an intracellular enzyme system. *Journal of Bacteriology,* **88**, 60-71.

Middelberg, A. P. J.; Lee, S. Y.; Martin, J.; Williams, D. R. G. and Chang, H. N. (1995). Size analysis of poly(3-hydroxybutyric acid) granules produced in recombinant *Escherichia coli. Biotechnology Letters*, **17**(2), 205-210.

Miller, J. (1972). *Experiments in molecular genetics.* Cold Spring Harbour Laboratory, New York, USA.

Miller, V. L. and Mekalanos, J. J. (1988). A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *Journal of Bacteriology*, **170**(6), 2575-2583.

Mitomo, H.; Barham, P. J. and Keller, A. (1987). Crystallisation and morphology of poly(β-hydroxybutyrate) and its copolymer. *Polymer Journal,* **19**, 1241-1253.

Molitoris, H. P.; Moss, S. T. and De Koning, G. J. M. (1996). Scanning electron microscopy of polyhydroxyalkanoate degradation by bacteria. *Applied Microbiology and Biotechnology*, **46**, 570-579.

Muller, H. M. and Seebach, D. (1993). Poly(hydroxyalkanoates): a fifth class of physiologically important organic biopolymers? *Angewandte Chemie, International Edition in English*, **32**, 477-502.

Narayan, R. (1988). Preparation of bio-based polymers for materials applications. *Applied Biochemistry and Biotechnology,* **17**, 7-22.

Oeding, V. and Schlegel, H. B. (1973). β-ketothiolase from *Hydrogenomonas eutropha* H16 and its significance in the regulation of poly-β-hydroxybutyrate metabolism. *Biochemical Journal*, **134**, 239-248.

Ostle, A. G. and Holt, J. G. (1982). Nile Blue A as a fluorescent stain for poly-β-hydroxybutyrate. *Applied and Environmental Microbiology*, **44**(1), 238-241.

Page, W. J. and Cornish, A. (1993). Growth of *Azotobacter vinelandii* UWD in fish peptone medium and simplified extraction of poly-beta-hydroxybutyrate. *Applied and Environmental Microbiology*, **59**(12), 4236-4244.

Page, W. J.; Manchak, J. and Rudy, B. (1992). Formation of poly(hydroxybutyrate-co-hydroxyvalerate) by *Azotobacter vinelandii* UWD. *Applied and Environmental Microbiology*, **58**(9), 2866-2873.

Pearson, W. R. (1990). Rapid and Sensitive Sequence Comparison with FASTP and FASTA. *Methods in Enzymology*, **183**, 63-98.

Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence analysis. *Proceedings of the National Academy of Sciences*, **85**, 2444-2448.

Pelczar, M. J. and Chan, E. C. S. (1977). *Laboratory Exercises in Microbiology*, McGraw-Hill Publishing Company Limited, London.

Peoples, O. P.; Masamune, S.; Walsh, C. T. and Sinskey, A. J. (1987). Biosynthetic thiolase from *Zoogloea ramigera*: III. Isolation and characterization of the structural gene. *The Journal of Biological Chemistry*, **262**, 97-102.

Peoples, O. P. and Sinskey, A. J. (1989a). Fine structural analysis of the *Zoogloea ramigera phbA-phbB* locus encoding beta-ketothiolase and acetoacetyl-CoA reductase: nucleotide sequence of *phbB*. *Molecular Microbiology*, **3**(3), 349-357.

Peoples, O. P. and Sinskey, A. J. (1989b). Poly-β-hydroxybutyrate biosynthesis in *Alcaligenes eutrophus* H16: characterisation of the genes encoding β-ketothiolase and acetoacetyl-CoA reductase. *The Journal of Biological Chemistry*, **264**(26), 15293-15297.

Peoples, O. P. and Sinskey, A. J. (1989c). Poly-β-hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus* H16. Identification and characterisation of the PHB polymerase gene (*phbC*). *The Journal of Biological Chemistry*, **264**, 15298-15303.

Pieper, U. and Steinbuchel, A. (1992). Identification, cloning and sequence analysis of the poly(3-hydroxyalkanoic acid) synthase gene of the gram-positive bacterium *Rhodococcus ruber*. *FEMS Microbiology Letters*, **75**(1), 73-79.

Preusting, H.; Hazenberg, W. and Witholt, B. (1993a). Continuous production of poly(3-hydroxyalkanoates) by *Pseudomonas oleovorans* in a high-cell-density, two-liquid-phase chemostat. *Enzyme and Microbial Technology*, **15**(4), 311-316.

Preusting, H.; Kingma, J. and Witholt, B. (1991). Physiology and polyester formation of *Pseudomonas oleovorans* in continuous two-liquid-phase cultures. *Enzyme and Microbial Technology*, **13**(10), 770-780.

Preusting, H.; van Houten, R.; Hoefs, A.; van Langenberghe, E. K.; Favre Bulle, O. and Witholt, B. (1993b). High cell density cultivation of *Pseudomonas oleovorans*: growth and production of poly (3-hydroxyalkanoates) in two-liquid phase batch and fed-batch systems. *Biotechnology and Bioengineering*, **41**(5), 550-556.

Ramsay, B. A.; Lomaliza, K.; Chavarie, C.; Dube, B.; Bataille, P. and Ramsay, J. A. (1990). Production of poly-(beta-hydroxybutyric-Co-beta-hydroxyvaleric) acids. *Applied and Environmental Microbiology*, **56**(7), 2093-2098.

Ramsay, B. A.; Saracovan, I.; Ramsay, J. A. and Marchessault, R. H. (1994). A method for the isolation of microorganisms producing extracellular long-side-chain poly(beta-hydroxyalkanoate) depolymerase. *Journal of Environmental Polymer Degradation*, **2**(1), 1-7.

Rapley, R. and Walker, M. (1992). PCR screening of DNA cloned into polylinker-containing vectors with M13 sequencing primers. *BioTechniques*, **12**(4), 516.

Repaske, R. and Repaske, A. C. (1976). Quantitative requirements for exponential growth of *Alcaligenes eutrophus*. *Applied and Environmental Microbiology*, **32**(4), 585-591.

Reynolds, E. S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *Journal of Cell Biology*, **17**, 208-212.

Rhee, Y. H.; Jang, J. H. and Rogers, P. L. (1992). Biopolymer production by an *Alcaligenes* sp. for biodegradable plastics. *Australian Biotechnology*, **2**(4), 230-232.

Ritchie, G. A. F. and Dawes, E. A. (1969). The non-involvement of acyl-carrier protein in poly-β-hydroxybutyrate synthesis in *Azotobacter beijerinckii*. *Biochemical Journal*, **112**, 803-805.

Ritchie, G. A. F.; Senior, P. J. and Dawes, E. A. (1971). The purification and characterisation of acetoacetyl-coenzyme A reductase from *Azotobacter beijerinckii*. *Biochemical Journal*, **121**(2), 309-316.

Rodrigues, M. F.; Valentin, H. E.; Berger, P. A.; Tran, M.; Asrar, J.; Gruys, K. J. and Steinbuchel, A. (2000). Polyhydroxyalkanoate accumulation in *Burkholderia* sp.: a molecular approach to elucidate the genes involved in the formation of two homopolymers consisting of short-chain-length 3-hydroxyalkanoic acids. *Applied Microbiology and Biotechnology*, **53**(4), 453-460.

Ryu, H. W.; Hahn, S. K.; Chang, Y. K. and Chang, H. N. (1997). Production of poly(3-hydroxybutyrate) by high cell density fed-batch culture of *Alcaligenes eutrophus* with phosphate limitation. *Biotechnology and Bioengineering*, **55**(1), 28-32.

Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold Spring Harbour Laboratory Press, New York, USA.

Sarkar, G.; Kapelner, S. and Somner, S. (1990). Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Research*, **18**(24), 7465.

Sawadogo, M. and van Dyke, M. W. (1991). A rapid method for the purification of deprotected oligodeoxynucleotides. *Nucleic Acids Research*, **19**(3), 674.

Scandola, M.; Pizzoli, M.; Ceccorulli, G.; Cesaro, A.; Paoletti, S. and Navarini, L. (1988). Viscoelastic and thermal properties of bacterial poly(D-(-)-β-hydroxybutyrate). *International Journal of Biological Macromolecules*, **10**, 373-377.

Schembri, M. A.; Bayly, R. C. and Davies, J. K. (1994). Cloning and analysis of the polyhydroxyalkanoic acid synthase gene from an *Acinetobacter* sp.: evidence that the gene is both plasmid and chromosomally located. *FEMS Microbiology Letters*, **118** (1-2), 145-152.

Schembri, M. A.; Bayly, R. C. and Davies, J. K. (1995). Phosphate concentration regulates transcription of the *Acinetobacter* polyhydroxyalkanoic acid biosynthetic genes. *Journal of Bacteriology*, **177**(15), 4501-4507.

Schlegel, H. G. (1990). *Alcaligenes eutrophus* and its scientific and industrial career. In: *Novel Biodegradable Microbial Polymers* (Ed. E. A. Dawes). Kluwer Academic Publishers, 133-141.

Schlegel, H. G.; Gottschalk, G. and Von Bartha, R. (1961). Formation and utilisation of poly-β-hydroxybutyric acid by knallgas bacteria (*Hydrogenomonas*). *Nature*, **191**, 463-465.

Schubert, P.; Kruger, N. and Steinbuchel, A. (1991). Molecular analysis of the *Alcaligenes eutrophus* poly(3-hydroxybutyrate) biosynthetic operon: identification of the N

terminus of poly(3-hydroxybutyrate) synthase and identification of the promoter. *Journal of Bacteriology*, **173**(1),168-175.

Schubert, P.; Steinbuchel, A. and Schlegel, H. G. (1988). Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly-β-hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. *Journal of Bacteriology*, **170**, 5837-5847.

Sim, S. J.; Snell, K. D.; Hogan, S. A.; Stubbe, J.; Rha, C. and Sinskey, A. J. (1997). PHA synthase activity controls the molecular weight and polydispersity of polyhydroxybutyrate in *vivo*. *Nature Biotechnology*, **15**, 63-67.

Sirockin, G. and Cullimore, S. (1969). *Practical Microbiology*. McGraw-Hill Publishing Company Limited, London.

Slater, S.; Gallaher, T. and Dennis, D. (1992). Production of poly-(3-hydroxybutyrate-co-3-hydroxyvalerate) in a recombinant *Escherichia coli* strain. *Applied and Environmental Microbiology*, **58**, 1089-1094.

Slater, S.; Houmiel, K. L.; Tran, M.; Mitsky, T. A.; Taylor, N. B.; Padgette, S. R. and Gruys, K. J. (1998). Multiple β-ketothiolases mediate poly(β-hydroxyalkanoate) copolymer synthesis in *Ralstonia eutropha*. *Journal of Bacteriology*, **180**(8), 1979-1987.

Slater, S. C.; Voige, W. H. and Dennis, D. E. (1988). Cloning and expression in *Escherichia coli* of the *Alcaligenes eutrophus* H16 poly-β-hydroxybutyrate biosynthetic pathway. *Journal of Bacteriology*, **170**, 4431-4436.

Smock, D. (1987). Are degradable plastics the answer to litter? *Plastics World*, **45**, 28-33.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, **98**(3), 503-517.

Steinbuchel, A. (1991). Polyhydroxyalkanoates. In: *Biomaterials: Novel Materials from Biological Sources* (Ed. D. Byrom). Stockton Press, New York, USA, 124-213.

Steinbuchel, A. (1996). PHB and other polyhydroxyalkanoic acids. In: *Biotechnology: a comprehensive treatise in 8 volumes* (Eds. H. J. Rehm and G. Reed). VCH Publishers, Weinheim, Germany, 403-464.

Steinbuchel, A.; Hustede, E.; Liebergesell, M.; Pieper, U.; Timm, A. and Valentin, H. (1992). Molecular basis for biosynthesis and accumulation of polyhydroxyalkanoic acid in bacteria. *Microbiology Reviews*, **103**, 217-230.

Steinbuchel, A. and Schlegel, H. G. (1989). Excretion of pyruvate by mutants of *Alcaligenes eutrophus*, which are impaired in the accumulation of poly(beta-hydroxybutyric acid) (PHB) under conditions permitting synthesis of PHB. *Applied Microbiology and Biotechnology*, **31**(2), 168-175.

Steinbuchel, A. and Schlegel, H. G. (1991). Physiology and molecular genetics of Poly(β-hydroxyalkanoic acid) synthesis in *Alcaligenes eutrophus*. *Molecular Microbiology*, **5**(3), 535-542.

Steinbuchel, A. and Valentin, H. E. (1995). Diversity of bacterial polyhydroxyalkanoic acids. *FEMS Microbiology Letters*, **128**, 219-228.

Suzuki, T.; Yamane, T. and Shimizu, S. (1986a). Mass production of poly-beta-hydroxybutyric acid by fed-batch culture with controlled carbon/nitrogen feeding. *Applied Microbiology and Biotechnology*, **24**(5), 370-374.

Suzuki, T.; Yamane, T. and Shimizu, S. (1986b). Mass production of poly-beta-hydroxybutyric acid by fully automatic fed-batch culture of methylotroph. *Applied Microbiology and Biotechnology*, **23**(5), 322-329.

Swift, G. (1993). Directions for environmentally biodegradable polymer research. *Accounts of Chemical Research*, **26**,105-110.

Tabor, S. and Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Sciences of the USA*, **82**(4), 1074-1078.

Tezuka, C.; Takeda, K. and Someya, J. (1980). Poly-beta-hydroxybutyrate of methane oxidizing bacteria. *Report of the Fermentation Research Institute* (Yatabe), **0**(54), 5-12.

Thomas, M. R. (1994). Simple, effective cleanup of DNA ligation reactions prior to electro-transformation of *E. coli*. *BioTechniques*, **16**(6), 988-990.

Thompson, J. D.; Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**(22), 4673-4680.

Timm, A. and Steinbuchel, A. (1992). Cloning and molecular analysis of the poly(3-hydroxyalkanoic acid) gene locus of *Pseudomonas aeruginosa* PAO1. *European Journal of Biochemistry*, **209**(1), 15-30.

Timm, A.; Wiese, S. and Steinbuchel, A. (1994). A general method for identification of polyhydroxyalkanoic acid synthase genes from pseudomonads belonging to the rRNA homology group I. *Applied Microbiology and Biotechnology*, **40**(5), 669-75.

Tombolini, R.; Povolo, S.; Buson, A.; Squartini, A. and Nuti, M. P. (1995). Poly-β-hydroxybutyrate (PHB) biosynthetic genes in *Rhizobium meliloti* 41. *Microbiology*, **141**, 2553-2559.

Tomita, K.; Saito, T. and Fukui, T. (1983). Bacterial metabolism of poly-β-hydroxybutyrate. In: *Biochemistry of Metabolic Processes* (Eds. D. L. F. Lennon; F. W. Stratman and R. N. Zahlten). Elsevier Science Publishing, Inc., New York, 353-366.

Ueda, S.; Matsumoto, S.; Takagi, A. and Yamane, T. (1992). Synthesis of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) from methanol and n-amyl alcohol by the methylotrophic bacteria *Paracoccus denitrificans* and *Methylobacterium extorquens*. *Applied and Environmental Microbiology*, **58**(11), 3574-3579.

Ueda, S.; Yabutani, T.; Maehara, A. and Yamane, T. (1996). Molecular analysis of the poly(3-hydroxyalkanoate) synthase gene from a methylotrophic bacterium, *Paracoccus denitrificans*. *Journal of Bacteriology*, **178**(3), 774-779.

Valentin, H. E. and Steinbuchel, A. (1993). Cloning and characterisation of the *Methylobacterium extorquens* polyhydroxyalkanoic-acid-synthase structural gene. *Applied Microbiology and Biotechnology*, **39**(3), 309-17.

van Wegen, R. J.; Lee, S. Y.; Thomas, C. J.; Williams, D. R. G. and Middelberg, A. P. J. (1997). Using recombinant *Escherichia coli* and dairy whey to produce poly(3-hydroxybutyrate). In: *Biochemical Engineering: Marching Toward the Century of Biotechnology: Proceedings of Asia-Pacific Biochemical Engineering Conference* (Eds. Z. Y. Shen; F. Ouyang; J. T. Yu and Z. A. Cao). Tsinghua University Press, Beijing, China, **1**, 622-625.

van Wegen, R. J.; Ling, Y. and Middelberg, A. P. J. (1998). Industrial production of polyhydroxyalkanotes using *Escherichia coli*: An economic analysis. *Transactions of the Institution of Chemical Engineers*, 76(A), 417-426.

Wang, F. and Lee, S. Y. (1997a). Poly(3-hydroxybutyrate) production with high productivity and high productivity and high polymer content by a fed-batch culture of *Alcaligenes latus* under nitrogen limitation. *Applied and Environmental Microbiology*, **63**(9), 3703-3706.

Wang, F. and Lee, S. Y. (1997b). Production of poly(3-hydroxybutyrate) by fed-batch culture of filamentation-suppressed recombinant *Escherichia coli*. *Applied and Environmental Microbiology*, **63**(12), 4765-4769.

Wegner, G. and Wagemann, K. (1994). Polymers and the environment—Current problems and future research. *Advanced Materials*, **6**(9), 629-634.

Wieczorek, R.; Steinbuchel, A. and Schmidt, B. (1996). Occurrence of polyhydroxyalkanoic acid granule-associated proteins related to the *Alcaligenes eutrophus* H16 GA24 protein in other bacteria. *FEMS Microbiology Letters*, **135**(1), 23-30.

Williams, S. F.; Palmer, M. A. J.; Peoples, O. P.; Walsh, C. T.; Sinskey, A. J. and Masamune, S. (1992). Biosynthetic thiolase from *Zoogloea ramigera*: Mutagenesis of the putative active-site base CYS-378 to SER-378 changes the partitioning of the acetyl S-enzyme intermediate. *Journal of Biological Chemistry*, **267**, 16041-16043.

Yabutani, T.; Maehara, A.; Ueda, S. and Yamane, T. (1995). Analysis of beta-ketothiolase and acetoacetyl-CoA reductase genes of a methylotrophic bacterium, *Paracoccus denitrificans*, and their expression in *Escherichia coli*. *FEMS Microbiology Letters*, **133** (1-2), 85-90.

Zhang, H.; Obias, V.; Gonyer, K. and Dennis, D. (1994). Production of polyhydroxyalkanoates in sucrose-utilising recombinant *Escherichia coli* and *Klebsiella* strains. *Applied and Environmental Microbiology*, **60**(4), 1198-1205.

Appendix A Buffers and solutions

- **Cosmid genomic library construction buffers and solutions**

SM buffer (for bacteriophage storage and dilution)

Per litre:

NaCl	5.8 g
$MgSO_4 \bullet 7H_2O$	2 g
1 M Tris•Cl (pH 7.5)	50 ml
2% gelatin	5 ml

Sterilise by autoclaving and store in 50-ml lots.

Maltose

Maltose was added to induce the expression of the lambda bacteriophage acceptor *LamB*.

Maltose	20 g
Water	100 ml

Sterilise by filtration. Add 1 ml of sterile 20% maltose solution for every 100 ml of growth medium.

- **DNA/RNA preparation/manipulation buffers and solutions**

*** Lysis solution for Bacterial chromosomal DNA preparation** (100 ml)

10% sarkosyl	50 ml
1M Tris (pH 8.0)	5 ml
0.25 M EDTA (pH 8.0)	25 ml
Sterile H_2O	20 ml

*** Solutions for plasmid mini preparation by alkaline lysis**

Solution 1

Tris•Cl (pH 8.0)	25 mM
EDTA (pH 8.0)	10 mM
Glucose	50 mM

Solution I can be prepared in batches of approximately 100 ml, autoclaved for 15 minutes at 10lb/sq.in., and stored at 4°C.

Solution 2

NaOH (freshly diluted from a 10 N stock)	0.2 N
SDS	1% (w/v)

Solution 3

5 M potassium acetate	60 ml
Glacial acetic acid	11.5 ml
H_2O	28.5 ml

The resulting solution is 3 M with respect to potassium and 5 M with respect to acetate.

Solution 4

Ammonium acetate	7.5M

* Buffers and solutions for RNA manipulation

Lysis buffer for RNA preparation (pH 5.5)

Sodium acetate	20 mM
EDTA	1 mM
SDS	0.5% (w/v)

RNA suspension buffer

Tris-HCl (pH 7.9)	40 mM
$MgCl_2$	6 mM
NaCl	10 mM
$CaCl_2$	10 mM

* Other Buffers and solutions for DNA/RNA manipulation

3 M pH 4.6 Sodium acetate (100 ml)

Sodium acetate	24.609 g
Glacier acetic acid	8 ml
Water	to a final volume of 100 ml

3 M pH 5.2 Sodium acetate buffer (100 ml)

Sodium acetate	40.84 g
Glacier acetic acid	8 ml
Water	to a final volume of 100 ml

3 M pH 5.6 potassium acetate buffer (100 ml)

potassium acetate	29.442 g
Glacier HAc	9 ml
Water	to a final volume of 100 ml

Tris-HCl buffered phenol

Phenol	500 ml
Tris-HCl, (1 M, pH 8.0)	500 ml
8-Hydroxyquinoline	500 mg

mixed thoroughly and allowed to equilibrate. The upper aqueous phase is removed by aspiration and further equilibrated twice with equal volumes of 1 M Tris-HCl (pH8.0), followed by three additions of an equal volume of 0.1 M Tris HCl (pH 8.0). The final upper aqueous layer is left above the phenol phase.

10 × TAE buffer (per litre)

Tris-base	60.55 g
Sodium acetate	14.3 g
Na_2EDTA	9.3 g

10 × TBE buffer (per litre)

Tris-base	54 g
Boric acid	27.6 g
EDTA	3.72 g

100 × TE buffer

Tris-HCl (pH 8.0)	1 M
EDTA	100 mM

TES buffer

1M Tris (pH8.0)	5 ml
0.25 M EDTA (pH8.0)	2 ml
5 M NaCl	1 ml
H_2O	92 ml

Tracking dye

Ficoll	15% (w/v)
Bromophenol blue	0.1% (w/v)
RNase A	0.1 mg/ml

- **10 × PCR buffer** (pH 8.3)

KCl	50 mM
Tris	10 mM
$MgCl_2$	2.5 mM

- **SDS-PAGE**

SDS lysing buffer

Tris (pH 6.8)	0.25 mM
SDS	2% (w/v)
Glycerol	10% (v/v)
β-mercaptoethanol	5% (v/v)
Bromphenol blue	15% (v/v)

Coomassie stain solution (per litre)

Coomassie brilliant blue R250	2.75 g
Ethanol	100 ml
Methanol	100 ml
Acetic acid	75 ml

Coomassie Destain solution (per litre)

Ethanol	100 ml
Methanol	100 ml
Acetic acid	75 ml

- **Southern hybridisation buffers**

Depurination solution

HCl	250 mM

Denaturation solution

NaOH	0.5 N
NaCl	1.5 M

Neutralisation solution

Tris-HCl (pH 7.5)	0.5 M
NaCl	3 M

20 × SSC buffer

NaCl	3 M
Sodium citrate (pH 7.0)	300 mM

5 × SSC buffer

NaCl	750 mM
Sodium citrate (pH 7.0)	75 mM

Prehybridisation solution (per litre)

Formamide	500 ml
SSPE	5 ×
Skim milk	10 g
SDS	70 g
Single stranded herring sperm DNA (Sigma)	250 mg

5 × SSPE

NaCl	0.75 M
Sodium phosphate buffer (pH 7.4)	0.44 M
EDTA	5 mM

Hybridisation buffer I (1×), pH 7.5

Tris	100 mM
NaCl	150 mM

Make up solution to 800 ml, pH 7.5. Add Milli Q H_2O to a final volume of 1,000 ml.

Hybridisation buffer II (1×)

Dissolve skim milk in hybridisation buffer I to a final concentration of 5% so as to make up hybridisation buffer II.

Hybridisation buffer III (1×)

Tris-HCl, pH 9.5 (+20 °C)	100 mM
NaCl	100 mM
$MgCl_2$	50 mM

Hybridisation buffer IV (1×)

TE buffer	1 ×

Appendix B Media

- **Glycerol-based medium for bacteria maintenance at –70 °C** (per 100 ml)

Glycerol	15 ml
Bacto peptone	1.0 g
Water to a final volume of	100 ml

- **LB medium** (pH 7.6) (per litre)

Bacto yeast extract	5.0 g
Bacto tryptone	20 g
$MgSO_4$	5.0 g
Water to a final volume of	1 litre

- **Nutrient broth** (pH 7.4 ± 0.2) (per litre)

Peptone	5.0 g
NaCl	5.0 g
Yeast extract	2.0 g
Beef extract	1.0 g
Water to a final volume of	1 litre

- **PHA medium** (Atlas, 1993)

Part A:

Composition per 900 ml

$K_2HPO_4 \bullet 3HO$	0.6 g
KH_2PO_4	0.2 g
$MgSO_4 \bullet 7H_2O$	0.2 g
$(NH_4)_2SO_4$	0.2 g

Dissolve the above components into Milli Q water to a final volume of 900 ml. pH7.2. Autoclave for 15 min at 121 °C. Cool to 25 °C.

Part B (per 100 ml):

Glucose	10 g

Dissolve glucose into Milli Q water to a final volume of 100 ml. Autoclave for 15 min at 121 °C. Cool to 25 °C.

Preparation of medium: Aseptically combine 900.0 ml of cooled, sterile Part A and 100 ml of cooled, sterile Part B. Mix thoroughly. Aseptically distribute into sterile tubes or flasks.

- **R-medium** (per litre) (Kim *et al.*, 1994a)

KH_2PO_4	13.5g
$(NH_4)_2HPO_4$	4.0 g
$MgSO_4 \bullet 7H_2O$	1.4 g
Citric acid	1.7 g
Trace metal solution	1.0 ml

Trace metal solution (per litre):

$FeSO_4 \bullet 7H2O$	10.0 g
$CaCl_2 \bullet 2H_2O$	2.0 g
$ZnSO_4 \bullet 7H_2O$	2.2 g
$MnSO_4 \bullet H_2O$	0.5 g
$CuSO_4 \bullet 5H_2O$	1.0 g
$(NH_4)_6Mo_7O_{24} \bullet 4H_2O$	0.1 g
$Na_2B_4O_7 \bullet 10H_2O$	0.02 g
5 M HCl to a final volume of	1 litre

- **SOC medium** (pH 7.0) (per litre) (Sambrook *et al.*,1989)

Bacto tryptone (Difco)	20 g
Bacto yeast extract (Difco)	5.0 g
NaCl	10 mM
KCl	2.5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	10 mM
Water to a final volume of	1 litre

- **Terrific Broth** (per litre) (Sambrook *et al.*,1989)

To 900 ml of deionised H_2O, add:

Bacto tryptone	12 g
Bacto yeast extract	24 g
glycerol	4 ml

Shake until the solutes have dissolved and sterilise by autoclaving for 20 minutes at 15lb/sq. in. on liquid cycle. Allow the solution to cool to 60 °C or less, and then add 100 ml of a sterile solution of 0.17 M KH_2PO_4, 0.72 M K_2HPO_4. (This solution is made by dissolving 2.31 g of KH_2PO_4 and 12.54 g of K_2HPO_4 in 90 ml of deionised H_2O. After the salts have dissolved, adjust the volume of the solution to 100 ml with deionised H_2O and sterilise by autoclaving for 20 minutes at 15 lb/sq. in. on liquid cycle.)

- **2 × YT Medium** (per litre) (Sambrook *et al.*,1989)

Bacto Tryptone	16 g
Yeast Extract	10 g

NaCl	5 g
Water to a final volume of	1 litre

Appendix C DNA sequence alignment analysis of *pha* genes

A: Multiple alignment of DNA sequence data for published *phaA* genes

Table A. Description of Genbank sequence entries used in the multiple alignment of *phaA* genes.

Accession No.	Source	Definition	Reference
A27012	*Chromatium vinosum*	PHA synthesis genes	Patent: WO 9302194-A 12 04-FEB-1993
A27001	*Chromatium vinosum*	*C.vinosum* complete sequence of fragment SE45	Patent: WO 9302194-A 1 04-FEB-1993
D49362	*Paracoccus denitrificans*	genes for beta-ketothiolase and acetoacetyl-CoA reductase	Yabutani *et al.* (1995)
I35761	Unknown		John (1997)
J04987	*R. eutropha* (strain H16)	beta-ketothiolase (*phbA*) and acetoacetyl-CoA reductase (*phbB*) genes	Peoples and Sinskey (1989b)
J02631	*Zoogloea ramigera*	thiolase gene	Peoples *et al.* (1987); and Williams *et al.* (1992)
L01112	*Chromatium vinosum* (strain D)	poly(3-hydroxybutyric acid) (*phbA*, *phbB*, *phbC*) genes	Liebergesell and Steinbuchel (1992)
L01113	*Thiocystis violacea* (strain 2311)	beta-ketothiolase and PHA synthase (*phbA* and *phbC*) genes	Liebergesell and Steinbuchel (1993)
S54369	*Thiocystis violacea*	*phbCAB operon*	Liebergesell and Steinbuchel (1993)
U17226	*Sinorhizobium meliloti* (strain 41)	beta-ketothiolase (*phbA*) and acetoacetyl CoA reductase (*phbB*) gene	Tombolini *et al.* (1995)
U88653	*Pseudomonas aeruginosa* (strain PAO1)	thiolase (*phaA*) gene	Direct Submission

```
Genbank accession
number:

J02631          ATGAGCACCCCGTCCATCGTCATCGCCAGCGC---CCGCACCGCGGTCGGTTCCTTCAAC
U17226          ATGAGCAATCCCTCGATCGTCATCGCCAGCGCGGCTCGCACGGCCGTCGGCTCCTTCAAC
D49362          ------ATGACCAAAGCCGTAATCGTTTCCGCTGCCCGCACCCCCGTCGGCAGCTTTCTG
J04987          ---ATGACTGA---CGTTGTCATCGTATCCGCCGCCCGCACCGCGGTCGGCAAGTTTGGC
I35761          ---ATGACTGA---CGTTGTCATCGTATCCGCCGCCCGCACCGCGGTCGGCAAGTTTGGC
A27012          ---ATGAGCGAGAACATCGTCATCGTCGACGCCGGCCGCAGTGCCATCGGAACCTTCGGC
A27001          ---ATGAGCGAGAACATCGTCATCGTCGACGCCGGCCGCAGTGCCATCGGAACCTTCGGC
L01112          ---ATGAGCGAGAACATCGTCATCGTCGACGCCGGCCGCAGTGCCATCGGAACCTTCGGC
S54369          ---ATGAGCGACACTATCGTTATCGTCGATGCGGGTCGGACCGCGATCGGAACCTTCGGC
L01113          ---ATGAGCGACACTATCGTTATCGTCGATGCGGGTCGGACCGCGATCGGAACCTTCGGC
U88653          ------ATGTCCGATATCGTCATCGTCGCCGGCGCCCGTACTCCCATGGGCGGCTTCCAG
                      *           ** ****     *     ** *   *  * **    **

J02631          GGCGCTTTCGCCAACACGCCCGCCCATGAACTCGGGGCGACCGTGATTTCGGCGGTTCTC
U17226          GGCGCCTTCGGCAATACTCTCGCGCACGAACTGGGTGCGGCCGCCATCAAGGCGGTGCTG
D49362          GGATCCTTTGCGAACCTGCCGGCGCATGAGCTTGGCGCCATCGTCCTGAAAGCGGTCGTC
J04987          GGCTCGCTGGCCAAGATCCCGGCACCGGAACTGGGTGCCGTGGTCATCAAGGCCGCGCTG
I35761          GGCTCGCTGGCCAAGATCCCGGCACCGGAACTGGGTGCCGTGGTCATCAAGGCCGCGCTG
A27012          GGCAGTCTGTCGTCACTCTCGGCCACCGAGATCGGCACCGCCGTGCTCAAGGGGCTGCTG
A27001          GGCAGTCTGTCGTCACTCTCGGCCACCGAGATCGGCACCGCCGTGCTCAAGGGGCTGCTG
L01112          GGCAGTCTGTCGTCACTCTCGGCCACCGAGATCGGCACCGCCGTGCTCAAGGGGCTGCTG
S54369          GGCGCACTCTCGGCCCTCCAGGCGACCGATATCGGCACGACCGTGCTCAAGGCCCTGATC
L01113          GGCGCACTCTCGGCCCTCCAGGCGACCGATATCGGCACGACCGTGCTCAAGGCCCTGATC
U88653          GGCAGCCTGGCCGGCGTTTCCGCGGTCGACCTGGGTGCCGTCGCGATCCGCGAAGCGGTG
                **     *              **    **  * **  *    *   *    *      *

J02631          GAGCGCGCGGGCGTTGCGGCGGGCGAGGTGAACGAGGTGATTCTCGGCCAGGTGCTGCCG
U17226          GAGCGGGCTGGCGTCGAAGCGGGCGAGGTGGACGAGGTGATCCTCGGTCAGGTGCTGCCG
D49362          GAGCGCGCCGGAATCGATCCTTCCGAAGTGTCCGAAACCATCCTGGGCCAGGTGCTGACC
J04987          GAGCGCGCCGGCGTCAAGCCGGAGCAGGTGAGCGAAGTCATCATGGGCCAGGTGCTGACC
I35761          GAGCGCGCCGGCGTCAAGCCGGAGCAGGTGAGCGAAGTCATCATGGGCCAGGTGCTGACC
A27012          GCGCGTACCGGACTCGCGCCGGAACAGATCGACGAGGTGATTCTCGGCCAGGTGCTGACC
A27001          GCGCGTACCGGACTCGCGCCGGAACAGATCGACGAGGTGATTCTCGGCCAGGTGCTGACC
L01112          GCGCGTACCGGACTCGCGCCGGAACAGATCGACGAGGTGATTCTCGGCCAGGTGCTGACC
S54369          GAGCGGACCGGAATCGCCCCCGAGCAGGTCAGCGAAGTTATTCTGGGGCAGGTGCTGACC
L01113          GAGCGGACCGGAATCGCCCCCGAGCAGGTCAGCGAAGTTATTCTGGGGCAGGTGCTGACC
U88653          CAGCGCGCCGGCATCGCCGCCGAGGACGTACAGGAAGTGATCATGGGTTGCGTCCTGCCC
                  ***  * **  *     *     *  *    **    **  * **    ** *** *

J02631          GCCGGCGAAGGCCAGAACCCGGCCCGCCAGGCCGCCATGAAGGCCGGCGTGCCGCAGGAG
U17226          GCGGGCGAGGGGCAGAATCCTGCGCGGCAAGCGGCGATGAAGGCCGGTCTCCCGCAGGAA
D49362          GCCGCCCAGGGCCAGAACCCCGCCCGCCAGGCGCATATCAAGGTTGGCTTACCGCGGGAA
J04987          GCCGGTTCGGGCCAGAACCCCGCACGCCAGGCCGCGATCAAGGCCGGCCTGCCGGCGATG
I35761          GCCGGTTCGGGCCAGAACCCCGCACGCCAGGCCGCGATCAAGGCCGGCCTGCCGGCGATG
A27012          GCCGGCGTGGGCCAGAACCCCGCCCGTCAGACCACGCTGCACGCGGGGCTACCGCATTCG
A27001          GCCGGCGTGGGCCAGAACCCCGCCCGTCAGACCACGCTGCACGCGGGGCTACCGCATTCG
L01112          GCCGGCGTGGGCCAGAACCCCGCCCGTCAGACCACGCTGCACGCGGGGCTACCGCATTCG
S54369          GCGGGCTGCGGTCAGAATCCGGCGCGTCAGACCACCTTGATGGCGGGGCTGCCGCACACG
L01113          GCGGGCTGCGGTCAGAATCCGGCGCGTCAGACCACCTTGATGGCGGGGCTGCCGCACACG
U88653          GCCGGCCTGAAGCAGGGCCCGGCCCGCCAGGCCGCGCTGGCGGCCGGCCTGCCCGCCGCC
                ** *         ***   ** ** ** **  *     *    *  **  * **

J02631          GCGACCGCCTGGGGCATGAACCAGCTTTGCGGCTCGGGCCTGCGCGCCGTCGCGCTCGGC
U17226          AAGACCGCCTGGGGCATGAACCAGCTTTGCGGCTCGGGCCTGCGCGCCGTAGCGCTCGGC
D49362          TCGGCGGCCTGGGTCATCAACCAGGTCTGCGGCTCGGGTCTGCGTACGGTGGCGCTGGCG
J04987          GTGCCGGCCATGACCATCAACAAGGTGTGCGGCTCGGGCCTGAAGGCCGTGATGCTGGCC
I35761          GTGCCGGCCATGACCATCAACAAGGTGTGCGGCTCGGGCCTGAAGGCCGTGATGCTGGCC
A27012          GTGCCGGCCATGACCATCAACAAGGTCTGCGGCAGCGGTCTGAAGGCGGTGCATCTGGCG
A27001          GTGCCGGCCATGACCATCAACAAGGTCTGCGGCAGCGGTCTGAAGGCGGTGCATCTGGCG
L01112          GTGCCGGCCATGACCATCAACAAGGTCTGCGGCAGCGGTCTGAAGGCGGTGCATCTGGCG
S54369          GTCCCGGCGATGACCATCAACAAGGTCTGCGGCAGTGGGCTCAAGGCGGTCCATCTGGCC
L01113          GTCCCGGCGATGACCATCAACAAGGTCTGCGGCAGTGGGCTCAAGGCGGTCCATCTGGCC
U88653          ACCGGCTGCACCACCATCAACAAGCTGTGCGGCTCGGGGATGAAGGCGGTGATGCTGGCC
                          *** *** ** * ******   **  *     * **    ** *

                      5'    ACCATCAACAAGGTGTGCGG    3'
Primer #4033                GG   G   C  C C
                             T          T
```

```
J02631          ATGCAGCAGATCGCCACGGGCGATGCGAGCATCATCGTCGCCGGCGGCATGGAATCCATG
U17226          ATGCAGCAGATCGCAACCGGCGATGCGAAGGTCATCGTTGCCGGCGGCATGGAGTCGATG
D49362          GCGCAGCAGGTTCTGCTGGGCGACGCCAGGATCGTCGTGGCCGGCGGCCAGGAATCCATG
J04987          GCCAACGCGATCATGGCGGGCGACGCCGAGATCGTGGTGGCCGGCGGCCAGGAAAACATG
I35761          GCCAACGCGATCATGGCGGGCGACGCCGAGATCGTGGTGGCCGGCGGCCAGGAAAACATG
A27012          ATGCAGGCCATCGCCTGCGGGGATGCCGACATCGTCATCGCCGGCGGTCAGGAGAGCATG
A27001          ATGCAGGCCATCGCCTGCGGGGATGCCGACATCGTCATCGCCGGCGGTCAGGAGAGCATG
L01112          ATGCAGGCCATCGCCTGCGGGGATGCCGACATCGTCATCGCCGGCGGTCAGGAGAGCATG
S54369          ATGCAGGCGGTTGCCTGTGGCGATGCGGAGATCGTGATCGCCGGTGGTCAAGAGAGCATG
L01113          ATGCAGGCGGTTGCCTGTGGCGATGCGGAGATCGTGATCGCCGGTGGTCAAGAGAGCATG
U88653          CACGACCTGCTCAAGGCTGGCACCAACCAGGTGATGGTCGCCGGTGGCATGGAAAGCATG
                    *     *        **             *  *  * ***** **     **     ***

J02631          TCCATGGCCCCGCATTGCG---CGCATCTG---GCCGGCGT-GAAGATGGGCGATTTCAA
U17226          TCGATGGCGCCGCATTGCG---CGCACCTGCGCGGCGGCGT-GAAGATGGGCGACTACAA
D49362          TCGCTGGCACCCCATGCCG---CCTATATC-GCGCCGGGGCAGAAGATGGGCGACATGAA
J04987          AGCGCCGCCCCGCACGTGCTGCCGGGCTCGCGCGATGGTTT-CCGCATGGGCGATGCCAA
I35761          AGCGCCGCCCCGCACGTGCTGCCGGGCTCGCGCGATGGTTT-CCGCATGGGCGATGCCAA
A27012          AGCCAGTCCTCGCACGTCCTGCCGCGTTCGCGCGACGGTCA-GCGCATGGGCGACTGGTC
A27001          AGCCAGTCCTCGCACGTCCTGCCGCGTTCGCGCGACGGTCA-GCGCATGGGCGACTGGTC
L01112          AGCCAGTCCTCGCACGTCCTGCCGCGTTCGCGCGACGGTCA-GCGCATGGGCGACTGGTC
S54369          AGTCAGTCGTCGCACGTCCTGCCGCGCTCGCGCGAGGGGCA-GCGCATGGGCGACTGGCC
L01113          AGTCAGTCGTCGCACGTCCTGCCGCGCTCGCGCGAGGGGCA-GCGCATGGGCGACTGGCC
U88653          TCCAACGCGCCCTATGTCCTGGAGAAAGCCCGCAGCGGCCT-GCGCATGGGCCACGGCGA
                      *  *  *                          **        ****** *

J02631          GATGATCGACACGATGATCAAGGACGGCCTGACCGACGCCTTCTACGGCTACCACATGGG
U17226          GATGATCGACACGATGATCAAGGACGGCCTGACGGATGCCTTCTACGGCTACCACATGGG
D49362          GATGCTCGACACCATGATCAAGGACGGGCTCTGGGACGCCTTCAACGATTACCACATGGG
J04987          GCTGGTCGACACCATGATCGTCGACGGCCTGTGGGACGTGTACAACCAGTACCACATGGG
I35761          GCTGGTCGACACCATGATCGTCGACGGCCTGTGGGACGTGTACAACCAGTACCACATGGG
A27012          GATGAAGGACACCATGATCGTCGACGGCCTCTGGGATGCCTTCAACAACTATCACATGGG
A27001          GATGAAGGACACCATGATCGTCGACGGCCTCTGGGATGCCTTCAACAACTATCACATGGG
L01112          GATGAAGGACACCATGATCGTCGACGGCCTCTGGGATGCCTTCAACAACTATCACATGGG
S54369          GATGAAGGACACCATGATCGTCGACGGCCTCTGGGATGCCTTCAATCAGTGCCACATGGG
L01113          GATGAAGGACACCATGATCGTCGACGGCCTCTGGGATGCCTTCAATCAGTGCCACATGGG
U88653          GATCAAGGACCACATGTTCCTCGACGGCCTGGAGGACGCCCGTACCGGCCGTTTGATGGG
                * *    ***   *** **   ***** **    ** *                  *****
                         3' TACTAGCAGCTGCCGGACAC  5' Primer #4034
                               A  GTC     C  GTG
                                  T

J02631          CACGACCGCCGAGAATGTCGCCAAGCAGTGGCAGCTTTCCCGCGACGAGCAGGACGCCTT
U17226          CATCACCGCCGAGAACGTTGCGCGGAAATGGCAGCTGACGCGCGAGGAACAGGACGAATT
D49362          CACCACGGCCGAAAACGTCGCCGGCAAATGGGAAATCTCGCGGGCCGAGCAGGACCAGTT
J04987          CATCACCGCCGAGAACGTGGCCAAGGAATACGGCATCACACGCGAGGCGCAGGATGAGTT
I35761          CATCACCGCCGAGAACGTGGCCAAGGAATACGGCATCACACGCGAGGCGCAGGATGAGTT
A27012          CACCACCGCCGAGAACATCGCCCAGAAGTACGGCTTTACGCGCGAGCAGCAGGACGCCTT
A27001          CACCACCGCCGAGAACATCGCCCAGAAGTACGGCTTTACGCGCGAGCAGCAGGACGCCTT
L01112          CACCACCGCCGAGAACATCGCCCAGAAGTACGGCTTTACGCGCGAGCAGCAGGACGCCTT
S54369          TGTCACGGCGGAGAACATCGCGAAGAAGTACGCCTTCACCCGCGAGGCGCAGGATGCGTT
L01113          TGTCACGGCGGAGAACATCGCGAAGAAGTACGCCTTCACCCGCGAGGCGCAGGATGCGTT
U88653          TTCGTTCGCCCAGGAGACCGCCGACAAGTACGGCGTGACCCGCGAGGAGATGGACGCCTA
                      **  *  *    **     * *      *  * ** *       ***    *

J02631          CGCCGTCGCCTCGCAGAACAAGGCCGAGGCCGCCCAGAAGGACGGCCGCTTCAAGGACGA
U17226          CGCGCTTGCCTCCCAGAACAAGGCGGAAGCGGCCCAGAAGGCCGGCCGTTTCGCCGACGA
D49362          CGCCGTCGCCTCGCAGAACAAGGCCGAGGCTGCGCAGAAGGCCGGCAAGTTCGCCGACGA
J04987          CGCCGTCGGCTCGCAGAACAAGGCCGAAGCCGCGCAGAAGGCCGGCAAGTTTGACGAAGA
I35761          CGCCGTCGGCTCGCAGAACAAGGCCGAAGCCGCGCAGAAGGCCGGCAAGTTTGACGAAGA
A27012          CGCCGCCGCCTCGCAGCAGAAGACCGAGGCCGCGCAGAAGGCTGGCCGCTTCCAGGACGA
A27001          CGCCGCCGCCTCGCAGCAGAAGACCGAGGCCGCGCAGAAGGCTGGCCGCTTCCAGGACGA
L01112          CGCCGCCGCCTCGCAGCAGAAGACCGAGGCCGCGCAGAAGGCTGGCCGCTTCCAGGACGA
S54369          CGCCGCGGCGTCCCAGCAGAAGGCGGAGGCCGCCATCCAATCCGGACGCTTTGCCGACGA
L01113          CGCCGCGGCGTCCCAGCAGAAGGCGGAGGCCGCCATCCAATCCGGACGCTTTGCCGACGA
U88653          TGCCATCGAATCGCTCAAGCGCGCCCAGGCCGCCATCGCCGATGGCTCGCTGGCGGCGGA
                 **    *  ** *   *     *  * ** **          **     *    *  **
```

```
J02631          GATCGTTCCCTTCATCGTCAAGGGCCGCAAGGGCGACATCACGG---TCGATGCCGACGA
U17226          GATCGTGCCTTTCGTCGTGAAGACGCGCAAGGGCGACGTAAACG---TGGATCAGGACGA
D49362          GATTGTGCCCGTGACCATCAAGTCCCGAAAGGGCGAGACGGTGG---TCGATGCCGATGA
J04987          GATCGTCCCGGTGCTGATCCCGCAGCGCAAGGGCGACCCGGTGGCCTTCAAGACCGACGA
I35761          GATCGTCCCGGTGCTGATCCCGCAGCGCAAGGGCGACCCGGTGGCCTTCAAGACCGACGA
A27012          GATCATTCCGATCGAGATTCCGCAGCGCAAGGGCGATCCGAAGGTGTTTGATGCCGACGA
A27001          GATCATTCCGATCGAGATTCCGCAGCGCAAGGGCGATCCGAAGGTGTTTGATGCCGACGA
L01112          GATCATTCCGATCGAGATTCCGCAGCGCAAGGGCGATCCGAAGGTGTTTGATGCCGACGA
S54369          GATCATCCCGGTGAGCATCCCGCAGCGCAAGGGCGATCCCCTGGTGTTCGACACCGACGA
L01113          GATCATCCCGGTGAGCATCCCGCAGCGCAAGGGCGATCCCCTGGTGTTCGACACCGACGA
U88653          AATCGTCCCGGTCACCGTCACCAGCCGCAAGGGCGAGAGCGTGG---TGAAGGACGACGA
                 **  * **  *     *        ** ********       *   *  *    ** **

J02631          ATATATCCGCCACGGCGCGACGCTCGATTCCATGGCGAAGCTCCGCCCGGCCTTCGACAA
U17226          GTACATCCGCCACGGTGCCACGCTGGATTCGATCGCAAAGCTCCGCCCGGCCTTCGACAA
D49362          ATACATCCGCCACGGCGCCACGCTGGAGGCAATGGAGAACGTGCGCCCCGCCTTTAGCAA
J04987          GTTCGTGCGCCAGGGCGCCACGCTGGACAGCATGTCCGGCCTCAAGCCCGCCTTCGACAA
I35761          GTTCGTGCGCCAGGGCGCCACGCTGGACAGCATGTCCGGCCTCAAGCCCGCCTTCGACAA
A27012          GTTCCCGCGTCACGGCACCACGGCCGAGAGTCTGGGCAAGCTGCGTCCGGCCTTCTCGAA
A27001          GTTCCCGCGTCACGGCACCACGGCCGAGAGTCTGGGCAAGCTGCGTCCGGCCTTCTCGAA
L01112          GTTCCCGCGTCACGGCACCACGGCCGAGAGTCTGGGCAAGCTGCGTCCGGCCTTCTCGAA
S54369          GTTTCCGCGTCCGGGCACCACGGCCGAAACGCTCGGCAGGCTGCGCCCCGCCTTCGACAA
L01113          GTTTCCGCGTCCGGGCACCACGGCCGAAACGCTCGGCAGGCTGCGCCCCGCCTTCGACAA
U88653          GCAACCG---CTGACCGCCAACCTGGAGAAGATCCCGAGCCTCAGGCCGGCGTTCCGCAA
                       *      * *     **     *      *     ** ** **     **

J02631          GGAAGGCACGGTGACGGCCGGCAACGCCTCCGGCCTCAATGACGGCGCGGCCGCGGCCCT
U17226          GGAAGGTACTGTGACCGCGGGCAACGCTTCGGGTCTCAACGACGGTGCCGCCGCGGCGCT
D49362          GGAGGGGACGGTGACGGCCGGCAATGCCTCGGGCCTGAACGACGGCGCCGCCGCCGTGCT
J04987          GGCCGGCACGGTGACCGCGGCCAACGCCTCGGGCCTGAACGACGGCGCCGCCGCGGTGGT
I35761          GGCCGGCACGGTGACCGCGGCCAACGCCTCGGGCCTGAACGACGGCGCCGCCGCGGTGGT
A27012          GGACGGCAGCGTCACGGCGGGTAACGCCTCCGGCATCAACGACGGGGCGGCCATGGTCGT
A27001          GGACGGCAGCGTCACGGCGGGTAACGCCTCCGGCATCAACGACGGGGCGGCCATGGTCGT
L01112          GGACGGCAGCGTCACGGCGGGTAACGCCTCCGGCATCAACGACGGGGCGGCCATGGTCGT
S54369          GCAAGGCACTGTGACCGCCGGCAACGCCTCCGGGATCAACGACGGCGCGGCCATGGTCGT
L01113          GCAAGGCACTGTGACCGCCGGCAACGCCTCCGGGATCAACGACGGCGCGGCCATGGTCGT
U88653          GGACGGTACCATCACCGCCGCCAATGCCAGCTCGATCTCCGACGGCGCCTCCGCGCTGGT
                *   ** *   * ** ** *  ** **        *    ***** **  **       *

J02631          CCTGATGAGCGAAGCGGAAGCCTCGCGCCGCGGCATCCAGCCGCTCGGCCGCATCGTTTC
U17226          CTTGATGACCGAGGCCGAGGCGGCCCGGCGCGGCATCCAGCCGCTTGCCCGCATCGTCTC
D49362          GGTCATGACCGAGGACGAGGCCGCTCGCCGCGGCCTGACGCCGCTGGCCCGCATCGCCTC
J04987          GGTGATGTCGGCGGCCAAGGCCAAGGAACTGGGCCTGACCCCGCTGGCCACGATCAAGAG
I35761          GGTGATGTCGGCGGCCAAGGCCAAGGAACTGGGCCTGACCCCGCTGGCCACGATCAAGAG
A27012          GGTGATGAAGGAGTCCAAGGCCAAGGAACTGGGTCTGAAGCCGATGGCGCGTCTGGTGGC
A27001          GGTGATGAAGGAGTCCAAGGCCAAGGAACTGGGTCTGAAGCCGATGGCGCGTCTGGTGGC
L01112          GGTGATGAAGGAGTCCAAGGCCAAGGAACTGGGTCTGAAGCCGATGGCGCGTCTGGTGGC
S54369          GGTGATGAAGGAATCCAAGGCCAAGGAACTGGGTTTGACGCCCATGGCGCGCCTGGTCGC
L01113          GGTGATGAAGGAATCCAAGGCCAAGGAACTGGGTTTGACGCCCATGGCGCGCCTGGTCGC
U88653          CCTGATGACCGCCGAGGAAGCCCAGCGCCGTGGCCTGAAGCCGCTGGCGCGAATCGTCGG
                  * ***   *      * **       *  **  *    **  * *      *

J02631          CTGGGCGACGGTCGGCGTCGATCCCAAGGTCATGGGCACCGGCCCGATCCCGGCCTCCCG
U17226          CTGGGCCACGGCAGGCGTCGACCCGCAGATCATGGGCACCGGCCCCATCCCCGCATCGCG
D49362          CTATGCGACCGCCGGTGTGGACCCGCAGATCATGGGCACCGGCCCGATCCCCGCCAGCCG
J04987          CTATGCCAACGCCGGTGTCGATCCCAAGGTGATGGGCATGGGCCCGGTGCCGGCCTCCAA
I35761          CTATGCCAACGCCGGTGTCGATCCCAAGGTGATGGGCATGGGCCCGGTGCCGGCCTCCAA
A27012          CTTCGCCAGCGCCGGTGTCGATCCGGCGATCATGGGGACGGGCCCGATCCCGGCGTCGAC
A27001          CTTCGCCAGCGCCGGTGTCGATCCGGCGATCATGGGGACGGGCCCGATCCCGGCGTCGAC
L01112          CTTCGCCAGCGCCGGTGTCGATCCGGCGATCATGGGGACGGGCCCGATCCCGGCGTCGAC
S54369          GTTCTCGAGCGCGGGCGTGGATCCGGCCATCATGGGCACCGGACCTATCCCGGCTTCGAC
L01113          GTTCTCGAGCGCGGGCGTGGATCCGGCCATCATGGGCACCGGACCTATCCCGGCTTCGAC
U88653          CCATGCCACCCAGAGCCAGGACCCGAGCGAGTTCACCCTGGCGCCGATCGGCGCGATGAC
                     * *      *    ** **        *       *  **  *    **

J02631          CAAGGCGCTCGAGCGCGCCGGCTGGAAGATCGGCGATCTCGACCTCGTGGAAGCCAACGA
U17226          CAAGGCCCTCGAAAAGGCCGGCTGGTCGGTCGCCGATATCGAGCTCGTGGAGGCGAACGA
D49362          CAAGGCGCTGGAAAAGGCCGGCTGGTCGGTCGGCGACCTGGACCTGGTCGAGGCGAACGA
J04987          GCGCGCCCTGTCGCGCGCCGAGTGGACCCCGCAAGACCTGGACCTGATGGAGATCAACGA
I35761          GCGCGCCCTGTCGCGCGCCGAGTGGACCCCGCAAGACCTGGACCTGATGGAGATCAACGA
A27012          CAAGTGCCTGGAGAAGGCCGGCTGGACCCCGGCGGATCTGGATCTGATCGAGGCCAACGA
A27001          CAAGTGCCTGGAGAAGGCCGGCTGGACCCCGGCGGATCTGGATCTGATCGAGGCCAACGA
L01112          CAAGTGCCTGGAGAAGGCCGGCTGGACCCCGGCGGATCTGGATCTGATCGAGGCCAACGA
S54369          CGATTGTCTGAAGAAGGCCGGCTGGGCCCCGGCTGATCTGGACCTTGTCGAGGCCAATGA
L01113          CGATTGTCTGAAGAAGGCCGGCTGGGCCCCGGCTGATCTGGACCTTGTCGAGGCCAATGA
U88653          CAACCTGTTCGCCAGGACCGGCTGGAGCAAGGACGACGTCGACCTGTTCGAGATCAACGA
                       *        ***  ***          **  * ** **  * **    ** **
```

```
J02631          AGCCTTCGCGGCGCAGGCCTGCGCGGTCAACAAGGACCTCGGCTGGGATCCGTCCATCGT
U17226          GGCTTTCGCGGCTCAGGCCTGCGCCGTCAACAAGGACCTCGGCTGGGATCCTTCGATCGT
D49362          GGCCTTCGCTGCCCAGGCCATCGCCGTGAACCGCGACATGGGCTGGGATCCCTCGATCGT
J04987          GGCCTTTGCCGCGCAGGCGCTGGCGGTGCACCAGCAGATGGGCTGGGACACCTCCAAGGT
I35761          GGCCTTTGCCGCGCAGGCGCTGGCGGTGCACCAGCAGATGGGCTGGGACACCTCCAAGGT
A27012          AGCCTTCGCCGCGCAGGCCATGTCGGTCAACCAGGACATGGGCTGGGATCTGTCCAAGGT
A27001          AGCCTTCGCCGCGCAGGCCATGTCGGTCAACCAGGACATGGGCTGGGATCTGTCCAAGGT
L01112          AGCCTTCGCCGCGCAGGCCATGTCGGTCAACCAGGACATGGGCTGGGATCTGTCCAAGGT
S54369          GGCCTTTGCCGCTCAGGCGATGTCCGTCAATCAGGAGATGGGCTGGGATCTCAGCAAGGT
L01113          GGCCTTTGCCGCTCAGGCGATGTCCGTCAATCAGGAGATGGGCTGGGATCTCAGCAAGGT
U88653          AGCCTTCGCCATGGTCACCATGCTCGCCATGCGCGAACACGGCCTCGATCATGCCAAGGT
                 ** ** **         *        *          *    ***   **        *  **

J02631          CAACGTCAACGGCGGTGCCATCGCCATCGGCCACCCGATCGGCGCGTCCGGCGCCCGCAT
U17226          CAACGTCAATGGCGGAGCGATCGCCATCGGCCATCCGATCGGTGCCTCCGGTGCCCGCGT
D49362          GAACGTGAACGGCGGCGCCATCGCCATCGGCCATCCCATCGGCGCCTCGGGCTGCCGCAT
J04987          CAATGTGAACGGCGGCGCCATCGCCATCGGCCACCCGATCGGCGCGTCGGGCTGCCGTAT
I35761          CAATGTGAACGGCGGCGCCATCGCCATCGGCCACCCGATCGGCGCGTCGGGCTGCCGTAT
A27012          CAACGTCAACGGCGGCGCCATCGCCATCGGTCATCCGATCGGCGCCTCCGGTGCGCGCGT
A27001          CAACGTCAACGGCGGCGCCATCGCCATCGGTCATCCGATCGGCGCCTCCGGTGCGCGCGT
L01112          CAACGTCAACGGCGGCGCCATCGCCATCGGTCATCCGATCGGCGCCTCCGGTGCGCGCGT
S54369          AAACGTCAATGGCGGAGCTATCGCCATCGGCCACCCGATCGGTGCCTCTGGTGCGCGTGT
L01113          AAACGTCAATGGCGGAGCTATCGCCATCGGCCACCCGATCGGTGCCTCTGGTGCGCGTGT
U88653          CAATGTCTATGGCGGCGCCTGCGCCCAGGGCCATCCGGTCGGTTCCACCGGCTCGCGGAT
                 ** **  * ***** **   ****   ** ** **  **** *  * **    **  *

J02631          CCTCAACACGCTCCTCTTCGAGATGAAGCGTCGCGGCGCCCGCAAGGGTCTCGCCACGCT
U17226          GCTGAACACGCTTCTTTTCGAAATGAAGCGGCGCGGCGTCTCCAAGGGGCTTGCCACCCT
D49362          CCTGAACACGCTGCTGTTCGAAATGCAGCGCCGCGACGCGAAAAAGGGCCTGGCCACGCT
J04987          CCTGGTGACGCTGCTGCACGAGATGAAGCGCCGTGACGCGAAGAAGGGCCTGGCCTCGCT
I35761          CCTGGTGACGCTGCTGCACGAGATGAAGCGCCGTGACGCGAAGAAGGGCCTGGCCTCGCT
A27012          GCTCGTGACCCTGCTCTATGAGATGCAGAAGCGCGACGCCAAGAAGGGTCTGGCGACGCT
A27001          GCTCGTGACCCTGCTCTATGAGATGCAGAAGCGCGACGCCAAGAAGGGTCTGGCGACGCT
L01112          GCTCGTGACCCTGCTCTATGAGATGCAGAAGCGCGACGCCAAGAAGGGTCTGGCGACGCT
S54369          GCTGGTCACCCTCCTCTATGAGATGCAGAAGCGCGATGCGAAGAAGGGTCTCGCAACCCT
L01113          GCTGGTCACCCTCCTCTATGAGATGCAGAAGCGCGATGCGAAGAAGGGTCTCGCAACCCT
U88653          CATCCTCACCCTGATCAACGCCCTGCGCCAGAAAGGCGGCAAGCGCGGCGTCGCCTCGCT
                  *    ** **  *    *   **         *  *        **  * **  * **

J02631          CTGCATCGGCGGCGGCATGGGCGTGGCGATGTGCATCGAGAGCCTTTAG---
U17226          GTGCATCGGCGGCGGCATGGGCGTCGCCATGTGCGTGGAACGCCTGTA----
D49362          GTGCATCGGCGGCGGCATGGGCGTCGCCCTGTGCCTCGAGCGCCCCTGA---
J04987          GTGCATCGGCGGCGGCATGGGCGTGGCGCTGGCAGTCGAGCGCAAATAA---
I35761          GTGCATCGGCGGCGGCATGGGCGTGGCGCTGGCAGTCGAGCGCAAATAA---
A27012          GTGCATCGGCGGCGGCCAGGGCGTGGCGCTGGCGGTCGAGCGGATG------
A27001          GTGCATCGGCGGCGGCCAGGGCGTGGCGCTGGCGGTCGAGCGGATGTGA---
L01112          GTGCATCGGCGGCGGCCAGGGCGTGGCGCTGGCGGTCGAGCGGATGTGA---
S54369          CTGCATCGGCGGCGGCCAGGGCGTGGCGTTGGCGGTCGAGCGTCTCTGA---
L01113          CTGCATCGGCGGCGGCCAGGGCGTGGCGTTGGCGGTCGAGCGTCTCTGA---
U88653          GTGCATCGGCGGCGGCGAAGCCACCGCCGTGGCGCTCGAACTGCTCTGA---
                 ***************   * *   **  **    * **
```

B: Multiple alignment of DNA sequence data for published *phaB* genes

Table B. Description of Genbank sequence entries used in the multiple alignment of *phaB* genes.

Accession No.	Source	Definition	Reference
A27012	*Allochromatium vinosum*	PHA synthesis genes	Patent: WO 9302194-A 12 04-FEB-1993
D49362	*Paracoccus denitrificans*	genes for beta-ketothiolase and acetoacetyl-CoA reductase	Yabutani *et al.* (1995)
I35762	Unknown		John (1997)
J04987	*R.eutropha* (strain H16)	beta-ketothiolase (*phbA*) and acetoacetyl-CoA reductase (*phbB*) genes	Peoples and Sinskey (1989b)
L01112	*Allochromatium vinosum* (strain D)	poly(3-hydroxybutyric acid) (*phbA*, *phbB*, *phbC*) genes	Liebergesell and Steinbuchel (1992)
U17226	*Sinorhizobium meliloti* (strain 41)	beta-ketothiolase (*phbA*) and acetoacetyl CoA reductase (*phbB*) gene	Tombolini *et al.* (1995)

```
Genbank accession
number:

A27012          ATGGCTC---GTATCGCACTCGTCACCGGCGGCATCGGCGGCATCGGCACTTCGATCTGC
L01112          ATGGCTC---GTATCGCACTCGTCACCGGCGGCATCGGCGGCATCGGCACTTCGATCTGC
I35762          ATGACTCAGCGCATTGCGTATGTGACCGGCGGCATGGGTGGTATCGGAACCGCCATTTGC
J04987          ATGACTCAGCGCATTGCGTATGTGACCGGCGGCATGGGTGGTATCGGAACCGCCATTTGC
D49362          ATGGCAA---AAGTGGCCCTTGTGACCGGCGGGTCGCGCGGCATCGGCGCGGCAATCTCG
U17226          ATGAGCA---GGGTAGCACTGGTAACGGGCGGATCCCGCGGCATTGGCGCTGCGATTTGC
                ***          * **    ** ** *****     * ** ** **  *  * ** *

A27012          ACACGCCTGGCAAAGGATGGCTGCACCGTCGTGGCGAACTGCCATCCGTCCGAGGCGGCC
L01112          ACACGCCTGGCAAAGGATGGCTGCACCGTCGTGGCGAACTGCCATCCGTCCGAGGCGGCC
I35762          CAGCGGCTGGCCAAGGATGGCTTTCGTGTGGTGGCCGGTTGCGGCCCCAACTCGCCG--C
J04987          CAGCGGCTGGCCAAGGATGGCTTTCGTGTGGTGGCCGGTTGCGGCCCCAACTCGCCG--C
D49362          AAGGCGCTGAAGGAGGCCGGCTATACGGTCGCCGCGAATTACG--CCGGCAATGACG---
U17226          GTGGCGCTGAAGGCTGCGGGCTACAAGGTGGCCGCAAACTATG--CCGGAAATGACG---
                      ***      *  ****     ** *  **    *     **       * **

A27012          GCCGCCGAAGAGTGGAAGCAGGCCCGTGCCGCCGAGGGGTTCGACATCGCCGTCTTCACC
L01112          GCCGCCGAAGAGTGGAAGCAGGCCCGTGCCGCCGAGGGGTTCGACATCGCCGTCTTCACC
I35762          GCCGC-GAAAAGTGGCTGGAGCAGCAGAAGGCCCTGGGCTTCGATTTCATTGCCTCGGAA
J04987          GCCGC-GAAAAGTGGCTGGAGCAGCAGAAGGCCCTGGGCTTCGATTTCATTGCCTCGGAA
D49362          --------ATGC-CGCCCGCGCCTT----TACCGAGGAAACCGGCATCAAGACCTACAAA
U17226          --------AGAG-GGCCAAGGCCTT----CGAGCAGGAAAGCGGCATTCCCGTCTACAAA
                        *     *     *              **    **   *      **

A27012          GCTGACGTGTCCTCGTTCGACGACAGCGCGCGCATGGTTCGCGAGATCACAGAGCAGGTC
L01112          GCTGACGTGTCCTCGTTCGACGACAGCGCGCGCATGGTTCGCGAGATCACAGAGCAGGTC
I35762          GGCAATGTGGCTGACTGGGACTCGACCAAGACCGCATTCGACAAGGTCAAGTCCGAGGTC
J04987          GGCAATGTGGCTGACTGGGACTCGACCAAGACCGCATTCGACAAGGTCAAGTCCGAGGTC
D49362          TGGTCGGTGGCGGATTACGACGCCTGCGCGGCGGGCATCAAGCAGGTGGAGGAGGAGCTT
U17226          TGGGACGTATCGAGCTATCAGGCCTGCGTCGATGGCATCGCCAGGGTCGAGGCCGACCTC
                      **  *    *   *      *          *      * *        *  *

A27012          GGTCCCATCGACATCCTGGTCAACTGTGCCGGCATCACCCGCGACAAGACCTTCAAGAAG
L01112          GGTCCCATCGACATCCTGGTCAACTGTGCCGGCATCACCCGCGACAAGACCTTCAAGAAG
I35762          GGCGAGGTTGATGTGCTGATCAACAACGCCGGTATCACCCGCGACGTGGTGTTCCGCAAG
J04987          GGCGAGGTTGATGTGCTGATCAACAACGCCGGTATCACCCGCGACGTGGTGTTCCGCAAG
D49362          GGCCCCATCGCCGTGCTGGTCAACAATGCCGGCATCACCCGCGACGCGATGTTCCACAAG
U17226          GGACCGGTCGACATCCTCGTCAACAATGCCGGCATCACCCGTGACGCCATGTTCCACAAG
                **     * *   * **  *****   ***** ******** ***      ***   ***
                           5'    ATCAACAACGCCGGTATCAC    3'
 Primer #4035                    G     TGT     C

A27012          ATGGAGCAGGCGCACTGGGAGGCCGTGATCAACGTCAACCTCAACAGCGTCTTCAACGTC
L01112          ATGGAGCAGGCGCACTGGGAGGCCGTGATCAACGTCAACCTCAACAGCGTCTTCAACGTC
I35762          ATGACCCGCGCCGACTGGGATGCGGTGATCGACACCAACCTGACCTCGCTGTTCAACGTC
J04987          ATGACCCGCGCCGACTGGGATGCGGTGATCGACACCAACCTGACCTCGCTGTTCAACGTC
D49362          ATGACCCCCCAGCAGTGGAAAGAGGTCATCGACACCAACCTGACGGGCCTGTTCAACATG
U17226          ATGACGCCGGAACAGTGGGGCGAAGTGATCGGCACCAATCTCACCGGCGTCTTCAACATG
                ***   *      * ***   *  ** ***  *  *** ** *      * ****** *

A27012          ACCCGTCAGGTGTGGGACGGGATGCTGGAGCGCGGCTTCGGGCGTATCATCAACATCTCG
L01112          ACCCGTCAGGTGTGGGACGGGATGCTGGAGCGCGGCTTCGGGCGTATCATCAACATCTCG
I35762          ACCAAGCAGGTGATCGACGGCATGGCCGACCGTGGCTGGGGCCGCATCGTCAACATCTCG
J04987          ACCAAGCAGGTGATCGACGGCATGGCCGACCGTGGCTGGGGCCGCATCGTCAACATCTCG
D49362          ACCCATCCGGTCTGGTCGGGCATGCGCGACCGCAAATACGGCCGTATCGTCAACATCAGC
U17226          ACGCATCCGCTGTGGTCGGGCATGCGCGACCGCGGCTTCGGCCGTATCGTCAACATCTCG
                **    * * *       ** ***   ** **    *  ** ** *** ********

A27012          TCGGTCAACGGTCAGCGCGGCCAGTTCGGTCAGGCCAACTATTCCGCCGCCAAGGCCGGT
L01112          TCGGTCAACGGTCAGCGCGGCCAGTTCGGTCAGGCCAACTATTCCGCCGCCAAGGCCGGT
I35762          TCGGTGAACGGGCAGAAGGGCCAGTTCGGCCAGACCAACTACTCCACCGCCAAGGCCGGC
J04987          TCGGTGAACGGGCAGAAGGGCCAGTTCGGCCAGACCAACTACTCCACCGCCAAGGCCGGC
D49362          TCGATCAACGGCCAGAAGGGCCAGGCCGGCCAGGCGAACTATTCGGCGGCCAAGGCAGGC
U17226          TCGATCAACGGCCAGAAGGGGCAGATGGGCCAGGTGAACTATTCCGCCGCCAAGGCCGGC
                *** * ***** ***   ** ***   ** ***   ***** **  * ******** **
```

```
A27012          ATGCACGGCTTCACCATGGCTCTGGCTCAGGAGGGTGCGTCCAAGGGCGTGACCGTCAAC
L01112          ATGCACGGCTTCACCATGGCTCTGGCTCAGGAGGGTGCGTCCAAGGGCGTGACCGTCAAC
I35762          CTGCATGGCTTCACCATGGCACTGGCGCAGGAAGTGGCGACCAAGGGCGTGACCGTCAAC
J04987          CTGCATGGCTTCACCATGGCACTGGCGCAGGAAGTGGCGACCAAGGGCGTGACCGTCAAC
D49362          GACCTGGGCTTTACCAAGGCGCTGGCCCAGGAAGGCGCCCGTGCCGGCATCACCGTGAAC
U17226          GATCTCGGCTTGACCAAGGCGCTGGCCCAGGAAGGGGCGGCGAAAGGGATCACCGTCAAC
                   *  ***** **** *** ***** ***** *  **        **  * ***** ***

A27012          ACCATCTCGCCCGGCTATGTCGAGACGGCCATGACCCTGGCGATGAACGACGATGTGCGC
L01112          ACCATCTCGCCCGGCTATGTCGAGACGGCCATGACCCTGGCGATGAACGACGATGTGCGC
I35762          ACGGTCTCTCCGGGCTATATCGCCACCGACATGGTCAAGGCGATCCGCCAGGACGTGCTC
J04987          ACGGTCTCTCCGGGCTATATCGCCACCGACATGGTCAAGGCGATCCGCCAGGACGTGCTC
D49362          GCGATCTGCCCCGGCTATATCGGCACCGAGATGGTCCGCGCCATCGACGAGAAGGTGCTG
U17226          GCGATTTGCCCCGGCTATATCGGCACCGAGATGGTGCGCGCCGTTCCGGAAAAGGTGCTC
                 *  * *  ** ****** ***  ** *  ***       **  *     *  * ****
                    5'       CCGATATAGCGGTGGCTGTA       3'
Primer #4036                       C   CC  C GC
                                   T

A27012          AACAGC---ATCATCAGCGGTATTCCGATGCGTCGCATGGCTC-AGCCTAATGAGATCGC
L01112          AACAGC---ATCATCAGCGGTATTCCGATGCGTCGCATGGCTC-AGCCTAATGAGATCGC
I35762          GACAAG---ATCGTCGCGACGATCCCGGTCAAGCGCCTGGGCC-TGCCGGAAGAGATCGC
J04987          GACAAG---ATCGTCGCGACGATCCCGGTCAAGCGCCTGGGCC-TGCCGGAAGAGATCGC
D49362          AACGAGGGCATCATCCCCCAGATCCCGGTCGC-CGCCTGGGCGGAGCCCGAGGAAATCGC
U17226          AACGAGCGGATCATTCCCCAGATACCCGTCGGACGCCTCGGCG-AGCCGGAGGAAGTGGC
                 **      *** *       ** **  *    *** * *     ***  * **  * **

A27012          CGCCGCCATCGCTTTCCTGGCCGGCGACGAGAGCGGTTATATGACGGGCGCCAATCTGC-
L01112          CGCCGCCATCGCTTTCCTGGCCGGCGACGAGAGCGGTTATATGACGGGCGCCAATCTGC-
I35762          CTCGATCTGCGCCTGGTTGTCGTCGGAGGAGTCCGGTTTCTCGACCGGCGCCGACTTCT-
J04987          CTCGATCTGCGCCTGGTTGTCGTCGGAGGAGTCCGGTTTCTCGACCGGCGCCGACTTCT-
D49362          GCGCTGCGTGGTCTTCCTGGCTTCCGAGGATGCCGGTTTCATCACCGGCTCGACCCATCA
U17226          ACGCTGCGTCGTGTTTCTCGCTTCCGACGACGCGGGCTTCATCACCGGCTCGACGATTT-
                      *   *  *   *  *    ** **    ** *     ** *** *

A27012          --CGGTCAACGGCGGTCTGTTCATGCATTGA
L01112          --CGGTCAACGGCGGTCTGTTCATGCATTGA
I35762          --CGCTCAACGGCGGCCTGCATATGGGCTGA
J04987          --CGCTCAACGGCGGCCTGCATATGGGCTGA
D49362          AGCGCCAAATGGCGGCCAGTTCTTCGTCTGA
U17226          --CGGCCAATGGCGGCCAGTACTTCGCCTGA
                  **   ** ***** * *    *    ***
```

C: Multiple alignment of DNA sequence data for published *phaC* genes

Table C. Description of Genbank sequence entries used in the multiple alignment of *phaC* genes.

Accession No.	Source	Definition	Reference
D43764	*Paracoccus denitrificans*	*phaC* gene for poly (3-hydroxyalkanoate) synthase	Ueda *et al.* (1996)
I35763	Unknown		Patent: US 5602321-A 25 11-FEB-1997
J05003	*R. eutropha* (strain H16)	poly-beta-hydroxybuterate-C (*phbC*) gene	Peoples and Sinskey (1989c)
L17049	*Rhodobacter sphaeroides*	polyhydroxyalkanoate synthase (*phaC*) gene	Hustede and Steinbuechel (1993)
L37761	*Acinetobacter* sp. (strain RA3849).	*phaA*, *phaB*, and *phaC* genes	Schembri *et al.* (1995)
M58445	*Pseudomonas oleovoran*	*phaC* gene	Huisman *et al.* (1991)
U04848	*Acinetobacter* sp. (strain RA3849)	PHA synthase (*phaC*) gene	Schembri *et al.* (1994)
U17227	*Sinorhizobium meliloti* (strain 41)	poly-beta-hydroxybutyrate synthase (*phbC*) gene	Tombolini *et al.* (1995)
U30612	*Rhizobium etli* (strain CE 3)	poly-beta-hydroxybutyrate synthase (*phaC*) gene	Cevallos *et al.* (1996)
U66242	*Zoogloea ramigera* (strain 115)	PHB polymerase (*phbC*) gene	Direct Submission
U78047	*Alcaligenes* sp. (strain SH-69)	PHA synthase (*phaC*) gene	Direct Submission
X66407	*Rhodococcus ruber* (strain NCIMB40126)	PHA-synthase (*phaC*) gene	Pieper and Steinbuchel (1992)
X66592	*Pseudomonas aeruginosa* (strain PAO1)	PHA-synthase1 (*phaC1*) and PHA-synthase2 (*phaC2*) genes	Timm and Steinbuchel (1992)
X97200	*Rhodobacter sphaeroides* (strain RV)	*phaC* gene	Direct Submission

```
Genbank accession
number:

U17227          ------------------------------------------------------------
U30612          ATGTACAACAAACGGATAAAAAGAGTGCTGCCGCCGGAGGAAATGGTGACCGACAGCAAG
J05003          ------------------------------------------------------------
I35763          ------------------------------------------------------------
U66242          ------------------------------------------------------------
U78047          ------------------------------------------------------------
L17049          ------------------------------------------------------------
X97200          ------------------------------------------------------------
D43764          ----------------------ATGGCCGCGCCCCCGCGCCCAGGCTGCCGCCCCGGCGGG
M58445a         ------------------------------------------------------------
X66592a         ------------------------------------------------------------
M58445b         ------------------------------------------------------------
X66592b         ------------------------------------------------------------
X66407          ------------------------------------------------------------
U04848          ------------------------------------------------------------
L37761          ------------------------------------------------------------

U17227          ------------------------------------------------------------
U30612          CAGGAGAGTGGCGGCCAGAAAAATGGCGACAAGACCGGTTTCGACGCGACCGATCTCAAA
J05003          ------------------------------------------------------------
I35763          ------------------------------------------------------------
U66242          ------------------------------------------------------------
U78047          ------------------------------------------------------------
L17049          ----------------------------------------------------ATGGCAACC
X97200          ----------------------------------------------------ATGGCAACC
D43764          GACCGGCCAGTCTGCCGGACTCGCTGCCGAGCCGCATCCAGCCCGAACACCGCCGCCTTC
M58445a         ------------------------------------------------------------
X66592a         ------------------------------------------------------------
M58445b         ------------------------------------------------------------
X66592b         ------------------------------------------------------------
X66407          ------------------------------------------------------------
U04848          ------------------------------------------------------------
L37761          ------------------------------------------------------------

U17227          ------------------------------------ATGGCTCGCGCGGCAGAGCAG
U30612          CCCTATCTGTTGAAGGATCCCGAGACCATGGCGATGAATTTCGCCCGGGCGCTCGAAAAT
J05003          ---------------ATGGCGACCGGCAAAGGC-GCGGCAGCTTCCACGCAGGAAGGCA
I35763          ---------------ATGGCGACCGGCAAAGGC-GCGGCAGCTTCCACGCAGGAAGGCA
U66242          ----------------------------------ATGAATTTGCCCGATCCGCAAGCC
U78047          ------------------------------------------------------------
L17049          GAAGAGCAGTCTCCGGGTTCCGGCCGTGACGCTCAGTTCGAGCGTCTGAACGCGAATCTC
X97200          GAAGAGCAGTCTCCGGGTTCCGGCCGTGACGCTCAGTTCGAGCGTCTGAACGCGAATCTC
D43764          GTCGAGGCGGCCTTCGGTCCCGGCAGC--CGCC---TCCCAACAGCTGGCCCAGAACATC
M58445a         ------------------------------------------------------------
X66592a         ------------------------------------------------------------
M58445b         ------------------------------------------------------------
X66592b         ------------------------------------------------------------
X66407          ------------------------------------------------------------
U04848          ----------------------ATGAACCCGAACTCATTTCAATTCAAAGAAAACATACT
L37761          ----------------------ATGAACCCGAACTCATTTCAATTCAAAGAAAACATACT

U17227          CTCGGAAAAGCCGCTTCCGCATGGCTCGCCCCCCGCGAAGCGGGCGAGAAGACGGATAGT
U30612          CTCGGCCAGGCCGCCTCGGCCTGGCTTGCGCCGCGCGAACGCGGCGAGATCACCGAAACG
J05003          AGTCCCAACCATTCAAGGTCACGCCGGGGCCATTCGATCCAGCCACATGGCTGGAA-TGG
I35763          AGTCCCAACCATTCAAGGTCACGCCGGGGCCATTCGATCCAGCCACATGGCTGGAA-TGG
U66242          ATTGCCAACGCCTGGATGTCCCAGGTGGGCGACCC----CAGCCA-ATGGCAATCC-TGG
U78047          ------ATGACTTCTGAATCATCCTGGGCCGAGAGCGCCCAGCAGTTCCAGCAGAT-TTT
L17049          ACCCGCATCGACGAGCTGTCGAAACGGCTGACGGCCGCTCTCACGAAGCGCAAACTGTCG
X97200          ACCCGCATCGACGAGCTGTCGAAACGGCTGACGGCCGCTCTCACGAAGCGCAAACTGTCC
D43764          GAGCGCATCGAATCGCTGACCCAGCGCCTGATCAGCGCGCTGGCGCAGCGCCGTCCCTCG
M58445a         ----------------------------ATGAGTAACAAGAACAACGATGAGCTGCAGCGG
X66592a         ----------------------------ATGAGTCAGAAGAACAATAACGAGCTTCCCAAG
M58445b         ----------------------------ATGAAAGACAAACCGGCCAAAGGAACGCCAACG
X66592b         ----------------------------ATGCGAGAAAAGCAGGAATCGGGTAGCGTGCCG
X66407          -------------------------------TTGCTCGACCACGTGCACAAGAAGTTGAAG
U04848          ACAATTTTTTTCTGTACATGATGACATCTGGAAAA-AATTACAAGAATTTTATTATGGGC
L37761          ACAATTTTTTTCTGTACATGATGACATCTGGAAAA-AATTACAAGAATTTTATTATGGGC
```

```
U17227          TTCGCCGAGCCGGTCTCCGACATGGTCAAGACCCTCTCCAAGGTCTCGGAATACTGGCTC
U30612          GCCATCGATCCGATGACCGACATGGTCAAGACGCTTTCCAAGATCAGCGAATACTGGATT
J05003          TCCCGCCAGTGGCAGGGCACTGAAGGCAACGGCCACGCGGCCGCGTCCGGCATTCCGGGC
I35763          TCCCGCCAGTGGCAGGGCACTGAAGGCAACGGCCACGCGGCCGCGTCCGGCATTCCGGGC
U66242          TTCAGCAAGGCGCCCACCACCGAGGCGAACCCGATGGCCACCATGTTGCAGGATATCGGC
U78047          CGGGCAAAGCTGGTCGCAGGCGTTGCAATCGTTTCAGAAACTGGACCTCGGGGCGCAGGT
L17049          GACCCCGCGCTGCACGGGCCCTCGGGCGACGTCTTCCTGAAGGCGATGACGGCCTACATG
X97200          GACCCCGCGTTGCACGGGCCCTCGAGCGACGTCTTCCTGAAGGCGATGACGGCCTACATG
D43764          AATCCCGGCGTCGAGATGCCGGGCCCCGACCTTTTCGCCACCGCGACCTCGGCCTGGATC
M58445a         CAGGCCTCGGAAAACACCCTGGGGCTGAACCCGGTCATCGGTATCCGCCGCAAAGACCTG
X66592a         CAAGCCGCGGAAAACACGCTGAACCTGAATCCGGTGATCGGCATCCGGGGCAAGGACCTG
M58445b         CTTCCCGCCACCAGCATGAACGTGCAGAACGCCATCCTCGGCCTGCGCGGTCGTGACCTG
X66592b         GTGCCCGCCGAGTTCATGAGTGCACAGAGCGCCATCGTCGGCCTGCGCGGCAAGGACCTG
X66407          TCGACCCTGGA--CC-CGATCGGCTGGGGTCCCGCAGT-GACGTCGGTGGCCG-GACGCG
U04848          AAAGCCCAATTAATGAGGCTTTGGCGCAGCTCAACAAAGAAGATATGTCTTTGTTCTTTG
L37761          AAAGCCCAATTAATGAGGCTTTGGCGCAGCTCAACAAAGAAGATATGTCTTTGTTCTTTG

U17227          TCCGACCCCCGGCGGACACTCGAAGCCCAGACCCATCTTCTCGGCAGCTTCTTCGATATG
U30612          TCCGATCCCCGCCGCACCTTCGAGGCGCAGACTCAGCTGATGTCGTCCTTCTTCGGCATC
J05003          CTGGATGCGCTGGCAGGCGTCAAGATCGCGCCGGCGCAGCTGGGTGATATCCAGCAGCGC
I35763          CTGGATGCGCTGGCAGGCGTCAAGATCGCGCCGGCGCAGCTGGGTGATATCCAGCAGCGC
U66242          ---GTTGCGCT--CAAACCGGAAG----CGATGGAGCAGCTGAAAAACGATTATCTGCGT
U78047          --ACCTGCCGCAGCCCCCCTGAAGCTGTCGCAGACCAAACTGCAGGCGTTGCAACAGCAA
L17049          GCCGAGATGATGCAGAACCCGGCCAAGATCCTCGAGCATCAGATCAGTTTCTGGGGCAAG
X97200          GCCGAGATGATGCAGAACCCGGCCAAGATCCTCGAGCATCAGATCAGTTTCTGGGGCAAG
D43764          AAGCTTCTGGCCGAGCAGCCCGAGCGGGTGATCGGCCAGCAGGTCAGCTATTGGGGCGAA
M58445a         TTGAGCTCGGCACGCACCGTGCTGCGCCAGGCCGTGCGCCAACCGCTGCACAGCGCCAAG
X66592a         CTCACCTCCGCGCGCATGGTCCTGCTCCAGGCGGTGCGCCAGCCGCTGCACAGCGCCAGG
M58445b         ATTTCCACGCTGCGCAATGTCAGCCGCCAAAGCCTGCGTCACCCGCTGCACACCGCACAT
X66592b         CTGACGACGGTCCGCAGCCTGGCTGTCCACGGCCTGCGCCAGCCGCTGCACAGTGCGCGG
X66407          CGGTCCGCAACCCTCAGGCCGTCACCGCCGCGACGGCGGAGTACGCGGGCCGGCTGGCGA
U04848          A-AGCACTATCTAAAAACCCAGCTCGCATGATGGAAATGCAATGGAGCTGGTGGCAAGGT
L37761          A-AGCACTATCTAAAAACCCAGCTCGCATGATGGAAATGCAATGGAGCTGGTGGCAAGGT

U17227          TGGTCGCGGACACTCCAGCGCATGGCAGCC--------GACGCCGTGGAGGACCCGGCCA
U30612          TGGATGCGCTCGATGCAGCGCATGCAGGGCACGCGTGGGATGCAGGGCGAGCCCCTGCCG
J05003          TACATG-AAGGACTTCTCAGCGCTGTGGCAGGCCAT--GGCCGAGGGCAAGGCCGAG--G
I35763          TACATG-AAGGACTTCTCAGCGCTGTGGCAGGCCAT--GGCCGAGGGCAAGGCCGAG--G
U66242          ----------GACTTCACCGCGTTGTGGCAGGATTT--TTTGGCTGGCAAGGC------G
U78047          TACCTC-AAGGAAGCCCAGGAACTGTGGGCGCAG-------GGCCTGCAGGGC--AC--C
L17049          AGCCTG-AAACATTACGTCGAGGCTCAGCACCAGCT--GGTGAAGGGCGAGCTGAAGCCG
X97200          AGCCTG-AAACATTACGTTGAGGCGCAGCATCAGCT--GGTGAAGGGCGAGCTGAAGCCG
D43764          ACCTTG-CGCCATTTCGCCGAGGCCCAGGCCGCCTT--TGCCCGCGGCACCGTGACGCCG
M58445a         CATGTG-GCCCACTTTGGCCTGGAGCTGAAGAACGT--GCTGCTGGGCAAGTCCAGCCTT
X66592a         CACGTG-GCGCATTTCAGCCTGGAGCTGAAGAACGT--CCTGCTCGGCCAGTCGGAGCTA
M58445b         CACCTG-TTGGCCCTGGGTGGCCAGCTGGGCCGGGT--GATACTGGGTGACACACCGCTT
X66592b         CACCTG-GTCGCCTTCGGAGGCCAGTTGGGCAAGGT--GCTGCTGGGCGACACCCTGCAC
X66407          -AGATC-CCGGCCGCGGCCACCCGCGTGTTCAACGC--GAACGATCCCGACGCACCGATG
U04848          CAAATACAAATCTACCAAAATGTGTTGATGCGCAGCGTGGCCAAAGATGTAGCACCATTT
L37761          CAAATACAAATCTACCAAAATGTGTTGATGCGCAGCGTGGCCAAAGATGTAGCACCATTT

U17227          ACC---TTCAGCAC----AACGACAAGCGCTTCGCCGACGAAGACTGGGTGAAGAACCCG
U30612          CCCGAGCCCGACACCCGCAAGGACAAGCGCTTTTCGGATGAGGATTGGCAGAAAAATCCG
J05003          CCACCGGTCCGC----TGCACGACCGGCGCTTCGCCGGCGACGCATGGCGCACCAACCTC
I35763          CCACCGGTCCGC----TGCACGACCGGCGCTTCGCCGGCGACGCATGGCGCACCAACCTC
U66242          CCAGCCGTCCA------G--CGACCG-CGCTTCAGCTCGGCAGCCTGGCAGGGCAATCCG
U78047          CCGGAGGTCAA----------GGACAAGCGTTTTGCCGGCGAAGGCTGGGCCAGCAACCCC
L17049          CCGCCGGACGTGACGCCGAAGGACCGCCGCTTCTCGAACCCGCTCTGGCAGACGCATCCC
X97200          CCGCCGGACGTGACGCCGAAGGACCGCCGCTTCTCGAACCCGCTCTGGCAGACGCATCCG
D43764          CCGCCCAGCGAAGGGCCGCGGGACCGGCGCTTTGCCAACCCGCTGTGGGAGGCGCATCCC
M58445a         GCCCCGGAAAG------CGACGACCGTCGCTTCAATGACCCGGCATGGAGCAACAACCCA
X66592a         CGCCCAGGCGA------TGACGACCGACGCTTTTCCGATCCGGCCTGGAGCCAGAATCCA
M58445b         CAGCCGAACCC------GCGCGATCCGCGCTTCAGCGACCCGACATGGAGCCAGAACCCG
X66592b         CAGCCGAACCC------ACAGGACGCCCGCTTCCAGGATCCATCCTGGCGCCTCAATCCC
X66407          CCGGTCGACCC------GCGGGACCGCCGCTTCTCCGACACCGCGTGGCAGGAGAACCCC
U04848          ATTCAGCCTGAAAG---TGGTGATCGTCGTTTTAACAGCCCATTATGGCAAGAACACCCA
L37761          ATTCAGCCTGAAAG---TGGTGATCGTCGTTTTAACAGCCCATTATGGCAAGAACACCCA
                                **    ** **             ***       * *
```

```
U17227          TTTTTCGATTTCATCCGCCAGGCCTACTTCGTCACCTCCGACTGGGCGGAGCGCATGGTG
U30612          TTCTTCGATTTCCTCCGCCAGGTCTATTTCGTCACGAGTGACTGGGTGGACAAGCTGGTG
J05003          CCATATCGCTTCGCTGCCGCGTTCTACCTGCTCAATGCGCGCGCCTTGACCGAGCTGGCC
I35763          CCATATCGCTTCGCTGCCGCGTTCTACCTGCTCAATGCGCGCGCCTTGACCGAGCTGGCC
U66242          ATGTCGGCCTTCAATGCCGCATCTTACCTGCTCAACGCCAAATTCCTCAGTGCCATGGTG
U78047          GTGGCGGCGTTTTCTGCGGCTGCCTACTTGCTCAACGCCCGAACGCTGATGGGCCTGGCC
L17049          TTCTTCAACTATCTCAAGCAGCAGTATCTGATGAACGCCGAGGCGGTGAATCAGGCCGTC
X97200          TTCTTCAACTATCTCAAGCAGCAGTATCTGATGAACGCCGAGGCGGTTAATCAGGCCGTC
D43764          TTCTTCAACTTCATCAAGCGGCAATACCAGATCAACGCCCAGGCCCTGCAGGAGGCGGCC
M58445a         CTTTACCGCCGCTACCTGCAAACCTATCTGGCCTGGCGCAAGGAGCTGCAGGACTGGATC
X66592a         CTGTACAAGCGCTACATGCAGACCTACCTGGCCTGGCGCAAGGAGCTGCACAGCTGGATC
M58445b         TTCTACCGGCGCGGCCTGCAAGCCTACCTGGCCTGGCAGAAGCAGACCCGCCTGTGGATC
X66592b         TTCTACCGGCGCACCCTGCAGGCCTACCTGGCGTGGCAGAAACAACTGCTCGCCTGGATC
X66407          GCGTACTTCTCGCTCCTGCAGAGCTACCTGGCCACGCGGGCCTACGTCGAGGAACTCACC
U04848          AATTTTGACTTGTTGTCACAGTCTTATTTACTGTTTAGCCAGTTAGTGCAAAACATGGTA
L37761          AATTTTGACTTGTTGTCACAGTCTTATTTACTGTTTAGCCAGTTAGTGCAAAACATGGTA
                                        **

U17227          AAGGACGCCGAGGGCCTTGACGATCATACCCGCCACAAGGCGGCCTTCTACGTTCGCCAG
U30612          TCGGAGACCGACGGCCTCGACGAGCACACCAAGCACAAGGCGGGATTCTACGTGAAGCAG
J05003          GATGCCGTCGAG---GCCGATGCCAAGACCCGCCAGCGCATCCGCTTCGCGATCTCGCAA
I35763          GATGCCGTCGAG---GCCGATGCCAAGACCCGCCAGCGCATCCGCTTCGCGATCTCGCAA
U66242          GAGGCGGTGGAC---ACCGCACCCCAGCAAAAGCAGAAAATACGCTTTGCCGTGCAGCAG
U78047          GAGGCGGTGGAG---GCCGATGAAAAAACCAAGGCCCGCATTCGCTTTGGCGTGGAGCAG
L17049          GAGGGGCTGGAGCATATCGAGCCGTCCGACAAGAAGCGGGTCGAATATTTCTCGCGCCAG
X97200          GAGGCGCTGGAGCATATCGAGCCTTCCGACAAGAAGCGGGTCGAGTATTTCTCGCGCCAG
D43764          AGCACACTGGACCTGCCCGAGATGACCGACCGGCGCCGGATCGAATGGTTCACCCGCCAG
M58445a         GGCAACAGCGAC---CTGTCGCCCCAGGACATCAGCCGCGGCCAGTTCGTCATCAACCTG
X66592a         AGCCACAGCGAC---CTGTCGCCGCAGGACATCAGTCGTGGCCAGTTCGTCATCAACCTG
M58445b         GAGGAAAGCCAC---CTGGACGACGATGACCGGGCCCGTGCGCACTTCCTGTTCAACCTG
X66592b         GACGAAAGCAAC---CTGGACTGCGACGATCGCGCCCGCGCCCGCTTCCTCGTCGCCTTG
X66407          GAGGCCGGCTCC---GGCGATCCGCTGCAGGACGGTAAGGCACGCCAGTTCGCGAACCTG
U04848          GATGTGGTCGAAGGTGTTCCAGACAAAGTTCGCTATCGTATTCACTTCTTTACCCGCCAA
L37761          GATGTGGTCGAAGGTGTTCCAGACAAAGTTCGCTATCGTATTCACTTCTTTACCCGCCAA

U17227          ATTGCCAGCGCTCTTTCCCCGACCAACTTCATCACGACGAATCCGCAGCT--CTATCGCG
U30612          ATCACGGCAGCCCTTTCGCCGAGCAACTTCATCGCTACCAACCCACAGCT--TTATCGCG
J05003          TGGGTCGATGCGATGTCGCCCGCCAACTTCCTTGCCACCAATCCCGAGGC--GCAGCGCC
I35763          TGGGTCGATGCGATGTCGCCCGCCAACTTCCTTGCCACCAATCCCGAGGC--GCAGCGCC
U66242          GTGATTGATGCCATGTCGCCCGCGAACTTCCTCGCCACCAACCCGGAAGC--GCAGCAAA
U78047          TGGATGGCCGCCATGGCGCCCAGCAACTTCCTCGCTTTCAACGCCGAGGC--GCAGAAGA
L17049          ATCGTCGATCTTTTCTCGCCCACGAACTTCTTCGGCACCAATCCCGACGC--GCTCGAAC
X97200          ATCGTCGATCTTTTCTCGCCCACCAACTTCTTCGGCACCAATCCCGACGC--GCTCGAGC
D43764          ATGATCGACATGATGGCGCCGACGAATTTTCTGGCCACCAATCCCGACGACAGCTGGAAA
M58445a         ATGACCGAAGCCATGGCTCCGACCAACACCCT-G--TCCAACCCGGCAGC--AGTCAAAC
X66592a         CTGACCGAGGCGATGTCGCCGACCAACAGCCT-G--AGCAACCCGGCGGC--GGTCAAGC
M58445b         ATCAACGATGCCCTGGCGCCAAGCAACTCGCT-G--CTCAACCCGCTGGC--GGTCAAGG
X66592b         CTCTCCGACGCCGTGGCACCCAGCAACAGCCT-G--ATCAATCCACTGGC--GTTAAAGG
X66407          ATGTTCGACGCGTTGGCGCCGTCGAACTTCCT-G--TGGAATCCCGGTGT--GCTCACCC
U04848          ATGATCAATGCGTTATCTCCAAGTAACTTTCTGTGGACTAACCCAGAAGT--GATTCAGC
L37761          ATGATCAATGCGTTATCTCCAAGTAACTTTCTGTGGACTAACCCAGAAGT--GATTCAGC
                         *  * **    **     *       **  *

U17227          AGACCGTGGCGTCGAGCGG-CGCCAATCTCGTGAAAGGCATGCAGATG------CTGGCG
U30612          AGACCATCGCGAGCAACGG-CGAAAACCTGGTGCGCGGCATGAAAATG------CTCGCC
J05003          TGCTGATCGAGTCGGGCGG-CGAA---TCGCTGCGTGCC--------------GGCGTG
I35763          TGCTGATCGAGTCGGGCGG-CGAA---TCGCTGCGTGCC--------------GGCGTG
U66242          AACTGATTGAAACCAAGGG-CGAGAGCCTGACGCGTGGCCTGGTCAATATGCTGGGCGAT
U78047          AAGCCATCGAGACCAAAGG-CGAGA--GTATTGCCAAG----------------GGCATG
L17049          GCGCCATCGCCACCGACGG-CGAGAGCCTGGTGCAGGGGCTGGAGAAT------CTCGTG
X97200          GCGCCATTGCCACCGATGG-CGAGAGCCTCGTGCAGGGGCTGGAGAAT------CTCGTG
D43764          AGGCGCTGGAGACCGAGGGACGAAAGCCTGGTCAGGGGCCTTGAGAAC------CTGGTG
M58445a         GCTTCTTCGAAACCGGCGG-CAAGAGCCTGCTCGATGGCCTGTCC---------AACCTG
X66592a         GCTTCTTCGAGACCGGCGG-CAAGAGCCTGCTGGACGGCCTCGGC---------CACCTG
M58445b         AACTGTTCAACAGCGGTGG-CCAGAGCCTGGTGCGCGGCGTGGCC---------CACCTG
X66592b         AACTGTTCAATACCGGCGG-GATCAGCCTGCTCAATGGCGTCCGC---------CACCTG
X66407          GCGCCTTCGAGACCGGCGG-GGCGAGCCTGCTGCGCGGCGCCCGC---------TACGCC
U04848          AAACTGTAGCTGAACAAGG-TGAAAACTTAGTCCGTGGCATGCAAGTT------TTCCAT
L37761          AAACTGTAGCTGAACAAGG-TGAAAACTTAGTCCGTGGCATGCAAGTT------TTCCAT
                      *          **
```

```
U17227          GAAGACATAGCCGCCGGGCGCGGCGAGCTTCGGCTCCGCCAGACGGACACCAGCAAGTTC
U30612          GAGGACATTGCTGCCGGAAAGGGCGAGCTTCGCCTTCGCCAGACCGACATGACGAAATTC
J05003          CGCAACATGATGGAAGACCTGACACGCGGCAAGATCTCGCAGACCGACGAGAGCGCGTTT
I35763          CGCAACATGATGGAAGACCTGACACGCGGCAAGATCTCGCAGACCGACGAGAGCGCGTTT
U66242          ATCAATATGCTGGGCGATATCAACAACGGCCATATCTCGCTGTCGGACGAATCGGCCTTT
U78047          CAGAACCTGCTGCACGACATCACGCAGGGCCATGTGTCGATGACCGACGAAAGCCTCTTC
L17049          CGCGACATCGAGGCCAACAACGGCGATCTGCTCGTCACGCTGGCCGACCCCGAGGCCTTT
X97200          CGCGACATCGAGGCCAACAAGGGCGATCTGCTCGTCACGCTGGCCGATCCCGAGGCCTTC
D43764          CGCGACGTCGAGCAGAACAGCGGCGAGCTGATCGTGTCGCTGGCCGACCGCGATGCCTTC
M58445a         GCCAAGGACCTGGTCAACAACGGTGGCATGCCCAGC---CAGGTGAACATGGACGCCTTC
X66592a         GCCAAGGACCTGGTGAACAACGGCGGGATGCCGAGC---CAGGTGGACATGGACGCCTTC
M58445b         CTCGATGACCTGCGCCACAATGACGGCCTGCCACGC---CAGGTCGACGAGCGCGCCTTC
X66592b         CTCGAAGACCTGGTGCACAACGGCGGCATGCCCAGC---CAGGTGAACAAGACCGCCTTC
X66407          GCGCACGACATCCTCAACCGCGGCGGCCTGCCGTTG---AAGGTGGATTCGGACGCGTTC
U04848          GATGATGTCATGAATAGCGGCAAGTATTTATCTATTCGCATGGTGAATAGCGACTCTTTC
L37761          GATGATGTCATGAATAGCGGCAAGTATTTATCTATTCGCATGGTGAATAGCGACTCTTTC
                    *                                       *    *          **

U17227          GCCATCGGAGAGAACATCGCGATCACTCCAGGCAAGGTGATCGCCCAGAACGATGTCTGC
U30612          GCCGTCGGGCGCGACATGGCGTTGACGCCGGGCAAGGTGATCGCCCAGAACGATATCTGC
J05003          GAGGTCGGCCGCAATGTCGCGGTGACCGAAGGCGCCGTGGTCTTCGAGAACGAGTACTTC
I35763          GAGGTCGGCCGCAATGTCGCGGTGACCGAAGGCGCCGTGGTCTTCGAGAACGAGTACTTC
U66242          GAAGTGGGCCGCAACCTGGCCATTACCCCGGGCACCGTGATTTACGAAAATCCGCTGTTC
U78047          GAGGTGGGCCGCAACGTGGCCACGACCGAGGGCGCGGTGGTGTACGAGAACGAACTGTTT
L17049          CAGGTGGGGCAGAACCTCGCCACCACCGAAGGGTCGGTCGTCTACCGCAACCGCATGTTC
X97200          CAGGTGGGCCAGAACCTCGCCACCACCGAAGGGTCGGTCGTCTACCGCAACCGCATGTTC
D43764          CGTGTGGGCGAGAACATCGGCACCACCGAGGGCACGGTGGTCGCGCGCACCAAGCTTTAC
M58445a         GAGGTGGGCAAGAACCTGGGCACCAGTGAAGGCGCCGTGGTGTACCGCAACGATGTGCTG
X66592a         GAGGTGGGCAAGAACCTGGCCACCACCGAGGGCGCCGTGGTGTTCCGCAACGACGTGCTG
M58445b         GAAGTGGGCGGCAACCTGGCCGCGACTGCCGGCGCCGTGGTGTTTCGCAACGAGCTGCTG
X66592b         GAGATCGGTCGCAACCTCGCCACCACGCAAGGCGCGGTGGTGTTCCGCAACGAGGTGCTG
X66407          ACCGTCGGCGAGAACCTCGCGGCCACACCCGGCAAGGTGGTCTTCCGCAACGACCTGATC
U04848          AGCTTGGGCAAAGATTTAGCTTACACCCCTGGTGCAGTCGTCTTTGAAAATGACATTTTC
L37761          AGCTTGGGCAAAGATTTAGCTTACACCCCTGGTGCAGTCGTCTTTGAAAATGACATTTTC
                    * **     *  * *     *     **    **  *       *

U17227          CAGGTGCTGCAATACGAGGCGAGCACCGAGACCGTGCTGAAGCGGCCGTTGCTCATCTGC
U30612          CAGATCATCCAGTACGAAGCCTCGACCGAGACGGTGCTGAAACGGCCATTGCTGATCTGC
J05003          CAGCTGTTGCAGTACAAGCCGCTGACCGACAAGGTGCACGCGCGCCCGCTGCTGATGGTG
I35763          CAGCTGTTGCAGTACAAGCCGCTGACCGACAAGGTGCACGCGCGCCCGCTGCTGATGGTG
U66242          CAGCTGATCCAGTACACGCCGACCACGCCGACGGTCAGCCAGCGCCCGCTGTTGATGGTG
U78047          CAGCTGCTCGAATACAAGCCCCTCACGGCCAAGGTGTACGAGCGGCCGTTTTTGCTGGTG
L17049          GAGCTGATCCAGTACAAGCCCACGACCGAGACGGTCCACGAGACGCCGCTGCTGATCTTT
X97200          GAGCTGATCCAGTACAAGCCCACGACCGAGACGGTCCACGAGACGCCGCTCCTGATCTTT
D43764          GAGCTGATCCAGTACAAGCCCACCACCGCGCAGGTGCATGAGATCCCGCTGGTGATCTTT
M58445a         GAGCTGATCCAGTACAAGCCCATCACCGAGCAGGTGCATGCCCGCCCGCTGCTGGTGGTG
X66592a         GAACTGATCCAGTACCGGCCGATCACCGAGTCGGTGCACGAACGCCCGCTGCTGGTGGTG
M58445b         GAACTGATCCAGTACAAGCCGATGAGCGAAAAGCAGCACGCCCGGCCACTGCTGGTGGTG
X66592b         GAGCTGATCCAGTACAAGCCGCTGGGCGAGCGCCAGTACGCCAAGCCCCTGCTGATCGTG
X66407          GAGCTGATCCAGTACGCGCCGCAGACCGAGCAGGTGCACGCGGTGCCGATCCTGGCCGCA
U04848          CAATTATTGCAATATGAAGCAACTACTGAAAATGTGTATCAAACCCCTATTCTAGTCGTA
L37761          CAATTATTGCAATATGAAGCAACTACTGAAAATGTGTATCAAACCCCTATTCTAGTCGTA
                 *  *  *  * **     *                        **  *  *

Primer (forward) 5' GTTGCAGTACAAGCCGCT 3' Lopez et al. (1997)
U17227          CCGCCCTGGATCAACAAATTCTACGTGCTGGACCTCAATCCGGAGAAGTCCTTCATCAAA
U30612          CCGCCCTGGATCAACAAGTTCTACATTCTCGACCTCAACCCGCAGAAATCCTTCATCAAA
J05003          CCGCCGTGCATCAACAAGTACTACATCCTGGACCTGCAGCCGGAGAGCTCGCTGGTGCGC
I35763          CCGCCGTGCATCAACAAGTACTACATCCTGGACCTGCAGCCGGAGAGCTCGCTGGTGCGC
U66242          CCGCCGTGCATCAACAAGTTCTACATCCTCGACCTGCAACCGGAAAATTCGCTGGTGCGC
U78047          CCGCCCTGCATCAACAAGTTCTACATCCTCGACCTGCAGCCCGAAAACTCGCTGATCCCG
L17049          CCGCCCTGGATCAACAAGTTCTACATCCTCGACCTCAAGCCGCAGAATTCCCTGCTGAAG
X97200          CCGCCCTGGATCAACAAGTTCTACATTCTCGATCTCAAGCCGCAGAACTCCCTGCTGAAG
D43764          CCGCCCTGGATCAACAAATTCTACATCCTCGACCTCAAGCCGCAGAACAGCCTGATCAAA
M58445a         CCGCCGCAGATCAACAAGTTCTACGTATTCGACCTGAGCCCGGAAAAGAGCCTGGCACGC
X66592a         CCGCCGCAGATCAACAAGTTCTACGTCTTCGACCTGTCGCCGGACAAGAGCCTGGCGCGC
M58445b         CCGCCACAGATCAACAAGTTCTACATCTTCGACCTCAGCTCGACCAACAGCTTCGTCCAG
X66592b         CCGCCGCAGATCAACAAGTACTACATCTTCGACCTGTCGCCGGAAAAGAGCTTCGTCCAG
X66407          CCGCCCTGGATCAACAAGTACTACATCCTCGACCTCGCGCCCGGCCGGAGCCTCGCCGAA
U04848          CCACCGTTTATCAATAAATATTATGTGCTGGATTTACGCGAACAAAACTCTTTAGTGAAC
L37761          CCACCGTTTATCAATAAATATTATGTGCTGGATTTACGCGAACAAAACTCTTTAGTGAAC
                ** **    ***** ** *  **  *  * **  *                    *
```

```
U17227          TGGGCTGTCGACCAGGGTCAGACGGTCTTCGTCATCTCCTGGGTAAACCCGGACGAACGC
U30612          TGGTGCGTCGACCAGGGGCAGACGGTCTTCGTCATTTCCTGGGTCAACCCGGATGGGCGC
J05003          CATGTGGTGGAGCAGGGACATACGGTGTTTCTGGTGTCGTGGCGCAATCCGGACGCCAGC
I35763          CATGTGGTGGAGCAGGGACATACGGTGTTTCTGGTGTCGTGGCGCAATCCGGACGCCAGC
U66242          TACGCGGTGGAGCAGGGCAACACCGTGTTCCTGATCTCGTGGAGCAATCCGGACAAGTCG
U78047          TACGCCGTCGAACAAGGCCACCGCACCTTTGAGGTGAGCTGGCGCAACCCCGACGACAGC
L17049          TGGCTGGTGGATCAGGGCTTCACGGTCTTCGTCGTCTCGTGGGTGAACCCCGACAAGAGC
X97200          TGGCTGGTGGATCAGGGCTTCACGGTCTTCGTCGTCTCCTGGGTGAACCCCGACAAGAGC
D43764          TGGATCGTGGACCAGGGCCATACGCTGTTCGTGGTGGCCTGGAAGAACCCCGACCCCAGC
M58445a         TACTGCCTGCGCTCGCAGCAGCAGACCTTCATCATCAGCTGGCGCAACCCGACCAAAGCC
X66592a         TTCTGCCTGCGCAACGGCGTGCAGACCTTCATCGTCAGTTGGCGCAACCCGACCAAGTCG
M58445b         TACATGCTCAAGAATGGCCTGCAGGTGTTCATGGTCAGCTGGCGCAACCCCGACCCGCGC
X66592b         TACGCCCTGAAGAACAACCTGCAGGTCTTCGTCATCAGTTGGCGCAACCCCGACGCCCAG
X66407          TGGGCGGTCCAGCACGGCCGAACCGTGTTCATGATCTCGTACCGGAACCCGGACGAGTCG
U04848          TGGTTGCGCCAGCAAGGTCATACAGTCTTTTTAATGTCATGGCGTAACCCAAATGCCGAA
L37761          TGGTTGCGCCAGCAAGGTCATACAGTCTTTTTAATGTCATGGCGTAACCCAAATGCCGAA
                                          **     *    *     ** **

U17227          CATGCCTCCAAGGACTGGGAAGCTTATGCACGCGAAGGCATAGGCTTCGCGCTTGATATC
U30612          CACGCCGAGAAGGACTGGGCCGCCTATGCCCGAGAGGGCATCGATTTCGCGCTGGAGACG
J05003          ATGGCCGGCAGCACCTGGGACGACTACATCGAGCACGCGGCCATCCGCGCCATCGAAGTC
I35763          ATGGCCGGCAGCACCTGGGACGACTACATCGAGCACGCGGCCATCCGCGCCATCGAAGTC
U66242          CTGGCCGGCACCACCTGGGACGACTACGTGGAGCAGGGCGTGATCGAAGCGATCCGCATC
U78047          CTGGCCCACAAGACCTGGGACGACTACGTGGAAGACGGCGCCATGGCTGCTATTGACGTG
L17049          TATGCCGGCATCGGCATGGACGACTACATCCGCGAAGGCTACATGCGCGCCATGGCCGAG
X97200          TATGCCGGCATCGGCATGGACGACTACATCCGCGATGGCTACATGCGCGCCATGGCCGAG
D43764          TATGGCGACACCGGCATGGACGATTACGTCAGCGC---CTATCTGGAGGTGATGGACCGG
M58445a         -CAGCGCGAATGGGGCCTGTCCAC--CTACATCGACGCGCTCAAGGAGGCGGTCGACGCG
X66592a         -CAGCGCGAATGGGGCCTGACCAC--CTATATCGAGGCGCTCAAGGAGGCCATCGAGGTA
M58445b         -CACCGCGAATGGGGCCTGTCCAG--CTACGTGCAGGCCCTGGAAGAAGCGCTCAACGCT
X66592b         -CACCGCGAATGGGGCCTGAGCAC--CTATGTCGAGGCCCTCGACCAGGCCATCGAGGTC
X66407          ATGCGGCACATCACGATGGACGACTACTACGTCGACGGCATCGCCACTGCGCTGGACGTG
U04848          CAGAAAGAATTGACTTTTGCCGATCTCATTACACAAGGTTCAGTGGAAGCTTTGCGTGTA
L37761          CAGAAAGAATTGACTTTTGCCGATCTCATTACACAAGGTTCAGTGGAAGCTTTGCGTGTA
                              *                                   *   *

U17227          ATCGAGCAGGCAACCGGCGAGCGCGAAGTCAATTCCATCGGCTATTGCGTCGGCGGGACG
U30612          ATCGAAAAGGCGACCGGGGAGAAGGAGGTCAACGCCGTCGGCTACTGTGTCGGCGGCACG
J05003          GCGCGCGACATCAGCGGCCAGGACAAGATCAACGTGCTCGGCTTCTGCGTGGGCGGCACC
I35763          GCGCGCGACATCAGCGGCCAGGACAAGATCAACGTGCTCGGCTTCTGCGTGGGCGGCACC
U66242          GTCCAGGACGTCAGCGGCCAGGACAAGCTGAACATGTTCGGCTTCTGCGTGGGCGGCACC
U78047          GTGCAGAACATCACCGGCGCCGAGCAGATTAACGCGCTGGGCTTTTGCATGGGGGGCACC
L17049          GTGCGCTCGATCACCCGGCAGAAGCAGATCAACGCGGTAGGCTATTGCATCGCGGGCACC
X97200          GTGCGCGCGATCACCCGGCAGAAGCAGATCAACGCGGTGGGCTACTGCATCGCGGGCACC
D43764          GTTCTGGATCTGACCGACCAGAAAAAGCTGAATGCGGTGGGCTATTGCATCGCCGGCACC
M58445a         GTGCTGGCGATTACCGGCAGCAAGGACCTGAACATGCTCGGTGCCTGCTCCGGCGGCATC
X66592a         GTCCTGTCGATCACCGGCAGCAAGGACCTCAACCTCCTCGGCGCCTGCTCCGGCGGGATC
M58445b         TGCCGCAGCATTAGCGGCAACCGCGACCCCAACCTGATGGGCGCCTGCGCCGGCGGCCTG
X66592b         AGCCGCGAGATCACCGGCAGCCGCAGCGTGAACCTGGCCGGCGCCTGCGCCGGCGGGCTC
X66407          GTCGAGGAGATCACCGGCTCGCCGAAGATCGAGGTGCTCTCCATCTGCCTCGGCGGCGCG
U04848          ATTGAAGAAATTACCGGTGAAAAAGAGGCCAACTGCATTGGCTACTGTATTGGTGGTACG
L37761          ATTGAAGAAATTACCGGTGAAAAAGAGGCCAACTGCATTGGCTACTGTATTGGTGGTACG
                      * *                 *                **    *  **

U17227          CTGCTTGCCGCCACCCTGGCGCTC-CATGCCGCCGAAGGCGAC--GAACGCATTCGCTCC
U30612          TTGCTCGCGGCAACGCTGGCGCTG-CACGCAAAGGAGAAGAAC--AAGCGGATCAAGACC
J05003          ATTGTCTCGACCGCGCTGGCGGTGCTGGCCG-CGCGCGGCGAG--CACCCGGCCGCCAGC
I35763          ATTGTCTCGACCGCGCTGGCGGTGCTGGCCG-CGCGCGGCGAG--CACCCGGCCGCCAGC
U66242          ATCGTTGCCACCGCACTGGCGGTACTGGCGG-CGCGTGGCCAG--CACCCGGCGGCCAGC
U78047          ATCCTCAGCAACGCGCTGGCGGTGCTGGCGG-CGCGTGGCGAC--GAGCCCGTGGCCAGT
L17049          ACGCTCACGCTGACGCTGGCGCACCTGCAGA-AGGCGGGCGAT--CCGTCCGTACGCTCG
X97200          ACGCTGACGCTGACGCTGGCCCACCTGCAGA-AGGCGGGCGAT--CCGTCCGTCCGCTCG
D43764          ACCCTGGCGCTGACCCCTGTCGTGCTGAAGC-AGCGCGGCGAC--GACCGGGTGAACGCG
M58445a         ACCTGCACGGCATTGGTCGGCCAC-TATGCCGCCCTCGGCGA--AAACAAGGTCAATGCC
X66592a         ACCACCGCGACCCTGGTCGGCCAC-TACGTGGCCAGCGGCGA--GAAGAAGGTCAACGCC
M58445b         ACCATGGCCGCACTGCAGGGCCACCTGCAGGCCAAGCACCAGCTGCGCCGGGTGCGCAGC
X66592b         ACCGTAGCCGCCTTGCTCGGCCACCTGCAGGTGCGCCGGCAACTGCGCAAGGTCAGTAGC
X66407          ATGGCCGCGATGGCCGCCGCGCGCGCGTTCG-CCGTCGGCGAC--AAGCGCGTGAGCGCC
U04848          TTACTTGCTGCGACTCAAGCCTATTACGTGGCAAAACGCCTGAAAAATCACGTAAAGTCT
L37761          TTACTTGCTGCGACTCAAGCCTATTACGTGGCAAAACGCCTGAAAAATCACGTAAAGTCT
                              *
```

```
U17227          GCGACGCTCTTCACCACGCAGGTGGATTTCACCCACGCCGGCGATCTCAAGGTTTTCGTG
U30612          GCCACGCTCTTCACCACTCAGGTCGATTTCACCCATGCGGGCGACCTCAAGGTCTTCGTC
J05003          GTCACGCTGCTGACCACGCTGCTGGACTTTGCCGACACGGGCATCCTCGACGTCTTTGTC
I35763          GTCACGCTGCTGACCACGCTGCTGGACTTTGCCGACACGGGCATCCTCGACGTCTTTGTC
U66242          CTGACCCTGCTGACCACCTTCCTCGACTTCAGCGACACCGG-GTGCTCGACGTCTT-GTC
U78047          GCCACCTTTCTTACCACGCTGATCGACTTCAGCGACACCGGCATCCTCGATGTGTTCATC
L17049          GCCACCTTCTTCACCACGCTCACCGACTTTTCGGACCCGGGTGAGGTGGGGGTGTTCCTC
X97200          GCCACCTTCTTCACCACGCTCACCGACTTCTCGGATCCGGGCGAGGTGGGGGTGTTCCTC
D43764          GCCACCTTCTTCACCGCGCTGACCGATTTCGCCGACCAGGGCGAGTTCACTGCCTATCTG
M58445a         CTGACCCTGCTGGTCAGCGTGCTGGACACCACCATGGACAACCAGGTCGCCCTGTTCGTC
X66592a         TTCACCCAACTGGTCAGCGTGCTCGACTTCGAACTGAATACCCAGGTCGCGCTGTTCGCC
M58445b         GCCACCTACCTGGTCAGCTTGCTGGACAGCAAGTTCGAAAGCCCCGCCAGCCTGTTCGCC
X66592b         GTCACCTACCTGGTCAGCCTGCTCGACAGCCAGATGGAAAGCCCGGCGATGCTCTTCGCC
X66407          TTCACGATGCTCAACACGCTGCTCGACTACAGCCAGGTCGGGGAACTGGGTCTGCTGACC
U04848          GCGACCTATATGGCCACCATTATCGACTTTGAAAACCCAGGCAGCTTAGGTGTATTTATT
L37761          GCGACCTATATGGCCACCATTATCGACTTTGAAAACCCAGGCAGCTTAGGTGTATTTATT
                   **     *   *          **

U17227          GACGACGACCAGATCCGCCACCTCGAG-------------GCCAATATGAGCGCCA-CC
U30612          GACGAGGAGCAACTGGCCGCGCTCGAA-------------GAGCATATGCAGGCGG-CC
J05003          GACGAGGGCCATGTGCAGTTGCGCGAGGCCACGCTGGGCGGCGGCGCCGGCGCGCCGTGC
I35763          GACGAGGGCCATGTGCAGTTGCGCGAGGCCACGCTGGGCGGCGGCGCCGGCGCGCCGTGC
U66242          GA-GAAACCCAGGTCGCGCTGCGTGAA-------------CAGCAATTGCGCGATG-GC
U78047          GACGAGGCCTTCGTCAAGTTCCGTGAG-------------ATGCAGATGGGCCATG-GC
L17049          AACGACGATTTCGTCGACGGGATCGAG-------------CGGCAGGTG-GCGGTGGAC
X97200          AACGACGATTTCGTCGACGGTATCGAG-------------CGGCAGGTG-GCGGTGGAC
D43764          CAGGAGGATTTCGTCTCAGGCATCGAG-------------GAGGAGGCG-GCGCGGACC
M58445a         GACGAGCAGACTTTGGAGGCCGCCAAG-------------CGCCACTCCTACCAGGCC
X66592a         GACGAGAAGACTCTGGAGGCCGCCAAG-------------CGTCGTTCCTACCAGTCC
M58445b         GACGAGCAGACCATCGAGGCCGCCAAG-------------CGCCGCTCCTACCAGCGC
X66592b         GACGAGCAGACCCTGGAGAGCAGCAAG-------------CGCCGCTCCTACCAGCAT
X66407          GACCCGGCGACGCTGGACCTCGTCGAG-------------TTCCGGATGCGGCAGCAG
U04848          AATGAACCTGTAGTGAGCGGTTTAGAA-------------AACCTGAACAATCAATTG
L37761          AATGAACCTGTAGTGAGCGGTTTAGAA-------------AACCTGAACAATCAATTG
                 *          *          *

U17227          GGCTACCTCGAAGGCTCGAAGATGGCGTCGGCCTTCAATATGCTCCGGGCTTCGGAACTG
U30612          GGCTATCTCGACGGTTCGAAGATGTCGATGGCTTTCAACATGCTGCGTGCGTCCGAGCTG
J05003          GCGCTGCTGCGCGGCCTTGAGCTGGCCAATACCTTCTCGTTCTTGCGCCCGAACGACCTG
I35763          GCGCTGCTGCGCGGCCTTGAGCTGGCCAATACCTTCTCGTTCTTGCGCCCGAACGACCTG
U66242          GGCCTGATGCCGGGCCGTGACCTGGCCTCGACCTTCTCGAGCCTGCGTCCGAACGACCTG
U78047          GGGCTGATGAAGGGACAGGACCTGGCGTCCACCTTCAGCTTCTTGCGCCCCAACGATCTG
L17049          GGGATCCTCGACAAGACCTTCATGTCGCGCACCTTCAGCTATCTGCGGTCGAACGACCTG
X97200          GGGATCCTCGACAAGACCTTCATGTCGCGCACCTTCAGCTACCTGCGCTCGAACGACCTG
D43764          GGCATCCTGGGCGCGCAGCTGATGACGCGCACCTTCAGCTTCCTGCGCGCCAACGACCTG
M58445a         GGTGTGCTCGAAGGCAGCGAGATGGCCAAGGTGTTCGCCTGGATGCGCCCCAACGACCTG
X66592a         GGCGTGCTGGAGGGCAAGGACATGGCCAAGGTGTTCGCCTGGATGCGCCCCAACGACCTG
M58445b         GGTGTGCTGGATGGCGCCGAGGTGGCGCGGATCTTCGCCTGGATGCGGCCCAACGACCTG
X66592b         GGCGTGCTGGACGGGCGCGACATGGCCAAGGTGTTCGCCTGGATGCGCCCCAACGACCTG
X66407          GGCTTCCTGTCCGGCAAGGAGATGGCCGGCAGCTTCGACATGATCCGCGCGAAGGACCTC
U04848          GGTTATTTCGATGGTCGTCAGTTGGCAGTTACCTTCAGTTTACTGCGTGAAAATACGCTG
L37761          GGTTATTTCGATGGTCGTCAGTTGGCAGTTACCTTCAGTTTACTGCGTGAAAATACGCTG
                *      *              ** *       ***       * **          **

U17227          ATCTGGCCATATTTCGTCAACAACTACCTCAAGGGCCAGGACCCCCTGCCCTTCGACCTG
U30612          ATCTGGCCTTATTTCGTCAACAGCTACCTCAAGGGCCAGGAGCCCCTGCCCTTCGACCTA
J05003          GTGTGGAACTACGTGGTCGACAACTACCTGAAGGGCAACACGCCGGTGCCGTTCGACCTG
I35763          GTGTGGAACTACGTGGTCGACAACTACCTGAAGGGCAACACGCCGGTGCCGTTCGACCTG
U66242          GTATGGAACTATGTGCAGTCGAACTACCTCAAAGGCAATGAGCCGGCGGCGTTTGACCTG
U78047          GTCTGGAACTACGTGGTGGGCAACTACCTCAAGGGCGAAACGCCGCCGCCGTTTGACCTG
L17049          ATCTATCAGCCGGCGATCAAGAGCTACATGATGGGCGAGGCGCCGCCGGCCTTCGACCTG
X97200          ATCTATCAGCCCGCCATCAAGAGCTACATGATGGGAGAGGCGCCCCCCGCCTTCGATCTG
D43764          GTCTGGGGGCCGGCGATCCGCAGCTACATGCTGGGCGAGACGCCGCCGGCCTTCGACCTG
M58445a         ATCTGGAACTACTGGGTCAACAACTACCTGCTCGGCAACGAGCCGCCGGTGTTCGACATC
X66592a         ATCTGGAACTACTGGGTCAACAACTACCTGCTCGGCAACCAGCCGCCGGCGTTCGACATC
M58445b         ATCTGGAACTACTGGGTCAACAACTACCTGCTCGGCAAGACACCACCAGCCTTCGACATC
X66592b         ATCTGGAACTACTGGGTCAACAACTACCTGCTCGGCAGGCAGCCGCCGGCGTTCGACATC
X66407          GTCTTCAACTACTGGGTCTCGCGGTGGATGAAGGGCGAGAAGCCGGCGGCCTTCGACATC
U04848          TACTGGAATTACTACATCGACAACTACTTAAAAGGTAAAGAACCTTCTGATTTTGATATT
L37761          TACTGGAATTACTACATCGACAACTACTTAAAAGGTAAAGAACCTTCTGATTTTGATATT
                   *                       *   *    **        **        ** **  *
```

```
U17227     CTTTACTGGAACTCCGATTCGACTCGGATGCCCGCGGCCAACCACTCTTTCTATCTGCGC
U30612     TTGTTCTGGAACGCCGATTCGACCCGCATGGCGGCGGCAAACCATGCCTTCTACCTTCGC
J05003     CTGTTCTGGAACGGCGACGCCACCAACCTGCCGGGGCCGTGGTACTGCTGGTACCTGCGC
I35763     CTGTTCTGGAACGGCGACGCCACCAACCTGCCGGGGCCGTGGTACTGCTGGTACCTGCGC
U66242     CTGTTCTGGAATTCGGACAGCACCAATTTGCCGGGCCCGATGTTCTGCTGGTACCTGCGC
U78047     CTGTACTGGAACAGCGACAGCACCAACCTGCCCGGTCCGTTCTACGCCTGGTACCTGCGT
L17049     CTCTACTGGAACGGAGACGGCACCAACCTGCCGGCGCAGATGGCGGTCGAATACCTGCGT
X97200     CTCTACTGGAACGGAGACGGCACCAACCTGCCGGCGCAGATGGCGGTCGAATATCTGCGC
D43764     CTGTTCTGGAACGGCGACGGCACCAACCTGCCCGGGCGCATGGCCGTGGAATACCTGCGC
M58445a    CTGTTCTGGAACAACGACACCACGCGCCTGCCGGCCGCCTTCCACGGC-G--ACCTGATC
X66592a    CTCTACTGGAACAACGACACCACGCGCCTGCCCGCCGCGCTGCACGGC-G--AGTTCGTC
M58445b    CTGTACTGGAACGCCGACAGCACGCGCCTGCCCGCGCGCTGCATGGC-G--ACCTGCTG
X66592b    CTCTACTGGAACAACGACAACACGCGGCTGCCCGCGGCGTTCCACGGC-G--AACTGCTC
X66407     CTCGCGTGGAACGAGGACAGCACGAGCATGCCCGCGGAGATGCACTCG-C--ACTACCTC
U04848     TTATATTGGAACAGCGATGGTACGAATATCCCTGCCAAAATTCATAATTTCTTATTGCGC
L37761     TTATATTGGAACAGCGATGGTACGAATATCCCTGCCAAAATTCATAATTTCTTATTGCGC
            *     *****     **     **     *  *  *

U17227     AACTGCTATCTGGAAAACAGGCTCTCCAG---GGGAGAGATG-ATGCTTGCCGGCCGCCG
U30612     AATTGCTATCTTCGCAACGCGCTGACGCA---GAACGAGATG-ATTCTCGACGGCAAGCG
J05003     CACACCTACCTGCAGAACGAGCTCAAGGTACCGGGCAAGCTG-ACCGTGTGCGGCGTGCC
I35763     CACACCTACCTGCAGAACGAGCTCAAGGTACCGGGCAAGCTG-ACCGTGTGCGGCGTGCC
U66242     AACACCTACCTGGAAAACAGCCTGAAAGTGCCGGGCAAGCTG-ACGGTGGCCGGCGAAAA
U78047     AATTTTTACCTGGAGAACAACCTGGTCAAACCAGGCAAGCTC-ACGGTGTGCGGCGAAAA
L17049     GGCCTGTGCCAGCAGGACCGGCTGGCGGG---CGGCACCTTC-CCGGTGCTGGGCTCGCC
X97200     GGCCTCTGCCAGCAGGACCGGCTGGCGGG---CGGCACCTTC-CCGGTGCTGGGCTCGCC
D43764     GGCCTGTGCCAGCAGAACCGCTTCGTCAA---GGAGGGGTTC-GATCTGATGGGCCACCG
M58445a    GAAATGTTCAAGAGCAACCCGCTGACCCGCCCGGACGCCCTG-GAGGTTTGCGGCACTCC
X66592a    GAACTGTTCAAGAGCAACCCGCTGAACCGCCCGGCGCCCTG-GAGGTCTCCGGCACGCC
M58445b    GACTTCTTCAAGCTCAACCCGCTGACCCACCCAGCCGGCCTG-GAGGTATGCGGCACACC
X66592b    GACCTGTTCAAGCACAACCCGCTGACCCGCCCGGGCGCGCTG-GAGGTCAGCGGGACCGC
X66407     CGGTCGCTCTACGGCCGCA-ACGAGCTCGCCGAGGGCCTCTACGTGCTCGACGGCCAGCC
U04848     AATTTGTATTTGAACAATGAATTGATTTCACCAAATGCCGTT-AAGGTTAACGGTGTGGG
L37761     AATTTGTATTTGAACAATGAATTGATTTCACCAAATGCCGTT-AAGGTTAACGGTGTGGG
                                                     *       *    **

U17227     CGTATCCCTCGGCGACGTCAAGATTCCCATCTACAATCTCGCTACGAAGGAGGATCACAT
U30612     CATATCTCTGAAAGACGTGAAGATCCCGATCTATAATCTCGCCACGCGCGAGGATCACAT
J05003     GGTGGACCTGGCCAGCATCGACGTGCCGACCTATATCTACGGCTCGCGCGAAGACCATAT
I35763     GGTGGACCTGGCCAGCATCGACGTGCCGACCTATATCTACGGCTCGCGCGAAGACCATAT
U66242     GATCGACCTCGGCCTGATCGACGCCCCGGCCTTCATCTACGGTTCGCGCGAAGACCACAT
U78047     GCTGGACCTGGGCAATCTCGATTTGCCGGTCTACATCTATGGCTCGCGCGAAGACCACAT
L17049     CGTGGGGCTGAAGGATGTGACGCTTCCCGTCTGCGCCATCGCCTGCGAGACCGACCATAT
X97200     CGTGGGCCTGAAGGATGTGACGCTGCCGGTCTGCGCCATCGCCTGCGAGACCGACCATAT
D43764     CCTGCATGTCGGCGACGTGACCGTGCCGCTTTGCGCCATCGCCTGCGAGACCGACCATAT
M58445a    GATCGACCTGAAACAGGTCAAATGCGACATCTACAGCCTTGCCGGCACCAACGACCACAT
X66592a    CATCGACCTGAAGCAGGTGACTTGCGACTTCTACTGTGTCGCCGGTCTGAACGACCACAT
M58445b    CATCGACCTGCAGAAGGTCGAGCTGGACAGTTTCACCGTGGCCGGCAGCAACGACCACAT
X66592b    GGTGGACCTGGGCAAGGTGGCGATCGACAGCTTCCACGTCGCCGGCATCACCGACCACAT
X66407     GCTGAACCTGCACGACATCGCGTGTGACACCTATGTCGTCGGCGCGATCAACGACCACAT
U04848     CTTGAATCTATCTCGTGTAAAAACACCAAGCTTCTTTATTGCGACGCAGGAAGACCATAT
L37761     CTTGAATCTATCTCGTGTAAAAACACCAAGCTTCTTTATTGCGACGCAGGAAGACCATAT
             *     *         *              *         *          ** ** **
                                   Primer (reverse)    5'  TCTGGTATA

U17227     CGCACCGGCGAAATCCGTGTTCCTCGGCAGCAGCAGCTTCGG---CGGCAAGGTG-----
U30612     CGCCCCCGCCAAGTCGGTTTTCCTCGGCAGCCGGTTCTTCGG---CGGCAAGGTG-----
J05003     CGTGCCGTGGACCGCGGCCTATGCCTCGACCGCGCTGCTGGC---GAACAAGCTG-----
I35763     CGTGCCGTGGACCGCGGCCTATGCCTCGACCGCGCTGCTGGC---GAACAAGCTG-----
U66242     CGTGCCGTGGATGTCGGCGTACGGTTCGCTCGACATCCTGAACCAGGGCAAGCCGGGCGC
U78047     CGTGCCGGCCACGGCCGCTTACGCATCCACCCAGGTGCTGCC---AGGCAAAAAG-----
L17049     CGCGCCGTGGAAAAGCAGCTTCAACGGCTTCCGTCAGTTCGGCTCGACCGACAAG-----
X97200     CGCGCCGTGGAAAAGCAGCTTCAACGGCTTCCGCCAGTTCGGCTCGACCGACAAG-----
D43764     CGCGCCCTGGAAGGACAGCTGGCGCGGCATCGCGCAGATGGGCTCCAGGGACAAG-----
M58445a    CACCCCGTGGCAGTCATGCTACCGCTCGGCGCACCTGTTCGG---CGGCAAGATC-----
X66592a    CACCCCCTGGGAGTCGTGCTACAAGTCGGCCAGGCTGCTGGG---TGGCAAGTGC-----
M58445b    CACCCCGTGGGATGCGGTGTACCGCTCGGCCTTGCTGCTGGG---TGGCGACCGG-----
X66592b    CACGCCCTGGGACGCGGTGTATCGCTCGGCCCTCCTGCTGGG---CGGCCAGCGC-----
X66407     CGTGCCCTGGACGTCGTCGTACCAGGCGGTGAACCTGCTCGG---GGGCGACGTG-----
U04848     CGCACTTTGGGATACTTGTTTCCGTGGCGCAGATTACTTGGG---TGGTGAATCA-----
L37761     CGCACTTTGGGATACTTGTTTCCGTGGCGCAGATTACTTGGG---TGGTGAATCA-----
           *    *                *                     *             *
           GCACGGCAC     3'  Lopez et al. (1997)
```

```
U17227          ----ACCTTCGTGCTCTCCGGCTCCGGGCACATCGCCGGTGTCGTCAACCCTCCGGCCCG
U30612          ----GAATTTGTTGTCACCGGCTCGGGACATATCGCCGGCGTCGTCAACCCGCCCGACAA
J05003          ----CGCTTCGTGCTGGGTGCGTCGGGCCATATCGCCGGTGTGATCAACCCGCCGGCCAA
I35763          ----CGCTTCGTGCTGGGTGCGTCGGGCCATATCGCCGGTGTGATCAACCCGCCGGCCAA
U66242          CAACCGCTTCGTGCTGGGCGCGTCCGGCCATATCGCCGGCGTGATCAACTCGGTGGCCAA
U78047          ----CGTTTTGTGATGGGCGCGTCAGGCCACATTGCTGGCGTGATCAACCCACCCGCCAA
L17049          ----ACCTTCATTCTCTCTCAATCGGGCCATGTGGCGGGCATCGTGAACCCGCCCAGCCG
X97200          ----ACCTTCATTCTTTCGGAATCGGGCCATGTGGCGGGCATCGTGAACCCGCCCAGCCG
D43764          ----ACCTTCATCCTGTCCGAATCGGGCCATATCGCCGGCATCGTCAACCCGCCCAGCAA
M58445a         ----GAGTTCGTGCTGTCCAACAGCGGCCACATCCAGAGCATCCTCAACCCGCCAGGCAA
X66592a         ----GAGTTCATCCTCTCCAACAGCGGTCACATCCAGAGCATCCTCAACCCACCGGGCAA
M58445b         ----CGCTTCGTGCTGGCCAACAGCGGGCACATCCAGAGCATCATCAACCCGCCCGGCAA
X66592b         ----CGCTTCATCCTGTCCAACAGCGGGCACATCCAGAGCATCCTCAACCCTCCCGGAAA
X66407          ----CGCTACGTGCTCACCAACGGCGGGCACGTCGCCGGTGCGGTGAACCCCCCGGGCAA
U04848          ----ACCTTGGTTTTAGGTGAATCTGGACACGTAGCAGGTATTGTCAATCCTCCAAGCCG
L37761          ----ACCTTGGTTTTAGGTGAATCTGGACACGTAGCAGGTATTGTCAATCCTCCAAGCCG
                        *   *  *          ** **  *     *     * **  *

U17227          AA---------GCAAGTATCAATACTGGACGGGAGGGGCG------------CCGAAGGG
U30612          GA---------GGAAATATCAATTCTGGACGGGCGGCCCG------------GCCAAGGG
J05003          GA---------ACAAGCGCAGCCACTGGACTAACGATG-CGCTG-----------CCGGA
I35763          GA---------ACAAGCGCAGCCACTGGACTAACGATG-CGCTG-----------CCGGA
U66242          GA---------ACAAGCGCACGTACTGGATCAACGACGGTGGCG-----------CCGCC
U78047          GG---------GCAAGCGCAGCCACTGGACTCGCGCCGACGGCAAG------TTCCCCGG
L17049          CA---------ACAAATACGGCCATTACACCAACGAGGGCCCGGCC------GGCACGCC
X97200          CA---------ACAAATACGGCCATTACACCAACGAGGGTCCTCTC------GACACACC
D43764          GA---------AGAAATACGGCCATTATACCTCGGACGCCGGTTTC------GGTCAGGG
M58445a         CC---------CCAAGGCGCGCTTCATGACCGGTGCCGAT------------CGCCCGGG
X66592a         CC---------CCAAGGCACGCTTCATGACCAATCCGGAA------------CTGCCCGC
M58445b         CC---------CCAAGGCCTACTACCTGGCCAACCCCAAG------------CTGTCCAG
X66592b         CC---------CCAAGGCCTGCTACTTCGAGAACGACAAG------------CTGAGCAG
X66407          GCGGGTGTGGTTCAAGGCCGTCGGCGCGCCGGACGCCGAGTCCGGCACCCCCCTGCCCGC
U04848          TA---------ATAAATACGGTTGCTACACCAATGCTGCC------------AAGTTTGA
L37761          TA---------ATAAATACGGTTGCTACACCAATGCTGCC------------AAGTTTGA
                           **

U17227          CGACATCGAGACCTGGATGGGTAAAGCGAAGGAGACGGCCGGGTCCTGGTGGCCGCATTG
U30612          CGAATACGAGACCTGGCTCGAGCAGGCGAGCGAGACGCCCGGATCATGGTGGCCACATTG
J05003          GTCGCCGCAGCAATGGCTGGCCGGCGCCATCGAGCATCACGGCAGCTGGTGGCCGGACTG
I35763          GTCGCCGCAGCAATGGCTGGCCGGCGCCATCGAGCATCACGGCAGCTGGTGGCCGGACTG
U66242          GATGCC-CAGGCCTGGTTCGATGGCGCGCAGGAAGTGCCGGGCAGCTGGTGGCCGCAATG
U78047          CACACTCGACCAATGGCTGGAAGGCGCCACCGAGCACCCCGGCAGCTGGTGGACCGACTG
L17049          GGAGTC---GTTCCGG---GAGGGGGCCGAGTTCCACGCGGGCTCCTGGTGGCCGCGCTG
X97200          CGCCGC---GTTCCGC---GAGGGGGCCGAGTTCCACGCAGGCTCCTGGTGGCCGCGCTG
D43764          CGAGCA---GCACTGGCTGGACAAGGCCAGCCATCACGAGGGCAGCTGGTGGGGCCGCTG
M58445a         TGACCCGGTGGCCTGGCAGGAAAACGCCACCAAGCATGCCGACTCCTGGTGGCTGCACTG
X66592a         CGAGCCCAAGGCCTGGCTGGAACAGGCCGGCAAGCACGCCGACTCGTGGTGGTTGCACTG
M58445b         CGACCCGCGTGCCTGGCTCCACGATGCCAAGCGCAGCGAAGGCAGCTGGTGGCCGTTGTG
X66592b         CGATCCACGCGCCTGGTACTACGACGCCAAGCGCGAAGAGGGCAGCTGGTGGCCGGTCTG
X66407          GGATCCGCAGGTCTGGGACGAGGCCGCCACCCGCTACGAGCACTCGTGGTGGGAGGACTG
U04848          AAATACCAAACAATGGCTAGATGGCGCAGAATATCACCCTGAATCTTGGTGGTTGCGCTG
L37761          AAATACCAAACAATGGCTAGATGGCGCAGAATATCACCCTGAATCTTGGTGGTTGCGCTG
                         *          **                      ******      **

U17227          GCAGGGTTGGGTCGAACGGCTCGACAAACGCAGGGTTCCGGCGCGG---AAGGCCGGAGG
U30612          GCAAGCCTGGATAGAGACGCATGACGGCAGACGCGTTGCAGCGCGC---AAGCCCGGCGG
J05003          GACCGCATGGCTGGCCGGGCAGGCCGGCGCGAAACGCGCCGCGCCCGCCAACTATGGCAA
I35763          GACCGCATGGCTGGCCGGGCAGGCCGGCGCGAAACGCGCCGCGCCCGCCAACTATGGCAA
U66242          GGCCGGGTTCCTGACCCAGCATGGCGGCAAGAAGGTCAAGCCCAAGGCCAAGCCCGGCAA
U78047          GTCAGGCTGGCTCAAGAGCCATGCGGGCAAGCAGATTGCAGCGCCCAAGGCCTATGGCAA
L17049          GGGCGCCTGGCTCGCCGAGCGATCGGGCAAGCAGGTCCCGGCGCGC-C--AGCCGGGCGA
X97200          GGGCGCCTGGCTGGCCGAGCGGTCGGGCAAGCAGGTCCCGGCGCGC-C--AGCCGGGCGA
D43764          GGGCGAATGGCTGGCCCGGCGGGCGGGGGGCATGGTCGATGCCCGC-G--ACCCGGGCGA
M58445a         GCAAAGCTGGCTGGGCGAGCGTGCCGGCGAGCTGGAAAAGGCGCCGACCCGCCTGGGCAA
X66592a         GCAGCAATGGCTGGCCGAACGCTCCGGCAAGACCCGCAAGGCGCCCGCCAGCCTGGGCAA
M58445b         GCTGGAGTGGATCACCGCGCGCTCCGGCCCGCTCAAGGCACCGCGCAGCGAACTGGGCAA
X66592b         GCTGGGCTGGCTGCAGGAGCGCTCGGGCGAGCTGGGCAACCCTGACTTCAACCTTGGCAG
X66407          GACGGCGTGGTCGAACAAGCGCGCCGGGGAACTGGTGGCGCCGCCGGCCA---TGGGGAG
U04848          GCAGGCATGGGTCACACCGTACACTGGTGAACAAGTCCCTGCCCGC---AACTTGGGTAA
L37761          GCAGGCATGGGTCACACCGTACACTGGTGAACAAGTCCCTGCCCGC---AACTTGGGTAA
                *      *                                   *             **
```

```
U17227          ------TCCGCTCAATTCCATCGAGGAAGCGCCCGGCTCCTACGTGCGCGTGCGCGCCTG
U30612          C---GATGCGCTGAACGCGATCGAAGAAGCACCGGGAAGTTATGTGATGGAACGCACCTG
J05003          ---TGCGCGCTATCGCGCAATCGAACCCGCGCCTGGGCGATACGTCAAAGCCAAGGCATG
I35763          ---TGCGCGCTATCGCGCAATCGAACCCGCGCCTGGGCGATACGTCAAAGCCAAGGCATG
U66242          ---CGCCCGCTACACCGCGATCGAGGCGGCGCCCGGCCGTTACGTCAAAGCCAAGGGCTG
U78047          AGGCACCAAATTCAAGGCCATCGAGCCCGCGCCTGGTCGCTACGTCAAGCAGAAAGCCTG
L17049          T---TCGAAACATCCCGAGCTCGCGCCGGCGCCCGGATCCTATGTGGCGGCGGTGGGCGG
X97200          T---TCGAAACATCCCGAGCTCGCGCCGGCGCCGGGATCCTATGTGGCGGCACCGGGCGG
D43764          -----GGG----CTTCGGCCCTGCGCCGGGCCTTTACGTCCACGAGCGGGCGTAA-----
M58445a         ---CCGTGCCTATGCCGCTGGCGAGGCATCCCCGGGCACCTACGTTCACGAGCGTTGA--
X66592a         ---CAAGACCTATCCGGCCGGCGAAGCCGCGCCCGGAACCTACGTGCATGAACGATGA--
M58445b         ---TGCCACCTACCCACCGCTGGGCCCCGCGCCGGGCACCTACGTGCTGACCCGATGA--
X66592b         ---CGCCGCGCATCCGCCCCTCGAAGCGGCCCCGGGCACCTACGTGCATATACGCTGA--
X66407          ---TACCGCCCATCCCCCGCTCGAGGATGCTCCGGGTACGTACGTCTTCAGCTGA-----
U04848          ---TGCGCAGTATCCAAGCATTGAAGCGGCACCGGGTCGCTATGTTTTGGTAAATTTATT
L37761          ---TGCGCAGTATCCAAGCATTGAAGCGGCACCGGGTCGCTATGTTTTGGTAAATTTATT
                                          *         *          * *

U17227          A-------------------------------
U30612          A-------------------------------
J05003          A-------------------------------
I35763          A-------------------------------
U66242          A-------------------------------
U78047          A-------------------------------
L17049          GGCTTGA---------------------------
X97200          GGCGGAAGTCGCCAAGCCTGCCGCGAGGTAA
D43764          -------------------------------
M58445a         -------------------------------
X66592a         -------------------------------
M58445b         -------------------------------
X66592b         -------------------------------
X66407          -------------------------------
U04848          CTAA----------------------------
L37761          CTAA----------------------------
```

Appendix D Kanamycin resistant gene from Genbank (accession No M17626)

```
          ▼ Nhe I
  1 CGATAAGCTA GCTTCACGCT GCCGCAAGCA CTCAGGGCGC AAGGGCTGCT AAAGGAAGCG
 61 GAACACGTAG AAAGCCAGTC CGCAGAAACG GTGCTGACCC CGGATGAATG TCAGCTACTG
                                  5'   CTGCAG-C CGGATGAATG TCAGCTA 3'
                                       Pst I     #3623 (F)

121 GGCTATCTGG ACAAGGGAAA ACGCAAGCGC AAAGAGAAAG CAGGTAGCTT GCAGTGGGCT
                                                 Promoter (?)(227-276)
181 TACATGGCGA TAGCTAGACT GGGCGGTTTT ATGGACAGCA AGCGAACCGG AATTGCCAGC
                                               Promoter (?)(280-329)
241 TGGGGCGCCC TCTGGTAAGG TTGGGAAGCC CTGCAAAGTA AACTGGATGG CTTTCTTGCC
                                                              RBS(Kanʳ)
301 GCCAAGGATC TGATGGCGCA GGGGATCAAG ATCTGATCAA GAGACAGGAT GAGGATCGTT
        Start (Kanʳ)
361 TCGCATGATT GAACAAGATG GATTGCACGC AGGTTCTCCG GCCGCTTGGG TGGAGAGGCT
421 ATTCGGCTAT GACTGGGCAC AACAGACAAT CGGCTGCTCT GATGCCGCCG TGTTCCGGCT
481 GTCAGCGCAG GGGCGCCCGG TTCTTTTTGT CAAGACCGAC CTGTCCGGTG CCCTGAATGA
541 ACTCCAAGAC GAGGCAGCGC GGCTATCGTG GCTGGCCACG ACGGGCGTTC CTTGCGCAGC
601 TGTGCTCGAC GTTGTCACTG AAGCGGGAAG GGACTGGCTG CTATTGGGCG AAGTGCCGGG
661 GCAGGATCTC CTGTCATCTC ACCTTGCTCC TGCCGAGAAA GTATCCATCA TGGCTGATGC
721 AATGCGGCGG CTGCATACGC TTGATCCGGC TACCTGCCCA TTCGACCACC AAGCGAAACA
781 TCGCATCGAG CGAGCACGTA CTCGGATGGA AGCCGGTCTT GTCGATCAGG ATGATCTGGA
841 CGAAGAGCAT CAGGGGCTCG CGCCAGCCGA ACTGTTCGCC AGGCTCAAGG CGCGGATGCC
901 CGACGGCGAG GATCTCGTCG TGACCCATGG CGATGCCTGC TTGCCGAATA TCATGGTGGA
961 AAATGGCCGC TTTTCTGGAT TCATCGACTG TGGCCGGCTG GGTGTGGCGG ACCGCTATCA
1021 GGACATAGCG TTGGCTACCC GTGATATTGC TGAAGAGCTT GGCGGCGAAT GGGCTGACCG
                                  #3722 (F)

1081 CTTCCTCGTG CTTTACGGTA TCGCCGCTCC CGATTCGCAG CGCATCGCCT TCTATCGCCT
                        Stop (Kanʳ)              ▼ Bst BI
1141 TCTTGACGAG TTCTTCTGAG CGGGACTCTG GGGTTCGAAA TGACCGACCA AGCGACGCCC
       3' TGCTC AAGAAGACTC GCC-GACGTC  5'
                 #3624 (R)       Pst I

1201 AACCTGCCAT CACGAGATTT CGATTCCACC GCCGCCTTCT ATGAAAGGTT GGGCTTCGGA
1261 ATCGTTTTCC GGGACGCCGG CTGGATGATC CTCCAGCGCG GGGATCTCAT GCTGGAGTTC
1321 TTCGCCCACC CCAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC
//
```

❖ Kanamycin resistant gene DNA fragment in pBSL15 (Alexeyev, 1995): *Nhe* I/*Bst* BI (7-1,175 bp).