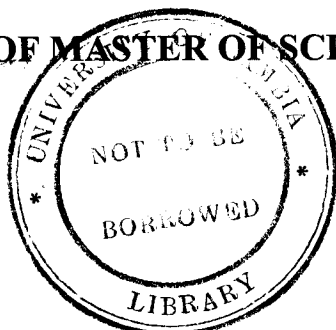


THE CHARACTERISATION OF AN ACETOPHENONE MONOOXYGENASE

BY

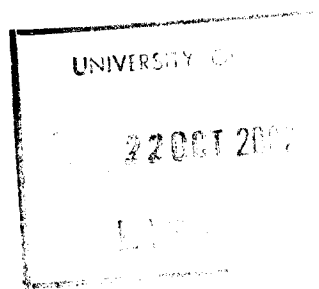
MOOLA SISEHO MUTONDO

**A DISSERTATION SUBMITTED TO THE UNIVERSITY OF
261734
ZAMBIA IN PARTIAL FULFILMENT OF THE REQUIREMENTS
OF THE DEGREE OF MASTER OF SCIENCE IN CHEMISTRY**



THE UNIVERSITY OF ZAMBIA

2002



M.Sc.
Thesis
2002
C

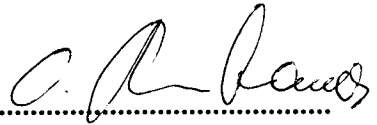
APPROVAL

THIS DISSERTATION BY Ms. M.S. Mutondo IS
APPROVED AS FULFILLING THE REQUIREMENT FOR THE AWARD OF
THE DEGREE OF MASTER OF Sciences BY
UNIVERSITY OF ZAMBIA.

NAME

SIGNATURE

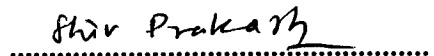
Univ.-Prof.Dr.Christian Sturmbauer
External Examiner



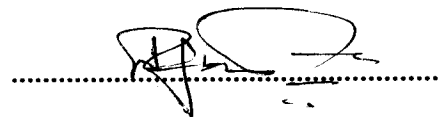
.....
Supervisor and Internal Examiner

.....

DR. S. PRAKASH
Internal Examiner



PROF. K.S. Mbatia
Dissertation Chairperson



DECLARATION

I hereby declare that this dissertation or any part of it has not previously been submitted for a degree in this or any other university.

Abstract

Catechols are important precursors in the production of pharmaceuticals but they are associated with many problems such as their instability and susceptibility to oxidation and polymerisation. In order to avoid these problems, acylcatechols are used. However, disadvantages exist in the industrial production of acylcatechols. The production of acylcatechols industrially involves chemical syntheses that require large amounts of hazardous peracids, and accumulation of large amounts of undesirable byproducts. The processes used also have low yields. A solution to these problems associated with the production of acylcatechols lies in the use of biocatalytic means of synthesising acylcatechols.

Baeyer-Villiger monooxygenases (BVMOs) have been found to have the ability to yield key chiral products of value in various chemoenzymatic syntheses used in industry. This study investigates the potential of the BVMOs to synthesise acylcatechols thereby avoiding the use of complex and hazardous procedures.

Preliminary screening of the ability of various bacteria to perform Baeyer-Villiger oxidations to produce acylcatechols showed that *Pseudomonas fluorescens* ACB and *Arthrobacter* sp. M5 have notable potential to catalyse such reactions.

In this study, the acetophenone monooxygenase produced by *Arthrobacter* sp. M5 was selected for further investigation. The acetophenone monooxygenase gene was cloned in *Escherichia coli* HB101 in order to enable sequence analysis studies of the

gene. This also facilitated greater manipulation of the gene in a host (*E. coli*) that is well studied and easy to apply in the overexpression of the acetophenone monooxygenase gene. Overexpression of the acetophenone monooxygenase gene is necessary because it allows for greater amounts of the acetophenone monooxygenase to be produced for optimisation studies that enable the enzyme to be tailored to the needs of industry. After optimisation, the acetophenone monooxygenase can be used to biocatalytically produce valuable acylcatechols thus avoiding the drawbacks associated with chemical syntheses.

A library of *Arthrobacter* sp. M5 total DNA was made in *E. coli* HB101 and was screened for Baeyer-Villiger monooxygenase activity using two colorimetric methods. The library was also screened using degenerate probes by Southern blotting.

ACKNOWLEDGEMENTS

I would like to thank the Chemistry Department at the University of Zambia for their motivation and their help in the compilation of this dissertation; Dr. L. E. Mumba for his invaluable time and effort as my supervisor; and Prof. D. B. Janssen for making it possible for me to study in his work group at the Rijks Universiteit Groningen, Holland.

I would also like to thank Nanne Kamerbeek, Rijks Universiteit , for his patient supervision and mentoring during my stay in Holland. I am grateful to Prof. C. E. Lovelace for her generous input and guidance.

TABLE OF CONTENTS

ABSTRACT	III
ACKNOWLEDGEMENTS.....	V
TABLE OF CONTENTS.....	VI
TABLE OF FIGURES.....	X
LIST OF TABLES	XIII
CHAPTER 1: INTRODUCTION.....	14
1.1 INTRODUCTION.....	14
CHAPTER 2: LITERATURE REVIEW	19
2.1 BAEYER-VILLIGER MONOOXYGENASES.....	19
2.2 OCCURRENCE IN NATURE	23
2.3 STRUCTURAL FEATURES OF BVMOs	23
2.4 SYNTHETIC APPLICATIONS OF BVMOs	27
2.4.1 <i>Enantioselective biotransformations</i>	27
2.4.2 <i>Kinetic resolution as a synthetic application</i>	29
2.5 THE STRUCTURE OF THE ACTIVE SITE OF BAEYER-VILLIGER MONOOXYGENASES	
31	
2.5.1 <i>Active site modelling of BVMOs</i>	31
2.5.2 <i>'Cubic space' BVMO models</i>	31
2.5.3 <i>Mechanism-based models of BVMOs</i>	34
2.6 ENZYMES	37

2.7	CLASSIFICATION OF ENZYMES	38
2.7.1	<i>Oxidoreductases:</i>	38
2.7.2	<i>Transferases:</i>	38
2.7.3	<i>Hydrolases:</i>	38
2.7.4	<i>Isomerases:</i>	38
2.7.5	<i>Lyases</i>	39
2.7.6	<i>Ligases</i>	39
2.8	OXIDATIONS.....	39
2.8.1	<i>Dioxygenases</i>	39
2.8.2	<i>Monooxygenases</i>	42
2.9	CLASSIFICATION OF FLAVOPROTEINS.....	46
2.9.1	<i>Simple flavoproteins</i>	46
2.9.2	<i>Flavoproteins with Auxiliary redox centres</i>	47
2.10	SUBSTRATE SPECIFICITY OF <i>ARTHROBACTER</i> SP. M5	48
2.11	DNA MANIPULATION PROTOCOLS.....	52
2.11.1	<i>Principle of the DNA packaging kit</i>	52
2.11.2	<i>Packaging substrates:</i>	53
2.11.3	<i>The Polymerase Chain Reaction</i>	53
2.11.4	<i>Restriction enzymes</i>	61
2.11.5	<i>Cloning technique</i>	63
2.12	SOUTHERN BLOTTING.....	68
2.12.1	<i>Complementarity and hybridisation</i>	68
2.12.2	<i>The blotting procedure</i>	71
2.13	<i>ARTHROBACTER</i> SP. M5.....	75
2.13.1	<i>Metabolism of substituted acetophenones</i>	75

2.13.2 Isolation and characterisation of a 4-chloroacetophenone-degrading culture.	76
CHAPTER 3: MATERIALS AND METHODS	77
3.1 MATERIALS.....	77
3.2 METHODS.....	81
3.2.1 Cultivation of bacteria	81
3.3 ISOLATION OF DNA	82
3.3.1 Isolation of total DNA from <i>Arthrobacter</i> sp. M5.....	82
3.3.2 Isolation of cosmid DNA (pLAFR3).....	83
3.3.3 pLAFR3 purification	84
3.4 PREPARATION OF DNA	85
3.4.1 Preparation of partial digest.....	85
3.4.2 Preparation of vector arms	85
3.4.3 Ligation of DNA fragments	86
3.4.4 Packaging of DNA.....	87
3.5 STORAGE OF LIBRARY	87
3.6 SCREENING OF LIBRARY	87
3.6.1 Screening method (A).....	87
3.6.2 Screening method (B.).....	88
3.6.3 Screening using probe method (C).....	88
3.7 PREPARATION OF GELS	89
3.8 MUTAGENESIS.....	90
3.8.1 PCR-based Site Directed Mutagenesis	90
3.9 DIGESTIONS.....	91
3.10 PREPARATION OF CELL-FREE EXTRACTS.....	93

3.11	TRANSFORMATION OF CELLS.....	93
3.12	MINIPREPARATIONS OF DNA.....	93
3.13	COLORIMETRIC ASSAYS.....	95
CHAPTER 4: RESULTS AND DISCUSSION		96
4.1	ISOLATION OF TOTAL DNA.....	96
4.2	ISOLATION OF COSMID VECTOR	98
4.3	PURIFICATION OF PLAFR3	102
4.4	PARTIAL DIGESTION OF <i>ARTHROBACTER SP. M5.</i>	103
4.5	PREPARATION OF VECTOR ARMS.	103
4.6	LIGATION OF DNA FRAGMENTS.....	104
4.7	GENE LIBRARY CONSTRUCTION.....	107
4.8	SCREENING THE GENE LIBRARY	115
4.9	DNA MINIPREPARATIONS.....	116
4.10	MUTANTS.....	121
4.11	SCREENING METHOD C: SOUTHERN BLOTTING.....	122
4.12	COMPARISON OF THE USE OF SOUTHERN BLOTTING AND PCR TO DETECT A GENE OF INTEREST.....	123
CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS.....		128
<u>REFERENCES</u>.....		131

TABLE OF FIGURES

Figure 2-1 Characterised pathway of progesterone metabolism in <i>Cylindrocarpon radicola</i> , containing two examples (denoted by *) of Baeyer-Villiger monooxygenase catalysed reactions	20
Figure 2-2 a) Intrasequential regeneration of nicotinamide nucleotide coenzymes using a coupled alcohol dehydrogenase (DH) and Baeyer-Villiger monooxygenase (MO). b) Chemical structure of NAD and NADH.....	22
Figure 2-3. Sequence data for Type 1 Baeyer-Villiger monooxygenases	25
Figure 2-4. Sequence data for Type 2 Baeyer-Villiger monooxygenases	26
Figure 2-5. (a) Enantiodivergent biotransformation of bicyclo [3.2.0] hept-2-en-6-one by cyclohexanone monooxygenase (CHMO) and 2,5-diketocamphane 1,2-monooxygenase. (b) Chemoenzymatic syntheses of various useful target molecules (multifidiene and viridiene, sarkomycin A and clerodin involving the biotransformation by (CHMO) of bicyclo [3.2.0]-type ketones.	28
Figure 2-6. (a) Kinetic resolution of 2'-substituted monocyclic ketones by Baeyer-Villiger monooxygenases. (b) Chemoenzymatic synthesis of (<i>R</i>)-(+)-lipoic acid by either 2-oxo-4, 5,5-trimethylcyclopent-3-enylacetyl-CoA monooxygenase (BVMO 1) or cyclopentanone monooxygenase (BVMO 2)	30
Figure 2-7 Active site model of cyclohexanone monooxygenase (CHMO) showing 'Cubic space' model proposed by Alphand and Furtoss . The 'forbidden zone' is indicated by dotted lines.	32
Figure 2-8 Active site model of cyclohexanone monooxygenase (CHMO) showing 'Cubic space' model proposed by Ottolina”.....	33
Figure 2-9 Mechanism-based model proposed by Taschner.....	34
Figure 2-10 Mechanism-based model proposed by Kelly’.....	36

Figure 2-11 a) An acetonide. (b) A tricyclic product of Diels-Alder reaction. (c) A cyclobutanone derivative as a product of a cycloaddition.	40
Figure 2-12 (a) The diol from trifluoromethylbenzene. (b) The corresponding acetonide (obtained from trifluoromethylbenzene).....	41
Figure 2-13. Compounds of interest formed by the action of monooxygenases.	43
Figure 2-14. A bicyclic compound obtained by macro-lactonisation followed by a Diels-Alder reaction on the biotransformation product.	45
Figure 2-15. (a) Oxidation of cyclohexylcyclohexane to give the diol by <i>Cunninghamella blakesleeana</i>	45
Figure 2-16. Substrate specificity of <i>Pseudomonas</i> ACB showing effect of para-, meta- and ortho- substitutions.....	50
Figure 2-17 Substrate specificity of <i>Pseudomonas</i> ACB showing para-substituent activities.	50
Figure 2-18 <i>Arthrobacter</i> sp. M5 activity for 2-hydroxyacetophenone, chlorinated acetophenones and methylated acetophenones.	51
Figure 2-19 The Polymerase Chain Reaction	54
Figure 2-20 Actions of restriction enzymes	62
Figure 2-21. Outline of cloning strategy.....	64
Figure 2-22 . Pathway for degradation of acetophenone (R=H) and chlorinated analogs (R=Cl) via the corresponding phenyl acetates, phenols and catechols based on the model of Cripps and Cripps <i>et al</i> . Degradation of 1- phenylethanols can follow the same pathway or occur by ring fission with retention of the side-chain.....	76
Figure 4-1 Agarose gel showing total DNA and Smart Ladder®	97
Figure 4-2. Agarose gel showing pLAFR3 DNA with concatemers.	99

Figure 4-3 Agarose gel showing pLAFR3 cosmid DNA after digestion with EcoRI.
Concatemerisation is not observed.....101

Figure 4-4 Agarose gel showing progressively less degree of digestion of
Arthrobacter sp M5 total DNA by SauIIIA restriction enzyme.105

Figure 4-5. Agarose gel showing selected sample #4 after preparation for ligation as
described in Section 3.106

Figure 4-6. Vector pLAFR3 after digestion with EcoRI and HindIII restriction
enzymes.....109

Figure 4-7. Vector pLAFR3 after digestion with alkaline phosphatase.....110

Figure 4-8. Vector pLAFR3 after phenol extraction.....111

Figure 4-9. Vector pLAFR3 after phenol extraction and ethanol precipitation at -20
°C (one of the unsuccessful attempts).....112

Figure 4-10. Vector pLAFR3 after phenol/chloroform extraction and ethanol
precipitation at -20 °C.113

Figure 4-11. Vector pLAFR3 and M5 DNA after ligation.114

Figure 4-12. Agarose gel showing results of a miniprep. Lanes with multiple bands
represent the presence of insert DNA.117

Figure 4-13. PCR to make a probe.....124

Figure 4-14 Gels for blotting.....125

Figure 4-15. Southern blots.....126

Figure 4-16 Smart Ladder®127

LIST OF TABLES

Table 2-1 <i>Escherichia coli</i> codon usage chart.	66
Table 3-1. Buffers and solutions used in the protocols.....	77
Table 3-2 Solutions used in isolation of DNA.....	78
Table 3-3 Enzymes, antibiotics and co-factors	79
Table 3-4 Vectors and bacteria	79
Table 3-5 Commercially available kits	79
Table 3-6 Bacterial strain properties	80
Table 3-7 Vector properties	80
Table 3-8 Enzyme digestions	92

CHAPTER 1: INTRODUCTION

1.1 Introduction

Baeyer-Villiger monooxygenases (BVMO's) have been recognised in industry for their ability to yield chiral products, which are of value in various industrial chemical syntheses. Catechols are important precursors in the production of pharmaceuticals but have many problems related to their use. The primary problem encountered with the use of catechols is their instability. They are susceptible to oxidation and polymerisation, which is a substantial drawback to their use in industrial chemical syntheses. Hence acylcatechols are used to avoid the problems caused by oxidation and polymerisation. However, a disadvantage also exists in the use of acylcatechols in industrial reactions. Large amounts of hazardous peracids and other undesirable by-products accumulate in industrial chemical syntheses using acylcatechols and this presents the added disadvantage in their use. A solution to this problem exists in the use of biocatalytic means to synthesise acylcatechols since this can be done without the above mentioned drawbacks.

In order to utilize biocatalysis to synthesise acylcatechols, it is necessary to do the following:

- a. Identify the Baeyer-Villiger oxidation and a suitable substrate to obtain the acylcatechol of interest as a product.
- b. Locate viable organisms that are capable of performing the desired Baeyer-Villiger oxidations that produce acylcatechols of interest.

- c. Clone the BVMO that is responsible for catalyzing the identified oxidation in an organism that will allow for further studies on the enzyme and subsequent overexpression to obtain larger quantities of the enzyme.

A survey of related literature shows that there are various microorganisms that can perform a Baeyer-Villiger oxidation using NAD(P)H-dependent monooxygenases containing flavine in the form of either FMN or FAD.

Preliminary studies of the ability of various bacterial strains to oxidise substituted acetophenones via acetophenone monooxygenase to produce their corresponding phenyl acetates showed that *Pseudomonas* ACB, *Arthrobacter* sp. M5 and *Alcaligenes* ACA possessed notable capability to catalyse such reactions. Studies of *Pseudomonas* ACB and *Alcaligenes* ACA were already in progress in our research group, therefore *Arthrobacter* sp. M5 was selected for further investigation in this study.

The substrate selected for use in this study had to have a dual function. It had to firstly be able to be used as the sole carbon source by *Arthrobacter* sp. M5 to facilitate the isolation of the bacteria. Secondly, the substrate had to be able to function as an indicator to signal the activity of a BVMO. This was necessary for the detection of the clone containing the BVMO amongst the other clones in the library devoid of BVMO activity. The substrate selected was para-nitroacetophenone. This compound was utilized as a sole carbon source in the isolation step of this study. In the presence of a BVMO, para-nitroacetophenone is oxidised to para-nitrophenol and acetic acid. Para-nitroacetophenone is colourless whereas para-nitrophenol is bright yellow, hence its

presence is easily detected. Therefore, the clone exhibiting BVMO activity could be located.

Cloning of the acetophenone BVMO was the major part of this study. To achieve this, the cloning strategy outlined in greater depth in Section 2.6.5.1 was utilised. This strategy involved the use of restriction enzymes (EcoRI, HindIII and BamHI) to prepare pLAFR3 vector arms and SauIIIA to prepare insert DNA isolated from *Arthrobacter* sp. M5 in the form of a partial digest. Ligating the vector arms and the insert DNA would result in a pool of DNA fragments each containing a section of the original *Arthrobacter* sp. M5 total DNA. These could then be inserted in *Escherichia coli* HB101 cells by phage infection. Once the library had been made, it could be screened using para-nitroacetophenone as a substrate. The presence of the clone containing the DNA encoding the acetophenone BVMO could be identified by the presence of a bright yellow colour (Figure 1-1).

Once the clone was isolated, further studies on the structure and active site of the enzyme could be carried out using overexpression as a means of obtaining larger quantities of the enzyme. Such studies could provide data enabling the optimisation of the enzyme for use in industry.

In the last part of this study, the esterase that catalyses the cleavage of phenyl acetate derivatives formed by acetophenone monooxygenase was investigated. This esterase gene (HapE) from *Pseudomonas* ACB had been cloned by N. Kamerbeek, a member of our research group, using the Polymerase Chain Reaction with degenerated primers based on the N-termini of the esterase and the monooxygenase. The structure of the

HapE gene product was studied by introducing point mutations into this gene. Point mutations introduced in the characteristic “motif” regions of the enzyme affect the folding and catalytic properties of the enzyme. These types of studies add to the knowledge currently available about the function of the motifs found in this family of enzymes.

Little is currently known about the mechanism of a biological Baeyer-Villiger reaction. Until now only one Baeyer-Villiger monooxygenase, cyclohexanone monooxygenase from *Acinetobacter* NCIMB 9871 has been cloned in *E. coli*. DNA sequences obtained from this study can be used to deduce the amino acid sequence of the enzyme and this in turn can provide information about related flavine-dependent enzymes when studied in conjunction with homology studies.

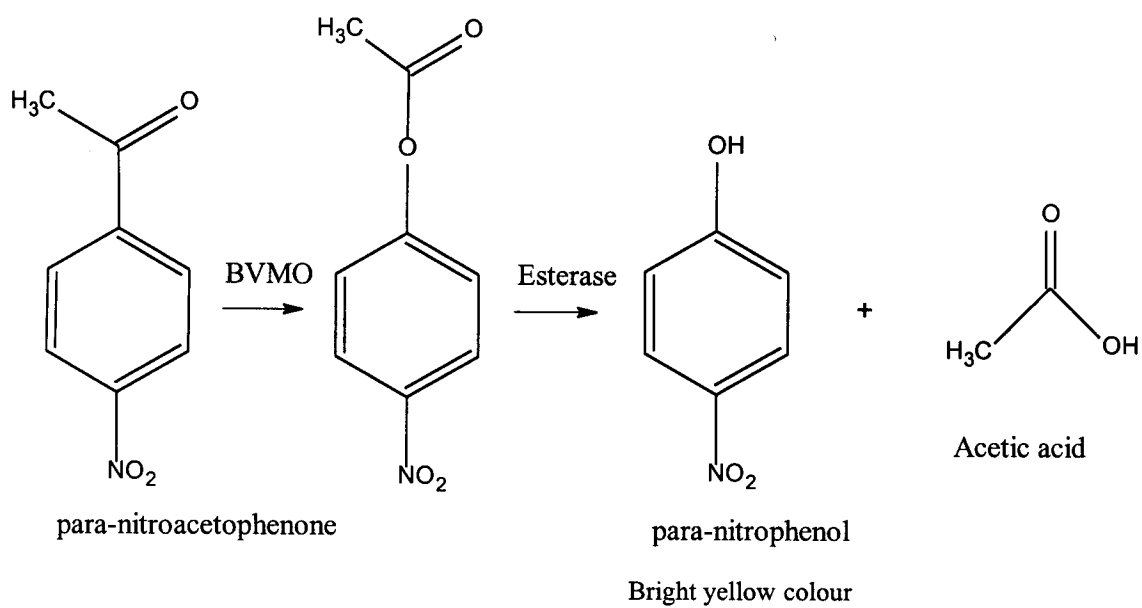


Figure 1-1 Reaction catalysed by acetophenone monooxygenase

CHAPTER 2:LITERATURE REVIEW

2.1 Baeyer-Villiger Monooxygenases

Baeyer-Villiger monooxygenases (BVMOs) (EC 1.14.13.x) are a group of flavoproteins with considerable potential for yielding key chiral products of value in the chemoenzymatic synthesis of a wide range of useful compounds such as neuroactive pharmaceuticals.¹. These enzymes exhibit the rare characteristic shared with notable biocatalysts such as ribulose biphosphate carboxylase/oxidase, of being able to catalyse two mechanistically different types of biochemical reactions, apparently in the same active site. Figure 2-1 shows two examples of Baeyer-Villiger monooxygenase catalysed reactions.

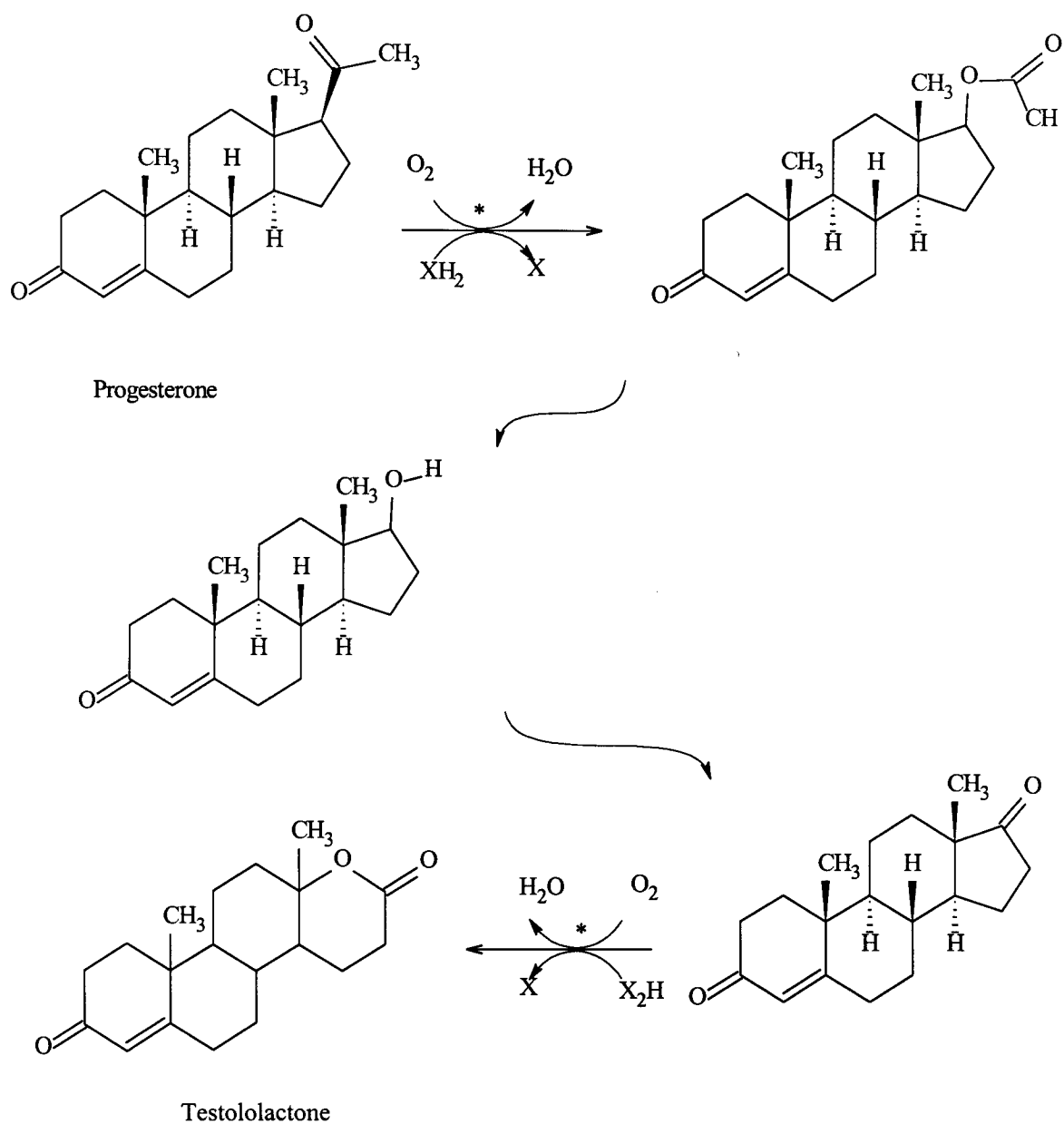


Figure 2-1 Characterised pathway of progesterone metabolism in *Cy lindrocarpon rad icicola*², containing two examples (denoted by *) of Baeyer-Villiger monooxygenase catalysed reactions

BVMOs were first recognised over 40 years ago in the initial era of steroid biotransformations. They were recognised as carbonyl monooxygenases with the ability to catalyse the nucleophilic oxygenation of ketones and aldehydes in a manner analogous to the established peracid-catalysed organic chemical reaction from which they take their name³. However it has been recently recognised that these enzymes can also catalyse the electrophilic oxygenation of various heteroatoms, as illustrated by their ability to form sulphoxides from organosulfides⁴. Both types of reaction can proceed with high selectivity. BVMOs utilise NAD(P)H as a coenzyme. The regeneration of NAD(P)H is often coupled to other enzymes as shown in Figure 2-2.

Cyclohexanone monooxygenase (CHMO) is the most studied and well understood BVMO; this enzyme purified from cyclohexanol-grown *Acinetobacter calcoaceticus* (NCIMB 9871) yields homochiral (S)-5-methyl-lactone enantiomeric excess (e.e.) = 98%] from 4-methylcyclohexanone⁵, and (R)-methyl phenyl sulphoxide (e.e.= 99%) from the equivalent sulphide⁶.

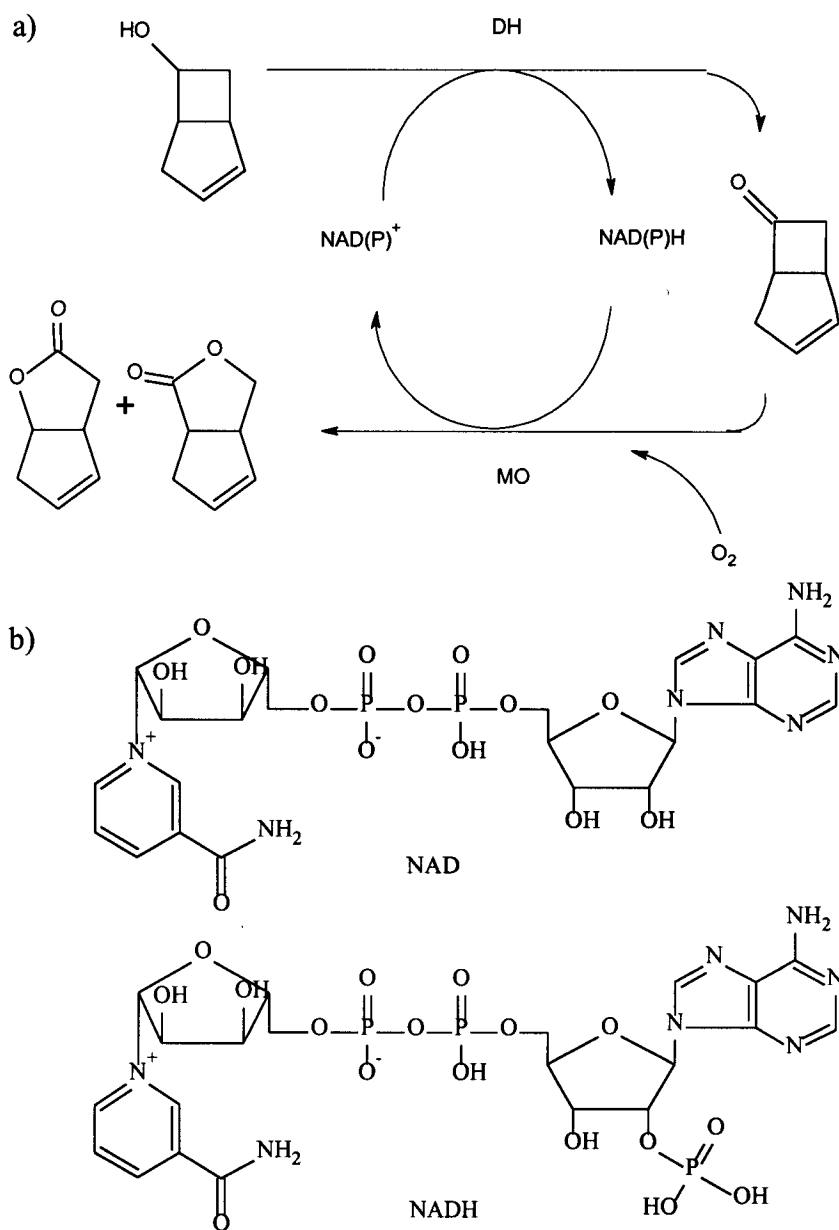


Figure 2-2

a) Intrasequential regeneration of nicotinamide nucleotide coenzymes using a coupled alcohol dehydrogenase (DH) and Baeyer-Villiger monooxygenase (MO)⁷.

b) Chemical structure of NAD and NADH

2.2 Occurrence in Nature

Microorganisms are the chief source of BVMOs in nature. Genera so far studied include *Acinetobacter*, *Pseudomonas*, *Xanthobacter*, *Rhodococcus* and *Norcardia*, and fungi of the genera *Curvularia*, *Dreschlera* and *Exophilia*. In bacteria, the natural role of these enzymes is to catalyse key steps in various catabolic oxidative pathways whereas their role in fungi is not well understood⁸.

When there is only one induced BVMO present in the microbial cell, it is possible to use washed-cell preparations of the biocatalyst to perform biotransformations. This poses a problem however, due to the presence of other enzymes in the cells such as active lactone hydrolases and esterases that prevent the build up of biotransformation products. Also whole cell preparations may have induced BVMOs of conflicting selectivities, which render the use of whole cell preparations impractical. However both these problems can be overcome by using pure enzyme preparations⁹.

2.3 Structural features of BVMOs

The best-studied BVMO to date as earlier mentioned is cyclohexanone monooxygenase from *A. Calcoaceticus*; therefore this is the model enzyme for this family of enzymes. The enzyme is monomeric with a molecular mass of 59kDa and its complete 3-dimensional structure remains to be solved. Work done by Trudgill¹⁰ concluded that the enzyme has only one active site. The enzyme has a pH optimum of 9 and utilises NADPH as coenzyme. Table 2-1 shows further characteristics of BVMOs purified from various organisms.

The complete sequence of amino acids has been determined only for CHMO ¹¹ but N-terminal sequence determination data for other BVMOs justifies the division of these flavoproteins into 2 groups: Type 1 (Figure 2-3) and Type 2 (Figure 2-4) enzymes.

Enzyme and source of micro-organism	Coenzyme	No. of proteins	Subunit composition	Molecular mass (kDa)	Mole coenzyme/ Mole protein	PH optimum
Cyclohexanone monooxygenase <i>Acinetobacter calcoaceticus</i> NCIMB 9871	NADPH	1	Single polypeptide chain	59	1 FAD	9.0
Cyclohexanone monooxygenase <i>Nocardia globerula</i> CL1	NADPH	1	Single polypeptide chain	53	1 FAD	8.4
Cyclohexanone monooxygenase <i>Xanthobacter sp</i>	NADPH	1	Single polypeptide chain	50	1 FMN	8.8
Cyclohexanone monooxygenase <i>Pseudomonas sp.</i> NCIMB 9872	NADPH	1	3-4 identical subunits	200	1 FAD	7.7
Tridecan-2-one monooxygenase <i>Pseudomonas cepacia</i>	NADPH	1	Two identical subunits	123	1 FAD	7.8-8.0
Diketocamphane monooxygenase <i>Pseudomonas putida</i> NCIMB 10007	NADPH	2	two identical subunits or one polypeptide chain	78 36	1 FMN	9.0
2-oxo- Δ^3 -4,5,5-trimethylcyclopentanyl-acetyl-Co-A monooxygenase	NADPH	1	Two identical subunits	106	1 FAD	7.8
Steroid monooxygenase <i>Cylindrocarpon radicola</i> ATCC11011	NADPH	1	Two identical subunits	155	1 FMN	8.5-9.0
Oxocineole monooxygenase <i>Pseudomonas flava</i> UQM 1742	NADPH	2	two identical subunits or one polypeptide chain	70 30		

(Taken from Willets A,1997)¹²

Table 2-1 Characteristics of Microbial BVMOs

Type 1: FAD- and NADPH- dependent

Binding motif for adenosine moiety of FAD

N-terminal.....5-25 amino acids.....G.x.G.x.x.G.....

Cyclopentanone monooxygenase from *Pseudomonas* sp. NCIMB 9871

T T M T T M T T E Q L G M N N S V N D K L D V L L I G A G

Cyclohexanone monooxygenase from *Acinetobacter calcoaceticus* NCIMB 9871

S Q K M D F D A I V I G G G F G G I

Steroid monooxygenase from *Cylindricarpon radicola* ATCC 11011

A E W A E E F D V L V V G A G A G G

Cyclohexanone monooxygenase from *Rhodococcus* sp. WT1

T A ? T I H ? V D A V V I G A G F G G I Y A V H K

2-Oxo-4,5,5-trimethylcyclopent-3-enylacetyl-CoA monooxygenase from *Pseudomonas putida* NCIMB 10007

N R A K S P A L D A V V I G A G V T G I Y Q A F L I

Figure 2-3. Sequence data for Type 1 Baeyer-Villiger monooxygenases¹²

3,6- Diketocamphane monooxygenase from *Pseudomonas putida* NCIMB 10007

A M E T G L I F H P Y M R P G R S A A Q T F D W G I K ? A

2, 5- Diketocampane monooxygenase from *Pseudomonas putida* NCIMB 10007

M Q A G F F G T P Y D I P T R T A R Q M

Luciferase (luxA) from *Vibrio harveyi* ATCC 33843

M K F G N F L L T Y Q P P E L S Q T E V M K R L V N L G

Luciferase (luxA) from *Vibrio fischeri* ATCC 7744

M K F G N I S F S Y Q P S G E

Luciferase (luxB) from *Vibrio harveyi* ATCC 33843

M K F G L F F L N F M F S K R S S

Luciferase (luxB) from *Vibrio fischeri* ATCC 7744

M K F G L F F L N F Q K D G I

Figure 2-4. Sequence data for Type 2 Baeyer-Villiger monooxygenases¹²

Type 1 enzymes are FAD binding and NADPH dependent, and both coenzymes as well as the substrates bind to the same polypeptide subunit; some of these enzymes are monomeric whilst others have a homo-oligomeric quaternary structure. Type 2 enzymes are FMN binding and NADH dependent and consist of 2 distinct types of subunits. One of these subunits binds the carbonyl substrate and the other binds the nicotinamide nucleotide, effectively serving as an NADH dehydrogenase.

An important feature of Type 1 enzymes is the presence of the characteristic GxGxxG consensus nucleotide sequence located near the N-terminus¹³. This sequence, observed in a number of FAD-binding enzymes, represents the binding site for the adenosine moiety of FAD. Immediately upstream from this consensus sequence is the invariant aspartic acid residue, which is the feature that distinguishes Type 1 BVMOs from other flavin-dependent enzymes. The functional significance of this residue remains to be established.

2.4 Synthetic applications of BVMOs

2.4.1 Enantioselective biotransformations

BVMOs display an impressive variety of nucleophilic oxygenations with great selectivity. With the exception of luciferases from *Vibrio fischeri* and *Photobacterium phosphoreum*¹⁴, all characterised BVMOs that have been examined perform an enantiodivergent biotransformation of the racemic ketone bicyclo[3.2.0]hept-2-en-6-one to yield two homochiral lactones (Figure 2-5a).

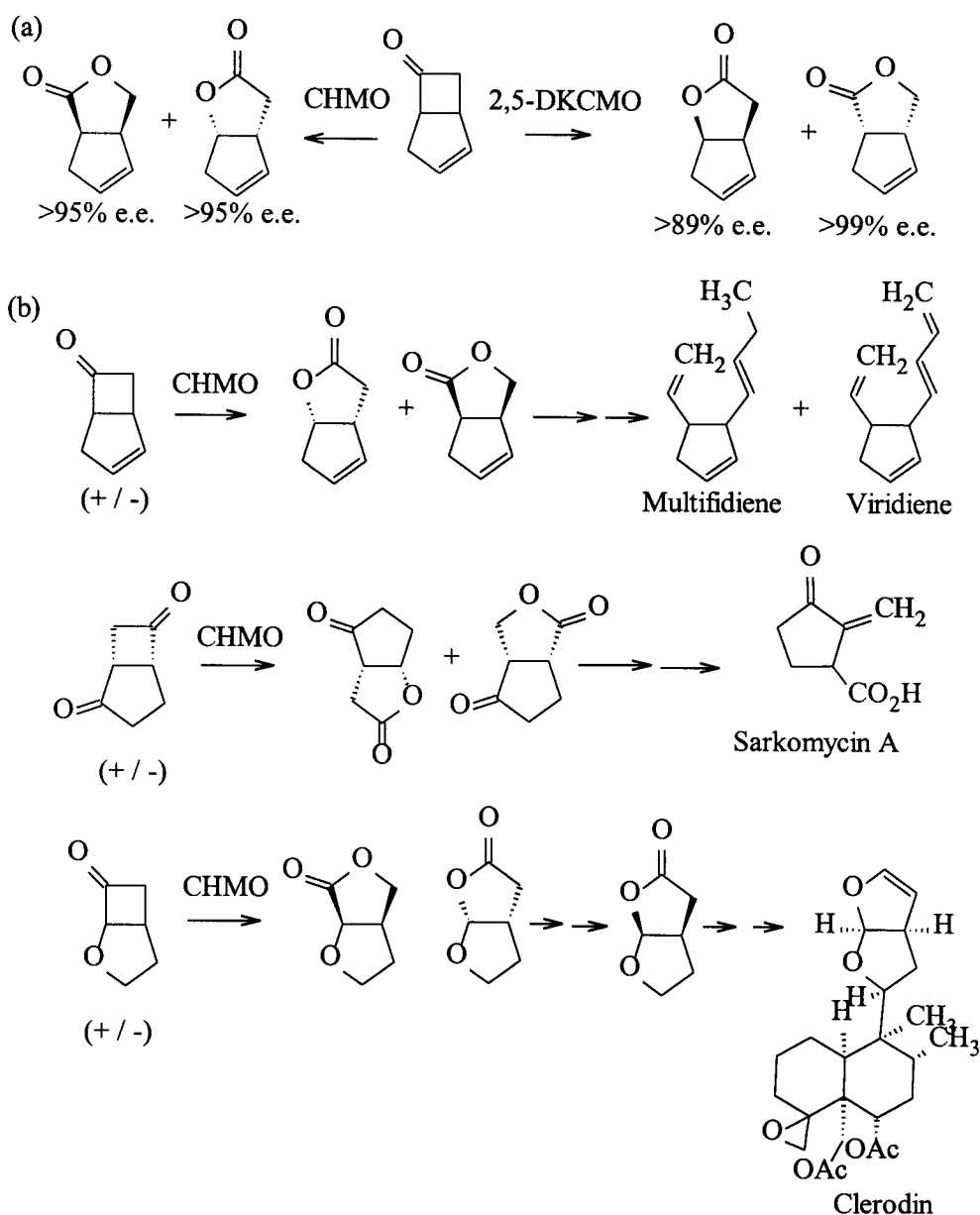


Figure 2-5. (a) Enantiodivergent biotransformation of bicyclo [3.2.0] hept-2-en-6-one by cyclohexanone monooxygenase (CHMO)¹⁵ and 2,5-diketocamphane 1,2-monooxygenase¹⁶. (b) Chemoenzymatic syntheses of various useful target molecules (multifidiene and viridiene¹⁷, sarkomycin A¹⁸ and clerodine¹⁹ involving the biotransformation by (CHMO) of bicyclo [3.2.0]-type ketones.

The significance of this finding is that it indicates that the two ketone enantiomers have been discriminated completely on the basis of regio-plus enantioselectivity (Figure 2-5a).

Such an outcome has been utilised industrially and improved to obtain all 4 possible bicyclic [3.3.0] lactones in homochiral form. These and equivalent biotransformations of various other bicyclic ketones containing furan and pyran rings have been used to yield chiral synthons that have in turn been incorporated into chemoenzymatic syntheses of useful compounds. Examples are pheromones multifidiene and viridiene, antibiotics such as sarkomycin A and the insect antifeedant clerodin (Figure 2-5b). Currently, it is apparent that such biotransformations are an extremely useful way of achieving asymmetric Baeyer-Villiger oxidations.

2.4.2 Kinetic resolution as a synthetic application

Using racemic substrates such as 2-bromo-3-hydroxy-substituted [3.2.0] ketones, various substituted [2.2.1] ketones²⁰ and 2'-substituted monocyclic ketones, CHMO (Figure 2-6a), and the two DKCMO isozymes (Figure 2-6b), perform very different types of regio-plus enantioselective biotransformations; each example yielding by kinetic resolution a single homochiral lactone. It has been observed that changing the nature or the length of the 2'-side-chain have significant effects on the outcome and yield of these biotransformations. Biotransformations of this type have been utilised to obtain key synthons in the synthesis of targets such as the potent antiviral carbocyclic nucleosides²¹.

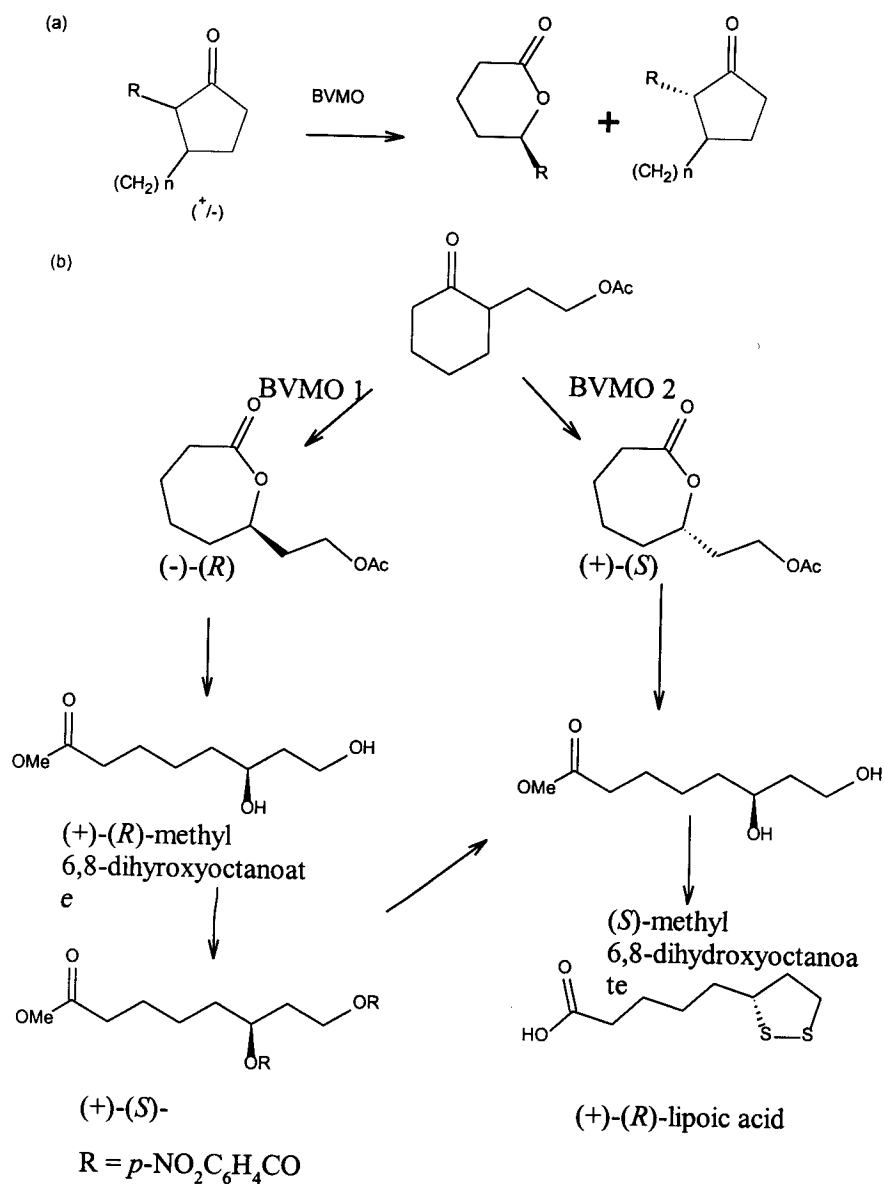


Figure 2-6. (a) Kinetic resolution of 2'-substituted monocyclic ketones by Baeyer-Villiger monooxygenases. (b) Chemoenzymatic synthesis of (*R*)-(+)-lipoic acid by either 2-oxo-4, 5,5-trimethylcyclopent-3-enylacetyl-CoA monooxygenase (BVMO 1²²) or cyclopentanone monooxygenase (BVMO 2²³)

2.5 The Structure of the Active site of Baeyer-Villiger monooxygenases

2.5.1 Active site modelling of BVMOs

Understanding the stereoselectivity of BVMOs is paramount in making predictions of likely products from untested substrates that are of commercial interest. To achieve this, it is necessary to know the three-dimensional structure of biocatalysts of commercial interest such as CHMO and DKCMO isozymes.

Joint programs to achieve this end have resulted in increased information about the three-dimensional structure. However, in the absence of full structures, models have been developed to explain the outcome of biotransformations performed by CHMO and DKCMO. Two approaches to achieve this are outlined below.

2.5.2 'Cubic space' BVMO models

In this approach, it has been assumed that a combination of stereochemical and stereoelectronic effects determines the outcome of events in the active site²⁴. Alphand proposed a model (Figure 2-7) based on a regular cubic spacing in which 'forbidden zones' exist. These forbidden zones represent regions of steric hindrance. The limits of the model were established by theoretically testing the positions that may be taken up by the intermediates in their transition state, during a biotransformation and matching them with the actual outcome.

The model of the CHMO active site proposed by Ottolina (Figure 2-8) was produced by using an approach similar to the 'Cubic space' approach used by Alphand¹. The

outcomes of both sulfoxidation and lactonisation biotransformations are only explained by this model.

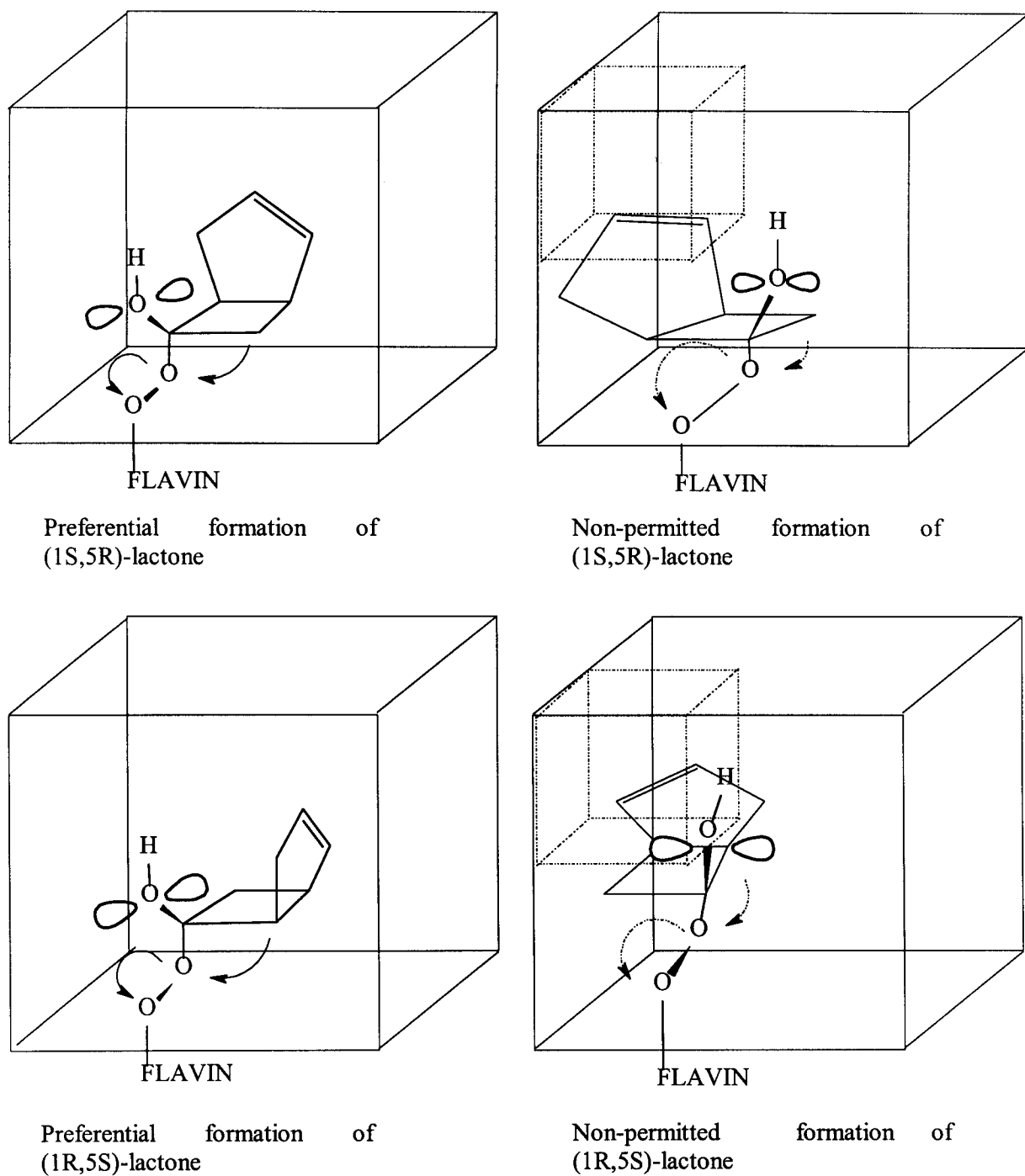


Figure 2-7 Active site model of cyclohexanone monooxygenase (CHMO) showing 'Cubic space' model proposed by Alphand and Furtoss¹. The 'forbidden zone' is indicated by dotted lines.

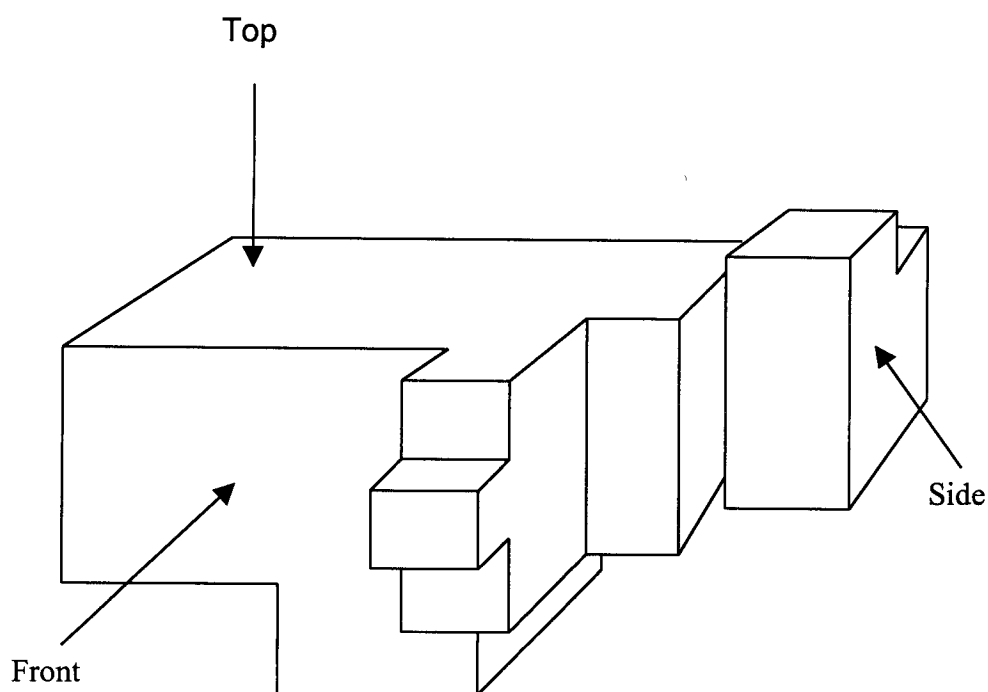
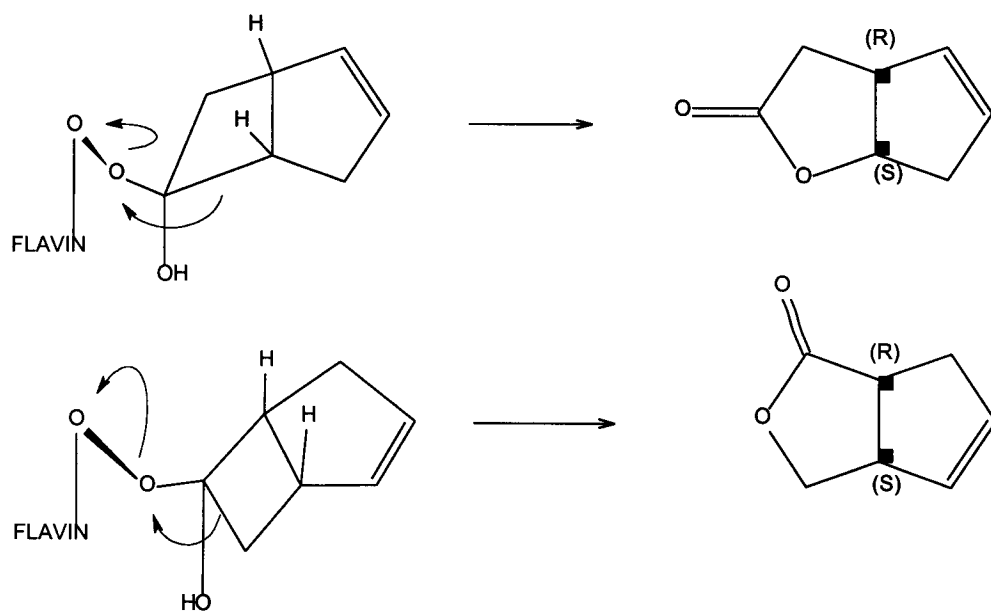


Figure 2-8 Active site model of cyclohexanone monooxygenase (CHMO) showing 'Cubic space' model proposed by Ottolina^{25,26,27}.

2.5.3 Mechanism-based models of BVMOs

Taschner used a mechanism-based approach to propose a model for the CHMO active site⁵. He pointed out that human and *E. coli* glutathione reductase are likely to have similar active sites. Highly conserved N-terminal sequences were observed in all three proteins and a similar BAB-domain containing the GxGxxG FAD binding motif⁴³. Starting from this point, Taschner developed a model that explained the observed outcomes of different *meso* ketones (Figure 2-9). An important factor in this model is that the hydroperoxide of the oxygenated FAD intermediate is assumed to be attached to the *re*-face of the isoalloxane ring of the flavin cofactor. Equally important is the approach of the ketone substrate towards the hydroperoxide from the dimethylbenzene moiety's direction.

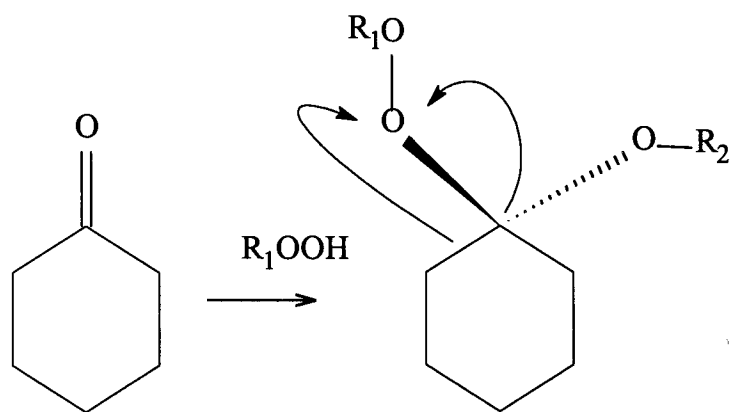


Configurational correlation to explain the biotransformation of bicyclo [3.2.0] hept-2-en-6-one by CHMO

Figure 2-9 Mechanism-based model proposed by Taschner⁵

Kelly proposed an alternative mechanism-based model (Figure 2-10) using data from the outcome of the biotransformation of an unsubstituted tricyclic ketone by CHMO²⁸. Enantioselectivity is proposed to be controlled by stereoelectronic factors, which constitute the main difference between this model and previous models. This model departs from Taschner's model in that the hydroxyperoxide in the transition state intermediate formed in the CHMO active site is attached to the *si*-face of the isoalloxane ring of the co-factor (FAD). Also, the ketone ring approaches the hydroperoxide from the direction of the diazine moiety.

Although Kelly's model explains the outcomes from tricyclic ketones, it fails to explain the outcomes of sulfoxidation of these enzymes²⁹ and why CHMO and DKCMO isozymes biotransform substituted cyclobutanones with the same enantioselectivity³⁰.



R_1 =flavin coenzyme
 R_2 =H

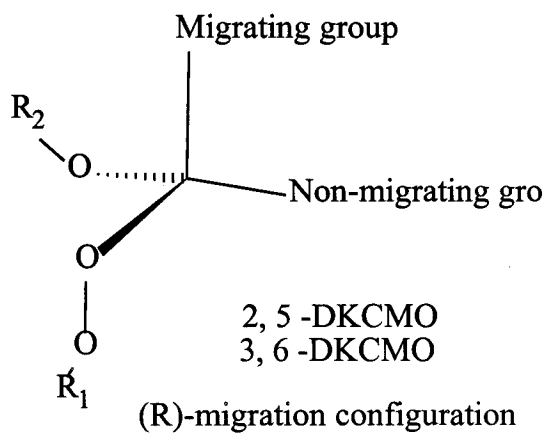
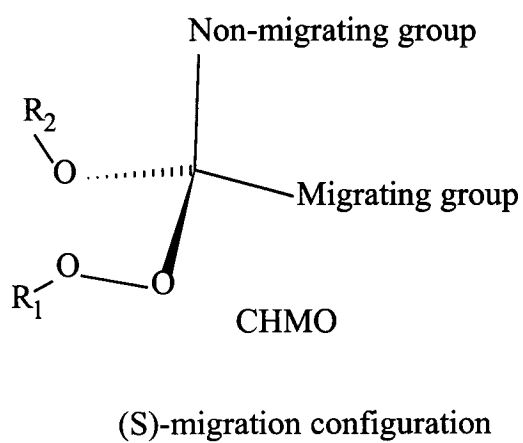


Figure 2-10 Mechanism-based model proposed by Kelly^{28,29}.

2.6 Enzymes

Enzymes are the catalysts, which make possible biochemical reactions³¹. Without enzymes, our body chemistry would not occur, and life would not exist. This illustrates the impressive power of enzymes as catalysts. They are also very specific. Amylase, a digestive enzyme, will hydrolyse starch, but not cellulose. Both molecules are polymers of glucose. They differ in the orientation of one bond at the junction of glucose units. Other enzymes can work effectively on a broader range of substrates

This broader specificity is useful in the case of an enzyme like pepsin, which is important in protein digestion. It can catalyse the hydrolysis of peptide (amide) bonds in a variety of proteins, which means that the body does not need to maintain a stock of more specific enzymes to tackle specific proteins.

These enzyme characteristics can be explained by their structure. Enzymes are almost all proteins. They are often globular proteins. Thus they can be described in terms of their primary, secondary, tertiary, and in many cases, quaternary structure. They are long chains of amino acid units held together by peptide bonds, looped and folded into secondary and tertiary (and often quaternary) structures by disulfide bonds, hydrophobic interactions, and salt bridges.

In addition, active enzymes usually involve cofactors. These are small molecules (sometimes inorganic ions) that are needed to complete the catalytically active structure of the enzyme. In such instances, the enzyme without the cofactor is called an apoenzyme, and the apoenzyme-cofactor complex is called a holoenzyme. The

protein chain of an apoenzyme can have functional groups on its side chains (R groups) that are important to its catalytic function, but other important functional groups are introduced by way of cofactors.

2.7 Classification of enzymes

Enzymes are classified according to the reactions they catalyze. In some cases, the terms used are fairly clear; in others, less so.

2.7.1 Oxidoreductases:

These are enzymes that catalyse the reduction or oxidation of a molecule. Any enzyme which catalyses a reduction have to also catalyse the reverse (oxidation) reaction, thus the double-barreled name "oxidoreductase."

2.7.2 Transferases:

These enzymes catalyse the transfer of a group of atoms from one molecule to another. A common example involves transfer of a phosphate between ATP and a sugar molecule.

2.7.3 Hydrolases:

These enzymes catalyse hydrolysis reactions. The hydrolysis of an ester would be an example of such a reaction.

2.7.4 Isomerases:

These enzymes catalyse the conversion of a molecule into an isomer. The *cis-trans* interconversion of maleate and fumarate is an example.

2.7.5 Lyases

Reactions that add a small molecule such as water or ammonia to a double bond (and the reverse, elimination, reactions) are catalysed by lyases.

2.7.6 Ligases

These enzymes catalyse reactions that make bonds to join together (ligate) smaller molecules to make larger ones.

2.8 Oxidations

Oxidation of aromatic and alicyclic molecules using monooxygenase and dioxygenase enzymes has attracted a considerable amount of attention from industries and scientists.

2.8.1 Dioxygenases

Various cycloaddition reactions can be carried out on the acetone [Figure 2-11(a)], including Diels-Alder [4+2] reactions to give a tricyclic compound (Figure 2-11(b)), [2+2] cycloaddition reactions to give cyclobutanone derivatives [Figure 2-11(c)]³² and the more unusual [6+4] cycloaddition reactions^{33, 34}. Not only is benzene biotransformed by *Pseudomonads* but so are many of its derivatives. Often the derivatives of benzene are obtained in an optically pure form and it has been observed that these diols have the 1(*S*), 2(*R*)-configuration. Trifluoromethylbenzene is biotransformed into the diol [Figure 2-12 (a)]; the corresponding acetone [Figure 2-12 (b)] can be converted into a wide range of derivatives through different chemical reactions^{35,36}. Interesting conversions involving 3-chlorocyclohexa-3, 5-diene-1, 2-diol and 3-methylcyclohexa-3, 5-diene-1,2-diol have also been described^{37, 38}.

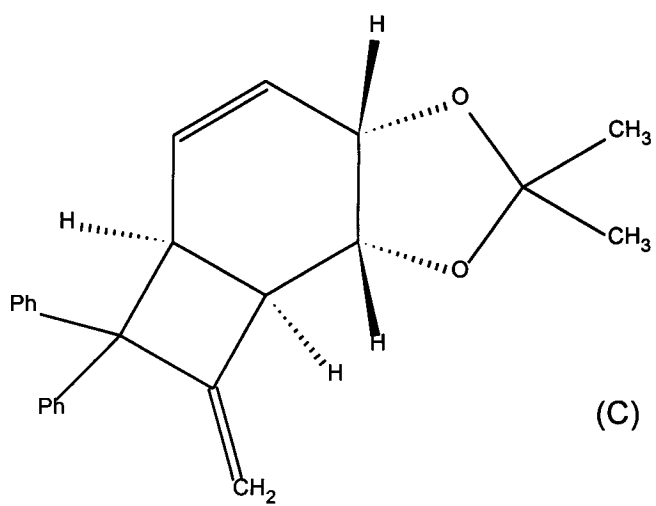
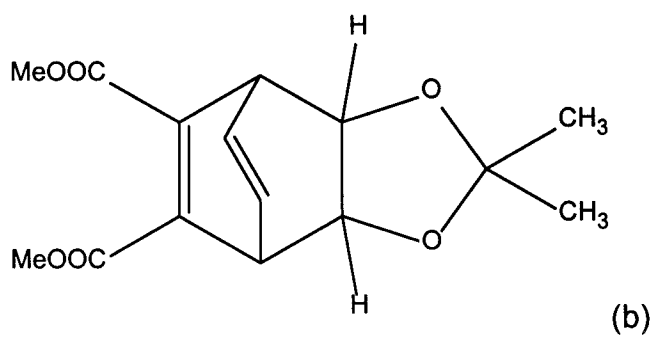
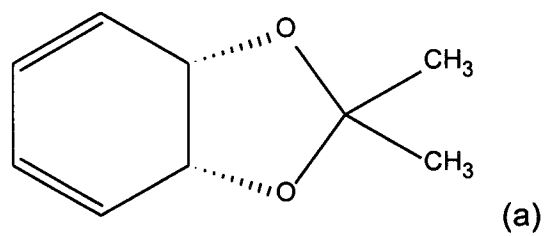


Figure 2-11 a) An acetonide. (b) A tricyclic product of Diels-Alder reaction. (c) A cyclobutanone derivative as a product of a cycloaddition.

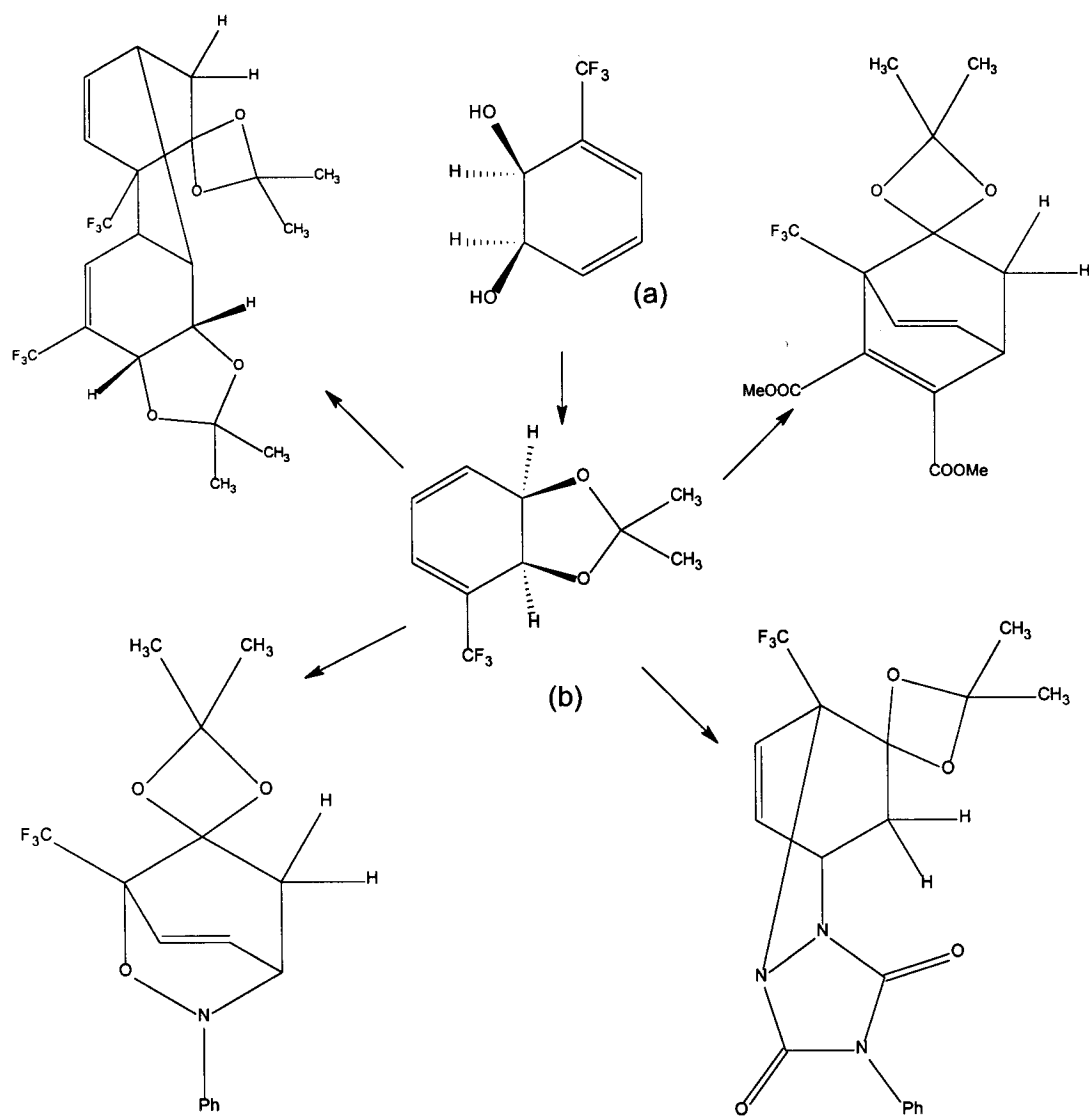


Figure 2-12 (a) The diol from trifluoromethylbenzene. (b) The corresponding acetone (obtained from trifluoromethylbenzene).

2.8.2 Monooxygenases

An oxidation reaction that is gaining an increasing amount of interest is the enzyme-catalysed conversion of a cyclic ketone into a lactone³⁹. It has been shown that (+/-)-7-anti-fluoro-5-endo-bromobicyclo[2.2.1]heptan-2-one is oxidized in an enantioselective manner using the *Acinetobacter* NCIMB 9871 to give optically pure lactone (Figure 2-13a) (30%) and it has been possible to recover optically active ketone (Figure 2-13b) (30%). Interestingly oxidation of the ketone (Figure 2-13b) with the chemical oxidant, meta-chloroperoxybenzoic acid gave the 3-oxabicyclooctan-2-one (Figure 2-13c) as the major product⁴⁰. The latter compound was converted into the carbocyclic nucleoside (Figure 2-13d), a compound possessing anti-HIV activity^{41, 42}.

The monooxygenase from the *Acinetobacter* can be purified and Baeyer-Villiger oxidations can be performed using the enzyme in association with its cofactor NADPH. It has been found that it is possible to convert a secondary alcohol into the corresponding lactone via the corresponding ketone by using a coupled enzyme system, which is *Thermoanaerobium brockii* dehydrogenase and *A. calcoaceticus* monooxygenase with *in situ* recycling of the cofactor⁴³. More recently a monooxygenase enzyme capable of using NADH as cofactor has been found, enabling cofactor recycling to be accomplished using formate dehydrogenase^{44, 45}.

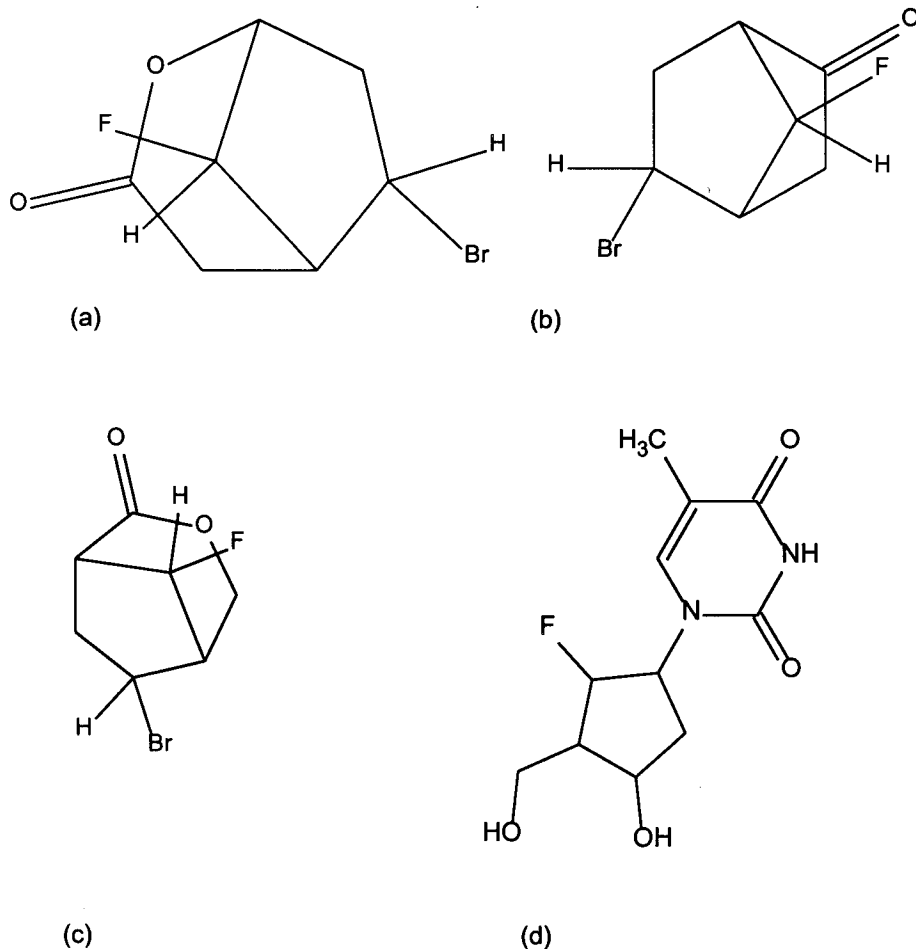


Figure 2-13. Compounds of interest formed by the action of monooxygenases.

- (a) Lactone formed from oxidation of (+/-)-7-anti-fluoro-5-endo-bromobicyclo [2.2.1] heptan-2-one by *Acinetobacter* NCIMB 9871.**
- (b) Optically active ketone recovered at 30% from *Acinetobacter* NCIMB 9871.**
- (c) 3-oxabicyclooctan-2-one produced from oxidation of ketone (b) with chemical oxidant meta-chloroperoxybenzoic acid as the major product.**
- (d) A carbocyclic nucleoside, obtained from conversion of 3-oxabicyclooctan-2-one, which has been shown to have HIV activity.**

Immobilised soya bean lipoxygenase-1 efficiently converts linoleic acid into 13(*S*)-hydroxyoctadecadienoic acid (HODE) with 78% yield in a few hours. The immobilised enzyme can be recovered and reused. The chemistry of 13-HODE has been investigated: For example, the bicyclic compound (Figure 2-14) has been obtained by macro-lactonisation and a Diels-Alder reaction on the biotransformation product⁴⁶.

The ability to introduce a desired functionality in an organic molecule at a position distant from pre-existing functionality is an attractive goal for both organic chemists and microbiologists alike. Some progress has been made towards providing chemical methods for remote hydroxylation⁴⁷ but, for the foreseeable future, microbiological methods seem to offer the most likely chances for success. The regioselective hydroxylation of steroids has been investigated in some depth, the formation of 11-hydroxyprogesterone from progesterone being one of the most noteworthy successes. The regio-selective (but not regiospecific) oxidation of cyclohexylcyclohexane to give the diol [Figure 2-15(a)] using *Cunninghamella blakesleeana* shows that even the most hydrophobic materials can be biotransformed⁴⁸. Spiro-heterocyclic compounds can allow the formation of useful derivatives, often in very high yield, and produce single compounds using *Beauveria sulfurescens*: for example the piperidine derivative [Figure 2-15(b)] furnishes the alcohol [Figure 2-15(c)] in 70% yield⁴⁹.

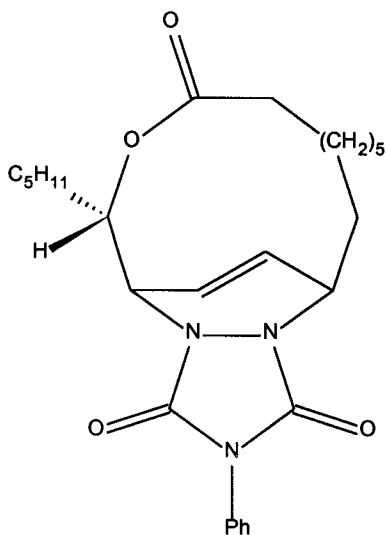
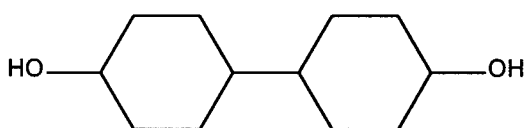
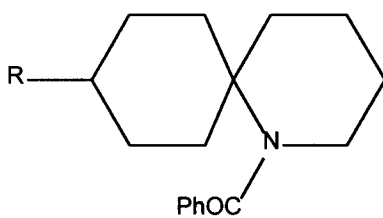


Figure 2-14. A bicyclic compound obtained by macro-lactonisation followed by a Diels-Alder reaction on the biotransformation product.



(a)



(b)

Figure 2-15. (a) Oxidation of cyclohexylcyclohexane to give the diol by *Cunninghamella blakesleeana*.

(b) When R= H, the piperidine derivative can be derivatised by *Beauveria sulfurescens* to yield the alcohol (where R = OH) in 70 % yield.

2.9 Classification of Flavoproteins

Flavoenzymes can be grouped into a relatively small number of classes where members of the same group share common properties differing from those of other classes such as the ability to use molecular oxygen as an electron acceptor, the nature of the auxiliary redox centres and the types of reactions catalysed.

2.9.1 Simple flavoproteins

2.9.1.1 Oxidases

Simple flavoproteins are classified on the basis of reactivity of the reduced enzyme with molecular oxygen. The oxidases such as glucose oxidase and D-amino acid oxidase react rapidly with oxygen to yield peroxide and oxidised flavoprotein in what has been shown to be a second order process without observable intermediates⁵⁰. The members of this class display the following properties:

- a) They all stabilise the red ionic flavin radical on one-electron reduction
- b) they all stabilise a flavin N (5)-sulfite adduct and the benzoquinoid anion form of 6- and 8-substituted hydroxy- and mercaptoflavins where the negative charge of the anionic flavin is localised in the N (1)-C(2)O region⁵¹.

These properties are interpreted as indicating the presence of a positively charged locus of the protein interacting with the pyrimidine ring of the flavin. This theory is supported by X-ray crystallography structures of flavoprotein oxidases that have been solved in recent years, e.g. glycolate oxidase⁵², glucose oxidase⁵³ and cholesterol oxidase⁵⁴. Glycolate oxidase shows the expected positively charged residue Lys230, and both glucose oxidase and cholesterol oxidase have histidine residues in the vicinity of the flavin pyrimidine ring. Solution studies of oxidases, where the native flavin has been replaced by flavins with reactive substituents showed in most cases

that the flavin was buried well inside the intact protein with only the region around the N (5)-position accessible to solvent⁵¹.

2.9.1.2 *Electron Transferases*

Enzymes of this class are all involved physiologically in single electron transfers. Examples are flavodoxin, ferredoxin-NADP reductase and NADPH-cytochrome P-450. In contrast to the oxidases, this group shows no tendency to stabilise the benzoquinoid forms of 6- and 8-substituted flavins. They also show no tendency to form flavin N(5)- sulfite adducts. Studies of solutions with chemically reactive flavins (e.g. 8-Cl-, 8-mercapto-, or 8-SCN-flavins) showed that the only part of the molecule accessible to the solvent is the benzene ring of the flavin⁵¹. Crystal structures available for flavodoxin⁵⁵ and ferredoxin-NADP⁺ reductase⁵⁶ support these conclusions.

2.9.1.3 *Flavoprotein monooxygenases*

In this class, the reduced enzyme reacts with O₂ to produce C (4) hydroperoxide intermediates. Enzymes in this class have NADH or NADPH as the physiological reductant. Examples of flavoprotein monooxygenases are bacterial luciferase and *p*-hydroxybenzoate hydroxylase.

2.9.2 Flavoproteins with Auxiliary redox centres

2.9.2.1 *Flavoprotein disulphide oxidoreductases*

Enzymes of this class contain a disulphide in close juxtaposition to the flavin, which in all cases is a FAD. Catalysis involves interaction of the flavin with a pyrimidine nucleotide, electron transfer through the flavin to the redox active disulphide/dithiol, and interaction of the latter with the second substrate⁵⁷. Examples of the

disulphide/dithiol second substrate are a lipoic acid derivative in lipoamide dehydrogenase, glutathione in glutathione reductase and thioredoxin in thioredoxin reductase.

2.9.2.2 *Heme-containing flavoproteins*

Enzymes belonging to this class are known as flavocytochromes. A well-studied example of a flavocytochrome is yeast lactate dehydrogenase (flavocytochrome b₂). This enzyme has a b-type cytochrome domain in addition to the FMN prosthetic group. Another example of a heme-containing flavoenzyme is the fusion complex of cytochrome P-450 BM-3.

2.9.2.3 *Metal-containing flavoproteins*

Xanthine oxidase is the most well known metal-containing flavoenzyme. This complex flavoenzyme has molybdenum and iron-sulphur centres as well as the FAD prosthetic group. Other examples of metal-containing flavoproteins are bacterial enzymes involved in the degradation of inactivated aromatic compounds such as benzene and catechol.

2.10 Substrate Specificity of *Arthrobacter* sp. M5

Previous studies of the substrate specificity of *P. fluorescens* sp. ACB and *Arthrobacter* sp. M5 with mono-substituted acetophenones and mono-substituted hydroxyacetophenones have revealed that these organisms have remarkably divergent substrate specificities⁵⁸. In the case of ACB, ortho- and meta- substitutions usually lead to a diminished activity compared to acetophenone, but the polar substituents hydroxy and amino on the ortho position give better substrates in the case of 4-chloro and 4-methylacetophenone (Figure 2-16). Para-substituent activities are summarised in (Figure 2-17). Trends are not evident. Halogen substituents do not show a trend

with respect to their size. Methoxy- is better than ethoxy- substitution but the reverse is true for their respective alkyl derivatives.

M5 shows generally broader substrate specificity in comparison with ACB. M5 has a notably low activity for 2-hydroxyacetophenone (the model substrate for ACB) and better activity for chlorinated and methylated acetophenones as (Figure 2-18) shows. Increasing the number of substituents on the ring leads to poorer substrates; chlorinated substrates are not accepted whereas methylated ones are. 4-hydroxy-3-methylacetophenone provided the best results.

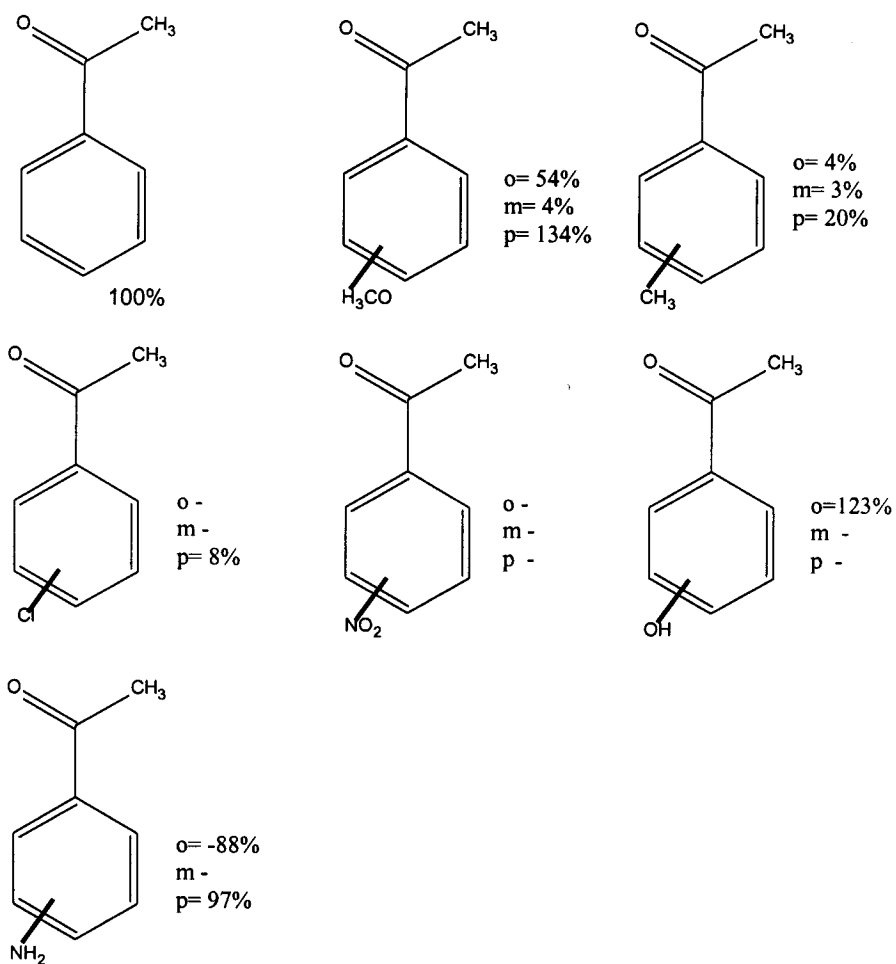


Figure 2-16. Substrate specificity of *Pseudomonas* ACB showing effect of para-, meta- and ortho- substitutions⁵⁸.

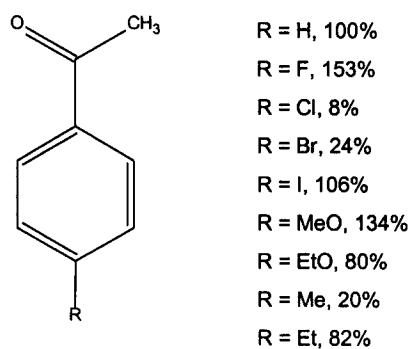


Figure 2-17 Substrate specificity of *Pseudomonas* ACB showing para-substituent activities⁵⁸.

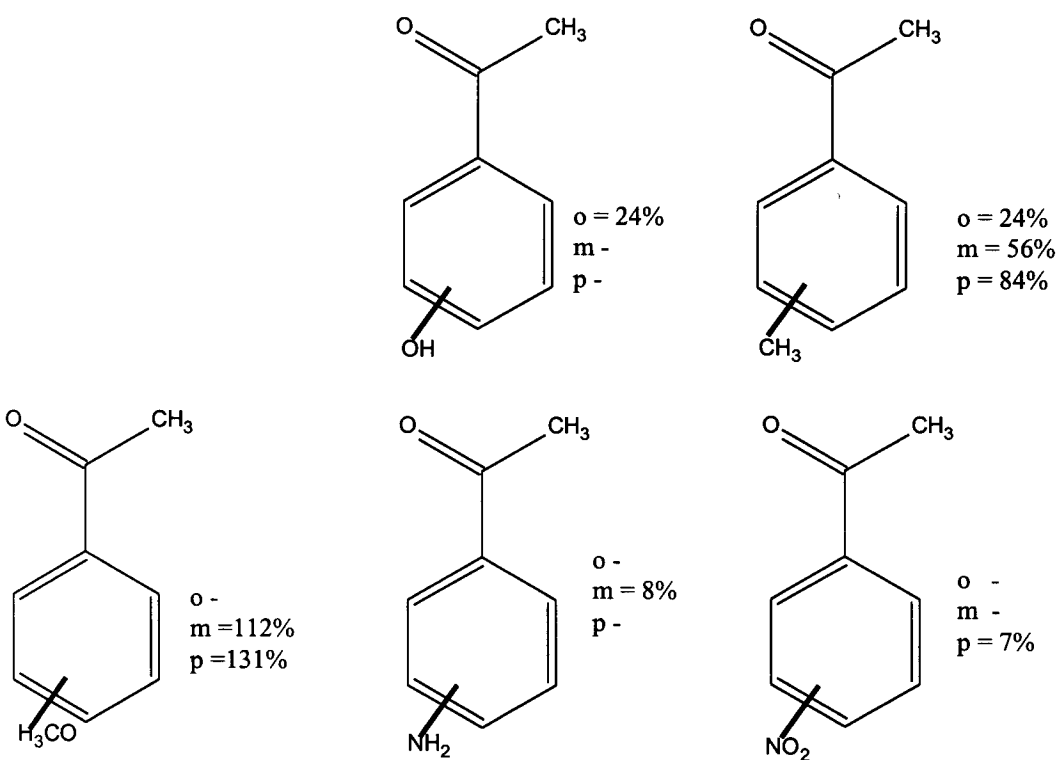


Figure 2-18 *Arthrobacter* sp. M5 shows low activity for 2-hydroxyacetophenone and better activity for chlorinated and methylated acetophenones⁵⁸.

2.11 DNA manipulation protocols

2.11.1 Principle of the DNA packaging kit

The DNA packaging kit consists of two combined extracts. The lysates, sonic extract (SE) and freeze thaw lysate (FTL) are prepared from two *E coli* lysogens which have complimentary mutations in genes required for the phage maturation pathway. The SE extract in the DNA packaging kit is prepared from the lysogen strain *E coli* NM759 which harbours an amber mutation in the gene of head component D leading to an accumulation of phage preheads. In addition this strain lacks restriction systems for methylated DNA.

The FTL extract is prepared from a lysogen mutated in the gene for head component E. The extract is complimentary to SE and provides gene D product and other components for phage maturation including phage tails.

Both lysogens contain an amber mutation in the lysis gene S to prevent lysis of host cells, and in the b2 deletion to prevent excision of the prophage after induction. The red 3 mutation present in the lambda genome in conjunction with recA mutation in the host prevent recombination. The extracts lack the deleterious Mcr-systems, which are known to restrict DNA sequences containing cytosine methylation. The Mrr system that recognise adenine methylation and HsdR system have also been deleted. When the appropriate exogenous DNA containing *cos* sites is added to the predisposed mixture of both extracts, the different components compliment each other. Single monomeric molecules are cleaved from the concatemers and packaged into mature phage particles, which can infect *E. coli* hosts with high efficiency.

2.11.2 Packaging substrates:

Optimal packaging of recombinant DNA is dependent on some special features of the substrate DNA. The DNA molecule must contain two cos sites at a distance of 38-52 kb apart, that is 75-105% of the length of the wild type genome. Monomeric molecules require a λ cos site at both ends.

Circular multimers as well as linear multimers are efficient packaging substrates; whereas circular monomers are packaged with poor efficiency. Therefore ligations of DNA to be packed *in vitro* should be carried out under conditions, which favour the formation of concatemers and minimize the formation of monomeric circular molecules (DNA concentrations greater than 0.2 $\mu\text{g}/\mu\text{l}$).

2.11.3 The Polymerase Chain Reaction⁵⁹

The chemistry of the polymerase chain reaction (PCR) depends on the complementarity of the DNA bases. When a molecule of DNA is sufficiently heated, the hydrogen bonds holding together the double helix are disrupted and the molecule unzips or "denatures" into single strands. If the DNA solution is allowed to cool, then complementary base pairs can reform (renature) and the original double helix is restored.

The polymerase chain reaction is a test tube system for DNA replication that allows a "target" DNA sequence to be selectively amplified, or enriched, several million-fold in just a few hours (Figure 2-19). Within a dividing cell, DNA replication involves a series of enzyme-mediated reactions, whose end result is a faithful copy of the entire

Polymerase Chain Reaction

There are 30-40 cycles of three steps(denaturation, annealing and extension):

Step 1. Denaturation of DNA for 1 minute at 94 °C

Step 2. Annealing of primers to target sequences in the desired direction: 45 seconds at 54 °C

Step 4. Extention of primers catalysed by Taq polymerase as it utilises dNTP's availabe in reaction medium: 2 minutes at 72 °C

Exponential amplification of a gene by PCR

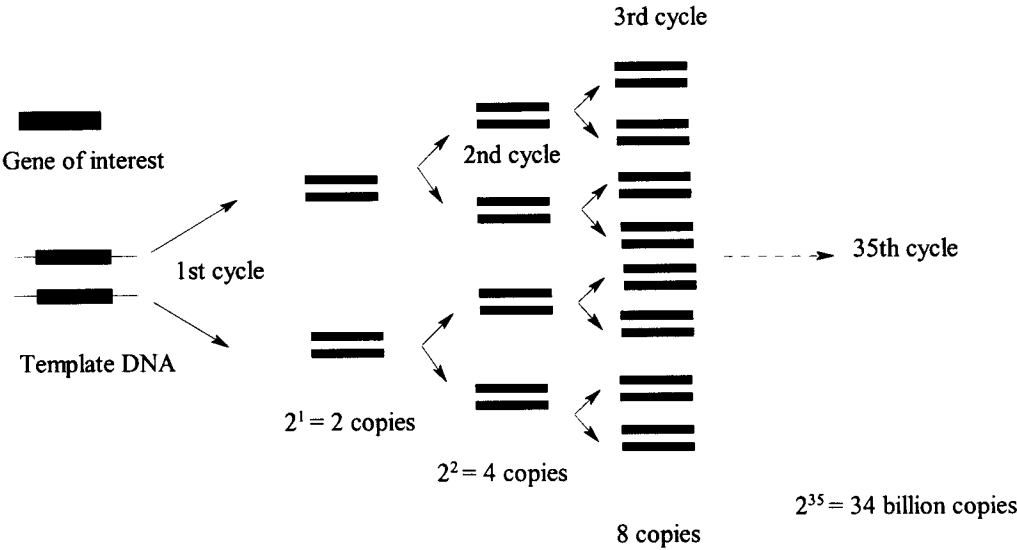


Figure 2-19 The Polymerase Chain Reaction⁵⁹

genome. Within a test tube, PCR uses just one indispensable enzyme - DNA polymerase - to amplify a specific fraction of the genome.

During cellular DNA replication, enzymes first unwind and denature the DNA double helix into single strands. Then, RNA polymerase synthesizes a short stretch of RNA complementary to one of the DNA strands at the start site of replication. This DNA/RNA heteroduplex acts as a priming site for the attachment of the DNA polymerase, which then produces the complementary DNA strand. During PCR, high temperature is used to separate the DNA molecules into single strands, and synthetic sequences of single-stranded DNA (20-30 nucleotides) serve as primers. Two different primer sequences are used to bracket the target region to be amplified. One primer is complementary to one DNA strand at the beginning of the target region; a second primer is complementary to a sequence on the opposite DNA strand at the end of the target region.

To perform a PCR reaction, a small quantity of the target DNA is added to a test tube with a buffered solution containing DNA polymerase, 2 oligonucleotide primers, the four deoxynucleotide building blocks of DNA, and the cofactor MgCl_2 . The PCR mixture is taken through replication cycles consisting of:

- a) One to several minutes at 94-96 °C, during which the DNA is denatured into single strands,
- b) One to several minutes at 50-65 °C, during which the primers hybridize or "anneal" (by way of hydrogen bonds) to their complementary sequences on either side of the target sequence,
- c) One to several minutes at 72 °C, during which the polymerase binds and extends a complementary DNA strand from each primer.

As amplification proceeds, the DNA sequence between the primers doubles after each cycle. Following thirty such cycles, a theoretical amplification factor of one billion is attained.

Two important innovations were responsible for automating PCR. First, a heat-stable DNA polymerase was isolated from the bacterium *Thermus aquaticus*, which inhabits *hot springs*. *This enzyme, called the Taq polymerase, remains active despite repeated heating during many cycles of amplification.* Second, DNA thermal cyclers were invented that use a computer to control the repetitive temperature changes required for PCR.

Following amplification, the PCR products are usually loaded into wells of an agarose gel and electrophoresed. Since PCR amplifications can generate microgram quantities of product, amplified fragments can be visualized easily following staining with a chemical stain such as ethidium bromide. While such amplifications are impressive, the importance of this method is that the amplification is selective - only the DNA sequence located between the primers is amplified exponentially. The rest of the DNA in the genome is not amplified and remains invisible in the gel.

2.11.3.1 Applications of PCR⁶⁰

PCR has proven a quick, reliable method for detecting all manner of mutations associated with genetic disease - from insertions, to deletions, to point mutations. Duchenne muscular dystrophy is an example of a genetic disease whose detection has been greatly simplified by the use of PCR. The human dystrophin gene, spread out over two million base pairs of DNA on the X chromosome, is the largest gene identified to date.

Boys afflicted with Duchenne muscular dystrophy have deletions in the protein coding regions (exons) of the dystrophin gene. The unusually large size of the gene makes it impractical to examine its entire length for mutations, so a technique called "multiplex PCR" is used to sample various regions of the gene from one end to the other. The technique involves simultaneous amplification from nine different sets of primers, all within the same test tube. Each set of primers is chosen to produce a different-sized amplification product from a different region of the dystrophin gene. Following amplification, the PCR products are analyzed by gel electrophoresis. Boys having a normal dystrophin gene will display nine different-sized amplification

products, while boys with deletions in the gene will be missing one or more of the PCR products.

PCR can also be used to detect the presence of unwanted genetic material, as in the case of a bacterial or viral infection. Conventional tests that involve the culture of microorganisms or use of antibodies can take weeks to complete or be tedious to perform. PCR offers a fast and simple alternative. For example, in the diagnosis of Acquired Immune Deficiency Syndrome (AIDS), PCR can be used to detect the small percentage of cells infected by the human immunodeficiency virus (HIV). DNA isolated from peripheral blood cells is added to a PCR reaction containing primers complementary to DNA sequences specific to HIV. Following amplification and gel electrophoresis, the presence of an appropriate-sized PCR product indicates the presence of HIV sequence and therefore, HIV infection.

The sensitivity of PCR is so great that signals may be obtained from degraded DNA samples and sometimes from individual cells. This ability and the inherent stability of DNA have combined to permit DNA to be amplified from some unusual sources, such as an extinct mammal called the quaga, an Egyptian mummy, and a three-million-year-old termite trapped in amber. This situation has, almost overnight, transformed ignored museum collections of biological specimens into treasure troves of genetic information. Evolutionary biologists are using these specimens and PCR to explore the genetic relatedness of organisms across species boundaries and now even across time.

When PCR is used with degraded DNA samples, it can synthesize an amplification product, even if the sample's average fragment size is less than the distance between the primer binding sites. During PCR, overlapping fragments within the target sequence can function as primers to generate full-length amplification products. This ability of PCR to utilize degraded DNA samples is of great interest to forensic scientists who must sometimes work with human cells in very poor condition. The technique has provided conclusive identifications in cases where conventional DNA typing has failed. Ironically, the greatest concern about the widespread use of PCR in forensic medicine is the technique's extreme sensitivity. Even miniscule amounts of DNA left over from previous amplifications can be reamplified, leading to an inconclusive result

2.11.3.2 Site-Directed Mutagenesis using PCR

In vitro site-directed mutagenesis is an invaluable technique for studying protein structure-function relationships, gene expression and vector modification⁶¹. The literature shows several methods, but many of these methods require single-stranded DNA as the template. The reason for this, historically, has been the need for separating the complementary strands to prevent reannealing. Use of PCR in site-directed mutagenesis accomplishes strand separation by using a denaturing step to separate the complementing strands and allowing efficient polymerization of the PCR primers. PCR site-directed methods thus allow site-specific mutations to be incorporated in virtually any double-stranded plasmid, eliminating the need for M13-based vectors or single-stranded rescue.

Several points should be mentioned concerning site-directed mutagenesis using PCR. First, it is often desirable to reduce the number of cycles during PCR when

performing PCR-based site-directed mutagenesis to prevent clonal expansion of any (undesired) second-site mutations. Limited cycling, which would result in reduced product yield, is offset by increasing the starting template concentration. Second, a selection must be used to reduce the number of parental molecules coming through the reaction. Third, in order to use a single PCR primer set, it is desirable to optimize the long PCR method. And fourth, because of the extended activity of some thermostable polymerases it is often necessary to incorporate an end-polishing step into the procedure prior to end-to-end ligation of the PCR-generated product containing the incorporated mutations in one or both PCR primers. The protocol used in this study provided an easy method for site-directed mutagenesis and accomplishes the above desired features by the incorporation of the following steps:

- a) Increase of template concentration to approximately 1000-fold over conventional PCR conditions
- b) *Reduction of the number of cycles from 25-30 to 5-10*
- c) Addition of the restriction endonuclease *DpnI* (recognition target sequence: 5-Gm6ATC-3, where the A residue is methylated) to select against parental DNA (DNA isolated from almost all common strains of *E. coli* is *Dam*-methylated at the sequence 5-GATC- 3)
- d) Utilisation of Taq Extender in the PCR mix for increased reliability for PCR to 10 kb
- e) Utilisation of *Pfu* DNA polymerase to polish the ends of the PCR product
- f) Efficient intramolecular ligation in the presence of T4 DNA ligase.

2.11.4 Restriction enzymes⁶²

Restriction enzymes are DNA-cutting enzymes found in bacteria (and harvested from them for use). Because they cut within the molecule, they are often called restriction endonucleases.

A restriction enzyme recognizes and cuts DNA only at a particular sequence of nucleotides. For example, the bacterium *Hemophilus aegypticus* produces an enzyme named HaeIII that cuts DNA wherever it encounters the sequence,

5'-GGCC-3'

3'-CCGG-5'

The cut is made between the adjacent G and C. This particular sequence occurs at 11 places in the circular DNA molecule of the virus phiX174. Thus treatment of this DNA with the enzyme produces 11 fragments, each with a precise length and nucleotide sequence. These fragments can be separated from one another and the sequence of each determined.

HaeIII and AluI cut straight across the double helix producing "blunt" ends.

However, many restriction enzymes cut in an offset fashion. The ends of the cut have an overhanging piece of single-stranded DNA. These are called "sticky ends" because they are able to form with any DNA molecule that contains the complementary sticky end. Any other source of DNA treated with the same enzyme will produce such molecules (Figure 2-20).

Mixed together, these molecules can join with each other by the base pairing between their sticky ends. The union can be made permanent by another enzyme, DNA ligase, that forms covalent bonds along the backbone of each strand. The result is a molecule of recombinant DNA (rDNA).

The ability to produce recombinant DNA molecules has not only revolutionized the study of genetics, but has laid the foundation for much of the biotechnology industry. The availability of human insulin (for diabetics), human factor VIII (for males with hemophilia A), and other proteins used in human therapy were made possible by recombinant DNA.

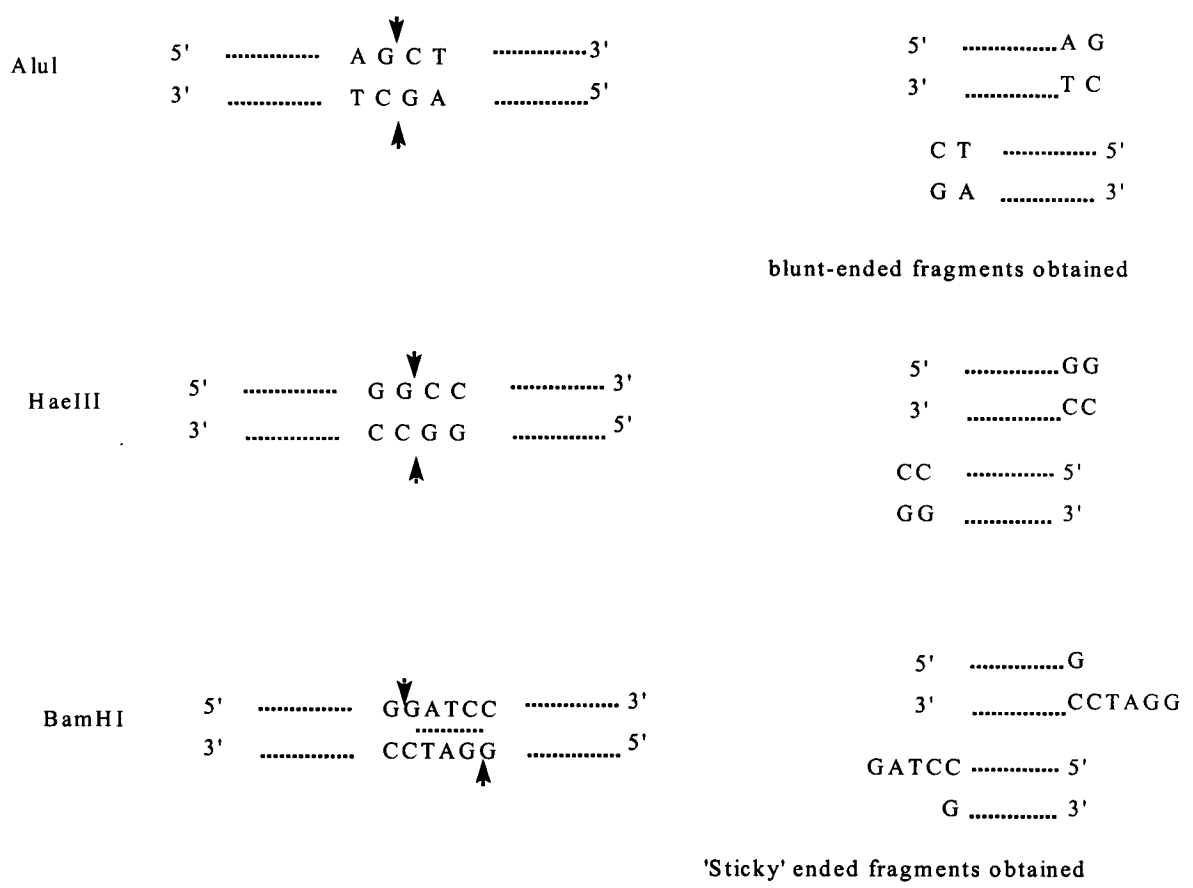


Figure 2-20 Actions of restriction enzymes⁶²

2.11.5 Cloning technique

2.11.5.1 Cloning strategy

Method A

The cosmid vector pLAFR3 is used to construct the genomic library (Figure 2-21). Two separate vector arms were made to reduce the concatemerisation. pLAFR3 is isolated and purified using caesium chloride gradient ultracentrifugation according to an adaptation the Ish-Horowicz method⁶³. Two batches of vector are digested with EcoRI and HindIII respectively followed by digestion with BamHI. A partial digest of M5 total DNA is made using SauIIIA to generate DNA fragments of adequate size (~24kb) for insertion. After ligation, the rDNA is packaged in phage heads using a kit from Boehringer Mannheim. Magnesium-cultured *E. coli* HB101 cells are then infected and plated on LB plates containing tetracycline for selective growth of transformed colonies.

Method B

An alternative method used by Gerrit Poelerens (personal communication, RUG, Netherlands) involves the omission of vector digestions with EcoRI and HindIII; instead pLAFR3 is digested with BamHI only. This method is useful because it involved fewer digestion steps, therefore reducing the amount of DNA lost with each digestion step. The drawback with this method is that vector could recircularise without any insert since the two ends are compatible as they result originally from the action of BamHI. This in turn led to lower insert frequency in the library. Ligation of the digested vector with M5 total DNA (from the partial digest) followed as in the previous method.

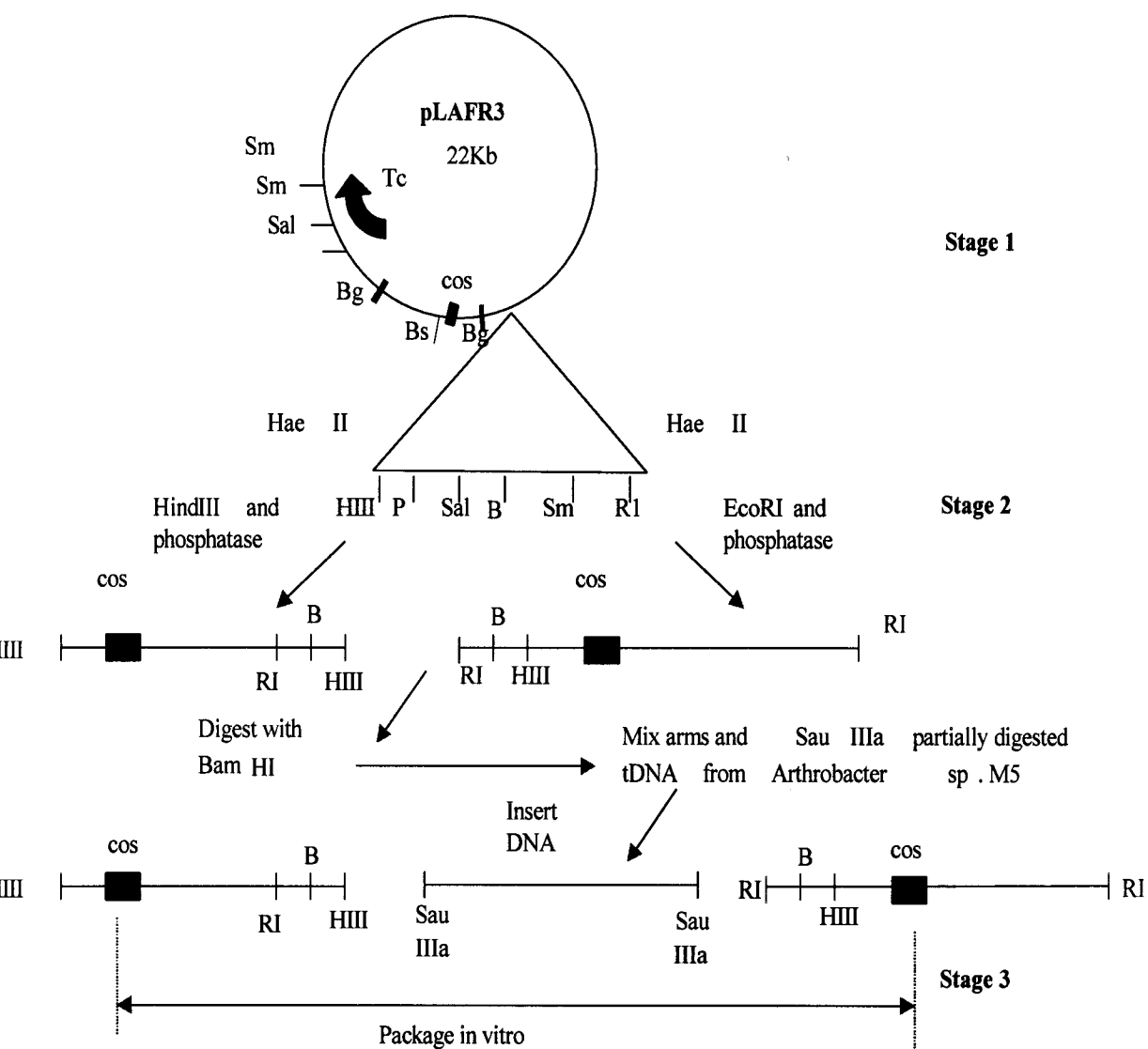


Figure 2-21. Outline of cloning strategy.

2.11.5.2 *Escherichia* codon usage

When designing the primers to be used in mutagenesis, it was necessary to use codons that occurred most commonly in *E. coli* in order to increase the expression levels of the introduced mutation. Referring to the codon usage chart in Table 2-2, it was possible to construct successful primers. This was achieved by firstly identifying the codon in the sequence that encodes the amino acid that will be mutated. The genetic code is degenerate meaning that more than one codon codes for the same amino acid. The use of particular codons to encode an amino acid vary among species. Hence the codon usage chart for *E. coli* made it possible to select a particular codon that was specific to *E. coli*. When the ideal codon was selected, it was then possible to introduce a point mutation that would be accepted by *E. coli*, thereby increasing the expression potential of the mutated gene product.

Table 2-2 *Escherichia coli* codon usage chart⁶⁴.

***Escherichia coli* : 4541860 codons)**

fields: [triplet] [frequency: per thousand] ([number])

UUU 22.0(100128)	UCU 9.3(42367)	UAU 16.7(75774)	UGU 5.2(23461)
UUC 16.5(74885)	UCC 8.9(40365)	UAC 12.3(55847)	UGC 6.3(28747)
UUA 13.8(62823)	UCA 7.9(35837)	UAA 2.0(9006)	UGA 1.0(4428)
UUG 13.3(60322)	UCG 8.7(39546)	UAG 0.3(1172)	UGG 14.5(65630)
CUU 11.3(51442)	CCU 7.2(32678)	CAU 12.7(57585)	CGU 20.7(93997)
CUC 10.6(48147)	CCC 5.4(24383)	CAC 9.6(43743)	CGC 21.1(96053)
CUA 4.0(18067)	CCA 8.5(38663)	CAA 14.8(67129)	CGA 3.7(16607)
CUG 50.9(231373)	CCG 22.3(101467)	CAG 28.8(130898)	CGG 5.7(25751)
AUU 29.9(135873)	ACU 9.5(43256)	AAU 18.7(84846)	AGU 9.1(41544)
AUC 24.6(111878)	ACC 22.7(103121)	AAC 21.6(98018)	AGC 15.6(70867)
AUA 5.3(24233)	ACA 7.9(35995)	AAA 34.4(156169)	AGA 2.7(12345)
AUG 27.2(123604)	ACG 14.0(63696)	AAG 11.4(51685)	AGG 1.6(7423)
GUU 19.1(86572)	GCU 16.2(73677)	GAU 32.3(146794)	GGU 25.1(114185)
GUC 14.8(67356)	GCC 25.0(113412)	GAC 19.3(87759)	GGC 28.6(130043)
GUA 11.2(51020)	GCA 20.6(93390)	GAA 39.5(179460)	GGA 8.6(39036)
GUG 25.5(115687)	GCG 32.2(146264)	GAG 18.5(83804)	GGG 11.1(50527)

Coding GC 51.37% 1st letter GC 58.50% 2nd letter GC 40.70% 3rd letter GC 54.90%

2.11.5.3 pET Vector System

The pET Vector System, developed by Studier and Moffat⁶⁵, is an effective tool for cloning and expression of recombinant proteins in *E. coli*. The cloned genes are expressed under the transcriptional control of bacteriophage T7 promoter. Transcriptional units cloned into the pET Vectors remain transcriptionally silent until the expression of a chromosomal copy of the T7 RNA polymerase gene is induced.

The three pET-5 Vectors which are derived from pBR322, contain the promoter and translational start site from bacteriophage T7 gene 10 (the major capsid protein of the T7 bacteriophage), BamHI and EcoRI cloning sites and the β -lactamase gene (for Amp^r). The BamHI and EcoRI cloning sites are preceded by a region which encodes an eleven amino acid sequence derived from the amino terminus of the T7 gene 10 protein. When DNA is cloned into either of these restriction sites, it is expressed as a fusion protein initiated at the amino terminal methionine of gene 10 peptide. The three pET-5 Vectors differ only in the reading frame relative to the BamHI site: pET-5a expresses from the GGA triplet, pET-5b expresses from the GAT triplet and pET-5c expresses from the ATC triplet of the GGATCC BamHI recognition sequence.

All three of pET-5 Vectors can also be used to express proteins without a fusion peptide. The gene 10 initiation codon is embedded in an Nde I restriction sequence which can be used as a cloning site. If the insert DNA also contains an Nde I site at the 5' end of the coding sequence, the bacteriophage initiation codon can be replaced by the initiation codon of the cloned DNA resulting in the elimination of the gene 10 terminal fusion.

Target genes are initially cloned into a pET-5 Vector in a strain, which does not carry the gene for T7 RNA polymerase (e.g. JM109). Following verification of the construct, the recombinant plasmid is then transferred into a DE3 host [eg. BL21 (DE3) pLysS or JM 109(DE3)]. DE3 strains carry a copy of the T7 RNA polymerase gene integrated into the host chromosome. Both BL21 (DE3) pLysS and JM 109(DE3) carry a copy of the T7 RNA polymerase gene under the control of the inducible lacUV5 promoter. The transformed cells are grown to mid-log phase and expression of the protein induced with IPTG. The desired protein product can accumulate to as much as 50% of the total cellular protein within two or three hours, due to the high transcriptional activity of T7 RNA polymerase.

2.12 Southern blotting

2.12.1 Complementarity and hybridisation

Molecular searches use one of several forms of complementarity to identify the macromolecules of interest among a large number of other molecules. Complementarity is the sequence-specific or shape-specific molecular recognition that occurs when two molecules bind together. For example: the two strands of a DNA double helix bind because they have complimentary sequences; also, an antibody binds to a region of a protein molecule because they have complimentary shapes.

Complementarity between a probe molecule and a target molecule can result in the formation of a probe-target complex. This complex can then be located if the probe molecules are tagged with radioactivity or an enzyme. The location of this complex can then be used to get information about the target molecule. In solution, hybrid

molecular complexes (usually called hybrids) of the following types can exist (other combinations are possible):

- a) DNA-DNA. A single-stranded DNA (ssDNA) probe molecule can form a double-stranded, base-paired hybrid with an ssDNA target if the probe sequence is the reverse complement of the target sequence.
- b) DNA-RNA. A single-stranded DNA (ssDNA) probe molecule can form a double-stranded, base-paired hybrid with an RNA (RNA is usually a single-strand) target if the probe sequence is the reverse complement of the target sequence.
- c) Protein-Protein. An antibody probe molecule (antibodies are proteins) can form a complex with a target protein molecule if the antibody's antigen-binding site can bind to an epitope (small antigenic region) on the target protein.

The first important feature of hybridization is that hybridization reactions are specific - the probes will only bind to targets with complimentary sequence (or, in the case of antibodies, sites with the correct 3-d shape). The second important feature of hybridization is that hybridization reactions will occur in the presence of large quantities of molecules similar but not identical to the target. That is, a probe can find one molecule of target in a mixture of millions of related but non-complementary molecules.

These properties make it possible to use hybridization to perform a molecular search for one DNA molecule, or one RNA molecule, or one protein molecule in a complex mixture containing many similar molecules. These techniques are necessary because a cell contains tens of thousands of genes, thousands of different mRNA species, and thousands of different proteins. When the cell is broken open to extract DNA, RNA,

or protein, the result is a complex mixture of all the cell's DNA, RNA, or protein. It is impossible to study a specific gene, RNA, or protein in such a mixture with techniques that cannot discriminate on the basis of sequence or shape. Hybridization techniques make it possible to pick out the molecule of interest from the complex mixture of cellular components and study it on its own.

Blots are named according to the target molecule.

Southern Blot:	DNA cut with restriction enzymes - probed with radioactive DNA.
----------------	---

Northern Blot:	RNA - probed with radioactive DNA or RNA
----------------	--

Western Blot:	Protein - probed with radioactive or enzymatically-tagged antibodies.
---------------	---

The formation of hybrids in solution is of little experimental value - if a solution of DNA is mixed with a solution of radioactive probe, the result is a radioactive solution. It is impossible to tell the hybrids from the non-hybridized molecules. For this reason, it is necessary first to physically separate the mixture of molecules to be probed on the basis of some convenient parameter.

These molecules must then be immobilized on a solid support, so that they will remain in position during probing and washing. The probe is then added, the non-

specifically bound probe is removed, and the probe is detected. The place where the probe is detected corresponds to the location of the immobilized target molecule.

In the case of Southern, Northern, and Western blots, the initial separation of molecules is done on the basis of molecular weight. (Cloning uses a different technique.)

2.12.2 The blotting procedure

2.12.2.1 *Gel electrophoresis*

This is a technique that separates molecules on the basis of their size. A slab of gel material is cast. Gels are usually cast from agarose or poly-acrylamide. These gels are solid and consist of a matrix of long thin molecules forming sub-microscopic pores. The size of the pores can be controlled by varying the chemical composition of the gel. The gel is cast and soaked with buffer. The gel is then set up for electrophoresis in a tank holding buffer and having electrodes to apply an electric field.

The pH and other buffer conditions are arranged so that the molecules being separated carry a net (-) charge so that they will be moved by the electric field from anode to the cathode. As they move through the gel, the larger molecules will be held up as they try to pass through the pores of the gel, while the smaller molecules will be impeded less and move faster. This results in a separation by size, with the larger molecules nearer the well and the smaller molecules farther away.

2.12.2.2 Transfer to Solid Support

After the DNA, RNA, or protein has been separated by molecular weight, it must be transferred to a solid support before hybridization. (Hybridization does not work well in a gel.) This transfer process is called blotting and that is why these hybridization techniques are called blots. Usually, the solid support is a sheet of nitrocellulose paper (sometimes called a filter because the sheets of nitrocellulose were originally used as filter paper), although other materials are sometimes used. DNA, RNA, and protein stick well to nitrocellulose in a sequence-independent manner. The DNA, RNA, or protein can be transferred to nitrocellulose in one of two ways:

- a) Electrophoresis, which takes advantage of the molecules' negative charge.
- b) Capillary blotting, where the molecules are transferred in a flow of buffer from wet filter paper to dry filter paper.

In a Southern Blot, the DNA molecules in the gel are double-stranded, so they must be made single stranded in order for the probe to hybridize to them. To do this, the DNA is transferred using a strongly alkaline buffer, which causes the DNA strands to separate. This process is called denaturation. DNA binds to the filter as single-stranded molecules. RNA and protein are run in the gels in a state that allows the probe to bind without this pre-treatment.

2.12.2.3 Blocking

At this point, the surface of the filter has the separated molecules on it, as well as many spaces between the lanes, etc., where no molecules have yet bound. If we added the probe directly to the filter now, the probe would stick to these blank parts of the

filter, like the molecules transferred from the gel did. This would result in a filter completely covered with probe that would make it impossible to locate the probe-target hybrids. For this reason, the filters are soaked in a blocking solution that contains a high concentration of DNA, RNA, or protein. This coats the filter and prevents the probe from sticking to the filter itself. During hybridization, the objective is for the probe to bind only to the target molecule.

2.12.2.4 Preparing the Probe

To prepare radioactive DNA probes for Southern blots and Northern blots, a restriction fragment of a plasmid containing the gene of interest must be made. The plasmid is digested with particular restriction enzymes and the digest is run on an agarose gel. Since a plasmid is usually less than 20 kb long, this results in 2 to 10 DNA fragments of different lengths. If the restriction map of the plasmid is known, the desired band can be identified on the gel. The band is then cut out of the gel and the DNA is extracted from it. Because the bands are well separated by the gel, the isolated DNA is a pure population of identical double-stranded DNA fragments.

The DNA restriction fragment (template) is then labeled by Random Hexamer Labeling⁶⁶.

- a) The template DNA is denatured - the strands are separated - by boiling.
- b) A mixture of DNA hexamers (6 nucleotides of ssDNA) containing all possible sequences is added to the denatured template and allowed to base-pair. They pair at many sites along each strand of DNA.

c) DNA polymerase is added along with dATP, dGTP, dTTP, and radioactive dCTP.

Usually, the phosphate bonded to the sugar (the α -phosphate, the one that is incorporated into the DNA strand) is synthesized from phosphorus-32 (^{32}P), which is radioactive.

d) The mixture is boiled to separate the strands and is ready for hybridization.

This produces a radioactive single-stranded DNA copy of both strands of the template for use as a probe.

2.12.2.5 Hybridisation

In all three blots, the labeled probe is added to the blocked filter in buffer and incubated for several hours to allow the probe molecules to find their targets.

2.12.2.6 Washing

After hybrids have formed between the probe and target, it is necessary to remove any probe that is on the filter that is not stuck to the target molecules. Because the nitrocellulose is absorbent, some of the probe soaks into the filter and must be removed. If it is not removed, the whole filter will be radioactive and the specific hybrids will be undetectable.

To do this, the filter is rinsed repeatedly in several changes of buffer to wash off any unhybridised probe. In Southern blots and Northern blots, hybrids can form between molecules with similar but not necessarily identical sequences (For example, the same gene from two different species). This property can be used to study genes from different organisms or genes that are mutated. The washing conditions can be varied so that hybrids with differing mismatch frequencies are maintained. This is referred to

as controlling the stringency: the higher the wash temperature, the more stringent the wash, the fewer mismatches per hybrid are allowed.

2.12.2.7 Detection of Probe-Target Hybrids

At this point, the sheet of nitrocellulose has spots of probe bound wherever the probe molecules could form hybrids with their targets. The filter looks like a blank sheet of paper and the probe must be detected. The probe-target hybrids can be detected by autoradiography using X-ray film if the probe is radioactive. If the probe is an antibody-enzyme conjugate, it can be detected by enzymatic development.

2.13 Arthrobacter sp. M5

2.13.1 Metabolism of substituted acetophenones.

Arthrobacter sp. degrade chlorosubstituted xanthenes and produce chlorinated 2-hydroxyacetophenones as by-products⁶⁷. Mono-, di- and trichlorinated acetophenones have also been formed from polychlorinated biphenyls (PCBs) by *Arthrobacter* sp^{68,69, 70}. Acetophenones have been observed as by-products of styrene and ethylbenzene degradation by a *Pseudomonas* strain⁷¹.

Investigations carried out by Cripps⁷² and Cripps *et al.*⁷³ on the microbial degradation of acetophenone showed that the first step is a biological type of Baeyer-Villiger oxidation (**Figure 2-22**). Various compounds (e.g. fluoroxanthene⁷⁴, cyclohexane-1, 2-diol⁷⁵, alkanes and alkenes^{76,77, 78}) are degraded via the same pathway. Phenyl acetate is the product of Baeyer-Villiger the oxidation and is further degraded via phenol and catechol.

2.13.2 Isolation and characterisation of a 4-chloroacetophenone-degrading culture.

According to studies done by Havel and Reineke⁷⁹, a growing culture was not attainable when the substrate was added directly to an enrichment culture. However, when 4-chloroacetophenone was added in vapour phase, it was possible to obtain a transferable growing culture. Consecutive sub-culturing onto fresh media enabled 8 morphologically different strains to be isolated on nutrient broth agar. None of the isolated pure cultures was able to grow with 4-chloroacetophenone as the sole carbon source. However, the 8 strains were systematically combined until growth was achieved. In this way, a defined culture of 2 strains was obtained. The 2 strains were tentatively identified as *Arthrobacter* sp. M5 and *Micrococcus*.

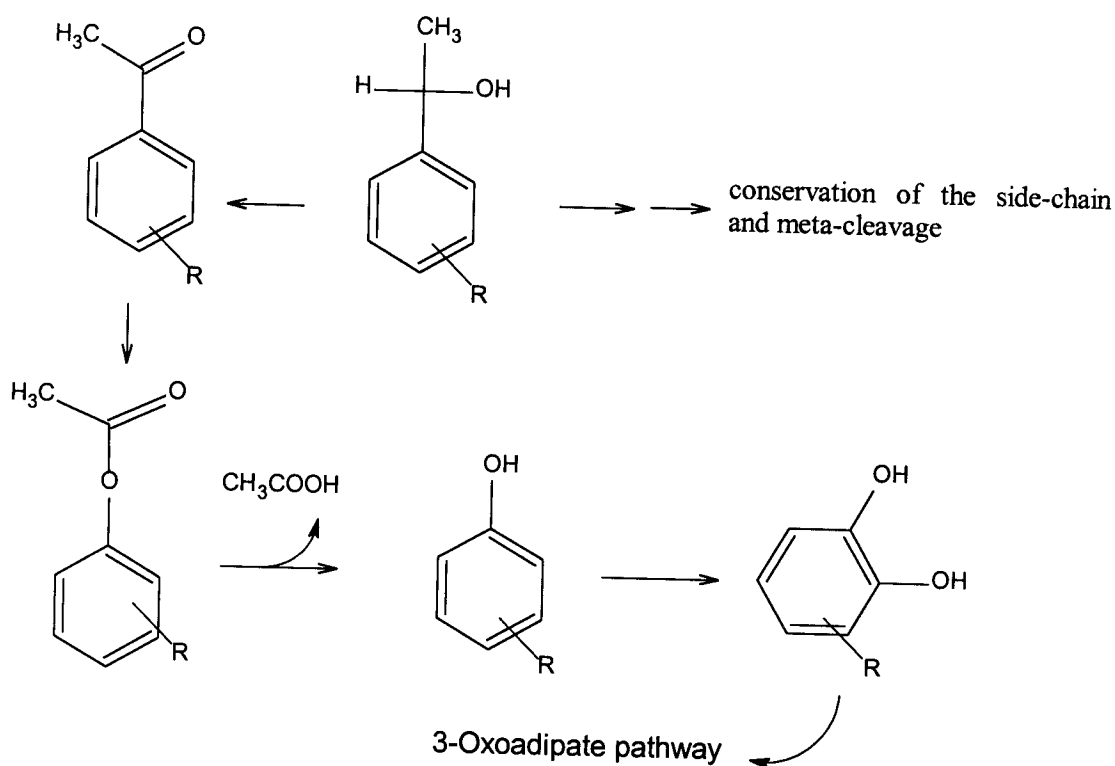


Figure 2-22⁷⁹. Pathway for degradation of acetophenone (R=H) and chlorinated analogs (R=Cl) via the corresponding phenyl acetates, phenols and catechols based on the model of Cripps⁷² and Cripps *et al.*⁷³ Degradation of 1-phenylethanols can follow the same pathway or occur by ring fission with retention of the side-chain.

CHAPTER 3: MATERIALS AND METHODS

3.1 Materials

Table 3-1. Buffers and solutions used in the protocols

Buffers	Solutions
1 M Tris-Cl pH 7.0	Tridest (triple distilled water)
20 mM Tris-Cl pH 7.5	Bidest (double distilled water)
25 mM Tris-Cl pH 8.0	0.1 mM p-nitroacetophenone
10 mM EDTA pH 8.0	50 mM glucose
Phosphate buffer pH 7	0.2 N NaOH
3 M sodium acetate pH 7.0	3.5 M CsCl ₂
3 M potassium acetate pH 4.8	0.5 M CaCl ₂
Alkaline phosphatase buffer pH 8.5	5 M potassium acetate
SM buffer pH 7	Glacial acetic acid
MM buffer pH 7	5 M NaCl
PCR buffer* pH 7	10 mM MgSO ₄
TAE buffer pH 7	0.5 M EDTA
Mutagenesis buffer* pH 7	Phenol
T ₁₀ E ₁ buffer (10 mM Tris: 1M EDTA) pH 7	Chloroform
Buffer A* pH 7	Ethidium bromide
Buffer B* pH 7	1% SDS
Buffer H* pH 7	Ethanol (705% and 96%)
	Isopropanol
	n-butanol
	0.2% maltose
	8% agarose
	Glycerol
*from Boehringer-Mannheim®	

Table 3-2 Solutions used in isolation of DNA

Solution I	Solution II	Solution III
50 mM glucose 25 mM Tris /Cl (pH 8.0) 10 mM EDTA (pH 8.0)	0.2 N NaOH (freshly diluted from a 10 N stock) 1% SDS	5 M potassium acetate 60ml glacial acetic acid 11.5ml H ₂ O 28.5ml
Comments		
Solution I was autoclaved at 10 LB/sq. In. For 15 minutes and stored at 4°C	Solution II was autoclaved at 10 lb in ⁻² . For 15 minutes and stored at 4°C	Solution III was 3 M with respect to potassium and 5 M with respect to acetate and autoclaved and stored as Solution I and II.

Table 3-3 Enzymes, antibiotics and co-factors

Enzymes	Antibiotics and Co-factors
Lysozyme [3.2.1.17]	Tetracycline*
RNAse [3.1.27.5]	Chloramphenicol*
SauIIIA [3.1.23.36]	Ampicillin*
Alkaline phosphatase [3.1.3.1]	
Eco RI [3.1.23.21]	1 mM ATP
Hind III [3.1.23.7]	250 µM dNTP
Bam HI [3.1.23.6]	NADPH
T ₄ ligase [6.5.1.3]	
Taq polymerase [6.5.1.x]	* from Boehringer-Mannheim®
DpnI DNA polymerase [2.7.7.x]	
Pfu DNA polymerase [2.7.7.x]	
Taq extender (Stratagene®) [6.5.1.x]	

Table 3-4 Vectors and bacteria

Vectors	Bacteria
pLYS5*	<i>Pseudomonas fluorescens</i> ACB
pet5a *	<i>Arthrobacter</i> sp M5
pLAFR3*	<i>E coli</i> HB101
*from Janssen frozen stock	<i>E coli</i> B121 DE23-pLYS5

Table 3-5 Commercially available kits

Name of Kit	Manufacturer
Quickchange mutagenesis kit	Stratagene, Germany
DIG labeling kit for southern blots	Boehringer-Mannheim®, Germany
DNA packaging kit	Boehringer-Mannheim®, Germany

Bacterial strains and vector properties

Table 3-6 Bacterial strain properties

Bacteria	Strain	References
<i>E. coli</i> strains	HB101	Boyer et al 1969 ⁸⁰
	BL21 DE23	
<i>Arthrobacter</i>	sp. M5	Keddie, R. M. and Jones, D., 1981 ⁸¹

Table 3-7 Vector properties

Vector	References
pLAFR3	Staskawicz, B. <i>et al.</i> , 1987 ⁸²
pLysS	Studier, F. W. <i>et al.</i> , 1990 ⁸³ Moffat, B.A. and Studier, F.W., 1987 ⁶⁵
pET3a	Rosenberg, A. H. <i>et al.</i> , 1987 ⁸⁴ Studier, F. W. <i>et al.</i> , 1990 ⁸⁵

3.2 Methods

3.2.1 Cultivation of bacteria

3.2.1.1 Cultivation and isolation

- a) *E. coli* HB101 was grown at 37 °C with shaking (200 rpm) in LB medium.
- b) *E. coli* HB101 with pLAFR3 was grown at 37 °C in LB medium containing tetracycline (Tc 12.5 mg/ml).
- c) *E. coli* B121 DE23-pLYS5 was grown at 37 °C in LB medium containing chloramphenicol (Chl 35 mg/ml).
- d) *E. coli* B121 DE23 with pLYS5 and pET5a-HapE was grown at 37 °C in LB medium containing chloramphenicol (Chl 35 mg/ml) and ampicillin (Amp 100 mg/ml).
- e) *Arthrobacter sp.* M5 isolated on acetophenone was cultivated in NB medium grown overnight at 37 °C with shaking (200 rpm).

3.2.1.2 Large cultures (for mutant cultivation)

Large cultures (1L) containing the appropriate antibiotic were inoculated with precultures and grown overnight at room temperature with shaking (200 rpm). For the precultures, 20ml LB amp + tet medium was inoculated with the appropriate cells and grown at 37° C with shaking (200 rpm). After growth, cells of large cultures were harvested by centrifugation at 7000 rpm for 15 minutes and resuspended in 50 mM phosphate buffer (pH 7).

3.3 Isolation of DNA

3.3.1 Isolation of total DNA from *Arthrobacter* sp. M5

4 Erlenmeyer of 100 ml NB medium were inoculated with 0.1 ml, 0.2 ml, 0.4 ml and 0.5 ml overnight culture respectively and grown overnight. All the cultures were harvested at once by centrifugation at 5000 rpm for 15 minutes. The pellet was resuspended in 5ml tridest with 50 ml 1M Tris/Cl (pH 8.4), 10 ml 0.5 M EDTA and 50 ml 5 M NaCl.

2 mg lysozyme was added and this was incubated for at least 60 minutes at 65 °C to lyse the cells. The sample showing the best lysis was used to proceed with the rest of the protocol. This was determined by inspecting the viscosity of the lysate in each test tube. The most viscous solution represented the sample showing greatest lysis.

To precipitate the protein, 0.8 ml 3 M sodium acetate pH 7.0 was added and the mixture was incubated at 65 °C for 120 minutes. Following this, 0.5 volumes phenol were added and this was incubated at 65 °C for 15 minutes at 4 °C. 0.5 volumes chloroform were then added at room temperature and the mixture was centrifuged for 10 minutes at 7500 rpm. The water phase was collected in a Greiner tube and the DNA was precipitated with 2 volumes ethanol and 0.1 volumes 3 M sodium acetate pH 5.2 for 30 minutes at 0 °C. The extraction with chloroform and phenol was done to purify the DNA. This was essential since the efficiency of restriction in subsequent stages depended greatly on the purity of the DNA.

The DNA precipitate was fished out using a Pasteur pipette (heated in a Bunsen flame to make a hook) and dissolved in 1-4 ml tridest.

3.3.2 Isolation of cosmid DNA (pLAFR3)

E. coli HB101 cells were harvested from 100 ml overnight culture by centrifugation at 6000 rpm. The cells were resuspended in 12 ml 50 mM glucose, 10mM EDTA, 25mM Tris/Cl (pH 8) and incubated for 5 minutes at room temperature. The cells were then divided into two batches of 6 ml each. 12 ml 1% SDS in 0.2 M NaOH was added to 6ml of resuspended cells and they were incubated at room temperature until the solution became clear, representing good lysis. 6ml 3M potassium acetate pH 4.8 was added; incubation for 5 minutes on ice followed. This was centrifuged twice for 10 minutes at 7500 rpm to remove as much protein as possible.

To precipitate the DNA, 0.6 volumes isopropanol were added and this was incubated for 30 minutes at -20° C following by centrifugation at 7500 rpm for 10 minutes. The pellet was washed once with 5ml 70% ethanol, dried at 50° C and redissolved in 1.0 ml T₁₀E₁ (10 M Tris buffer to 1 M EDTA).

6 µl RNase (10 mg/ml) was added the DNA was incubated for 30 minutes at 37 °C. The DNA was cleaned using phenol/chloroform extraction: an equal volume of phenol was added to the DNA solution in an Eppendorf tube and the contents were inverted 5 times to ensure adequate mixing of the 2 phases. The tube was then centrifuged at 14 000 rpm in a desktop centrifuge for two minutes. The upper water phase containing the DNA was pipetted off and put in a clean Eppendorf tube, leaving the impurities at the interface of the two phases. Following this, an equal volume of 1:1 phenol/chloroform solution (V/V) was added and the procedure repeated. Then finally, an equal volume of chloroform was added and the extraction repeated.

3.3.3 pLAFR3 purification

Further purification of DNA was carried out using cesium chloride gradient ultracentrifugation in order to obtain an isolate of pLAFR3 cosmid, which would be susceptible to the restrictions that followed in subsequent stages. Also, the cosmid has a known tendency of forming concatemers i.e. many cosmids joint together to form long linear strands that did not cut well. This treatment was used to minimise this.

1.35g CsCl was added to 1000 ml pLAFR3 (in T₁₀E₁). The DNA was diluted if it was not dissolving well. 30ml ethidium bromide (10 mg/ml) was added. The supernatant was transferred to a polycarbonate tube using blunt-ended tips. 1.4 ml 3.5 M CsCl was added to the polycarbonate tube taking care not to allow the two layers to mix. The mixture was centrifuged overnight at 80,000 rpm at 20° C in a fixed-angle rotor.

The lower band of the two observed (using UV light) was collected. The ethidium bromide was removed by several extractions with n-butanol saturated with water. The DNA was diluted by adding 4 volumes of water. The DNA was precipitated by adding 2.5 volumes 96% ethanol and incubating at -20 °C for 30 minutes. This was spun down for 15 minutes and the pellet was washed twice with 70% ethanol and dried, then redissolved in 50 ml T₁₀E₁ buffer.

3.4 Preparation of DNA⁸⁶

3.4.1 Preparation of partial digest

Eight Eppendorf tubes were prepared as illustrated in fig. 3.1. 100 µl containing 10 µg DNA was added to the first tube. Tubes 2-7 had 50 µl containing 50 µg DNA added to them. Tube 8 remained empty. 0.05 units SauIII A in the appropriate buffer (buffer B manufactured by Boehringer Mannheim) was added to tube one and mixed. 50 µl of the resulting solution was transferred from tube one to tube two. The contents of tube two were mixed followed by transfer of 50 µl to tube three and so on, until all tubes had 50 µl DNA. The tubes were incubated for one hour at 37° C, and the reaction was stopped by placing the tubes on ice. The DNA was analysed on 0.8% agarose and the sample showing the onset of digestion was selected to proceed with. The particular sample which showed the onset digestion varied from the 4th sample to the 6th sample. 0.2 volumes alkaline phosphatase buffer (10x) and 1 unit alkaline phosphatase were added; the volume was adjusted to twice its original amount using double distilled water. This was incubated at 37° C for 15 minutes and the sample was purified by phenol/chloroform extraction. The DNA was precipitated with ethanol at -20° C for 30 minutes. The DNA was redissolved in 10 µl tridest.

3.4.2 Preparation of vector arms⁸⁶

Ten micrograms pLAFR3 was digested with EcoRI and 10 µg pLAFR3 was digested with HindIII. 0.2 volumes alkaline phosphatase buffer (10x) and 0.5 units alkaline phosphatase were added; the volume was adjusted to twice the volume that the previous restriction had been carried out in. This was incubated for 30 minutes at 37 °C and the alkaline phosphatase was inactivated by incubating the reaction mixtures

at 75° C for 10 minutes. The DNA was purified using phenol/chloroform extraction. The DNA was precipitate with 2.5 volumes ethanol and 0.1 volume 3 M sodium acetate pH 5.2, for 30 minutes at -20 °C. The pellet was washed twice with 70% ethanol and then dissolved in 10 ml tridest. Both samples were then digested with BamHI and purified with phenol/chloroform extraction. The DNA was precipitate as before and the pellet was dissolved in 10 ml tridest.

3.4.3 Ligation of DNA fragments

Ligation was carried out in standard conditions; the reaction mixture contained 1 unit T4-ligase, 1 mM ATP and 1:1:1 molar ratio of insert DNA: BamHI /EcoRI digested vector: BamHI/HindIII digested vector. The ligation was carried out overnight at 16 °C.

Infection was carried out according to the manufacturer's instructions of the DNA packaging kit used by Boehringer Mannheim, following the protocol as described Sambrook *et al*⁸⁶.

20 ml of LB medium supplemented with 0.2% maltose and 10 mM MgSO₄ was inoculated with a single colony of *E coli* HB101 and grown at 37 °C with vigorous shaking for 4-5 hours until the OD₆₀₀ had reached 0.8-1.0.

The bacteria was centrifuged for 10 minutes at 3500 rpm (4 °C). The cells were diluted by gently resuspending the pellet in 40 ml 10 mM MgSO₄ (twice the original volume.) The resuspended bacteria were expected to have an OD₆₀₀ of 0.5.

The bacteria were now ready for infection.

3.4.4 Packaging of DNA

The desired number of packaging extracts were thawed at room temperature and placed immediately on ice when they were beginning to thaw.

As soon as the extract was thawed, the DNA to be packaged was immediately added.

The amount of the DNA added was 0.1 μ l- 5 μ l in a maximum volume of 4 μ l.

This was mixed gently with a pipette, avoiding air bubbles. The vial was then centrifuged for a few seconds at 14 000 xg.

The packaging mixture was then incubated for 2 hours at room temperature (20-22 °C). 450 μ l SM buffer and 20 μ l chloroform were added. The chloroform had the function of causing protein material to precipitate, forming debris in the mixture. The debris was centrifuged for a few seconds. The supernatant was then ready for plating or could be stored for a few weeks. However, the highest efficiencies were obtained from using freshly prepared solutions as the titre dropped several fold after storage.

3.5 Storage of library

The genomic library was stored as frozen glycerol stocks in microtitre plates.

3.6 Screening of library

3.6.1 Screening method (A)

0.1 mM para-nitroacetophenone in MM buffer pH 7 was used as substrate for the acetophenone monooxygenase. Microtitre plates were filled with 150 μ l of MM buffer supplemented with substrate and were inoculated with clone cell mass and

incubated overnight at room temperature; the following day they were inspected for the indicative colour change to bright yellow.

3.6.2 Screening method (B.)

Microtitre plates were filled with 150 µl Mm buffer supplemented with substrate and inoculated with 50 µl of growing cell suspension of clones. The clones were then incubated at 30° C with shaking (200 rpm) for at least four hours to allow them to grow, and hopefully express their insert DNA. 10 µl NaOH 5M was then added to increase the pH, and encourage colour formation in wells containing the acetophenone monooxygenase product.

3.6.3 Screening using probe method (C)

Another screening method was used as an alternative to the first colorimetric method of screening. This involved using the known esterase sequence from *Pseudomonas* ACB to create a probe that could be used to screen the total DNA of *Arthrobacter* sp. M5.

PCR mixture:

The following were put in a PCR vial:

5 µl PCR buffer (10x concentrated)

5 µl PCR DIG labelling mix

1 µl primer 1 [*Pseudomonas* ACB acetophenone monooxygenase 5'→3']

1 µl primer 2 [*Pseudomonas* ACB acetophenone monooxygenase 3'→5']

1 µl template [total DNA from *Pseudomonas* ACB]

1 µl TAQ polymerase

36 µl tridest

The total volume was 50 µl. The mixture was subjected to PCR in the following conditions:

30 cycles:

5 minutes at 94°C for denaturation of DNA.

1 minute at 56°C

1 minute at 72°C (elongation)

The probe was stored at -20°C and denatured at 95°C for 10 minutes before use.

Arthrobacter sp M5 and *Pseudomonas* ACB total DNA samples were digested with BamHI and HindIII to cut the genomes into smaller bands that may contain the desired gene. The samples were run on agarose gel overnight at 10 mA to obtain well-resolved bands. 15µL of digested DNA was loaded into each well. The gel was then used in a southern blot

3.7 Preparation of gels

0.8 g low melting point agarose was dissolved in 100 ml TAE buffer by heating using a microwave for a few minutes to make 0.8% agarose gel. 1ul ethidium bromide was added to 25 ml of warm agarose solution, thoroughly mixed and poured into a mould. A comb was placed in the warm solution in order to form the loading wells when the gel had set upon cooling.

3.8 Mutagenesis

Quickchange site directed mutagenesis kit by Stratagene was used to design and prepare the primers according to the manufacture's instructions in order to introduce two point mutations (Asp483-Asn and Gly490-Ala) into HapE from *Pseudomonas* ACB cloned in pET5a expression vector. The codon usage chart of *E coli* was consulted when choosing the codons for alanine and asparagine.

3.8.1 PCR-based Site Directed Mutagenesis

Plasmid template DNA (approximately 0.5 μ mole) was added to a PCR cocktail containing, in 25 μ l of 1x mutagenesis buffer: (20 mM Tris HCl, pH 7.5; 8 mM MgCl₂; 40 μ g/ml BSA); 12-20 μ mole of each primer (one of which must contain a 5'-phosphate), 250 μ M each dNTP, 2.5 U *Taq* DNA polymerase, 2.5 U of *Taq* Extender (Stratagene).

The PCR cycling parameters used were 1 cycle of:

4 minutes at 94 °C,

2 minutes at 50 °C,

2 min at 72 °C;

this was followed by 5-10 cycles of

1 minute at 94 °C,

2 minutes at 54 °C

1 minute at 72 °C (step 1).

The parental template DNA and the linear, mutagenesis-primer incorporating newly synthesized DNA were treated with *DpnI* (10 U) and *Pfu* DNA polymerase (2.5U). This resulted in the *DpnI* digestion of the *in vivo* methylated parental template and hybrid DNA and the removal, by *Pfu* DNA polymerase, of the *Taq* DNA polymerase-extended base(s) on the linear PCR product. The mixture was incubated at 37 °C for 30 minutes and then transferred to 72 °C for an additional 30 minutes (step 2). Mutagenesis buffer (1x, 115 µl, containing 0.5 mM ATP) was added to the *DpnI*-digested, *Pfu* DNA polymerase-polished PCR products. The solution was mixed and 10 µl was removed to a new Eppendorf tube and T4 DNA ligase (2-4 U) was added.

The ligation was incubated for greater than 60 min at 37 °C (step 3).

The treated solution was transformed into competent *E. coli* (step 4).

3.9 Digestions

All restriction enzyme digestions were carried out as described by Sambrook *et al*⁸⁶, using appropriate buffers according to the manufacture's instructions (Boehringer Mannheim[®]) as illustrated below:

Table 3-8 Enzyme digestions

Enzyme	Buffer pH	Buffer	Enzyme Concentration in Units per μL	Incubation Period in minutes	Temperature In $^{\circ}\text{C}$
EcoRI	7	Buffer H	10	60	37
HindIII	7	Buffer B	10	60	37
BamHI	7	Buffer B	10	90	37
SauIIIA	7	Buffer A	Varies in partial digest	30	37
Alkaline phosphatase	8.5	Alkaline phosphatase buffer	1	30 (10 minutes at 65°C to inactivate enzyme)	37
Taq polymerase	7	PCR buffer	1	Depends on cycle parameters	Depends on cycle parameters
RNAse	7	Tridest	6 μl of 10 mg/ml sol	30	37
Ligase	7	Ligation buffer	1	overnight	16

3.10 Preparation of cell-free extracts

Cell free extracts were prepared by suspending the harvested cells of interest in phosphate buffer and then using a French press to mechanically shear the cells and thereby release the contents of the cells. The cells were stored on ice before and after lysis.

3.11 Transformation of cells

Two 20 ml cultures of *E. coli* B121 DE23-pLYS5 cells were grown to an OD₆₀₀ of 0.3-0.4 at 37 °C with appropriate antibiotics and shaking (at 200 rpm). 1 ml of cells was harvested by centrifugation at 14000 rpm for 30 seconds in Eppendorf tubes and resuspended in 0.5 ml ice cold CaCl₂. The cells were allowed to stand on ice for 30 minutes. The cells were harvested and carefully resuspended in 0.1 ml CaCl₂ rendering the cells competent for 30 minutes. 1 ml of 10x diluted vector was added to each batch and this was left on ice for 30 minutes following which the cells were heat-shocked at 42 °C for 90 seconds.

400 ml LB medium was added and this was incubated for 45 minutes at 37 °C. The cells were harvested and plated on LB medium containing the appropriate antibiotics for selection of transformants.

3.12 Minipreparations of DNA

Minipreparations of DNA were obtained by using alkaline lysis method. A single bacterial colony was transferred into 2 ml LB medium containing the appropriate

antibiotic in a loosely capped 15ml tube. The culture was incubated overnight at 37 °C with vigorous shaking.

1.5 ml of the culture was poured into an Eppendorf tube and centrifuged at 14 000 g for 30 seconds, at 4 °C. The remainder of the culture was stored at 4 °C. The medium in the Eppendorf tube was removed by aspiration, leaving the bacterial pellet as dry as possible. The pellet was then resuspended in 100 µl ice-cold solution I by vigorous vortexing. Following vortexing, 200 µl of freshly prepared solution II was added. The contents of the tube were mixed by inverting the tube several times rapidly. It was essential that the entire surface of the tube came into contact with solution II. Vortexing was not done to avoid shearing of DNA. The tubes were stored on ice. Then 150 µl of ice cold Solution III were added. The tube was inverted to mix the contents and stored on ice for 3-5 minutes. Centrifugation followed at 14000 g for 5 minutes at 4 °C and the supernatant was transferred into a fresh tube.

The double-stranded DNA was precipitated with 2 volumes of 70% ethanol at room temperature. The mixture was allowed to stand for 2 minutes at room temperature after mixing. The mixture was centrifuged again at 14000 g for 5 minutes (4 °C). The supernatant was removed by gentle aspiration and the tube was allowed to stand in an inverted position to drain off any remaining moisture. The pellet was rinsed off with 1ml 70% ethanol and the liquid removed by aspiration. It was allowed to air-dry for 10 minutes. The DNA was redissolved in 50 µl T₁E₁₀ buffer (pH 8) and vortexed briefly. The DNA was stored at -20 °C.

3.13 Colorimetric Assays

The concentrations necessary for a color change to occur were determined by carrying out a simple test. 10 wells of a microtitre plate were filled with 50 mM phosphate buffer (pH 7) and a (0-100 mM) range concentration of p-nitroacetophenol. The total volume of each well was 150 μ L. A loopful of *Arthrobacter sp.* M5 was added to hydrolyse the p-nitroacetophenone biocatalytically to p-nitrophenol, which would result in a yellow colour. This test indicated the threshold concentration of p-nitrophenol necessary for a visible yellow colour.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Isolation of total DNA

Total DNA of *Arthrobacter sp.* M5 was isolated using the method of alkaline lysis described in Section 3.3.1. The initial isolate was found to contain a lot of RNA, and protein, which could be seen on the agarose gel as a dense white region towards the end of the gel (Figure 4-1). Since total DNA has a relatively high mass, it was expected to run at the top end of the gel. The RNA was removed by digesting the isolate with RNase as described in Section 3.3.1. The protein was removed by salt precipitation.

Observations during isolation:

Arthrobacter sp. M5 was cultivated as outlined in Section 3.2.1.1. The cell pellet obtained upon centrifugation had a slight orange colouration, which could be due to the presence of carotenoid molecules and iron-containing enzymes. Lysis of the cells was carried out according to Section 3.3. and a viscous solution was obtained which indicated that the cells had undergone sufficient lysis; hence the cell contents were now suspended in the solution.

The protein content was removed by salt precipitation. The solution obtained had a cloudy appearance due to the aggregation of precipitated protein. Centrifugation was used to separate the protein matter, which formed a loose pellet along one side of the centrifuge tube. The centrifugation was repeated to ensure that as much protein precipitate as possible was removed.

The solution that remained appeared as though the components were not completely miscible. This was due to the high salt concentration present in the solution. Extraction with phenol and chloroform purified the DNA, which was dissolved in the upper water phase. This purification was important for the operation of restriction enzymes in proceeding steps. The DNA precipitation was carried out using 70% ethanol and salt according to Section 3.3.1. This yielded a solution in which the DNA was suspended as a thread, which was fished out using a Pasteur pipette that was modified into a hook using a Bunsen flame. The DNA was redissolved in T₁E₁₀ buffer (pH 7) and stored at -20 °C for later use.

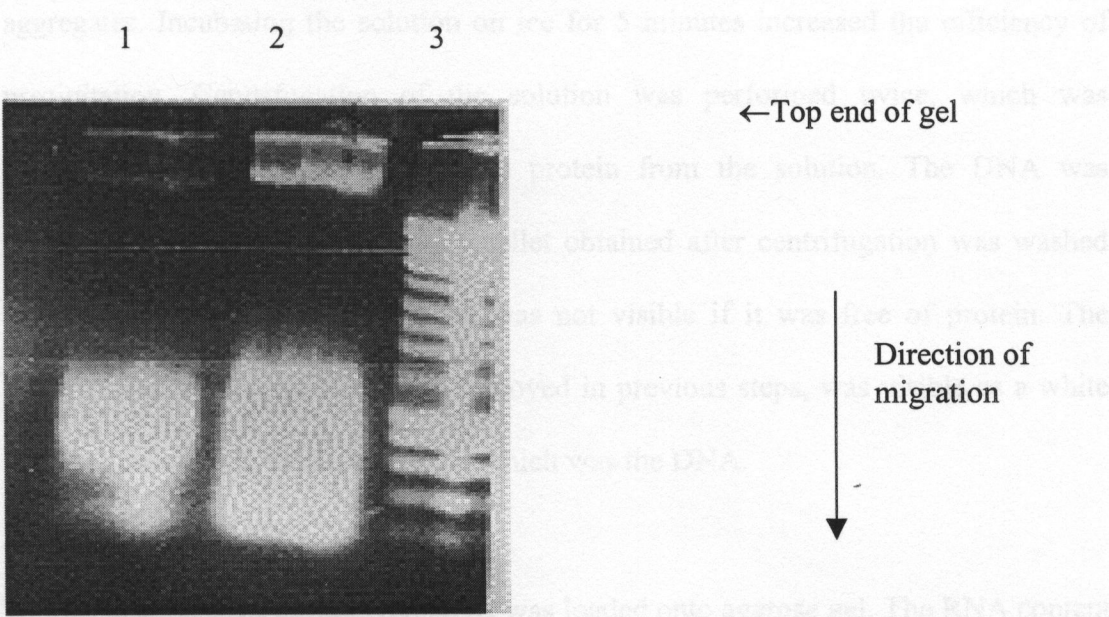


Figure 4-1 Agarose gel showing total DNA and Smart Ladder[®]

Running conditions: 100mA; Running time: 30 minutes; TEA buffer: pH 8.0

Lane 1: total DNA of *Arthrobacter sp.* M5 sample 1

Lane 2: total DNA of *Arthrobacter sp.* M5 sample 2

Lane 3: Smart Ladder[®]

(Figure 4-16 shows the sizes of individual bands on the Smart Ladder[®])

4.2 Isolation of cosmid vector

The cosmid vector pLAFR3 was isolated using the method outlined in Section 3.3.2. The *E coli* HB101 cells harvested from the overnight culture grown as described in Section 3.2.1.1 formed an off-white pellet with a plaque upon centrifugation. SDS was used to lyse the cells yielding an opaque solution, which became clear as lysis progressed. A little sliminess in the solution was also observed with lysis. Addition of sodium acetate resulted in the precipitation of the protein component as white cloudy aggregates. Incubating the solution on ice for 5 minutes increased the efficiency of precipitation. Centrifugation of the solution was performed twice, which was sufficient to separate the precipitated protein from the solution. The DNA was precipitated using isopropanol. The pellet obtained after centrifugation was washed using ethanol and dried. The pellet was not visible if it was free of protein. The protein, which may not have been removed in previous steps, was visible as a white residue surrounding the clear region, which was the DNA.

At this stage, 2 µl of redissolved DNA was loaded onto agarose gel. The RNA content was visible as a thick smear towards the bottom of the gel (fig4.2). RNase was used to remove this component. Phenol/ chloroform extraction was used to remove remnants of protein, which appeared as white flocculation at the interface of the solution. The RNase was also removed at this point with the protein residue.

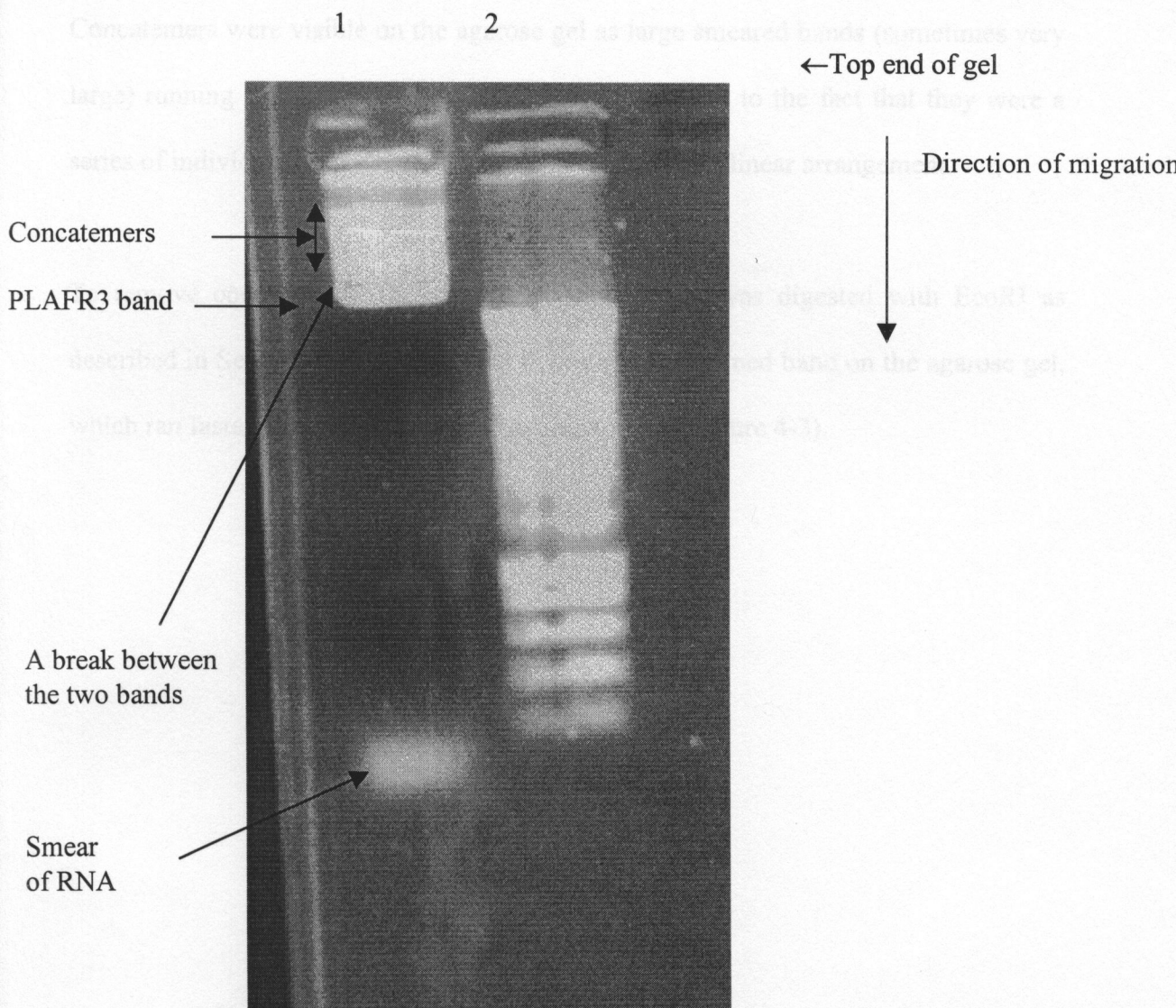


Figure 4-2. Agarose gel showing pLAFR3 DNA with concatemers.

Running voltage: 100mA

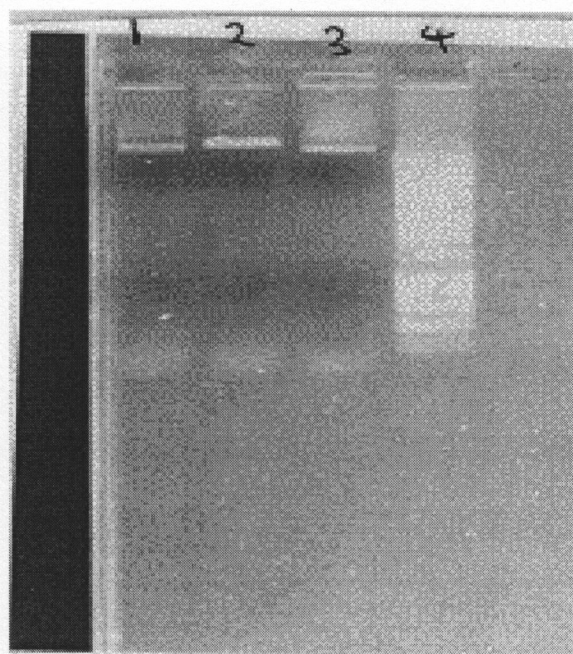
Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: pLAFR3 with concatemers

Concatemers were visible on the agarose gel as large smeared bands (sometimes very large) running slower than the pLAFR3 DNA band due to the fact that they were a series of individual vector molecules joined together in a linear arrangement.

To remove concatemers, the pLAFR3 cosmid DNA was digested with EcoRI as described in Section 3.9. This resulted in one clearly defined band on the agarose gel, which ran faster than the concatemers at about 24 kb (Figure 4-3).



←Top end of gel

Direction of migration



Figure 4-3 Agarose gel showing pLAFR3 cosmid DNA after digestion with EcoRI. Concatemerisation is not observed.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: pLAFR3 after short digestion

Lane 2: pLAFR3 after digestion, no concatemers observed

Lane 3: pLAFR3 with concatemers before digestion

4.3 pLAFR3 purification using caesium chloride.

The pLAFR3 DNA was purified according to Section 3.3.3. When caesium chloride was added to the redissolved pLAFR3, the solution became colder indicating that an endothermic reaction had taken place. Inverting the tube several times with care dissolved the caesium chloride more efficiently. A clear solution was obtained. Ethidium bromide was added resulting in a deep pink solution. When 1.35 M caesium chloride solution was added carefully drop by drop, the difference in the densities of the two solutions became apparent. This was observed as 2 phases in the solution; a clear upper phase and a pink lower phase.

After overnight centrifugation, UV light was used to view the pLAFR3 bands. Two bands were visible. The upper band was made up of open chains of pLAFR3, which were less dense. This affected migration in the caesium chloride gradient. The open chains became stationary upper end of the solution where the density of the solution was less. The lower band consisted of closed circularised pLAFR3, which formed denser structures that became stationary at the lower end of the tube where the density of the solution was higher. A slight agitation of the solution caused the bands to be less defined hence great caution was used in collecting the desired lower band.

The ethidium bromide was removed using n-butanol saturated in water resulting in a clear solution without any pink colour. The DNA was diluted with water precipitated using ethanol. The final clear pellet obtained was washed, dried and redissolved in T₁₀E₁ buffer and stored at -20 °C until required.

4.4 Partial digestion of *Arthrobacter sp.* M5.

The partial digest was prepared according to Section 3.4.1. The increasing extent of digestion in each tube was clearly visible when 2 µl of solution from each tube was loaded onto gel (Figure 4-4). This enabled the selection of tube 4 for subsequent steps in the procedure. Tube 4 was selected because it showed a smear meaning that various sizes of DNA segments were present in the band. It also ran a little lower down the gel than the other visible bands meaning that the DNA present had been sufficiently cleaved to create open linear fragments that could be used in subsequent stages. The DNA present in the undigested sample was tangled to form a dense mass, which did not migrate as far down the gel as the untangled digested tube 4.

Alkaline phosphatase was used to remove the phosphate group at the sticky ends of the DNA fragments. This was done to stop circularisation of the fragments at their sticky ends. This was also important because it ensured that the insert DNA would bind with greater frequency to the correct ends of the vector arms.

4.5 Preparation of vector arms.

The vector arms were prepared as described in Section 3.4.2. When the DNA was purified initially using phenol/ chloroform extraction, the water phase containing the DNA repeatedly formed a bubble within the phenol/ chloroform phases. This resulted in loss of DNA because the bubble was difficult to collect without contaminating the DNA with a phenol/ chloroform residue. This was overcome by diluting the DNA

with water (up to 100 μ l) hence much less amounts of DNA were lost when a bubble would form. Losses in DNA were mainly encountered during phenol/ chloroform extraction.

4.6 Ligation of DNA fragments.

Ligation was carried out as outlined in Section 3.4.3; 2 μ l of the ligation mixture was loaded onto agarose gel. A single band was observed (Figure 4-5). Infection was carried out using magnesium cultured *E coli* HB101⁸⁶. An average optical density at a wavelength of 600nm was used to harvest the cells. It was observed that cells would only reach this optical density in 4-5 hours if they had been grown from a freshly prepared plate culture from frozen stocks. Hence when cells obtained from an old plate (a week or more) were used to inoculate the magnesium culture, the growth was very slow and sometimes was not observed at all. Therefore fresh plates of *E. coli* HB101 were prepared for each inoculation (Section 3.2.1).

The first infection using cloning Method A (Section 2.11.5) yielded a disappointingly low number of transformants. Hence further attempts were made with little improvement in the amount of transformants obtained. Therefore Method B (Section 2.11.5) was used to prepare the vector resulting in a large number of transformants. A reduced aliquot of infection mixture was used to plate the bacteria in order to obtain colonies that were separated enough to facilitate efficient screening.

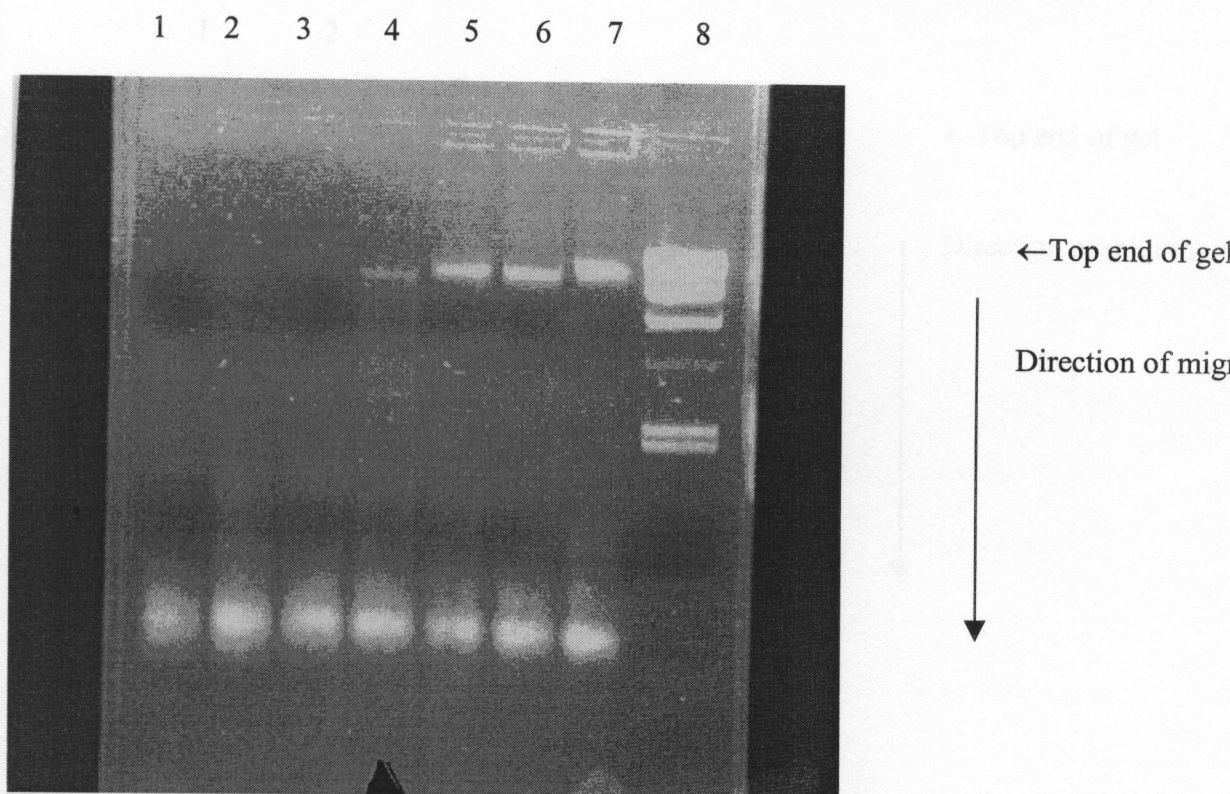


Figure 4-4 Agarose gel showing progressively less degree of digestion of *Arthrobacter sp* M5 total DNA by SauIII A restriction enzyme.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1-6: pLAFR3 after partial digestion with SauIII A

Lane 7: undigested pLAFR3

Lane 8: Lambda/HindIII ladder

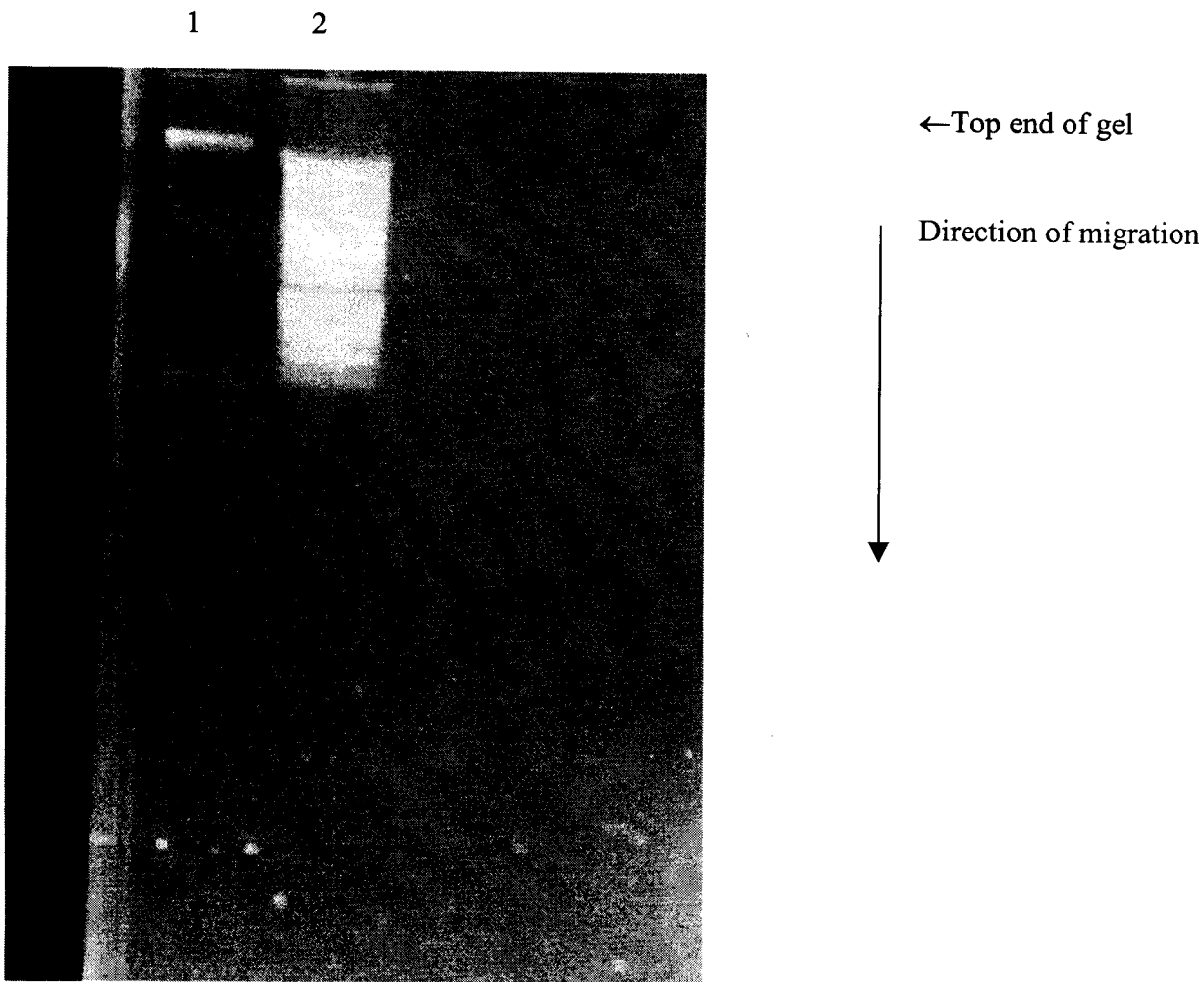


Figure 4-5. Agarose gel showing selected sample #4 after preparation for ligation as described in Section 3.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: Sample 4 after preparations for ligation

Lane 2: Smart Ladder[®]

4.7 Gene library construction

Method A as outlined in the cloning strategy (Section 3.6.1) was attempted initially but yielded no results, meaning that the DNA visible on gel was too little in concentration to use in the proceeding ligation procedure due to losses of vector DNA experienced with each phenol/chloroform extraction and ethanol precipitation. The inadequate concentration of the DNA on gel was determined visually by comparing the intensity of the DNA bands of the vector to those of the ladder used. The concentration of DNA in the bands of the ladder (Smart Ladder®) was able to be determined from the product information provided by the manufacturer.

At each stage in the vector preparation, gels were run to check the concentration levels of the DNA. The following figures show the progress of each step, illustrating the decreasing concentration of DNA with each extraction step, which ultimately lead to the inadequacy in the vector preparatory step experienced with method A described in Section 3.6.1. Figure 4-6 shows the vector pLAFR3 after digestion with EcoRI and HindIII restriction enzymes. Figure 4-7 shows EcoRI-digested vector pLAFR3 and HindII-digested vector pLAFR3 after they had undergone digestion with alkaline phosphatase. Figure 4-8 shows the subsequent step of phenol extraction. It is possible to see the decrease in DNA concentration that occurred as a result of the extraction process. Figure 4-9 shows one of the initial attempts at completing the necessary steps of constructing the library. This shows the substantial losses in DNA concentration that resulted in an unsuccessful attempt. Protein smears are visible at the bottom of the gel but the DNA concentration is clearly very low after the final precipitation.

Admittedly, the results shown in Figures 4-6 to 4-9 were partly due to inexperience in handling DNA. Therefore the DNA concentration remaining at the end of the procedure was inadequate to proceed with ligation.

Figure 4-10 shows all the components of the ligation mixture obtained from the above steps on one gel. This was done to enable a comparison between the concentrations of each component to be made. The ideal concentration ratio of insert DNA to each of the two vector arms was approximately 1:1:1.

After a successful ligation procedure, the hybrid DNA composed of vector pLAFR3 and *Arthrobacter* sp. M5 DNA was visible on gel as one neat clear band. This is shown in Figure 4-11.

When the ligation mix from this Method A was used to infect competent *E. coli* HB101 cells, the number of colonies of transformants obtained was very low.

Therefore Method (B.) as outlined in the cloning strategy (Section 3.6.2) was attempted from which a genomic library with 20% insert was constructed. The fewer steps involved were advantageous.

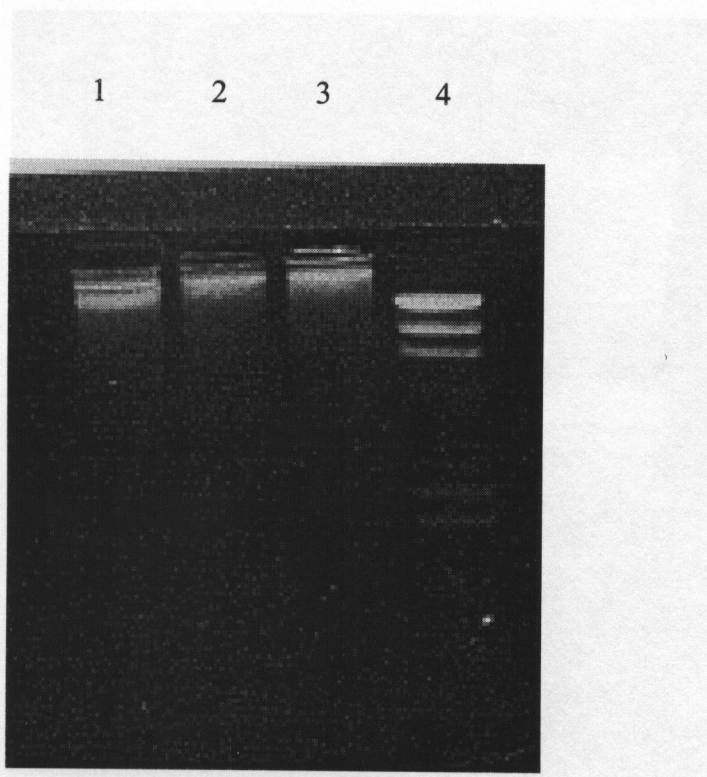


Figure 4-6. Vector pLAFR3 after digestion with EcoRI and HindIII restriction enzymes.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: undigested pLAFR3

Lane 2: pLAFR3 after digestion with HindIII

Lane 3: pLAFR3 after digestion with EcoRI

Lane 4: Lambda/HindIII ladder

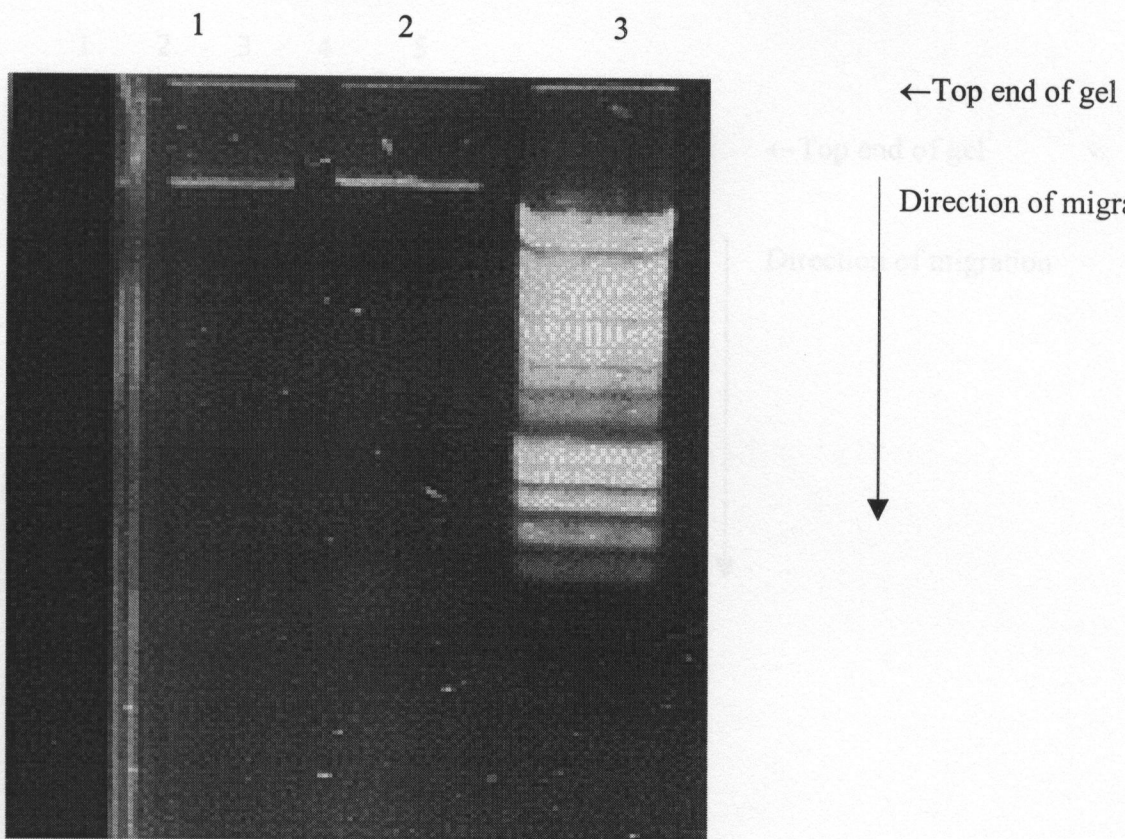


Figure 4-7. Vector pLAFR3 after digestion with alkaline phosphatase.

Running voltage: 100mA

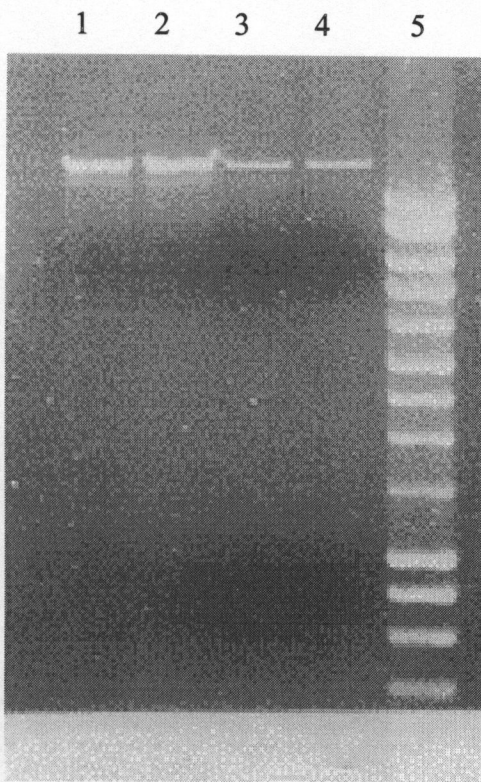
Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: EcoRI digested pLAFR3 after alkaline phosphatase digestion

Lane 2: HindIII digested pLAFR3 after alkaline phosphatase digestion

Lane 3: Smart Ladder[®]



←Top end of gel

Direction of migration

Figure 4-8. Vector pLAFR3 after phenol extraction.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: EcoRI-digested pLAFR3 before extraction

Lane 2: HindIII-digested pLAFR3 before extraction

Lane 3: EcoRI-digested pLAFR3 after extraction and precipitation

Lane 4: HindIII-digested pLAFR3 after extraction and precipitation

Lane 5: Smart Ladder[®]

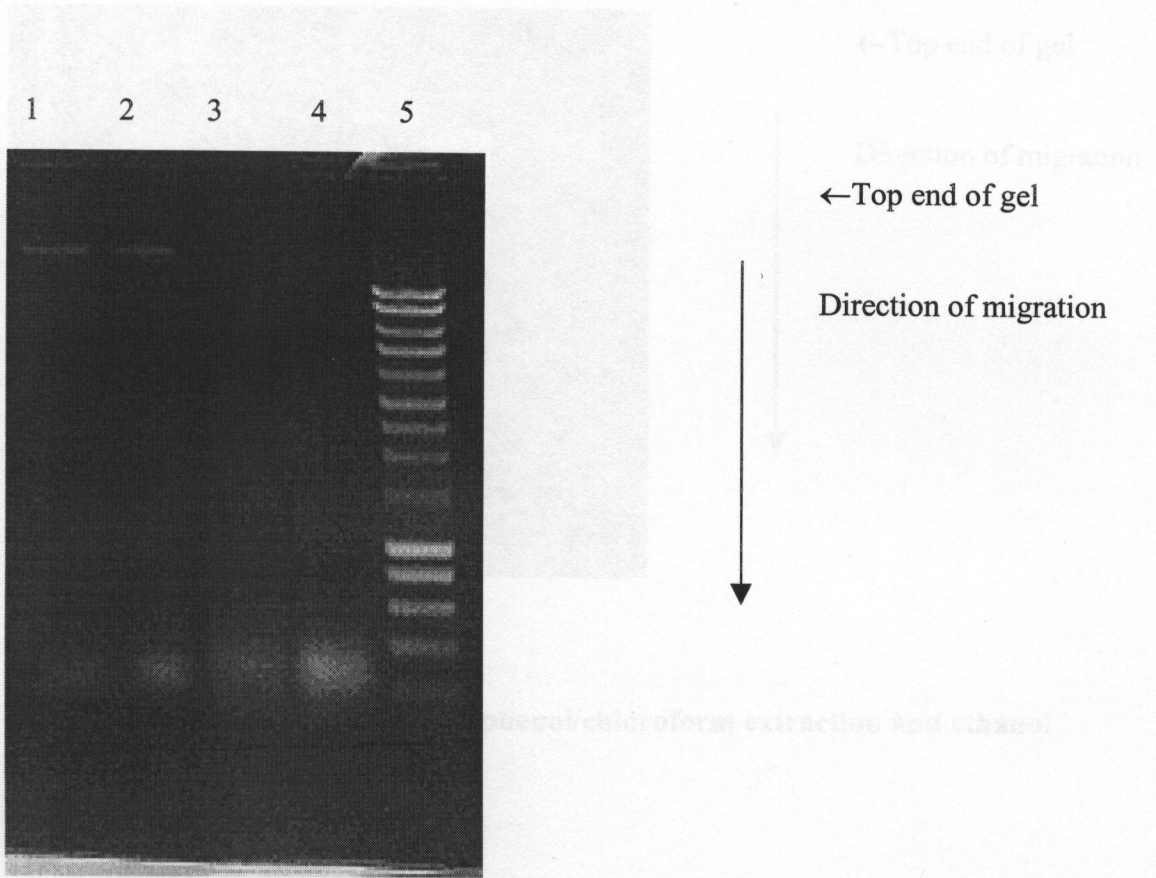


Figure 4-9. Vector pLAFR3 after phenol extraction and ethanol precipitation at -20 °C (one of the unsuccessful attempts).

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: EcoRI-digested pLAFR3 after extraction

Lane 2: HindIII-digested pLAFR3 after extraction

Lane 3: EcoRI-digested pLAFR3 after precipitation

Lane 4: HindIII-digested pLAFR3 after precipitation

Lane 5: Smart Ladder®

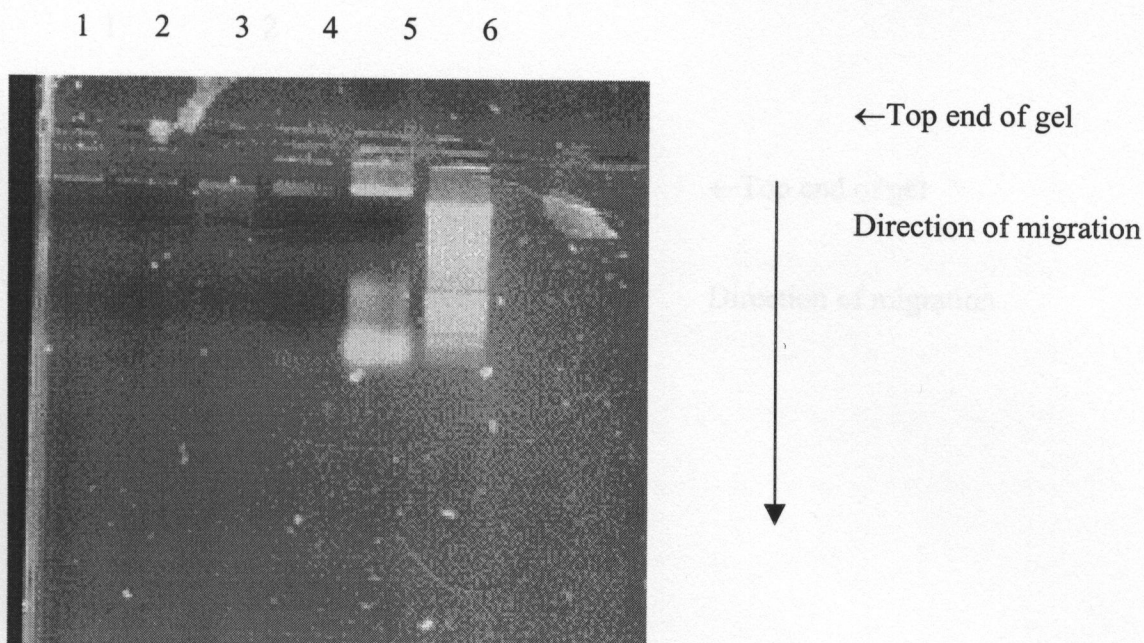


Figure 4-10. Vector pLAFR3 after phenol/chloroform extraction and ethanol precipitation at -20 °C.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: EcoRI-digested pLAFR3 after final extraction

Lane 2: HindIII-digested pLAFR3 after final extraction

Lane 3: Partially digested insert DNA

Lane 4: Undigested pLAFR3

Lane 5: Undigested M5 total DNA

Lane 6: Smart Ladder[®]

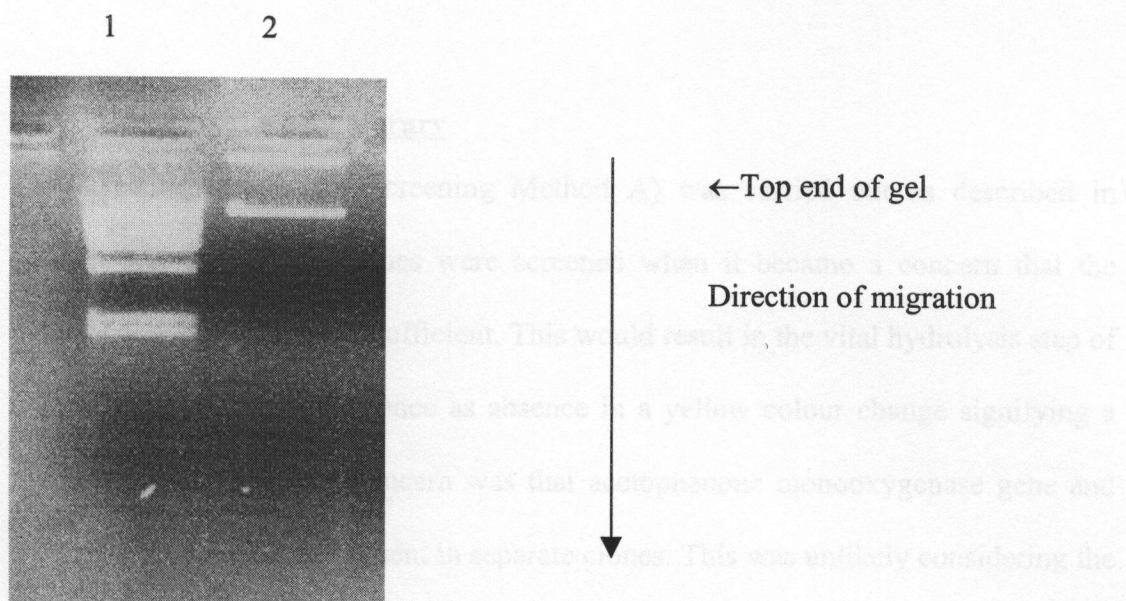


Figure 4-11. Vector pLAFR3 and M5 DNA after ligation.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lane 1: Lambda/HindIII marker

Lane 2: Ligated vector and insert DNA

4.8 Screening the gene library

The screening procedure (screening Method A) was carried out as described in Section 3.6. Over 2000 clones were screened when it became a concern that the esterase activity may not be sufficient. This would result in the vital hydrolysis step of the reaction not occurring hence as absence in a yellow colour change signifying a positive clone. An added concern was that acetophenone monooxygenase gene and the esterase gene may be present in separate clones. This was unlikely considering the close proximity of the two genes in many other species but it was still taken into consideration at this stage.

In light of this, screening Method B (Section 3.6.2.) was developed and used to screen the library afresh. Two clones were found to produce a strong yellow colouration. Surprisingly, the colour was not stable and diminished to the original beige colour of the original cell suspension with fifteen minutes. The colour may not have been stable due to the possibility that the yellow compound may have been rapidly used by another pathway as an alternative substrate, thereby impeding its accumulation.

The two clones were used to make cell-free extracts according to Section 3.10. The extracts were tested for acetophenone monooxygenase activity and were found to be inactive. Hence they were found to be false positives.

4.9 DNA Minipreparations.

The infection of competent cells was carried out as described in Section 3.4.3. The presence of insert was investigated by performing minipreps on transformed colonies as described in Section 3.12. After prepping 10 colonies, the insert became visible on gel as the presence of bands that ran faster than the vector after digestion with a restriction enzyme such as EcoRI (Figure 4-12).

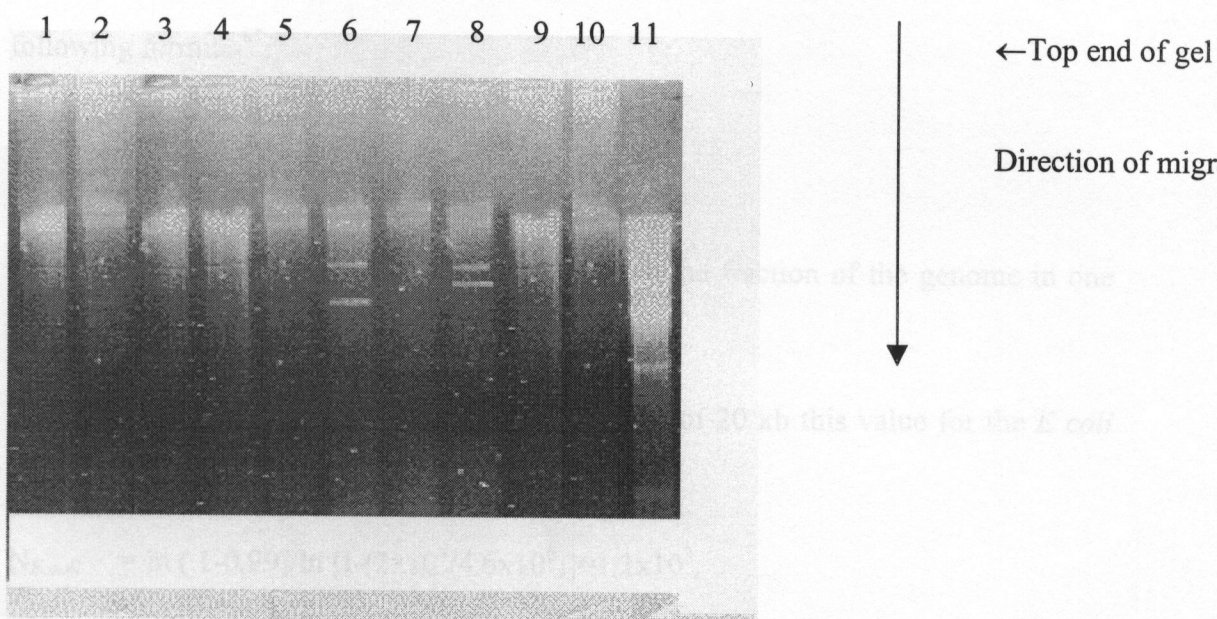


Figure 4-12. Agarose gel showing results of a miniprep. Lanes with multiple bands represent the presence of insert DNA.

Running voltage: 100mA

Running time: 30 minutes

TEA buffer, pH 8.0

Lanes 1-10: Plasmid DNA

Lanes 6 and 8 showing presence of multiple bands (presence of insert DNA)

Lane 11: Lambda/HindIII marker

The library was found to require the screening of 3000 colonies in order to statistically have a degree of certainty of finding the clone harbouring the acetophenone monooxygenase gene of interest. This figure was arrived at by using the following formula⁸⁷:

$$N = \ln(1-P) / \ln(1-f)$$

Where P is the probability that is desired and f is the fraction of the genome in one insert.

Therefore for a probability of 0.99 with insert size of 20 kb this value for the *E coli* genome is

$$N_{E.coli} = \ln(1-0.99) / \ln[1-(2 \times 10^4 / 4.6 \times 10^6)] = 1.1 \times 10^3$$

Where I = insert size (20kb)

G= genome size (4 x 10³ Kb)

N= independent clones needed.

P= probability of a given sequence being represented in this number of clones (99%)

Therefore,

N= 1.1x10³ colonies with 100% insert

For a library with 20% insert,

1.1 x10³ x 5= at least 5000 colonies.

Primarily, screening was performed as outlined in the Methods chapter (i.e. Method A. Section 3.6.1). However this was modified due to concern that the esterase activity may not be sufficient to perform the hydrolysis step of the reaction to give the yellow colour change. Also in the slightly unlikely event of the acetophenone monooxygenase and the esterase being present in separate clones, the second step of the reaction would not be achieved and no colour change would be observed.

In light of this, Method B (Section 3.6.2) was used to locate two clones that eventually turned out to be false positives. A strong yellow colour was observed as expected but this colour was not stable, and diminished to the original colour within fifteen minutes.

The instability of this colour was thought to be due to the low expression of the cloned acetophenone monooxygenase gene. To investigate this, an increased cell mass was used in the screening procedure used as described in Section 3.6.2. Another amendment to the screening procedure was the addition of 10 mM of sodium hydroxide (10 M) to each well in the screening plates in order to increase the pH of the screening environment. This was done to try and counter the effect of chemical hydrolysis, which occurred at acidic to neutral pH values.

The acetophenone monooxygenase gene was also known to occur close to an esterase gene⁸⁸ hence the possibility existed of there being an esterase responsible for the degradation of the yellow coloured product.

The finding that the colour was unstable was also surprising since the simple control tests carried out as outlined in Section 3.13 showed that the yellow colour was stable even at low concentrations (10 mM) *in vitro*.

The two clones were miniprepmed to check for insert as described in Section 3.12. They were also used to make cell free extract to test for acetophenone monooxygenase activity as described in Section 3.10. No activity was seen hence it was concluded that the clones were false positives.

Closer scrutiny of the screening technique revealed that *E. coli* cells were able to degrade para-nitroacetophenone to produce the desired colour when their cell mass was sufficient. This was observed by increasing the cell mass used in screening. This was most probably due to *E. coli*'s ability to utilise another degradation pathway that was not previously known. Therefore the screening strategy was abandoned because using this method would require the use of increased cell mass in order to achieve enough expression in each well to give a positive reaction.

Further attempts were made to make a genomic library with higher insert frequency and were nearly successful due to thorough restriction of cosmid DNA resulting in the absence of concatemers. This was achieved by incubating the restrictions for double the length of time used previously. This overcame the problem of DNA that was difficult to restrict.

It was also found that improved handling of DNA was an important factor.

Ironically, a 50% insert genomic library was lost due to preliminary minipreps that yielded DNA that was difficult to restrict (DNA that either did not respond to restriction enzymes or needed long periods of incubation), leading to an incorrect assumption that there was no insert. One factor that caused the DNA to be difficult to restrict was that if the DNA was dissolved in T₁E₁₀ buffer (pH 7), the EDTA could have interfered with the action of the enzyme. This was because the EDTA in the solution could cause chelation of cofactors at the active site of the enzymes leading to unsuccessful digestion of DNA. Also if the DNA was not sufficiently clean (free of RNA and Protein), the restriction was not successful as the impurities caused the restriction enzymes to be inhibited.

4.10 Mutants

The two selected mutants of pET5a-HapE namely Asp483-Asn and Gly490-Ala were successfully constructed. Sequence data confirmed the introduction of the two mutations. The pET5a vectors isolated from the transformants were sequenced by Piet Tepstra (Groningen Main Hospital, Netherlands) The Stratagene[®] kit used outlined clearly the important factors taken into consideration in the design of primers such as GC content, primer length and position of the mutation in the primer. The codon usage chart of *E coli* revealed the most accepted codons to use for each mutation.

HapE Asp483-Asn

Primer 1(a 31-mer)

5'- CCGAGCACGAATTCTAACTTGATCGTGTATGG-3'

Primer 2(a 31-mer)

5'- CCATACACGATCAAGTTAAATTCGTGCTCGG-3'

HapE Gly490-Ala

Primer 1(a 27-mer)

5'- GTGTATGGCACGGCGTTCCATGCCTCG-3'

Primer 2(a 27-mer)

5'-CGAGGCATGGAACGCCGTGCCATACAC-3'

4.11 Screening method C: Southern blotting.

A third method outlined in Section 3.6.3 was used to screen the *Arthrobacter sp.* M5 genome for the acetophenone monooxygenase gene. A probe was successfully made by PCR using part of the known esterase sequence (N. Kamerbeek). The esterase gene was amplified by PCR using two probes, which were designed by N. Kamerbeek (Figure 4-13). The total DNA of *Arthrobacter sp.* M5 and *Pseudomonas* ACB was restricted into resolved bands and the gels obtained were able to be used for Southern blotting (Figure 4-14).

Three Southern blots were made in which the stringency was varied in order to increase the chance of binding between the probe and the total DNA. The blots were reflective of the stringency used and therefore displayed increased unspecific binding as the stringency was increased. The blots were inconclusive, as it was difficult to distinguish the bands clearly (Figure 4-15). Unfortunately, the only binding that seemed to occur was between the undigested total DNA and the probe. A likely reason for the lack of binding between the probe and the digested DNA bands was that there was insufficient homology between the *Pseudomonas* ACB and *Arthrobacter sp.* M5 Esterase genes for binding to occur.

4.12 Comparison of the use of Southern blotting and PCR to detect a gene of interest.

PCR was much quicker and easier to use than Southern blotting to locate the target of interest but was fairly stringent and as a technique, had the drawback of sensitivity to base pair mismatch. In contrast, the stringency of Southern blotting could be controlled allowing for greater base pair mismatch. Southern blotting was found to be much more labour-intensive and took a longer time to get the result. It was also potentially hazardous due to some of the reagents used. Despite these drawbacks, Southern blotting had considerably more advantages than PCR when applied to locating a gene of interest. Southern blotting allowed for the detection of multiple copy or related genes and it also made it possible to determine if a gene of interest is a single locus, or member of a related family of loci (i.e. paralogous). Also, Southern blotting could be performed on DNA of an unknown sequence. These attributes of Southern blotting were very advantageous because the sequence of the monooxygenase (target gene) was unknown and was suspected to be a member of a related family, which included the known esterase, hence the esterase sequence was used to make a probe.

It is interesting to note that when trying to identify or detect a gene in "organism X" using a sequence from "organism Y", there is higher probability of getting a band on a Southern blot (that can be later used to help clone the gene of interest) than when using PCR if the genes are less than 90% identical. Even with 90%, it may be necessary to have the sequence of more than one homologue to identify highly conserved regions to design primers with enough homology to match the gene (if it is even present) in the target DNA. The intensity of positive band(s) in the Southern blot may give an idea of how close the target is to the gene of interest.

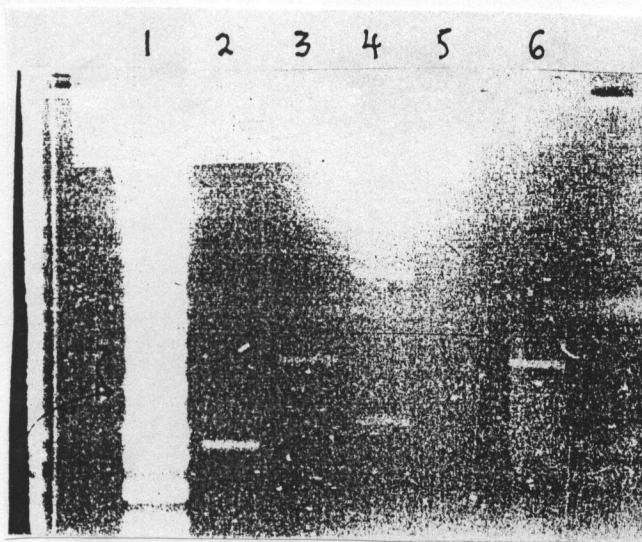


Figure 4-13. PCR to make a probe.

Lane 1: Smart ladder®

Lane 2: amplification due to unspecific binding of probe

Lane 3 : correct amplification of probe. The band runs at the same position as the control (esterase gene) in lane 6.

Lane 4: pLAFR3

Lane 5: Contains no sample

Lane 6: control esterase gene

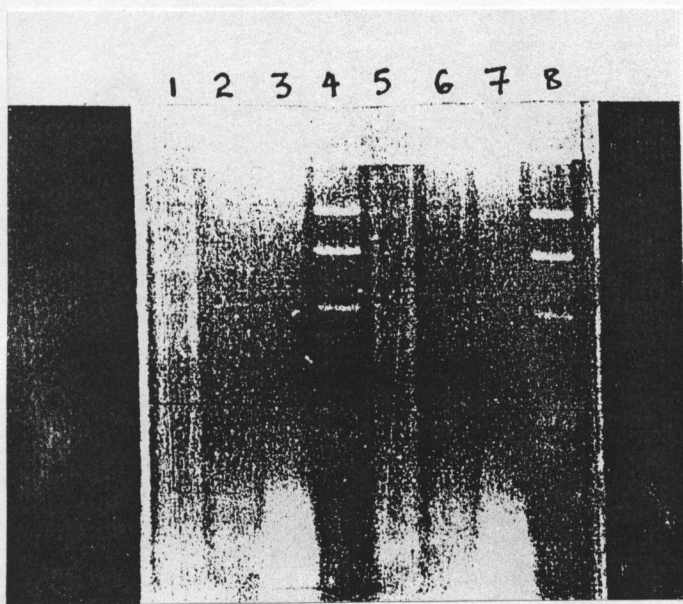


Figure 4-13. Gel for blotting (more stringent)

and B (less stringent). Both plates show the following details:

Lane 1 and 5: *Arthrobacter* sp. M5 total DNA digested with BamHI

Figure 4-14 Gel for blotting

Lane 1 and 5: *Arthrobacter* sp. M5 total DNA digested with BamHI

Lane 2 and 6: *Arthrobacter* sp. M5 total DNA digested with HindIII

Lane 3 and 7: *Pseudomonas* ACB total DNA

Lane 4 and 8: Lambda phage marker

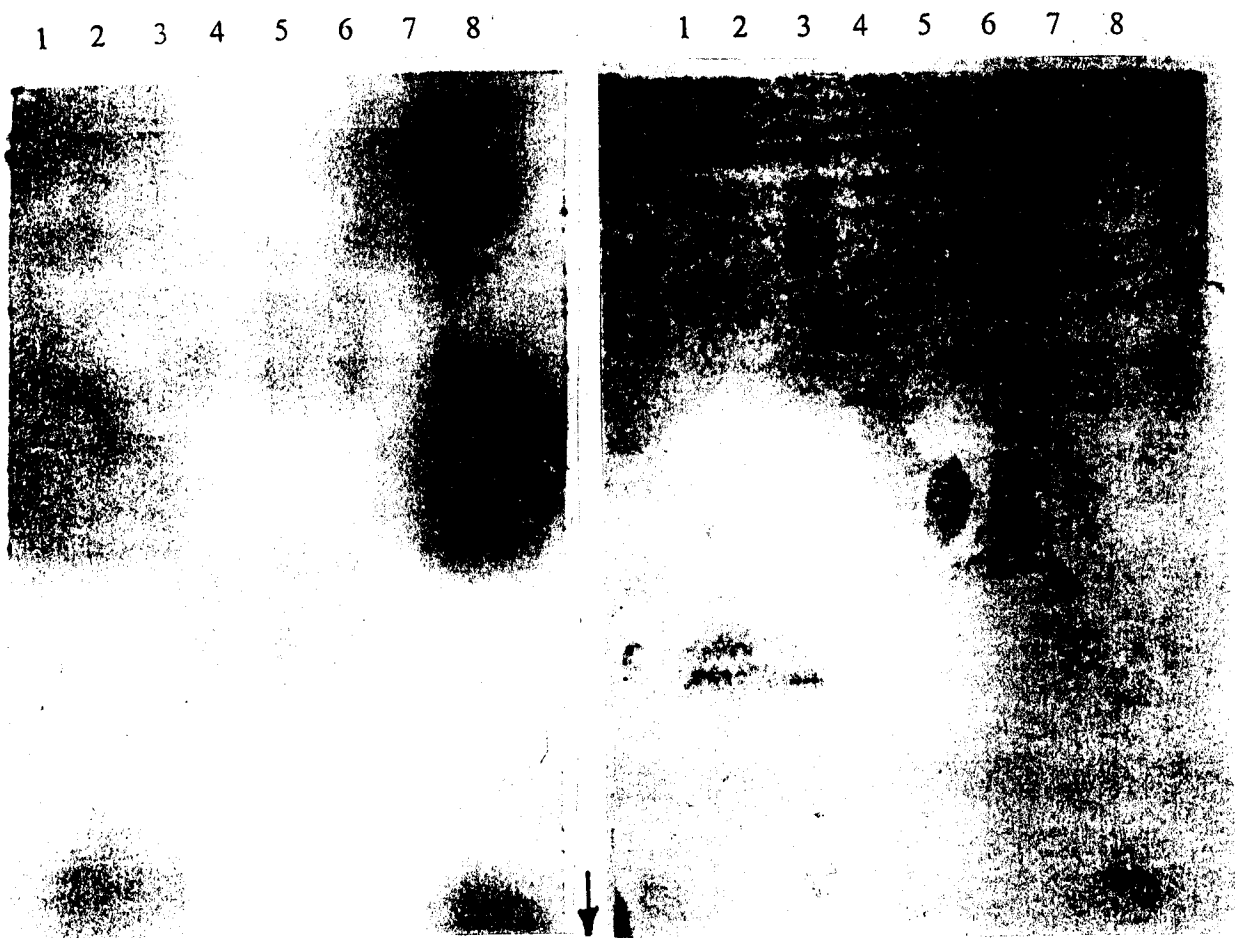


Figure 4-15. Photographic plates A and B of Southern blots A (more stringent) and B (less stringent). Both plates show the following details:

Lane 1 and 5: *Arthrobacter* sp. M5 total DNA digested with *Bam*HI

Lane 2 and 6: *Arthrobacter* sp. M5 total DNA digested with *Hind*III

Lane 3 and 7: *Pseudomonas* ACB total DNA

Lane 4 and 8: Lambda phage marker

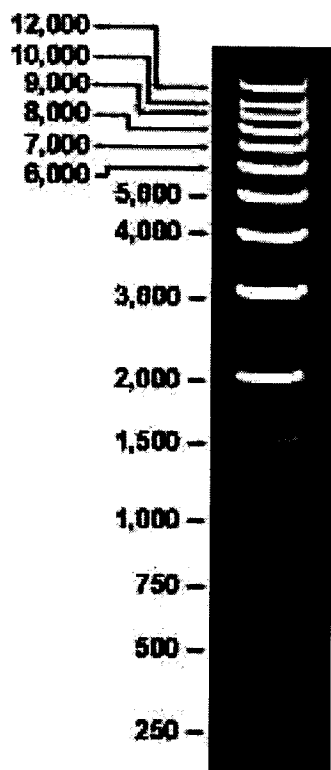


Figure 4-16 Smart Ladder®

The Smart Ladder® contains fragments from 1 kb to 10 kb at 1-kb increments, as well as bands at 250 bp, 500 bp, 750 bp, 1500 bp and 12 kb. The bands at the ladder's upper end are brighter than those at the lower end, with the transition to higher intensity at the 2-kb band serving as a useful reference point

Chapter 5: CONCLUSIONS AND RECOMMENDATIONS

The aim of this project was to create a cosmid library of *Arthrobacter sp.* M5 in *E. coli* HB101 and use this library to clone the acetophenone monooxygenase gene. This was achieved in part. A library was created but the gene was not located. This was due to unsuccessful screening techniques. As mentioned previously, screening Method (A) and Method (B) were unsuccessful in locating the acetophenone monooxygenase clone. A possible explanation for this is that an alternative pathway for the degradation of para-nitroacetophenone could be present in *E. coli* cells. *E. coli* cells in high concentration that had not undergone transformation were experimentally observed to be able to bring about the desired colour-change of colourless to yellow.

This was a problem because it was suspected that the expression of the cloned acetophenone monooxygenase was low and needed increased cell mass to bring about a colour change. Hence increasing the cell mass to increase expression also increased the interference from the alternative degradation pathway. More time was required to improve and develop this screening technique because unanticipated problems were encountered.

Screening Method (C) may have been successful if more time was available to achieve the correct stringency in order to attain a balance between unspecific binding of the probe to the total DNA and specific binding to the corresponding DNA. Even if the correct stringency is achieved, the probe was made from *Pseudomonas* ACB esterase sequences of DNA therefore there exists the possibility that binding between

the probe and the *Arthrobacter* sp. M5 total DNA would occur with difficulty because of inherent differences in the DNA of both species.

It was also the aim of this project to construct mutants of the associated esterase gene HapE that had already been cloned. Two point mutations (Asp483-Asn and Gly490-Ala) were introduced into HapE. This was confirmed by sequence data. The success of this section of the study has made it possible for further work to be carried out using these mutants to investigate the nature of the esterase active site.

It is recommended that screening the library should be attempted using colony blotting to locate the clone. This method was successful in cloning the HapE gene. It is a labour intensive method but is less prone to failure when performed with no time constraint.

It is also further recommended that the colorimetric screening techniques should be developed further to avoid interference from alternative degradation pathways present in the *E. coli* cells. This could be achieved by finding a compound, which can be used as an indicator for BVMO activity at low substrate concentrations.

Further work that can be carried out on the HapE mutants includes studies to ascertain the effect of different types of point mutations on the catalytic properties of the enzyme. The polarity and charge of amino acids in the motif regions of the enzyme are usually critical to the folding and catalytic mechanism of an enzyme. Such studies reveal the level of conservation of certain amino acids in the enzyme enabling their functional importance to be deduced.

Further development of screening techniques can be done with the aim of finding other compounds that can be used as colour indicators. Probes that can be used to find the acetophenone monooxygenase gene can be constructed using DNA from other members of the family of Baeyer-Villiger monooxygenases. This can be done in an attempt to achieve closer correlation between the DNA from different species that lowers the binding affinity of the different DNAs to each other.

References

-
- ¹ Alphand, V. and Furtoss, R. (1995) in *Enzyme Catalysis in Organic Synthesis* (Drauz, K. and Waldmann, H., eds), pp. 745-772, VCH
- ² Fried, J., Thoma, R. W. and Klingsberg, A. (1953) *J. Am. Chem. Soc.* **75**: 5764-5767
- ³ Baeyer, A. and Villiger, V. (1899) *Ber. Dtsch. Chem.* **32**: 3625-3633
- ⁴ Latham, J. A. and Walsh, C. (1987) *J. AM. Chem. Soc.* **109**: 3421-3427
- ⁵ Taschner, M. J., Peddada, L., Cyr, P., Chen, Q-Z. and Black, D.j. (1992) in *Microbial Reagents in Organic Synthesis* (Serve, S., ed.), pp. 347-360, Kluwer Academic.
- ⁶ Carrea, G. *et al.* (1992) *Tetrahedron Asymmetry* **3**: 1063-1068
- ⁷ Willets, A., Knowles, C. J., Levitt, M. S., Roberts, S. M., Sandey, H. and Shipston, N. F. (1991) *J. Chem. Soc., Perkin Trans. 1* 1608-1610
- ⁸ Trudgill, P. W. (1978) in *Developments in Biodegradation of Hydrocarbons-1* (Watkinson, R. J., ed.), pp. 47-84, Applied Science
- ⁹ Alphand, A., Archelas, A. and Furtoss, R. (1990) *Biocatalysis* **3**: 73-73
- ¹⁰ Donoghue, N. A., Norris, D. B. and Trudgill, P. W. (1976) *Eur. J. Biochem.* **63**: 175-192
- ¹¹ Chen, Y-C. J., Peoples, O. P. and Walsh, C. T. (1988) *J. Bacteriol.* **170**: 781-789
- ¹² G. Grogan, PhD Thesis, University of Exeter, 1995
- ¹³ Wierenga, R. K., Knowles, C. J., Levitt, M. S., Roberts, S. M., Sandey, H. and Shipston, N. F. (1991) *J. Chem. Soc., Perkin Trans 1* 1608-1610
- ¹⁴ Villa, R. and Willets, A. *J. Mol. Catalysis B: Enzymatic* (1997)
- ¹⁵ Alphand, V., Archelas, A. and Furtoss, R. (1989) *Tetrahedron Lett.* **30**, 3663-3664

-
- ¹⁶ Grogan, G., Roberts, S. M. and Willets, A. (1993) *J. Chem. Soc., Chem. Commun.* 699-701
- ¹⁷ Carnell, A. J., Roberts, S. M., Sik, V. and Willets, A. (1991) *J. Chem. Soc., Perkin Trans. 1* 2385-2389
- ¹⁸ Konigsberger, K. and Griengl, H. (1994) *Bioorg. Med. Chem.* **2**: 595-604
- ¹⁹ Petit, F. and Furtoss, R. (1995) *Synthesis* 1517-1520
- ²⁰ Gagnon, R., Grogan, G., Roberts, S. M. and Willets, A. (1995) *J. Chem. Soc., Perkin Trans. 1* 1505-1511
- ²¹ Levitt, M. S., Newton, r. f., Roberts, S. M. and Willets, A. (1990) *J. Chem. Soc., Chem. Commun.* 616-620
- ²² Adger, B. *et al.* (1995) *J. Chem. Soc., Chem. Commun.* 1563-1564
- ²³ Bes, M. T., Villa, R., Roberts, S. M., Wan, P. and Willets, A (1996) *J. Mol. Catalysis B: Enzymatic* **1**: 127-134
- ²⁴ Jones, J. B. and Jakovac, I. J. (1982) *Can. J. Chem.* **60**: 19-20
- ²⁵ Ottolina, G. *et al.* (1995) *Tetrahedron Asymmetry* **6**: 1375-1386
- ²⁶ Colonna, S., Gaggero, N., Pasta, P. and Ottolina, G. (1996) *J. Chem. Soc., Chem Commun.* 2303-2307
- ²⁷ Ottolina, G., Carrea, G., Colonna, S. and Ruckemann, A. (1996) *Tetrahedron Asymmetry* **7**: 1123-1136
- ²⁸ Kelly, D. P. (1996) *Tetrahedron Asymmetry* **7**: 1149-1152
- ²⁹ Kelly, D. P., Knowles, C. J., Hadhi, J. G., Taylor, I. N. and Wright, M. A. (1996) *Tetrahedron Asymmetry* **7**: 365-368
- ³⁰ Gagnon, R. *et al.* (1995) *J. Chem. Soc., Perkin Trans. 1* 2527-2528
- ³¹ Yudkin, M. and Offord, R. (1975) *Comprehensible Biochemistry*, Longman Group Limited, pp 109-135.

-
- ³² Pittol, C. A. , Roberts, S.M. Sutton P. W. and Williams, J. O. (1994), *J. Chem. Soc. Chem. Commun.* 803.
- ³³ Cotterill, I. C., Roberts, S. M. and Williams, J. O. (1988) *J. Chem. Soc., Chem. Commun.* 1628.
- ³⁴ Mahon, M. F., Molloy, K., Pittol, C. A., Pryce, R. J., Roberts, S. M., Ryback, G., Sik, V., Williams, J. O. and Winders, J. A. (1991) *J. Chem. Soc., Perkin Trans. 1*, 1255.
- ³⁵ Pittol, C. A., Pryce, R. J., Roberts, S. M., Ryback, G. and Williams, J. O. (1989) *J. Chem. Soc., Perkin Trans. 1*, 160.
- ³⁶ Downing, W., Latouche R., Pittol, C. A., Pryce, R. J. Roberts, S.M., Ryback, G. and Williams, J. O. (1990), *J. Chem. Soc. Perkin Trans 1*, 2613
- ³⁷ Hudlicky, T., Luna, H., Price, J. D. and Rulin, F. (1989) *Tetrahedron Lett.* **30**: 4053
- ³⁸ Carless, H. A. J. (1992) *J. Chem. Soc., Chem. Commun.* 234
- ³⁹ Walsh, C. T. and Chen, Y.-C. J. (1988) *Angew. Chem. Int. Ed.* **27**: 333
- ⁴⁰ Levitt, M. S., Newton, R.F., Roberts, S. M. and Willetts, A. J. (1990) *J. Chem. Soc., Chem. Commun.* 619
- ⁴¹ Fletcher, C. A., Hilpert, H., Myers, P. L., Roberts, S. M. and Storer, R. (1989) *J. Chem. Soc., Chem. Commun.* 1707
- ⁴² Highcock, R. M., Hilpert, H., Myers, P. L., Roberts, S. M. and Storer, R. (1991) *J. Chem. Soc., Perkin Trans. 1*, 1127
- ⁴³ Willets, A. J., Knowles, C. J., Levitt, M. S., Roberts, S. M., Sandey, H. and Shipston, N. F. (1991) *J. Chem. Soc., Perkin Trans. 1*, 1608
- ⁴⁴ Grogan, G., Roberts, S. M. and Willetts, A. J. (1993) *J. Chem. Soc., Chem. Commun.* 699

-
- ⁴⁵ Gagnon, R., Grogan, G., Levitt, M. S., Roberts, S. M., Wan, P. W. H., and Willetts, A. J. (1994) *J. Chem. Soc. Perkin Trans 1*, 2537
- ⁴⁶ Maguire, N., Read, G. and Roberts, S. M. (1994) *J. Chem. Res.*, S, 376 (M,2065)
- ⁴⁷ Sarneski, J. E., Michos, D., Thorp, H. H., Brudwig, G. W., Didiuk, M., Poon, T., Blewitt, J. and Crabtree, R. H. (1991) *Tetrahedron Lett.* **32**: 1153
- ⁴⁸ Davies, H. G., Dawson, M. J., Lawrence, G. C., Mayall, J., Noble, D., Roberts, S. M., Turner, M. K. and Wall, W. F. (1986) *Tetrahedron Lett.* **27**: 1089
- ⁴⁹ Floyd, N., Munyemana, F., Roberts, S. M. and Willetts, A. J. (1993) *J. Chem. Soc. Perkin Trans. 1*, 881
- ⁵⁰ Massey, V. (1994) *J. Biol. Chem.* **269**: 22459-22462
- ⁵¹ Ghisla, S., and Massey, V. (1986) *Biochem. J.* **239**: 1-12.
- ⁵² Lindqvist, Y. (1989) *J. Mol. Biol.* **209**: 151-166
- ⁵³ Hecht, H. J., Kalisz, H. M., Hendle, J., Schmidt, R. D., and Schomburg, D. (1993) *J. Mol. Biol.* **229**: 153-172
- ⁵⁴ Vrieling, A., Lloyd, L. F., and Blow, D. M. (1991) *J. Mol. Biol.* **219**: 533-544
- ⁵⁵ Ludwig, M. L., and Luschinsky, C. L., *Chemistry and Biochemistry of flavoenzymes* (Müller, F., ed) Vol. III pp. 427-466, CRC Press, U.S.A. (1992)
- ⁵⁶ Karplus, P. A., Daniels, M. J., and Herriot, J. R. (1990) *Science*. **251**: 60-66 (1990)
- ⁵⁷ Williams, C. H. Jr. *Chemistry and Biochemistry of flavoenzymes* (Müller, F., ed) Vol. III pp. 121-221, CRC Press, U.S.A., 1992.
- ⁵⁸ Van der Ven *et al.* (1995) *Microbial Degradation of Aromatic Compounds*, RUG internal study.
- ⁵⁹ Andy Vierstaete, Rijks Universiteit Gent, Belgium,
<http://allserv.rug.ac.be/~avierstr/pdf/PCR.pdf>, accessed on December 7, 2001
- ⁶⁰ Octavian Henegariu, Yale University

<http://info.med.yale.edu/genetics/ward/tavi/pA.html> , accessed on July 30 2000

⁶¹ Weiner, M. P., Gackstetter, T., Costa, G. L., Bauer, J. C. and Kretz, K. A.

<http://www.nwfsc.noaa.gov/protocols/pcr-mut.html> from *Molecular Biology:*

Current Innovations and Future Trends. Eds. A.M. Griffin and H.G.Griffin. 1995

Horizon Scientific Press, PO Box 1, Wymondham, Norfolk, U.K. Accessed on June

14, 2000

⁶² John W. Kimball, Harvard University,

<http://www.ultranet.com/~jkimball/BiologyPages/R/RestrictionEnzymes.html>

accessed on August 15, 2000

⁶³ Ish-Horowicz, D. and Burke, J. F. (1981) *Nucl. Acids. Res.* **9**, 2989

⁶⁴ Weizmann Institute of Science, Department of Biological Services

http://bioinfo.weizmann.ac.il:3456/kegg/codon_table/codon_eco.html , accessed

on July 20, 2000

⁶⁵ Moffat, B. A. and Studier, F. W., (1987) *Cell.* **49**: 221

⁶⁶ Karl Clark, Double Twist Inc.

http://www.protocol-online.net/molbio/DNA/dna_labeling.htm,

accessed on April 3, 2001

⁶⁷ Tomasek, P. H., and Crawford, R. L. (1986) *Bacteriol.* **167**: 818-827,

⁶⁸ Barton, M. R., and Crawford. R. L. (1988) *Appl. Environ. Microbiol.* **54**: 594-595

⁶⁹ Bedard, D. L., and Haberl, M. L (1990). *Microb. Ecol.*, **20**: 87-102

⁷⁰ Bedard, D. L., Haberl, M. L., May, R. J., and Brennan, M. J. (1987) *Appl. Environ. Microbiol.* **53**: 1103-1112

⁷¹ Utkin, I. B., Yakimov, M. M., Mateeva, L. N., Kozlyak, E. I., Rogozhin, I. S., Solomon, Z. G., Bezborodov, A. M. (1991) *FEMS. Microbiol. Lett.* **77**: 237-242

⁷² Cripps, R. E. (1975) *Biochem J.* **152**: 233-241

- ⁷³ Cripps, R. E., Trudgill, P. W., and Whateley, J. G. (1978) *Eur. J. Biochem.* **86**: 175-186
- ⁷⁴ Weissenfels, W. D., Beyer, M. Klein, J. and Rhem, H. J. (1991) *Appl. Microbiol. Biotechnol.* **34**:528-535
- ⁷⁵ Davey, J. F. and Trudgill, P. W. (1977) *Eur. J. Biochem.* **74**: 115-127
- ⁷⁶ Allen, J. E., Forney, F. W. and Markovetz, A.J. (1970) *Lipids.* **6**: 488-452
- ⁷⁷ Allen, J. E. and Markovetz, A.J. (1970) *J. Bacteriol.* **103**: 426-434
- ⁷⁸ Forney, F. W. and Markovetz, A.J. (1970) *J. Bacteriol.* **102**: 281-282
- ⁷⁹ Havel, J. and Reineke, W. (1993) *Appl. Environ. Microbiol.* **59**: 2706-2712
- ⁸⁰ Boyer, H.W. and Roulland-Dussoix, D., (1969) *J. Mol. Biol.* **41**: 459-465
- ⁸¹ Keddie, R. M. and Jones, D., (1981) Saprophytic, aerobic coryneform bacteria, P. 1838-1878. In Starr, M. P., Stolp, H., Truper, H. G., Balows, A. and Schlegel, H. G. (ed.), The prokaryotes. Springer-Verlag, Berlin.
- ⁸² Staskawicz, B. et al., (1987) *J. Bacteriol.* **169** 2: 5789-5794
- ⁸³ Studier, F.W. et al., (1990) *Methods Enzymol.* **185**: 60
- ⁸⁴ Rosenberg A. H. et al., (1987) *Gene.* **56**: 125-135
- ⁸⁵ Studier F. W. et al., (1990) *Meth. Enzymol.* **185**: 60-89
- ⁸⁶ Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn. Coldspring Harbour Laboratory Press, Cold Spring Harbour, New York.
- ⁸⁷ Recombinant DNA Principles and Methodologies. (1998) Edited by J. J. Greene and B. R. Venigalla. Marcel Dekker Inc. New York.
- ⁸⁸ Nanne Kamerbeek, personal communication

UNIVERSITY OF ZAMBIA LIBRARY

261734

