Development of a Novel High-Throughput Screening Assay and its Application to Racemases

Jemma Morgan

Doctor of Philosophy
2008

Abstract

A high-throughput hydrogen peroxide-based colorimetric screen used to detect oxidase activity was extended to detect racemase activity through the production of the specific substrate for an enantioselective oxidase enzyme. A two-plasmid system (encoding the racemase and oxidase) was shown to be successful and could have possible applications for many different classes of enzymes that produce an oxidase substrate.

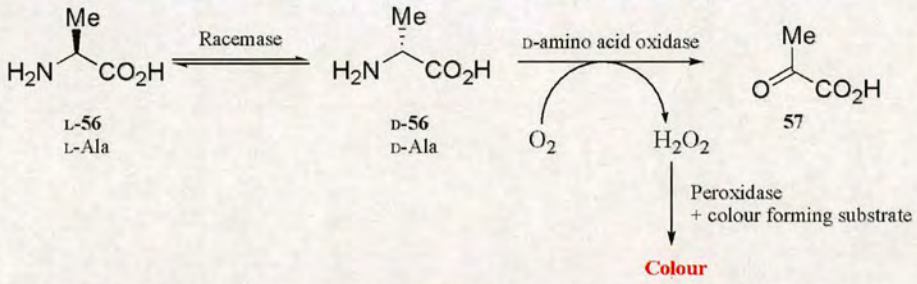


Figure 1: The screen for racemase activity using an oxidase/peroxidase cascade

In order to validate the screening technology, three amino acid racemase genes were cloned from genomic DNA and placed into expression vectors. The screen was used to confirm the substrate specificities of each racemase. The amino acid racemase genes were then subjected to random mutagenesis using the mutator strain method and error-prone polymerase chain reaction. Libraries of the variant racemase enzymes were screened for novel activities towards substrates not accepted by the wild type enzymes: L-arginine; L-lysine and L-leucine.

Novel activity was discovered in selected *Streptomyces coelicolor* alanine racemase variants. Sequencing revealed that the best variant had three point mutations, I195T, N223D and I374N. Computer modelling suggested that the I374N mutation was a key mutation and so the variant containing the double mutation (I195T and N223D) was prepared. Partial purification of the wild type enzyme and the variants containing the double and the triple mutations was carried out to compare the

substrate specificities. The I195T/N223D variant was shown to be ten times more active towards L-arginine that the wild type enzyme, and the variant containing all three mutations was shown to be 100 times more active towards L-arginine than the wild type enzyme. Both variants displayed novel activity towards L-arginine only.

This work has demonstrated and validated the novel screen for racemase activity and was successfully used to select variants with altered substrate specificity. The system should then be applicable to a range of screening applications and be a useful tool in the development of new biocatalysts.

Acknowledgements

First I would like to thank my supervisor Professor Nick Turner who gave me the opportunity to work in Biocatalysis, both in the UK and in Japan, who supervised this work, gave useful advice and encouragement throughout the project, and a big thank you for the final checking of this manuscript.

I would like to thank Dr Rob Speight, my industrial supervisor, who has been a great help and inspiration throughout the entire project, from buying the drinks on my first day, the many suggestions through the ups and downs of the project, and finally the penultimate proof-reading.

For the essential everyday support I would like to thank the rest of the Ingenza team, especially Dr Ian Fotheringham, Dr Nick Oswald, Dr Alison Arnold and Dr Ian Archer for providing guidance and tuition in many areas, not all scientific! A big thank you to all members, past and present, of the Turner/Flitsch group who have been friends and made the time in Edinburgh and then in Manchester enjoyable, especially Dr Kirk Malone who gave help with the molecular modelling.

A big thank you to John, for being there throughout the last four years and encouraging me to sit and write this thesis when the end seemed so far away, I am eternally grateful.

A big thank you to all my family, there are too many to name you all! You have all supported me so well and I can now finally finish studying and focus on a 'real' job.

The last thank you has to be for my Dad, without whom, I can honestly say, I would never have made it this far.

Abbreviations

4-AAP	4-Aminoantipyrine
AARAC	Amino acid racemase
ABA	Aminobutyric acid
Amp	Ampicillin
Arg	Arginine
Bn	Billion
BCA	Bicinchoninic acid
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
BVMO	Baeyer-Villiger monooxygenase
Cam	Chloramphenicol
CASTing	Complete active-site saturation test
CD	Circular dichroism
CE	Capillary electrophoresis
CFE	Cell free extract
Conc	Concentration
CV	Column volume
DAAO	D-amino acid oxidase
DAB	3,3'Diaminobenzidine
DCM	Dichloromethane
dNTP	Deoxy-nucleoside triphosphate
dH_2O	Distilled water
DKR	Dynamic kinetic resolution
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
DTT	Dithiothreitol
E	Enantiomeric ratio / or statistical similarity score in BLAST searching
e.e.	Enantiomeric excess
EDTA	Diaminoethanetetraacetic acid
EMDee	Enzymatic method for determining enantiomeric excess

epPCR	Error-prone PCR
EtOH	Ethanol
FACS	Fluorescence activated cell sorting
FAD	Flavin adenine dinucleotide (oxidised state)
$FADH_2$	Flavin adenine dinucleotide (reduced state)
FDA	Food and Drug Administration
gDNA	Genomic DNA
HIC	Hydrophobic interaction chromatography
HNL	Hydroxynitrile lyase
HPLC	High performance liquid chromatography
HRP	Horse radish peroxidase
ITCHY	Iterative truncation for the creation of hybrid enzymes
IPTG	Isopropyl-β-D-thiogalactopyranoside
IVC	*In vitro* compartmentalisation
Kan	Kanamycin
KCl	Potassium chloride
kDa	kilo Dalton
LB	Luria-Bertani medium
Lys	Lysine
MAO	Monoamine oxidase
α-MBA	alpha-Methylbenzylamine
MCS	Multiple cloning site
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
Min	Minute
MW	Molecular weight
MWCO	Molecular weight cut off
NaCl	Sodium chloride
Ni-NTA	Nickel-nitrilotriacetic acid
NEB	New England Bioloabs
OD	Optical density
OPA	*o*-Phthaldialdehyde

OPD	*o*-Phenylenediamine
pkDAAO	Pig kidney D-amino acid oxidase
PCR	Polymerase chain reaction
PLP	Pyridoxal 5'phosphate
PMS	Phenazine metasulfate
PMSF	Phenylmethylsulfonyl fluoride
Rac	Racemic
RACHITT	Random chimeragenesis on transient templates
ResQ	ResourceQ
RPM	Revolutions per minute
SDS	Sodium dodecyl sulfate
SDS PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
Ser	Serine
StEP	Staggered extention PCR
SOC	Super optimal broth-catabolite repression effect
TAE	Tris acetate
TBHBA	2,4,6-Tribromo hydroxybenzoic acid
Thr	Threonine
Tris	Tris(hydroxymethyl)aminomethane
TPQ	2,4,5-trihydroxyphenylalanine quinine
tvDAAO	*Trigonopsis variabilis* D-amino acid oxidase
Tyr	Tyrosine
UV	Ultraviolet
v/v	Volume by volume
w/v	Weight by volume

Contents

2. Results and Discussion

4. Experimental

1. Introduction

1.1 Biocatalysis

Biocatalysis is a rapidly expanding area of science that combines chemistry, biology, chemical and biological engineering. It is an enabling technology that offers unique options and solutions to problems faced in organic synthesis.[1] The broadness of application is supported and driven by the large number of enzymes rapidly becoming available and the range of reactions that they can carry out selectively and under ambient conditions. For example, a review in 2003 noted that 75 % of chemical racemisation techniques depend on harsh reaction conditions such as thermal racemisation, catalysis using strong acids or bases, or proceed *via* configurationally labile intermediates.[2] In contrast, racemase enzymes have optimum working temperatures of 20-45°C, can work at atmospheric pressure and prefer a slightly basic pH (7.5-9.0).[3] Developments in molecular biology, such as recombinant technologies, have provided access to many of the biocatalysts and a number are employed industrially for the manufacture of products such as lactose-free milk (>100 000 metric tons per year), acrylamide (>10 000 tons per year) and cocoa butter (>10 000 tons per year).[4]

Biocatalysts are frequently employed to differentiate between enantiomeric substrates. The industrial route by Lonza to both enantiomers of *N*-Cbz-proline is achieved from the racemate using acylases with different activities from *Arthrobacter* sp. and *Achromobacter xylosooxidans,*[5] Figure 1.1.

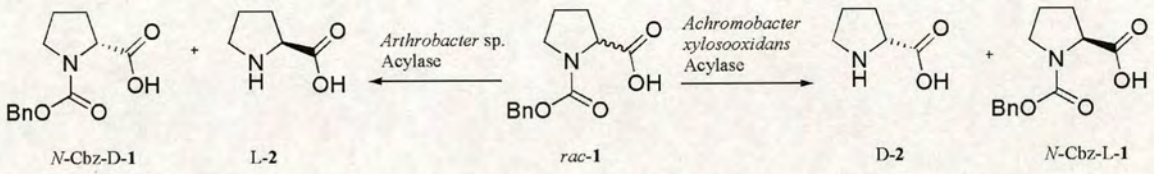
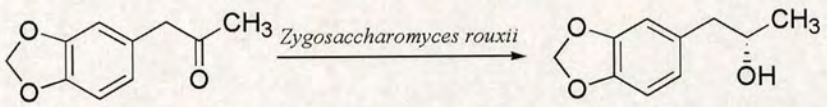


Figure 1.1: The separation of enantiomers of N-Cbz-proline using different acylases

As well as discriminating between different enantiomers, enzymes are able to synthesise enantiomerically pure products from prochiral starting materials. The reduction of carbonyl compounds to alcohols is achieved by chemists using metal reducing agents such as sodium borohydride. However, the metal reduction produces racemates unless chiral agents are included in the reaction mixture and problems can occur when other functionality is present in the starting material. Dehydrogenases can reduce ketones to produce chiral products with good e.e.s. The reduction of **3** (Figure 1.2) is carried out using *Zygosaccharomyces rouxii* to produce the alcohol (*S*)-(+)-**4** in 95 % yield and 99.9 % e.e..[6] (*S*)-(+)-**4** is used as an intermediate in the synthesis of (-)-talampanel, a compound with anti-epileptic properties.



3 (*S*)-(+)-**4**

Figure 1.2: The reduction of ketone 3 using a dehydrogenase

Although many enzymes and microbial sources have been discovered for carrying out a wide range of synthetic reactions,[7] there is not yet a biocatalyst available for every possible chemical reaction. Where possible biocatalysts have been identified, for example, by screening microbial libraries or analysing commercially available enzymes, they are often not suitable in the desired application as they may have limited substrate specificities, slow reaction times and/or poor stability in an industrial environment. Directed evolution is a tool that can be used to improve many properties of these biocatalysts and optimise them for the desired application.

1.2 Principles of Directed Evolution

Directed evolution is a technique that is carried out at the molecular level to alter the gene sequences of enzymes and ultimately the phenotypes. It has been used to broaden the substrate specificities of enzymes,[8, 9] to introduce enantioselectivity,[10] to improve stability under various conditions[11] and to completely change the reaction that the enzyme performs.[12] The principle of directed evolution relies upon the generation of many mutant genes and the subsequent screening or selection of the resulting enzyme variants for properties of interest as shown in Figure 1.3.

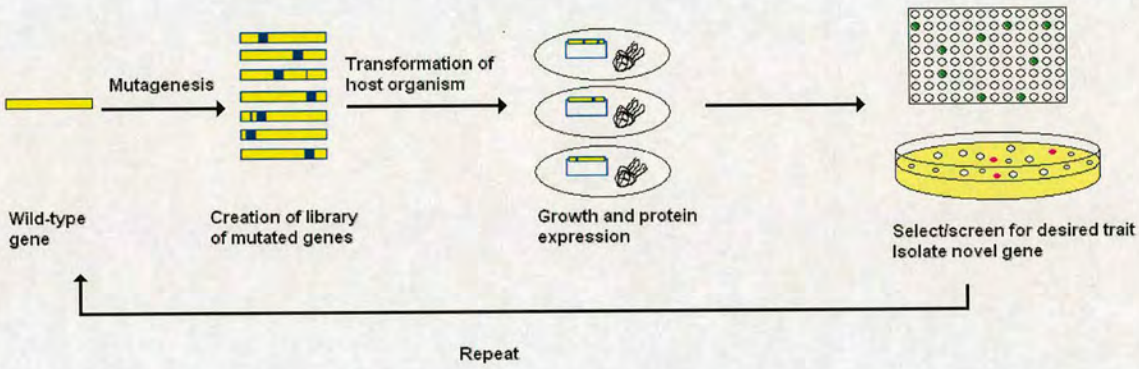


Figure 1.3: Principles of Directed Evolution

The parent gene is subjected to a random mutation process, such as error prone PCR. The resulting plasmid library is then used to transform an expression host, such as *Escherichia coli* and the resultant cells are grown and screened for activity using 'life or death' selection methods or a solid or solution phase assay.[13]

Iterative cycles of the mutagenesis/screening process can be used to produce an optimised enzyme. There are currently two main approaches to altering enzyme function: rational enzyme design and random mutagenesis. Despite some successes,[12, 14] rational enzyme design is limited by the complicated nature of enzymes, as well as by a general

lack of knowledge of enzyme/substrate recognition and catalysis. Quite often unpredictable changes have led to an improvement in enzyme function.[15] A limitation of rational design is that it is necessary to have a structural model to analyse. More frequently, random mutagenesis methods are employed to identify residues of importance for catalysis or enzyme stability, which are then subjected to site-directed mutagenesis.

1.3 Random mutagenesis

The first key step in directed evolution is to create a library of mutated genes. Many techniques have been used over the last 15 years to carry out mutagenesis[16] and new methods are constantly being developed.[17-19] There are numerous methods that can be presented for creation of mutated gene libraries and the most common and more established methods are briefly presented here. Random mutagenesis can occur at points along the gene resulting in single bases mutations, or over a much wider sequence space using gene recombination that results in longer strings of sequence being replaced with a homologous gene sequence. The two most commonly used techniques for point mutations are error-prone PCR and mutator strains.

1.3.1 Error-prone PCR

The most commonly used technique for random point mutagenesis is error-prone PCR (epPCR). During a PCR, different DNA polymerases incorporate mismatched nucleotides into the gene during synthesis at different rates. The DNA polymerase *Taq* makes approximately 8 errors per 10^6 base pairs, whereas *Pfu* hotstart polymerase has a lower error rate of 1 per 10^6 base pairs.[20] These natural error rates are too low to introduce enough mutations to produce a library of genes but they can be increased using various conditions.[21, 22]

One method to increase error rates during PCR is to introduce different concentrations of Mn^{2+} (to replace the Mg^{2+} natural cofactor of the polymerase) to aid mispairing; another method is to include biased ratios of the dNTPs to encourage mis-incorporation of bases. However, specific types of mutation are more common that others which leads to a bias in the library. For example, when analysing DNA bases, a pyrimidine (C or T) is more likely to be substituted by another pyrimidine, whereas a purine (A or G) will be more commonly substituted by a purine. A second bias to consider is that when a mutation is introduced in the early stages of amplification, all subsequent daughter strands will

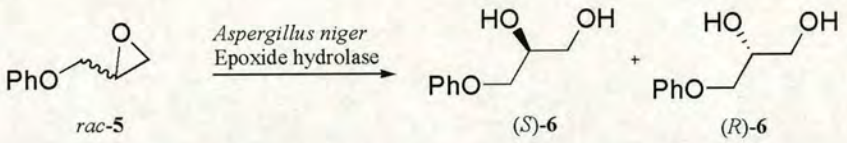
contain that mutation, leading to an over-representation in the final library. Finally, when translating the DNA, there is the codon bias that arises due to the nature of the genetic code. For example, three bases constitute a codon and if only one of those bases is changed at a time, there are only a limited number of amino acids that the variant codon can represent. One Ala codon is GCT, if the last base (T) is substituted for any other, the codon still represents Ala, if the second base (C) is varied the possible resulting amino acids are Ser, Pro and Thr, if the first base (G) is changed then the resulting amino acids are Val, Asp, Glu and Gly. Therefore, only seven other residues are obtainable starting with the GCT codon and substituting one base.

Companies such as BD Biosciences and Stratagene offer commercially available epPCR kits that are designed to overcome some of the biases and enable simple control of the error rate. The Stratagene kit uses a modified polymerase and the Diversify™ kit from BD Biosciences uses a combination of manganese addition, dNTP ratios and amplification cycles to produce variant genes with as little bias as possible.

The final size of the epPCR library is limited by the subsequent sub-cloning steps, such as the purification of the PCR fragments, restriction digestion, ligation and transformation. A procedure has been developed using the Quikchange site-directed mutagenesis protocol from Stratagene[23] in order to eliminate the ligation step. In this procedure, the epPCR fragment products are employed as primers in a whole plasmid PCR of the wild type template DNA. Therefore, the mutations that occur in the epPCR are incorporated into the plasmid DNA. The wild type template DNA can be eliminated after amplification by digestion with *Dpn*I, the restriction enzyme that recognises and digests methylated DNA only. DNA is methylated when replicated in *E. coli* cells, DNA synthesised during a PCR is not.

EpPCR was used as the mutagenesis technique for improving the enantioselectivity of an epoxide hydrolase from *Aspergillus niger* towards glycidyl phenyl ether.[24] The

enantiomeric ratio (E value) was improved from 4.6 to 10.8 in favour of the (*S*)-product in just one round of epPCR, Figure 1.4.



rac-**5** (*S*)-**6** (*R*)-**6**

Figure 1.4: The epoxide hydrolase catalysed opening of rac-5

1.3.2 Mutator Strains

XL1-Red is a mutator strain of *E. coli* designed by Stratagene that lacks efficient enzymes for DNA repair.[25] When the new DNA is synthesised prior to cell replication and division, the DNA sequence is checked for errors by the DNA polymerase. Mutations *mutT* and *mutS* are present in the DNA polymerase in the XL1-Red strain that lower the efficiency of the DNA repair. There is also a mutation in a subunit of DNA polymerase III (*mutD*) that increases the mutation rate during replication. Therefore, when the cells are growing and dividing, any mistakes made are not corrected and the error is carried forward into the daughter cells. The candidate DNA for directed evolution is used to transform the XL1-Red cells and the cells allowed to grow for set time periods, the longer the growth, the more mutations are introduced. Analysis for novel activities is then carried out. However, the mutations also occur throughout the plasmid DNA, not just in the gene of interest. Therefore, mutations may occur in the antibiotic resistant gene, or in the promoter region, that may affect the ability of the plasmid DNA to be sustained in the host and produce the enzymes of interest.

The mutator strain method of evolution is one of the simplest random mutation strategies and has been employed in a number of successful directed evolution experiments. One example of a successful directed evolution is the improvement in enantioselectivity of a monoamine oxidase from *Aspergillus niger* towards (*S*)-α-methylbenzylamine (MBA) with multiple cycles of mutagenesis using the XL1-Red mutator strain of *E.coli,* (Stratagene) and a colorimetric screen.[10, 26, 27] The V_{max} value was increased 100-times

from $1.29x10^{-2}$ μmol/min/mg in the wild type, to 1.61 μmol/min/mg for the new variant. The corresponding values for the (*R*)-enantiomer were $1.42x10^{-3}$ μmol/min/mg for the wild type enzyme and $2.55x10^{-2}$ μmol/min/mg for the new mutant. The greatest V_{max} increase was seen with the same variant versus the wild type enzyme when using 1-methyl-1,2,3,4-tetrahydroisoquinoline (MTQ) as substrate. The values for the wild type on (*R)* and (*S)*-MTQ were $6.18x10^{-4}$ and $1.48x10^{-3}$ μmol/min/mg respectively, in the new mutant these values are $2.27x10^{-2}$ and 1.99 respectively.

1.3.3 Oligonuceotide based methods, random mutations at specific points

Both epPCR and the mutator strain introduce random mutations at random points in the gene. Using oligonucleotide methods the area(s) of randomisation can be specified in a process known as saturation mutagenesis. For example, if a number of residues were hypothesised as important for defining the active site and so the substrate specificity, it is possible to design degenerate primers that substitute the codons in the area of interest to result in every other possible residue. Libraries of those variants can then undergo activity screening to determine the effects of those substitutions.

Site specific randomisation has recently been employed in the method of Combinatorial Active site Saturation Testing, abbreviated to CASTing.[9, 28, 29] Areas of potential importance for controlling substrate specificity were identified using computer modelling and subjected to randomisation in a step-wise fashion. The gene of the most active variant, discovered after one site was randomised, was used as template for the proceeding mutagenesis experiment. To compare CASTing to other methods of mutagenesis, the same gene for the epoxide hydrolase shown in Figure 1.4 and improved using epPCR, was subjected to CASTing.[29] By screening a total number of 20,000 clones from random step-wise mutagenesis of five areas, the E value of the reaction was improved from 4.6 to 115 in favour of the (*S*)-product. CASTing was also used to evolve a Baeyer-Villiger cyclopentanone monooxygenase (CPMO) to improve the resulting e.e.

in the oxidation of 4-methoxycyclohexanone.[28] The e.e. was improved from 46 % to 92 % in favour of the (*R*)-enantiomer.

1.3.4 Gene Shuffling, a recombination approach

Gene shuffling is a method frequently used to achieve gene recombination and therefore increased diversity and was first introduced in 1994.[30] The technique generates gene variants from a set of homologous genes by digestion and recombination.[22] Sets of homologous genes are digested with DNase I to create DNA fragments, which are then reassembled in a PCR reaction without using separate primers. The fragments of the genes are denatured at a high temperature, followed by a temperature reduction that allows DNA fragments with complementary areas (but not necessarily from the same parent gene) to anneal. A polymerase extension then reconstructs areas of the genes where there is no fragment overlap. As the original genes used are homologous, recombination can occur between two different genes, creating hybrids that can be analysed for novel activities.[30]

Gene shuffling was used to evolve an alanine racemase with a higher expression level from two biosynthetic alanine racemases, one from *E. coli* and the other from *Salmonella typhimurium.*[31] After one round of shuffling, a novel alanine racemase was produced that had three times the specific activity, and was expressed twice as much in *E. coli* as the parental enzymes. The alanine racemase activities of the *E. coli* and *S. typhimurium* parental racemases were 9.9 and 2.3 units mg^{-1}, whereas the best chimera had a specific activity of 25.6 units mg^{-1}.

1.3.5 Other methods of recombination

Techniques, such as staggered extention PCR (stEP), random chimeragenesis on transient templates (RACHITT) and iterative truncation for the creation of hybrid enzymes (ITCHY), have been developed to produce libraries of chimeric enzymes.[22] Both RACHITT and ITCHY rely on the fragmentation of the parent genes and

reassembly. However in stEP, homologous genes are combined and used whole in a PCR reaction that has very short extension times. The principle behind the technique is that only a very short gene sequence of template 1 is amplified before a rise in temperature causes the primer and the template to separate. During the next annealing process, the new elongated primer can bind to template 2 and amplify a further short gene sequence.

RACHITT involves the creation of the complementary strand of one of the genes of interest. The other genes are then fragmented and the two DNA strands separated. The fragments are then annealed to the first whole gene and any gaps are removed using a polymerase. The template strand is then removed and the novel chimeric gene amplified to produce a whole gene.

With all stEP and RACHITT, it is necessary to have sets of homologous genes whereas ITCHY allows recombination of genes with little similarity. The genes of interest are digested with the same restriction enzymes and then the fragments are ligated together to form novel variants that can be screened for traits of interest.

There are many different techniques used to create libraries of genes, both at random sites and at specific sites (many more than covered here) and the method chosen depends on the researchers aims and the information available concerning the gene of interest.

1.4 Screening for novel activities- Current methods

The first key step in directed evolution is mutagenesis and in the second step it is essential to analyse the variant libraries for new activities. Just as many methods have been developed for the library creation, there are many methods that have been developed for screening/selection to detect improved or novel enzymatic activities.[32] Screens for novel enzyme activity also have application for analysing microbial isolates and libraries for various enzymatic conversions. Bacterial and fungal species are constantly being isolated from the environment and being able to characterise them in terms of their reactivities aids the area of biotechnology greatly.

The analysis of large libraries of variant enzymes has long been seen as the bottleneck in directed evolution.[33] Although there are always new techniques and ideas of ways to produce vast gene libraries (Section 1.3), the development of screens to analyse those libraries in a high-throughput format has not been so rapid, even though the outcome of directed evolution experiments is critically dependent on how that library is screened. Over the last few years numerous groups have published new ideas and introduced novel concepts for library screening,[34] both on agar plates and in micro titre-plate format and a number of those techniques are presented and discussed here.

1.4.1 Screening versus Selection

There are two concepts for detecting novel enzyme activities: Selection and Screening. Selection links the characteristic of interest directly to the ability of the expression host to survive. For example, the alanine racemase chimeras created in the DNA shuffling mutagenesis experiment presented in Section 1.3.4 were analysed using a selection screen. As well as evolving racemases with higher expression levels, Ju *et al* were trying to evolve a serine racemase.[31] The group made an *E. coli* serine auxotroph strain that was unable to grow on minimal media unless a source of L-serine was present and then they introduced D-serine into the minimal media. If the variant racemase expressed in

the *E. coli* was able to racemise D-serine, the cell would have a source of L-serine and be able to grow. The phenotypic selection method of library analysis greatly reduces the final number of variant enzymes that need to be analysed in detail as only those colonies that grow can carry out the reaction of interest. However, the use of the selection procedure is limited to screening only for reactions that are of direct biological relevance or that can be indirectly linked to a selectable phenotype. Another disadvantage is that microbes are highly adaptable and can deal rapidly with significant changes in environmental conditions and therefore the incidence of false positives is increased.

The more common method of analysing variant enzyme libraries is to screen for the reaction of interest where the product(s) or by-product(s) are easily identified, either visually or by instrumental analysis.

1.4.2 Agar plate-based assays

The simplest and most high-throughput method to screen for novel activities is to monitor the cells when they are growing on an agar plate. A single 8 cm diameter plate can hold up to 3,000 separate individual colonies and the ability to screen for activity at this level enables the rapid analysis of vast enzyme libraries. Currently, there are three approaches to screening on agar plates:[35] (i) by detecting a shift in pH; (ii) by the detection of a halo formed around a growing colony, indicating cleavage of an insoluble substrate; or (iii) by detecting the formation of a coloured or fluorescent product.

1.4.2.1 Agar plate-based assays: pH

Screens using pH are used when the reaction of interest generates an acidic or basic product that causes a shift in the pH in the area around the host colony. An indicator dye highlights where the reaction is taking place by changing colour. In 2002, Littlechild *et al* developed a screening method for Baeyer-Villiger Monooxygenases (BVMOs) using a pH indicator solution and an esterase.[36] Once grown, colonies expressing the BVMOs

were lifted using a filter paper method and placed onto fresh agar containing the substrate for the cyclohexanone monooxygenase **7**, Figure 1.5.

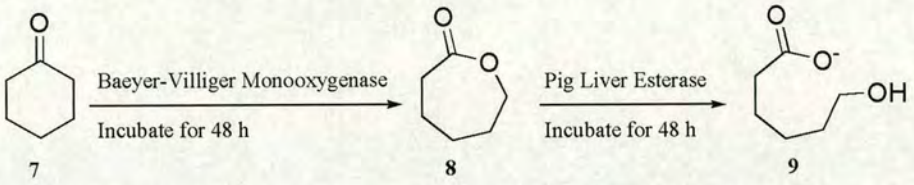


Figure 1.5: The detection of BVMO activity using Pig Liver Esterase

After incubation for 48 h to allow sufficient build up of **8**, the filter papers were sprayed with pig liver esterase. After a second 48 h incubation, the colonies were sprayed with a full range pH indicator solution to detect formation of acid **9**. Colour differences indicated the presence of a BVMO that was able to oxidise **7**. A range of substrates were tested to determine the applicability of the screen as the pig liver esterase substrate specificity determines the substrates that can be used.

Bornscheuer first detected the novel activity of a variant pig liver esterase from *Pseudomonas fluorescens* (generated using the mutator strain method, Section 1.3.2) on agar plates using pH, in 1998.[37] The activity of the esterase was evolved towards **10**, an intermediate in the synthesis of Epithilon A (Figure 1.6) and activity was identified by detection of an increase in medium acidity.

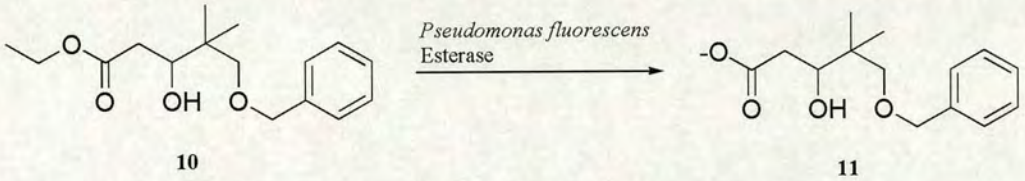


Figure 1.6: First use of pH change detection to monitor esterase activity

More recently, Schwab *et al* combined the shift in pH with a fluorescent indicator.[38] Colonies expressing an esterase from *Pseudomonas fluorescens* were grown on a nylon membrane. The membrane was then transferred onto a gel containing esterase substrate,

fluorescein and buffer. The acid generated in the ester hydrolysis was used to quench the fluorescence of fluorescein. A camera was used to monitor fluorescence and a decrease was shown to correspond to esterase activity in the colonies.

The use of pH based assay methods on agar plates is limited to detection of enzymes that produce, or can be coupled to a second enzyme to produce, acidic or basic products.

1.4.2.2 Agar plate-based assays: Halo

Halo assays are used to detect hydrolytic enzyme activity by cleavage of an insoluble substrate distributed around the growth medium of the colony. Cleavage of the substrate results in solubility and a visible zone clearing in the turbid medium. Halo assay methods are commonly used for lipases and proteases.[35] Initial work with zone clearing detection methods were carried out using suspensions of casein or skim milk proteins to detect protease activity.[39] Reetz discusses the use of tributyrin, an insoluble compound, in a screen for lipase activity.[40] This group identified active lipases by visual analysis of colonies growing on solid media containing tributyrin that were surrounded by a clear zone.

In 2004, a group in Korea reported the development and use of a zone clearing, plate-based screen to detect polysaccharide-degrading activity.[41] Different insoluble coloured forms of amylose, xylan and hydroxyethyl-cellulose (HE-cellulose) were prepared and distributed through the solid growth media of various microorganisms, isolated from soil and sludge. Insoluble amylose-blue, amylose-red, xylan-blue, HE-cellulose-blue and HE-cellulose orange were prepared using the dyes Cibacron Blue 3GA, Cibacron Brilliant Red 3B-A and reactive Orange 14. Microorganisms displaying amylase, xylanase and/or HE-cellulase activities were identified by appearance of red, blue and/or orange haloes around the growing colony.

Zone clearing provides an indication of the expression and activities of the enzymes of interest as the size and clarity of the halo is a measure of activity. The colorimetric halo assay described above was also shown to detect different activities on one plate by using combinations of the coloured substrates. In summary, zone clearing can be a very quick test for detecting hydrolytic enzyme activities and can give an indication of both the expression level and activity.

1.4.2.3 Agar plate-based assays: Colorimetric and Fluorescent

The two agar plated based methods for screening enzyme activities presented so far provide rapid detection of hydrolytic enzymes and those that produce acidic or basic products (or can be coupled to enzymes that can). In the halo assay it is crucial that the enzyme converts an insoluble compound into a soluble product. The criteria limit the applicability of the halo and the pH based screening methods to more diverse enzymes. However, there are many other reactions that enzymes carry out that are desirable to detect.

The most common plate-based assays rely on colorimetric or fluorescent detection. In a recent overview of plate-based assay methods by Turner, eight out of the nine examples presented identified enzyme activity using colorimetric or fluorescent methods.[33] The method relies on the incorporation of a chromophore into the enzyme substrate, or a cascade reaction that results in colour formation.

The most well-known plate-based colorimetric screen is the use of 5-bromo-4-chloro-3-indoyl-β-D-galactose (X-gal) **12** for detecting galactosidase activity (Figure 1.7) and is an example of the substrate for the enzyme containing a chromophore. Upon cleavage of the sugar from the heterocycle, compound **13** is released and rapidly dimerises to form an insoluble blue compound **14.** The cleavage of X-gal is used widely in molecular biology for Blue/White screening. The principle behind the screening is that the gene for the α-protein of galactosidase is on a cloning plasmid and the β-protein is encoded by

the *E. coli* host. When both proteins are successfully produced, they combine to create an active galactosidase that will cleave **12** to produce a blue colour in a growing colony. Disruption of the α-fragment gene on the cloning plasmid by successful insertion of another gene prevents production of the α–protein and so does not result in active galactosidase, therefore in the presence of X-gal the colony will remain white. Therefore plasmids with a gene successfully inserted are easily identified.

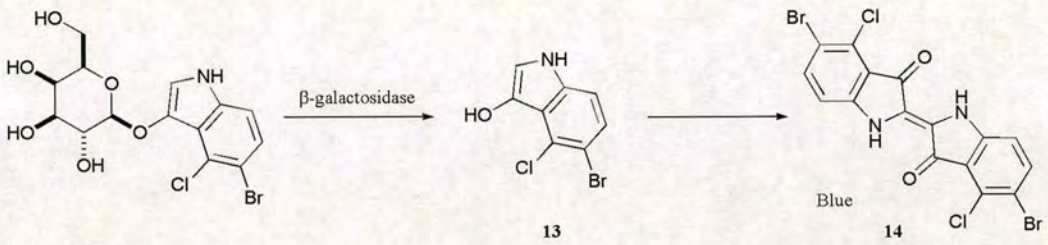
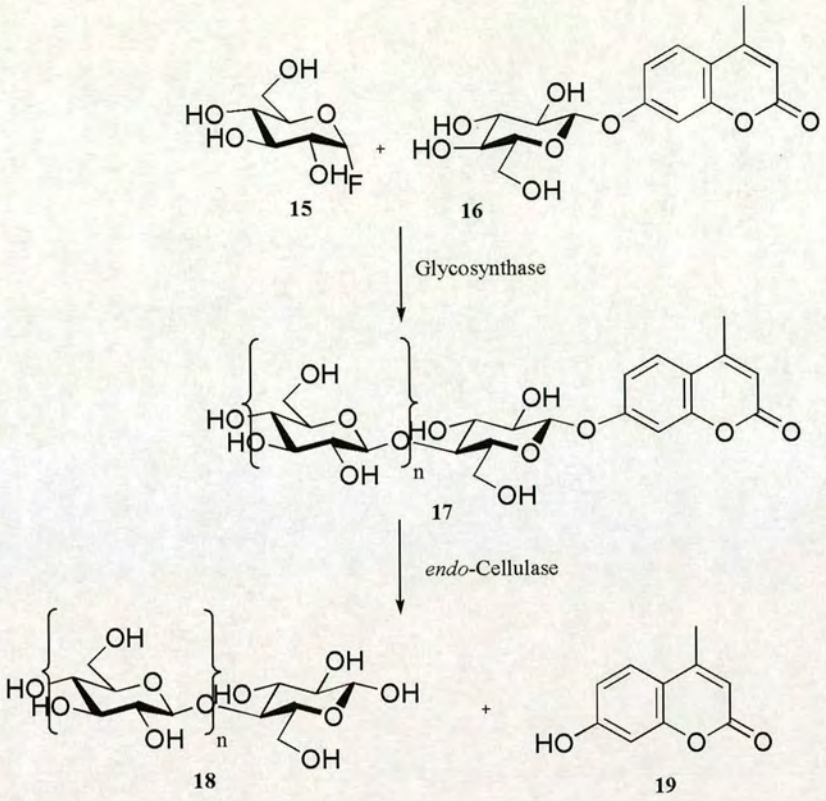


Figure 1.7: Cleavage of X-gal 12 to form blue compound 14

Recently, a very similar colorimetric screen was used in the directed evolution of β-galactosidase towards β-fucosidase activity using 5-bromo-4-chloro-3-indoyl-β-D-fucopyranosidase (X-fuc) instead of X-gal.[42] Due to release of the same heterocycle **13**, formation of the blue compound could be used to detect β-fucosidase activity.

The use of a screening reaction utilising release of fluorescence was developed in the screening for new glycosynthases.[8, 43] The sugar acceptor for the novel glycosynthase enzyme contained a fluorophore which was then incorporated into the glycosynthase product. The *endo*-cellulase was then used to cleave and release the fluorescent side chain, 4-methylumbelliferone **19** and indirectly confirm active glycosynthase, Figure 1.8.



15 16 17 18 19

Figure 1.8: Detection of glycosynthase activity by release of 4-methylumbelliferone 19

In the above techniques, the substrates containing the chromophores are analogs of the natural substrates or the substrate of interest and therefore the selected enzymes may not display the same final activities towards the substrate of interest. A different approach is to use a cascade reaction that produces colour or fluorescence where the enzyme acts on the actual substrate of interest.

In a screen of microbes for the deracemisation of α-amino acids,[44] Ohta *et al* used the conversion of nitroblue-tetrazonium chloride to formazan (blue) to detect the formation of L-Phe (from D-Phe) via reaction with an L-Phe dehydrogenase and regeneration of the NAD^+ cofactor using phenazine metasulfate (PMS), Figure 1.9.

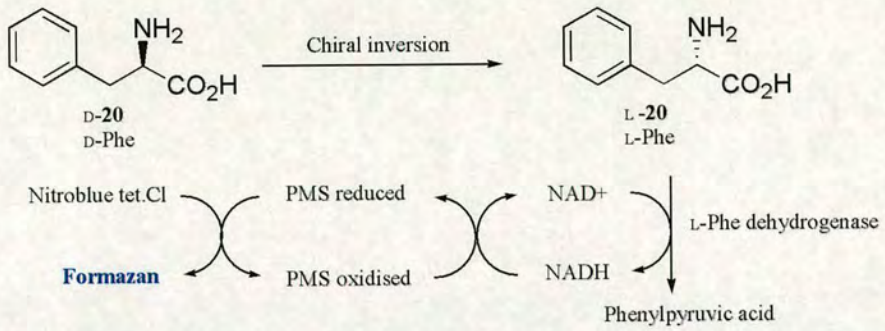
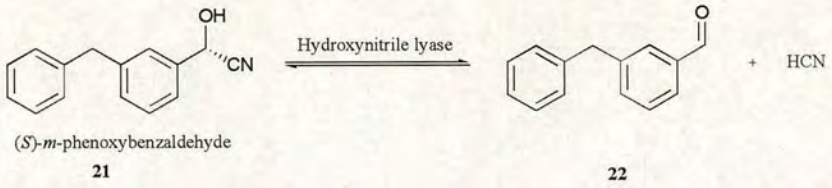


Figure 1.9: Detection of microbial chiral inversion via a cascade reaction

Fifty strains with unknown chiral inversion activities were analysed using the colorimetric plate-based assay and nine were found that could produce L-Phe from D-Phe (the route was shown to proceed via the keto acid and not racemisation with a racemase). The best candidate microbe for D-Phe racemisation was identified as *Nocardia diaphariozonaria*.

The cascade system of enzyme activity identification does not necessarily rely on secondary enzymatic reactions. A recent example (2007) used a copper containing solution to identify the HCN by-product of a hydroxynitrile lyase (HNL).[45] Colonies expressing variant hydroxynitrile lyases were grown on a membrane on an agar plate. The membrane was then removed from the agar and soaked with substrate **21,** Figure 1.10. A second filter paper previously soaked with a chloroform solution of copper ethylacetoacetate and 4,4'-methylenebis[*N,N*-dimethylaniline] and dried, was placed on top of the colony covered membrane. The hydroxynitrile lyases that successfully converted **21** to **22** were identified by formation of HCN and the resulting blue colour (the oxidation product of the 4,4'-methylenebis[*N,N*-dimethylaniline] base in the presence of Cu^{II}).



(*S*)-*m*-phenoxybenzaldehyde
21

22

Figure 1.10: Hydroxynitrile lyase conversion of 21 to 22 and HCN

The advantage of the assay is that it is not substrate specific but relies on the formation of gaseous HCN. Therefore any hydroxynitrile lyase reaction could, in theory, be monitored using this assay.

1.4.3 High-throughput screening in microtitre plates

Agar plate-based assays are efficient for looking at the same reaction in many thousands of colonies and providing a Yes/No answer for activity. For quantitative assays, small volume samples can be analysed in 96, 384 or 1536 well microtitre plates and recently Verenium have developed GigaMatrix® ultra-highthroughput screening technology with up to one million wells in a plate with the same footprint as the 96-well plates.[46] Whole cells, cell lysates and pure enzymes can be assayed for small variations in activity. Microtitre plate-based assays rely on the detection of a coloured or fluorescent product formed either by an enzyme coupled screen, coloured product release during reaction, or coordination to a metal. Recent examples of screens developed in each of these areas are discussed below and demonstrate that high-throughput screening is still very much an area of developing research.

1.4.3.1 Enzyme coupled assays

The majority of enzymatic reactions do not produce any detectable changes in UV/Vis or fluorescence during substrate turnover. A common method of assay is to couple the enzyme of interest with a second enzyme that is $NAD(P)^+$-dependent. In agar plate-based assays, the change in oxidation state of the $NAD(P)^+$ cofactor can be revealed by coupling to nitroblue tetrazolium and phenazine methosulfate (PMS), Figure 1.9.

However, in microtitre plates, NAD(P)-dependent enzymes are easily detectable by a change in absorbance at 340 nm during cofactor oxidation/reduction.

In 2001, Bornscheuer developed a method for screening for hydrolases by coupling the hydrolysis to an NADH-dependent L-malate dehydrogenase.[47] The acetic acid produced during the reaction of interest was used stoichiometrically in the conversion of NADH to NAD^+, Figure 1.11. The acetic acid is captured by an acetyl-CoA synthase which then produces acetyl-CoA. A citrate synthase then catalyses the subsequent reaction between acetyl-CoA and oxaloacetate, where oxaloacetate is formed from L-malate using an NADH-dependent dehydrogenase. The screen assay is widely used by food industry for detecting acetic acid in food products. The esterase from *Pseudomonas fluorescens* was used as the test enzyme and activity was successfully detected. A method was also developed for determining E values using initial rates of reaction using the same screen.

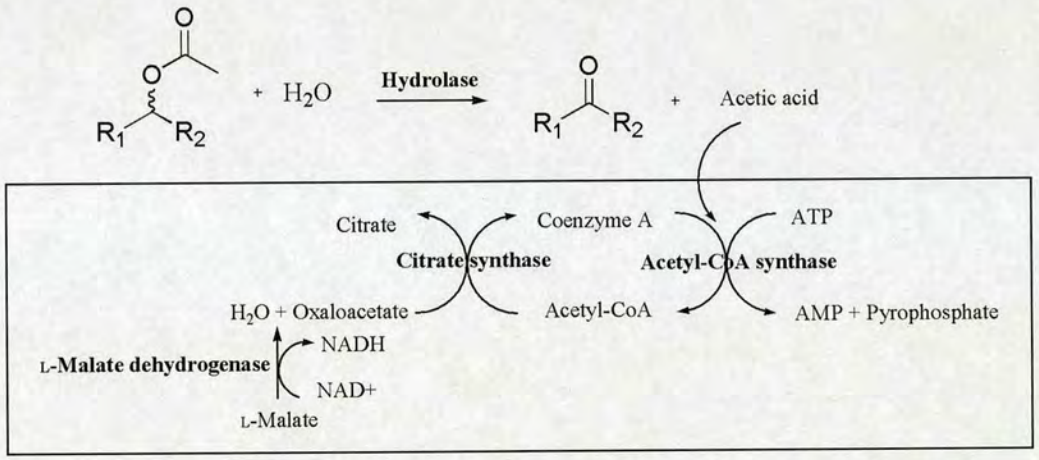


Figure 1.11: NAD^+ coupled screen to detect hydrolase activity

NAD(P)-dependent enzymes have also been used in other high-throughput screens to determine enantioselectivities. In 2001, an enzymatic method for determining enantiomeric excess (EMDee) was developed by monitoring the formation of NADPH by an alcohol dehydrogenase in a 384-well plate using a UV/fluorescent reader, Figure

1.12.[48] Here the e.e. of the (*S*)-isomer in a mixture of (*R*) and (*S*)-**23**, produced by a metal catalysed reaction, was measured using an (*S*)-alcohol dehydrogenase.

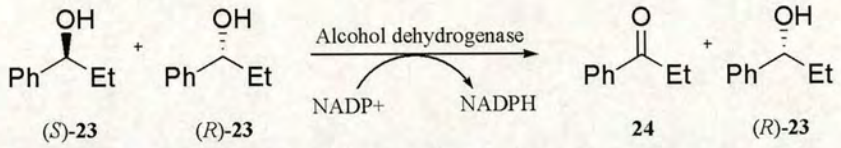


Figure 1.12: Using an alcohol dehydrogenase to determine e.e. values

The rate of oxidation of a range of known ratios of (*R*)- and (*S*)-**23** by the alcohol dehydrogenase was measured to make a standard plot of relative rates versus e.e., to which the unknown sample was compared. The method was effective but only worked when the total concentration of **23** was known. In 2004, the method was further developed using two alcohol dehydrogenases with different enantioselectivities to allow the determination of the e.e. of a range of chiral alcohols of unknown concentrations.[49]

1.4.3.2 Coloured or fluorescent product detection

The simplest way to detect enzyme reaction is by formation of a coloured or fluorescent product. In 2005, an assay was developed to detect BVMO activity in microtitre plates using fluorescence.[50] Standard substrates for the BVMOs were substituted with a fluorescent compound such as umbelliferone **27**, for example the cyclohexanone-umbelliferone **25**, and upon oxidation the product hydrolysed and was released to give a fluorescent signal, Figure 1.13.

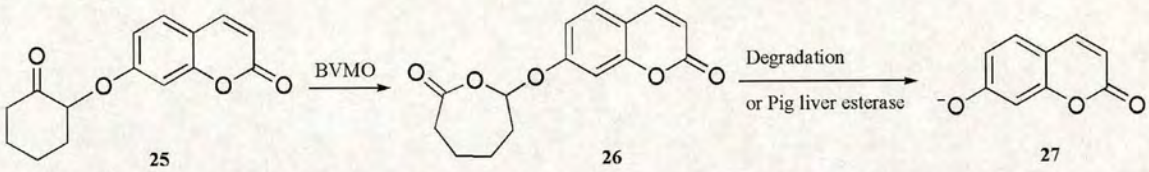
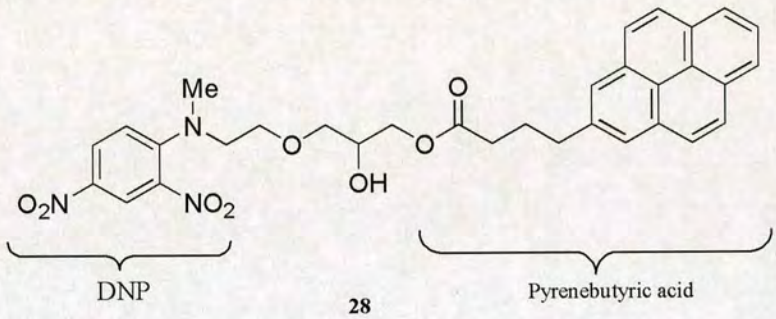


Figure 1.13: BVMO assay using release of fluorescence to detect activity

A range of cyclic and aliphatic substrates were successfully used and the fluorescent signal was increased in the presence of pig liver esterase.

In 2006, the same group developed an assay for lipases and esterases using fluorescence resonance energy transfer (FRET).[51] Pyrenebutyric acid, a fluorescent reporter, was covalently linked using various spacers to a dinitrophenylamino (DNP) group, Figure 1.14. As the emission spectrum of the pyrene overlaps with the emission spectrum of the DNP, the fluorescence was quenched. When the pyrene and the DNP were separated from each other by hydrolysis using the lipase or esterase, fluorescence was observed. Different spacer groups were used to test the substrate specificities of the lipases and esterases. Thirty-six commercially available enzymes were tested against fifteen substrates in a microtitre plate assay and compound **28**, Figure 1.14, was identified as an efficient reporter system for lipase activity under basic conditions.



28

Figure 1.14: FRET substrate for analysing Lipase/Esterase activity

An EMDee method has also been developed using fluorescence to determine e.e.s of allylic acetates using a lipase by the group that developed the original EMDee assay using NADPH-dependent enzymes.[52] The hydrolysis of **29** by a lipase released acetate that protonated *p*-nitrophenol **31** present in the assay solution, Figure 1.15. The protonation resulted in a change in the absorbance at 405 nm for the *p*-nitrophenol, which was measured for a range of allylic acetate enantiomeric ratios. The e.e. of the acetates was then plotted against relative initial rates and could be used to determine e.e.s of unknown samples.

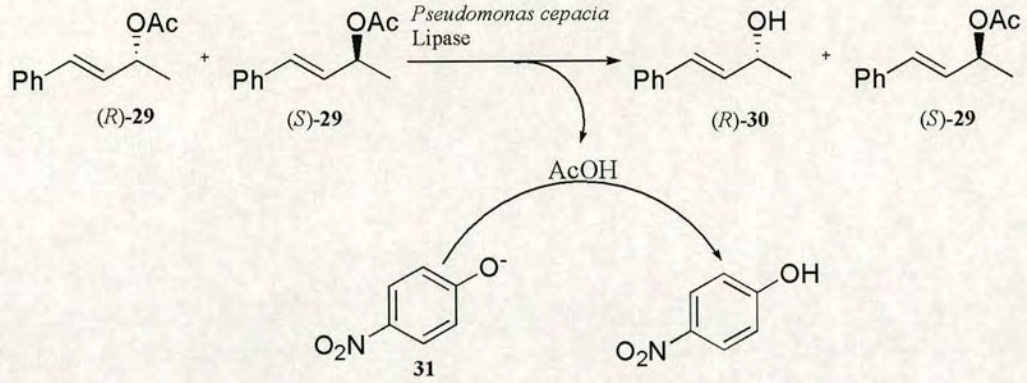


Figure 1.15: Using p-nitrophenol to detect lipase activity

The advantage of this method is that it is not substrate dependent since it detects the release of acetate, a by-product of many lipase reactions. Eighty-eight samples were analysed in twenty-five minutes providing a rapid indication of lipase activity and the e.e. of the reaction.

1.4.3.3 Detection of enzyme activity using metal coordination

In 2004, Hwang and Kim developed a method to detect ω-transaminase activity using a copper sulfate/methanol system.[53] In the ω-transaminase reaction, an amino donor **32** and an amino acceptor **33** were converted into a ketone **34** and an amino acid **35**. The resulting amino acid was shown to form a complex with the copper ion in the copper sulfate to produce a blue colour, which was quantified by measurement at 595 nm, Figure 1.16. This method was used to detect a novel transaminase with three fold improved reactivity towards 3-amino-3-phenylpropionic acid compared to the wild type.

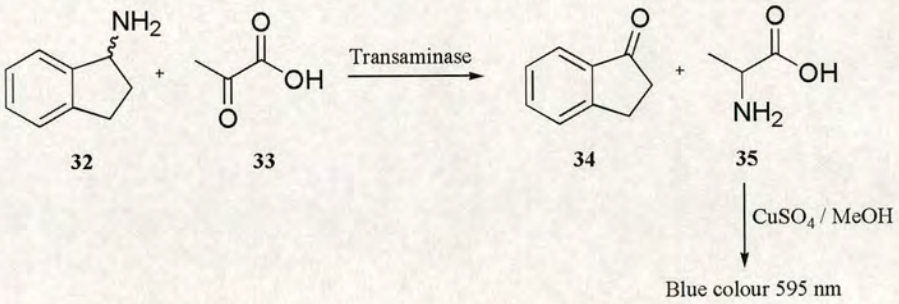


Figure 1.16: Using a copper staining solution to detect amino acid and transaminase activity

A similar copper staining method was also used for the detection of lipase activity using dimedone.[54] A range of lipase substrates were synthesised that contained dimedone, for example **36**, shown in Figure 1.17. Upon enzyme-catalysed hydrolysis, the dimedone was released **38** and bound to Cu^{II} present in the microtitre plates as $Cu(OAc)_2$. The dimedone-Cu adduct was green, absorbed at 417 nm and was detected using a UV/Vis plate reader.

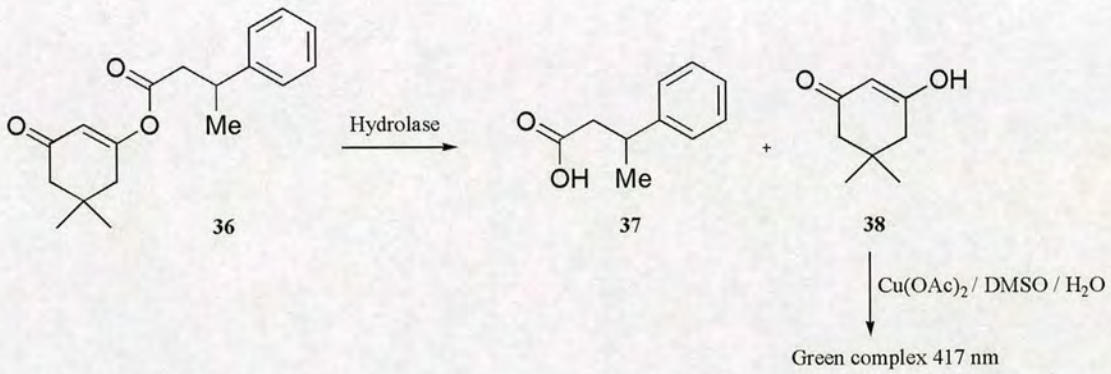


Figure 1.17: Hydrolase screening by detection of the copper-dimedone adduct

Eight lipases and two esterases were used to validate the screen using substrate **36.** The commercially available lipases displayed activity comparable to activities measured by HPLC. However crude enzyme preparations displayed a high level of background activity and so the assay could only be used with pure enzyme preparations.

1.4.4 High-throughput screening using other media

Enzyme screening systems have also been developed to monitor chemical reactions where the reagents were insoluble in water. In 2002, Berkowitz *et al* developed a biphasic system in which the chemical conversion occurred in an upper organic layer[55] (intramolecular allylic aminations of **39** using transition metal catalysts), Figure 1.18. Ethanol was produced as a by-product that diffused into the lower aqueous layer. The alcohol was then oxidised twice by the NADH dependent enzymes alcohol dehydrogenase and aldehyde dehydrogenase. The formation of NADH was monitored at

340 nm. It was shown that six reactions with a reaction time of ten minutes could be screened simultaneously.

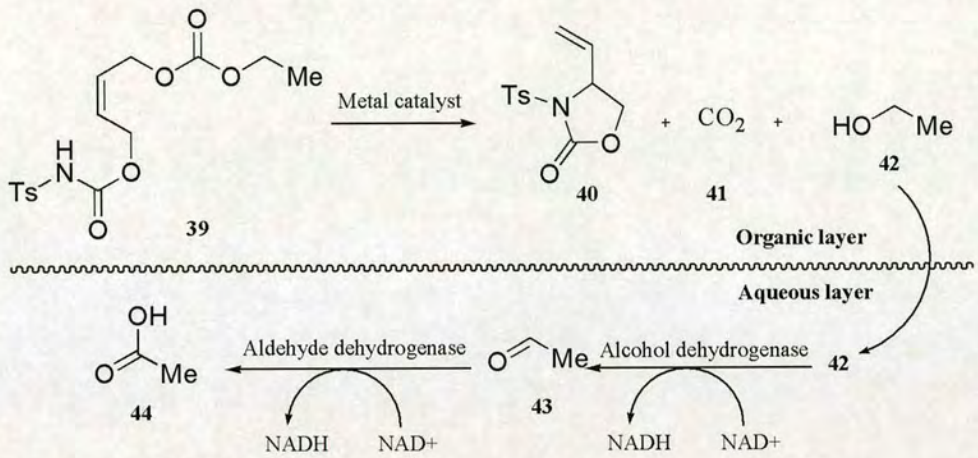


Figure 1.18: Using NAD^+ dependent hydrogenases to detect intramolecular allylic aminations

Another approach to overcoming solubility problems was reported in 2005 by Babiak and Reymond.[56] In this assay, fluorogenic substrates were adsorbed onto silica gel plates in order to detect lipase, esterase, glycosidase and protease activities. The large contact area between the enzyme and the substrate aimed to overcome the poor water solubility of the chromogenic/fluorogenic substrates of the enzymes. Silica gel plates and reverse phase C18 thin layer chromatography (RP-C18 TLC) plates were investigated and it was shown that twenty different substrates containing esters and acyloxymethyl ethers of umbelliferone could be distributed on one plate. The substrates were transferred onto the plates dissolved in dichloromethane and then the plate was dried. A 1 µl solution of a lipase was spotted onto the plate and successful reaction resulted in a blue fluorescent mark. The group showed that 144 reactions could be observed on a 5x5 cm plate, offering a practical alternative to screening on microtitre plates where the required enzyme volumes are much larger (5-200 µl). Glycosidase activity was also observed using water soluble substrates in the same manner.

Glass slides have also been used to attach substrates to analyse enzyme activities. In a study by Gronux and Reymond in 2006, lipases were subjected to substrate profiling using many different substrates covalently attached to a glass slide.[57] The substrates were designed so that upon lipase cleavage, diol **46** remained. Chemical manipulation of the substrates after lipase hydrolysis on the glass slides allowed red fluorescent detection of the diols using rhodamine sulfohydrazide **48**, Figure 1.19. On a glass plate 12x3 cm, 324 spots of lipase were placed analysing twelve pure substrates and numerous mixtures of the substrates.

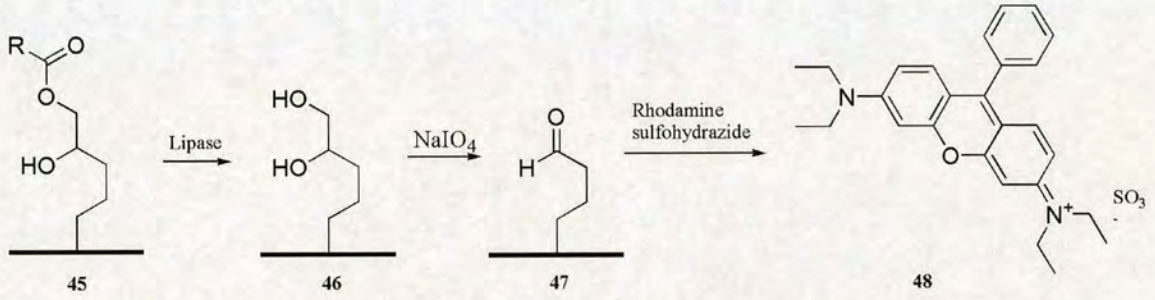
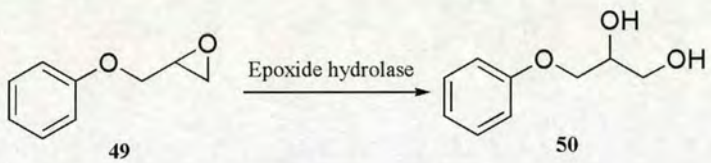


Figure 1.19: Lipase detection using red fluorescence of substrates on glass slides

The technique provides a practical alternative to screening in microtitre plates. However the screen still requires optimisation as pure enzyme preparations are needed at present. Only lipase cleavages that produced 1,2-diols were detected and the screen is not enantioselective.

A miniaturized screen recently published by the Reetz group has potential to function in high-throughput with lower reagent consumption. Epoxide hydrolase variants were used as the test enzymes and enantioselectivities were analysed using mini-channels on a microchip electrophoresis device.[58] Four channels were present on each chip enabling it to be used for one enzyme and three substrates, or three enzymes and one substrate. The enzyme and substrate were pulled and mixed through the chip channels using either a vacuum or voltage. A chiral selector compound pumped through the chip enabled UV-laser analysis to distinguish between enantiomers. These researchers showed that

reactions were possible using purified enzymes, cell lysates and whole cells. Epoxide hydrolase variants generated by directed evolution were investigated for the conversion of epoxide **49** to diol **50** and E value increases from 4 (wild type) to 101 were measured, Figure 1.20. The data obtained using the chip was comparable with yields and e.e.s determined using HPLC.



49 50

Figure 1.20: Epoxide Hydrolase ring opening reaction used to investigate the use of chips for screening

Another screening method that has been developed by Griffiths and Tawfik is the use of emulsions to compartmentalize genes.[59] Water-in-oil emulsions were designed to contain all the necessary components for gene expression. Therefore proteins could be expressed and contained in the same vesicle as the coding gene, achieving *in vitro* compartmentalization (IVC). The size of the emulsions was designed so that each vesicle of water statistically contained one gene. Any additives needed in the water compartments were introduced by addition to the oil. The IVC approach has been used to examine variants of methyltransferases, RNA ligases, β-galactosidases, phosphotriesterases and Diels-Alder ribozymes. A number of screening and gene selection systems have been designed based on the properties of the proteins of interest. A fluorescence quenched substrate was added to the IVC and expression of an enzyme that could cleave the substrate and produce fluorescence was detected using flow cytometry. The fluorescent compartments could then be broken open to isolate the active enzyme-coding genes. A second method to examine the expressed proteins was by using gene-linked-substrates. The desired conversion of the substrate altered the properties of the gene which allowed subsequent isolation. This approach was used to screen for methyltransferases, where any active enzymes would methylate the gene. By breaking all the compartments and combining the genes, the non-methylated DNA could be

removed by digestion, leaving the methylated genes that coded for active enzyme. Subsequent amplification using the polymerase chain reaction created many copies of the genes of interest for sequencing or further expression.

The IVC expression and screening method provides an alternative to screening enzymes in a cell without the complications of other proteins and cell contents. The method was shown to be very high-throughput as the volumes of the vesicles were in the femto-litre range.

1.5 Screening for novel activities- Developing a new method

Many of the screens discussed in section 1.4 were designed specifically for detecting one reaction, or the conversion of a specific substrate, with the development of each being significantly time consuming. Therefore there would be great advantages to a screen that was readily adaptable to a range of reactions. In 2006, Sutherland *et al* decided to approach screening in a different manner and design one simple screen for many reactions that could be built upon like a 'biosynthetic tree'.[60] Their idea is shown in Figure 1.21. If a screen could be developed for the conversion of **Y** to **Z**, then that screen could be used for the detection of the conversions of **V**, **W**, or **X** to **Y**.

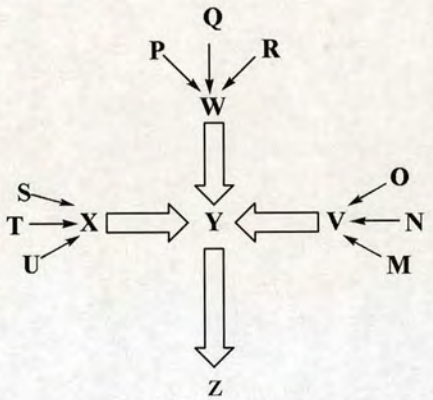


Figure 1.21: The biosynthetic tree approach for enzyme screening

Once a method of detection for **X**, **W** and **V** was established, a screen for all conversions of **M-U** to **Y** could be developed using the same reaction cascade.

In 2003, the idea for the current project was to design a screen based on a similar idea. The monoamine oxidase (MAO) work carried out by the Turner group was based on a colorimetric enzyme cascade system,[10, 27] shown in Figure 1.22. The screen ultimately detects the formation of hydrogen peroxide, an oxidation by-product.

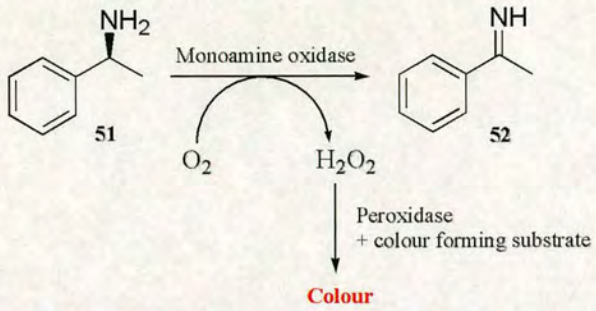


Figure 1.22: Detecting peroxide formation in the oxidation of amines

Amine oxidases are classified into two groups according to which cofactor they possess and are widely distributed in Nature. Type I are copper dependent and require 2,4,5-trihydroxyphenylalanine quinine (TPQ) as a cofactor, whereas type II are flavin dependent. The cofactor required by the porcine kidney DAAOx and the MAO in the Turner study was flavin adenine nucleotide **53**, FAD, Figure 1.23. There are two principal steps in the catalytic cycle of MAO catalysed oxidation of an amine substrate with the FAD cofactor. The first of these steps is the oxidation of the substrate with concurrent reduction of FAD to $FADH_2$, **53** to **54**. The protons are removed from the substrate. Second, is the regeneration of $FADH_2$ to FAD by molecular oxygen, releasing hydrogen peroxide and completing the redox reaction, **54** to **53**.

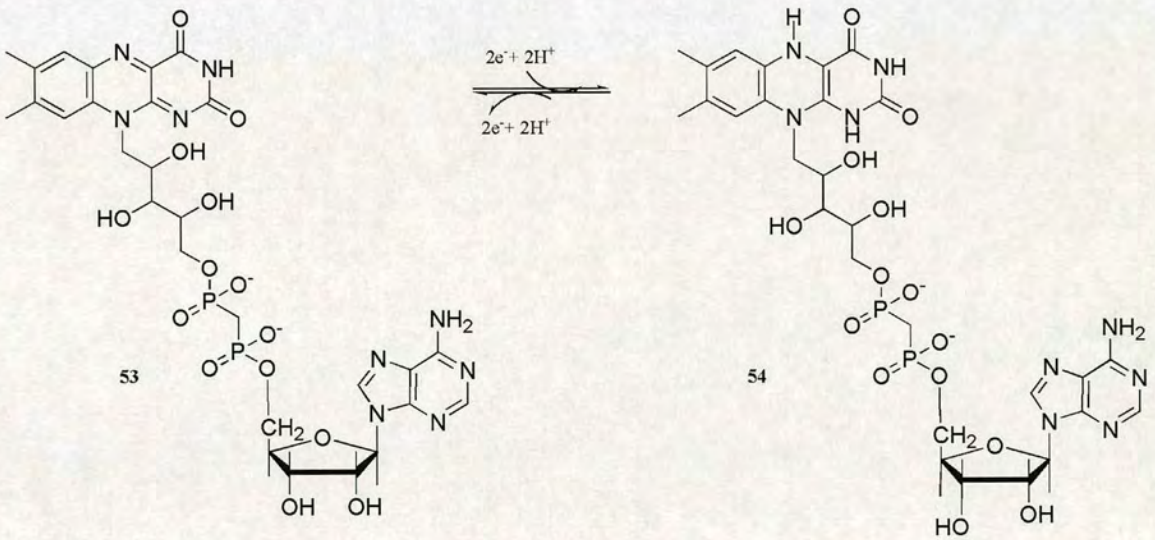


Figure 1.23: The reduction of FAD (53) to $FADH_2$ (54)

In the assay, peroxidase is introduced and uses the hydrogen peroxide to oxidise another substrate, for example, 3,3'-diaminobenzidine **55**, Figure 1.24, which polymerises upon oxidation to form an insoluble red/brown compound allowing easy identification of active colonies.[10] In the solution phase, 4-aminoantipyrine and 2,4,6-tribromohydroxybenzoic acid can be used to form a pink soluble dye product that is detected using a microtitre plate reader at 510 nm.

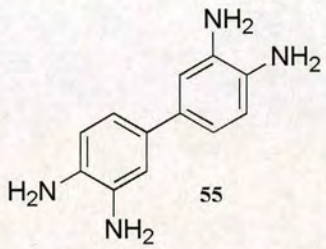


Figure 1.24: Diaminobenzidine

The peroxidase catalysed oxidation forming a coloured compound using hydrogen peroxide can be seen as the conversion of **Y** to **Z**, Figure 1.21. The oxidation of amines to produce hydrogen peroxide by MAOs can be compared to the conversion of **X** to **Y**. There are also amino acid oxidases and alcohol oxidases that are FAD-dependent, such as the amino acid oxidase from *Trigonopsis variabilis*, or the alcohol oxidase from *Pichia pinus*. The oxidation of amino acids by the amino acid oxidases to generate hydrogen peroxide can be seen as the conversion of **W** to **Y**. Therefore, any reaction that makes an amine could be detected using the amine oxidase and the peroxidase **X, Y, Z**; any reaction that produces an amino acid can be detected using the amino acid oxidase and the peroxidase **W, Y, Z**; and any reaction that produces an alcohol can be detected using an alcohol oxidase and the peroxidase **V, Y, Z**. The new coupled assay tree is presented in Figure 1.25.

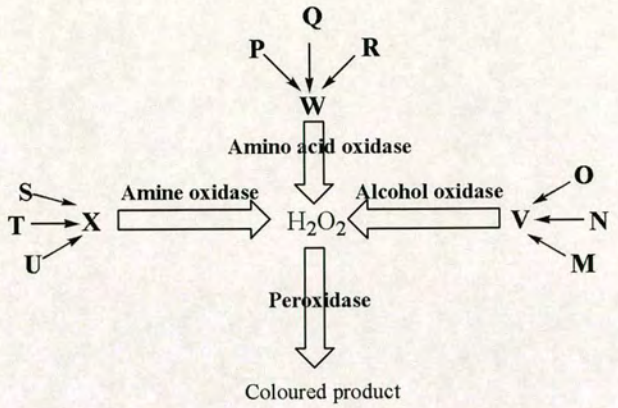


Figure 1.25: The coupled assay model based on the detection of peroxide

There are many examples of enzymes that require FAD as a cofactor but amine and amino acid oxidases will be used as the examples in this work. Enzymes that produce amines include transaminases and imine reductases, and enzymes that produce amino acids include esterases and amino acid dehydrogenases. The tree can be extended to include some possible enzymatic syntheses of amines and amino acids, as shown in Figure.1.26.

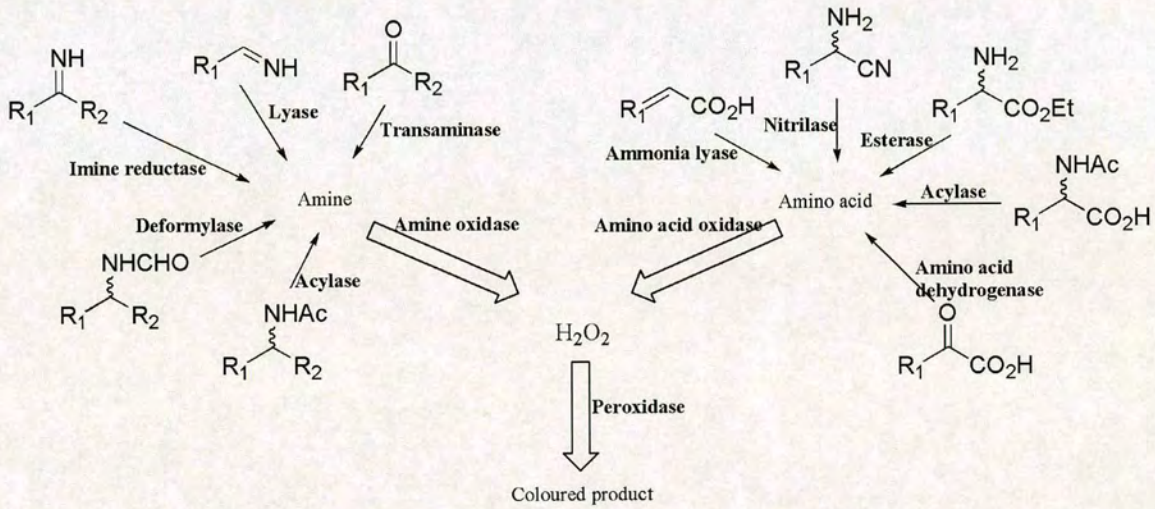


Figure 1.26: Possible enzymatic routes to amines and amino acids amenable to a peroxide detection screen

As presented in Figure 1.26, a wide range of enzymatic reactions can be screened for using the peroxide/peroxidase method. The MAO enzyme had been evolved to be highly enantioselective for the (*S*)-enantiomers of amines therefore the screen has the advantage of enantiodiscrimination. Amino acid oxidases also generally display high enantioselectivity, for example the oxidase from pig kidney, commercially available from Sigma, has been shown to preferentially oxidise D-enantiomers.[61] The novel screen not only has the capability of screening for amines or amino acids, but enantiospecific amine and amino acid products.

In this work, it was decided to use the amino acid oxidase and the peroxidase system to develop the screen and test its applicability. It was also decided to select a system that would test the ability of the screen to determine enantioselectivity. However, using oxidases and peroxidases to detect secondary enzyme activity has been reported previously: D-amino acid oxidase, peroxidase and 4-aminoantipyrine (4-AAP) were used successfully to screen for thermostable amino acid amidases that could produce D-Phe from D-phenylalaninamide.[62] Amino acid racemases have also been analysed using a D-amino acid oxidase, peroxidase and 4-AAP to detect the conversion of L-Ala to D-Ala.[63]

The novel aspect of the screen developed here was to introduce the gene for the amino acid oxidase into the cells producing the enzyme of interest using a second plasmid, followed by the co-expression of the enzyme of interest and the oxidase to successfully detect activity. By introducing the oxidase gene into the cell on a second plasmid, any enzyme that can produce a substrate for the oxidase can be screened (alcohols using the alcohol oxidase, amines using the amine oxidase).

Successful co-expression of the D-amino acid oxidase from *T. variabilis* with a second enzyme in one cell has been achieved for the bioconversion of cephalosporin C to 7-aminocephalosporanic acid. In the first experiment, a fusion protein was made with the oxidase and glutaryl-7-aminocephalosporanic acid acylase.[64] In the second, both the

genes were cloned into one plasmid.[65] It was reported that the oxidase and the acylase were active and could turnover substrate.

Co-expression from two plasmids for enzyme screening has also been employed for analysing glycosynthases (Figure 1.9). In 2001, Withers *et al* described the co-expression of a glycosynthase with an *endo*-cellulase, which reported the glycosynthase activity by release of fluorescence. Both genes were on separate plamids.[43] Later these workers placed both genes on one plasmid and were able to detect glycosynthase activity.[8]

1.6 Racemase Enzymes

Amino acid racemases were chosen as the class of enzyme of interest for development of a novel screen. Amino acid racemases catalyse the interconversion of the two enantiomers of an amino acid. The advantage of using amino acid racemases for the screen was that the enzymes do not require any additional substrates for amino acid conversion, keeping the screen simple for initial development work. The proton on the α-carbon is removed and reprotonation occurs using the amino acid side chains already present in the active site. The principle of the amino acid racemase screen was to introduce the L-amino acid and follow conversion by the oxidation of the D-amino acid product and subsequent production of a coloured compound by the peroxidase (Figure 1.27).

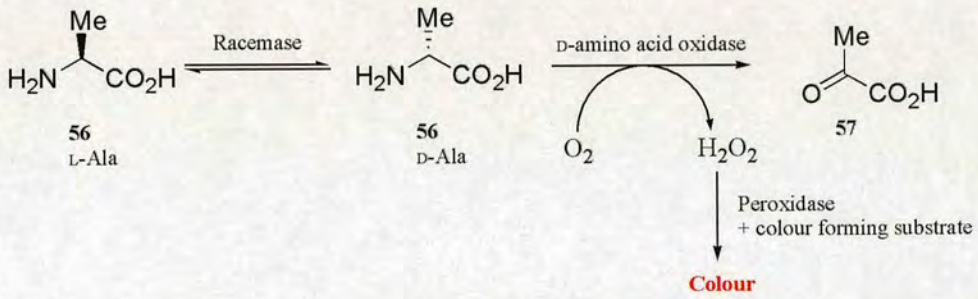


Figure 1.27: Principle of the amino acid racemase screen

L-amino acids are most commonly found in Nature and are mostly used in protein synthesis and as a carbon and nitrogen source. For example in *Bacillus subtilis*, L-Ala is converted by an L-alanine dehydrogenase into pyruvate and ammonia as a carbon and nitrogen source.[66, 67] D-amino acids are not as common but can be found in bacterial cell walls (D-Ala and D-Glu),[68] D-ser has been found to act as a neuromodulator in mammals[69] and D-Ala has been shown to play a role in marine animals in osmotic stress.[70] The less common D-enantiomers are generally produced using racemase enzymes. There are two main categories of amino acid racemases, those that are pyridoxal phosphate (PLP) dependent and those that are PLP-independent.[71] PLP is the

active form of pyridoxine, more commonly known as vitamin B_6 and is the coenzyme required by many enzymes that metabolise amino acids, (racemases, transaminases and decarboxylases). Pyridoxal phosphate is covalently bound to the side chain of a lysine residue in the enzymes as an imine. Imines are useful functional groups for catalysis as they are formed in a reversible process involving nucleophilic attack on the carbonyl by the primary amine in a weakly acidic environment. As formation is easily reversible they are valuable intermediates as the PLP is not bound strongly to the enzyme and can react with other primary amine groups.

As one aim of the project is to broaden the substrate specificity of the enzyme to include unnatural substrates, the amino acid racemase with the widest native substrate specificity would be the most appropriate starting point for mutagenesis. In the literature many amino acid racemases are documented[2] and racemases from one species in particular act on a wide range of substrates.[72] The broad substrate specificity enzymes are PLP-dependent amino acid racemases from *Pseudomonas putida.*[71, 73, 74] This enzyme was chosen as the test candidate for the novel screen, as a number of amino acid racemisations could be analysed and used to demonstrate the use of the oxidase reaction for detecting a number of substrates.

1.6.1 Understanding amino acid racemases: The mechanism

The majority of PLP-dependent enzymes function in the following way: the bond that is to be broken (for example, C-H in racemases, C-O in decarboxylases) is held parallel to the π bond system of the PLP. The nitrogen in the PLP is usually protonated creating an electron sink. In Figure 1.28, the π system is displayed with the amino acid (red) held with the C-H bond parallel to the π bonds.

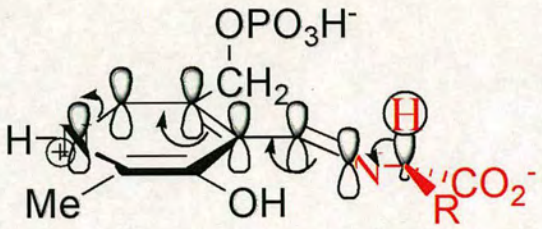


Figure 1.28: The PLP pi system with the amino acid C-H bond parallel

When the proton is abstracted, a negative charge is left on the amino acid carbon that can be delocalised through the π system and the resulting quinoid intermediate is relatively stable. Due to the stability of the resulting intermediate, the pKa of the proton on the α-carbon of the PLP-amino acid substrate is therefore reduced compared to the amino acid in solution. The activation energy to deprotonate is lower making it easier to racemise the asymmetric centre using a basic group. Proton abstraction occurs to give a new imine between the nitrogen and the carbon of the amino acid and an enamine in the heterocyclic ring. The mechanism of a PLP-dependent racemase is shown in Figure 1.29.[75-78] It has been shown that the rate determining step in the racemisation of alanine is the binding and release of L-Ala, not the deprotonation of the α-carbon.[79]

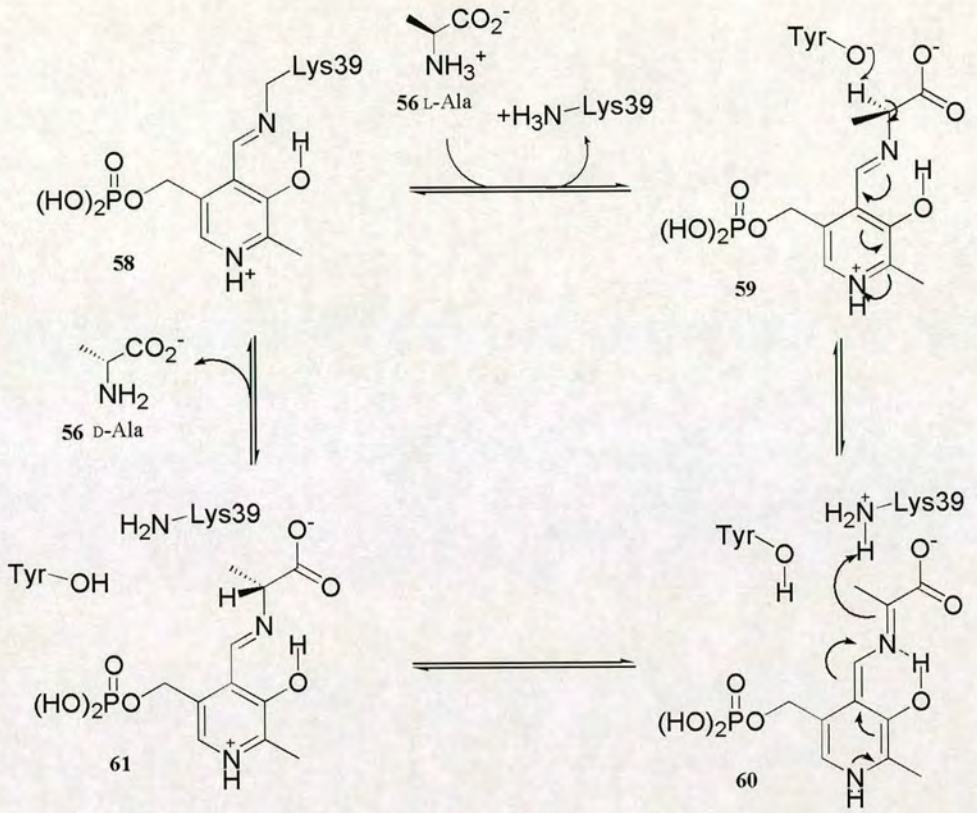


Figure 1.29: The racemase mechanism using the electron sink of the PLP

Structure **58** shows the PLP covalently bound to the Lys residue in the active site of the racemase. When an amino acid (L-Ala **56**) is introduced, a new imine bond is formed between the amino acid and the PLP **59.** The amino acid is held so that the C-H bond is parallel to the pi bond system of the ring. A tyrosine in the active site is in the correct position to abstract the proton and create the asymmetric centre on the amino acid 60. Re-aromatisation is the driving force to re-protonate the α-carbon. In transaminases the abstracted proton is transferred to the carbon α to the pyridine ring, transamination has been seen to occur with racemases at low pHs.[80, 81] Reprotonation can occur from either face of the planar structure. If reprotonation occurs using the phenolic –OH, then the resulting amino acid will be L-. If reprotonation occurs on the opposite face using Lys-NH_3^+ (as shown) the amino acid is in the D- conformation **61**. The Lys-NH_2 is reprotonated (mechanism unknown, proposed to occur using water molecules)[82] and the Tyr-OH deprotonated. The amino acid is released from the PLP complex (imine bond

formation is easily reversible) and the PLP covalently bonds to the Lys residue. All steps are reversible and racemisation can proceed in the opposite direction, that is from D-Ala to L-Ala.

The second mechanism that has been proposed does not rely on the electron sink of the PLP. In most PLP-dependent enzymes, there is an acidic residue in the active site close to the pyridinium nitrogen that allows protonation. However, in amino acid racemases there is an arginine residue in this position. The large positively charged guanidinium head group of the arginine does not allow protonation of the N and two research groups have carried out computer modeling studies which support another possible mechanism for racemisation.[83, 84] However, it has been calculated computationally that the acidity of the proton on the α-carbon is still significantly reduced by the PLP and solvation effects.[84] The alternate mechanism uses the acid group of the amino acid in the deprotonation of the α-carbon,[85] (Figure 1.30) and has been proposed by a group that, prior to publication in 2002, carried out many of the mechanistic studies of alanine racemases used to determine the first mechanism using the protonated PLP nitrogen. Once the amino acid has displaced the Lys side chain and formed an imine with the PLP, the acidic group of the amino acid abstracts a proton from the tyrosine –OH, which in turn abstracts the proton from the amino acid α carbon in a six-membered transition state **63**. The proton shuttle can move in the opposite direction to release the L-Ala, or the C-C bond of the amino acid can rotate and move the deprotonated α-carbon in a position to abstract a proton from the Lys side chain **64**. The transition state of the second proton abstraction is also six-membered **65**. The Lys side chain then displaces the amino acid from the PLP and D-Ala **56** is released. Again, each step is reversible and D-Ala can be racemised to L-Ala using the opposite route.

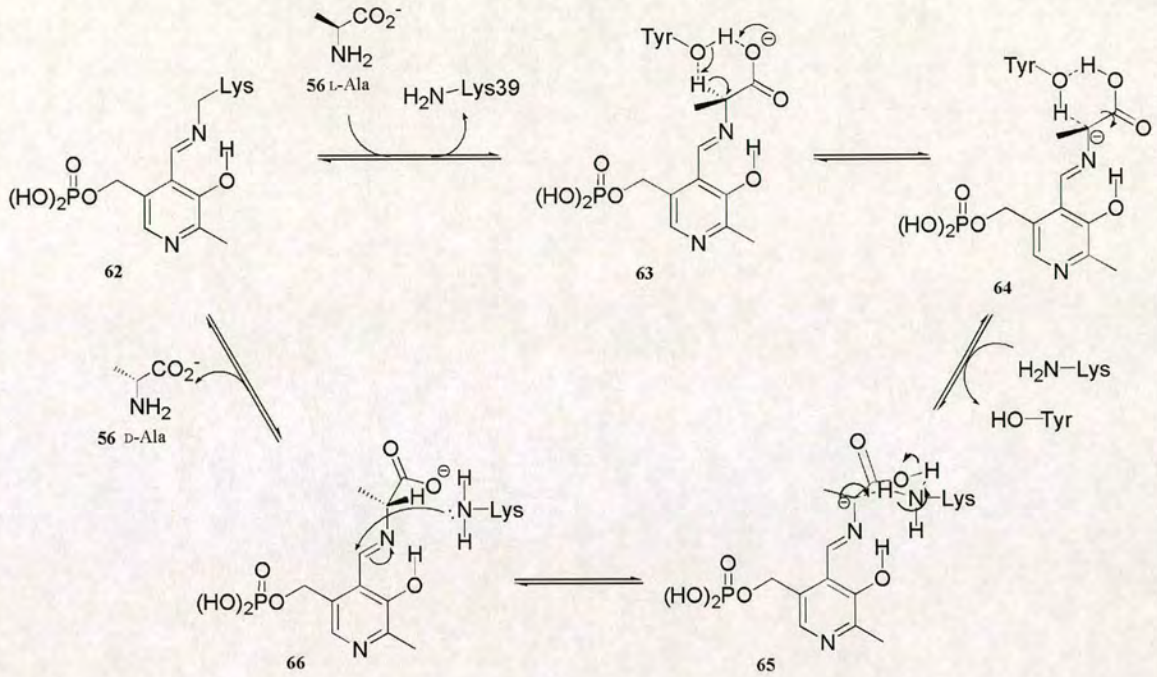


Figure 1.30: Alternate mechanism of racemisation

Various studies have proven that, regardless of the mechanism, the Tyr residue abstracts the proton from L-Ala,[76, 86, 87] whereas Lys abstracts the proton from D-Ala.[77, 88] Fenn *et al* carried out site directed mutagenesis on the alanine racemase of *Geobacillus stearothermophilus* to remove the Tyr residue and saw that the racemase was inactivated[76] and Watanabe *et al* carried out similar studies with the Lys residue.[77] Alanine racemases have been shown to exist as dimers containing two active sites, created on the boundary between the monomers. The important Lys residue belongs to one monomer and the catalytically important Tyr residue belongs to the other.[89]

1.6.2 Current amino acid racemase screens

The amino acid racemase from *Pseudomonas* sp. was chosen as the starting candidate to develop the high-throughput screen. A literature search revealed that all high-throughput solid phase assays for amino acid racemases described to date relied on a selection method. In 1993, a range of bacteria were investigated for threonine racemase activity by growing on minimal media containing D-Thr as the sole nitrogen source.[90] *P. putida*

ATCC 17642 was identified and shown to have threonine racemase activity. Asano *et al* also used the selection method when examining 200 microorganisms for D-Phe racemisation.[74] The microorganisms were grown using D-Phe as the only carbon and nitrogen source.

A number of coupled enzymatic screens to detect racemase activity have been used in microtitre plate format. Screening using NAD^+/NADH has been employed to detect conversions of D-Ala to L-Ala and L-Ala to D-Ala.[91, 92] An NADH dependent L-Ala dehydrogenase was used to detect L-Ala produced by the racemase and a D-amino acid transferase was coupled to an NADH dependent lactate dehydrogenase to detect D-Ala (Figure 1.31). In both studies, the conversion of NAD^+ to NADH, or NADH to NAD^+ was monitored at 340 nm and used to examine racemase activity.

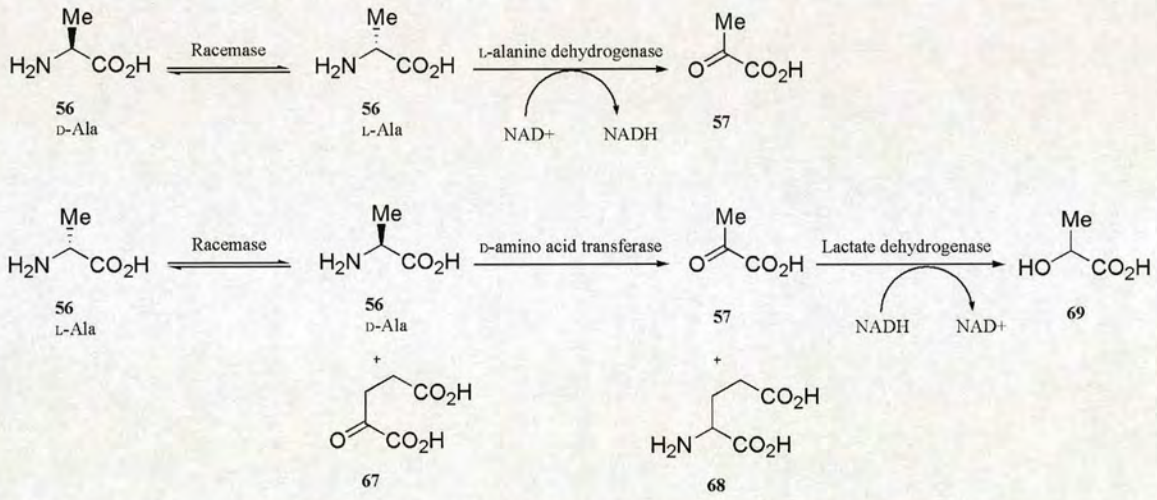


Figure 1.31: NAD^+/NADH enzyme coupled racemase screens

Racemisation of amino acids has also been followed by detection of peroxide using a peroxidase and a colour forming substrate. The racemisation of serine by an alanine racemase from *Geobacillus stearothermophilus* was measured using the following step-wise method:[93] purified racemase enzyme was added to L-Ser, the reaction was incubated for twenty minutes and boiled to deactivate the enzyme, a sample was then

removed and added to a solution of D-amino acid oxidase, peroxidase and o-phenylenediamine (OPD) and incubated. The developed colour (polymerised OPD) was measured at 492 nm and compared to a standard curve generated previously using known concentrations of D-Ser, Figure 1.32.

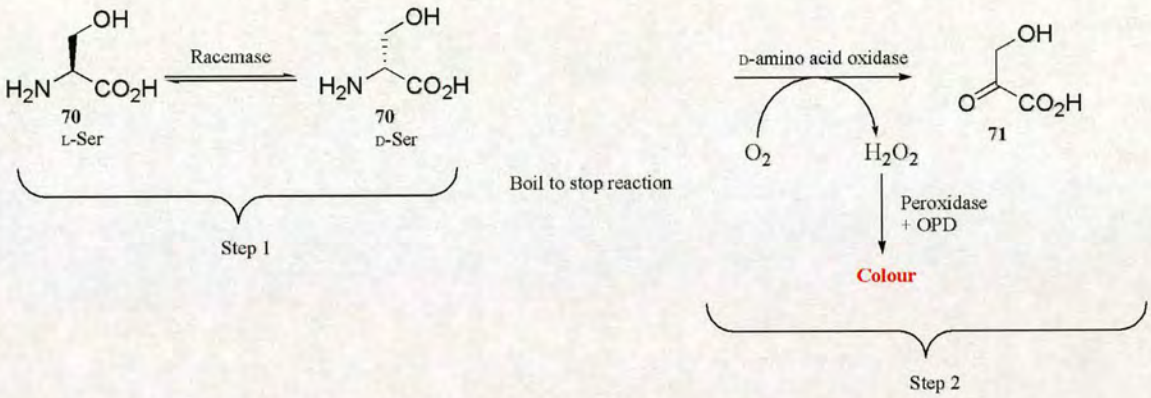


Figure 1.32: The stepwise detection of serine racemase activity

A similar method of racemase detection was carried out in 2005 using the cell free extract of variant racemases created by shuffling the alanine racemases from *E. coli* and *Salmonella typhimurium.*[31] The racemase activity was monitored by carrying out reactions with L-Ser or L-Ala for 30 or 10 minutes respectively, deactivating the racemase using HCl and adding the remaining mixture of amino acids to a cocktail of D-amino acid oxidase, peroxidase, 4-AAP and *N*-ethyl-*N*-(2-hydroxy-3-sulfopropyl)-m-toluidine. The exact same method was used in 2006 by Ono *et al* to examine alanine racemase activity.[63]

Non-enzymatic routes to detect amino acid racemase activities included the analysis of amino acid enantiomers by chiral HPLC,[3] the development of a HPLC/mass spectrometry high-throughput method (56 min per 96 well plate) by DSM,[94] using deuterated water followed by monitoring the NMR signal of the α-proton of the amino acid and methods based on optical rotations.[90] A medium-throughput assay was developed by Schonfeld and Bornscheuer in 2004 for partially purified glutamate

racemases in microtitre plate format.[95] Cultures grown in deep well plates were subjected to a His-tag 96 well plate purification and transferred to a liquid handler for addition to L-Glu. The decrease in optical rotation of the solution was followed and related to specific activity. It was estimated that three 96-well plates could be examined per day.

Circular dichroism (CD) has also been employed recently to measure amino acid racemase activity.[96] The CD spectrums of 1 mM L-Ala, 1 mM D-Ala and a range of mixtures of the enantiomers were measured between 205 and 215 nm. The spectrums were integrated and plotted against the concentration of L-Ala to generate a standard curve. Samples from a D- to L-Ala racemisation reaction were then analysed by CD and the integrated spectra compared to the standard to determine L-Ala concentrations.

1.7 Summary and Aims

Directed evolution has been developed as a powerful tool to alter the characteristics of proteins and make them useful catalysts in organic chemistry. Despite the many developments of techniques for directed evolution, general high-throughput screens are not as numerous. Frequently, screens are developed for a specific reaction and this step can be very time consuming.

The goal in this project was to establish a screening method that could be applied in many directed evolution experiments that also had the advantage of being enantioselective. Amino acid racemases were good candidates for the development of the novel two-plasmid screen as current solid phase amino acid racemase assays rely on selection methods only. The solution phase screening using the oxidase/peroxidase system has been carried out in a step-wise fashion and using a continuous method.

As candidates for biotransformations, amino acid racemases have the added advantage that the cofactor (PLP) is regenerated after each catalytic cycle (although the mechanism has not been confirmed) and is covalently attached to the enzyme when not involved in catalysis (therefore not free in solution and separated from the enzyme). Amino acid racemases do not require any additional substrates for enzymatic action and so with the introduction of an amino acid, the process of racemisation should start immediately.

Once a test candidate enzyme was chosen and the gene cloned, the first aim was to establish a method for the screen and show that it could be used to detect wild type racemase activity. Then, to show that the assay could be used in conjunction with directed evolution experiments to detect novel activities from a background of inactivity, the cloned racemase genes were subjected to mutation using a variety of methods and screened.

2. Results and Discussion

2.1 Identification, isolation and expression of amino acid racemases

The ultimate aim of this project was to develop and demonstrate the application of a new screening method, for which amino acid racemase enzymes were chosen as the test candidates. This chapter presents the identification of possible racemase candidates, the subsequent gene cloning and enzyme expression experiments. The first racemase chosen was the amino acid racemase from *P. putida*, as the broad substrate specificity is well documented.[1] In the first section of the chapter, the cloning and sequencing results of the *P. putida* racemase from genomic DNA is presented.

In the second section, amino acid racemases with unknown or narrow substrate specificities were selected for use in directed evolution studies to prove the application of the two-plasmid screen to identify novel specificities. Initial selection of the enzymes was carried out using a protein sequence probing method, with the *P. putida* sequence, to identify similar proteins. The bioinformatics search algorithm tool, Basic Local Alignment Search Tool (BLAST) was used to identify protein sequences similar to the *P. putida* racemase. BLAST works on the principle that the sequences consist of segments of aligned information separated by areas of no alignment. The BLAST algorithm allows the user to modify parameters to refine the search, in particular the scoring matrix BLOSUM (BLOcks SUbstitution Matrix), which can be changed to evaluate proteins with various levels of homology. BLOSUM 62 is usually set as a default, but to widen the search to include more distantly related proteins the matrix can be set to 45, or if looking at relatively short evolutionary distances, BLOSUM 90 is preferable. The sequences returned by the BLAST program are given an E value, which is a statistical score giving an indication of how closely related the query and result sequence are.

Scores below 0.01 indicate matches, with the smaller the number, the more similar the sequences. It is unlikely that matches occur by chance.

Following identification of the candidate racemases, the genes were isolated and inserted into expression vectors. Restriction digests were carried out to confirm the plasmid constructions. The genes were sequenced, compared to the reported sequences and translated to analyse the protein sequences. Finally, the expression of the racemases was tested in a range of *E. coli* host cells.

2.1.1 Cloning and characterisation of the broad substrate specificity racemase from *P. putida*: PCR of the racemase gene from gDNA and subsequent cloning into pET16b

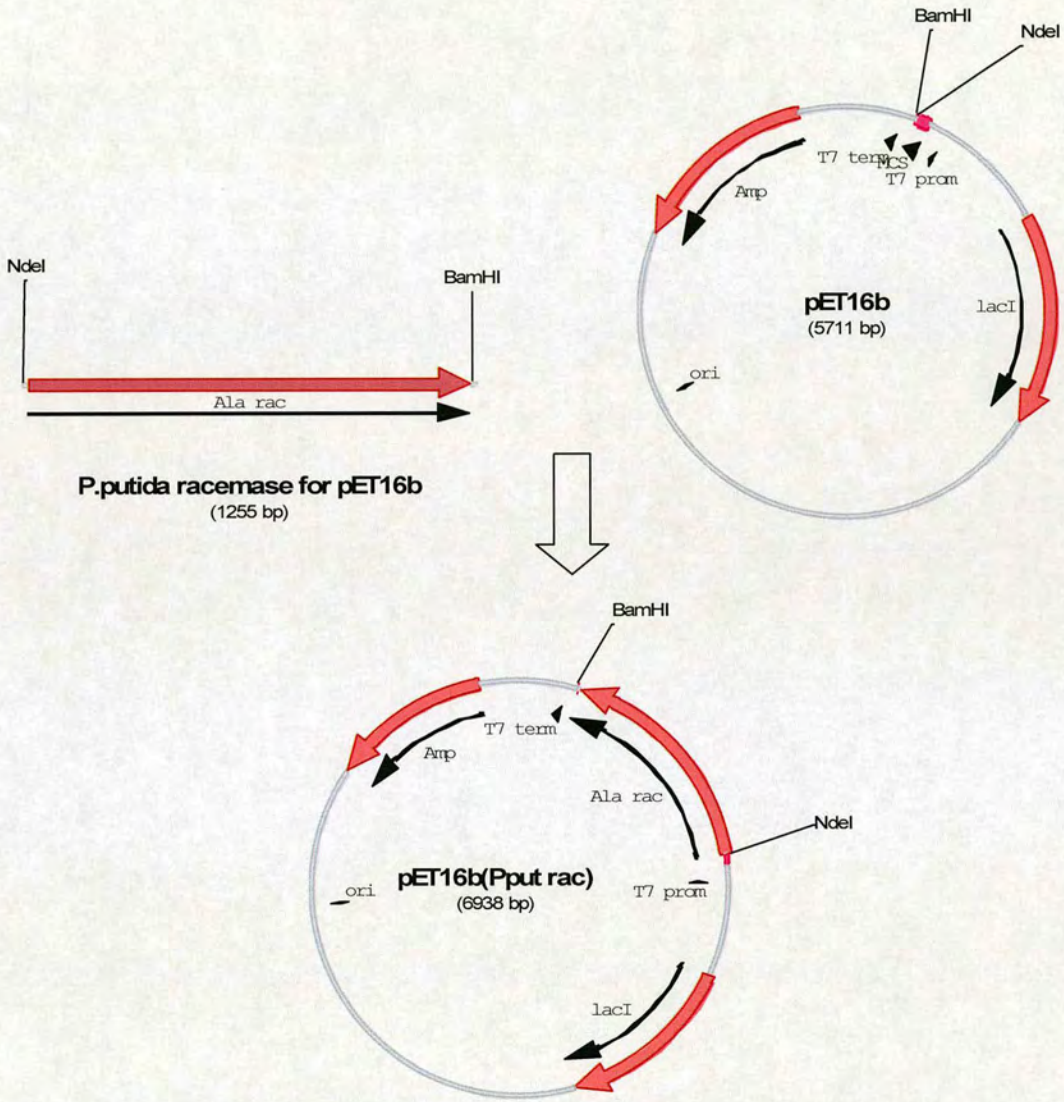


Figure 2.1.1: Cloning plan to insert the racemase gene into pET16b

The gDNA from *P. putida* was prepared as described in section 4.2.3 and the putative gene was obtained by PCR. Primers were designed to introduce an *Nde*I cleavage site at the 5' terminus of the gene and a *Bam*HI site at the 3' terminus. Gel electrophoresis was used to analyse and isolate the PCR product and a band of approximately 1.2 kbp was

excised from the gel. The PCR product and pET16b plasmid DNA were subjected to a double restriction digest using *Nde*I and *Bam*HI and both were purified by gel electrophoresis. The gene insert was ligated with the linearised pET16b and the resultant plasmid used to transform into *E. coli* BL21 (DE3). Only low numbers of small colonies were found on the plates after an overnight growth and so the incubation period was extended for 24 h. The number of colonies increased slightly after the extended incubation period. Colonies were selected and cultures were grown for 36-48 h (as 12-24 h rarely resulted in visible cell growth) and subjected to plasmid DNA extraction. However, no plasmid DNA could be isolated.

A number of ligation/transformation experiments were performed to ensure that there was not a contamination and fresh LB/Agar/Amp plates were prepared to ensure that there was no problem with the antibiotic. Miniprep kits from different suppliers (Roche, Sigma and Qiagen) were used to eliminate suspicions of any contaminants. The same results occurred each time.

It was a possibility that the enzyme might have a toxic effect on the cell as it could be depleting L-amino acids that are essential for cell production and growth, explaining the slow growth seen on the plates and in the cultures. When the enzyme becomes too toxic for the cell, the cell is unable to grow and replicate properly. *E. coli* without this plasmid is able to grow rapidly and dominate the culture when the majority of the ampicillin antibiotic is degraded by a secreted β-lactamase, which occurs by the time the culture is becoming slightly turbid (OD_{600} of approximately 0.2).[97] It had been observed previously that pET16b can produce low level enzyme expression, even in the absence of expression inducers.[98] An alternative plasmid, pTTQ18,[99] was then used here as protein expression is strongly regulated by the binding of a variant of the *lac* repressor to the *tac* operator, (pET16b contains the T7 promoter). The *tac* promoter is a hybrid of the *lac* and the *trp* promoters and is a strong promoter as it matches the consensus sequence

for the *E. coli* RNA polymerase. The repressor is inhibited by isopropyl-β-D-thiogalactopyranoside (IPTG) and this can be used to switch on the protein expression.

2.1.2 Cloning of *P. putida* racemase gene into pTTQ18

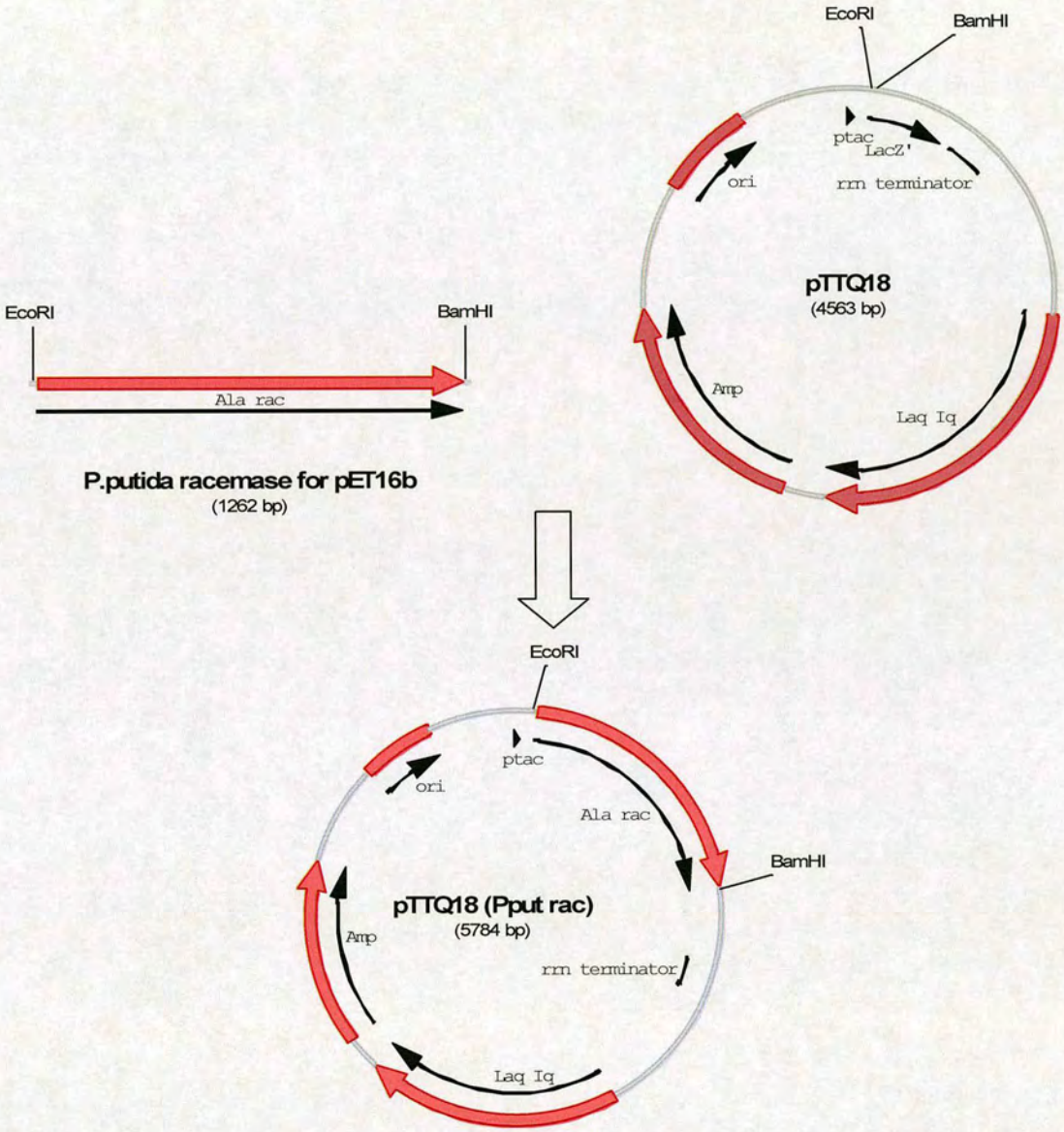


Figure 2.1.2: Cloning plan to insert the racemase gene into pTTQ18

Oligonucleotides were designed to amplify the gene from *P. putida* gDNA by PCR with *Eco*RI and *Bam*HI cleavage sites on the 5' and 3' gene termini respectively. The PCR

product was purified by gel electrophoresis and subjected to a double digest. When the digest was visualised on an agarose gel, smaller than expected bands were seen, (Figure 2.1.3).

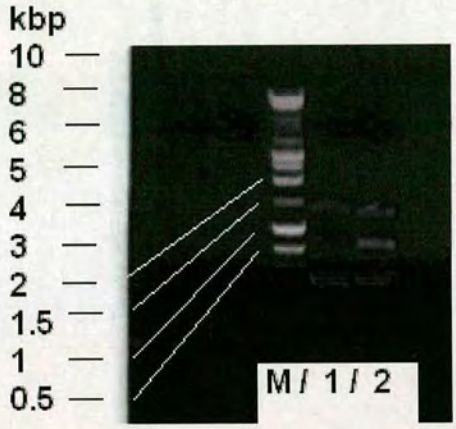


Figure 2.1.3: An agarose gel showing the racemase containing plasmid after digest with EcoRI M= Ladder, Lane 1= Digested PCR product, Lane 2=Repeat of digest

The size of the predicted racemase gene is 1.2 kbp and the largest band in lanes 2 and 3 is approximately this size. However, there are two bands at 900 bp and 300 bp that were not seen on the gel when the PCR product was purified before digestion. The *Eco*RI was presumed to have cleaved the gene unexpectedly as this effect was not seen when using *Bam*HI and this gene for cloning into pET16b. The PCR and digestion were repeated to demonstrate that the initial PCR had not introduced a mutation and an extra *Eco*RI cleavage site in the gene. For the digestion, *Eco*RI from a different source was tested with the same result and this ruled out contamination. Different buffers were also tested to ensure *Eco*RI was not exhibiting *Eco*RI star activity. The logical conclusion was that the gene must have an unreported *Eco*RI site in the gene and alternative restriction sites for cloning would have to be used.

2.1.3 Cloning of racemase gene into pINGE2 to make pJLM101

The vector pINGE2 (Ingenza Ltd) is a derivative of pTTQ18 where a second multiple cloning site has been inserted and this was chosen in a third strategy to clone the racemase gene.

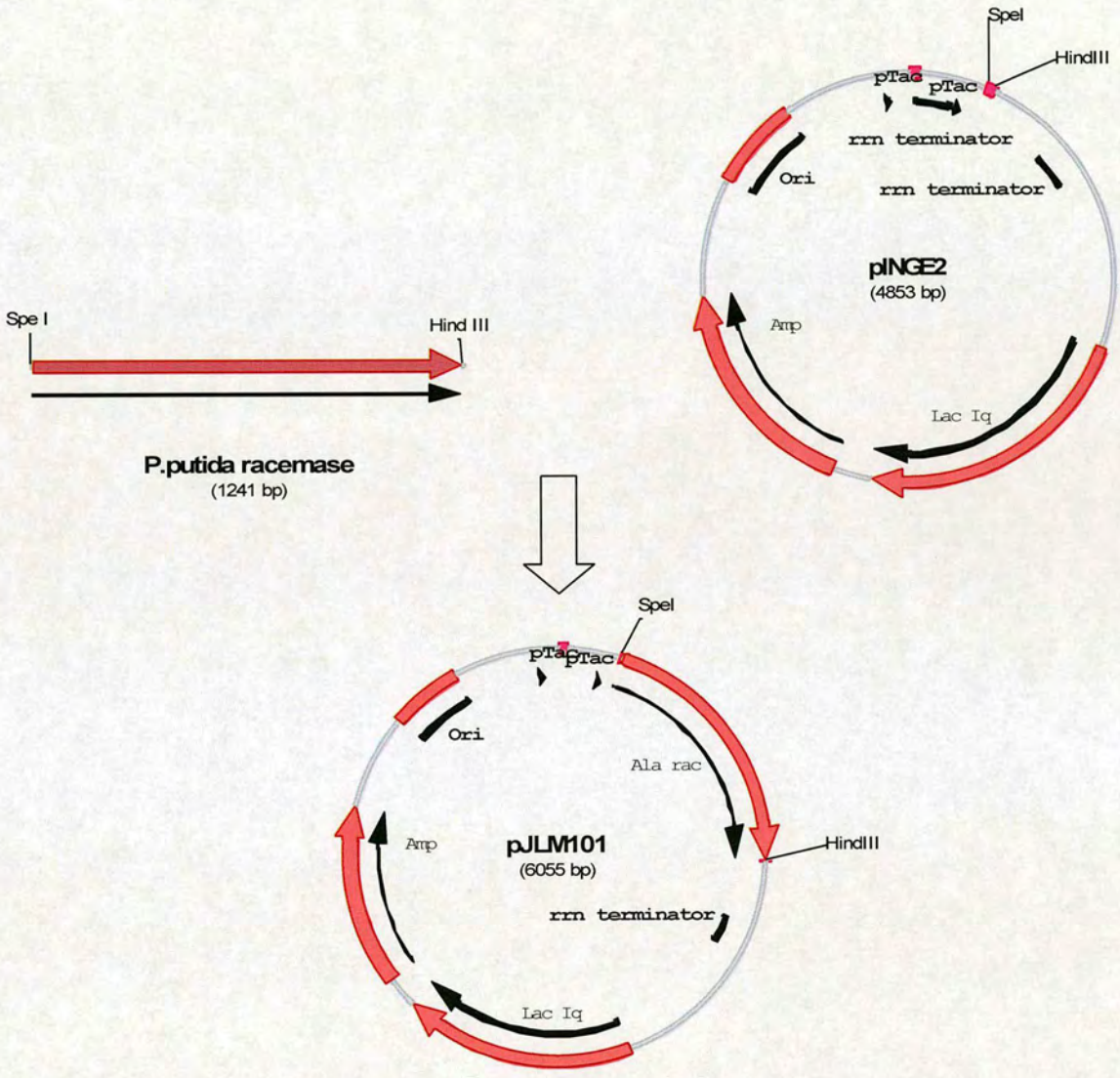


Figure 2.1.4: Cloning plan to insert the racemase gene into pINGE2

The gene was amplified from the *P. putida* gDNA with a *Spe*I cleavage site introduced at the 5' terminus and a *Hind*III site at the 3' terminus. A double digest of pINGE2 and the PCR product was carried out and the purified fragments were ligated and used to

transform *E. coli* Top10 cells. The plasmid DNA was isolated and subjected to a diagnostic digest to confirm insertion and placement of the gene in pTTQ18, (Figure 2.1.5a and b).

a)

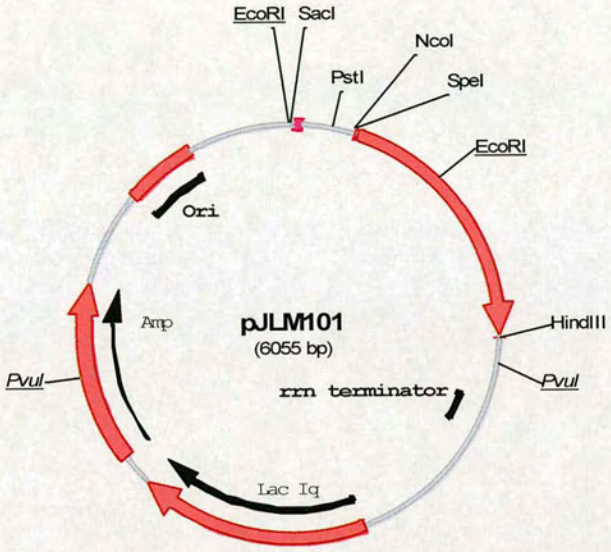
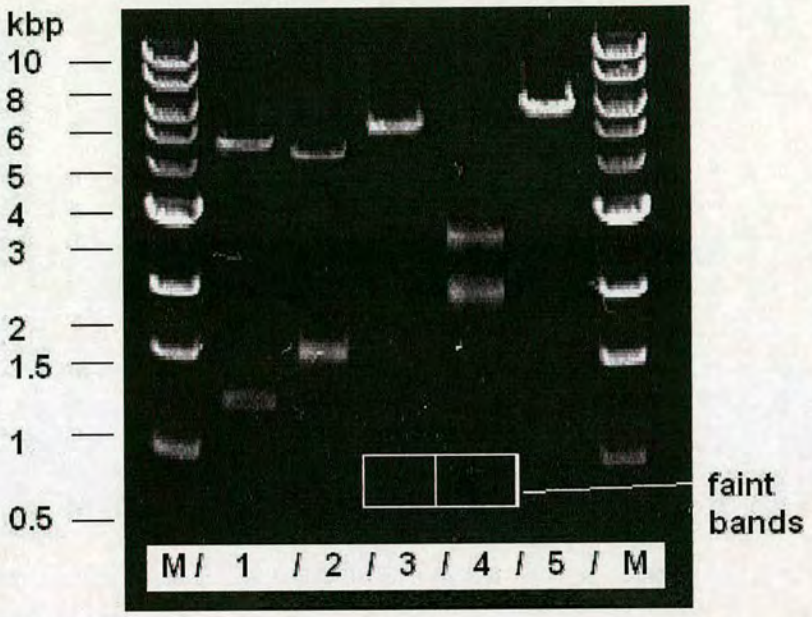


Figure 2.1.5a:A plasmid map of pJLM101 displaying restriction sites used in a diagnostic digest

b)



b: An agarose gel analysis of diagnostic restriction digest. The expected bands are as follows: Lane M = 1 kb Marker (Biorad); Lane 1 = double digest with SpeI/HindIII showing the expected bands at 1.2 and 4.8 kbp; Lane 2 = double digest with SacI/HindIII, showing the expected bands at 1.5 and 4.5 kbp; Lane 3 = digest with EcoRI only, showing the expected bands at 0.9 and 5.1 kbp; Lane 4 = double digest with PvuI/NcoI, showing the expected bands at 0.9, 2.4 and 2.7 kbp; Lane 5 = Digest with SpeI only, showing the expected band at 6.0 kbp corresponding to linearised vector

The plasmid construct was digested with *Spe*I/*Hind*III to excise the racemase gene and confirm the size. Digestion with *Sac*I/*Hind*III confirmed the presence of the gene in the second MCS of the vector. *Eco*RI restriction confirmed the presence of two *Eco*RI sites. Digestion with *Pvu*I/*Nco*I was used to confirm presence of the amp resistance gene. *Spe*I was used to linearise the vector and electrophoresis confirmed the total predicted size.

2.1.4 Sequencing results of racemase gene in pJLM101

The pJLM101 construct was analysed by sequencing PCR. The sequences were aligned with the reported gene sequence, (Appendix 1). 114 base changes between the reported sequence and the cloned gene were found. Excluding the first three bases added as a result of the cloning strategy, there are a total of 20 amino acid differences. The majority

of the base differences are in the third base of the codon and so are generally redundant. The amino acid variations are as follows: N35T; T36A; V48I; I83V; L129V; G132A; A150L; A155E; T160P; M165L; A180T; I191V; A273S; R274H; K278Q; H286R; S311A; V327A; E386D; and N400S. All numbers refer to the *P. putida* literature sequence. It is unlikely that these changes were introduced by PCR given the fidelity of *Pfu* hotstart polymerase. The likelihood is that the literature sequence was from a different strain of *P. putida* than that used for the PCR.

One of the single base differences introduced a sequence that is recognised by the restriction enzyme *Eco*RI. The expected GAATTT was found by sequencing to be GAATTC (positions 433 to 438, with reference to the *P. putida* literature sequence), explaining the observations when trying to digest the gene with this enzyme (Section 2.1.2).

2.1.5 Identification and isolation of other racemase genes:

Using BLAST to identify other candidate racemases

A variety of racemases were cloned in parallel. Using the analysis tools available on the TIGR website,[100] a BLAST search was carried out using the literature *P. putida* protein as the query sequence, focusing on databases of sources of genomic DNA (gDNA) that were readily available. Three other racemases were selected as possible candidates, using the scoring matrix BLOSUM 62. For a summary of the BLAST search results see Table 2.1. A sequence alignment of the protein sequences is shown in Appendix 2.

Gene	**Organism**	**Similarity *(%)***	**Identity*(%)***	**Gene Length**	**E-value**
	P. putida	Query sequence		1227 bp	-
putative *alr*	*S. coelicolor*	189/384 *(49)*	113/384 *(29)*	1152 bp	2.7×10^{-32}
alr	*B. subtilis*	171/347 *(49)*	98/347 *(28)*	1041 bp	4.5×10^{-34}
dal	*B. subtilis*	177/369 *(48)*	110/369 *(30)*	1107 bp	2.5×10^{-31}

Table 2.1: Results of the BLAST search

The racemase from *P. putida* is well documented as a broad substrate specificity racemase. The *dal* racemase from *B. subtilis* was one of the first racemases to be isolated and has been well studied.[101] It has narrow substrate specificity, with alanine as the only reported substrate. *Dal*$^-$ mutants grow in minimal media, but not in the presence of L-Ala, suggesting that the *dal*$^-$ mutant contains a second racemase, expression of which is repressed by L-Ala.[66] The *B. subtilis alr* has only been proposed to express a racemase enzyme and this has never been isolated or studied and the role maybe different to that in other bacteria containing 2 racemases. In bacteria such as *E. coli*, there is an alanine racemase constitutively expressed at a low level to provide D-Ala for the bacterial cell wall, from the *dadX* gene[102] (this is similar to the *dal* gene from *B. subtilis)*. In *E. coli* there is a second racemase that is induced in the presence of L-Ala, to convert excess L-Ala to D-Ala. The D-Ala is then converted by a D-specific alanine dehydrogenase to pyruvate and ammonia as a carbon and nitrogen source and also to provide energy. *B. subtilis* contains an L-specific alanine dehydrogenase, removing the need for the second *alr* racemase.

The racemase from *S. coelicolor* is an uncharacterised racemase which is assigned as a 'putative' alanine racemase in the genome databases (www.tigr.org) and there is no literature describing the study of this enzyme. This is also the only alanine racemase found in the *Streptomyces coelicolor* genome using a name search and the only gene that is returned by the BLAST matrix when using the *P. putida* racemase sequence as a probe (using www.tigr.org and www.ncbi.nih.nlm.gov) to search the genome.

In summary, a broad substrate specificity racemase (from *P. putida*), a well studied racemase from *B. subtilis* (*dal*) and two racemases were identified for future study.

2.1.6 Isolation and cloning of the racemases from *B. subtilis* and *S. coelicolor*

2.1.6.1 The construction of pJLM201: the alanine racemase from *S. coelicolor*

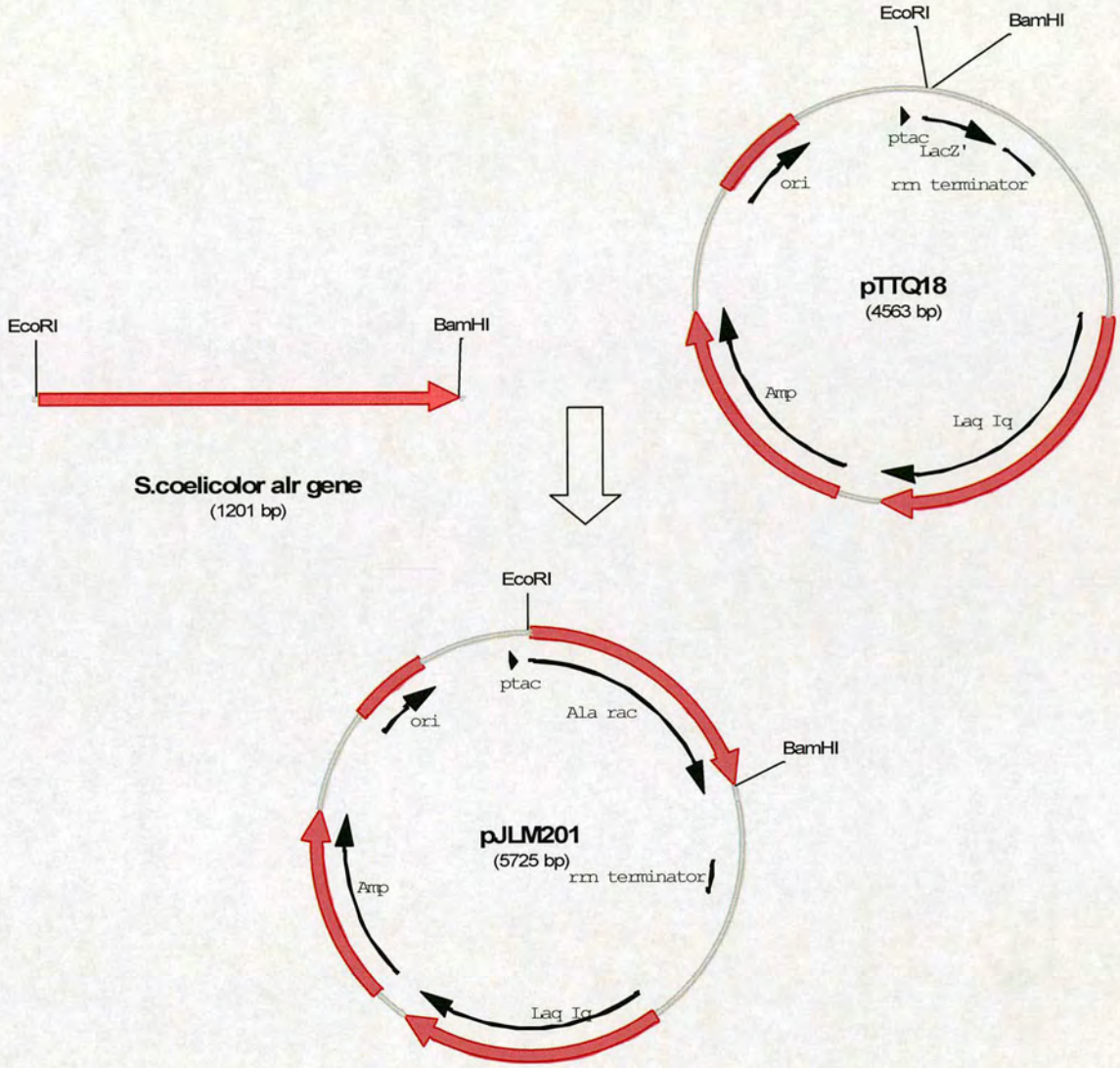


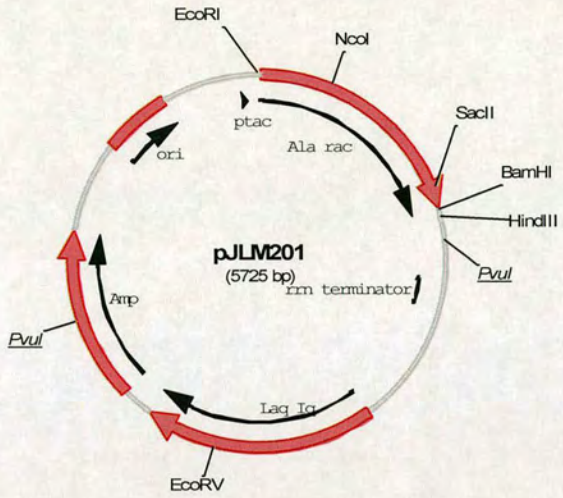
Figure 2.1.6: The cloning plan to create pJLM201 using pTTQ18 and the racemase gene

The plasmid pTTQ18 was chosen as the host vector as the *P. putida* racemase was in pINGE2 (pTTQ18 with a second MCS). To have all racemases in similar constructs would reduce the variability in expression that may occur from using different systems. The gene was amplified from the gDNA by PCR, introducing an *Eco*RI cleavage site at the 5' terminus and a *Bam*HI site at the 3' terminus. The PCR product was inserted into

the corresponding sites of pTTQ18 and the resultant plasmid used to transform *E. coli* Top10 cells. Following plasmid purification, a restriction digest analysis was carried out (Figure 2.1.7). A gene insert similar to the size of the *alr* gene was shown by digestion with *Eco*RI/*Bam*HI and gel electrophoresis. Digestion with *Hind*III/*Nco*I confirmed the racemase gene was inserted into the plasmid in the correct orientation. Digestion with *Pvu*I was used to check the insert size. Placement of the racemase gene with respect to the LacIq gene was confirmed with digestion by *Eco*RV/*Sac*II. A single digest with *Eco*RI linearised and confirmed the size of the plasmid.

The pJLM201 construct was subjected to DNA sequencing. The sequences were aligned with the reported gene sequence and are shown in Appendix 3. The translation of the two sequences is also shown. There are only two base differences between the sequences (not including the two extra codons introduced by the cloning strategy at the start of the gene). Both of the base changes are the third base in the codon and have no effect on the sequence of the translated protein.

a)



b)

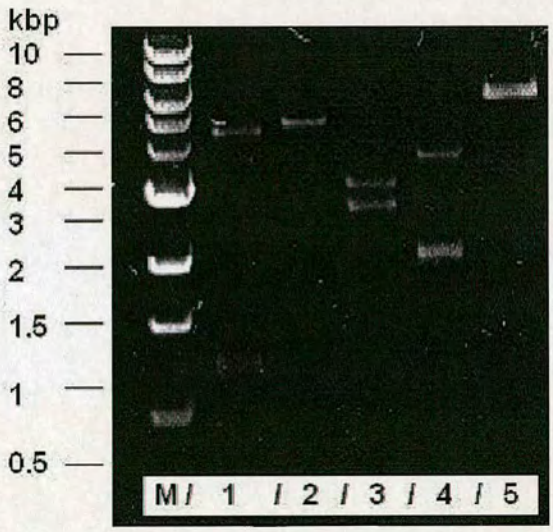


Figure 2.1.7: Diagnostic digest of pJLM201

a: Plasmid map displaying restriction sites used for diagmostic digest

b: Agarose gel analysis of diagnostic restriction digest. Expected band sizes are as follows: Lane M = 1 kb Marker (Biorad); Lane 1 = double digest with EcoRI/BamHI showing the expected bands at 1.2 and 4.5 kbp; Lane 2 = double digest with HindIII/NcoI showing the expected bands at 0.85 and 4.9 kbp; Lane 3 = single digest with PvuI showing the expected bands at 2.7 and 3.0 kbp; Lane 4 = double digest with EcoRV/SacII showing the expected bands at 2.0 and 3.7 kbp, Lane 5 = digest with EcoRI showing the expected band at 5.7 kbp for the linearised vector

2.1.6.2 The construction of pJLM301: the Ala racemase (*dal*) from *B. subtilis*

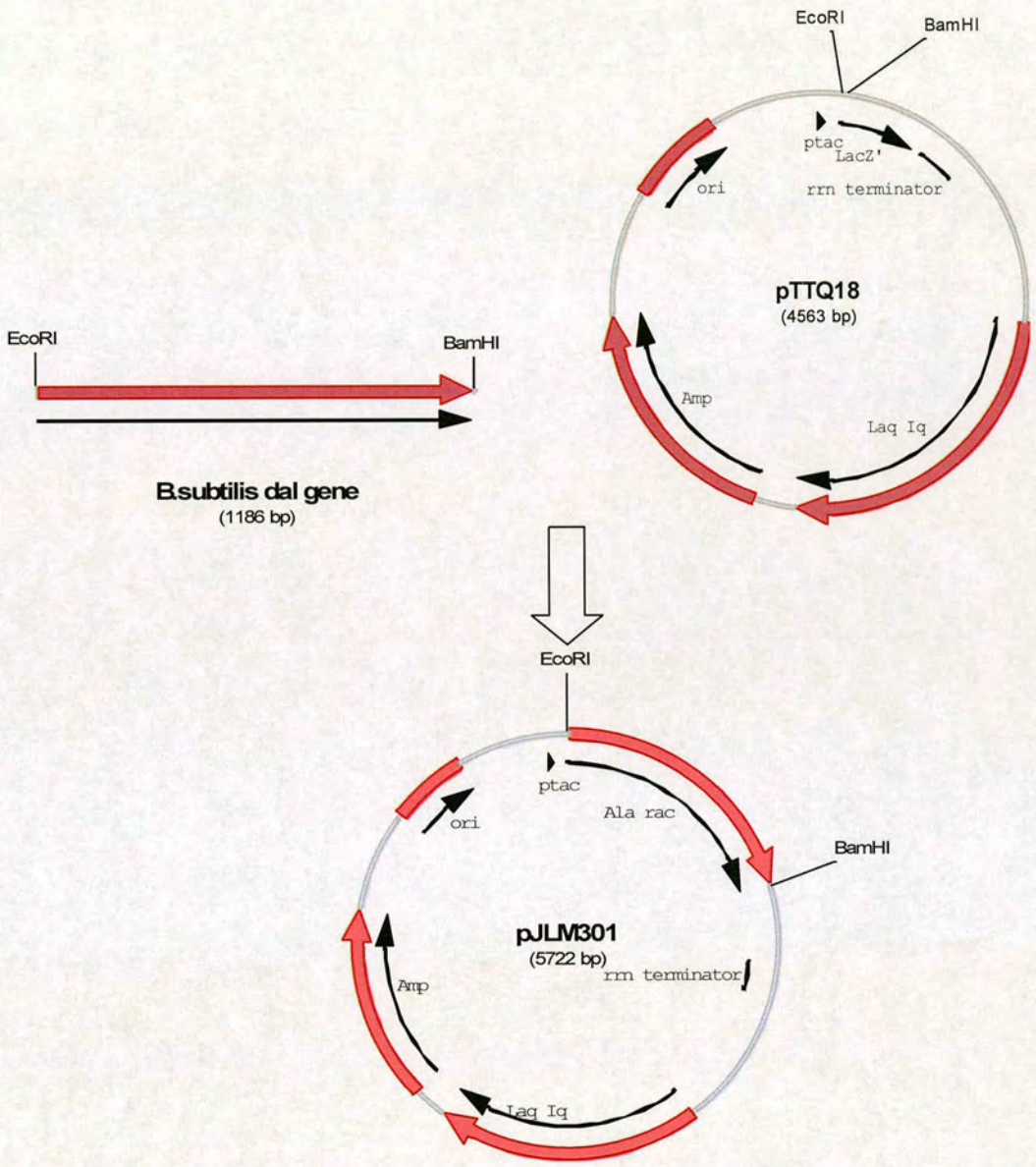


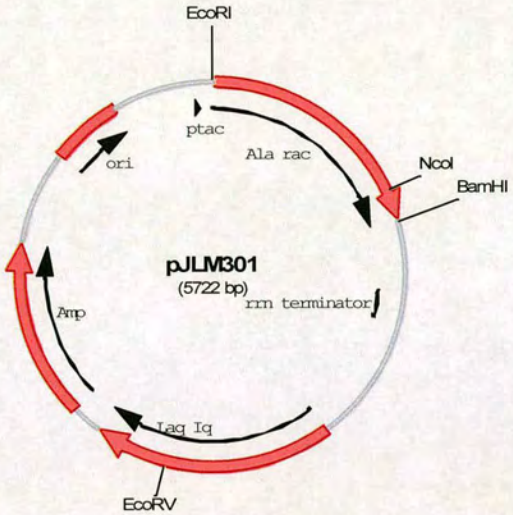
Figure 2.1.8: The cloning plan to create pJLM301 from pTTQ18 and the racemase gene

The *dal* gene was obtained from gDNA using PCR. An *Eco*RI cleavage site was introduced at the 5' terminus and a *Bam*HI site at the 3' terminus. The PCR product was purified, digested and ligated into pTTQ18 using the corresponding restriction sites. *E.*

coli Top10 cells were used to host the new plasmid. After plasmid purification, an insert equal to the size of the racemase gene was detected by restriction digest and gel electrophoresis. The plasmid was digested with *Bam*HI/*Eco*RI to excise the gene and confirm the size. Digestion with *Eco*RV/*Bam*HI confirmed the placement of the multiple cloning site with respect to the $lacI^q$ gene and restriction with *Eco*RV/*Nco*I confirmed the gene of interest was inserted in the site in the orientation predicted (with the C-terminus of the gene at the *Bam*HI site). Finally a digest with *Eco*RI was carried out to linearise the vector to confirm the size. The results of this digest are shown in Figure 2.1.9.

The gene was sequenced to confirm insertion of the *dal* gene and the alignment of the cloned gene with the expected sequence is shown in Appendix 4. There is only one base change at the end of the sequence when compared to the literature sequence (not including the codons introduced by cloning). The base change does not change the resulting amino acid as it is at the third base (more commonly known as the wobble position).

a)



b)

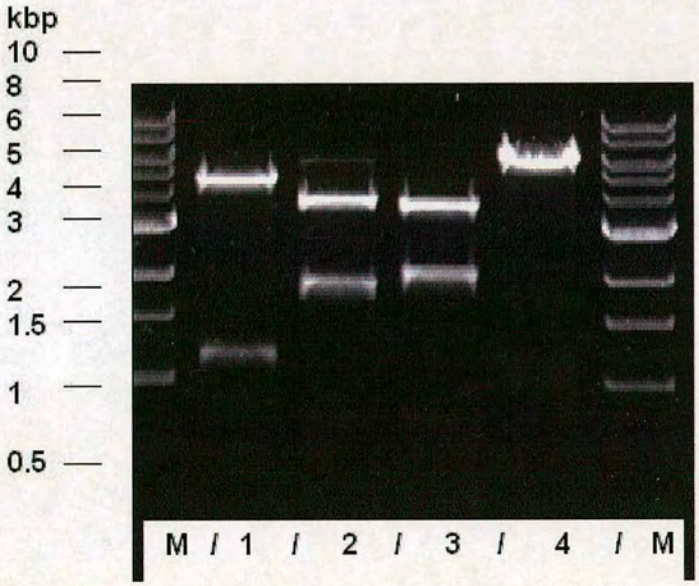


Figure 2.1.9: Diagnostic digest of pJLM301

a: Plasmid map displaying restriction sites used for a diagnostic digest

b: Agarose gel of diagnostic restriction digest with expected bands as follows: Lane 1= double digest with EcoRI/BamHI showing the expected bands at 1.2 kbp and 5.5 kbp; Lane 2= double digest with EcoRV/BamHI showing the expected bands at 2.0 and 3.7 kbp; Lane 3= double digest with EcoRV/NcoI showing the expected bands at 2.7 kbp and 3.0 kbp; Lane 4= digest with EcoRI only showing the expected band at 5.7 kbp corresponding to linearised vector

2.1.6.3 The construction of pJLM302: Plasmid encoding *alr* from *B. subtilis*

The PCR amplification of this gene from gDNA was unsuccessful. A range of annealing temperatures were tested (40-60˚C) combined with various detergents and solvents to aid primer binding, (DMSO, BSA and glycerol). The primer and dNTP concentrations were varied and different polymerase enzymes (*Taq* beads, *Pfu* hotstart and Advantage®-HF2 (Clonetech)) were employed but no successful DNA amplification was achieved. Multiple DNA fragments between 0.5 and 2.0 kbp were seen on gels at a range of annealing temperatures (42-52˚C) with 10 % glycerol added and at lower annealing temperatures (42 and 45˚C) with 1 % BSA added. The next step would have been to redesign the primers. The Sigma Genosys Oligo calculator was used to analyse the primers used. They were predicted to have weak and very weak secondary structures and thus very low probability of forming primer dimers. The calculated melting temperatures were within 5˚C of each other (64 and 67˚C). Therefore no problems were expected from the primers. The observed DNA fragment smears suggested non- primer binding and longer primers may have eliminated this problem. However, experiments on this gene were ended to focus on the other cloned racemases.

2.1.7 Expression tests of the racemases

The predicted sizes of the racemase monomers were calculated[103] from the DNA sequencing results. The predicted weights were 44.0 kDa for the *P. putida* racemase, 41.2 kDa for the *S. coelicolor* racemase and 43.6 kDa for the *B. subtilis dal* racemase.

2.1.7.1 Transformation of *E. coli* BL21(DE3) with pJLM101 and expression of the *P. putida* racemase

The plasmid pJLM101, with pTTQ18 as a control, was used to transform *E. coli* BL21(DE3). After an overnight growth there were few pJLM101 colonies visible. The plate was left for another 48 h. Colonies were picked and used to inoculate LB/Amp media. The three cultures were induced with 1 mM IPTG at OD_{600} 0.4-0.6 and harvested after 16 h growth. The expression of the racemase was tested with lysed cells using the

solution phase assay. No activity towards alanine or amino butyric acid was seen. An SDS-PAGE gel of the pTTQ18 and pJLM101 containing-cultures is shown in Figure 2.1.10 with an equivalent amount of total protein loaded into each lane. Analysis of the gel revealed no obvious over-expression of the racemase compared to the control. Due to the problems seen previously with the pET16b construct in *E. coli* BL21(DE3) and the proposal that basal expression of the racemase prior to induction was potentially toxic to the expression host, it was decided to try expression in *E. coli* Top10 and *E. coli* BL21(DE3)pLysS cells. The *E. coli* Top10 cell line was designed for production of high amounts of plasmid DNA and is not as efficient at expression as *E. coli* BL21(DE3).[104] *E. coli* BL21(DE3)pLysS was designed for the expression of toxic proteins as it contains a second plasmid (pLysS) that expresses a protein to inhibit T7 RNA polymerase, preventing basal expression of the protein of interest.[105]

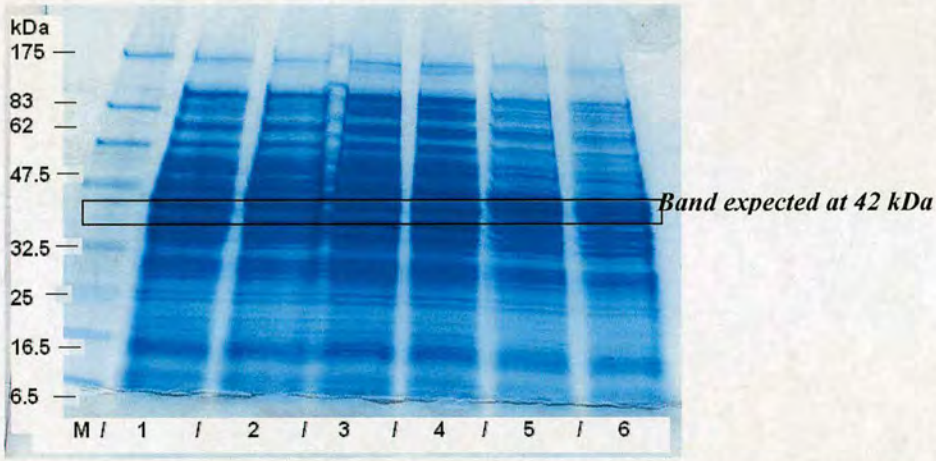


Figure 2.1.10:An SDS gel showing expression of P. putida racemase from pJLM101 and the pTTQ18 negative control. The contents of the lanes were as follows: M= Protein Ladder; Lanes 1-2 = 0h induction for pTTQ18 and pJLM101 respectively; Lanes 3-4 = 1h post-induction for pTTQ18 and pJLM101 respectively; Lanes 5-6= 16 h post-induction for pTTQ18 and pJLM101 respectively

2.1.7.2 Expression of the racemases in *E. coli* Top10 cells

The constructs were used to transform *E. coli* Top10 and two colonies from each were selected, grown and protein expression was induced with 1 mM IPTG when the OD of

the cultures was approximately 0.6. Samples were removed for analysis at 0 h, 1 h and 18 h post induction by SDS-PAGE (Figure 2.1.11).

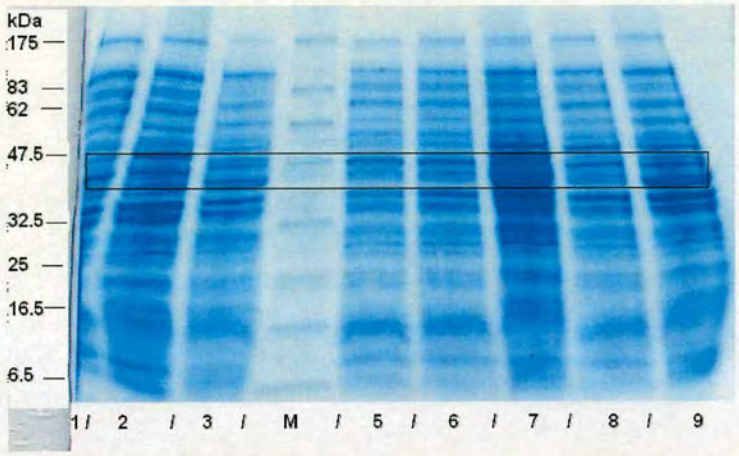


Figure 2.1.11a: An SDS gel showing the P. putida racemase expression in E. coli Top10 cells. The area of interest corresponding to the expected size of the racemase is boxed. The contents of the lanes are as follows: M= Ladder; Lanes 1-3 = correspond to the pTTQ18 Culture 1 (negative control) at 0 h, 1 h and 18 h post-induction respectively; Lanes 5-7 = P. putida racemase expression culture 1 at 0 h, 1 h, 18 h post-induction respectively; Lanes 8-9 = second P. putida racemase expression culture at 1 h and 18 h post-induction respectively

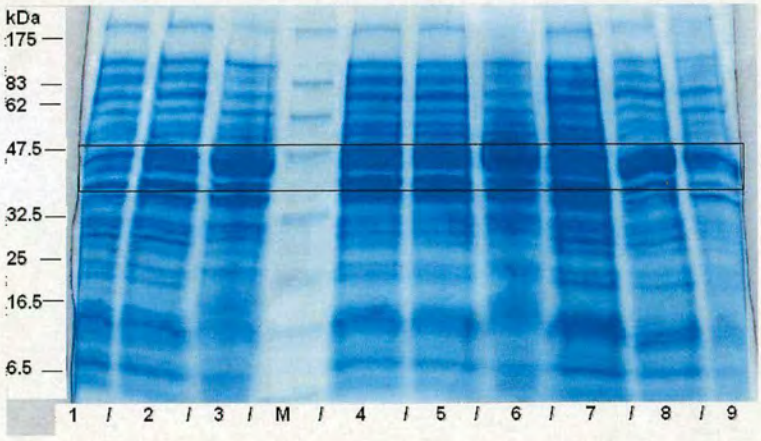


Figure 2.1.11b: An SDS gel showing racemase expression in E. coli Top10 cells. The area of interest corresponding to the expected racemase size is boxed. The contents of the lanes are as follows: M = Ladder; Lanes 1-3 = the S. coelicolor racemase containing culture at 0 h, 1 h and 18 h post-induction respectively; Lanes 4-6 = the second S. coelicolor racemase containing culture at 0 h, 1 h and 18 h post-induction respectively; Lanes 7-9 = the B. subtilis racemase containing culture at 0 h, 1 h and 18 h post-induction respectively

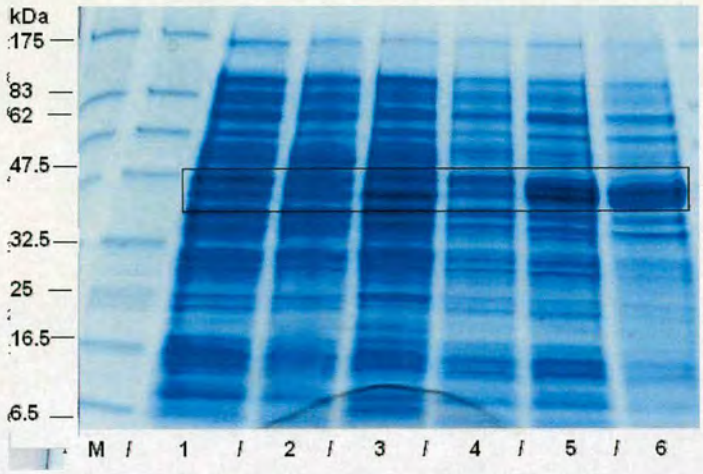


Figure 2.1.11c: An SDS gel showing expression of the racemases in E. coli Top10 cells. The area of interest corresponding to the expected racemase size is boxed. The contents of the lanes are as follows: M = Ladder; Lanes 1-3 = the second negative control culture at 0 h, 1 h and 18 h post-induction respectively; Lanes 4-6 = the second B. subtilis racemase containing culture at 0 h, 1 h and 18 h post-induction respectively

It is possible to see expression of the racemases in the 18 h, post induction lanes. As presented in 2.1.7.1, attempting to express the racemases in *E. coli* BL21 strains seemed to result in slow growth and eventual death of the host cells. The possible lower expression of the racemases in *E. coli* Top10 resulted in longer cell growth periods and visible protein bands by SDS-PAGE.

2.1.7.3 Expression tests in *E. coli* BL21(DE3)pLysS

Since large scale cultures were required for analysis and assay, *E. coli* BL21(DE3)pLysS was investigated as a host cell line for racemase over-expression. The racemase from *S. coelicolor* was chosen as the test candidate as expression in *E. coli* Top10 cells was high enough to visualise on an SDS gel (Figure 2.1.11b). It might be suggested that racemase expression is primarily toxic to the host when it is growing, as this is when the D-Ala and D-Glu are essential for cell wall biosynthesis, therefore it was decided to induce expression when the cells had reached late exponential phase of growth.

Initial experiments using *E. coli* BL21(DE3)pLysS[pJLM201] revealed that freshly transformed cells are essential. Glycerol stocks of pLysS cells were shown to be unstable and over time resulted in degradation of plasmid DNA. Plasmid DNA extracted from older glycerol stocks produced long DNA fragment smears when analysed on an agarose gel. Plasmids containing the racemase gene were extracted from freshly prepared *E. coli* Top10 and *E. coli* BL21(DE3)pLysS cells and were subjected to a restriction digest to extract the gene. It was shown that after a restriction digest to excise the racemase gene, both showed a single band corresponding to the size of the racemase gene but the pLysS cells also produced a DNA fragment smear. There is a second plasmid (the pLysS) in BL21(DE3)pLysS cells to express T7 lysozyme and the second antibiotic degrading enzyme (chloramphenicol acetyl transferase). It is suggested that the DNA smear in the freshly prepared plasmid is the degraded pLysS and that over time the racemase plasmid may also be destroyed. It was also shown that colonies picked from LB/agar plates that had been stored at 4°C for more than 2 days gave variable and unreliable results when grown and induced.

Overnight cultures were grown at 37°C from freshly transformed colonies, in LB/Amp/Cam/PLP (PLP concentration = 0.5 μM). A 2 % inoculum was used for five 50 ml cultures (containing the same antibiotics and PLP at the same concentration). When the OD at 600 nm reached approximately 1, two of the flasks were placed at 26°C, To two flasks at each temperature, 0.1 mM and 1 mM IPTG was added. The final flask at 37°C was not induced. 1 mM L-Ala was also added to provide substrate for the induced racemase so that it did not deplete amino acids essential for the host.

Flask (1): 37°C + 1 mM L-Ala
(2): 37°C + 1 mM L-Ala and 0.1 mM IPTG
(3): 37°C + 1 mM L-Ala and 1 mM IPTG
(4): 26°C + 1 mM L-Ala and 0.1 mM IPTG
(5): 26°C + 1 mM L-Ala and 1 mM IPTG

Aliquots from each culture were removed at the point of induction (t = 0) and at t = 1, 2 and 17 h post induction. The samples were centrifuged and resuspended to a final OD_{600} of 10, then 40 µl of the cell suspension was used to test for amino acid racemase activity using the solution phase assay (Section 2.2) with L-2-amino butyric acid (ABA). The results can be seen in Figure 2.1.12. 15 µl of each sample was loaded onto an SDS gel for analysis, see Figure 2.1.13.

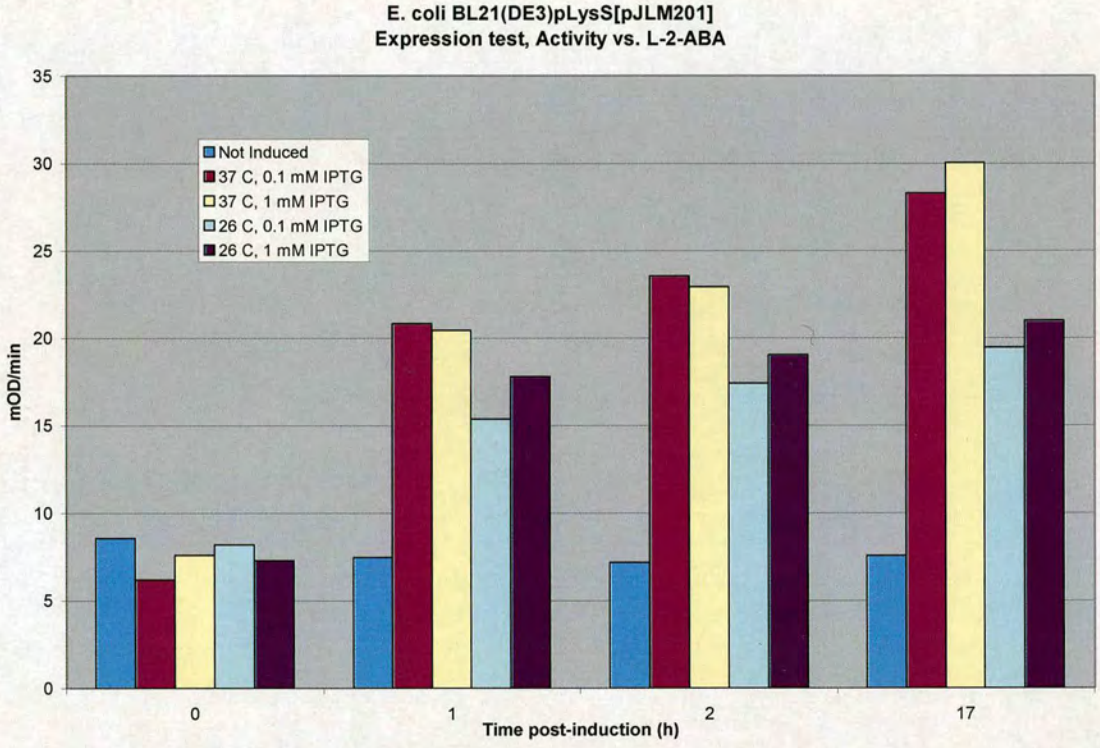


Figure 2.1.12: Histogram showing the amount of activity of the racemase enzyme at different times post induction at different growth temperatures with increasing IPTG concentrations

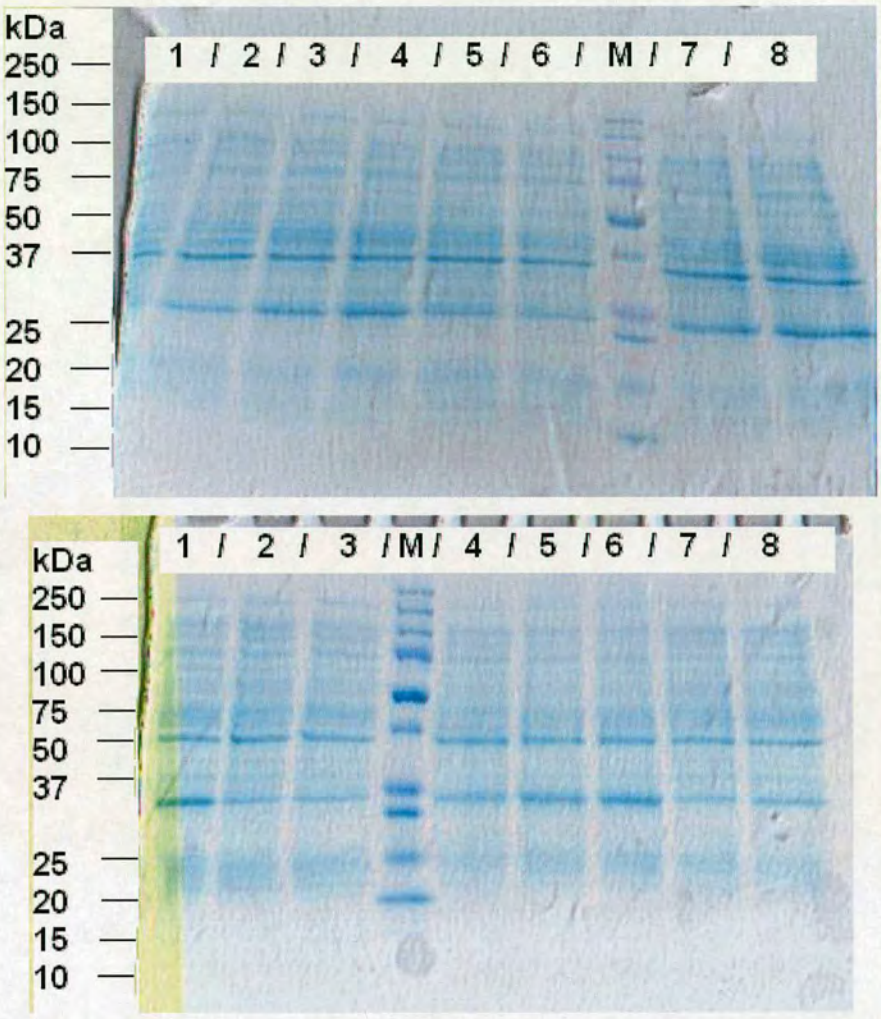


Figure 2.1.13: SDS gels showing the S. coelicolor expression in E. coli BL21(DE3) pLysS post induction at different growth temperatures with increasing IPTG concentrations. Band corresponding to racemase expected at 42 kDa. The conditions for each of the flasks was as follows: (1)= 37°C; (2)= 37°C 0.1 mM IPTG; (3)= 37°C 1 mM IPTG; (4)= 26°C 0.1 mM IPTG; (5)= 26°C 1 mM IPTG. The lanes in the gel contain CFE from the following flasks: Gel 1: Lane 1 = cell contents at time of induction; Lanes 2-6 = CFE of flasks (1)-(5) respectively at 1 h post-induction; Lanes 7-8 = Flasks (1)-(2) respectively at 2 h post-induction. Gel 2: Lanes 1-3 = CFE of flasks (3)-(5) respectively at 2 h post-induction; Lanes 4-8 = CFE of flasks (1)-(5) respectively at 17 h post-induction.

It was found that conditions (2) and (3), which are 37°C 0.1 mM and 1 mM IPTG respectively, produced the most active racemase. However conditions (2) also produced the highest final density of cells at 600 nm at cell harvest (17 h post-induction).

Condition (2) was therefore used for all large culture preparations. However, it was still not possible to positively identify over-expressed racemase on an SDS-PAGE gel (Figure 2.1.13).

2.1.8 Discussion

The aim of the work in this chapter was to identify and clone a number of amino acid racemases for the development of a screen and for directed evolution towards different substrates.

Plasmids containing the genes of three of the four racemases identified were successfully produced (pJLM101, pJLM201 and pJLM301). The fourth gene, *alr* from *B. subtilis,* was never successfully amplified using the designed primers. It is possible that the strain of *B. subtilis* in the lab has a slightly different sequence for the *alr* gene, compared to the published sequence and so the primers could not bind properly to amplify the gene, or there was non-specific binding. The next approach would be to redesign the primers and elongate them to enable more sites for binding recognition. A second approach could involve designing primers where the third base in the codon was undefined, (N) and would enable primer binding if the sequence was not as published in those positions.

The three racemases that were successfully cloned provided a good basis for the development of the screen to detect racemase activity, as one is a known broad substrate specificity racemase, one has high specificity for alanine and the final racemase is putative and has never been characterised.

The poor and inconsistent cell growth observed when cloning the broad substrate specificity racemase gene into pET16b, led to the proposal that the racemase may be toxic to the host when expressed. Attempts to isolate the plasmid of pET16b containing the cloned *P. putida* racemase failed, indicating that the cells had lost the plasmid.

However, this gene was cloned into pTTQ18, under the control of a different expression promoter, without problems. One explanation for the observed problems could be that as D-Ala and D-Glu are essential components in the bacterial cell wall,[67, 68] and if the racemase was expressed at a low level when the *E. coli* cells were just starting to grow the concentration of available D-Ala and D-glu would be reduced, causing inefficient cell wall synthesis.

The toxicity of this racemase in *E. coli* would also explain the poor cell growth encountered when trying to over-express the enzyme in *E. coli* BL21(DE3). Expression carried out using *E. coli* Top10 cells was more successful and the suggestion that can be made here is that the cell line was not designed as an expression system and so racemase expression levels are lower in growing cells compared to the *E. coli* BL21(DE3). For larger cultures, the best results for the most active racemase were obtained when using *E. coli* BL21(DE3)pLysS cells and inducing with an IPTG concentration of 0.1 mM when the cells were at the end of the exponential growth phase. The lower amount of IPTG meant that the racemases were expressed at a lower level and the cells could grow to a higher density.

The *P. putida* racemase was over-expressed in BL21(DE3)pLysS cells but not in BL21(DE3) cells, the difference between the two cell lines being the plasmid pLysS. The pLysS plasmid constitutively expresses a low amount of T7 lysozyme, an inhibitor of the T7 RNA polymerase used for T7 expression systems such as the pET vector series. The lack of functioning polymerase prevents basal expression of the cloned gene, until IPTG is added and the amount of expressed T7 RNA polymerase increases. In the plasmids constructed here containing the racemase genes, the promoter system is *tac* rather than T7. There is no T7 RNA polymerase to inhibit, yet expression was more successful and resulted in active racemase, Figure 2.1.13. Comparison of the two strains for expression was repeated later in these studies and confirmed that there was little racemase expression in BL21(DE3) but expression was observed in BL21(DE3)pLysS.

The observed instability of the plasmids in stored *E. coli* BL21(DE3)pLysS cells may be due to the T7 lysozyme. The lysozyme is contained within the cell, but upon storage the cell may die, releasing the lysozyme and resulting in lysis of surrounding cells. This would be the start of a chain reaction, as surrounding cells would then lyse and release their lysozyme.

2. Results and Discussion

2.2 Development of a high-throughput screen to detect racemase activity

The ability to screen for enzyme activity is essential for directed evolution experiments. It is generally accepted that the major limiting factor in the application of directed evolution approaches is the availability of suitable high-throughput screens or selection systems for the enzyme activity of interest. The current methods of screening for racemase enzymes described in the introduction are not very high-throughput. The following chapter presents the development of an assay to detect racemase activity, both in the solution phase and on solid phase using agar-based screening methods. Both versions of the screen rely on the detection of hydrogen peroxide, generated by the enantioselective oxidation of an amino acid by an amino acid oxidase. The principle of the racemase screen was to introduce the unreactive amino acid enantiomer for the oxidase, so that if the racemase was able to racemise the substrate, the correct enantiomer for oxidation would be produced, oxidation could occur and the hydrogen peroxide could be detected using HRP and a colorimetric substrate, Figure 2.2.1.

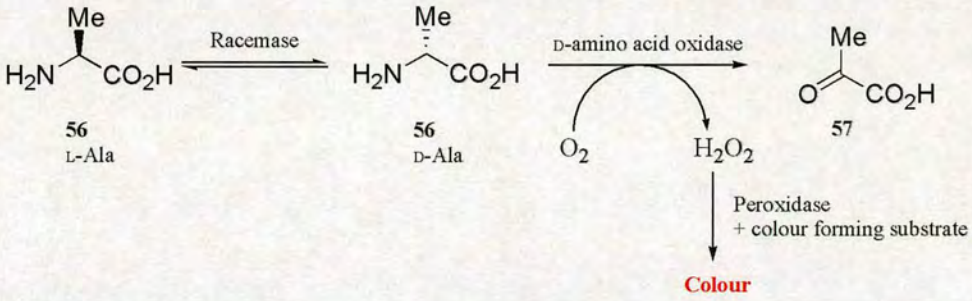


Figure 2.2.1: The racemase detection assay involving production of hydrogen peroxide that can be detected colorimetrically.

The first section of this chapter describes the use of the assay in the solution phase using commercially available racemase and amino acid oxidase enzymes. The screen was then tested using the cloned amino acid racemases and optimised to ensure reproducible results. The solution phase assay was used to obtain a substrate specificity profile of the expressed racemases against a panel of substrates. As the assay is an indirect detection of racemase activity, it had to assumed that the H_2O_2 detected by the colorimetric assay is produced as a consequence of enzymatic racemisation. To further confirm racemase activity and the principle of the solution phase using a direct method, HPLC analysis of an amino acid racemisation reaction was carried out.

Analysis of racemase activity in the solution phase is preferable when working with small numbers of samples. However, the aim was to be able to screen large numbers of variants in a high-throughput screen. A common technique is the use of agar plate-based screening methods. The numbers of variants that can be screened using these methods are limited by the number of agar plates used, each plate (15 cm diameter) can hold approximately 6 000 to 8 000 colonies when covered. The same reaction principles from the solution phase screen were applied in the solid phase, using the oxidase to oxidise the racemase product and produce hydrogen peroxide, which in turn could be detected using a peroxidase and a colour forming substrate. The development of the solid phase screen is presented, from a simple 'pour-over' reaction mixture, to a two-plasmid screen, capable of screening large numbers of variant enzymes using a range of oxidases.

2.2.1 The solution phase assay using commercially available enzymes

The performance of the screen was tested initially using commercially available; amino acid racemase (from *Geobacillus stearothermophilus*), the D-amino acid oxidase (from porcine kidney) and the horse radish peroxidase. (For preparation see Section 4.6.1). Assays were performed at pH 8.0 as the optimum D-amino acid oxidase performance has been reported to be pH 8.5[61] and the optimum working pH for racemase activity is

between pH 8.0 and 10.0.[106] However the optimum activity for the peroxidase has been reported between pH 5 and 6.[106] Using pH 8.0 enabled optimum activity of the racemase and slightly less than optimal activity of the oxidase. The HRP was added in excess to compensate for the pH being higher than for optimum activity.

The commercial enzymes were diluted with dH_2O, according to the enzyme unit activity reported by Sigma towards alanine. Into each well on a 96-well micro-titre plate containing 190 µl assay mixture (including HRP), 10 µl of racemase/oxidase was added in increasing concentrations, shown in Table 2.2.1. The plate was then incubated at 37°C and the reaction monitored by measuring the absorbance of the coloured product at 510 nm. Absorbances were plotted as a function of time. A picture of the plate after reaction can be seen in Figure 2.2.2, reactions were compared by measuring the greatest increase in absorption per min and these results are also shown in Table 2.2.1.

(mOD/min)	**1**	**2**	**3**	**4**
	0 U Oxidase	**1 U Oxidase**	**2 U Oxidase**	**3 U Oxidase**
A: 0 U Racemase	0	0	0	0
B: 1 U Racemase	0	3.9	28	79
C: 2 U Racemase	0	12	168	288
D: 3 U Racemase	0	53	325	446

Table 2.2.1: The racemase:oxidase unit ratios used to test the solution phase assay and the mOD/min values for colour formation

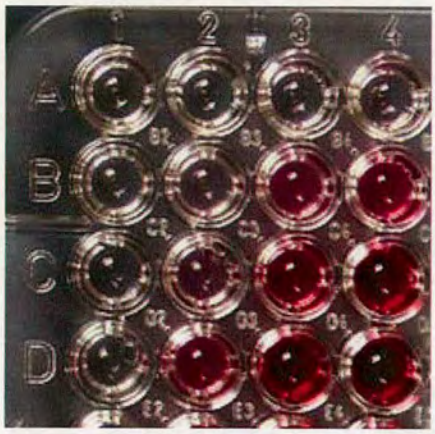


Figure 2.2.2: Picture of plate after assay showing gradual increase in colour formation with increasing oxidase and racemase concentrations

Wells containing both the commercial racemase and oxidase resulted in colour change and an increase in either racemase or oxidase resulted in an increase in the rate of formation of coloured product. The controls containing no racemase and/or no oxidase did not show any colour formation during the reaction time. It is necessary to use a negative control in all assays as the dye product forms naturally due to the light sensitivity of the reagent, 4-amino antipyrine. In conclusion, the solution phase assay was shown to detect racemase activity using the oxidase/peroxidase coupled screen.

2.2.2 The solution phase assay for the cloned racemases

The racemases from the three constructs pJLM101, pJLM201 and pJLM301 were expressed in *E. coli* Top10 cells and the cell free extracts subjected to a solution phase assay against increasing concentrations of L-Alanine. The cells were harvested and lysed to obtain the CFE prior to assay (Section 4.5.1.1). The CFE was then added to assay mixture containing HRP, L-Ala and the amino acid oxidase, (Section 4.6.1.3).

Negative controls for each Ala concentration lacking the CFE were assayed and the results were subtracted from the racemase data. The results are displayed in Figure 2.2.3.

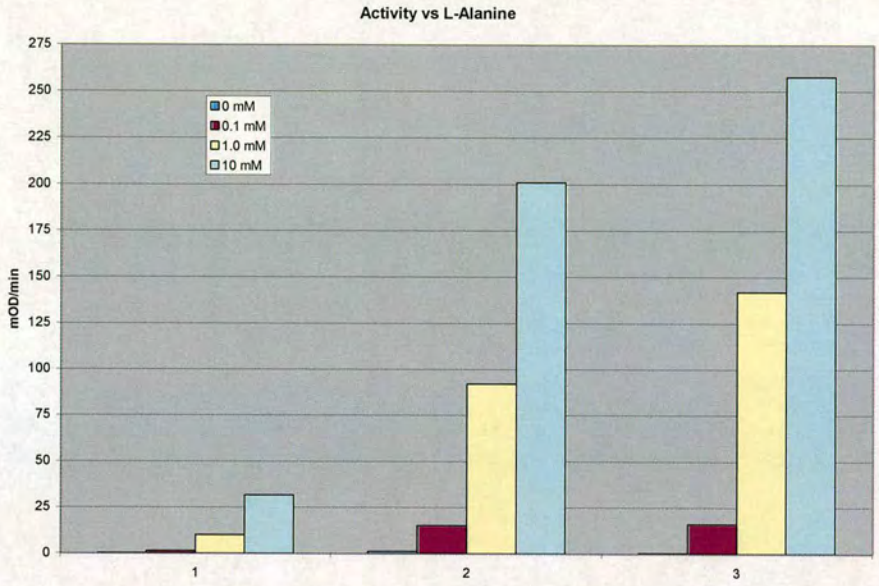


Figure 2.2.3: Histogram showing CFE racemase activity towards different concentrations of L-Ala. The culture expressing the P. putida racemase is shown in 1; The culture expressing the S. coelicolor racemase is shown as 2; The culture expressing the B. subtilis racemase is shown in 3.

There was little colour formation in wells containing the CFE and no substrate. The *E. coli* cells expressing the racemase from *B. subtilis* (pJLM301) showed the highest activity and the lowest activity was seen for the cells expressing the *P. putida* broad specificity racemase (pJLM101).

The solution phase assay was shown to detect racemase activity in expressed cells but the results observed above may have been due to limited oxidase activity. Therefore an experiment was performed to verify that the amount of oxidase used in the screen (0.2 U) was sufficient to detect 0.1 mM, 1.0 mM and 10 mM D-Ala, without being rate limiting and the results can be seen in Figure 2.2.4.

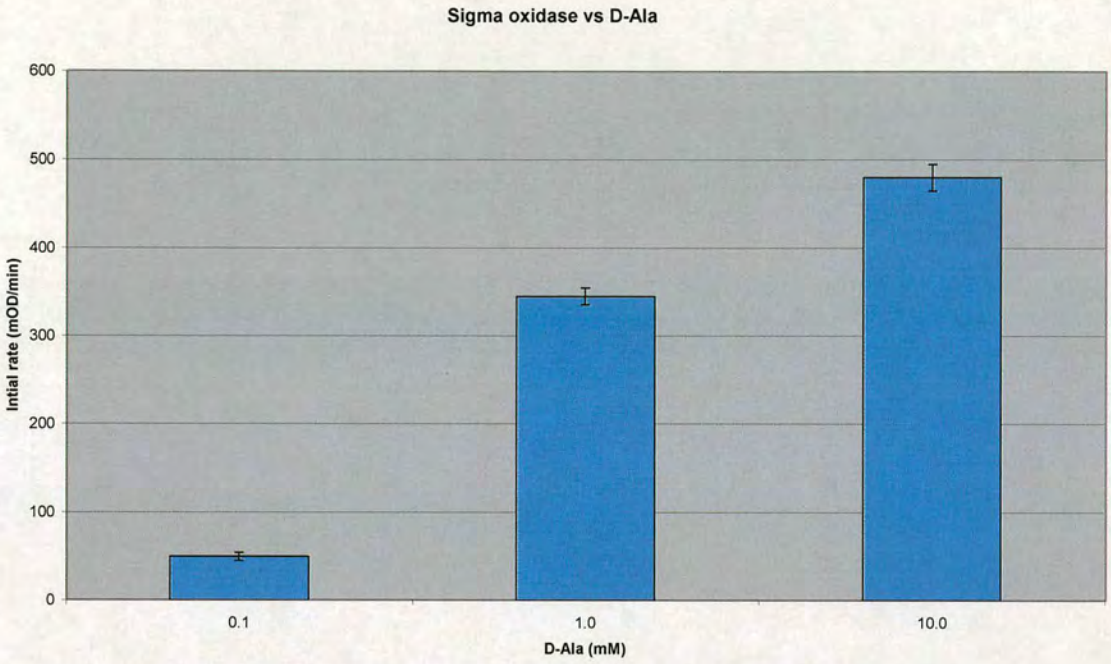


Figure 2.2.4: Maximum mOD/min values obtainable using 0.2 U Oxidase

The values obtained for oxidation of 0.1, 1.0 and 10 mM D-Ala using 0.2 U of oxidase are 2-3 times greater than those obtained when combined with the racemase and racemase substrate. This confirms that the oxidase and HRP were capable of detecting the maximum amount of D-Ala that could be present after racemisation by the racemase.

2.2.3 Determining substrate specificities of the racemases using the solution phase assay

In order to determine the substrate specificity of the racemases, a solution phase assay was carried out using the CFEs of the expressed racemases against a range of amino acid substrates. All cells were harvested and lysed to obtain the CFE, which was then assayed using the panel of amino acid substrate assays. The reactions were all carried out in triplicate and the average taken. Cells containing the plasmid with no racemase gene (pTTQ18) were also assayed to provide a negative control. All averaged results for ten amino acids are shown in Figure 2.2.5.

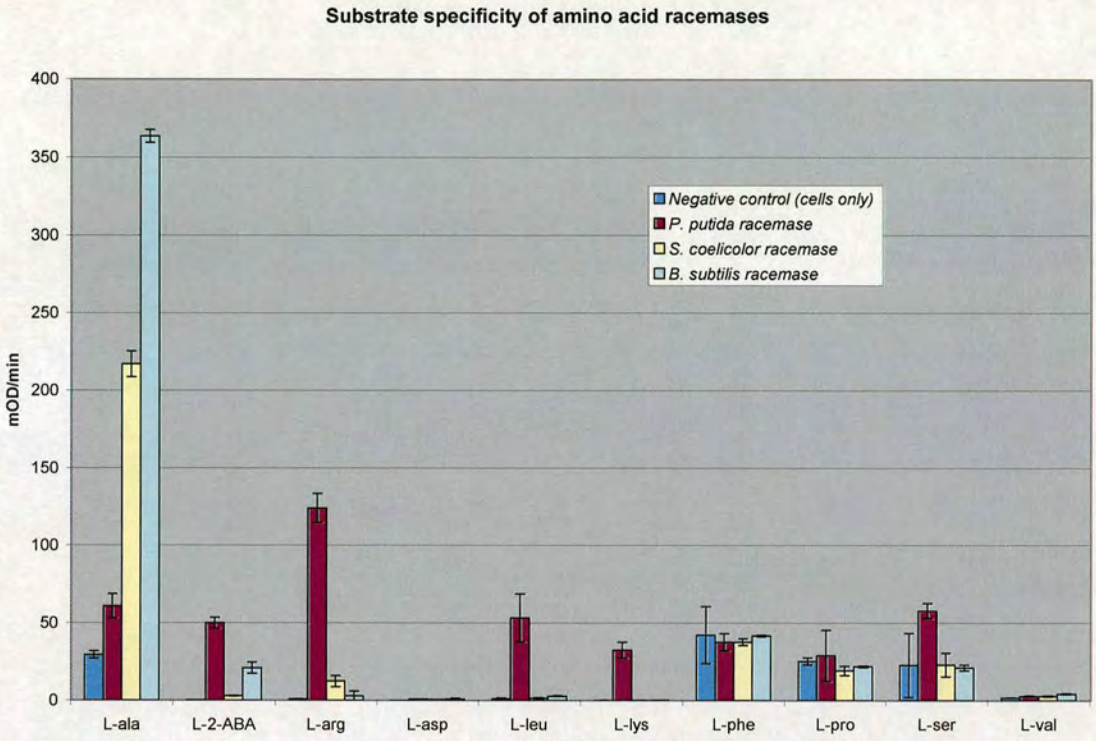


Figure 2.2.5: Substrate specificity of racemase activity, including pTTQ18 negative control

The negative control is displayed in Figure 2.2.5 to indicate the substrate specificity of the host *E. coli* racemases. The racemase from *P. putida* has activity towards alanine, amino-butyric acid, arginine, leucine, lysine and serine, but no activity towards aspartate, phenylalanine, proline and valine. This data correlates very well with the relative substrate specificity recorded by Lim et al in 1993,[90] the only discrepancy was that they observed a higher activity towards lysine.

The substrate specificity of the putative racemase from *S. coelicolor* towards the substrates screened was shown to be very narrow, with low activity with L-2-ABA and high activity with L-alanine. Similar results were obtained for the Dal racemase from *B. subtilis*, which was very active with L-alanine and slightly active with L-2-ABA. This correlates well with the information previously recorded for the Dal racemase.[101]

The substrate specificity of the *E. coli* host racemases is visible in the negative control. Activity against alanine, phenylalanine, proline and serine was observed. To analyse behaviour of the expressed racemases in future experiments, an amino acid that is active against the expressed racemases but not against the host racemases should be used. The amino acid 2-ABA is a good candidate as no activity was observed against this substrate with the negative control, but activity was seen with all the expressed racemases.

It is possible in these assays that there is a small quantity of the D-amino acid in the L-amino acid source, manufacturers labelled the L-amino acids as 99% pure. However, small contaminants of the D-enantiomer can produce false results in the assay. Assays carried out later (Section 2.7.2) suggested that the activity displayed in Figure 2.2.5 by the host *E. coli* racemases towards substrates phenylalanine and proline were indeed false and may have been a result of D-Phe and D-Pro present in the L-Phe and L-Pro sources respectively.

2.2.4 Optimisation of the solution phase screen

At this stage an analysis of the solution phase assay was carried out to minimise errors. It was noticed in the previous assay that the data range within each triplicate could be made smaller and more reproducible. The enzyme volumes added to the assay mixture were susceptible to great errors, for example, if the required volume addition of oxidase was 5 µl, to pipette 4.8 µl, into one well and 5.2 µl meant a difference between the wells of 0.4 µl. To reduce these errors and improve reproducibility, it was decided to increase the total volumes added.

It was decided to prepare a two times concentrated assay mixture and aliquot 100 µl per well. This left 100 µl for the addition of enzymes, CFE and/or water. For all further experiments the remaining volume consisted of 10 µl HRP 1mg/ml (volume doubled), 10-20 µl DAAO 0.2 U (previously 5 µl) and 50-70 µl CFE. Working with larger

volumes minimised pipetting errors and gave more reliable results. Assay data then measured in triplicate was shown to be comparable using the new technique. However, the relative substrate specificity data obtained previously was shown to be accurate.

2.2.5 HPLC analysis of racemase activity

To verify that the colour formation in the solution phase assay was a result of substrate racemisation, the reaction of L-2-ABA with whole cells expressing each racemase was followed by HPLC.

2.2.5.1 Preparation and analysis of standard solutions

To enable analysis of a racemisation experiment, standards of 10 mM L-, D- and *rac*-2-ABA were prepared. The *rac*-2-ABA was prepared and diluted to examine reproducibility of sample preparation (labelled A1). Diluting A1 to 2.5 mM gave sample A2, A2 was then used to make A3 (1.25 mM *rac*-2-ABA) and A3 used to make A4 (0.625 mM). Each standard (A2, A3 and A4) was sampled in triplicate, mixed in the HPLC injector loop for derivatisation (Section 4.5.3.5) and injected onto the column. 2 mM L-2-ABA and 2 mM D-2-ABA standards were prepared and each one analysed twice on the HPLC to confirm retention times. The peak areas returned for all readings of A2, A3 and A4 and for the single enantiomers are shown in Figure 2.2.6. The average, standard deviation, standard error and relative standard deviation (%RSD, =100x (standard deviation/average)) were calculated to analyse the reproducibility of the machine and the sampling. The HPLC sampling and derivatisation varies between 0.7 and 2.8 %RSD.

Standard curves for the peak areas against concentrations were plotted and lines of best fit were drawn with error bars to determine concentrations from peak areas. The plot of both D- and L- peak areas can be seen in Figure 2.2.7. The equation to determine the L-2-ABA concentration (x) from peak area (y) is y=1416.8x. The equation to determine the D-2-ABA concentration (x) from peak area (y) is y=1470.3x.

	A2 sample		A3 sample		A4 sample			
2-ABA	*Racemic*						D-2-ABA	L-2-ABA
Total concn (mM)	2.5		1.25		0.625			
Individual enantiomer concn (mM)	1.25		0.625		0.3125		2	2
	Peak areas							
	L	D	L	D	L	D		
Injection 1	1761	1775	878	879	452	452	2974	2780
Injection 2	1799	1813	884	889	444	443	3003	2892
Injection 3	1732	1743	905	913	448	447		
Average peak area	1764	1777	889	894	448	447	2989	2836
Standard dev	33.6	35.0	14.2	17.5	4.0	4.5	20.5	79.2
Standard error	19.4	20.2	8.2	10.1	2.3	2.6	14.5	56.0
%RSD	1.9	2.0	1.6	2.0	0.9	1.0	0.7	2.8

Figure 2.2.6: HPLC standard peak areas and error analysis

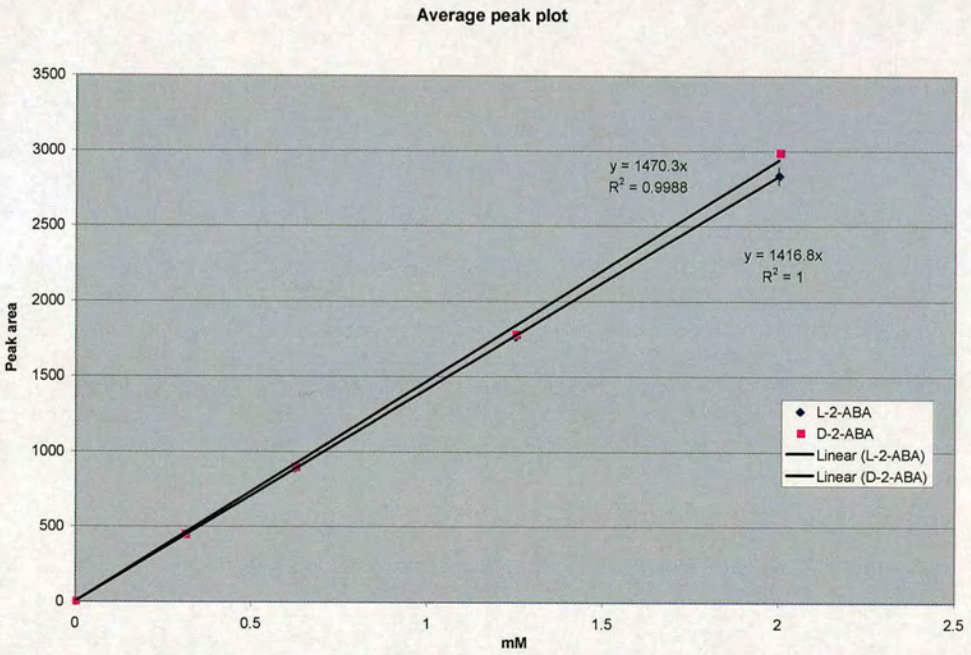


Figure 2.2.7: A graph showing the results of the HPLC peak areas at known concentrations of ABA, including error bars. The lines of best fit are shown and the equations allow peak areas to be used to calculate the unknown 2-ABA concentrations

2.2.5.2 Analysis of racemisation reactions by racemase-expressing cells

E. coli BL21(DE3) pLysS cells expressing each of the three racemases (from *P. putida*, from *S. coelicolor* and *B. subtilis*) were harvested 16 h after induction (Section 4.4.5). The racemisation of 12 mM D- and L-2-ABA was attempted using 5 % (w/v) cells in phosphate buffer, pH 8.0 (Section 4.6.3.3) in duplicate. Samples were removed at 0.1 h and 15 h. HPLC analysis of the enantiomers was carried out and e.e.s were calculated for each time point using the peak area and the standard equations and are plotted in Figures 2.2.8 and 2.2.9. After 15 h, the e.e. had fallen to between 0 % and 3 % D-2-ABA, Table 2.2.2, indicating that successful racemisation had occurred.

The racemisation of 2-ABA was shown to be successful at substrate concentrations of 12 mM D- and 12 mM L-2-ABA. In both cases, there was a small e.e. of the L-conformer after 15 h. The racemase from *B. subtilis* appeared to be the most active and produced an e.e. of 17 % and 4 % in 0.1 h, starting from 99 % D- and L-2-ABA respectively.

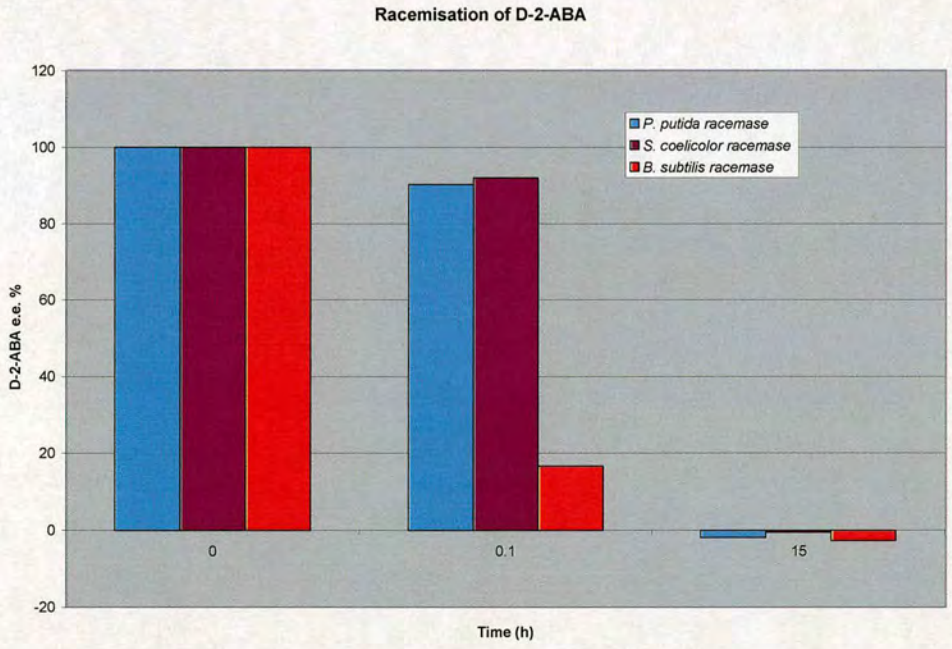


Figure 2.2.8: Histogram showing the e.e. values in the racemisation of D-2-ABA using cells expressing the racemases. The % e.e. of D-2-ABA is plotted as a function of time

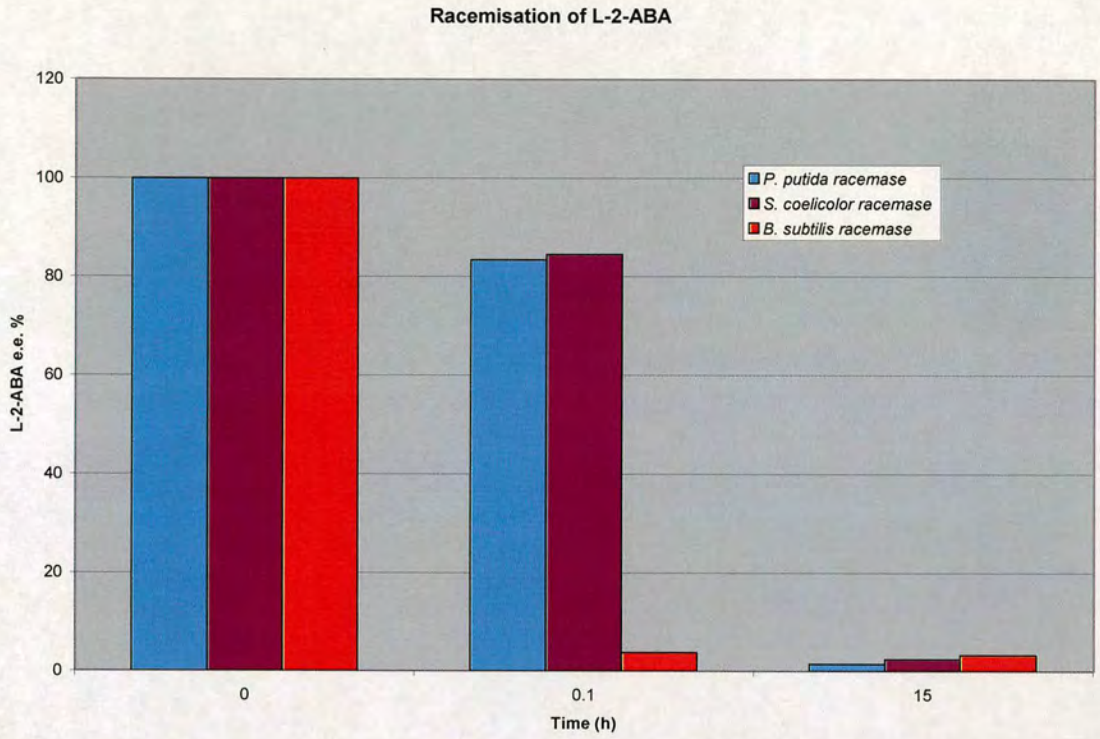


Figure 2.2.9: Histogram showing the e.e. values in the racemisation of L-2-ABA using cells expressing the racemases. The % e.e. of L-2-ABA is plotted as a function of time

2.2.5.3 Observations

In all cases racemase activity was demonstrated and there was a slight e.e. of L-2-ABA after 15 h. The rate of enzyme-catalysed racemisation from L→D–amino acids has been proposed to be faster than the conversion of D→L-amino acids.[107] The difference in conversion rates was seen when the e.e.s at t = 0.1 h for all enzymes studied were compared for both D- and L-2-ABA racemisations, as L-2-ABA was racemised faster.

In summary, it has been shown that racemisation of the substrate was occurring, with all racemases and importantly, that the correct substrate for a D-amino acid oxidase is being formed in the racemase screen. Once the screening had been explored in the solution phase, it was possible to begin investigation of the use of the screen in the solid phase.

2.2.6 Evaluation of the solid phase assay with the cloned and expressed racemases

The solution phase assay was shown to be successful in detecting amino acid racemase activity. However, the solution phase assay was only medium-throughput (96 different samples per plate, each plate was analysed for between 30 and 60 min). For the application of the developed screen in directed evolution, it was crucial to be able to assay large numbers of samples more quickly. Therefore a solid phase assay was developed for detecting racemase activity using the same principles as the solution phase assay, *i.e* to employ cells expressing the racemase and add the L-amino acid substrate, the D-amino acid oxidase and the hydrogen peroxide detecting reagents to identify active enzymes.

Colonies of *E.* coli cells expressing each of the racemases were grown on nitrocellulose membranes on LB/agar plates. The membranes allowed the colonies to grow on normal agar-containing media and facilitated the transfer of the colonies to the assay mixture. Following incubation of the membranes at –20°C to partially lyse the cells, each membrane was cut in half. An assay mixture containing 10 mM L-Ala, DAAO, HRP, colour forming substrate (DAB) buffered at pH 8.0 was poured onto one half membrane and onto the other half, the same assay mix without the L-Ala substrate. The latter was a control to ensure colour change was not the result of a possible contaminant or side reaction in the assay mix, but a result of conversion of L- to D- alanine. The results can be seen in Figure 2.2.10.

The expected red/brown colour was observed over the membranes expressing racemase. No colour formation was seen in the plates lacking substrate, indicating that the colour must have been the result of a reaction between the colonies on the membranes and the provided substrate. However, the colour had diffused throughout the agarose and was not localised over the colonies as required to allow identification of individual colonies with the activity of interest. It seems likely that the racemase was indeed racemising L-alanine to produce D-alanine, which subsequently diffused through the agarose faster than it was oxidised. As the DAAO and colorimetric substrates were distributed evenly

throughout the assay plate, colour formation occurred where the D-alanine diffused and not only over individual colonies as desired. The D-alanine was seen to diffuse up to 1 cm from the edge of the membrane, as revealed by coloured product.

(i) positive control, D-Ala (ii) negative control, no sub (iii) B. subtilis racemase, no sub (iv) with L-Ala

v) S. coelicolor racemase, L-Ala (vi) P. putida racemase, no sub (vi) with L-Ala

Figure 2.2.10: The solid phase assay performed on all constructs

An attempt to reduce the colour distribution was made by using 10 and 100 times less L-alanine (1 and 0.1 mM) but this approach did not produce definite areas of colour over specific colonies. Lowering the concentration of L-Ala also reduced the intensity of the colour and would ultimately lower the ability of the screen to detect activity. The addition of more DAAO in the assay mixture would have made the cost of the high-throughput screen unfavourable for extensive use.

Another approach to solve the problem was considered, involving the co-expression of the racemase and the oxidase in the same cell, therefore localising oxidation in the colony so that even if any D-alanine diffused through the agarose, it would not result in colour as oxidation would occur only in the cell.

2.2.7 Development of a two-plasmid screen

It was decided to introduce an amino acid oxidase into the cell on a second plasmid. If both genes were on the same plasmid, problems would occur when making variants using the XL1-Red mutator strain as this could mutate the oxidase gene also. Putting the oxidase on a second plasmid that is only introduced for screening reduced the possibility that any racemase activity would be missed due to a deactivated oxidase. Additionally, the plasmid bearing the oxidase gene could be easily removed for the scale up of racemisation reactions.

The co-expression of two enzymes from separate plasmids in one cell requires that the plasmids have compatible origins of replication and different antibiotic resistance markers. Most commonly used *E. coli* plasmids contain the ColE1 replication origin, or a derivative of, which results in a high number of copies of the plasmid per cell (300-500) and the racemase genes were constructed on such a plasmid.[99, 108] Compatible origins of replication, such as pI5a result in a low plasmid copy number (10-12 per cell). It was decided to leave the racemase on the high copy plasmid and introduce the oxidase on a low copy plasmid and select for cells containing both plasmids using two antibiotics. The promoter systems were also chosen to be different to enable selective control of the expression of both enzymes. Therefore, a low copy number plasmid containing an oxidase gene and an antibiotic resistance gene other than for ampicillin was created.

2.2.7.1 Creation of plasmid pJLM501 containing an amino acid oxidase gene

Step 1: Combining a heat inducible promoter with a Kan resistant gene

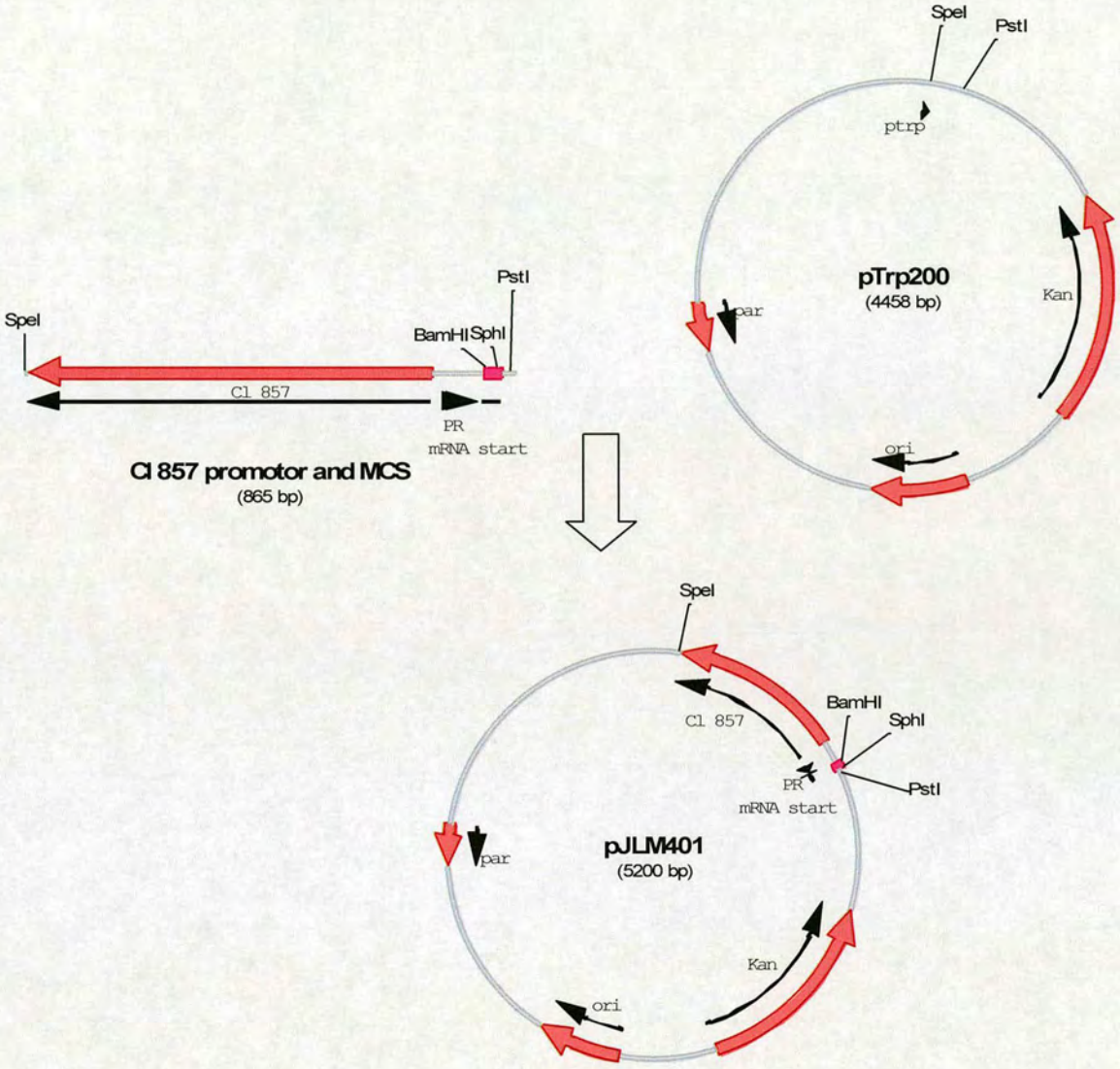


Figure 2.2.11: Creation of pJLM401

The heat inducible promoter and MCS from pPOT3 was inserted into Trp200

A heat inducible phage promoter (CI 857) was obtained from the plasmid pPOT3[109] by PCR amplification (to introduce *Spe*I and *Pst*I restriction sites at 5' and 3' ends respectively) and inserted into pTrp200 to make pJLM401, Figure 2.2.11.

Step 2: Insertion of a D-amino acid oxidase gene under the control of the heat inducible promoter

The D-amino acid oxidase gene, from *Trigonopsis variabilis,* was amplified from the plasmid pRES151 (Ingenza Ltd). The *Bam*HI and *Pst*I digested gene and plasmid pJLM401 were ligated to make pJLM402, Figure 2.2.12. Digestion of pJLM402 confirmed introduction of the oxidase gene.

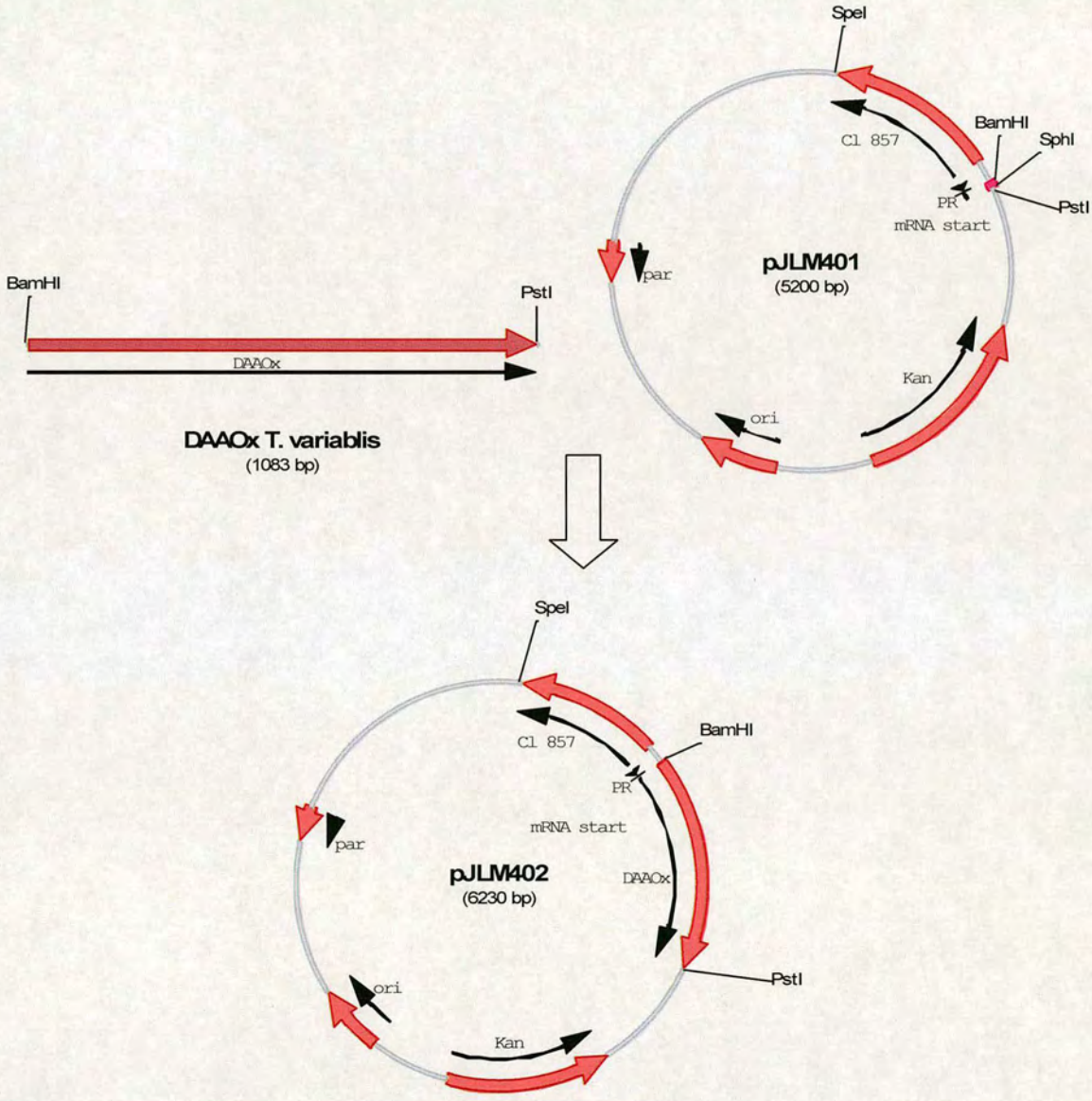


Figure 2.2.12: Creation of pJLM402 by insertion of the oxidase gene into pJLM401

Step 3: Transferring the oxidase gene/kan resistant gene/heat inducible promoter cassette into pJLM402

The entire heat inducible promoter and oxidase were excised from pJLM402 using *Spe*I and *Pst*I, Figure 2.2.12, as pTrp200 and subsequently pJLM402 were high copy number plasmids. The promoter and oxidase gene fragment were inserted into the low copy number plasmid pTrp338, to make pGL12, (Figure 2.2.13).

Step 4: Removing one of two *Bam*HI sites from pGL12 to enable future replacement of the oxidase gene

The high-throughput screen was developed to detect novel amino acid racemase activities and relied on the oxidation of the amino acids by an amino acid oxidase. For the wider application of the screen to detect activities such as ketone reductase activity, an alcohol oxidase would be needed. The restriction site at the beginning of the amino acid oxidase gene was *Bam*HI. However, there was a second *Bam*HI restriction site in the plasmid, after the heat inducible promoter. To enable replacement of the amino acid oxidase gene, the plasmid was subjected to a Quikchange PCR to remove the *Bam*HI site in the pGL12 vector backbone, (Section 4.2.14.4). The modified low copy plasmid, containing the oxidase gene, a heat inducible promoter and a kanamycin resistant coding gene, was called pJLM501, Figure 2.2.13.

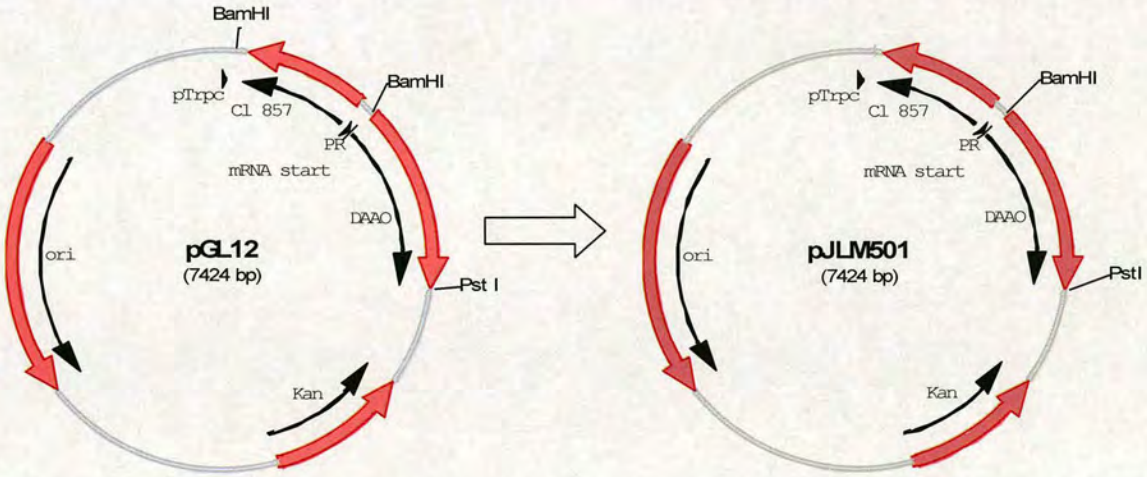


Figure 2.2.13: The creation of pJLM501 by removal of a BamHI site from the pGL12 backbone

The removal of the *Bam*HI site was confirmed by sequencing of the pGL12 and pJLM501 plasmids and by a diagnostic digest of pJLM501, Figure 2.2.14. The plasmid was digested with *Bam*HI/*Pst*I to excise the oxidase gene, with *Eco*RI to confirm the placement of the heat inducible promoter, with *Sac*I to confirm the presence of the oxidase/promoter cassette, with *Nde*I to check the host vector and with *Bam*HI to linearise the vector and confirm the size. The digest and plasmid map can be seen in Figure 2.2.14 and all bands revealed were of the expected size. Once the construction of pJLM501 was confirmed it was possible to try the co-transformation and expression and repeat the solid phase assay of the racemases.

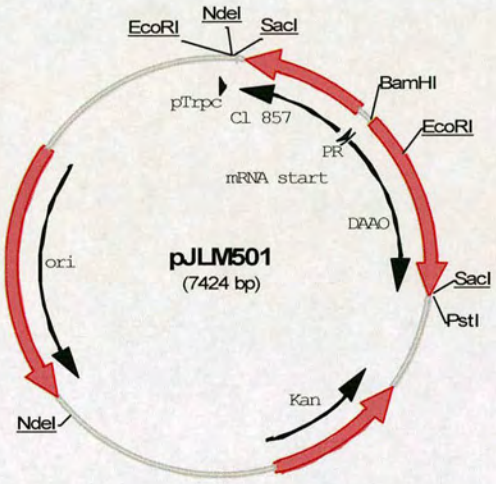
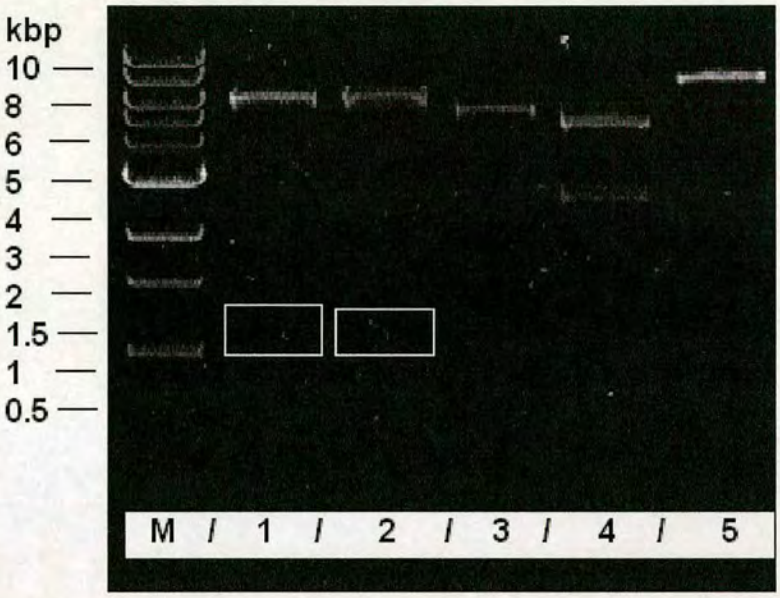


Figure 2.2.14 a) Plasmid map displaying restriction sites, double sites are underlined



b) An agarose gel of the diagnostic restriction digest. The expected bands are as follows: Lane M= 1kb Marker; Lane 1= double digest using BamHI/PstI showing expected bands at 1.1 kbp and 6.3 kbp; Lane 2= digest using EcoRI only showing expected bands at 1.2 kbp and 6.2 kbp; Lane 3= digest using SacI only showing expected bands at 2.0 kbp and 5.4 kbp; Lane 4= digest with NdeI only showing expected bands at 2.8 kbp and 4.6 kbp; Lane 7= digest with BamHI only showing expected band at 7.4 kbp corresponding to linearised vector.

Figure 2.2.14: Diagnostic digest of pJLM501

2.2.7.2 Using the two-plasmid screen to detect racemase activity

The plasmids containing the racemase and oxidase genes were used to transform *E. coli* Top10 as described in Section 4.4.4 and plated onto membranes laid on LB/agar plate containing kanamycin and ampicillin to select for cells containing both plasmids. The plates were incubated overnight at 30°C, after which the membranes were placed onto fresh plates containing IPTG (1 mM) and incubated at 37°C for a further 18 h to induce racemase and oxidase expression. The membranes were then assayed as described in Section 4.6.2.2 against L-alanine and the results can be seen in Figure 2.2.15. The top row of the diagram in Figure 2.2.15 shows a positive control with D-Ala as substrate to confirm expressed and active oxidase and all as expected, showed colour. The second

row shows all constructs with no substrate added, no colour is expected, or seen. The third row shows all with L-Ala added. The three with racemase are dark brown and the negative control with pTTQ18 in place of the vectors containing the racemase genes is a light brown in colour. A picture of the final plate (*B. subtilis* racemase) is magnified to display the colonies more clearly.

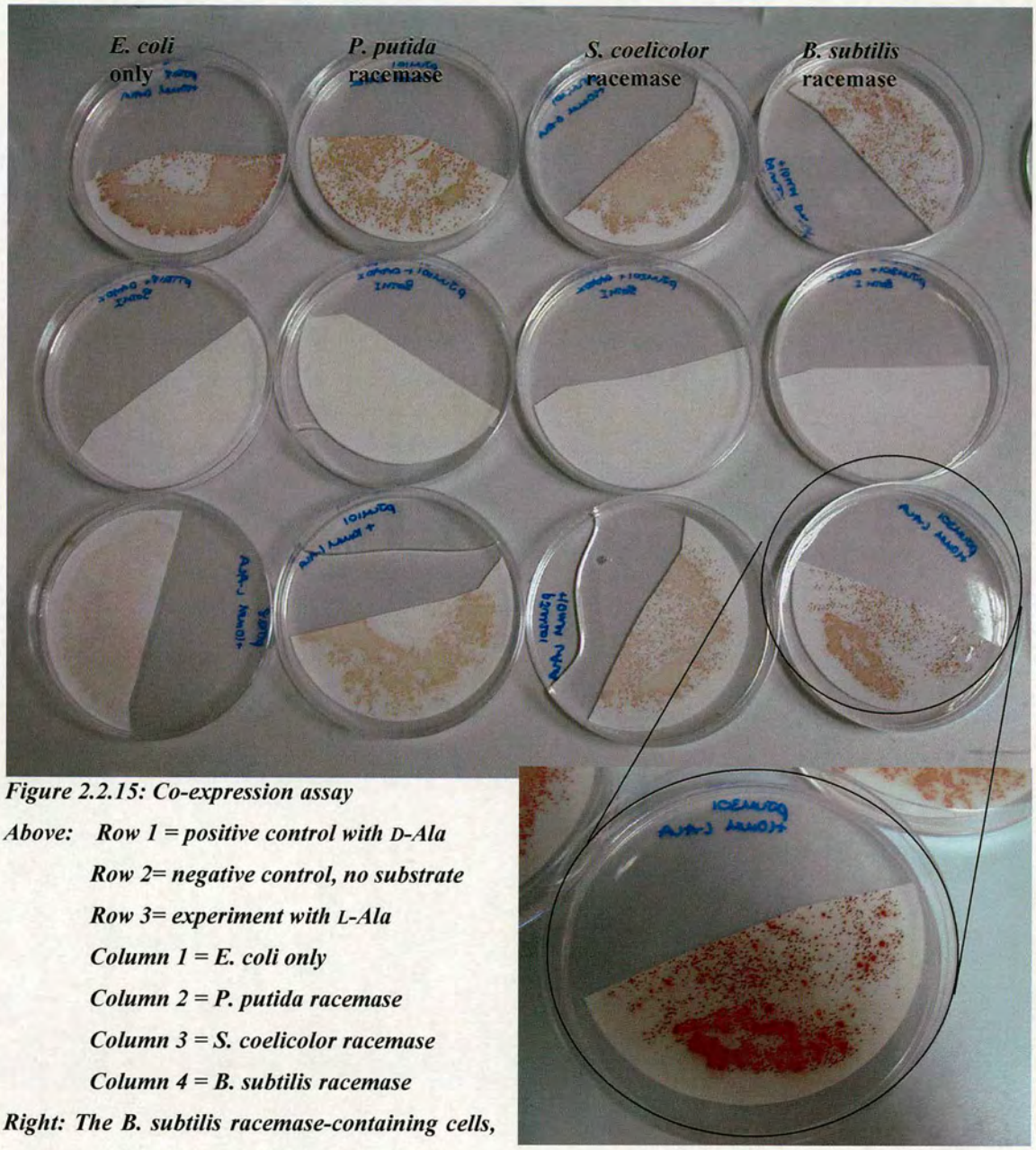


Figure 2.2.15: Co-expression assay

Above: ***Row 1 = positive control with D-Ala***

Row 2= negative control, no substrate

Row 3= experiment with L-Ala

Column 1 = E. coli only

Column 2 = P. putida racemase

Column 3 = S. coelicolor racemase

Column 4 = B. subtilis racemase

Right: The B. subtilis racemase-containing cells, assayed with L-Ala, magnified

The modified screening procedure now allowed individual colonies expressing racemase to be easily identified. The positive control showed that the D-amino acid oxidase was expressed and active and rapidly oxidised 10 mM D-Ala (the maximum possible concentration of D-Ala present in the racemisation experiments was 5 mM, and would occur only if the 10 mM L-Ala was quantitatively racemised). Therefore the concentration of expressed oxidase was sufficient to rapidly oxidise any D-Ala formed. The negative control showed no colour change in the absence of substrate, ruling out other hydrogen peroxide-forming reactions. In the presence of L-alanine, the colour appeared gradually, (the magnified picture, Figure 2.2.15, was taken an hour after the picture with all the plates).

2.2.8 Conclusions

The solution phase assay was developed and used for the detection of alanine racemase activity, initially with commercially available enzymes and then with cells expressing racemase. HPLC analysis was used to verify racemisation independently of the oxidase and confirmed that the expressed racemases were producing a racemate from a single enantiomer. The two-plasmid screen was shown to be successful in detecting alanine racemase activity in specific colonies. The oxidase gene was present on a low copy plasmid and so there should be less of the oxidase enzyme present in the cell, compared to the racemase. The screening revealed that this was not a problem, as the positive control with 10 mM D-Ala turned dark brown almost immediately. The colonies expressing racemase turned brown more slowly demonstrating that the oxidase was not rate limiting.

It was noted that the *E. coli* cells containing pTTQ18 (without racemase gene) and oxidase gene also turned brown, but more slowly than the cells heterologously expressing racemase. The background activity is likely due to the endogenous alanine racemases in *E. coli*. In all assays it was necessary to have a negative control to compare the background *E. coli* racemase activity to the expressed racemases under investigation.

2.2.9 Future work

The two-plasmid system has many possible applications as a high-throughput screen. The broad substrate specificity of the *T. variabilis* oxidase, investigated in Section 2.3, demonstrates that this oxidase could be used to detect a range of racemase substrates.

In addition, there are many other oxidases that could be used in place of the *T. variabilis* amino acid oxidase. For example, there are enantiospecific amine oxidases such as the 'D5' mutant developed by Turner and coworkers that is specific for a range of amines such as (*S*)-α-methyl benzylamine.[10] It would be possible to probe genomic libraries for transaminase activity that could convert alanine and acetophenone into α-methyl benzylamine by transformation of the library into competent cells containing the amine oxidase. There are other oxidases, such as alcohol oxidases and a glucose oxidase that could be employed here. Using the current system, the amino acid oxidase can be replaced with another oxidase by simple cloning into the *Bam*HI/*Pst*I sites of pJLM501. Any libraries or enzymes of interest must be on an antibiotic resistant (other than kanamycin) and IPTG induced plasmid.

The advantage of the oxidase on a second plasmid is that it is independent of the enzyme that is being screened. The oxidase can be easily introduced and just as easily removed.

2. Results and Discussion

2.3 Analysis of the behaviour of the amino acid oxidases used

In previous chapters, all assays involving the D-amino acid oxidase have been focused on the activity of the amino acid racemases, with only a brief investigation into the maximum oxidation rates (mOD/min) possible using different concentrations of D-alanine (Section 2.2.2). A more in depth study was needed to understand the behaviour of the oxidases used throughout this work, in order to identify where reaction rates are limited by the racemase and where they are limited by the oxidase. It is important to know the substrate specificity of the oxidase to be able to choose substrates that can be used to probe racemase activity. In a reported investigation into the specificity, inhibition and reaction sequence of a D-amino acid oxidase from pig,[61] it was noted that some L-amino acids were competitive inhibitors. Inhibition by L-amino acids would present major problems for the racemase screening as the L-enantiomer is initially supplied.

In this chapter the three oxidases used throughout the screening process were investigated. First, the substrate specificity of the D-amino acid oxidase from *T. variabilis* is presented, as this was the enzyme used for the two plasmid screen.

There were two candidate oxidases available for the solution phase screen, the commercially available oxidase from porcine kidney and a partially purified *T. variabilis* oxidase supplied by Ingenza. Analysis of the behaviour of these two oxidases was carried out to determine which was the most appropriate oxidase for the screen, that is, which was least affected by the presence of L-amino acids. The amount of oxidase required, so not to be rate limiting at a variety of substrate concentrations, was also explored. Finally, the substrate specificity of the pig kidney D-amino acid oxidase (pkDAAO) was examined.

2.3.1 Substrate specificity of the *T. variabilis* amino acid oxidase (tvDAAO)

The substrate specificity of the oxidase was determined against a panel of substrates using CFE of over-expressed oxidase, generated using *E. coli* as the host organism. The substrate specificity profile is shown in Figure 2.3.1. It can be seen that a range of substrates are accepted by this oxidase and so a range of substrates can be used to examine racemase activity.

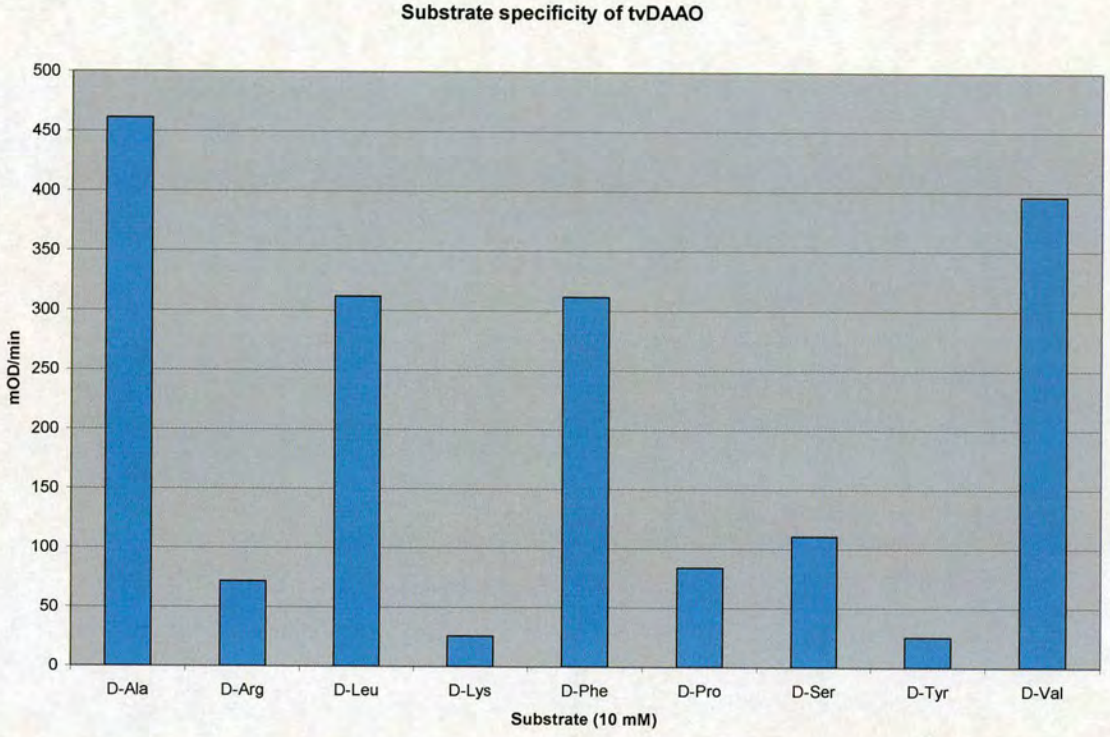


Figure 2.3.1: Substrate specificity of the tvDAAO

2.3.2 Behaviour of the tvDAAO in the presence of D- and L-Arg

An ammonium sulfate precipitate of the CFE containing the partially purified tvDAAO, over-expressed in *E. coli,* was supplied for use in the solution phase screen. L-Arg was the substrate chosen for racemase evolution (as this was a known substrate for the broad substrate specificity racemase, but not a substrate of the other two) and so the behaviour of the oxidase towards Arg was investigated. Two experiments were carried out: The first was to examine the ability of the oxidase to detect one concentration of D-Arg in the

presence of increasing concentrations of L-Arg. The second mimicked the changing Arg concentrations in the assay, where the ratio of D:L-Arg was varied from 0:10 to 10:0, with the total substrate D/L-Arg concentration maintained at 10 mM

2.3.2.1 Inhibition studies of L-Arg

The activity of the tvDAAO towards 5 mM D-Arg was measured in the presence of different amounts of L-Arg from (0 to 10 mM). The supplied ammonium sulfate tvDAAO pellet (100 mg) was resuspended in 10 ml water and 20 µl was used for the reaction, (assay conditions Section 4.6.1.4). The reaction was followed by formation of the coloured product, detected at 510 nm and the results can be seen in Figure 2.3.2. The rate values (mOD/min) for the inhibition reaction can be seen in Figure 2.3.3. A blank control was measured using water instead of substrate. The time delay in adding the reagents and measuring the OD accounts for the OD readings at t=0.

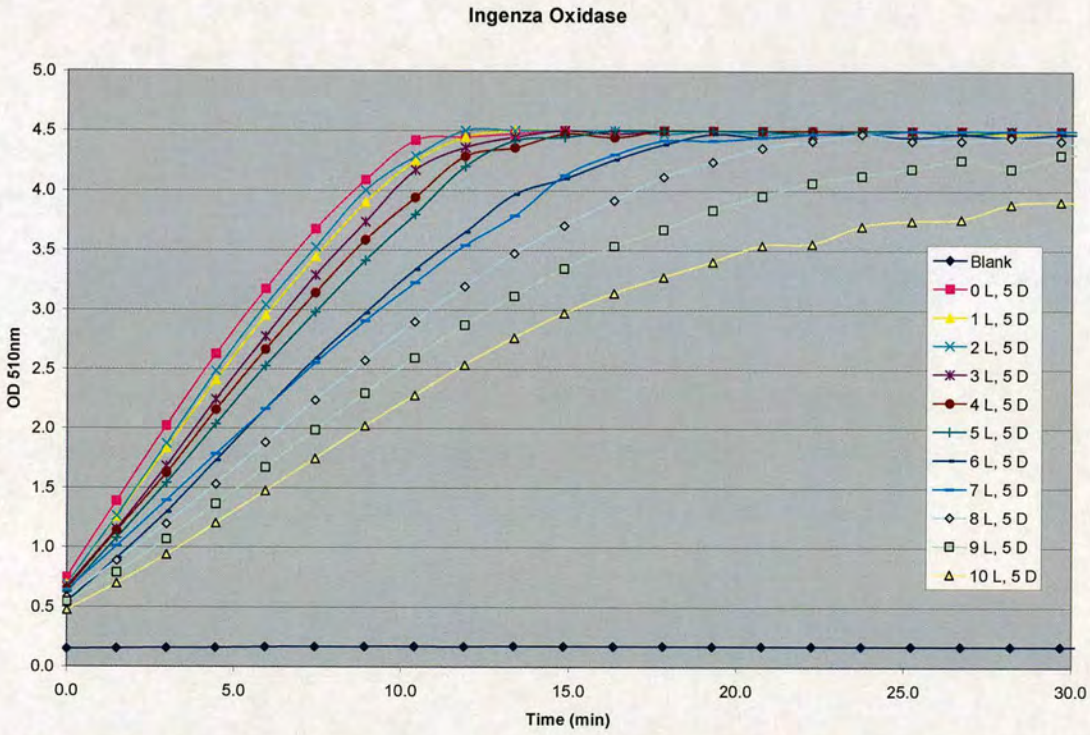


Figure 2.3.2: The oxidation of 5 mM D-Arg by the tvDAAO in the presence of increasing amounts of L-Arg

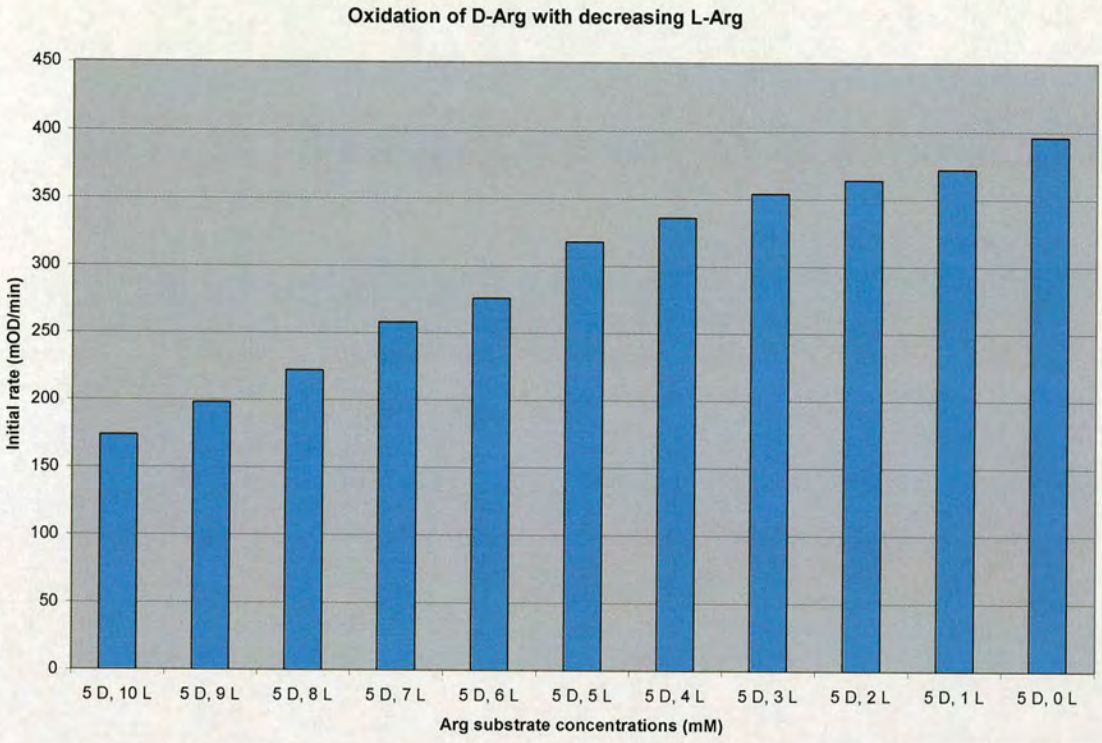


Figure 2.3.3: The increasing initial rates of oxidation of 5 mM D-Arg, in the presence of decreasing amounts of L-Arg

This analysis reveals that the tvDAAO is inhibited by increasing amounts of L-Arg. The rate of oxidation of a constant concentration of D-Arg increases as the amount of L-Arg decreases. When there is no L-Arg present, the initial rate of oxidation is almost 400 mOD/min, compared with 175 mOD/min in the presence of 10 mM L-Arg. The rate of oxidation is more than halved by the introduction of 10 mM L-Arg.

2.3.2.2 The ability to detect D-Arg in the presence of L-Arg

The initial rate of oxidation of D-Arg at a range of concentrations (0 and 10 mM) and then in the presence of L-Arg, was measured. For example, the initial rate of oxidation of 3 mM D-Arg was measured, followed by the initial rate of oxidation of 3 mM D-Arg in the presence of 7 mM L- Arg. The total concentration of D-Arg and L- Arg was equal to

10 mM, as this was the concentration used for assaying the racemases. The comparison of the rates of oxidation can be seen in Figure 2.3.4.

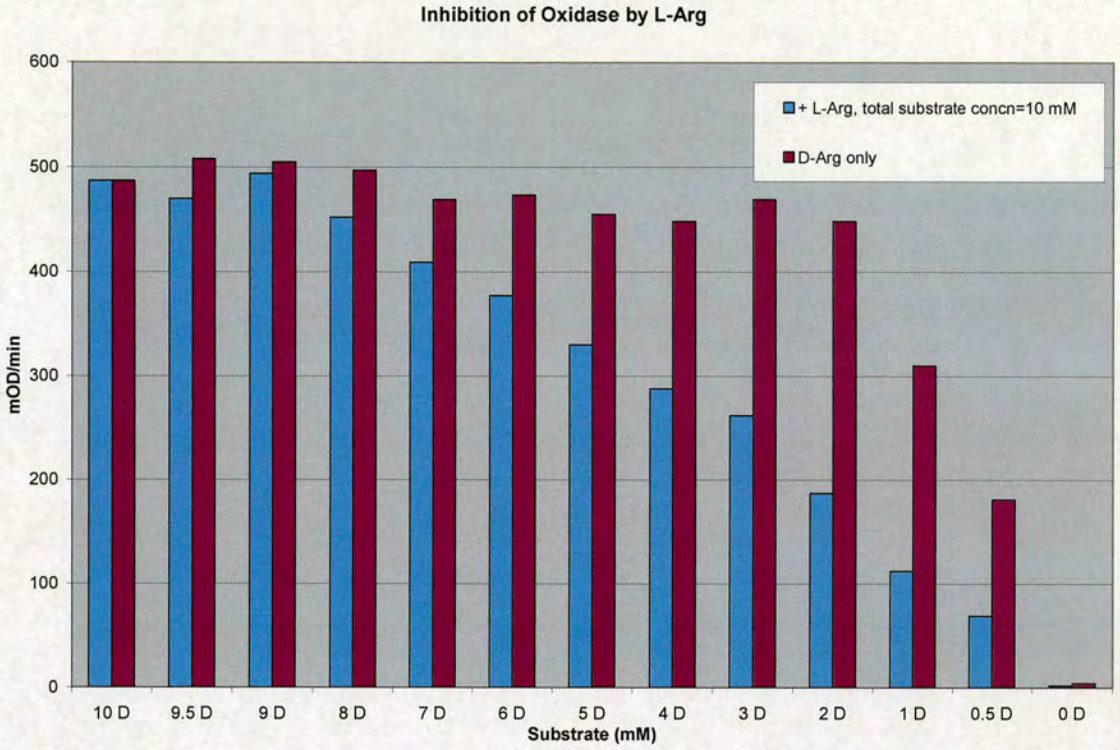


Figure 2.3.4: Initial rates of oxidation of D-Arg by tvDAAO and in the presence of L-Arg

It is clearly shown that the oxidation of D-Arg is affected by L-Arg. For the screen to be successful, the oxidation of D-Arg cannot be rate limiting otherwise the racemase rate cannot be determined. For example, the racemase screening is carried out with 10 mM L-Arg and if the racemase converts 0.5 mM of the L- to D-Arg, then the rate that should be measured is 181 mOD/min, (detection of 0.5 mM D-Arg). However due to the presence of the remaining L-Arg, the rate measured is only 70 mOD/min, (Figure 2.3.4, measurements at 0.5 D). The tvDAAO is therefore not an ideal candidate to identify racemase activity as the rate of oxidation is variable in the presence of L-Arg.

2.3.3 Behaviour of the pkDAAO in the presence of D- and L-Arg

The oxidase used in the initial stages of screen development (Section 2.2) was the commercially available amino acid oxidase from porcine kidney. To analyse the behaviour of this oxidase in the presence of L- and D-Arg, the same experiments used to investigate the tvDAAO were carried out.

2.3.3.1 Inhibition studies of L-Arg

The rate of oxidation of 5 mM D-Arg using the pkDAAO was measured in the presence of different amounts of L-Arg from (0 to 10 mM), (assay conditions Section 4.6.1.4, using 0.2 U D-amino acid oxidase). The reaction was followed by formation of the coloured product, detected at 510 nm, the results of which can be seen in Figure 2.3.5. A blank control was also measured without any substrate. The rate values (mOD/min) for the inhibition reaction can be seen in Figure 2.3.6.

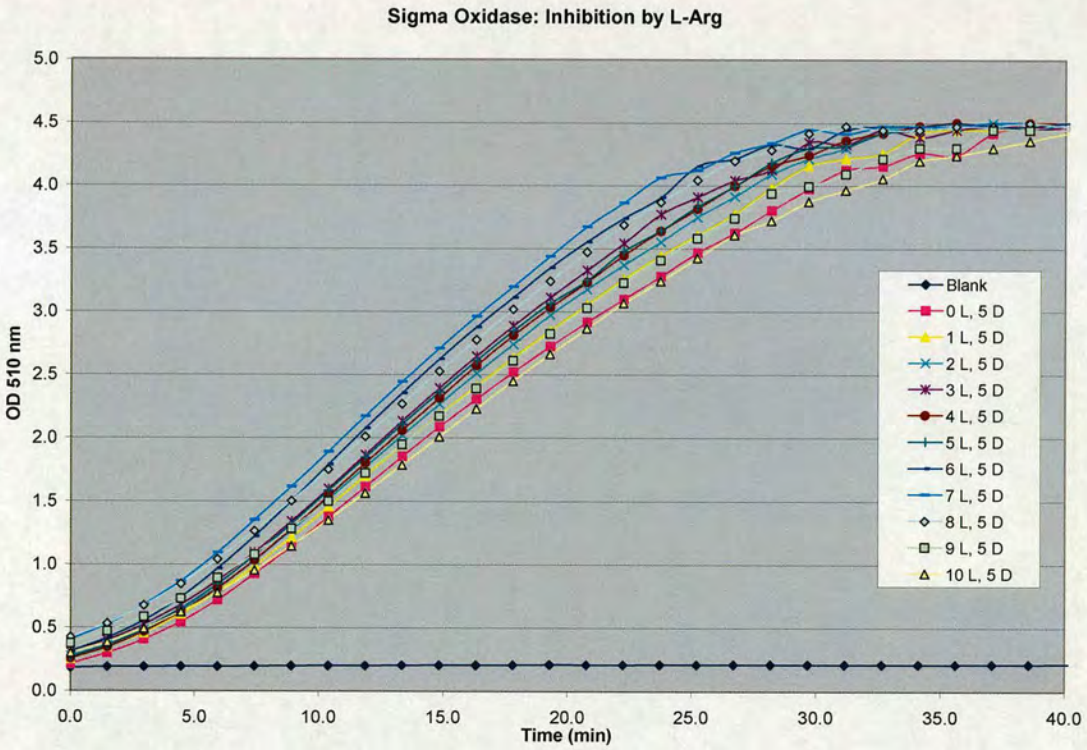


Figure 2.3.5: The rate of oxidation of 5 mM D-Arg in the presence of increasing amounts of L-Arg

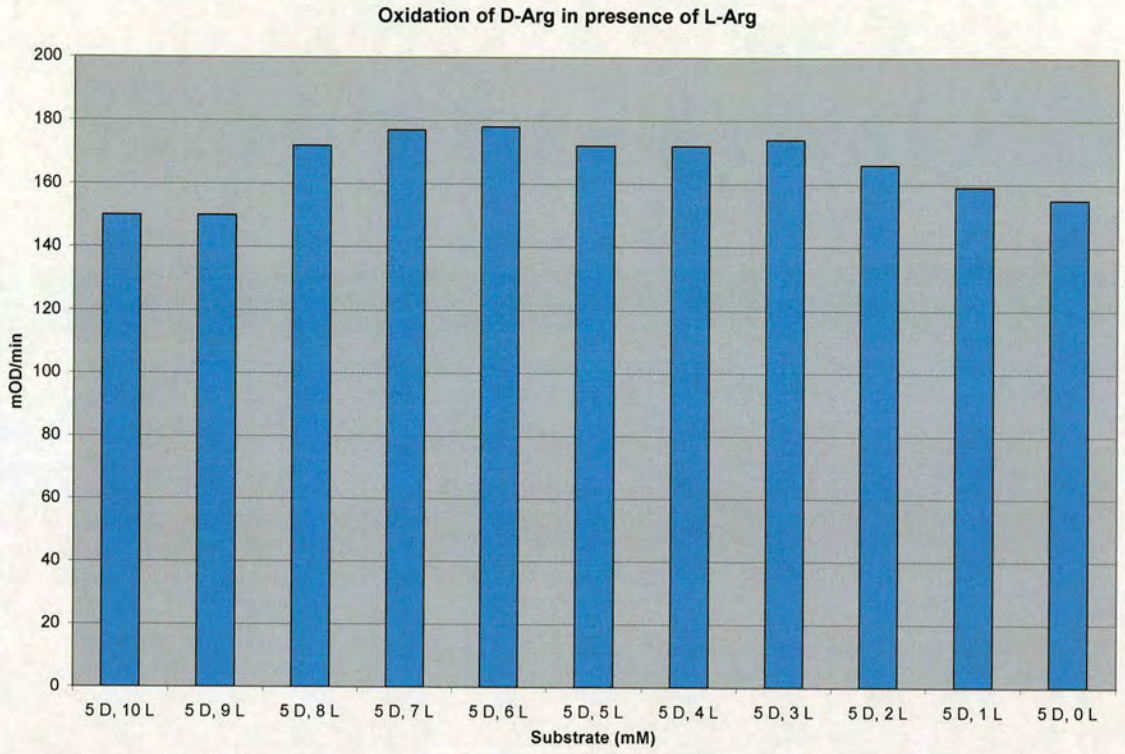


Figure 2.3.6: The initial rates of oxidation of 5 mM D-Arg in the presence of L-Arg

The oxidation carried out in the presence of various concentrations of L-Arg revealed that this oxidase is not greatly affected by fluctuating concentrations of L-Arg. The rates of oxidation of 5 mM D-Arg do not vary as much (between 150 and 178 mOD/min), compared with the tvDAAO rates (181 to 400 mOD/min).

2.3.3.2 The ability to detect D-Arg in the presence of L-Arg

It was decided to also analyse the rates of oxidation of D-Arg at a range of concentrations (0 and 10 mM) and then in the presence of L-Arg, (as for the tvDAAO experiments). For example, the initial rate of oxidation of 3 mM D-Arg was measured, followed by the initial rate of oxidation of 3 mM D- Arg in the presence of 7 mM L- Arg. The total concentration of D- Arg and L- Arg was equal to 10 mM, as this was the concentration used for assaying the racemases. The comparison of the rates of oxidation can be seen in Figure 2.3.7.

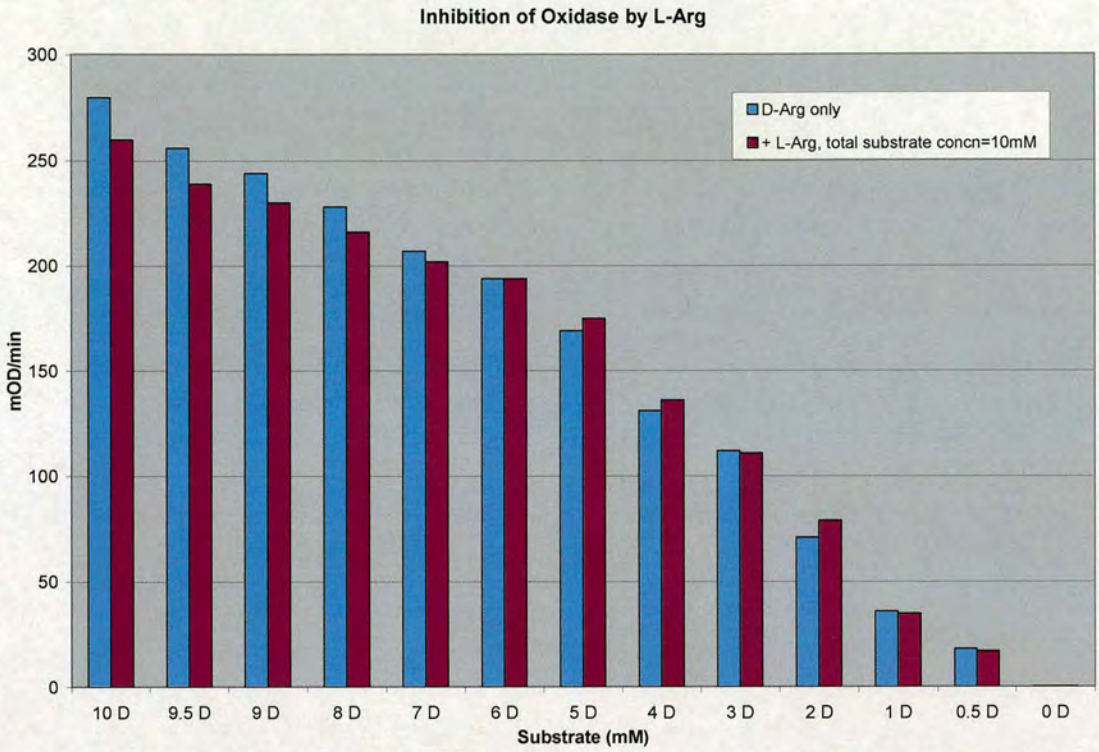


Figure 2.3.7: Initial rates of oxidation of D-Arg by tvDAAO and in the presence of L-Arg

The data for the pkDAAO suggests that it is not inhibited by L-Arg. The values for oxidation rate of each concentration of D-Arg do not vary too much from the values obtained with L-Arg present.

2.3.3.3 Determining the amount of pkDAAO required for maximum rates of oxidation of D-Arg

Although compared to the data for the tvDAAO the oxidation of the pkDAAO is not affected by the presence of L-Arg, the final rates obtained for pkDAAO oxidation at each of the D-Arg concentrations are much lower than those obtained with the tvDAAO and maybe implied that not enough of the pkDAAO was used. For example at 10 mM D-Arg (no L-) the rate obtained with the tvDAAO was almost 500 mOD/min, (Figure 2.3.4) and with the pkDAAO it was 280 mOD/min (Figure 2.3.7). An investigation was carried out at different concentrations of L-Arg (12.5, 25 and 50 mM), using increasing

concentrations of pkDAAO (0, 0.05, 0.1, 0.2, 0.4, 0.7, 0.9 U) to determine the concentration of oxidase required for high rates of D-Arg oxidation. The rates that were obtained are shown in Figure 2.3.8.

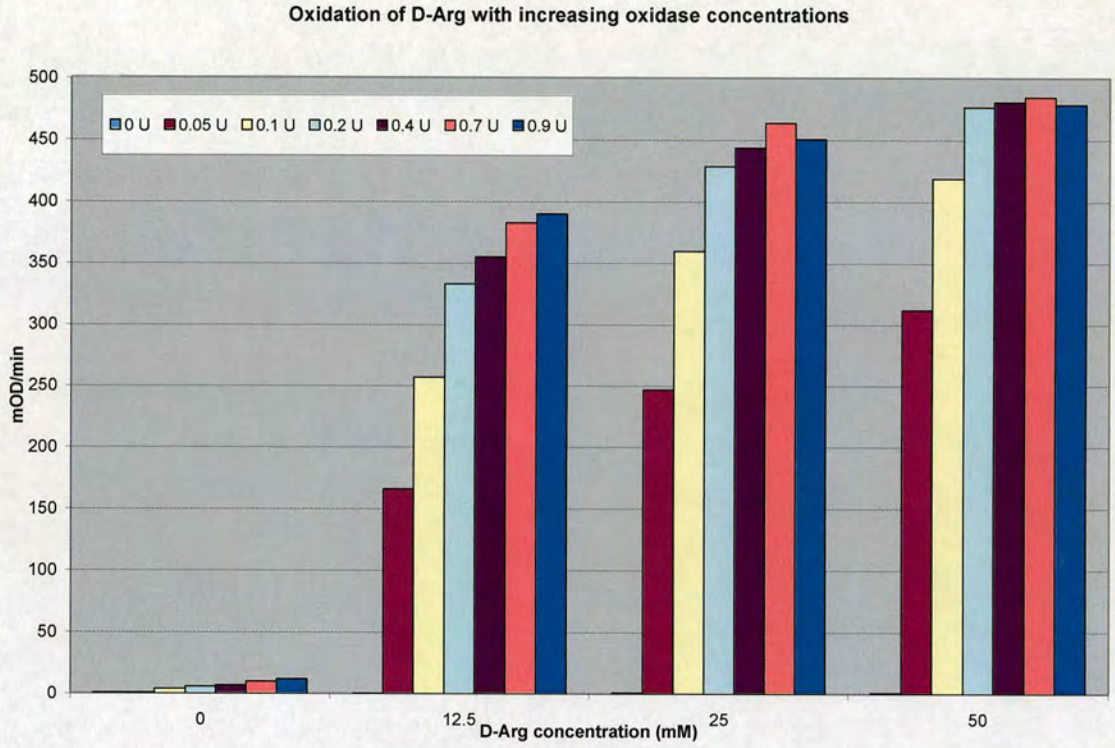


Figure 2.3.8: The oxidation of D-Arg using increasing pkDAAO enzyme units

At 12.5 mM D-Arg, the experiment revealed that the difference in rate using 0.2 U compared to 0.4 U of D-amino acid oxidase is 20 mOD/min. The difference between the rates when using 0.2 and 0.1 U oxidase is much greater, (76 mOD/min). At 25 mM D-Arg, 0.2 U of oxidase produced a rate very similar to 0.4 U, 429 and 444 mOD/min respectively, but was much more different than for 0.1 U (360 mOD/min). At 50 mM D-Arg, the same was observed, using 0.2 U oxidase was much faster than when using 0.1 U, but not very different that when using 0.4 U. It was decided to use 0.4 U of oxidase for experiments containing 10 mM substrate as the rate of oxidation was starting to plateau and using 0.7 U would almost double the price of the oxidation reaction.

2.3.3.4 Substrate specificity of the pkDAAO

In order to use the pkDAAO to screen for racemase activity, it was important to understand the substrate specificity of the oxidase. Only an amino acid that is accepted by the oxidase can be used to screen the racemase. A range of D-amino acids were analysed using the liquid phase assay (Section 4.6.1.4). 0.4 U of oxidase was used for each substrate. Each reaction was run in triplicate, with a negative control containing no oxidase and the data was averaged. The final concentration of all substrates was 10 mM. Where the D-amino acid was only available as a racemate, a total concentration of 20 mM was added, to provide 10 mM of the D-enantiomer. The substrate specificity is shown in Figure 2.3.9.

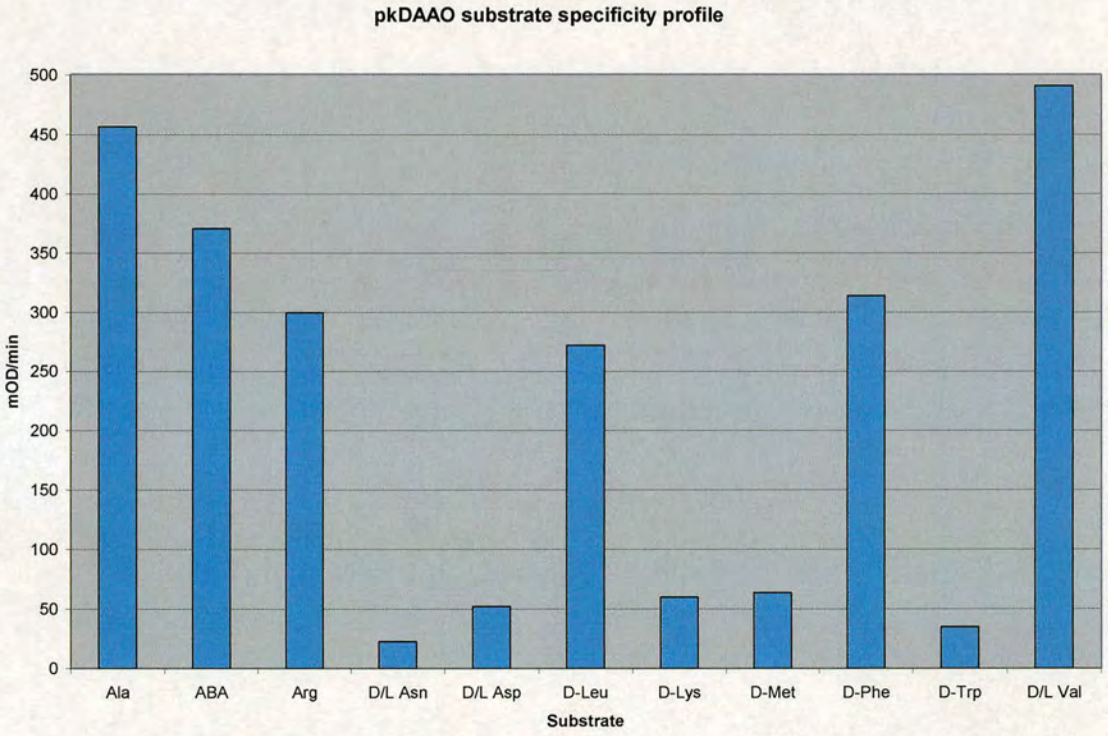


Figure 2.3.9: Substrate specificity of the pkDAAO

The substrate specificity revealed that the oxidase has a wide substrate specificity, accepting Ala, 2-ABA, Arg, Leu, Phe and Val, with lower activity towards Asn, Asp, Lys, Met and Trp.

2.3.4 Conclusions

For the purposes of screening in the solution phase, it has been shown that the pkDAAO is preferable to the tvDAAO, as the rate of oxidation using the former is not affected by presence of L-Arg. For the screen to be useful and provide accurate rates of racemisation, the rate of oxidation must not be rate limiting. When developing the screen for oxidase activity, the amount of peroxidase was designed to be in large excess to provide accurate oxidation rates and in the same way, the screen for racemase activity must not be limited by the oxidation rate.[27] The oxidase from pkDAAO has been shown to not be greatly affected by L-Arg and so is the better candidate for screening in the solution phase.

The substrate specificity of the pkDAAO was then recorded using a panel of substrates. It was shown that all substrates that were screened later in this work (Ala, 2-ABA and Arg) are accepted by this oxidase.

2. Results and Discussion

2.4 First attempt at mutation and screening of the racemase genes

Random mutation of the racemase genes

The solid and solution phase screens have been shown to be successful at identifying amino acid racemase activity. However, the amino acid racemases expressed and tested on the solid phase were all wild type and had identical activities. The next stage in developing a novel screening method was to show that the screen could be used to detect activity against a background of inactivity. It was decided to evolve the racemases to accept new substrates and use the screen to isolate and examine possible novel enzymes.

Examining the substrate specificities of the three cloned racemases (Section 2.2.3) it can be seen that there are a number of substrates that are accepted by the broad substrate specificity racemase, but not by the other racemases. It is therefore possible to use PLP-dependent enzymes to racemise substrates other than alanine and 2-amino butyric acid and the aim here was to evolve the narrow substrate specificity racemases towards one or more of these substrates. There are three substrates that are accepted by the broad substrate specificity racemase but not by the other racemases: Arg, Leu and Lys. The substrate specificity of the tvDAAO (from the two-plasmid screen) is shown in Section 2.3.1 and shows that the oxidase accepts arginine, leucine and lysine. It was decided to create a variant library of enzymes and screen for activity towards Arg, Leu and Lys.

This chapter presents the unsuccessful attempts to generate libraries of the racemase genes using the XL1-Red method. The first attempt revealed the presence of small quantities of the plasmid encoding the broad substrate specificity racemase in the plasmid preparations of the other racemases. Despite this problem, the experiment still represented the first example of the selection of racemase activity from a small

proportion of a mixed racemase population and was the first validation of the screening methodology. The second and third attempts produced very low mutation frequencies and revealed a high proportion of the wild type sequence.

2.4.1 Creation and analysis of variant library one

2.4.1.1 Library synthesis and screening

The plasmids containing the narrow substrate specificity racemase genes were used to transform XL1-Red cells (Section 4.4.3, Method 1). The library of variants was screened using the two-plasmid solid phase assay (Sections 4.4.4 and 4.6.2.2). For each racemase, ten plates, each covered with 2-3000 colonies, were assayed using L-Arg as probe substrate. Eight hits were detected from the plates expressing the *S. coelicolor* racemase library and four hits were seen on the plates expressing the *B. subtilis dal* racemase library. The possible variants were re-assayed to confirm activity towards arginine and the substrate specificities determined.

2.4.1.2 Variant enzyme results

The variant genes were sequenced and it was found that the variants were not novel enzymes, but the broad substrate specificity racemase from *P. putida.* The results indicated that a small amount of pJLM101 DNA was present in the pJLM201 and pJLM301 stocks.

2.4.1.3 Conclusions

The two-plasmid solid phase screen was used to detect contamination of the narrow substrate specificity racemases with a small amount of the broad substrate specificity racemase. Approximately 25 000 colonies were screened for each racemase. Eight colonies were selected from the *S. coelicolor* library due to their positive effect on the solid phase screen and were shown to be the *P. putida* racemase. Four hits were found in the *B. subtilis dal* library. The screen has been shown to be very powerful as it detected the *P. putida* racemase when only present in 0.016 % and 0.032 % of the colonies (for

the *B. subtilis dal* and the *S. coelicolor* racemases respectively). The stocks of the racemases were discarded as a result of the screen and new stocks prepared.

2.4.2 Creation and analysis of variant libraries two and three

2.4.2.1 Library synthesis and screening

Libraries of the racemase genes were prepared using the same method as for library 1. The mutation rate was found by sequencing to be very low and so a third method of mutator strain growth was attempted as detailed in Section 4.4.3, Method 2.

2.4.2.2 Analysis of the variant libraries

Ten colonies from each racemase gene library (*S. coelicolor* and *B. subtilis*) were grown, the DNA extracted and the gene sequenced. From the ten colonies in the *S. coelicolor* library, three were seen to have one base mutation in the gene, the other seven were wild type racemase. For the ten colonies of the gene library from *B. subtilis*, only two were shown to have one mutation in the whole gene and the remainder were wild type.

2.4.3 Conclusions

The two-plasmid screen has been used to detect activity against a background of inactivity, by detecting the broad substrate specificity racemase when mixed in with the other racemases at very low concentrations. The experiment is analogous to an experiment where low concentrations of an enzyme with known activity are purposely doped with inactive enzymes and the different activities of the enzymes detected.

The generation of variant racemase genes was not successful using the XL1-Red mutator method. The mutation rate was too low to produce hit enzymes. It was shown that the mutation rate for the *S. coelicolor* racemase gene was 3 per 12 000 bp (1 per 4 000 bp). It was decided to use another method to introduce random mutations in the racemase genes. Error-prone PCR was chosen as only the gene of interest is mutated and the mutation frequency can be increased depending on the conditions used.

2. Results and Discussion

2.5 Successful attempt at evolving the racemase genes

Random mutation of the racemase genes using error-prone PCR

Error-prone PCR (epPCR) was the second method chosen to randomly mutate the racemase genes as using the mutator strain did not produce good quality libraries for screening for novel racemase activities. An introduction to epPCR and the principles behind the technique are presented in Section 1.3.1.

Following variant racemase synthesis using epPCR, the racemase libraries were screened for novel racemase activity using Arg, Leu and Lys, and a range of variants were detected. The initial analysis of the variants and the attempted re-cloning to place the genes into a plasmid designed for enzyme purification is presented. The mutations were analysed using a 3D model, generated from the known structure of the *S. lavendulae* alanine racemase. The experiments to remove a mutation proposed to be important for the novel activity is shown to enable further evaluation of that amino acid role.

2.5.1 The synthesis of a variant library of the *S. coelicolor* racemase gene, attempt 1

The Diversify^{TM} PCR random mutagenesis kit from BD Biosciences was used to create racemase gene libraries according to Section 4.2.15. Three reactions were carried out simultaneously to introduce an average of 2.7, 4.8 and 8.1 mutations respectively per kbp. The epPCR products were prepared for integration into the plasmid pJLM201 (containing the *S. coelicolor* racemase gene) using whole plasmid PCR, by restriction digestion and removal of single strand DNA termini using the Klenow fragment of DNA polymerase I. The procedure ensured that both strands were fully complementary to each other and the target vector (except for the mutated bases), a prerequisite for successful

whole plasmid PCR. After removal of the template DNA with *Dpn*I, the resulting plasmids were used to transform *E. coli* XL1-Blue cells. After cell growth, the plasmid DNA was isolated to obtain the libraries of racemase genes.

2.5.2 Analysis of the variant library of the *S. coelicolor* racemase gene

To validate the library of racemase genes, a number of colonies were obtained, the plasmid DNA isolated and sequenced. A large number of wild type racemase sequences were obtained (90 %). After discussion, it was proposed that the wild type plasmid DNA was not efficiently removed after the whole plasmid PCR. The first possible explanation was that the template DNA was synthesised using *E. coli* Top10 cells, which may not efficiently produce methylated DNA. The cells are mrr, hsdRMS and mcrBC deficient, which means they lack the enzymes to digest non-methylated DNA, denoted in the genotype Δ(*mrr-hsdRMS-mcrBC*). Therefore there is no selective pressure on the cells to methylate any synthesised DNA. The second explanation was that the template DNA used was not freshly prepared; it had been stored for a number of weeks and it was proposed that the DNA is gradually demethylated upon storage. The combination of these two factors would result in a very low concentration of methylated DNA and subsequent digestion with *Dpn*I would not remove all the template plasmid. It was decided to repeat the epPCR using template DNA that had been freshly prepared using BL21(DE3).

2.5.3 The synthesis of a variant library of the *S. coelicolor* racemase gene, attempt 2

Using the Diversify™ PCR random mutagenesis kit, one reaction was carried out according to Section 4.2.15. to introduce approximately 4.8 mutations per kbp. The epPCR products were blunt ended and then used as primers in a whole plasmid PCR of pJLM201 (plasmid containing wild type *S. coelicolor* racemase gene). After removal of the template DNA, the resulting plasmids were used to transform *E. coli* XL1-Blue cells. After cell growth, the plasmid DNA was isolated to obtain the library of racemase genes.

2.5.4 Analysis of a variant library of the *S. coelicolor* racemase genes

To validate the library of racemase genes, nine colonies were obtained, the plasmid DNA isolated and sequenced. Four out of the nine genes contained mutations, insertions and/or deletions, therefore 55 % of the library was wild type racemase. The alignment of the gene sequences is shown in Appendix 5. A summary of the mutations of the nine randomly chosen colonies is presented in Table 2.5.1.

Colony	Base Mutations	Amino Acid Mutations
A	A46G	Asp13Gly
	C510T	Arg168Trp
	C578T	Pro unchanged
	A1126T	Thr373Ser
	C1176 deleted	Frame shift
B		No changes
C	G238A	Ala74Thr
	A269T	Pro unchanged
	T insert 619	Frame shift
	T662C	No change
D		No changes
E	G65A	Arg unchanged
	T372A	Trp122Arg
	G447T	Gly147Cys
	C711T	His unchanged
	T839A	Val227Asp
	C894T	Thr unchanged
F		No changes
G	G263A	Ala unchanged
	C505T	Ala166Val
	A985T	Thr326Ser
H		Unchanged
I		Unchanged

Table 2.5.1: A summary of the base mutations from nine randomly selected colonies from the epPCR library. Effects on the amino acids are shown in column 3.

A total of sixteen base changes were found (not including the insertion or deletion), resulting in nine amino acid mutations. There were a total of ten transitions and six transversions. The average mutation rate was 4.5 per gene. The library contained a large number of mutations and therefore could be screened for novel activity.

2.5.5 Screening the *S. coelicolor* racemase gene variant library

The library was combined with the oxidase for the co-transformation of five aliquots of *E. coli* Top10 cells according to Section 4.4.4. The library was screened in the solid phase for activity against a combination of substrates, L-Arg, L-Leu and L-Lys. The following six control reactions were also carried out

Negative control:

1) WT racemase with L-Arg/Leu/Lys substrate mixture
   -to check that the WT racemase did not react with any of the substrates

Positive controls:

2) WT racemase with L-Arg/Leu/Lys substrate mixture, plus L-Ala
   -confirmed assay mixture prepared correctly
3) WT racemase with L-Arg/Leu/Lys substrate mixture, plus D-Ala
   -confirmed oxidase was working
4) WT racemase with L-Arg/Leu/Lys substrate mixture, plus D-Arg
   -checked that substrate mix did not inhibit oxidation of D-Arg
5) WT racemase with L-Arg/Leu/Lys substrate mixture, plus D-Leu
   -checked that substrate mix did not inhibit oxidation of D-Leu
6) WT racemase with L-Arg/Leu/Lys substrate mixture, plus D-Lys
   -checked that substrate mix did not inhibit oxidation of D-Lys

The variant library was screened and a total of 101 hits were detected, an individual colony colouration can be seen in Figure 2.5.1a. The hits were picked from the plates,

re-grown, plated onto membranes and re-assayed against the L-Arg/Leu/Lys substrate mixture (without separation of the racemase from the oxidase), Figure 2.5.1b.

Figure 2.5.1: Activity hits towards L-Arg.
a) Individual hit in library screen. b) Re-screen positive activities towards L-Arg.

Forty-two hits were detected by colour formation over the membranes within three hours of assay. The plasmids containing the racemase genes were isolated (the plasmid containing the oxidase gene was removed by digestion using restriction enzymes that are known not to cut the racemase plasmid, Section 4.6.2.3) and used to transform *E. coli* Top10 cells for assay in the solution phase. At the transformation stage, three hits did not grow, reducing the total to thirty-nine.

2.5.6 Solution phase screening of possible novel *S. coelicolor* racemases

The CFE from each of the thirty-nine over-expressed possible racemase variants were assayed against a panel of L-amino acids (Ala, Arg, Leu and Lys) to determine their individual substrate specificities. The reactions were very slow and so the plates containing the assay mixtures were sealed to prevent evaporation of the assay mixture and to eliminate light. The plates were then incubated at 37°C and left overnight. In the morning the wells were examined to see which had turned pink, indicating racemisation. A total of thirteen hits were chosen on their ability to produce a positive result in the solution phase assay against the individual substrates, see Table 2.5.2. The majority of the variants were active with L-Arg, only one was active with Leu and Lys.

Substrate	Variant reference
L-Arg	21, 36, 48, 56, 69, 70, 77, 81, 82, 88, 89, 93, 94
L-Leu	69
L-Lys	69

Table 2.5.2: Successful assay reaction of individual hits with individual substrates

2.5.7 Sequence analysis of variants displaying novel racemase activity

The variants listed in Table 2.5.2 were investigated further by forward and reverse sequencing of the racemase gene. All except variant 36 were successfully sequenced. Repeated attempts to sequence variant 36 failed.

The gene sequences were translated to reveal the mutated residues. An alignment of the twelve amino acid sequences is shown in Appendix 6. All twelve of the sequences were identical to eachother and contained three mutations compared to the parent racemase. The three mutations were I195T, N223D and I374N, as a result of the same base mutations. Therefore, the twelve variants were the same racemase picked twelve different times from the library. The variant that displayed activity towards all three substrates was also shown to be the triple mutant. Therefore the assay results were false for variant 69 against L-Leu and L-Lys, and a repeat of the solution phase screen of variant 69 confirmed that it did not have activity for L-Leu and L-Lys. From this point forward, these variants were treated as one and labelled triple mutant (3M).

2.5.8 Modelling of the triple mutant racemase

To better understand the effects of the mutations in 3M, a molecular model of the triple mutant was created based on the structure of a similar alanine racemase from *S. lavendulae*. An alignment of the protein sequences revealed that the *S. coelicolor* and the *S. lavendulae* have 72 % identity and 79 % similarity, Appendix 7. The software used was Accelrys Discover Studio (DS) and the PDB of the *S. lavendulae* used was 1vft. The

file showed the alanine racemase crystallised structure with cycloserine, an inhibitor bound to the PLP cofactor. The original *S. lavendulae* amino acid sequence was subjected to individual mutations using Accelrys DS to match the *S. coelicolor* wild type sequence and the global energy of the structure minimised. The racemase exists as a homodimer and contains two active sites, situated at the interface of the monomers, Figure 2.5.2. Residues from both monomers are involved in catalysis at each active site.

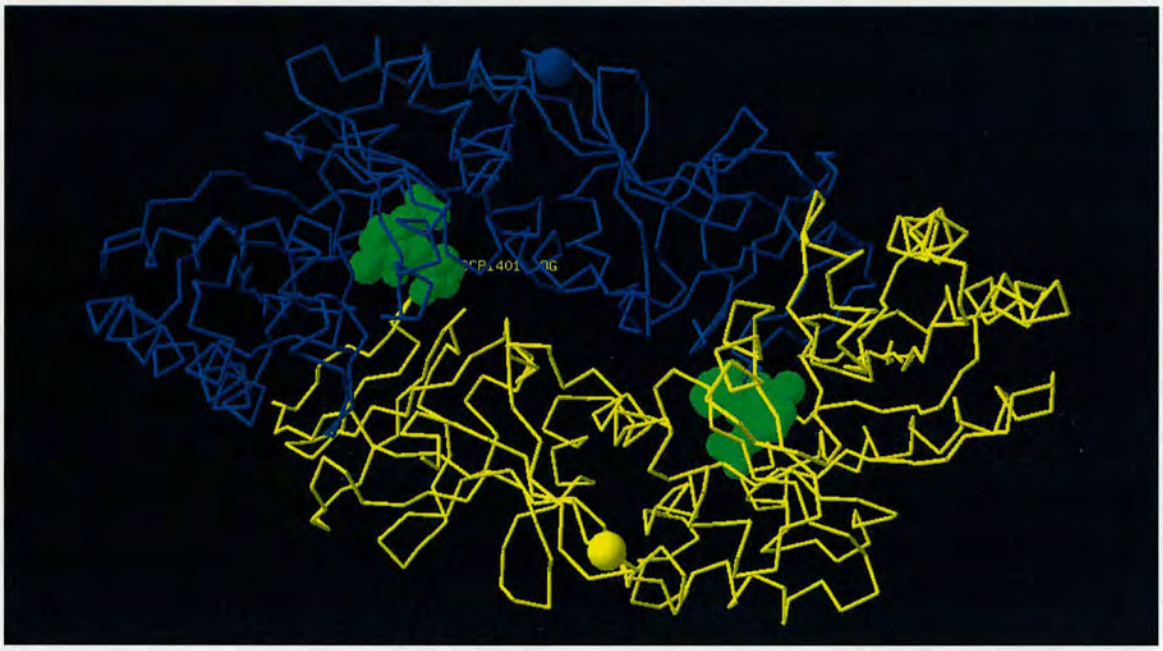


Figure 2.5.2: An α- carbon backbone of the alanine racemase dimer PDB 1vft. The individual monomers are coloured blue and yellow with balls indicating C-termini. A space fill model of the PLP cofactor bound in the active sites is coloured green

2.5.8.1 I195T

The isoleucine at position 195 was identified in the model (also Ile in the *S. lavendulae* racemase sequence) and mutated to threonine. The five residues either side of the *S. lavendulae* racemase corresponding Ile195 are PGHPS**I**RLQLD and in the 3M the sequence is PGHPS**T**AAQLT (Appendix 7). The structure was energy minimized and investigated. It was seen that the residue was located at the terminus of an α-helix approximately 15 Å from the bottom of the active site (Figure 2.5.3) and approximately

9 Å from the second N223 mutation, Figure 2.5.4. The residue was too far from the active site to have an easily rationalised effect on the racemase mechanism. Long range effects may be possible and the preceding sequence reveals that this area is almost identical to that of the *S. lavendulae* racemase. However, analysing the proceeding residues, only 2/5 are identical (LQLDA in *S. lavendulae* and AQLTR in *S. coelicolor*) implying that this region is not critical for racemase function.

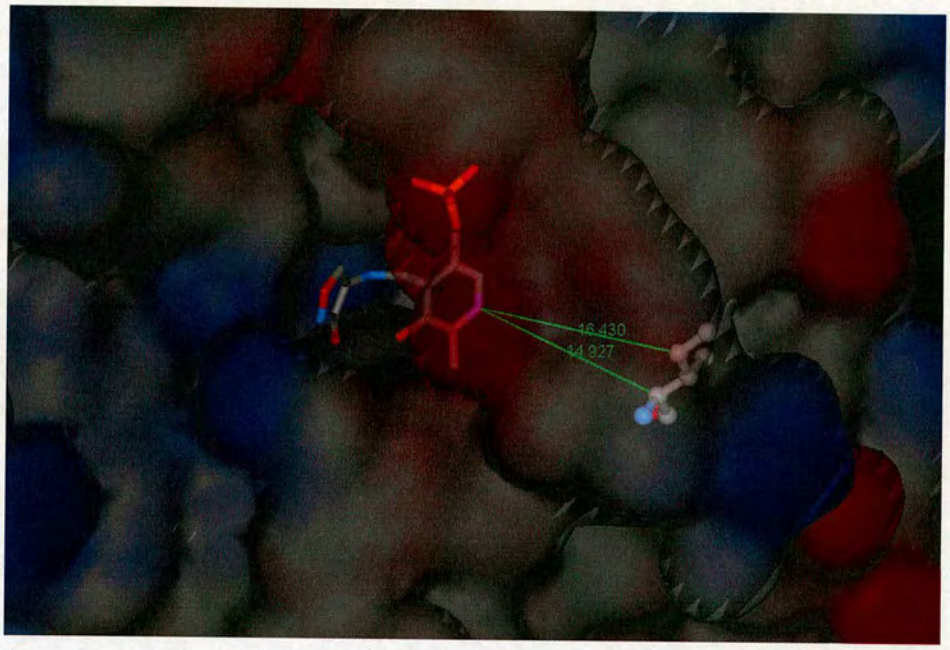


Figure 2.5.3: A snapshot of the model showing a surface map with charged regions shaded. The distance of the I195Tmutation (ball and stick) from active site is labelled. The cofactor is shown in solid line and the entrance to the active site can be seen

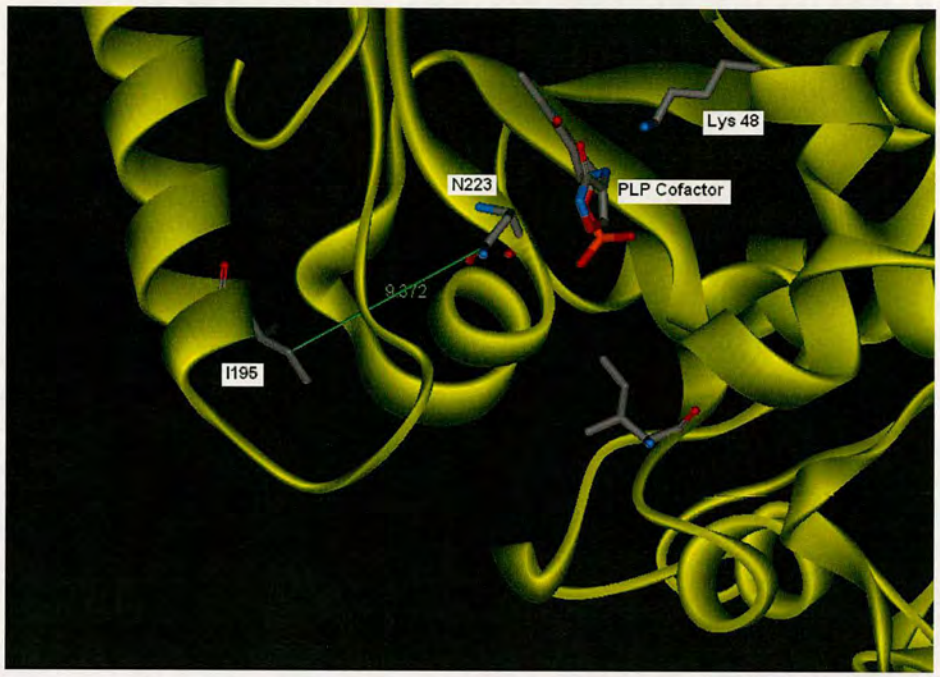


Figure 2.5.4: A carbon backbone ribbon picture of the racemase monomer. I195 is shown near the terminus of an α-helix 9.4 Å from the N223 residue found at the bottom of the active site

2.5.8.2 N223D

As shown in Figure 2.5.4, the residue 223D could be found near the bottom of the active site. Distance measurements using the modeling software showed that the distance from the arginine found directly underneath the PLP cofactor to the residue was 3 Å, Figure 2.5.5. It might be proposed that D223 could have formed a catalytic triad with neighbouring H183 and Y285 (the latter is the base that abstracts a proton from the amino acid during racemisation). However, the stereochemistry required for a catalytic triad is not present since it would be necessary to have the acid group of D223 pointed directly at the histidine proton.

It is possible that the new acidic group of D223 interacts with the positive guanidinium head group of R239, so disturbing the H-bond between the latter and the nitrogen in the pyridinium ring of the PLP, facilitating racemisation of substrates other than alanine. In PLP-dependent racemase enzymes, it has been proposed that the positively charged Arg underneath the PLP cofactor prevents protonation of the pyridinium nitrogen (normally

found with other PLP-dependent enzymes). The lack of a positively charged nitrogen means that the negative charge formed during proton abstraction of the amino acid could not be stabilized through the pyridinium ring, so preventing by-product formation (see Introduction, Section 1.6.1 for detailed racemase mechanism). If the new acidic group disturbed the H-bond between R239 and the pyridinium nitrogen, an alternate racemase mechanism could occur, explaining the novel racemase activity.

The sequence region containing the N223D mutation is highly conserved in alanine racemase enzymes, suggesting that this area has in important role in the racemase function,[93] (see Appendix 2 for protein sequence alignment). The area has been shown to support the phosphate substituent of PLP.[110]

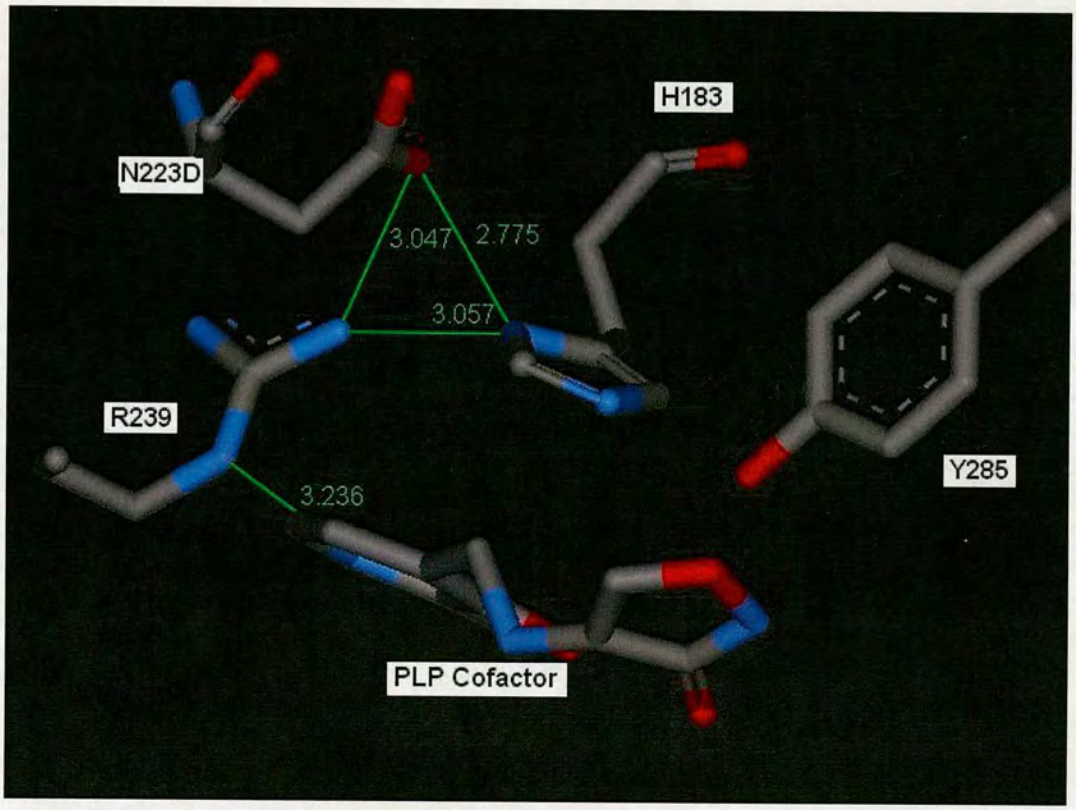
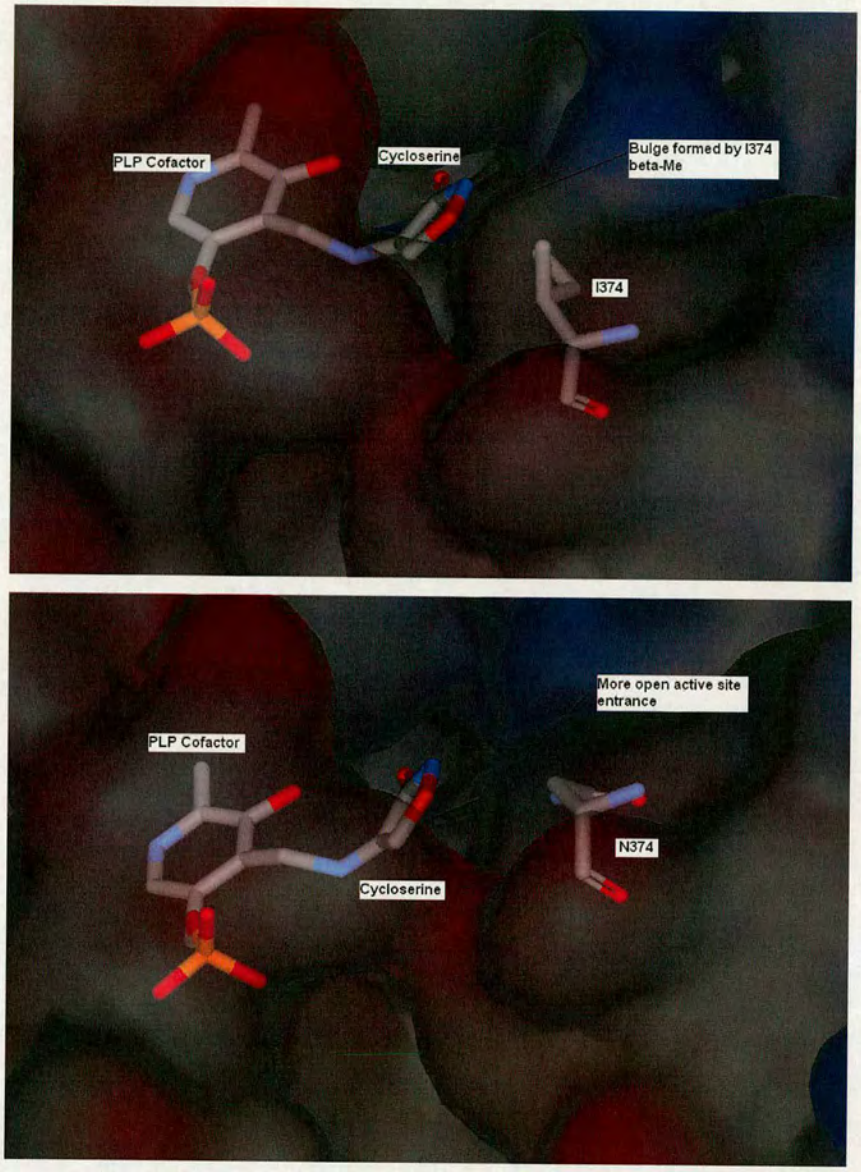


Figure 2.5.5: A close up view of key residues in solid line format in the active site of the alanine racemase. The PLP cofactor is shown and the distance to the Arg residue at the bottom of the active site is selected. Neighbouring residues that may form a catalytic triad are shown which include the N223D

2.5.8.3 I374N

Modelling revealed that the I374N mutation was shown to have altered the entrance to the active site of the racemase. The β-Me of the Ile amino acid was shown to stick out and restrict the active site entrance, Figure 2.5.6a. As there is no substitution of the β - carbon in Asn, replacing Ile with Asn removed the obstruction and opened the active site, Figure 2.5.6b. The surface of the protein was coloured by electrostatic potential and a red (negative) area can be seen close to the active site entrance. The negative charge would attract a positive group, possibly the positively charged guanidinium group of Arg and the larger entrance to the active site would allow passage of the amino acid, binding with the PLP cofactor and racemisation.

Site-directed mutagenesis carried out in 2002 on another alanine racemase from *G. stearothermophilus* found that the C-terminus region lines one side of the active site, binds the phosphate group of the PLP and was proposed to control substrate specificity.[93] The group mutated the Tyr354 and found that the racemase accepted new substrates by opening the active site to allow larger amino acid side chains to enter. Comparing the sequence of the *S. coelicolor* racemase and the racemase in the paper, it can be seen that the Ile residue mutated here, is two residues away from the corresponding Tyr residue and has been shown by modelling to play a role in opening the active site.



a) The presence of the Ile side chain restricts the entrance to the active site

b) Lack of the β-Me group in I374N opens the entrance

Figure 2.5.6: Solid surface maps displaying the entrance to the active site with charged regions shaded. The PLP cpfactor is highlighted in solid line format with the inhibitor cycloserine bound. The residue I374 or I374N is shown in solid line format around the entrance to the active site..

2.5.9 Removing the I374N mutation from 3M to create a double mutant (2M)

The I374N mutation was proposed to be important in the evolution of the racemase towards arginine. To confirm the importance, it was decided to remove this mutation

from the gene and create the double mutant I195T, N223D. The nucleotide change that resulted in the I374N mutation was shown to occur after a *Sac*II restriction site in the racemase gene. In order to remove this mutation from the triple mutant gene, a fragment of the plasmid between the *Sac*II and *Eco*RV sites was removed by restriction digestion and replaced with the identical fragment from the wild type gene, Figure 2.5.7. The racemase gene was then sequenced to confirm removal of the nucleotide change that caused the I374N mutation.

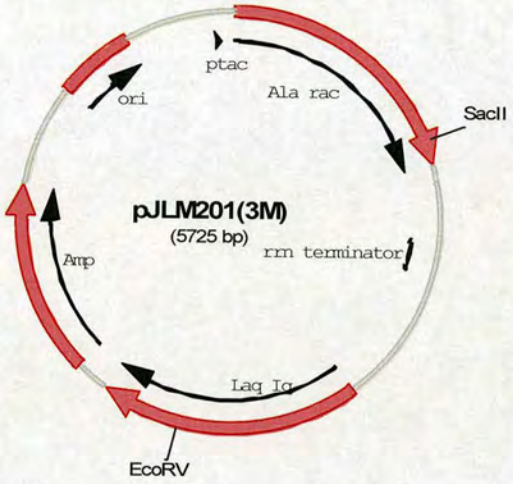


Figure 2.5.7: The plasmid containing the 3M racemase gene with the restriction sites used to remove the I374N mutation labelled

The variant 2M was assayed on the solid phase using the two-plasmid screen for L-Arg racemase activity. The wild type gene and the triple mutant were assayed as controls. As expected, there was no colour formation (indicating racemase activity) for the wild type racemase. The plate containing the 3M racemase went red after 30 min and the 2M variant turned red after 2 h incubation at room temperature.

The assay results indicated that 2M was active towards L-Arg. The delay in colour formation may have been due to a lower expression of the racemase, or it may be due to the double mutant being less active than the 3M racemase. Any difference in reaction rates will be verified with purified protein in the next chapter.

2.5.10 Attempts to clone the racemase genes into a vector to facilitate purification

It was decided to over-express and purify the wild type racemase and the variants in order to obtain kinetic data to compare the overall change in reaction rates. In order to purify the racemases using a simple technique, it was decided to clone the racemase genes into vectors that would enable His-tagging of the expressed protein. A simple metal chelating column could then be used in a one-step purification.

The vector chosen was pET16b as this would attach a His-tag onto the protein at the N-terminus. It was expected that the base level expression of the racemase during growth previously observed with the broad substrate specificity racemase in pET16b would not be a problem due to the narrower substrate specificity of the *S. coelicolor* racemase.

PCR was used to amplify the wild type racemase gene and the 3M racemase gene to enable insertion into the pET16b vector using the *Nde*I and *Bam*HI restriction sites. The amplification was carried out using the same PCR conditions originally used to amplify the *S. coelicolor* racemase gene from the gDNA, Sections 4.2.10 and 4.2.11.2.

A PCR product was observed using the SyBr Safe DNA stain and the UV transilluminator, indication successful amplification. However, subsequent restriction digestion of the PCR products and the pET16b vector and ligation/transformation reactions were never successful. The process was repeated many times varying the polymerase source, the restriction digestion enzymes, the digestion time, the pET16b source, the ligase and ligase source, the cells transformed, the antibiotics used and the template DNA used for amplification. None of these variables resulted in successful recloning of the racemase genes into the pET16b plasmid. To proceed with activity analysis, it was decided to continue using the racemase genes in their original vectors and attempt purification using a variety of separation methods.

2.5.11 Conclusions

The results presented in this chapter summarise the work carried out to create a variant library of *S. coelicolor* racemase genes and the subsequent screening of that library.

Screening the library against a mixture of substrates (L-Arg, Leu and Lys) resulted in 101 possible variants that were eventually focused to show one variant type, the triple mutant. Analysis of the mutations was carried out by modelling the structures and this work is presented in the next chapter, 2.6.

In order to analyse the different rates at which the variants can racemise Arg and Ala, it was decided to purify the racemases. Unfortunately, cloning problems prevented the use of a simple purification system using a His-tag. In order to proceed with the purification, it was decided to attempt purification using different chromatographic techniques. The investigation of a purification route is presented in Chapter 2.7.

The cloning problems experienced were later identified as being the result of the new DNA staining method introduced in the Manchester lab. SyBr Safe was used instead of the possible carcinogen ethidium bromide. The original manual supplied with the SyBr Safe does not mention any possible problems. However, the second manual supplied noted that SyBr Safe should not be used with the UV transilluminators as the DNA is degraded and difficult to recover. It was also discovered that it is very difficult to remove the SyBr safe from the DNA sample upon purification, causing problems with subsequent restriction digestion and ligation reactions. Switching to the SyBr Safe recommended illuminator resolved the majority of the cloning problems. However, this solution was realised too late for recloning to be done during these Ph.D studies.

During the synthesis of the *S. coelicolor* racemase gene variant library, the synthesis of the *dal* racemase gene library from *B. subtilis* was carried out. The initial gene library synthesised using the XL1-Red mutator method was shown to contain a high amount of

wild type (80 %) and very few mutations per gene (an average of 0.6 base mutations per gene). The epPCR technique was then used and the variant library was confirmed by sequencing ten random colonies. The epPCR was shown to have introduced an average of three mutations per gene, 40 % of the library was the wild type gene. In summary, using the epPCR technique to produce variant gene libraries was shown to be much more efficient than the XL1-Red method. The epPCR method also has the convenience of being faster as a library can be developed in three days (Day 1, epPCR and the whole plasmid PCR. Day 2, transformation. Day 3, DNA library isolated), whereas the XL1-Red method can take up to one week.

2. Results and Discussion

2.6 Attempted purification of the *S. coelicolor* putative racemase and variants

This chapter presents the development of a purification protocol for the *S. coelicolor* wild type racemase and the 2M and 3M variants. Purification of the racemase from *S. lavendulae* alanine racemase, which has 72 % identity and 79 % similarity to the *S. coelicolor* racemase, was carried out using a polyHis-tag at the C-terminus.[89] Using a similar approach was not possible here due to the problems experienced with cloning techniques so I decided to attempt purification using chromatography and separate the protein according to specific properties. The 3M variants was chosen as the candidate for purification as it has novel Arg racemase activity and so could be distinguished from the *E. coli* racemases present. Fractions could be assayed for the arg racemisation activity using the oxidase/peroxidase screen, whereas if using the wild type racemase, the Ala racemase activity of the expression host would be indistinguishable from the protein of interest.

The first method attempted was metal affinity chromatography. In this method, a column containing the chelating agent nitriloacetic acid (NTA) is incubated with a metal-containing solution (for example Ni^{2+}or Cu^{2+}) to form coordination compounds. The cell lysate is then passed through the column and any proteins with exposed His residues can bind to the metal, whilst other proteins pass through the column. The bound protein is then eluted by adding increasing concentrations of imidazole. It was suggested that although there were no additional histidines at either terminus of the racemase, the structure may have displayed enough histidines on the surface to bind to an affinity column for partial purification.

The second purification technique used was ammonium sulfate precipitation. The principle behind this technique is that solubilities of proteins vary according to the ionic strength of the solution they are in and this phenomenon can be used to separate proteins. As ionic strength increases, so does the protein solubility. However, once the maximum solubility is reached, increasing the ionic strength further then decreases the protein solubility and eventually it will precipitate. As the solubilities of individual proteins are different, changing ionic strength causes proteins to precipitate at different points. Precipitated proteins can be re-dissolved by addition of a low salt buffer. To investigate the precipitation of the racemase of interest, aliquots of CFE were subjected to increasing ammonium sulfate concentrations and the resulting supernatants and precipitates were assayed for racemase activity, using the oxidase/peroxidase screen.

The second chromatographic method investigated was anionic exchange. For this method, the isoelectric point (pI) of the protein of interest can be used to separate it from proteins with different pIs. The pI is the pH at which the protein has no overall charge and this is dependent on the surface residues. At a pH below the pI, a protein will carry a net positive charge and at a pH above the pI, a protein will have a net negative charge. Using an online pI prediction tool,[103] the theoretical pI of the *S. coelicolor* racemase was proposed to be 5.78. Therefore, at pH 8.0, the protein will have an overall negative charge and anion exchange chromatography can be used. Proteins with no net charge and proteins with pIs above 8 (positively charged) will not bind to an anion exchange column. The bound proteins are then eluted from the column by increasing the salt concentration, as the salt ions compete with the proteins for the charged compounds in the column.

The last chromatographic technique explored was hydrophobic interaction chromatography (HIC), where proteins are separated according to their surface hydrophobicity. Each protein displays different residues on the surface and this affects the behaviour of the protein with a hydrophobic medium. High salt concentrations

promote interactions between the protein and the column, causing retention. Any protein that has no, or small, hydrophobic areas on the surface will pass straight through the column. By lowering the salt concentration, proteins can be gradually eluted as the hydrophobic interactions become weakened. Different HIC columns contain different hydrophobic mediums and so can be used for many proteins with various hydrophobic properties.

After investigating the different purification methods and alternate orders of application, a route to partially purified racemase was developed.

2.6.1 Metal affinity chromatography

2.6.1.1 Cu^{2+} affinity purification

The 3M racemase was expressed in *E.coli* BL21(DE3)pLysS, according to the conditions determined in Section 2.1.7.3. It was decided to attempt metal affinity chromatography using copper, as this is reported adsorb proteins displaying one surface histidine (Amersham Biosciences HiTrap Chelating HP instruction booklet). A cell pellet containing the 3M racemase, of approximately 1.3 g was lysed using the BugBuster protocol, Section 4.5.1. The CFE was loaded onto the prepared Cu-NTA, containing Buffer A (50 mM Phosphate buffer, pH 8.0, 300 mM KCl). The bound protein was eluted from the column using an increasing concentration of Elution buffer B (Buffer A + 1 M imidazole). The resulting chromatogram is shown in Figure 2.6.1.

The total protein eluted in the main peak in Figure 2.6.1 was not very high. The fraction was assayed against L-Arg in the solution phase, as was the CFE and the non-binding proteins obtained when loading the CFE. The CFE and the non-binding proteins were seen to contain most of the Arg racemase activity. The Cu^{2+} was not binding the protein of interest and therefore a different metal was tested.

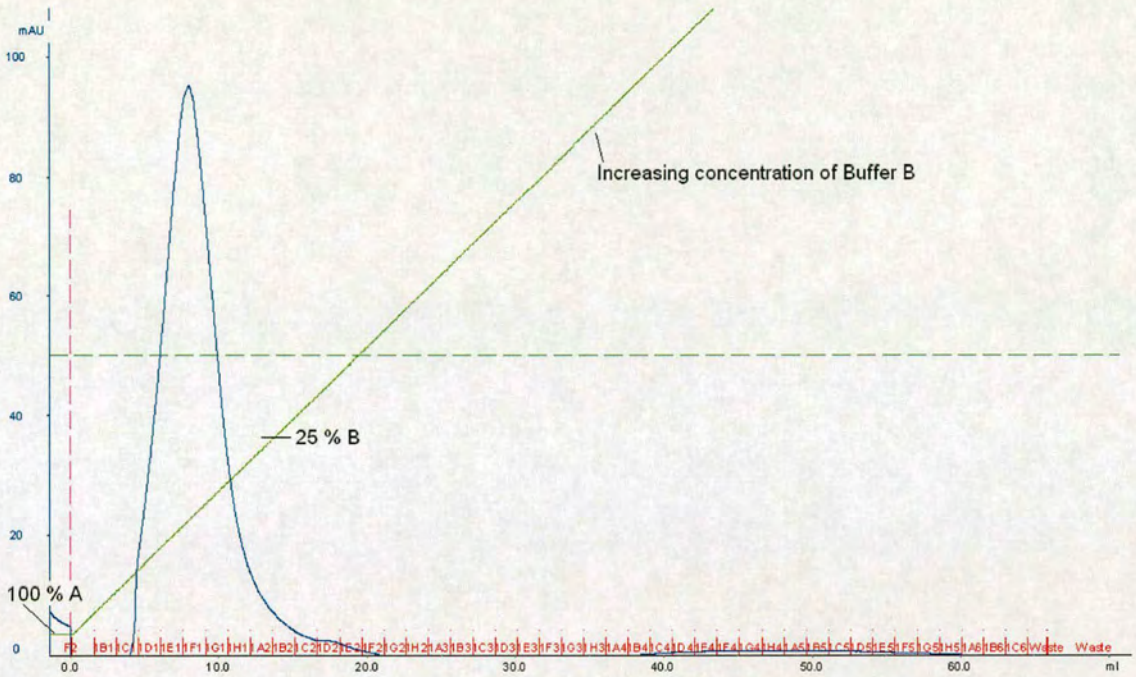


Figure 2.6.1: Cu^{2+} affinity purification of the racemase 3M. The increasing concentration of imidazole (Buffer B) is shown in green. The main protein peak starts to elute at 6 % B

2.6.1.2 Ni^{2+} affinity purification

As Cu^{2+} for the affinity purification was not successful, the metal Ni^{2+} was employed. The column and CFE were prepared as stated for the Cu^{2+} column and the same buffers and conditions were used to elute any bound protein. The resulting chromatogram is shown in Figure 2.6.2. All fractions were assayed for the racemisation of L-Arg using the oxidase/peroxidase screen, the results are shown in Figure 2.6.3. It was shown that the majority of the racemase activity was in the CFE and the non-binding proteins and a small amount in fractions C2/D2/E2.

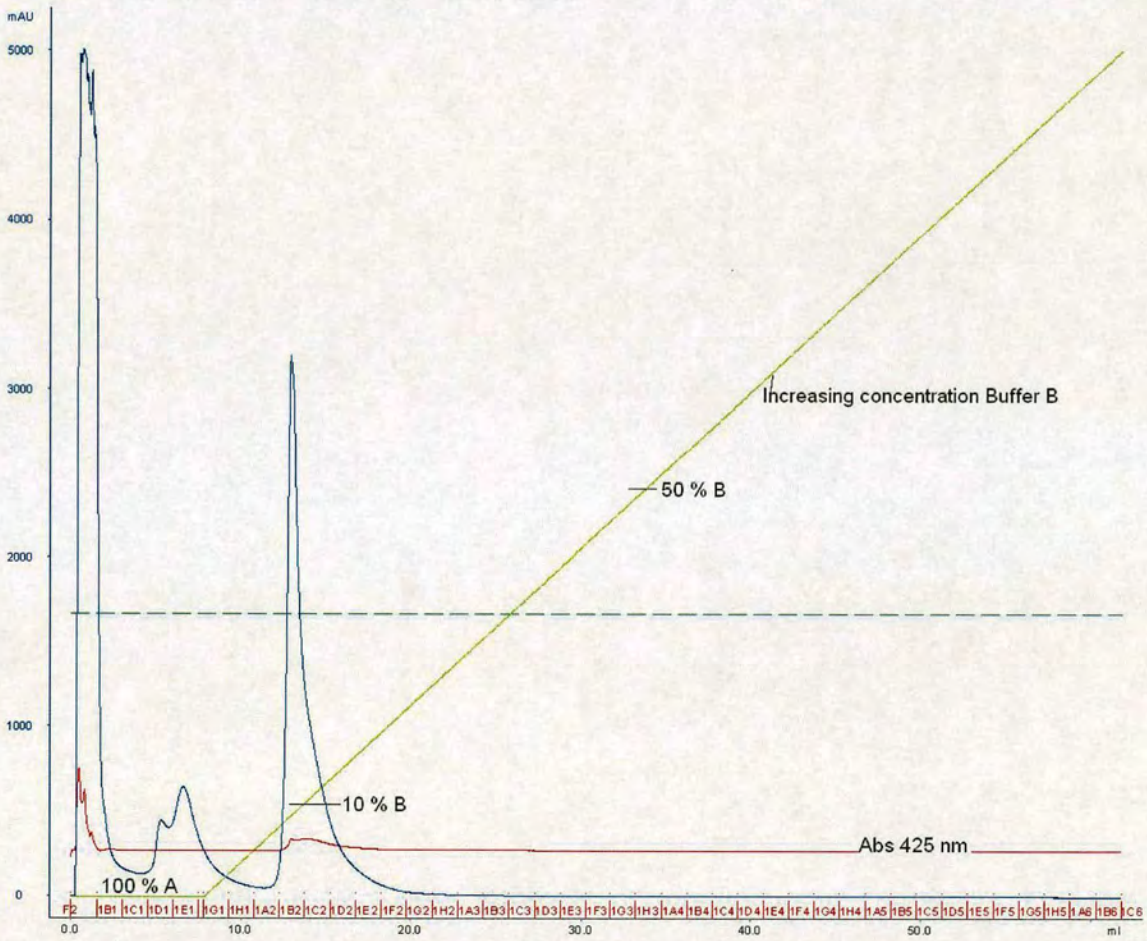


Figure 2.6.2: Ni^{2+} affinity purification of the racemase 3M. The increasing concentration of imidazole (Buffer B) is shown in green. The protein peak starts to elute at approximately 6 % B.

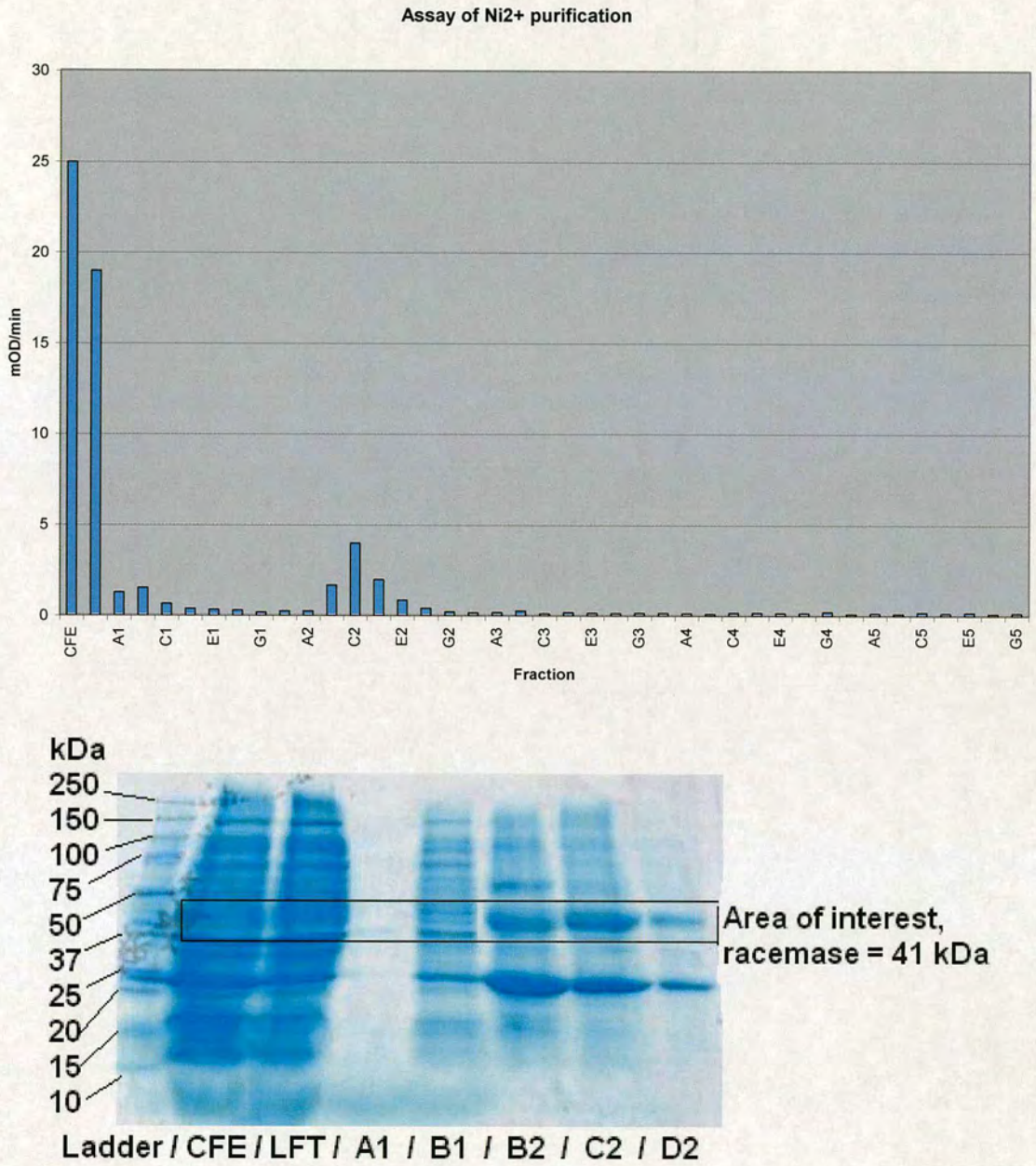


Figure 2.6.3: Assay of all peaks obtained from Ni-NTA purification for the racemisation of L-Arg. The CFE and non-binding proteins (LFT) are shown in the first two columns. An SDS-PAGE analysis of the key fractions is shown.

From the assay, it could be seen that the majority of the racemase was eluting with the non-binding proteins and a small amount in the peaks B2-E2. A partial purification was

observed, but fractions showing racemase activity were shown to still contain many other proteins. As the assay of the CFE and the flow-through confirmed that the majority of the racemase was not binding to the column, a different purification technique was needed.

2.6.2 Ammonium sulfate precipitation

The principles of ammonium sulfate precipitation were presented at the beginning of this chapter. This technique was chosen to eliminate unwanted proteins from the CFE, prior to column chromatography.

To determine how much ammonium sulfate was needed to separate the racemase from other proteins, a 2 g cell pellet containing the expressed racemase 3M, was lysed using the BugBuster protocol, Section 4.5.1. The CFE was divided into nine 1 ml aliquots and ammonium sulfate added to obtain 0, 20, 30 , 40 , 50, 60, 70, 80 and 100 % saturation. The amount of solid ammonium sulfate added per sample was calculated using the table in Figure 2.6.4. Blank samples containing water in the place of CFE were also tested to determine if the ammonium sulfate had any positive effect on the racemase assay.

	Percentage saturation at 0°																
Initial concentration	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
of ammonium sulfate	Solid ammonium sulfate (grams) to be added to 1 liter of solution																
0	106	134	164	194	226	258	291	326	361	398	436	476	516	559	603	650	697

Figure 2.6.4: Amount of ammonium sulfate to be added (mg/ml or g/l) to obtain different percent saturations.[111]

Each precipitation was carried out at 4°C, whilst stirring, over 4 h. The precipitate was then isolated by centrifugation. The pellets and supernatants from each saturation percent were assayed for L-Arg racemase activity using the oxidase/peroxidase screen. The results of the assay can be seen in Figure 2.6.5. The supernatants and pellets were very difficult to separate from each other in the 80 and 100 % saturations and accounts

for the increased activity visible in the supernatant. The blank samples confirm that the ammonium sulfate does not have any positive effect on the solution phase assay. It can be seen that a large proportion of the racemase activity is present in the pellets at 50, 60 and 70 % saturation. It was decided to add solid ammonium sulfate to 40 % saturation and discard the resulting pellet, then add solid ammonium sulfate to 70 % saturation, harvest the pellet and resuspend in buffer to obtain the maximum amount of racemase. To verify that this was the best approach, CFE containing the 3M racemase was lysed, subjected to a 40 % ammonium sulfate cut, followed by a 70 % cut and the two pelleted samples were resuspended and assayed for racemase activity towards L-Arg. The 41-70 % saturated sample was shown to be six times more active towards L-Arg than the 0-40 % sample (18 mOD/min vs. 3 mOD/min).

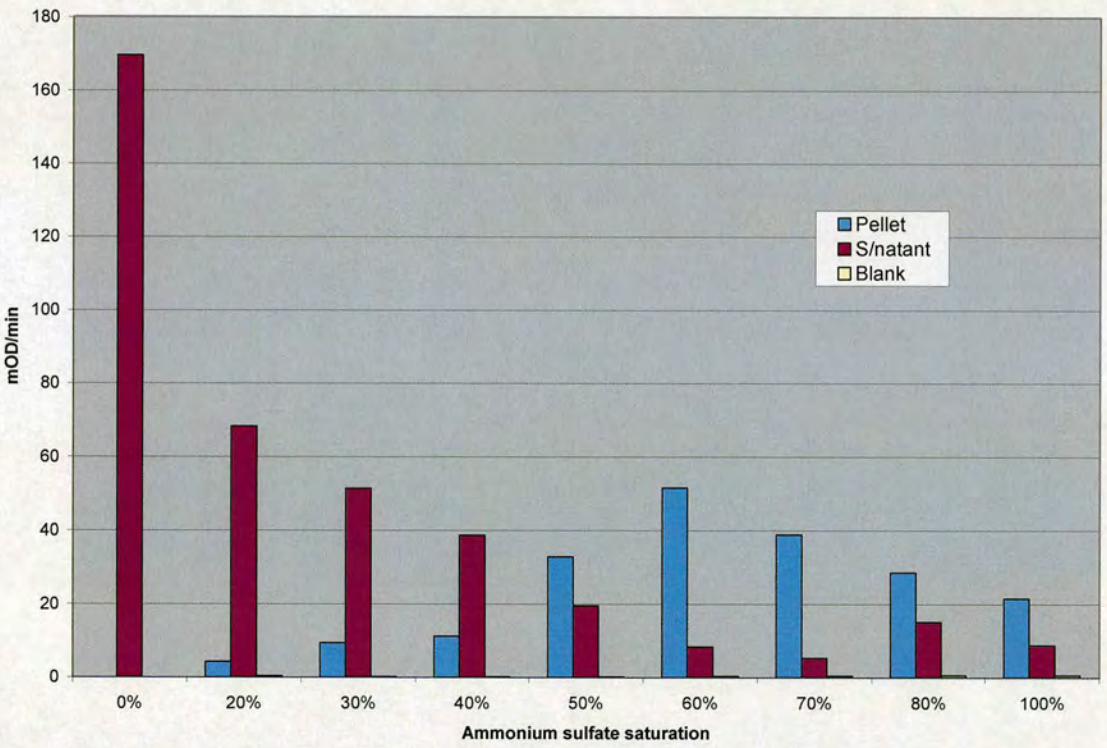


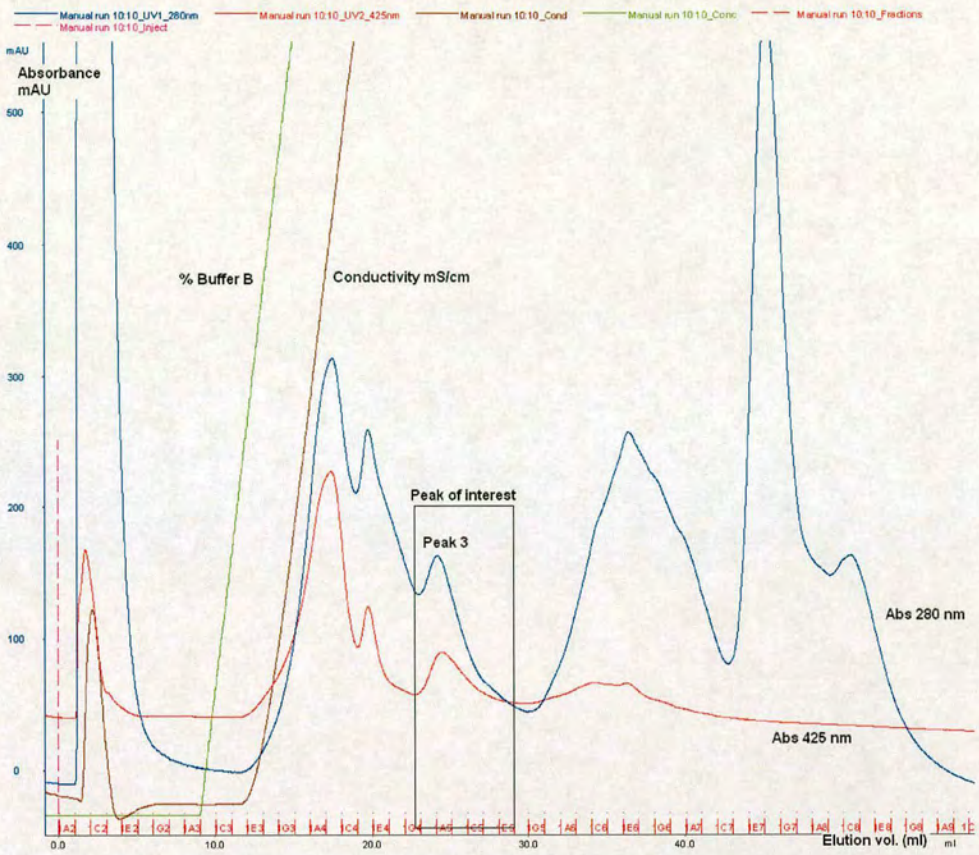
Figure 2.6.5: The racemase activities of the supernatants, pellets and blanks measured at each percent ammonium sulfate saturation

For the ammonium sulfate saturation experiments, the 0-40 % pellet was initially isolated by addition of solid ammonium sulfate. However, it was decided to use saturated ammonium sulfate solution, to add drop-wise to the stirring CFE, to avoid areas of high concentrations around the solid salt that may cause unwanted precipitation.

2.6.3 Anion exchange chromatography

Anion exchange was the next purification method tested. The CFE containing the 3M racemase was prepared (Section 4.5.1) and subjected to a 0-40 % ammonium sulfate precipitation to remove unwanted proteins, followed by a 41-70 % precipitation to obtain the racemase-containing sample. The precipitated proteins were resuspended in Buffer A (50 mM Tris.HCl, pH 8.0, 50 mM NaCl and 50 μM PLP) and any excess ammonium sulfate was removed using a PD10 column (Section 4.5.3). Purification of the sample was then tested using a range of anion exchange columns: ResourceQ, Q sepharose fast flow; Q sepharose XL; DEAE sepharose fast flow; and ANX sepharose 4 fast flow. The protein samples were applied to each of the columns and eluted according to Section 4.5.7.2, using Elution Buffer 50 mM Tris.HCl, pH 8.0, 50 mM NaCl and 50 μM PLP, 1 M NaCl.

The ResourceQ (ResQ) column provided the best separation of the peak (Figure 2.6.6, peak 3, eluted at 18 mS/cm) that displayed racemase activity towards L-Arg. Colour formation in assay solution without substrate was also measured and subtracted from the L-Arg racemase data. A plot of the chromatogram and an SDS-PAGE analysis of peak 3 are shown in Figure 2.6.6. A partial purification of the racemase was successful. However, it is unknown which of the bands in the 41 kDa area of the gel correspond to the racemase. It was decided to attempt a second column to purify the racemase further.



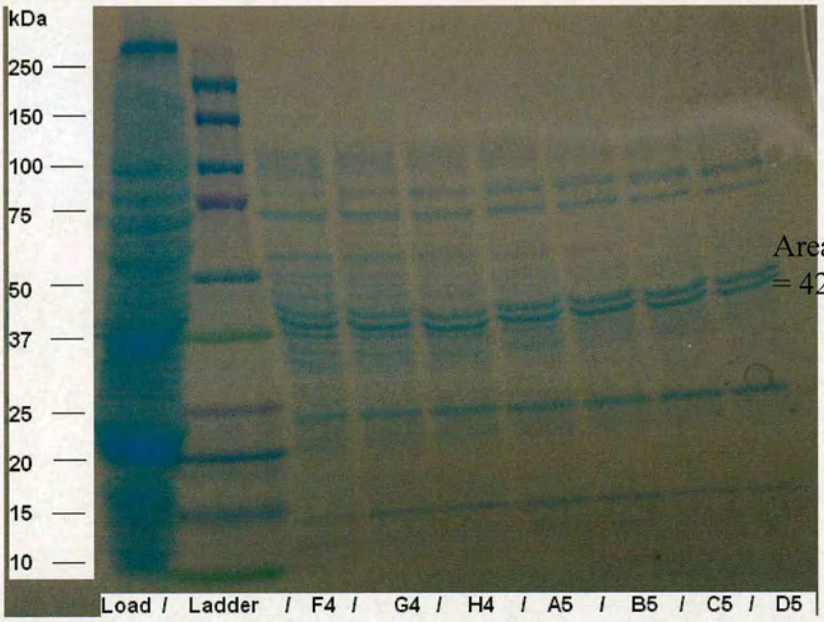


Figure 2.6.6: The ResQ separation of proteins contained in the 41-70 % precipitate of the racemase CFE. Peak 3 eluted at 18 mS/cm and was shown to contain L-Arg racemase activity. An SDS-PAGE analysis of the fractions in peak 3 is shown. The Load is the protein sample prior to ResQ purification

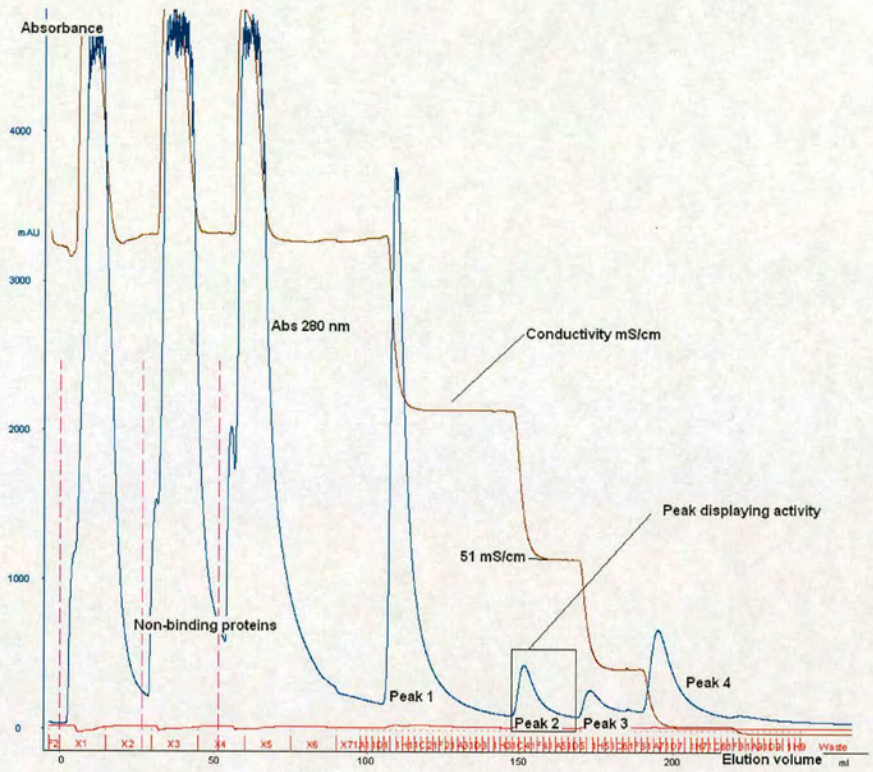
2.6.4 Hydrophobic interaction chromatography

Hydrophobic interaction chromatography (HIC) relies on protein binding to a hydrophobic medium in a high salt solution. It is a technique ideal for following an ammonium sulfate precipitation as the sample can be applied to the column without the need for desalting. A range of columns were tested for the purification of the racemase: Phenyl FF (high and low sub); Phenyl HP; Butyl FF, Butyl-S FF; and Octyl FF. The Phenyl HP gave the best separation of the peaks displaying racemase activity and was investigated for further separations.

The CFE containing the 3M racemase was prepared (Section 4.5.1) and subjected to a 41-70 % ammonium sulfate precipitation to obtain the racemase-containing sample. The precipitated proteins were resuspended in Buffer A (50 mM Tris.HCl, pH 8.0, 50 mM NaCl, 1 M $(NH_4)_2SO_4$ and 50 μM PLP), loaded onto the column and eluted with an increasing concentration of Elution buffer 50 mM Tris.HCl, pH 8.0, 50 mM NaCl and 50 μM PLP, (decreasing salt concentration). The elution gradient was performed step-wise starting with 100 % Buffer A, then 60 %, 30 %, 10 %, 1 % and 0 %. The chromatogram can be seen in Figure 2.6.7. The CFE, non-binding proteins and fractions from each of the four peaks were assayed for L-Arg racemase activity. Colour formation in assay solution without substrate was also measured and subtracted from the L-Arg racemase data. Peak 2 was shown contain active racemase. The fractions were concentrated using a 50 kDa MWCO concentrator to aid purification and remove any excess ammonium sulfate, washed three times with 5 ml Elution buffer, to obtain a final volume of 2 ml.

It was then decided to purify this sample further using the ResQ anionic exchange chromatography.

a)



b)

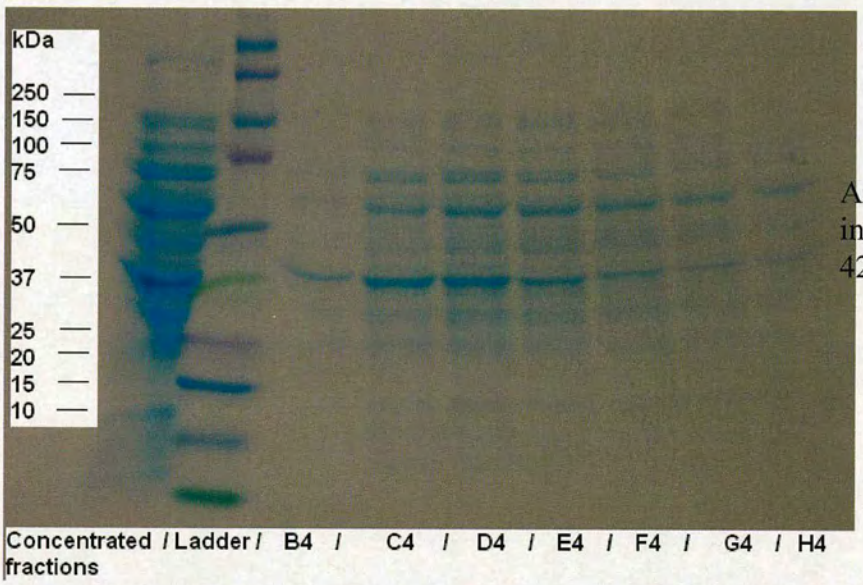


Figure 2.6.7: a) The separation of proteins using HIC. Peak 2 was shown to contain the L-Arg racemase activity and b) the fractions were analysed by SDS-PAGE and concentrated

2.6.5 Anionic exchange chromatography after HIC

The partially purified racemase sample from Section 2.6.4 was applied to the ResQ column using the buffers and conditions reported in Section 2.6.3. The racemase-containing peak was eluted with a step-wise gradient of increasing NaCl concentration. The racemase fraction was eluted at 10 % Elution buffer (12 mS/cm). The chromatogram can be seen in Figure 2.6.8.

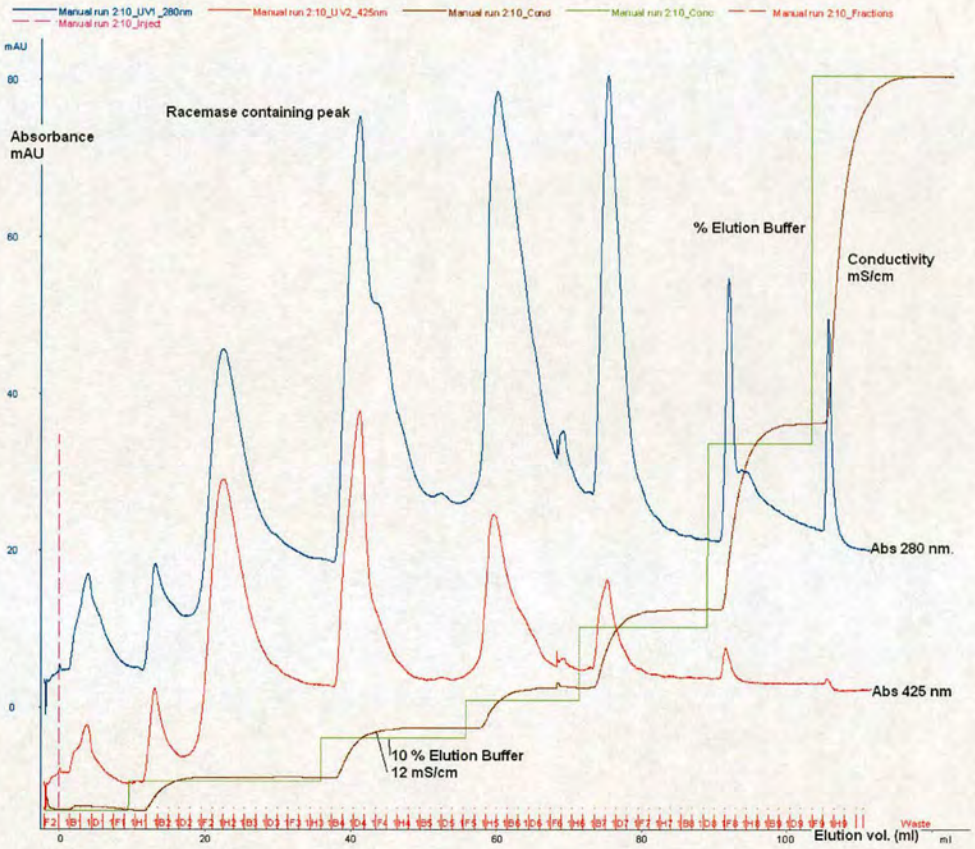


Figure 2.6.8: Protein separation using anionic interaction chromatography after ammonium sulfate fractionation and HIC. Peak displaying racemase activity is shown

All fractions corresponding to peaks were collected, concentrated to 0.5 ml and assayed for L-Arg racemase activity. Colour formation in assay solution without substrate was also measured and subtracted from the L-Arg racemase data. The racemase-containing peak was shown to elute at 10 % elution buffer with a conductivity of 12 mS/cm.

The activities of 50 µl sample aliquots of the racemase-containing peaks after the HIC and the ResQ purifications were shown to be 5.9 and 1.1 mOD/min respectively. However, the total volumes of the samples were 4 and 0.5 ml; therefore the total activities per sample were 236 and 11 mOD/min after the HIC and the ResQ steps respectively. Much racemase activity is lost during the ResQ purification (little activity was seen in the other peaks). The gradual loss of activity was observed throughout all trial purifications. Initially, racemase activity could be observed in the CFE using the solution phase assay within 30 min. However, after the first column purification the activity was not visible until assay for a few hours and fractions from the second column were assayed overnight in order to see racemase activity. Samples of the CFE, the ammonium sulfate precipitation, the concentrated racemase-containing fractions from the HIC and the concentrated racemase-containing fractions from the ResQ were analysed by SDS-PAGE and can be seen in Figure 2.6.9.

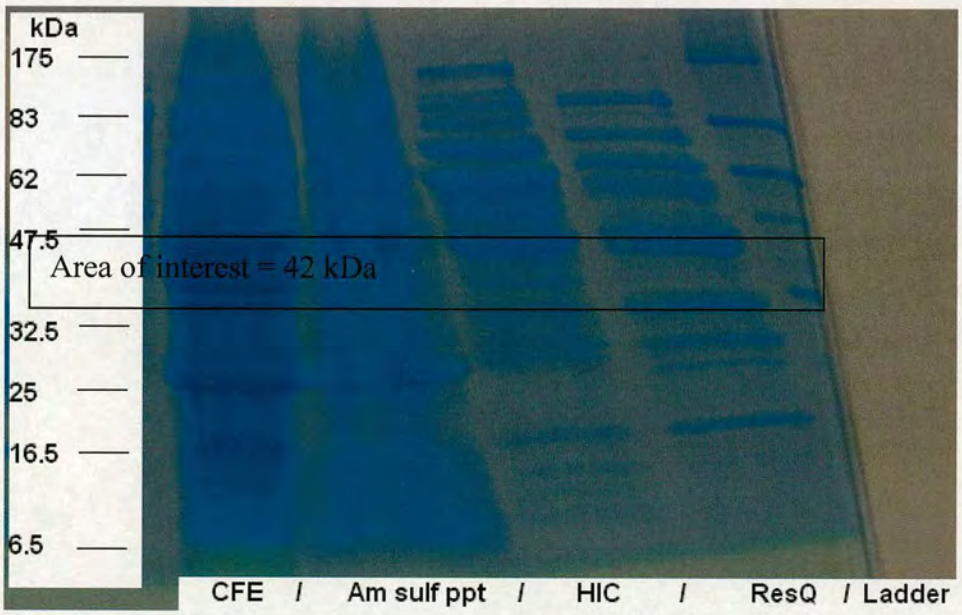


Figure 2.6.9: SDS-PAGE analysis of the purification of the triple mutant racemase (41 kDa)

2.6.6 Conclusions

The purification process as adopted for the racemase appears to slowly render the enzyme inactive. After the purification steps presented so far (ammonium sulfate precipitation, HIC and anionic exchange chromatography), the total racemase activity in the sample was 11 mOD/min, compared with 236 mOD/min after the HIC and 2045 mOD/min in the CFE. Comparing the purity of the racemase using the SDS-PAGE analysis in Figure 2.6.9, it was seen that there were many unwanted proteins present; the loss of activity after the HIC and ResQ separation did not correspond to a substantial purification of the racemase as desired.

It was decided to partially purify the wild type racemase, the 2M and the 3M racemases using the ammonium sulfate precipitation and the HIC and then determine the substrate specificities for each. Each racemase would be expressed under exactly the same conditions, the same mass of cell pellets would be purified and the total protein present in the final samples would be determined. Analysis by SDS-PAGE would confirm the ratios of proteins present and allow some comparison of the substrate specificity data.

Due to time constraints it was not possible to investigate a new purification route, one of the major problems with the method was that the racemase was not over-expressed well enough in the *E. coli* BL21(DE3)pLysS cells to see a large excess of the racemase by SDS-PAGE analysis. The expression of the 3M racemase was attempted in *E. coli* BL21(DE3) but was unsuccessful as there appeared to be less racemase activity than when using *E. coli* BL21(DE3)pLysS. Ideally, the genes would be cloned into different vectors that allow better expression and more efficient purification by affinity chromatography, such as metal affinity using a polyhis-tag. Using this method, the racemases could be purified more rapidly and avoid the loss of activity observed in the current route. The screen has been successfully employed to determine a purification route for the racemases. It was possible to assay fractions from each step of the purification for L-Arg racemase activity.

2. Results and Discussion

2.7 Partial purification and substrate specificity determination of the wild type racemase and variants

In Chapter 2.6, a route to the partially purified wild type racemase, the double and the triple mutant racemases was explored. The CFE was prepared using BugBuster, an ammonium sulfate precipitation was employed, followed by HIC. It was decided to determine the substrate specificities of each using the partially purified preparations and the data is presented here. It is shown that compared to the wild type racemase, the only other amino acid that the double and the triple mutant racemases had novel activity towards was L-Arg, whilst the activity towards the wild type substrate L-Ala decreases.

A 6 g wet cell weight pellet grown and harvested according to the conditions described in Section 2.1.7.3, was used for the purification of the wild type racemase from *S. coelicolor*, the double mutant and the triple mutant racemases. To measure and compare the substrate specificity of the host *E. coli* racemases, vector without racemase gene (pTTQ18) was also used to transform *E. coli* BL21(DE3)pLysS and a 6 g cell pellet was obtained using the identical expression protocol. The cell free extract from each of the cell pellets was obtained using BugBuster with added lysonase, protease inhibitor cocktail and PMSF (Section 4.5.1.2).

The CFEs were subjected to an ammonium sulfate precipitation using a saturated ammonium sulfate solution (Section 4.5.1). The precipitated proteins were resuspended in Buffer A (50 mM Tris.HCl, pH 8.0, 50 mM NaCl, 1 M $(NH_4)_2SO_4$ and 50 μM PLP), filtered and loaded onto the Phenyl HP column pre-equilibrated with the same buffer. The non-binding proteins were removed with 40 % Elution buffer (50 mM Tris.HCl, pH 8.0, 50 mM NaCl and 50 μM PLP). The racemase-containing proteins were eluted with

70 % Elution buffer, combined and concentrated using a 50 kDa MWCO concentrator to approximately 1 ml. The sample was washed three times on the concentrator with 5 ml Elution buffer and the resulting 1 ml sample was diluted to 3.5 ml using the Elution buffer. At this stage, samples were removed and diluted 1:10 to enable protein concentration determination using the BCA assay kit (Section 4.5.6). The results are shown in Table 2.7.1.

Partially purified sample	**Protein concentration mg/ml**
E. coli host only	4.4
Wild type *S. coelicolor* racemase	3.7
Double mutant	4.8
Triple mutant	5.7

Table 2.7.1: Protein concentrations of the partially purified samples

The samples were analysed by SDS-PAGE to visually compare the protein content in each, Figure 2.7.1. In the samples containing expressed racemase (not the *E. coli* sample), are the same proteins expressed in similar ratios. Due to the increasing band strength, it can be seen that total protein concentration increases from the WT to the triple mutant, as recorded in Table 2.7.1. The *E. coli* host sample contains almost exactly the same proteins in the same ratios as the other samples, except for a more heavy protein band at approximately 47.5 kDa. This band does not correspond to either of the *E. coli* racemases, as the dadX racemase is 38.8 kDa and the alr racemase is 39.2 kDa.

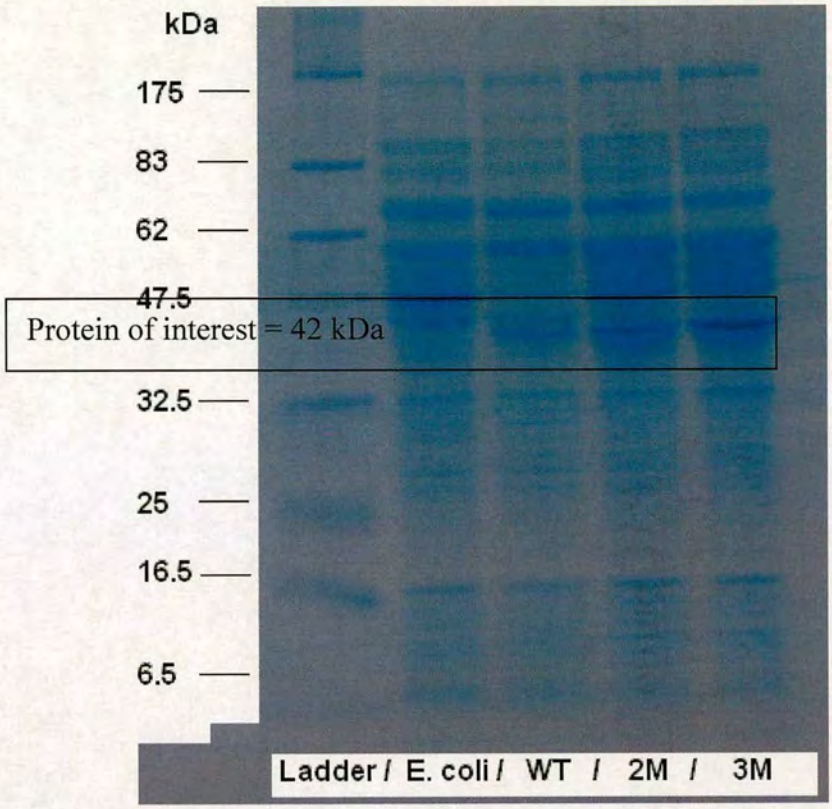


Figure 2.7.1: SDS-PAGE analysis of the samples used to determine the substrate specificities of the racemases. E. coli refers to the expression host, WT is the S. coelicolor wild type racemase, 2M refers to the double mutant, 3M refers to the triple mutant

The partially purified racemases were then assayed using the oxidase/peroxidase screen against a panel of L-amino acids to determine their substrate specificities, 50 µl was used to assay each substrate. The assays were performed in triplicate, with blank assays run containing water instead of the racemase preparation. The latter was measured to subtract any colour formation resulting from reaction of the oxidase or the HRP with the assay mixture. The triplicate measurements were averaged and the blank/water data subtracted. The maximum rates of colour formation were then determined per µg protein. The data is presented in Figure 2.7.2.

It is observed that as the substrate specificity of the racemases is broadened to accept Arg, the activity towards Ala decreases and the activity towards 2-ABA increases. It has

been seen in the evolution of other enzymes towards novel substrates, for example the evolution of a monoamine oxidase towards α-methylbenzylamine that as the substrate specificity is broadened, the activity towards the wild type substrate decreases.[10] Arginine is much larger than alanine, see Figure 2.7.3, and as the enzyme has been evolved towards the larger substrate, the specificity for alanine has decreased. 2-ABA is one carbon larger than alanine, Figure 2.7.3, and the specificity for this substrate has increased. It is possible that either (or both) of the I195T and N223D mutations have created an opening in the active site in the region where the side chains of the amino acids bind.

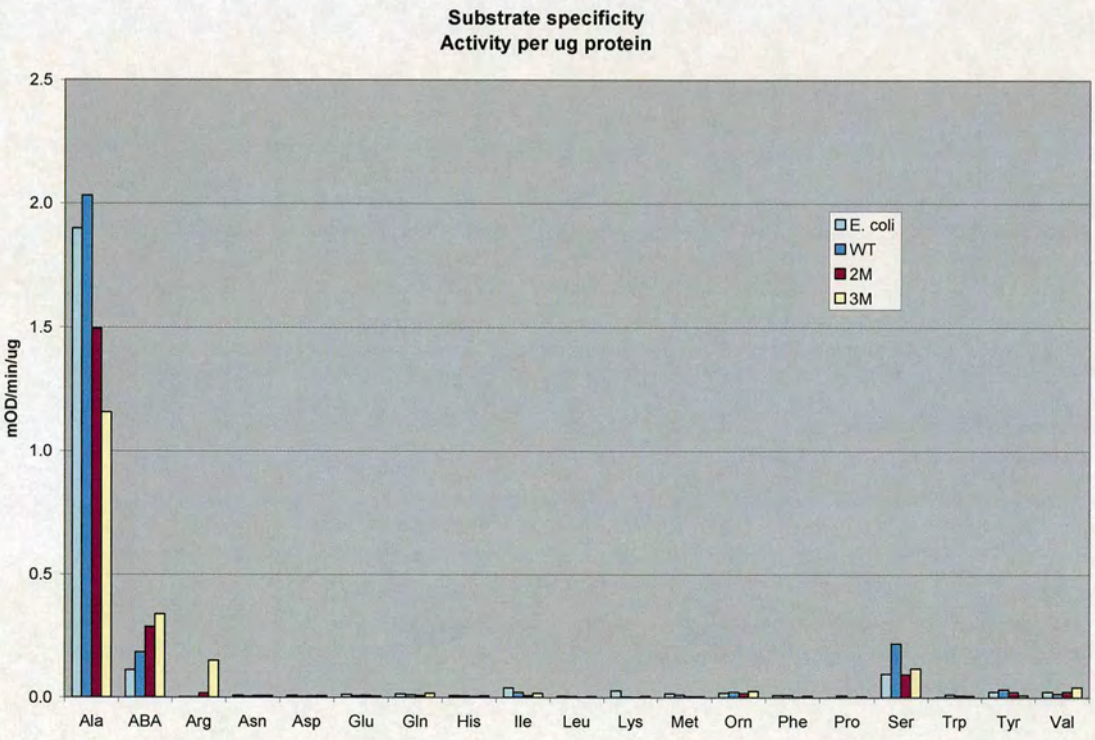


Figure 2.7.2: The substrate specificities of the E. coli racemase, the wild type S. coelicolor racemase and the double (2M) and triple (3M) variants

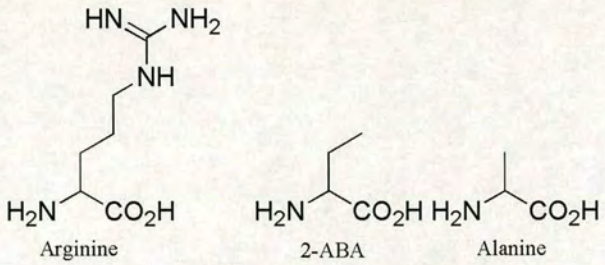


Figure 2.7.3: Structures of Arg, 2-ABA and Ala

From the substrate specificity data presented in Figure 2.7.2, it can be seen that the only substrate that the variants have novel activity for is arginine. The activities of the *E. coli* racemase and the wild type *S. coelicolor* racemase for arginine were found to be 0.003 and 0.002 mOD/min/μg respectively. The activity of the double mutant towards arginine was 0.017 mOD/min/μg and for the triple mutant, 0.151 mOD/min/μg. The double mutant is ten times more active towards arginine than the wild type and almost ten times less active than the 3M. The data implies that the triple variant contains a mutation that is important for racemase activity, I374N. However, the double mutant does have some increased activity towards arginine, implying that either or both of the mutations N222D and I195T aid racemisation of the novel substrate.

2. Results and Discussion

2.8 Conclusions and Future Work

The current project has developed a new high-throughput screen that can be applied to many different classes of enzymes. This is similar to the biosynthetic tree screening model proposed by Sutherland.[60] The screen relies on the *in vivo* production of an enantioselective compound by the enzyme of interest, which is then a substrate for an enantioselective FAD-dependent oxidase, also expressed *in vivo*.

Amino acid racemases were chosen as candidate enzymes to study the two-plasmid screen as they have a covalently bound cofactor and they do not require any additional substrates for enzyme action. Using amino acid racemases also tested the enantioselective part of the screen, since if the oxidase was not able to distinguish well between enantiomers many false positive results would have been generated. Three amino acid racemase genes were isolated from gDNA and cloned into vectors for subsequent enzyme analysis. The racemase from *S. coelicolor* is documented in genomic research sites as putative. The work carried out here into the substrate specificity of this racemase is the first carried out according to the current literature and it can now be claimed that this gene encodes a racemase for which alanine appears to be the best substrate.

Once amino acid racemase libraries were successfully generated, the first round of screening of the *S. coelicolor* racemase library identified enzymes with possible novel activity towards arginine. Subsequent, more detailed analysis of the variant activity confirmed that the novel screen had identified an alanine racemase that accepted arginine as a substrate. Comparison with the literature revealed that the I374N mutation was in the same region as a previously published study that had shown that site-directed mutagenesis of a neighbouring residue increased the substrate specificity.[93] The

published work had not mutated the racemase to be active towards arginine. However, the work carried out here introduced a residue that opened the entrance to the active site to allow larger amino acids entry.

The work carried out in this project showed that the two-plasmid screening system could be used to analyse variant enzyme libraries and detect novel activity from a background of inactivity.

The two-plasmid screen is versatile and could now be developed for a range of enzyme screening projects. Future work could see competent cells made containing the second plasmid with the D-amino acid oxidase gene. When a reaction of interest produces a D-amino acid, the genes to be screened can be cloned and used to transform the competent cells. It would also be possible to screen for enzymes that produced L-amino acids as broad substrate L-amino acid oxidases have been discovered.[112] If the reaction of interest produces an amine, the enantioselective amine oxidase developed by Turner and co-workers could be used, again stored in competent cells on a plasmid.[10] Thus a whole screening system based on FAD-dependent oxidases could be made available when screening for reactions that create simple organic molecules.

3. Materials

3.1 Materials for Microbiology

3.1.1 LB Medium

Per litre: Tryptone 10 g, Yeast Extract 5 g, NaCl 10 g. Sterilise by autoclaving, 121°C for 20 min

3.1.2 LB/agar Medium

Per litre: Tryptone 10 g, Yeast Extract 5 g, NaCl 10 g, Agar 15 g. Sterilise by autoclaving, 121°C for 20 min

3.1.3 SOC Medium

Per litre: Yeast extract 0.5 % w/v, Tryptone 2.0 % w/v, NaCl 10 mM, KCl 2.5 mM, $MgCl_2$ 10 mM, $MgSO_4$ 10 mM, Glucose 20 mM. Sterilise by autoclaving, 121°C for 20 min

3.1.4 Antibiotic Stock Solutions

Ampicillin	50 mg/ml dH_2O
Chloramphenicol	34 mg/ml EtOH
Kanamycin	50 mg/ml dH_2O

Sterilise by 0.2 μM filtration. Stock solutions are 1000 times the working concentrations.

3.1.5 IPTG Stock Solution

IPTG	1 M in dH_2O

Sterilise by 0.2 μM filtration.

3.1.6 Spheroplast Preparation Buffer

Sucrose 20 mM in dH_2O, Lysozyme to a final concentration 0.2 mg/ml

3.2 Materials for Molecular Biology

3.2.1 Bacterial Resuspension /Lysis Buffer –Genomic DNA extraction

$Na_2EDTA.2H_2O$ 1.861 g, Tris Base 0.606 g, Dissolve in 80 ml dH_2O. Add 10 % Tween-20 5 ml, 10 % Triton X-100 5 ml. pH to 8.0 with HCl. Final volume to 100 ml with dH_2O. Sterilise by 0.2 µM filtration.

3.2.2 Bacterial Deproteinisation Buffer –Genomic DNA extraction

Guanidine HCl 28.659 g dissolved in 70 ml dH_2O, added 100 % Tween-20 20 ml and made up to a final volume of 100 ml with dH_2O and sterilised by 0.2 µM filtration.

3.2.3 Agarose Gel Sample Loading Sample Buffer

Glycerol 30 % (v/v), EDTA 20 mM, Bromophenol blue 0.2 % (w/v)

3.2.4 TAE (Tris acetate EDTA buffer) 50x

Tris base 242g, acetic acid 57.1ml, 0.5M EDTA pH 8.0 100ml

3.3 Materials for Protein Purification

3.3.1 Phosphate buffer

A 1 M solution of the dibasic KH_2PO4 was prepared. A 1 M solution of the monobasic K_2HPO4 was prepared. For pH 8.0, 94 ml of the monobasic solution was mixed with 6 ml of the dibasic solution. For pH 7.0, 61.5 ml of the monobasic solution was mixed with 38.5 ml of the dibasic solution.

3.3.2 Ammonium Sulfate Precipitation

A saturated ammonium sulfate solution was made by adding solid ammonium sulfate to water until no more dissolved.

3.3.3 Cu and Ni Affinity Chromatography

Activating metal solutions: 0.1 M $CuSO_4$, 0.1 M $NiSO_4$

Binding Buffer: pH 8.0 50 mM Phosphate buffer, 300 mM KCl

Elution Buffer: pH 8.0 50 mM Phosphate buffer, 300 mM KCl, 1 M Imidazole

Column Stripping Buffer: 0.5 M EDTA

3.3.4 Anion Exchange Chromatography

Start Buffer: 20 mM Phosphate buffer, pH 8.0, or pH 7.0

Elution Buffer: 20mM Phosphate buffer, pH 8.0 or pH 7.0, 1M NaCl

Following optimisation

Start Buffer: 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 50 μM PLP.

Elution Buffer: As start buffer, +1 M NaCl

3.3.5 Hydrophobic Interaction Chromatography

Start Buffer: 20 mM Phosphate buffer, 50 μM PLP, pH 7.0 or 8.0, 1 M Ammonium sulfate

Elution Buffer: 20 mM Phosphate buffer, 50 μM PLP, pH 7.0, or pH 8.0

Following optimisation

Start Buffer: 50 mM Tris-HCl buffer, 50 μM PLP, pH 8.0, 1 M Ammonium sulfate

Elution Buffer: 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 50 μM PLP

3.3.6 SDS PAGE Running Buffer

Tris/HCl pH 8.3 50 mM, Glycine 380 mM, SDS 2 % (w/v)

3.3.7 Protein gel sample loading buffer

150 mM Tris.HCl pH 6.8, SDS 1 % (w/v), Glycerol 50 % (v/v), Bromophenol blue 0.01 % (w/v), or Laemmli Sample Buffer (BioRad, 161-0737)

3.3.8 Western Blotting Buffer

10 mM 3-(cyclohexylamino)-1-propane sulfonic acid (CAPS) pH 11, 10 % v/v methanol.

3.3.9 Coomassie blue staining solution

Ethanol 50 %, acetic acid 10 %, Coomassie blue R250 0.2 % (w/v)

3.3.10 De-staining solution

Ethanol 25 %, acetic acid 10 %

4. Experimental

4.1 Identifying alternative racemases

The sequence of the *P. putida* putative gene was obtained from www.tigr.org. A BLAST search was carried out using this as the query sequence on all gDNA stocks held in the group. The search matrix used was the default BLOSUM62. Sequences returned by the program were given an expected, or E-, value providing a measure of the significance of an alignment. This statistical value reveals how closely matching the query sequence is to the returned sequence. Results returned with E values less than 0.01 were examined, as these sequences were unlikely to be related by chance alone.

4.2 DNA Analysis and Manipulation

4.2.1 Horizontal agarose gel electrophoresis

1x TAE agarose was heated until the agarose was completely dissolved. It was cooled to approximately 45°C before the addition of ethidium bromide, final concentration approximately 50 μg/L (or SYBR® Safe stain, Invitrogen), and allowed to set. DNA samples containing loading buffer were loaded into the wells alongside a DNA molecular weight marker. The gel was run at 100 V and visualised with UV radiation on a trans-illuminator.

4.2.2 Purification of DNA fragments using gel electrophoresis

Bands were cut under UV radiation from the gel rapidly to avoid over-exposure of the DNA. Bands were purified using the QIAquick gel extraction kit (QIAGEN) according to the manufacturer's protocol.

4.2.3 Purification of genomic DNA

50 ml cultures of *P. putida* ATCC47064 at approx OD_{600} 2 were harvested at 6 000 rpm, 15 min and subjected to DNA purification using the QIAGEN blood and cell culture DNA extraction kit. Eluted DNA was stored at –20°C.

4.2.4 Purification of plasmid DNA

3-5ml bacterial cultures were harvested and the plasmid DNA extracted using the Qiagen QIAprep spin miniprep kits. Eluted DNA was stored at –20°C.

4.2.5 DNA Quantification

DNA concentrations were determined by spectroscopy (Eppendorf BioSpectrometer) at a wavelength of 260 nm. Purity was assayed on the basis of the A_{260}/A_{280} ratio, where a value of approximately 1.8 corresponds to pure dsDNA.

4.2.6 Restriction endonuclease digestion of DNA

All enzymes used were purchased from New England Biolabs Inc. (NEB) and were purchased with the appropriate buffer (10x concentrate). Digests were carried out at 37°C for a minimum of 1 h, maximum of 4 h, or overnight at room temperature. Double digests were performed in the appropriate buffer as recommended by NEB, table 4.5.

Enzymes in digest	**Buffer Recommended**
*Nde*I, *Bam*HI	*Bam*HI
*Eco*RI, *Bam*HI	*Eco*RI
*Spe*I, *Hind*III	Buffer 2
*Hind*III, *Eco*RI	*Eco*RI

Table 4.5: Buffers for double digests

4.2.7 Vectors used

The following vectors were used for cloning: pET16b, pTTQ18, and pINGE2. All were already held in the laboratory by colleagues. pINGE2 was obtained from Ingenza Ltd and is a modified version of pTTQ18, containing a second multiple cloning site.

4.2.8 Dephosphorylation of plasmid DNA 5' termini

To reduce the possibility of vector re-ligation, shrimp alkaline phosphatase (SAP) from Promega was used to dephosphorylate the 5' termini of all vectors after digest. Reaction buffer was supplied 10x concentrate. Reactions were performed, according to the protocol supplied, at 37°C for 15 min followed by heat-inactivation of the phosphatase at 65°C for 15 min. The reaction mixture was cleaned using the QIagen PCR Cleanup kit.

4.2.9 DNA Ligation

Trial ligations were carried out and the optimum vector:insert ratio found for a 1.2 kbp insert and a 6 kbp vector was found to be 3:1. All subsequent ligations were performed with a vector: insert ratio of approximately 3:1. Total reaction volumes varied between 10-20 µl according to DNA concentrations. T4 DNA ligase from NEB was used, supplied with reaction buffer (10x concentrate). Ligations were left at room temperature for a minimum of 1 h before use in transformations.

4.2.10 Oligonucleotides for Racemase cloning

Primers to amplify the desired gene sequences were designed and ordered from Thermo Electron® and Sigma Genosys®. Sequences shown in tables 4.1-4

For Vector	Cleavage site	Sequence 5'→ 3'
pET16b	*Nde*I	GCT CTA GAC ATA TGC CCT TTC GCC GTA CCC TTC TGG
pET16b	*Bam*HI	ATA GGA TCC ATC ATC AGT CGA CGA GTA TCT TCG GG
pTTQ18	*Eco*RI	GAC CCC GTT ATG AAT TCT ATG CCC TTT CGC CGT ACC
pTTQ18	*Bam*HI	Same as for vector pET16b
pINGE2	*Spe*I	GCT TAT CTA GAC TAG TCC CTT TCG CCG TAC CCT TCT GG
pINGE2	*Hind*III	CGC AAG CTT CAT CAG TCG ACG AGT ATC TTC GGG TTG G

Table 4.1:Primers for cloning P.putida putative racemase gene

For Vector	Cleavage site	Sequence 5' → 3'
pTTQ18	*Eco*RI	GGT TAT CTA GGA ATT CGA TAA AGC TTT GCC GAG AAG TC
pTTQ18	*Bam*HI	ATC GGA TCC TCA TCA CTA CAG GTA TAA TAC TGG AG

Table 4.2: Primers for cloning B. subtilis alr gene

For Vector	Cleavage site	Sequence 5' → 3'
pTTQ18	*Eco*RI	GCT TAT CTA GGA ATT CGA GCA CAA AAC CTT TT TAC AGA GAT ACG
pTTQ18	*Bam*HI	ATA GGA TCC TCA TCA TTA ATT GCT TAT ATT TAC CTG C

Table 4.3: Primers for cloning B. subtilis dal gene

For Vector	Cleavage site	Sequence 5' → 3'
pTTQ18	*Eco*RI	GCT TAT CTA GGA ATT CGA GCG AGA CAA CTG CCC GGC GG
pTTQ18	*Bam*HI	ATA GGA TCC TCA TCA TTC GTT GAC GTA GAC GCG CGG

Table 4.4: Primers for cloning S. coelicolor alr gene

4.2.11 Polymerase Chain Reaction (PCR)

PCR was used to isolate the racemase genes from their respective genomic DNA sources. The Pfu Turbo® Hotstart polymerase from Stratagene® was used for all

reactions. Deoxynucleoside triphosphates (dNTP's) were supplied from Sigma® containing 10 mM of each of the four deoxynucleotides, recommended amount 1 µl per 50 µl reaction volume.

4.2.11.1 PCR of putative racemase gene from *P. putida*

PCR reagents:

Template DNA	40 ng
Primer forw (10 µM)	3 µl
Primer rev (10 µM)	3 µl
dNTPs	3 µl
Reaction buffer (10x)	5 µl
Pfu Turbo® Hotstart polymerase	1 µl
DMSO	5 µl
dH_2O	to 50 µl

The reaction mixture was prepared on ice and the polymerase was added immediately before the PCR tube was transferred to the PCR machine. From a large PCR screen the following program was successful:

(1) 95°C	10 min
(2) 95°C	2 min
(3) 50°C	1 min 30
(4) 72°C	2 min 30
(5) Go to (2) rpt x30	
(6) 72°C	10 min
(7) 6°C	Hold

Results analysed by horizontal agarose gel electrophoresis, 4.2.1. (During PCR screen annealing temps of 50< and <60 produced bands of correct size)

4.2.11.2 PCR of putative alanine racemase from *S. coelicolor*

A screen of different concentrations of dNTP's/DNA/primers over a range of annealing temperatures resulted in identical reagents and conditions needed for amplification as those in section 4.2.7.2.

4.2.11.3 PCR of *dal* gene from *B. subtilis*

Initial PCR reactions were carried using colonies of *B. subtilis,* but was unsuccessful. Genomic DNA was prepared and PCR was succesful.

Template DNA	40 ng
Primer forw (10 μM)	3 μl
Primer rev (10 μM)	3 μl
dNTPs	2 μl
Reaction buffer (10x)	5 μl
Pfu Turbo® Hotstart polymerase	1 μl
DMSO	2.5 μl
dH_2O	to 50 μl

The same programme used was identical to that in section 4.7.2.2.

4.2.12 Creation of plasmids pJLM101, pJLM201 and pJLM301

The amplified *P. putida* racemase gene (1.2 kbp) was inserted into pINGE2 using *Spe*I and *Hind*III restriction sites to create pJLM101, (6.0 kbp). The amplified *S. coelicolor* putative racemase *alr* gene (1.2 kbp) was inserted into pTTQ18 using *Eco*RI and *Bam*HI restriction sites to create pJLM201 (5.7 kbp). The *B. subtilis* alanine racemase *dal* gene (1.2 kbp) was inserted into pTTQ18 using *Eco*RI and *Bam*HI cloning sites to create pJLM301 (5.7 kbp).

4.2.13 Sequencing PCR of pJLM101, pJLM201, and pJLM301

Sequencing primers (Sigma) were designed to align with the DNA sequence prior to the gene start sequence, then at approximately 300bp, 600 bp, and 900 bp from the start. The length of the sequence results allowed sufficient overlap to align. Reverse primers were also designed to copy the complementary strand at the same positions.

Vector	**Annealing position**	**Sequence 5'→ 3'**
pJLM101	Prior to start	GGCCTTCTGCAGTTCTGGCA
pJLM101	300 bp	TGCCCTGCGTGGCGGTGGC
pJLM101	600 bp	CAAGCTGGTCGCGCTGATGAC
pJLM101	900 bp	CCGTGGGCTATGACCGCACC
pJLM101	900 bp (redesigned)	CCGTCGGTTACGACCGCACC
pJLM201	Prior to start	GTTGACAATTAATCATCGGC
pJLM201	300 bp	TCATGTGCTGGCTGTGGACGC
pJLM201	600 bp	CGGGAGATGACGGCCTACGC
pJLM201	900 bp	GCTCGGCTACGCGGACGGCA
pJLM301	Prior to start	GTTGACAATTAATCATCGGC
pJLM301	300 bp	GGCAATCGCTGCTGAGTATGACG
pJLM301	600 bp	CCGCTGCCGTTAAAGAATCTAATGGTCC
pJLM301	900 bp	GGGAAAACGCCTGAAAATTGCCG

Plasmid DNA was prepared using *E.*coli Top10 cells, section 4.2.4. The reagents were prepared in a 0.2 ml PCR tube:

ABI Prism BigDye Terminator v3.1 Ready reaction mix	2 µl
Sequencing primer (1 pmol/µl)	3.2 µl
Template DNA	150-300 ng
dH_2O	up to 10 µl

PCR Programme:

(1)	95°C	1 min
(2)	96°C	30 s
(3)	50°C	30 s
(4)	60°C	4 min
(5)	Go to (2) rpt x30	
(6)	72°C	10 min
(7)	6°C	Hold

Samples were then taken to the SBS Sequencing Service, Ashworth Laboratories, at the University of Edinburgh.

4.2.14 Creation of Low Copy Number Oxidase Plasmids

4.2.14.1 PCR of pPOT3 heat inducible promoter

Forward (5' CGCCACTAGTACCTTGCCGATCAGCCAAACG) and reverse (5' GGCCTGCAGGGGATCTTCAGTCAGTCAGC) primers were purchased from Sigma Genosys to amplify the heat inducible promoter from pPOT3 and introduce *Spe*I and *Pst*I restriction sites (respectively, underlined) for cloning. The PCR was carried out as follows:

Template DNA	40 ng
Primer forw (10 μM)	3 μl
Primer rev (10 μM)	3 μl
dNTPs	1 μl
Reaction buffer (10x)	5 μl
Pfu Turbo® Hotstart polymerase	1 μl
dH_2O	to 50 μl

PCR programme:

(1)	95°C	10 min
(2)	95°C	1 min 30
(3)	50°C	1 min 30
(4)	72°C	1 min 30
(5)	Go to (2) rpt x30	
(6)	72°C	10 min
(7)	6°C	Hold

4.2.14.2 Creation of pJLM401

The PCR product (approximately 900 bp) was gel purified and digested with *Spe*I and *Pst*I. The vector pTrp200 was digested with the same enzymes and dephosphorylated. Ligation and subsequent transformation into Top10 cells produced colonies that were grown in 10 ml LB/Kan. DNA isolation and digest using *Spe*I and *Pst*I confirmed the creation of pJLM401.

4.2.14.3 Creation of pJLM402

The D-amino acid oxidase gene from *T. variablis* was donated by Ingenza in the plasmid pPOT3(DAAO). The primers used for oxidase amplification were also obtained from Ingenza and the PCR was carried out as follows:

Template DNA	40 ng
Primer forw (10 µM)	3 µl
Primer rev (10 µM)	3 µl
dNTPs	1 µl
Reaction buffer (10x)	5 µl
Pfu Turbo® Hotstart polymerase	1 µl
dH_2O	to 50 µl

PCR programme:

(1)	95°C	10 min
(2)	95°C	1 min
(3)	50°C	1 min 30
(4)	72°C	1 min 30
(5)	Go to (2) rpt x30	
(6)	72°C	10 min
(7)	6°C	Hold

The resulting PCR band was purified by excising from an agarose gel after electrophoresis. The insert and pJLM401 were digested using *Bam*HI and *Pst*I. pJLM401 was dephosphorylated. Ligation and transformation into Top10 cells afforded colonies that were grown in 10 ml LB/Kan. DNA isolation and digest using *Bam*HI and *Pst*I, and sequencing confirmed the creation of pJLM402.

The primers used for sequencing were purchased from Sigma Genosys,
*Bam*HI site prior to oxidase gene: ACGCAAGTTCTTAAGAAGGAGATATAC
Sequencing carried out as detailed in section 4.2.13.

4.2.14.4 From pGL12 to pJLM501 using Quikchange

The heat inducible promoter and oxidase gene were excised from pJLM402 as one length of DNA and placed into pTrp338 by Ingenza to create a low copy vector containing the oxidase.

To remove one *Bam*HI site from pGL12 to enable removal of the amino acid oxidase gene and replacement with any other gene using the other *Bam*HI site, the following primers were purchased from Sigma:

Forward GAGCTCGGTACCCGGCCATCCTTGAGGAAATACTTACCC and reverse GGGTAAGTATTTCCTCAAGGATGGCCGGGTACCGAGCTC. The Quikchange PCR reagents were mixed as follows, as described by the Stratagene Quikchange manual.

Template DNA	40 ng
Primer forw (final 125 ng)	10.5 µl
Primer rev (final 125 ng)	10.5 µl
dNTPs	1 µl
Reaction buffer (10x)	5 µl
Pfu Turbo® Hotstart polymerase	1 µl
dH_2O	to 50 µl

PCR programme:

(1) 95°C	30 s
(2) 95°C	30 s
(3) 55°C	1 min
(4) 68°C	7 min
(5) Go to (2) rpt x18	
(6) 68°C	10 min
(7) 6°C	Hold

*Dpn*I restriction enzyme (1µl) was added to the crude PCR product, the reaction was mixed by gentle pipetting and spun for 1 min at maximum speed in a micro-centrifuge. The digestion of methylated DNA was allowed to proceed at 37°C, 200 rpm for 1 h. 1 µl of this mixture was then added to *E. coli* XL1-blue chemically competent cells (Stratagene) on ice. After 30 min incubation on ice, the cells were heat shocked at 42°C for 45 seconds and placed back on ice for 2 min. Pre-warmed LB (500 µl) was added and incubated at 200 rpm, 37°C for 1 hour. The transformation mixture (100 µl) was

spread onto a LB agar plate containing kanamycin. Colonies picked the following day were used to inoculate 10 ml LB/Kan. Purification of the plasmid DNA and digest using *Bam*HI confirmed the removal of one of the *Bam*HI sites, and sequencing confirmed this.

The primers used for sequencing were purchased from Sigma Genosys,
*Bam*HI site prior to Oxidase gene: ACGCAAGTTCTTAAGAAGGAGATATAC
Area of *Bam*HI site removed by Quikchange: TCTGGCGGTGATAATGGTTGC
Sequencing carried out as detailed in section 4.2.13

4.2.15 Error-Prone PCR (epPCR) and QuikChange PCR

4.2.15.1 epPCR and preparation of primers for Quikchange

Diversify™ PCR Random Mutagenesis Kit from BD Biosciences (Catalogue number K1830-1) was purchased and the protocol followed to produce approximately 2.7, 4.8, and 8.1 mutations per kbp. epPCR was carried out on all cloned racemases. Primers were the same as those used in section 4.2.10 for initial racemase amplification. The conditions are shown:

Mutations per 1 kbp	2.7	4.8	8.1
PCR Grade Water	38 µl	35 µl	32 µl
10X Titanium *Taq* Buffer	5 µl	5 µl	5 µl
$MnSO_4$ (8 mM)	2 µl	4 µl	4 µl
dGTP (2 mM)	1 µl	2 µl	5 µl
50x Diversify dNTP mix	1 µl	1 µl	1 µl
Primer mix (10 µM each primer)	1 µl	1 µl	1 µl
Template DNA (1 ng/µl)	1 µl	1 µl	1 µl
TITANIUM *Taq* Polymerase	1 µl	1 µl	1 µl
Total Volume	50 µl	50 µl	50 µl

PCR programme:

(1)	94°C	30 s
(2)	94°C	30 s
(3)	68°C	1 min 30
(4)	Go to (2) rpt x25	
(5)	68°C	1 min
(7)	4°C	Hold

The PCR product was purified by gel electrophoresis and digested using the restriction sites introduced by the primers. After digestion, the PCR product/ Quikchange primers were blunt ended using Klenow.

epPCR fragment, digested and purified	60 µl
Klenow	2 µl
Buffer 2	7 µl
dNTP's 2 mM	0.9 µl

Alternatively, the Klenow was added to the digest without purification. This reaction mixture was left at room temperature for 20 min, and the DNA purified using the Qiagen PCR Cleanup kit. The resulting Quikchange PCR primers were used immediately or stored at -20°C.

4.2.15.2 Quikchange Whole Plasmid PCR to create mutant libraries

Template DNA to be used was always freshly prepared using BL21(DE3) *E. coli* cells (or using any cells that are *dam*+).

Template DNA	100 ng
Primers	125 ng
dNTP's (100 mM each)	0.5 µl
Pfu Buffer	5 µl
DMSO	5 µl
dH_2O	to a final volume of 49 µl
	Add 1 µl *Pfu* Hotstart

PCR programme:

(1) 95°C	30 s
(2) 95°C	30 s
(3) 55°C	1 min
(4) 68°C	8 min
(5) Go to (2). Rpt x 18	
(5) 68°C	10 min
(7) 4°C	Hold

The crude PCR product had *Dpn*I restriction enzyme (1µl) added and the reaction was mixed by gentle pipetting and spun for 1 min at maximum speed in a micro-centrifuge. The digestion of methylated DNA was allowed to proceed at 37°C, 200 rpm for 1 h. 1 µl of this mixture was then added to *E. coli* XL1-blue chemically competent cells (Stratagene), thawed on ice. After 30 min incubation on ice, the cells were heat shocked at 42°C for 45 s and placed back on ice for 2 min. Pre-warmed LB (500 µl) was added and incubated at 200 rpm, 37°C for 1 h. 2 x 200 µl of the transformation was spread onto LB agar plates containing antibiotic.

2 ml LB/Amp was added to each plate the following day and the colonies scraped from the agar with a plastic sterile loop. The library of variant DNA was isolated from the colonies using the Qiagen miniprep kit and stored at -20˚C.

4.3 Growth of *P. putida* for genomic DNA extraction

Glycerol stocks of *P.putida* ATCC 47054 were available at -80°C in the laboratory. Streaks were made onto Nutrient Broth/agar plates and incubated at 30°C overnight. Single colonies were used to inoculate 500 ml Nutrient broth and grown at 30°C, 200 rpm, in baffled flasks. Genomic DNA was extracted, section 4.2.3 and stored at 4°C.

4.4 Growth and Transformation of *E. coli*

4.4.1 Transformation of *E.coli* BL21(DE3) with plasmid DNA by heat shock

Batches of one shots (50 µl) competent cells were supplied by Invitrogen™. BL21(DE3) cells were used when transforming pET16b(racemase) vectors and BL21(DE3)pLysS and Top10 cells were used for all pINGE2 and pTTQ18 plasmids.

(1) Cells were thawed on ice. 1-2 µl of plasmid DNA was added and mixed by tapping the tube.

(2) After incubation on ice for 30 min, the cells were placed in a water bath at 42°C for 30 s exactly and quickly returned to the ice.

(3) After 2 min, 250 µl of pre-warmed SOC (42°C) was added and cells were left at 37°C, 200 rpm for 1 h.

(4) Cells were then plated onto pre-warmed (37°C) LB/Agar/Antibiotic and incubated at 37°C overnight.

When transforming libraries, 900 µl SOC was added, and 250 µl was spread onto larger LB plates.

4.4.2 Growth of *E. coli*

From the overnight plates, single colonies were picked and used to inoculate 10ml LB/Amp at 200 rpm, 37°C. These cultures were used for DNA analysis and manipulation. When larger culture volumes were required, 100 µl of the starter culture was placed into 50-250 ml LB/Amp. For protein analysis, the OD_{600} was measured until it was between 0.4 and 0.6 and protein expression was induced by the addition of IPTG (final concentration of 1 mM).

Cells were harvested by centrifugation; volumes of 1.5 ml at 13 000 rpm for 2 min using a desktop centrifuge, medium sized cultures (<50 ml) at 4000 g for 10 min using a Sorvall swing bucket rotor, and larger volumes (>50 ml) at 6000 g for 10 min. The supernatant was removed and the cell pellet resuspended in 100 mM phosphate buffer, pH 8.0. This was stored at –80°C.

4.4.3 Transformation of *E.coli* XL1-Red with plasmid DNA by heat shock

XL1-Red competent cells were supplied from Stratagene®.

Method 1

(1) 100 µl cells were thawed on ice. 1.7µl of β-mercaptoethanol was added to each shot (giving a final concentration of 25mM). 1-2 µl of plasmid DNA was added and mixed by tapping the tube.

(2) After incubation on ice for 30 min, the cells were placed in a water bath at 42°C for 45 s exactly and quickly returned to the ice.

(3) After 2 min, 950 µl of pre-warmed SOC (42°C) was added and cells were left at 37°C, 200 rpm for 1 h.

(4) 50 µl Cells were then plated onto pre-warmed (37°C) LB/Agar/Antibiotic and incubated at 37°C overnight.

(5) 9ml LB/ampicillin was added to the remaining transformation, and grown at 37°C overnight. 100 µl of overnight culture was used to inoculate a second 10 ml LB/ampicillin, and grown overnight. The sub-culturing was repeated up to 6 times. 3 ml

of final culture was harvested and the plasmid DNA extracted to produce the first library of variants.

Method 2 (from 'Methods in Molecular Biology, Vol 231, Directed Evolution Library Creation. Methods and Protocols)

(1) 50 µl Cells were thawed on ice. β-mercaptoethanol was added, giving a final concentration of 25 mM. After 10 min, 1-2 µl of plasmid DNA was added and mixed by tapping the tube.

(2) After incubation on ice for 30 min, the cells were placed in a water bath at 42°C for 45 s exactly and returned to the ice for 2 min.

(3) 1 ml of pre-warmed SOC (42°C) was added and cells were left at 37°C, 200 rpm for 1 h.

(4) 200 µl Cells were then plated onto pre-warmed (37°C) LB/Agar/Antibiotic and incubated at 37°C overnight.

(5) 2 ml LB/Amp was added to the plate the following day and the colonies scraped from the agar with a plastic sterile loop. This was added to 10 ml LB/Amp, and the OD_{600} measured. The OD was then diluted to approximately 0.0005 in 200 ml and grown overnight. When the OD of the culture reached 2-2.5 (after 12-24 h), the cells were again diluted to produce an OD600 of 0.0005. The sub-culturing was repeated up to 6 times. 3 ml of final culture was harvested and the plasmid DNA extracted to produce the library of variants.

4.4.4 Co-transformation of Racemase and Oxidase for screening

One shot Top10 cells from Invitrogen™ were used. The normal protocol for transformation was followed. However, when adding the plasmid DNA to the thawed cells, equal concentrations of racemase and oxidase DNA must be added, with a total volume added of 5 µl or less.

When transforming libraries, 900 µl SOC was added, and 250 µl aliquots were spread onto larger (15 cm diameter) LB plates.

4.5 Protein Purification and Analysis

Unless otherwise stated, all work was carried out at 4°C

4.5.1 Obtaining the CFE

BugBuster Protein Extraction Reagent (Novagen, 71370-4) was used. To 1 g (wet cell weight) cell pellet, 5 ml BugBuster reagent was added. 10 µl Lysonase (provided with BugBuster), 50 µl Protease inhibitor cocktail (Sigma, P8849), and 0.3 mM (final concentration) PMSF were added and the cell resuspended. The mixture was left spinning at room temperature for 10-15 min, and the cell debris was removed by centrifugation at 16000 rpm, 4°C for 20 min. The CFE was placed immediately at 4°C.

4.5.2 Ammonium sulfate precipitation

4.5.2.1 Removing the protein pellet up to 40 % saturation

Method 1

The CFE was placed at 4°C on a stirrer plate. Solid ammonium sulfate was added slowly to the stirring solution, to ensure complete dissolution after each addition. Once the desired percent saturation was reached, the sample was centrifuged at 4000 rpm for 30 minutes to obtain a '0-40 %' pellet. The supernatant was removed and subjected to a second ammonium sulfate precipitation

Method 2

To minimise areas of high concentration of ammonium sulfate salts, producing unnecessary precipitation, a saturated ammonium sulfate solution was prepared. To a stirring CFE solution at 4°C, saturated ammonium sulfate was added dropwise, until the

final saturation was 40 %. The sample was centrifuged and the supernatant removed for subsequent analysis, or further precipitation.

4.5.2.2 Obtaining the 40-70 % ammonium sulfate precipitated protein

To the stirring supernatant from 4.5.2.1, solid ammonium sulfate was added slowly until 70 % saturation had been reached. The sample was centrifuged at 4000 rpm for 30 min, and the supernatant discarded. The pellet was resuspended in the minimum amount of buffer, required for loading onto the proceeding column, and desalted using a PD10 column

4.5.3 PD10 desalting

Desalting of the freshly purified protein from the ammonium sulfate precipitation was carried out to prepare the protein for anion exchange chromatography. Disposable Amersham Bioscience PD-10 desalting columns were used (Amersham Biosciences). The column was equilibrated with 25 ml of the proceeding column equilibrating buffer. Exactly 2.5 ml of the protein sample was loaded onto the top of the column and the flow through discarded. 3.5 ml of the same buffer was applied to the top of the column and the flow through collected.

4.5.4 SDS-PAGE Analysis

Analysis of the purity of protein and expression samples was done using SDS PAGE. Pre-cast gels were purchased (BioRad, SDS 10-15% polyacrylamide gel) and either New England Biolabs markers (P77085) or BioRad Kaleidoscope protein markers (161-0375) were used to analyse protein weights. The molecular weights of the NEB standard proteins are 175, 83, 62, 47.5, 32.5, 25, 16.5 and 6.5 kDa, and the BioRad standards are 250, 150, 100, 75, 50, 37, 25, 20, 15, 10 kDa. 20 µl of the protein sample was mixed with 20 µl of the Laemmli loading buffer (BioRad, 161-0737), incubated at 95°C, and centrifuged at max speed in a desktop centrifuge for 1 min. 10-30 µl of protein sample was loaded onto the SDS PAGE gel, already submerged in SDS PAGE running buffer. The electrophoresis was carried out at 150 V for 90 min. The gel was placed into

staining solution for a minimum of 1 h (preferably overnight) and then de-stained to develop the bands.

4.5.5 Western Blot

A western blot was carried out to transfer protein for sequencing. An SDS gel was run as described in 4.5.4. However, prior to staining, the gel was washed thoroughly with dH_2O, and then with Blotting Buffer. The gel was placed into a semi-dry blotting apparatus, next to a polyvinylidiene fluoride (PVDF) membrane, and sandwiched between two 5 mm filter papers. The filter papers were prepared by soaking in HPLC-grade methanol for 5 min, followed by equilibration in blotting buffer for a further 5 min. The electrophoresis was carried out at 150 mA for 2 h. The PVDF membrane was stained for 20 s with Coomasie blue, and immediately placed into 50 % methanol to de-stain. The gel was also stained to assess efficiency of the protein transfer. Once de-stained, the PVDF membrane was air-dried and the band corresponding to the protein of interest was cut out. The sample was then sent for N-terminal sequencing, to Dr Jeff Keen at the Proteomics Facility, Institute of Membrane and Systems Biology LIGHT Laboratories, University of Leeds, Leeds, LS2 9JT.

4.5.6 BCA Protein Assay- Protein Quantification

To determine protein concentrations, a BCA Protein Assay kit from Pierce Biotechnology (23227) was used. A standard curve was produced for each assay as the colour development measured has no end point. Diluted albumin standards (BSA) were prepared at concentrations of 2, 1.5, 1.0, 0.75, 0.50, 0.25, 0.125, 0.025, and 0 mg/ml. In triplicate, 25 µl of each of the standards, and the unknown samples were added to 200 µl of the 'working reagent' supplied, in a micro-titre plate. The plate was mixed on the plate reader shaker for 30 s, and then covered and incubated at 37°C for 30 min. The plate was allowed to cool to room temperature, (approx 10 min), and the absorbance at 562 nm measured. Using the standards, a plot of concentration vs. absorbance was

made, and the absorbencies of the unknown samples were used to estimate the protein concentrations.

4.5.7 Akta Purification

All chromatography was carried out using an Akta Explorer from Amersham Biosciences. The Akta was housed in a chilled cabinet, therefore all runs were carried out at 4°C. When not in use, all Akta tubing contained 20 % ethanol. Weekly washes with 0.1 M NaOH were carried out to remove any residual protein and salts.

4.5.7.1 Ni^{2+} and Cu^{2+} affinity Chromatography

1 ml Hi-trap chelating HP columns were used (Amersham 17-0408-01). When not in use, columns were stored in 20 % ethanol at 4°C. To use, columns were manually washed with 5 ml dH_2O, charged with 5 ml metal salt solution, and washed with equilibration buffer (5 ml). The filtered CFE was loaded and the column attached to the Akta. The flow rate was set to 1 ml/min and the machine was used to create a linear gradient from 0 to 100 % elution buffer. The fraction content was recorded using UV absorbance at 280 and 425 nm (280 nm for general protein content, and 425 to analyse the PLP cofactor). Once the run was complete, the fractions of interest were assayed in the solution phase.

4.5.7.2 Anion Exchange Chromatography

A Hi-trap Ion exchange selection kit was tested, (Amersham 17-6002-33). After screening, 1 ml ResourceQ columns from Amersham Biosciences (17-1177-01) were used, containing a quaternary ammonium bonded phase. Columns were stored at 4°C.

To use, the column was attached to the machine and washed with 10 ml dH_2O at a flow rate of 2 ml/min. The column back pressure for the columns in the selection kit was set at 0.3 MPa. The column back pressure for the ResourceQ column was set at 1.5 MPa. Once washed, the columns were equilibrated with start buffer. The protein sample was filtered (0.2 μM) and loaded onto the column using the sample loop (2.5 ml, or 10 ml,

sample size ideally half of the loop capacity). The proteins were eluted with a gradient increasing to 100 % elution buffer (increasing salt gradient). The fractions were analysed by measuring absorbance at 280 and 425 nm. Fractions of interest were assayed using the solution phase assay. Once the run was finished, the column was washed with start buffer, then dH_2O, and finally 20 % ethanol.

4.5.7.3 Hydrophobic Interaction Chromatography

A Hi-Trap HIC selection kit was purchased from Amersham Biosciences (11-0034-53). After screening, 5 ml Hi-trap Phenyl Sepharose HP Columns were purchased (17-5195-01). When not in use, columns were stored in ethanol, at 4°C.

To use, the column was attached to the machine and washed with 10 ml dH_2O at a flow rate of 2 ml/min. The column back pressure for all columns was set at 0.3 MPa. Once washed, the columns were equilibrated with elution buffer (low salt), followed by start buffer (high salt). The protein sample was filtered (0.2 μM) and loaded onto the column using the sample loop (2.5 ml, or 10 ml, sample size ideally half of the loop capacity). The proteins were eluted with a gradient increasing to 100 % elution buffer (decreasing salt gradient). The fractions were analysed by measuring absorbance at 280 and 425 nm. Fractions of interest were assayed using the solution phase assay. Once the run was finished, the column was washed with elution buffer (low salt), then dH_2O, and finally 20 % ethanol.

4.6 Assays for Racemase Activity

4.6.1 The solution phase assay

4.6.1.1 Proof of principle assay

Alanine racemase and D-amino acid oxidase were obtained from Sigma, (Alanine racemase from *G. stearothermophilus*, A8936, D-amino acid oxidase from Porcine kidney, A1914). The stock 50 ml chromogenic assay mixture contained the following:

4-aminoantipyrine (4-AAP), 1 M	37.5 µl (final conc 0.75 mM)
2 % (w/v) 2,4,6-tribromo-3-hydroxybenzoic acid (TBHBA) in DMSO	500 µl (final conc 0.02 % (w/v))
Potassium phosphate buffer 1 M, pH 8.0	5 ml
Substrate, L-Alanine	10 mM
MilliQ H_2O	up to 50 ml

Assay mixture, racemase, oxidase, and horseradish peroxidase were then added to a 96 well plate and incubated at 37°C. The reaction was monitored by measuring the absorbance of the colorimetric product (ε = 29400 M^{-1}) at 510 nm with a plate reader (Molecular Devices, Versa Max tunable micro-titre plate reader).

4.6.1.2 Sensitivity of D-amino acid oxidase in the solution phase assay

According to 4.6.1.1, assay mixtures containing the following substrate concentrations were made:

L-Ala (mM)	10	9	8	7	6	5	4	3	2	1	0
D-Ala (mM)	0	1	2	3	4	5	6	7	8	9	10

Oxidase, HRP, and water were added to the wells to a final volume of 200 µl. The reaction was followed on the Molecular Devices micro-titre plate reader.

4.6.1.3 Determining activities of cells expressing racemase

The whole cells obtained by 4.2.2 were used for the assay to evaluate potential racemase activity. To determine if any racemase activity was present, 180 µl of the assay mix was added to single wells in a U-shaped 96-well micro-titre plate containing,

Horse radish peroxidase, 1mg/ml	5 µl
D-amino acid oxidase	5 µl (to 0.2 U per well)

Whole cells in phosphate buffer	10 µl

The plate was then incubated at 37°C and the reaction monitored by measuring the absorbance of the colorimetric product (ε = 29400 M^{-1}) at 510 nm with a plate reader (Molecular Devices, Versa Max tunable micro-plate reader). Time points were taken every 30 s over a period of 1 h 30 min and absorbance plotted as a function of time.

4.6.1.4 Optimisation of the solution phase assay

A smaller amount of assay solution of 2x concentrate was used to minimise pipetting errors and ensure stock solution was re-made frequently.

4-aminoantipyrine (4-AAP), 1 M	15 µl (final concentration 0.75 mM)
2 % (w/v) 2,4,6-tribromo-3-hydroxybenzoic acid (TBHBA) in DMSO	200 µl (final conc 0.02% (w/v))
Potassium phosphate buffer 1 M, pH 8.0	2 ml
Substrate, L-Alanine	20 mM
MilliQ H_2O	up to 10 ml

Use 100 µl per well in the assay. The amount of HRP (1 mg/ml) used was increased to 10 µl, and the D-amino acid oxidase (purchased from Sigma) was made up to 20 U/ml (0.2 U/10 µl, according to the activity reported by the vendors). The remaining 80 µl is available for whole cell suspensions/CFE/water, allowing for minimal significant pipetting errors.

4.6.2 The solid phase assay

4.6.2.1 Small scale screening of one enzyme

After transformation and cell recovery (section 4.4.1) the cells were plated onto 7.5 cm diameter nitrocellulose membranes laid over LB/Amp/Agar plates. These were incubated overnight at 37°C to allow colonies to form. The membrane was removed

from the plates and placed at –20°C for a minimum of 24 h to allow partial cell lysis. An assay mix was prepared to the following recipe (recipe is for four plates):

1 x 3,3'-diaminobenzidine tablet (Sigma®)	
Potassium phosphate buffer 1 mM	40 ml
Screening substrate (usually L-amino acid)	10 mM
Horse radish peroxidase (10 µl per plate of 4 mg/ml solution)	40 µl
Porcine kidney D-amino acid oxidase (1 U per plate)	4 U

4 x 10 ml volumes of molten 2 % agarose (w/v in dH_2O) at approximately 55°C were prepared. To each 10 ml of 2 % agarose, 10 ml of assay mix was added and poured over membrane. The plates were left at room temperature to allow any colour to develop.

4.6.2.2 Screening cells containing the racemase and the oxidase

After transformation of library DNA, 250 µl recovered cells were spread onto 15 cm diameter nitrocellulose membranes on LB/Amp/Kan. These were incubated overnight at 30°C. The following day, the membrane was lifted from the LB/antibiotic plate, laid on an LB/antibiotic/IPTG containing plate and incubated overnight at 37°C to induce expression of the racemase and the oxidase. The membrane was then placed at -20°C overnight to partially lyse the cells. The assay mix was prepared as follows:

To cover 2 plates of 15 cm diameter

1 x 3,3'-diaminobenzidine tablet (Sigma®)	
Potassium phosphate buffer 1 mM	50 ml
Screening substrate (usually L-amino acid)	10 mM
Horse radish peroxidase (50 µl per plate of 1 mg/ml solution)	100 µl

2 x 25 ml volumes of molten 2 % agarose (w/v in dH_2O) at approximately 55°C were prepared. To each aliquot of 2 % agarose, 25 ml of assay mix was added and poured over membrane. The plates were left at room temperature to allow any colour to

develop. Hits were picked and placed into 10 ml LB/Amp/Kan and left to grow overnight at 37°C.

4.6.2.3 Re-assay of Co-transformed hits

Hits from overnight cultures were mini-prepped to obtain a mixture of oxidase and racemase DNA. Initially the separation of the two plasmids was carried out by gel electrophoresis, followed by excision of the smaller band from the gel, corresponding to the racemase, and purification. After optimisation, the DNA mixture was subjected to a digest with SpeI and NdeI, cutting the oxidase only, 4 times.

The purified racemase was then co-transformed using fresh oxidase to confirm that novel activity was not due to the former oxidase.

4.6.3 HPLC analysis of 2-ABA Racemisation

All HPLC methods/sample preparation/derivatisation were from Ingenza

The instrument used was an Agilent 1200 series HPLC comprising a vacuum degasser, quaternary pump, auto-sampler, thermo-stated column compartment and diode array detector (DAD). The column used was a Phenomenex HyperClone ODS (C18) 5u 150 x 4.6 mm, at 25°C. The flow rate was 1 ml/min. The gradient was isocratic, and detection measured at 338 nm.

4.6.3.1 Preparation of OPA/*N*-isobutyryl-L-cysteine reagent

Phthaldialdehyde (0.05 g) and *N*-isobutyryl-L-cysteine (0.110 g) were dissolved in 0.5 ml methanol. The final volume was made up to 5 ml using 0.4 N potassium borate buffer, pH 10. This solution was used immediately.

4.6.3.2 Column equilibration

The column was washed for 30 min with 20 % 10 mM potassium phosphate buffer pH 7.0 (Mobile phase A), 80 % acetonitrile (Mobile phase B), followed by equilibration with 84 % A and 16 % B. Column temperature was 25°C.

4.6.3.3 2-Amino butyric acid (2-ABA) racemisation

A 0.5 g pellet of cells expressing racemase was resuspended in 10 ml 10 mM potassium phosphate buffer. 12 mM (*R*)-amino butyric acid (2-ABA) or 12 mM (*S*)-2-ABA was added. Immediately a 1 ml sample was removed. The reaction was placed at 37°C for 15 h and a second 1 ml sample removed.

4.6.3.4 Sample preparation for HPLC analysis

A 1ml sample was heated at 95°C for 10 min, and then centrifuged at 13 000 rpm for 2 min. 250 µl of the supernatant was added to 750 µl HPLC grade water. The solution was then filtered through a 0.2 µM filter into an HPLC vial

4.6.3.5 Derivatisation protocol

Using the Agilent 1200, 16 µl 0.4 N potassium borate buffer, pH 10 was drawn into the mixing loop, followed by 16 µl of the OPA/*N*-isobutyryl-L-cysteine reagent. 8 µl of the 2-ABA-containing sample was drawn into the loop, and mixed for 6 cycles. The mixture was then held in the loop for 2 min before injection onto the column. Retention time for (*S*)-2-ABA was 4.2 min, for (*R*)-2-ABA 5.5 min, and for the OPA reagent 10.8 min at 338 nm.

1. Muller, M., Chemical diversity through biotransformations. *Current Opinion in Biotechnology* **2004,** 15, 591-598.
2. Schnell, B.; Faber, K.; Kroutil, W., Enzymatic racemisation and its application to synthetic biotransformations. *Advanced Synthesis & Catalysis* **2003,** 345, (6-7), 653-666.
3. Ishiwata, K. I.; Fukuhara, N.; Shimada, M.; Makiguchi, N.; Soda, K., Enzymatic Production of L-Tryptophan from Dl-Serine and Indole by a Coupled Reaction of Tryptophan Synthase and Amino-Acid Racemase. *Biotechnology and Applied Biochemistry* **1990,** 12, (2), 141-149.
4. Bommarius, A. S.; Riebel, B. R., *Biocatalysis*. 1 ed.; Wiley-VCH Verlag GmbH & Co.: Weinheim, 2004.
5. Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Kesseler, M.; Sturmer, R.; Zelinski, T., Industrial Methods of the Production of Optically Active Intermediates. *Angew. Chem. Int. Ed.* **2004,** 43, 788-824.
6. Vicenzi, J. T.; Zmijewski, M. J.; Reinhard, M. R.; Landen, B. E.; Muth, W. L.; Marler, P. G., Large-scale stereoselective enzymatic ketone reduction with *in situ* product removal via polymeric adsorbent resins. *Enzyme and Microbial Technology* **1997,** 20, (7), 494-499.
7. Patel, R. N., Biocatalysis: Synthesis of chiral intermediates for drugs. *Current Opinion in Drug Discovery and Development* **2006,** 9, (6), 741-764.
8. Kim, Y. W.; Lee, S. S.; Warren, R. A. J.; Withers, S. G., Directed evolution of a glycosynthase from Agrobacterium sp increases its catalytic activity dramatically and expands its substrate repertoire. *Journal of Biological Chemistry* **2004,** 279, (41), 42787-42793.
9. Reetz, M. T.; Carballeira, J. D.; Peyralans, J.; Hobenreich, H.; Maichele, A.; Vogel, A., Expanding the Substrate Scope of Enzymes: Combining Mutations Obtained by CASTing. *Chemistry: A European Journal* **2006,** 12, 6031-6038.
10. Carr, R.; Alexeeva, M.; Enright, A.; Eve, T. S. C.; Dawson, M. J.; Turner, N. J., Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity. *Angewandte Chemie-International Edition* **2003,** 42, (39), 4807-4810.
11. Yokoigawa, K.; Okubo, Y.; Soda, K.; Misono, H., Improvement in thermostability and psychrophilicity of psychrophilic alanine racemase by site-directed mutagenesis. *Journal of Molecular Catalysis B-Enzymatic* **2003,** 23, (2-6), 389-395.
12. Seebeck, F. P.; Hilvert, D., Conversion of a PLP-dependent racemase into an aldolase by a single active site mutation. *Journal of the American Chemical Society* **2003,** 125, (34), 10158-10159.
13. Williams, G. J.; Nelson, A. S.; Berry, A., Directed evolution of enzymes for biocatalysis and the life sciences. *Cellular and Molecular Life Sciences* **2004,** 61, (24), 3034-3046.
14. Yow, G. Y.; Watanabe, A.; Yoshimura, T.; Esaki, N., Conversion of the catalytic specificity of alanine racemase to a D-amino acid aminotransferase activity by a double active-site mutation. *Journal of Molecular Catalysis B-Enzymatic* **2003,** 23, (2-6), 311-319.

15. Arnold, F. H., Combinatorial and Computational Challenges for Biocatalyst Design. *Nature* **2001,** 409, 253-257.
16. Petrounia, I. P.; Arnold, F. H., Designed Evolution of Enzymatic Properties. *Current Opinion in Biotechnology* **2000,** 11, (4), 325-330.
17. Otten, L. G.; Quax, W. J., Directed Evolution: Selecting today's biocatalysts. *Biomolecular Engineering* **2005,** 22, 1-9.
18. Lutz, S.; Patrick, W. M., Novel Methods for Directed Evolution of Enzymes: Quality, not Quantity. *Current Opinion in Biotechnology* **2004,** 15, (4), 291-297.
19. Jaeger, K.-E.; Eggert, T., Enantioselective Biocatalysis Optimized by Directed Evolution
Current Opinion in Biotechnology **2004,** 15, (4), 305-313.
20. Stratagene *PfuTurbo® Hotstart DNA Polymerase Instruction Manual.*
21. Cadwell, R. C.; Joyce, G. F., Mutagenic PCR. *PCR Methods and Applications* **1994,** 3, (6), 5136-5140.
22. Neylon, C., Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Research* **2004,** 32, (4), 1448-1459.
23. Stratagene, Quikchange Site-directed Mutagenesis kit. In.
24. Reetz, M. T.; Torre, C.; Elpper, A.; Lohmer, R.; Hermes, M.; Brunner, B.; Maichale, A.; Bocola, M.; Arand, M.; Cronin, A.; Genzel, Y.; Archelas, A.; Furstoss, R., Enhancing the enantioselectivity of an epoxide hydrolase by directed evolution. *Organic Letters* **2004,** 6, (2), 177-180.
25. Stratagene http://www.stratagene.com/manuals/200129.pdf.
26. Alexeeva, M.; Carr, R.; Turner, N. J., Directed evolution of enzymes: new biocatalysts for asymmetric synthesis. *Organic & Biomolecular Chemistry* **2003,** 1, (23), 4133-4137.
27. Alexeeva, M.; Enright, A.; Dawson, M. J.; Mahmoudian, M.; Turner, N. J., Deracemization of alpha-methylbenzylamine using an enzyme obtained by in vitro evolution. *Angewandte Chemie-International Edition* **2002,** 41, (17), 3177-+.
28. Clouthier, C. M.; Kayser, M. M.; Reetz, M. T., Designing new Baeyer-Villiger monooxygenases using restricted CASTing. *Journal of Organic Chemistry* **2006,** 71, (22), 8431-8437.
29. Reetz, M. T.; Wang, L. W.; Bocola, M., Directed evolution of enantioselective enzymes: Iterative cycles of CASTing for probing protein-sequence space. *Angewandte Chemie-International Edition* **2006,** 45, (8), 1236-1241.
30. Stemmer, W. P. C., Rapid Evolution of a Protein *In vitro* by DNA Shuffling. *Nature* **1994,** 370, 389-391.
31. Ju, J.; Misano, H.; Ohnishi, K., Directed evolution of bacterial alanine racemases with higher expression level. *Journal of Bioscience and Bioengineering* **2005,** 100, (3), 246-254.
32. Reymond, J.-L.; Babiak, P., Screening Systems. In *White Biotechnology*, Springer-Verlag: Berlin Heidelberg, 2006; Vol. 105, pp 31-58.
33. Turner, N. J., Agar Plate-based Assays. In *Enzyme Assays*, Reymond, J. L., Ed. WILEY-VCH Weinheim, Germany, 2006; pp 139-161.

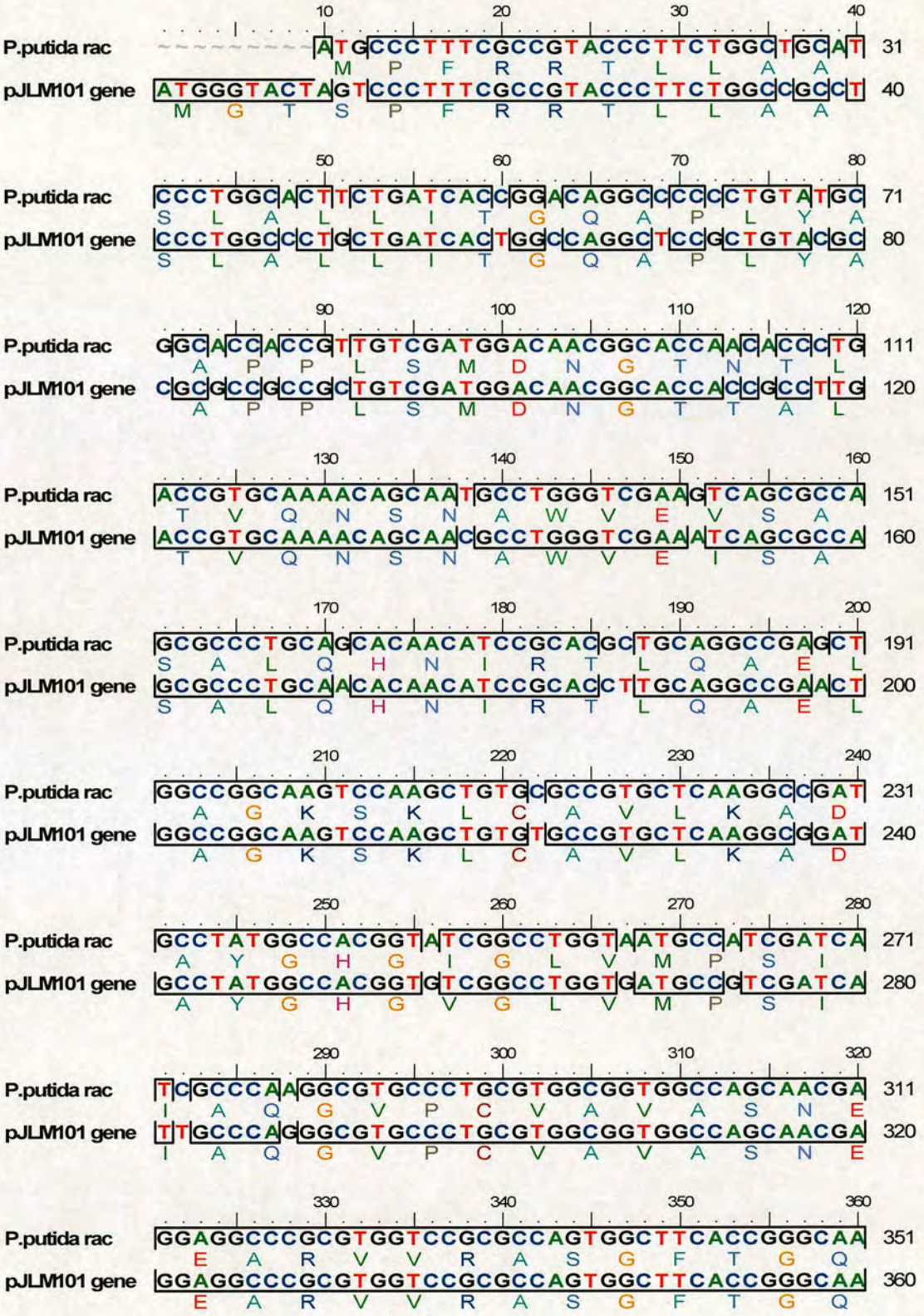
34. Olsen, M.; Iverson, B.; Georgiou, G., High-throughput screening of enzyme libraries. *Current Opinion in Biotechnology* **2000,** 11, 331-337.
35. *Directed Molecular Evolution of Proteins.* Wiley-VCH Verlag GmbH: Weinheim, 2002.
36. Watts, A. B.; Beecher, J.; Whitcher, C. S.; Littlechild, J. A., A method for screening Baeyer-Villiger monooxygenase activity against monocylcic ketones. *Biocatalysis and Biotransformation* **2002,** 20, (3), 209-215.
37. Bornscheuer, U. T.; Altenbuchner, J.; Meyer, H. H., *Biotechnology and Bioengineering* **1998,** 58, 554-559.
38. Pohn, B.; Gerlach, J.; Schneideler, M.; Katz, H.; Uray, M.; Bischof, H.; Klimant, I.; Schwab, H., Micro-colony array based high-throughput platform for enzyme library screening. *Journal of Biotechnology* **2007,** 129, 162-170.
39. Cowan, D. A.; Daniel, R. M., *Journal of Biochemical and Biophysical Methods* **1982,** 6, 31-37.
40. Reetz, M. T., Combinatorial and evolution-based methods in the creation of enantioselective catalysts. *Angewandte Chemie-International Edition* **2001,** 40, (2), 284-310.
41. Ten, L. N.; Im, W.-T.; Kim, M.-K.; Kang, M. S.; Lee, S.-T., Development of a plate technique for screening of polysaccharide-degrading micoorganisms by using a mixture of insoluble chromogenic substrates. *Journal of Microbiological Methods* **2004,** 56, 375-382.
42. Parikh, M. R.; Matsumura, I., Site-saturation mutagenesis is more efficient that DNA shuffling for the directed evolution of b-fucosidase from b-galatosidase. *Journal of Molecular Biology* **2005,** 352, 621-628.
43. Mayer, C.; Jakeman, D. L.; Mah, M.; Karjala, G.; Gal, L.; Warren, R. A. J.; Withers, S. G., Directed evolution of new glycosynthases from Agrobacterium beta-glucosidase: a general screen to detect enzymes for oligosaccharide synthesis. *Chemistry & Biology* **2001,** 8, (5), 437-443.
44. Kato, D. I.; Miyamoto, K.; Ohta, H., Microbial deracemization of alpha-amino acids. *Journal of Molecular Catalysis B-Enzymatic* **2005,** 32, (4), 157-165.
45. Krammer, B.; Rumbold, K.; Tschemmernegg, M.; Pochlauer, P.; Schwab, H., A novel screening method for hydroxynitrile lyase activity suitable for highthrouput screening. *Journal of Biotechnology* **2007,** 129, 151-161.
46.http://www.verenium.com/Pages/Technology/EnzymeTech/TechEnyScreening.html
47. Baumann, M.; Sturmer, R.; Bornscheuer, U. T., A high-throughput-screening method for the identification of active and enantioselective hydrolases. *Angew. Chem. Int. Ed.* **2001,** 40, (22), 4201-4204.
48. Abato, P.; Seto, C. T., EMDee: An enzymatic method for determining enantiomeric excess. *Journal of the American Chemical Society* **2001,** 123, (37), 9206-9207.
49. Li, Z.; Butikofer, L.; Witholt, B., High-throughput measurement of the enantiomeric excess of chiral alcohols by using two enzymes. *Angewandte Chemie-International Edition* **2004,** 43, (13), 1698-1702.

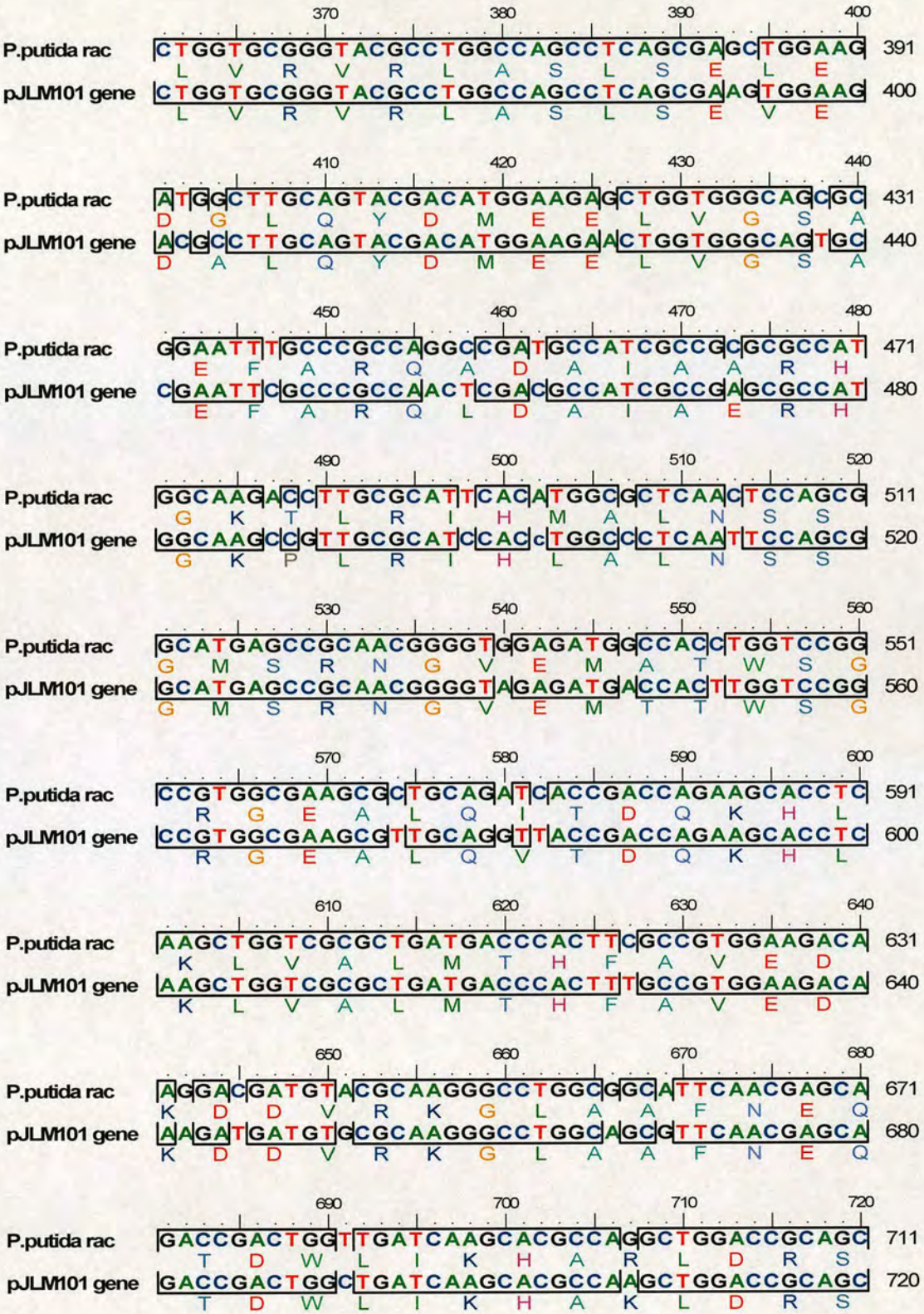
50. Sicard, R.; Chen, L. S.; Marsaioli, A. J.; Reymond, J.-L., A Fluorescence-based assay for Baeyer-Villiger Monooxygenases, hydroxylases and lactonases. *Advances in Synthesis and Catalysis* **2005,** 347, 1041-1050.
51. Yang, Y.; Babiak, P.; Reymond, J.-L., Low background FRET-substrates for lipases and esterases suitbale for high-throughput screening under basic (pH11) conditions. *Organic and Biomolecular Chemistry* **2006,** 4, 1746-154.
52. Onaran, M. B.; Seto, C. T., Using a lipase as a high-throughput screening method for measuring the enantiomeric excess of allylic acetates. *Journal of Organic Chemistry* **2003,** 68, (21), 8136-8141.
53. Hwang, B.-Y.; Kim, B.-G., High-throughput screening method for the identification of active and enantioselective w-transaminases. *Enzyme and Microbial Technology* **2004,** 34, 429-436.
54. Humphrey, C. E.; Easson, M. A. M.; Turner, N. J., Dimedone esters as novel hydrolase substrates and the application in the colorimetric detection of lipase and esterase activity. *ChemBioChem* **2004,** 5, 1144-1148.
55. Berkowitz, D. B.; Bose, M.; Choi, S., In situ enzymatic screening (ISES): A tool for catalyst discovery and reaction development. *Angewandte Chemie-International Edition* **2002,** 41, (9), 1603-1607.
56. Babiak, P.; Reymond, J.-L., A high-throughput, low-volume enzyme assay on solid support. *Analytical Chemistry* **2005,** 77, (2), 373-377.
57. Grognux, J.; Reymond, J.-L., A red-fluorescent microarray for lipase fingerprinting. *Molecular Biosystems* **2006,** 2, 492-498.
58. Belder, D.; Ludwig, M.; Wang, L.-W.; Reetz, M. T., Enantioselective catalysis and analysis on a chip. *Angew. Chem. Int. Ed.* **2006,** 45, 2463-2466.
59. Kelly, B. T.; Baret, J.-C.; Taly, V.; Griffiths, A. D., Miniaturising chemistry and biology in microdroplets. *Chemical Communications* **2007**, 1773-1788.
60. Krebs, G.; Hugonet, L.; Sutherland, J. D., Substrate ambiguity and catalytic promiscuity within a bacterial proteome probed by an easy phenotypic screen for aldehydes. *Angewandte Chemie-International Edition* **2006,** 45, (2), 301-305.
61. Dixon, M.; Kleppe, K., D-Amino Acid Oxidase .2. Specificity Competitive Inhibition and Reaction Sequence. *Biochimica Et Biophysica Acta* **1965,** 96, (3), 368-&.
62. Komeda, H.; Ishikawa, N.; Asano, Y., Enhancement of the thermostability and catalytic activity of D-stereospecific amino-acid amidase from *Ochrobactrum anthropi* SV3 by directed evolution. *Journal of Molecular Catalysis B: Enzymatic* **2003,** 21, 283-290.
63. Ono, K.; Yanagida, K.; Oikawa, T.; Ogawa, T.; Soda, K., Alanine racemase of alfalfa seedlings (Medicago sativa L.): First evidence for the presence of an amino acid racemase in plants. *Phytochemistry* **2006,** 67, (9), 856-860.
64. Luo, H.; Li, Q.; Yu, H. M.; Shen, Z. Y., Construction and application of fusion proteins of D-amino acid oxidase and glutaryl-7-aminocephalosporanic acid acylase for direct bioconversion of cephalosporin C to 7-aminocephalosporanic acid. *Biotechnology Letters* **2004,** 26, (11), 939-945.

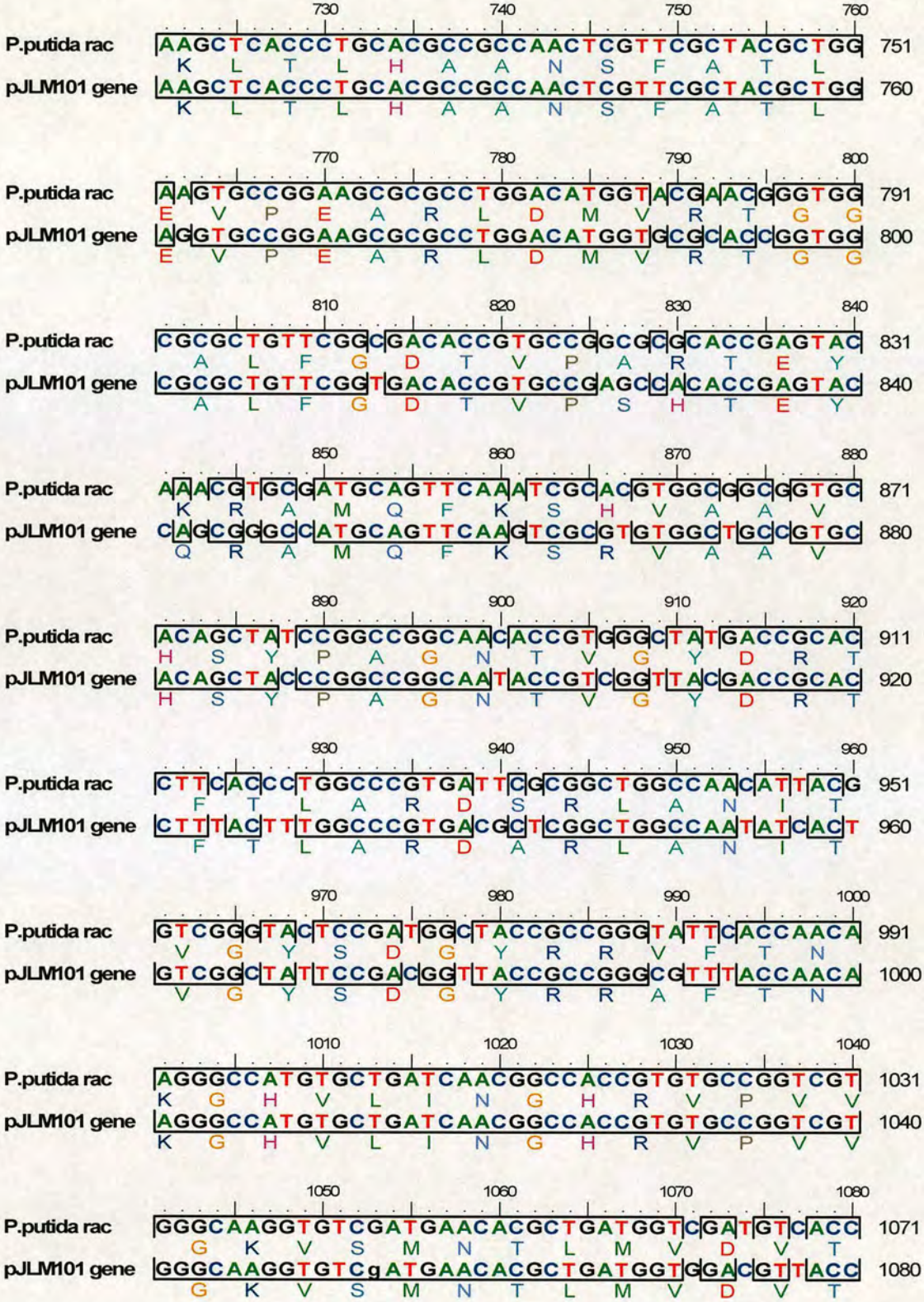
65. Luo, H.; Yu, H. M.; Li, Q.; Shen, Z. Y., Cloning and co-expression of D-amino acid oxidase and glutaryl-7-aminocephalosporanic acid acylase genes in Escherichia coli. *Enzyme and Microbial Technology* **2004,** 35, (6-7), 514-518.
66. Ferrari, E.; Henner, D. J.; Yang, M. Y., Isolation of an Alanine Racemase Gene from Bacillus-Subtilis and Its Use for Plasmid Maintenance in B-Subtilis. *Bio-Technology* **1985,** 3, (11), 1003-1007.
67. Palumbo, E.; Favier, C. F.; Deghorain, M.; Cocconcelli, P. S.; Grangette, C.; Mercenier, A.; Vaughan, E. E.; Hols, P., Knockout of the alanine racemase gene in *Lactobacillus planatarum* results in septation defects and cell wall perforation. *FEMS Microbiology Letters* **2004,** 233, (1), 131-138.
68. Hols, P.; Defrenne, C.; Ferain, T.; Derzelle, S.; Delplace, B.; Delcour, J., The alanine racemase gene is essential for growth of Lactobacillus plantarum. *Journal of Bacteriology* **1997,** 179, (11), 3804-3807.
69. Wolosker, H.; Sheth, K. N.; Takahashi, M.; Mothet, J.; jr, R. O. B.; Ferris, C. D.; Snyder, S. H., Purification of serine racemase: biosynthesis of the neuromodulator D-serine. *Proc. Natl. Acad. Sci USA* **1999,** 96, 721-725.
70. Uo, T.; Ueda, M.; Nishiyama, T.; Yoshimura, T.; Esaki, N., Purification and characterization of alanine racemase from hepatopancreas of black-tiger prawn, Penaeus monodon. *Journal of Molecular Catalysis B-Enzymatic* **2001,** 12, (1-6), 137-144.
71. Yoshimura, T.; Esaki, N., Amino acid racemases: Functions and mechanisms. *Journal of Bioscience and Bioengineering* **2003,** 96, (2), 103-109.
72. Y.-H. Lim, K. Y., N. Esaki, and K. Soda, A New Amino Acid Racemase with Threonine a-Epimerase Activity from *Pseudomonas putida*: Purification and Characterisation. *J. Bacteriol.* **1993,** 175, (13), 4213-4217.
73. Adams, E.; Mukherje.Kl; Dunathan, H. C., Alanine Racemase of Pseudomonas - Observations on Substrate and Inhibitor Specificity. *Archives of Biochemistry and Biophysics* **1974,** 165, (1), 126-132.
74. Asano, Y.; Endo, K., Amino-Acid Racemase with Broad Substrate-Specificity, Its Properties and Use in Phenylalanine Racemization. *Applied Microbiology and Biotechnology* **1988,** 29, (6), 523-527.
75. Sun, S. X.; Toney, M. D., Evidence for a two-base mechanism involving tyrosine-265 from arginine-219 mutants of alanine racemase. *Biochemistry* **1999,** 38, (13), 4058-4065.
76. Fenn, T. D.; Holyoak, T.; Stamper, G. F.; Ringe, D., Effect of a Y265F mutant on the transamination-based cycloserine inactivation of alanine racemase. *Biochemistry* **2005,** 44, (14), 5317-5327.
77. Watababe, A.; Yoshimura, T.; Lim, Y. H.; Kurokawa, Y.; Soda, K.; Esaki, N., Stereochemistry of the hydrogen abstraction from pyridoxamine phosphate catalyzed by alanine racemase of *Bacillus stearothermophilus*. *Journal of Molecular Catalysis B: Enzymatic* **2001,** 12, 145-150.
78. Watanabe, A.; Yoshimura, T.; Lim, Y. H.; Kurokawa, Y.; Soda, K.; Esaki, N., Stereochemistry of the hydrogen abstraction from pyridoxamine phosphate catalyzed by alanine racemase of Bacillus stearothermophilus. *Journal of Molecular Catalysis B-Enzymatic* **2001,** 12, (1-6), 145-150.

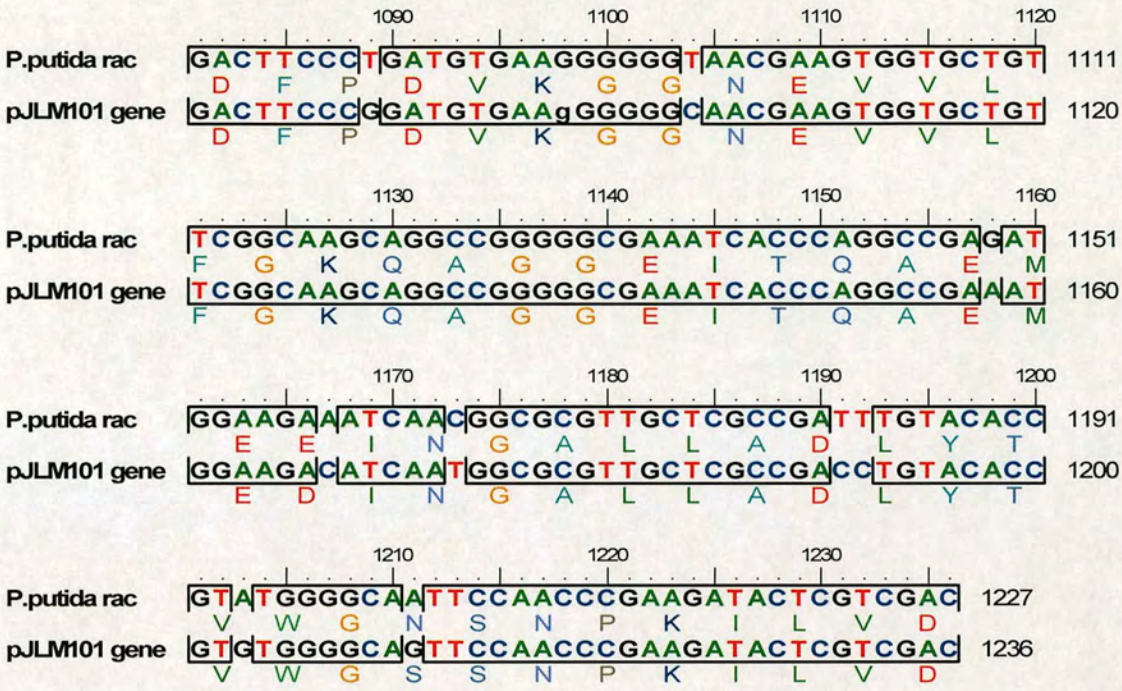
79. Faraci, W. S.; Walsh, C. T., Racemization of Alanine by the Alanine Racemases from Salmonella-Typhimurium and Bacillus-Stearothermophilus - Energetic Reaction Profiles. *Biochemistry* **1988,** 27, (9), 3267-3276.
80. Fenn, T. D.; Stamper, G. F.; Morollo, A. A.; Ringe, D., A side reaction of alanine racemase: Transamination of cycloserine. *Biochemistry* **2003,** 42, (19), 5775-5783.
81. Lim, Y. H.; Yoshimura, T.; Kurokawa, Y.; Esaki, N.; Soda, K., Nonstereospecific transamination catalyzed by pyridoxal phosphate-dependent amino acid racemases of broad substrate specificity. *Journal of Biological Chemistry* **1998,** 273, (7), 4001-4005.
82. Mustata, G.; Briggs, J. M., Cluster analysis of water molecules in alanine racemase and their putative structural role. *Protein Engineering Design & Selection* **2004,** 17, (3), 223-234.
83. Major, D. T.; Gao, J. L., A combined quantum mechanical and molecular mechanical study of the reaction mechanism and alpha-amino acidity in alanine racemase. *Journal of the American Chemical Society* **2006,** 128, (50), 16345-16357.
84. Major, D. T.; Nam, K.; Gao, J. L., Transition state stabilization and alpha-amino carbon acidity in alanine racemase. *Journal of the American Chemical Society* **2006,** 128, (25), 8114-8115.
85. Watanabe, A.; Yoshimura, T.; Mikami, B.; Hayashi, H.; Kagamiyama, H.; Esaki, N., Reaction mechanism of alanine racemase from Bacillus stearothermophilus - X-ray crystallographic studies of the enzyme bound with N-(5 '-phosphopyridoxyl)alanine. *Journal of Biological Chemistry* **2002,** 277, (21), 19166-19172.
86. Watanabe, A.; Yoshimura, T.; Mikami, B.; Esaki, N., Tyrosine 265 of alanine racemase serves as a base abstracting alpha-hydrogen from L-alanine: The counterpart residue to lysine 39 specific to D-alanine. *Journal of Biochemistry* **1999,** 126, (4), 781-786.
87. Watanabe, A.; Kurokawa, Y.; Yoshimura, T.; Esaki, N., Role of tyrosine 265 of alanine racemase from Bacillus stearothermophilus. *Journal of Biochemistry* **1999,** 125, (6), 987-990.
88. Ondrechen, M. J.; Briggs, J. M.; McCammon, J. A., A model for enzyme-substrate interaction in alanine racemase. *Journal of the American Chemical Society* **2001,** 123, (12), 2830-2834.
89. Noda, M.; Matoba, Y.; Kumagai, T.; Sugiyama, M., Structural evidence that alanine racemase from a D-cycloserine-producing microorganism exhibits resistance to its own product. *Journal of Biological Chemistry* **2004,** 279, (44), 46153-46161.
90. Lim, Y. H.; Yokoigawa, K.; Esaki, N.; Soda, K., A New Amino-Acid Racemase with Threonine Alpha-Epimerase Activity from Pseudomonas-Putida - Purification and Characterization. *Journal of Bacteriology* **1993,** 175, (13), 4213-4217.
91. Kurokawa, Y.; Watanabe, A.; Yoshimura, T.; Esaki, N.; Soda, K., Transamination as a side-reaction catalyzed by alanine racemase of Bacillus stearothermophilus. *Journal of Biochemistry* **1998,** 124, (6), 1163-1169.
92. Watanabe, A.; Kurokawa, Y.; Yoshimura, T.; Kurihara, T.; Soda, K.; Esaki, N., Role of lysine 39 of alanine racemase for Bacillus stearothermophilus that binds

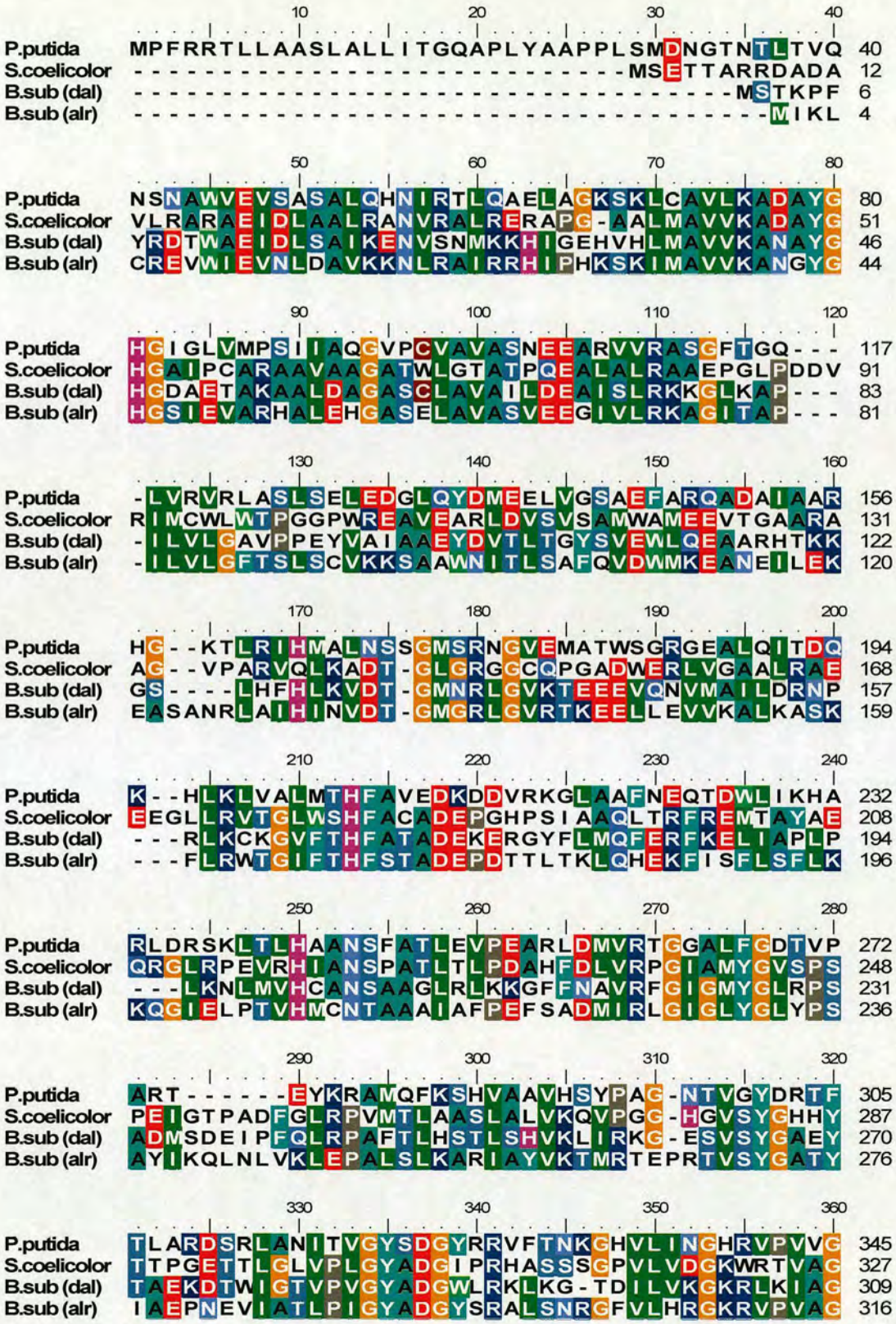
pyridoxal 5 '-phosphate. Chemical rescue studies of Lys(39) -> Ala mutant. (vol 274, pg 4189, 1999). *Journal of Biological Chemistry* **1999,** 274, (42), 30334-30334.
93. Patrick, W. M.; Weisner, J.; Blackburn, J. M., Site-directed mutagenesis of Tyr354 in Geobacillus stearothermophilus alanine racemase identifies a role in controlling substrate specificity and a possible role in the evolution of antibiotic resistance. *Chembiochem* **2002,** 3, (8), 789-792.
94. Henderickx, H. J. W.; Duchateau, A. L. L.; Raemakers-Franken, P. C., Chiral liquid chromatography-mass spectrometry for high-throughput screening of enzymatic racemase activity. *Journal of Chromatography A* **2003,** 1020, (1), 69-74.
95. Schonfeld, D. L.; Bornscheuer, U. T., Polarimetric assay for the medium-throughput determination of alpha-amino acid racemase activity. *Analytical Chemistry* **2004,** 76, (4), 1184-1188.
96. Noda, M.; Matoba, Y.; Kumagai, T.; Sugiyama, M., A novel assay method for an amino acid racemase reaction based on circular dichroism. *Biochemical Journal* **2005,** 389, 491-496.
97. Novagen, *pET System Manual*. 11th ed.
98. PhD thesis, Alexeeva, M. University of Edinburgh, Edinburgh, 2003.
99. *Gene* **1987,** 51.
100. www.tigr.org
101. Thornton, M. P.; Johnston, R. B., Studies on Alanine Racemase from Bacillus Subtilis. *Federation Proceedings* **1967,** 26, (2), 843-&.
102. Wild, J.; Hennig, J.; Lobocka, M.; Walczak, W.; Klopotowski, T., Identification of the Dadx Gene Coding for the Predominant Isozyme of Alanine Racemase in Escherichia-Coli-K12. *Molecular & General Genetics* **1985,** 198, (2), 315-322.
103. www.expasy.ch
104. Invitrogen *One Shot Top10 Competent Cells. Version M*; 2004.
105. Invitrogen *One Shot BL21(DE3) pLysS Competent Cells. Version G*; 2002.
106. www.brenda.uni-koeln.de
107. Spies, M. A.; Toney, M. D., Multiple hydrogen kinetic isotope effects for enzymes catalyzing exchange with solvent: Application to alanine racemase. *Biochemistry* **2003,** 42, (17), 5099-5107.
108. http://www1.qiagen.com/resources/info/plasmid_specifications.aspx#1
109. Fotheringham, I. G. Transaminase biotransformation process. 6197558, 2001.
110. LeMagueres, P.; Im, H.; Dvorak, A.; Strych, U.; Benedik, M.; Krause, K. L., Crystal structure at 1.45 angstrom resolution of alanine racemase from a pathogenic bacterium, Pseudomonas aeruginosa, contains both internal and external aldimine forms. *Biochemistry* **2003,** 42, (50), 14752-14761.
111. Dawson, R. M. C.; Elliott, D. C.; Jones, K. M., *Data for Biochemical Research*. 2nd ed.; Oxford University Press: London, 1969.
112. Geueke, B.; Hummel, W., A new bacterial L-amino acid oxidase with a broad substrate specificity: purification and characterization. *Enzyme and Microbial Technology* **2002,** 31, (1-2), 77-87.

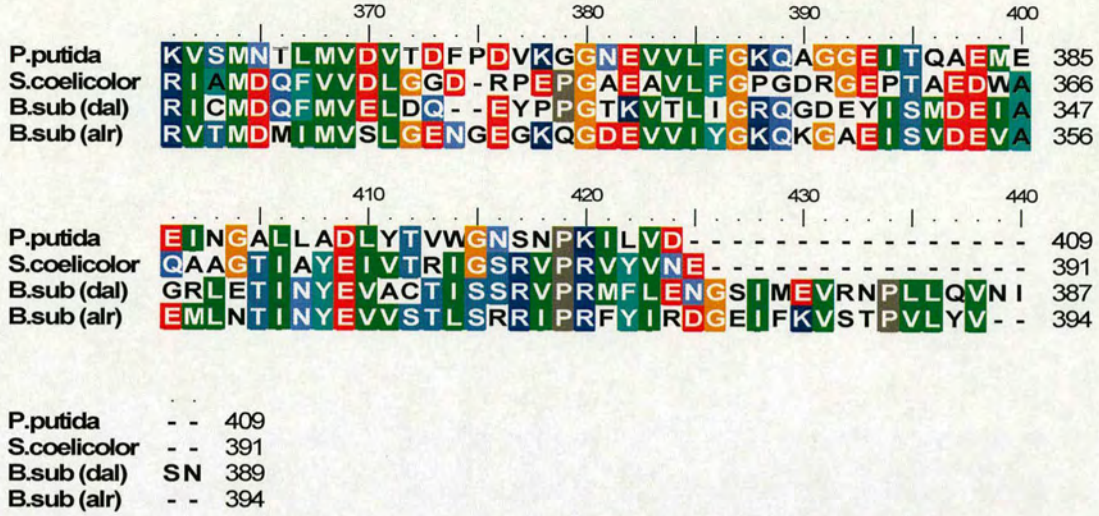

P.putida rac
pJLM101 gene
10 20 30 40
ATGCCCTTTCGCCGTACCCTTCTGGCTGCAT 31
M P F R R T L L A A
ATGGGTACTAGTCCCTTTCGCCGTACCCTTCTGGCCGCCCT 40
M G T S P F R R T L L A A
50 60 70 80
CCCTGGCACTTCTGATCACCGGACAGGCCCCCTGTATGC 71
S L A L L I T G Q A P L Y A
CCCTGGCCCTGCTGATCACTGGCCAGGCTCCGCTGTACGC 80
S L A L L I T G Q A P L Y A
90 100 110 120
GGCACCACCGTTGTCGATGGACAACGGCACCAACACCCTG 111
A P P L S M D N G T N T L
CGCGCCGCCGCTGTCGATGGACAACGGCACCACCGCCTTG 120
A P P L S M D N G T T A L
130 140 150 160
ACCGTGCAAAACAGCAATGCCTGGGTCGAAGTCAGCGCCA 151
T V Q N S N A W V E V S A
ACCGTGCAAAACAGCAACGCCTGGGTCGAAATCAGCGCCA 160
T V Q N S N A W V E I S A
170 180 190 200
GCGCCCTGCAGCACAACATCCGCACGCTGCAGGCCGAGCT 191
S A L Q H N I R T L Q A E L
GCGCCCTGCAACACAACATCCGCACCTTGCAGGCCGAACT 200
S A L Q H N I R T L Q A E L
210 220 230 240
GGCCGGCAAGTCCAAGCTGTGCGCCGTGCTCAAGGCCGAT 231
A G K S K L C A V L K A D
GGCCGGCAAGTCCAAGCTGTGTGCCGTGCTCAAGGCGGAT 240
A G K S K L C A V L K A D
250 260 270 280
GCCTATGGCCACGGTATCGGCCTGGTAATGCCATCGATCA 271
A Y G H G I G L V M P S I
GCCTATGGCCACGGTGTCGGCCTGGTGATGCCGTCGATCA 280
A Y G H G V G L V M P S I
290 300 310 320
TCGCCCAAGGCGTGCCCTGCGTGGCGGTGGCCAGCAACGA 311
I A Q G V P C V A V A S N E
TTGCCCAGGGCGTGCCCTGCGTGGCGGTGGCCAGCAACGA 320
I A Q G V P C V A V A S N E
330 340 350 360
GGAGGCCCGCGTGGTCCGCGCCAGTGGCTTCACCGGGCAA 351
E A R V V R A S G F T G Q
GGAGGCCCGCGTGGTCCGCGCCAGTGGCTTCACCGGGCAA 360
E A R V V R A S G F T G Q

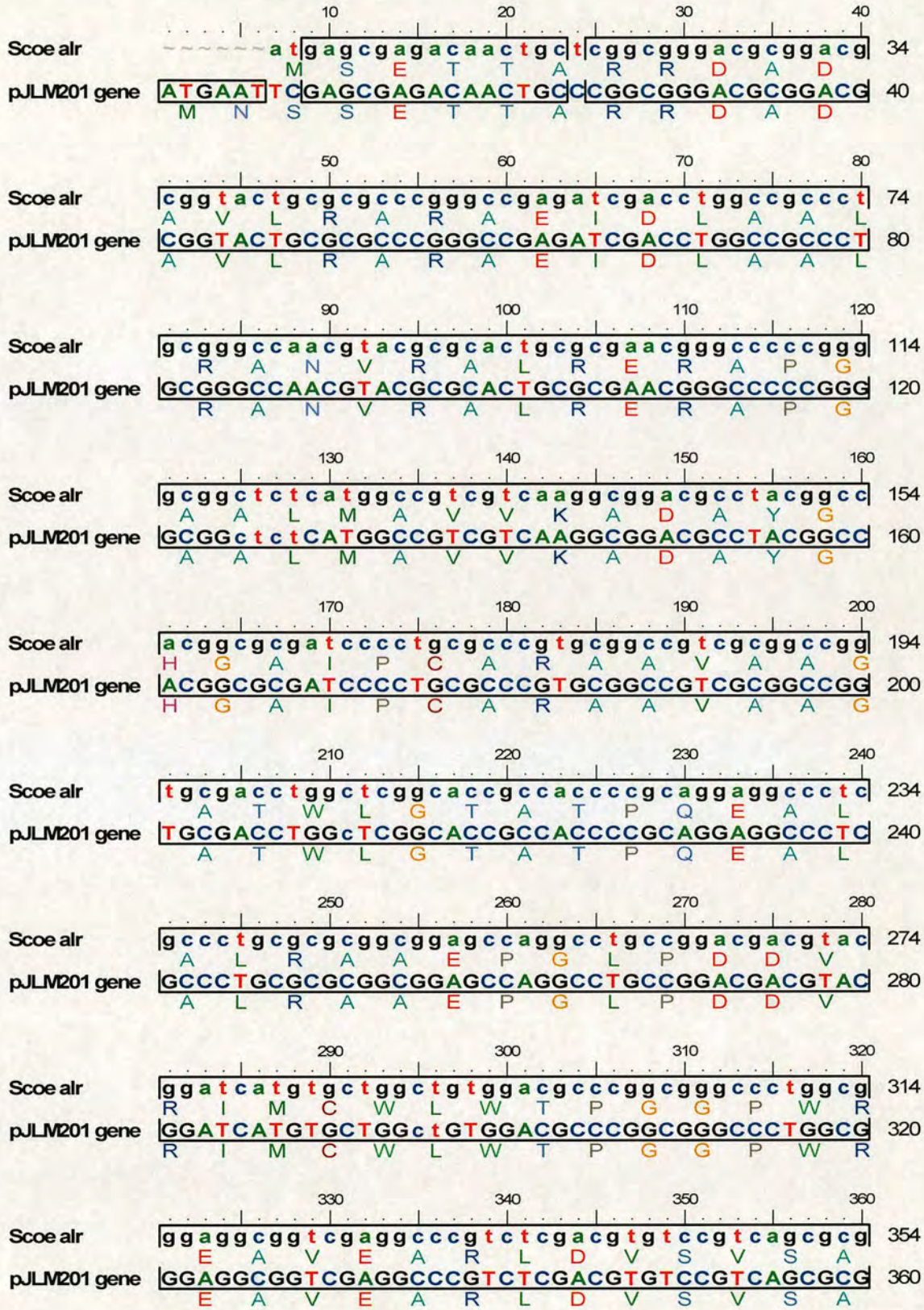

370 380 390 400
P.putida rac CTGGTGCGGGTACGCCTGGCCAGCCTCAGCGAGCTGGAAG 391
L V R V R L A S L S E L E
pJLM101 gene CTGGTGCGGGTACGCCTGGCCAGCCTCAGCGAAGTGGAAG 400
L V R V R L A S L S E V E
410 420 430 440
P.putida rac ATGGCTTGCAGTACGACATGGAAGAGCTGGTGGGCAGCGC 431
D G L Q Y D M E E L V G S A
pJLM101 gene ACGCCTTGCAGTACGACATGGAAGAACTGGTGGGCAGTGC 440
D A L Q Y D M E E L V G S A
450 460 470 480
P.putida rac GGAATTTGCCCGCCAGGCCGATGCCATCGCCGCGCGCCAT 471
E F A R Q A D A I A A R H
pJLM101 gene CGAATTCGCCCGCCAACTCGACGCCATCGCCGAGCGCCAT 480
E F A R Q L D A I A E R H
490 500 510 520
P.putida rac GGCAAGACCTTGCGCATTCACATGGCGCTCAACTCCAGCG 511
G K T L R I H M A L N S S
pJLM101 gene GGCAAGCCCGTTGCGCATCCACCTGGCCCTCAATTCCAGCG 520
G K P L R I H L A L N S S
530 540 550 560
P.putida rac GCATGAGCCGCAACGGGGTGGAGATGGCCACCTGGTCCGG 551
G M S R N G V E M A T W S G
pJLM101 gene GCATGAGCCGCAACGGGGTAGAGATGACCACTTGGTCCGG 560
G M S R N G V E M T T W S G
570 580 590 600
P.putida rac CCGTGGCGAAGCGCTGCAGATCACCGACCAGAAGCACCTC 591
R G E A L Q I T D Q K H L
pJLM101 gene CCGTGGCGAAGCGTTGCAGGTTACCGACCAGAAGCACCTC 600
R G E A L Q V T D Q K H L
610 620 630 640
P.putida rac AAGCTGGTCGCGCTGATGACCCACTTCGCCGTGGAAGACA 631
K L V A L M T H F A V E D
pJLM101 gene AAGCTGGTCGCGCTGATGACCCACTTTGCCGTGGAAGACA 640
K L V A L M T H F A V E D
650 660 670 680
P.putida rac AGGACGATGTACGCAAGGGCCTGGCGGCATTCAACGAGCA 671
K D D V R K G L A A F N E Q
pJLM101 gene AAGATGATGTGCGCAAGGGCCTGGCAGCGTTCAACGAGCA 680
K D D V R K G L A A F N E Q
690 700 710 720
P.putida rac GACCGACTGGTTGATCAAGCACGCCAGGCTGGACCGCAGC 711
T D W L I K H A R L D R S
pJLM101 gene GACCGACTGGCTGATCAAGCACGCCAAGCTGGACCGCAGC 720
T D W L I K H A K L D R S

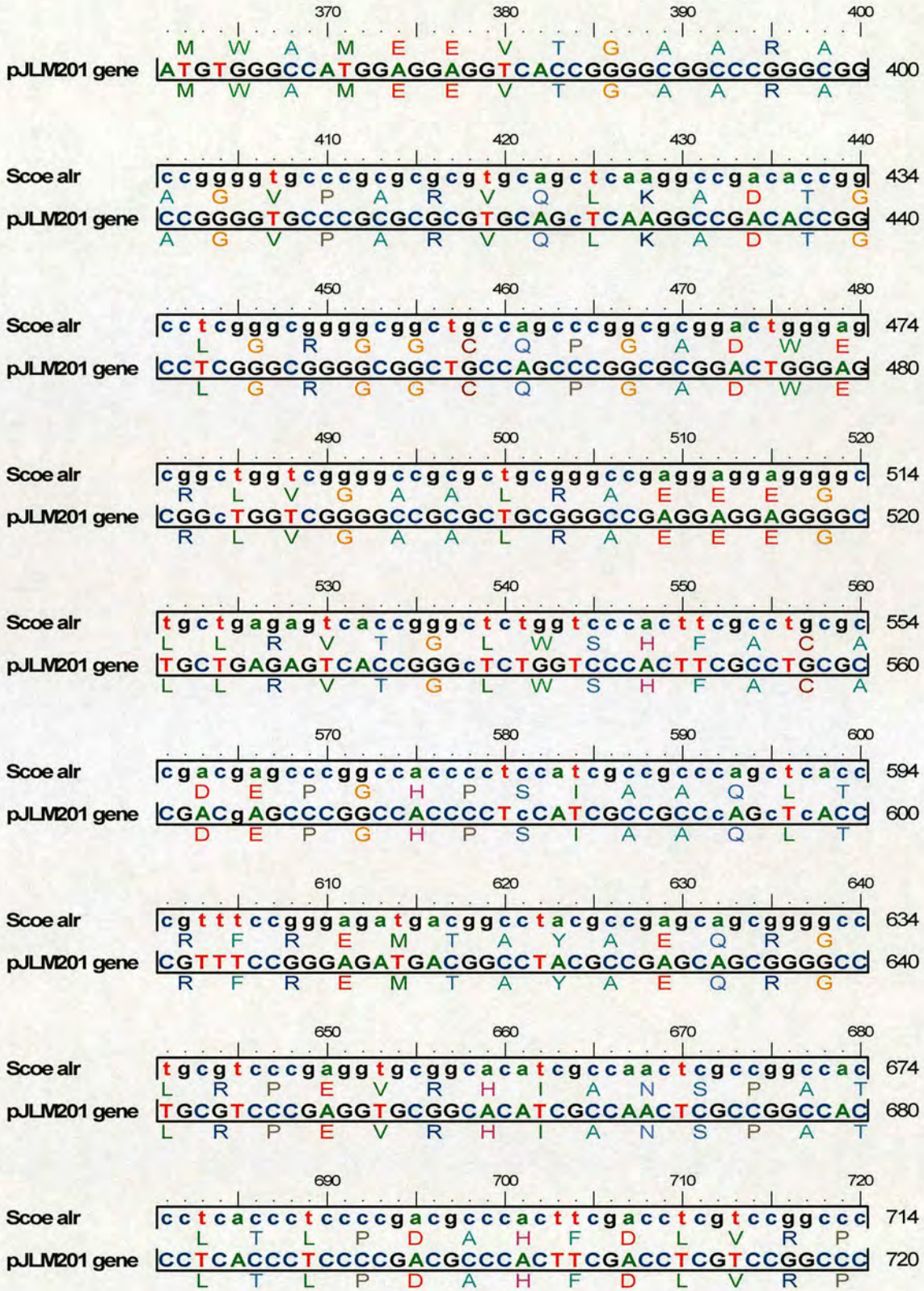

730 740 750 760
P.putida rac AAGCTCACCCTGCACGCCGCCAACTCGTTCGCTACGCTGG 751
K L T L H A A N S F A T L
pJLM101 gene AAGCTCACCCTGCACGCCGCCAACTCGTTCGCTACGCTGG 760
K L T L H A A N S F A T L
770 780 790 800
P.putida rac AAGTGCCGGAAGCGCGCCTGGACATGGTACGAACGGGTGG 791
E V P E A R L D M V R T G G
pJLM101 gene AGGTGCCGGAAGCGCGCCTGGACATGGTGCGCACCGGTGG 800
E V P E A R L D M V R T G G
810 820 830 840
P.putida rac CGCGCTGTTCGGCGACACCGTGCCGGCGCGCACCGAGTAC 831
A L F G D T V P A R T E Y
pJLM101 gene CGCGCTGTTCGGTGACACCGTGCCGAGCCACACCGAGTAC 840
A L F G D T V P S H T E Y
850 860 870 880
P.putida rac AAACGTGCGATGCAGTTCAAATCGCACGTGGCGGCGGTGC 871
K R A M Q F K S H V A A V
pJLM101 gene CAGCGGGCCATGCAGTTCAAGTCGCGTGTGGCTGCCGTGC 880
Q R A M Q F K S R V A A V
890 900 910 920
P.putida rac ACAGCTATCCGGCCGGCAACACCGTGGGCTATGACCGCAC 911
H S Y P A G N T V G Y D R T
pJLM101 gene ACAGCTACCCGGCCGGCAATACCGTCGGTTACGACCGCAC 920
H S Y P A G N T V G Y D R T
930 940 950 960
P.putida rac CTTCACCCTGGCCCGTGATTCGCGGCTGGCCAACATTACG 951
F T L A R D S R L A N I T
pJLM101 gene CTTTACTTTGGCCCGTGACGCTCGGCTGGCCAATATCACT 960
F T L A R D A R L A N I T
970 980 990 1000
P.putida rac GTCGGGTACTCCGATGGCTACCGCCGGGTATTCACCAACA 991
V G Y S D G Y R R V F T N
pJLM101 gene GTCGGCTATTCCGACGGTTACCGCCGGGCGTTTACCAACA 1000
V G Y S D G Y R R A F T N
1010 1020 1030 1040
P.putida rac AGGGCCATGTGCTGATCAACGGCCACCGTGTGCCGGTCGT 1031
K G H V L I N G H R V P V V
pJLM101 gene AGGGCCATGTGCTGATCAACGGCCACCGTGTGCCGGTCGT 1040
K G H V L I N G H R V P V V
1050 1060 1070 1080
P.putida rac GGGCAAGGTGTCGATGAACACGCTGATGGTCGATGTCACC 1071
G K V S M N T L M V D V T
pJLM101 gene GGGCAAGGTGTCGATGAACACGCTGATGGTGGACGTTACC 1080
G K V S M N T L M V D V T

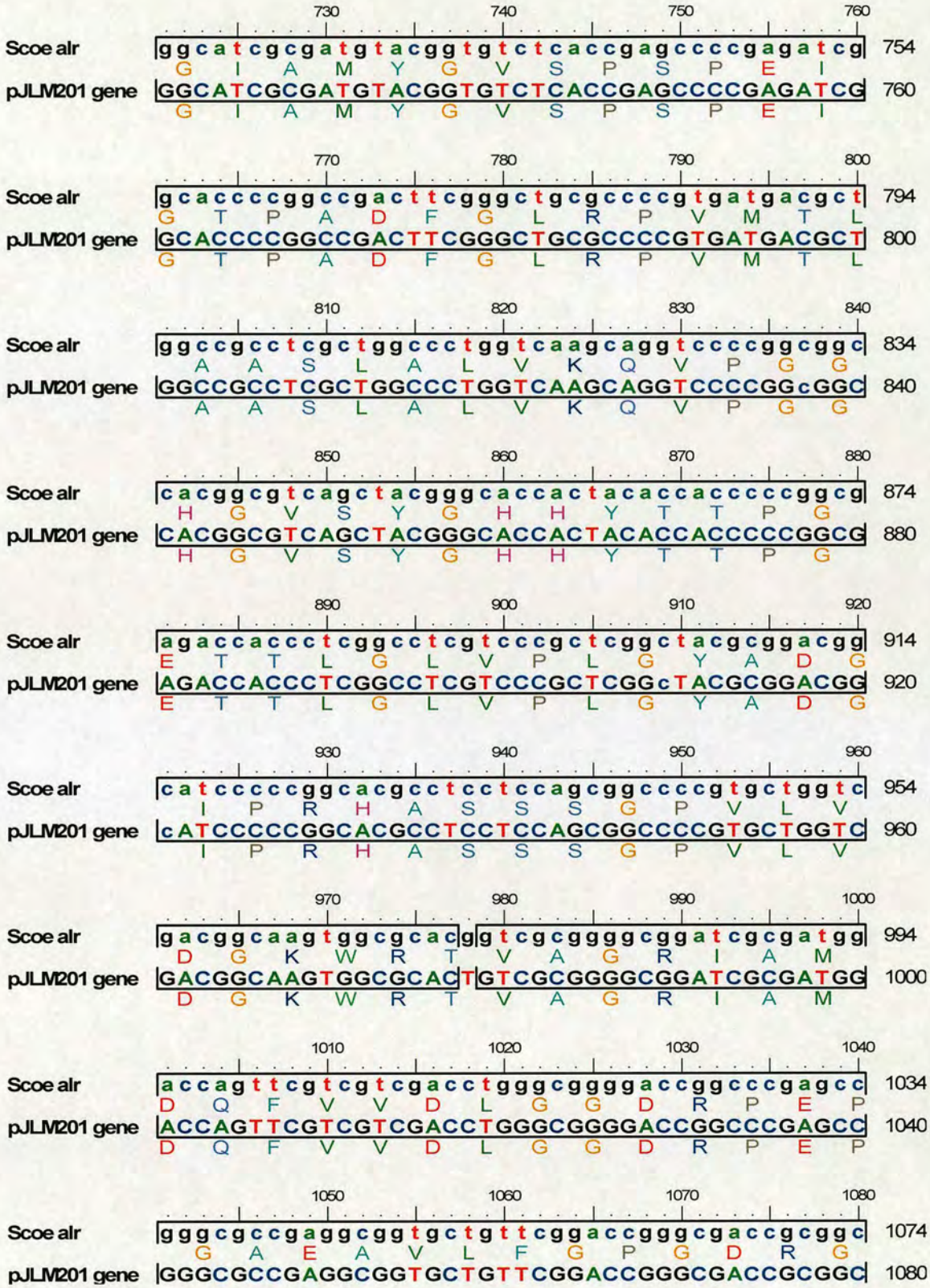

1090
1100
1110
1120
P.putida rac
GACTTCCCTGATGTGAAGGGGGGTAACGAAGTGGTGCTGT 1111
D F P D V K G G N E V V L
pJLM101 gene
GACTTCCCGGATGTGAAgGGGGGCAACGAAGTGGTGCTGT 1120
D F P D V K G G N E V V L
1130
1140
1150
1160
P.putida rac
TCGGCAAGCAGGCCGGGGGCGAAATCACCCAGGCCGAGAT 1151
F G K Q A G G E I T Q A E M
pJLM101 gene
TCGGCAAGCAGGCCGGGGGCGAAATCACCCAGGCCGAAAT 1160
F G K Q A G G E I T Q A E M
1170
1180
1190
1200
P.putida rac
GGAAGAAATCAACGGCGCGTTGCTCGCCGATTTGTACACC 1191
E E I N G A L L A D L Y T
pJLM101 gene
GGAAGACATCAATGGCGCGTTGCTCGCCGACCTGTACACC 1200
E D I N G A L L A D L Y T
1210
1220
1230
P.putida rac
GTATGGGGCAATTCCAACCCGAAGATACTCGTCGAC 1227
V W G N S N P K I L V D
pJLM101 gene
GTGTGGGGCAGTTCCAACCCGAAGATACTCGTCGAC 1236
V W G S S N P K I L V D

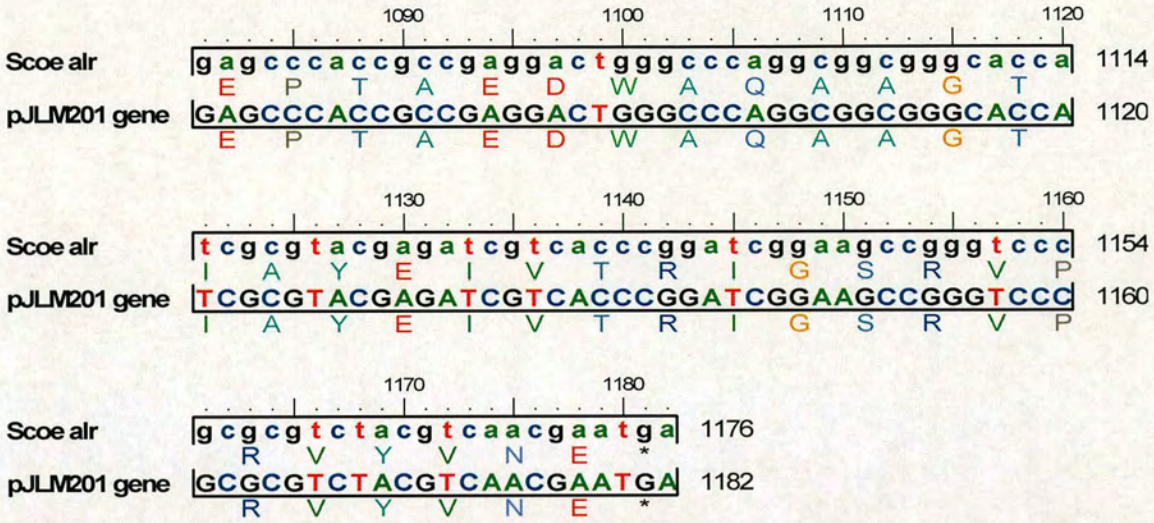

10 20 30 40
P.putida MPFRRTLLAASLALLITGQAPLYAAPPLSMDNGTNTLTVQ 40
S.coelicolor ---------------------------MSETTARRDADA 12
B.sub (dal) -------------------------------MSTKPF 6
B.sub (alr) ---------------------------------MIKL 4
50 60 70 80
P.putida NSNAWVEVSASALQHNIRTLQAELAGKSKLCAVLKADAYG 80
S.coelicolor VLRARAEIDLAALRANVRALRERAPG-AALMAVVKADAYG 51
B.sub (dal) YRDTWAEIDLSAIKENVSNMKKHIGEHVHLMAVVKANAYG 46
B.sub (alr) CREVWIEVNLDAVKKNLRAIRRHIPHKSKIMAVVKANGYG 44
90 100 110 120
P.putida HGIGLVMPSIIAQGVPCVAVASNEEARVVRASGFTGQ--- 117
S.coelicolor HGAIPCARAAVAAGATWLGTATPQEALALRAAEPGLPDDV 91
B.sub (dal) HGDAETAKAALDAGASCLAVAILDEAISLRKKGLKAP--- 83
B.sub (alr) HGSIEVARHALEHGASELAVASVEEGIVLRKAGITAP--- 81
130 140 150 160
P.putida -LVRVRLASLSELEDGLQYDMEELVGSAEFARQADAIAAR 156
S.coelicolor RIMCWLWTPGGPWREAVEARLDVSVSAMWAMEEVTGAARA 131
B.sub (dal) -ILVLGAVPPEYVAIAAEYDVTLTGYSVEWLQEAARHTKK 122
B.sub (alr) -ILVLGFTSLSCVKKSAAWNITLSAFQVDWMKEANEILEK 120
170 180 190 200
P.putida HG--KTLRIHMALNSSGMSRNGVEMATWSGRGEALQITDQ 194
S.coelicolor AG--VPARVQLKADT-GLGRGGCQPGADWERLVGAALRAE 168
B.sub (dal) GS----LHFHLKVDT-GMNRLGVKTEEEVQNVMAILDRNP 157
B.sub (alr) EASANRLAIHINVDT-GMGRLGVRTKEELLEVVKALKASK 159
210 220 230 240
P.putida K--HLKLVALMTHFAVEDKDDVRKGLAAFNEQTDWLIKHA 232
S.coelicolor EEGLLRVTGLWSHFACADEPGHPSIAAQLTRFREMTAYAE 208
B.sub (dal) ---RLKCKGVFTHFATADEKERGYFLMQFERFKELIAPLP 194
B.sub (alr) ---FLRWTGIFTHFSTADEPDTTLTKLQHEKFISFLSFLK 196
250 260 270 280
P.putida RLDRSKLTLHAANSFATLEVPEARLDMVRTGGALFGDTVP 272
S.coelicolor QRGLRPEVRHIANSPATLTLPDAHFDLVRPGIAMYGVSPS 248
B.sub (dal) ---LKNLMVHCANSAAGLRLKKGFFNAVRFGIGMYGLRPS 231
B.sub (alr) KQGIELPTVHMCNTAAAIAFPEFSADMIRLGIGLYGLYPS 236
290 300 310 320
P.putida ART------EYKRAMQFKSHVAAVHSYPAG-NTVGYDRTF 305
S.coelicolor PEIGTPADFGLRPVMTLAASLALVKQVPGG-HGVSYGHHY 287
B.sub (dal) ADMSDEIPFQLRPAFTLHSTLSHVKLIRKG-ESVSYGAEY 270
B.sub (alr) AYIKQLNLVKLEPALSLKARIAYVKTMRTEPRTVSYGATY 276
330 340 350 360
P.putida TLARDSRLANITVGYSDGYRRVFTNKGHVLINGHRVPVVG 345
S.coelicolor TTPGETTLGLVPLGYADGIPRHASSSGPVLVDGKWRTVAG 327
B.sub (dal) TAEKDTWIGTVPVGYADGWLRKLKG-TDILVKGKRLKIAG 309
B.sub (alr) IAEPNEVIATLPIGYADGYSRALSNRGFVLHRGKRVPVAG 316

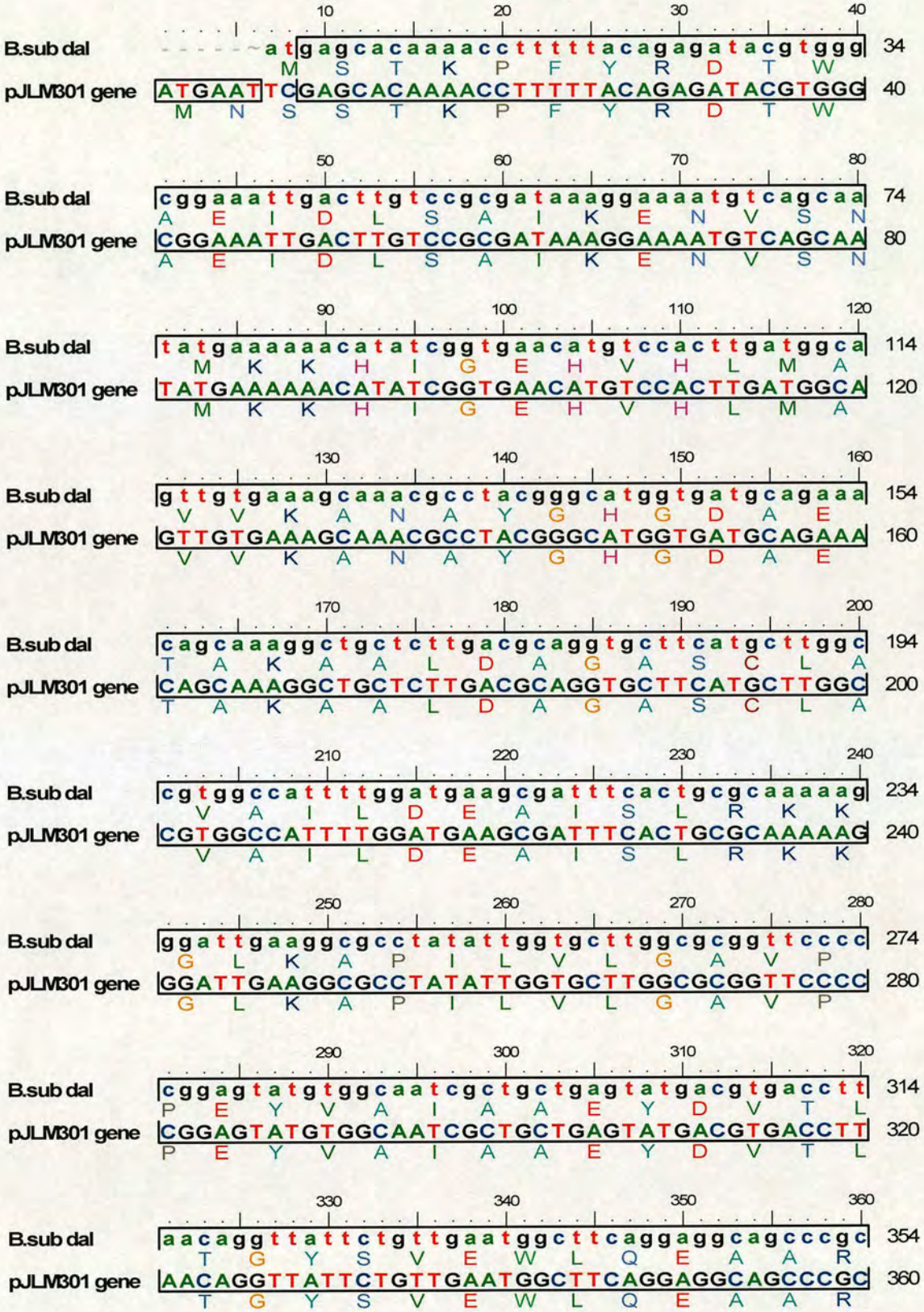

370
380
390
400
P.putida KVSMNTLMVDVTDFPDVKGGNEVVLFGKQAGGEITQAEME 385
S.coelicolor RIAMDQFVVDLGGD-RPEPGAEAVLFGPGDRGEPTAEDWA 366
B.sub (dal) RICMDQFMVELDQ--EYPPGTKVTLIGRQGDEYISMDEIA 347
B.sub (alr) RVTMDMIMVSLGENGEGKQGDEVVIYGKQKGAEISVDEVA 356
410
420
430
440
P.putida EINGALLADLYTVWGNSNPKILVD---------------- 409
S.coelicolor QAAGTIAYEIVTRIGSRVPRVYVNE---------------- 391
B.sub (dal) GRLETINYEVACTISSRVPRMFLENGSIMEVRNPLLQVNI 387
B.sub (alr) EMLNTINYEVVSTLSRRIPRFYIRDGEIFKVSTPVLYV-- 394
P.putida -- 409
S.coelicolor -- 391
B.sub (dal) SN 389
B.sub (alr) -- 394

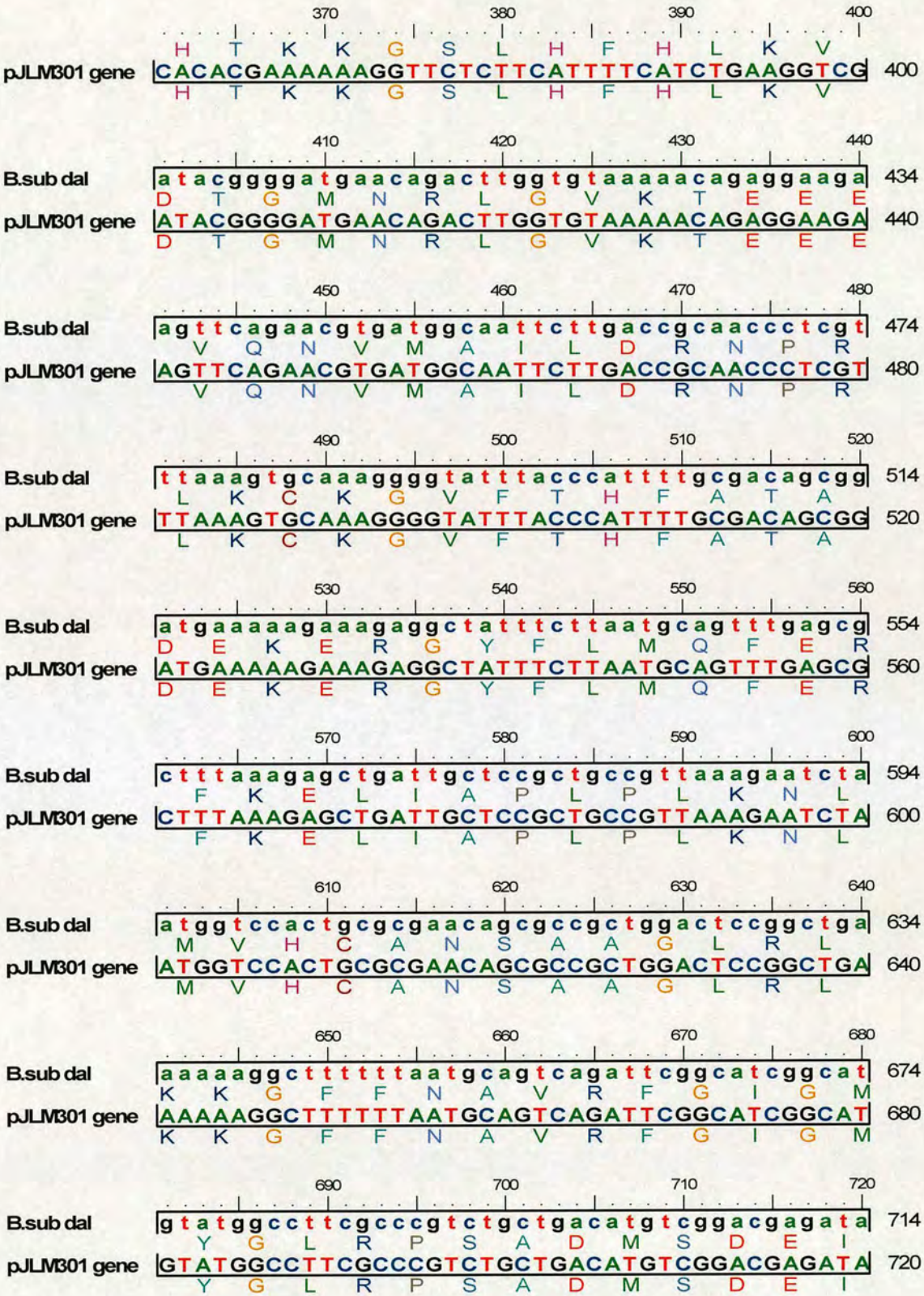

Scoe alr
pJLM201 gene
10 20 30 40
atgagcgagacaactgctcggcgggacgcggacg
M S E T T A R R D A D
ATGAATTCGAGCGAGACAACTGCCCGGCGGGACGCGGACG
M N S S E T T A R R D A D
34
40
50 60 70 80
cggtactgcgcgcccgggccgagatcgacctggccgccct
A V L R A R A E I D L A A L
CGGTACTGCGCGCCCGGGCCGAGATCGACCTGGCCGCCCT
A V L R A R A E I D L A A L
74
80
90 100 110 120
gcgggccaacgtacgcgcactgcgcgaacgggccccgggg
R A N V R A L R E R A P G
GCGGGCCAACGTACGCGCACTGCGCGAACGGGCCCCGGGG
R A N V R A L R E R A P G
114
120
130 140 150 160
gcggctctcatggccgtcgtcaaggcggacgcctacggcc
A A L M A V V K A D A Y G
GCGGctctCATGGCCGTCGTCAAGGCGGACGCCTACGGCC
A A L M A V V K A D A Y G
154
160
170 180 190 200
acggcgcgatccccctgcgcccgtgcggccgtcgcggccgg
H G A I P C A R A A V A A G
ACGGCGCGATCCCCCTGCGCCCGTGCGGCCGTCGCGGCCGG
H G A I P C A R A A V A A G
194
200
210 220 230 240
tgcgacctggctcggcaccgccaccccgcaggaggccctc
A T W L G T A T P Q E A L
TGCGACCTGGcTCGGCACCGCCACCCCGCAGGAGGCCCTC
A T W L G T A T P Q E A L
234
240
250 260 270 280
gccctgcgcgcggcggagccaggcctgccggacgacgtac
A L R A A E P G L P D D V
GCCCTGCGCGCGGCGGAGCCAGGCCTGCCGGACGACGTAC
A L R A A E P G L P D D V
274
280
290 300 310 320
ggatcatgtgctggctgtggacgcccggcgggcccctggcg
R I M C W L W T P G G P W R
GGATCATGTGCTGGctGTGGACGCCCGGCGGGCCCTGGCG
R I M C W L W T P G G P W R
314
320
330 340 350 360
ggaggcggtcgaggcccgtctcgacgtgtccgtcagcgcg
E A V E A R L D V S V S A
GGAGGCGGTCGAGGCCCGTCTCGACGTGTCCGTCAGCGCG
E A V E A R L D V S V S A
354
360

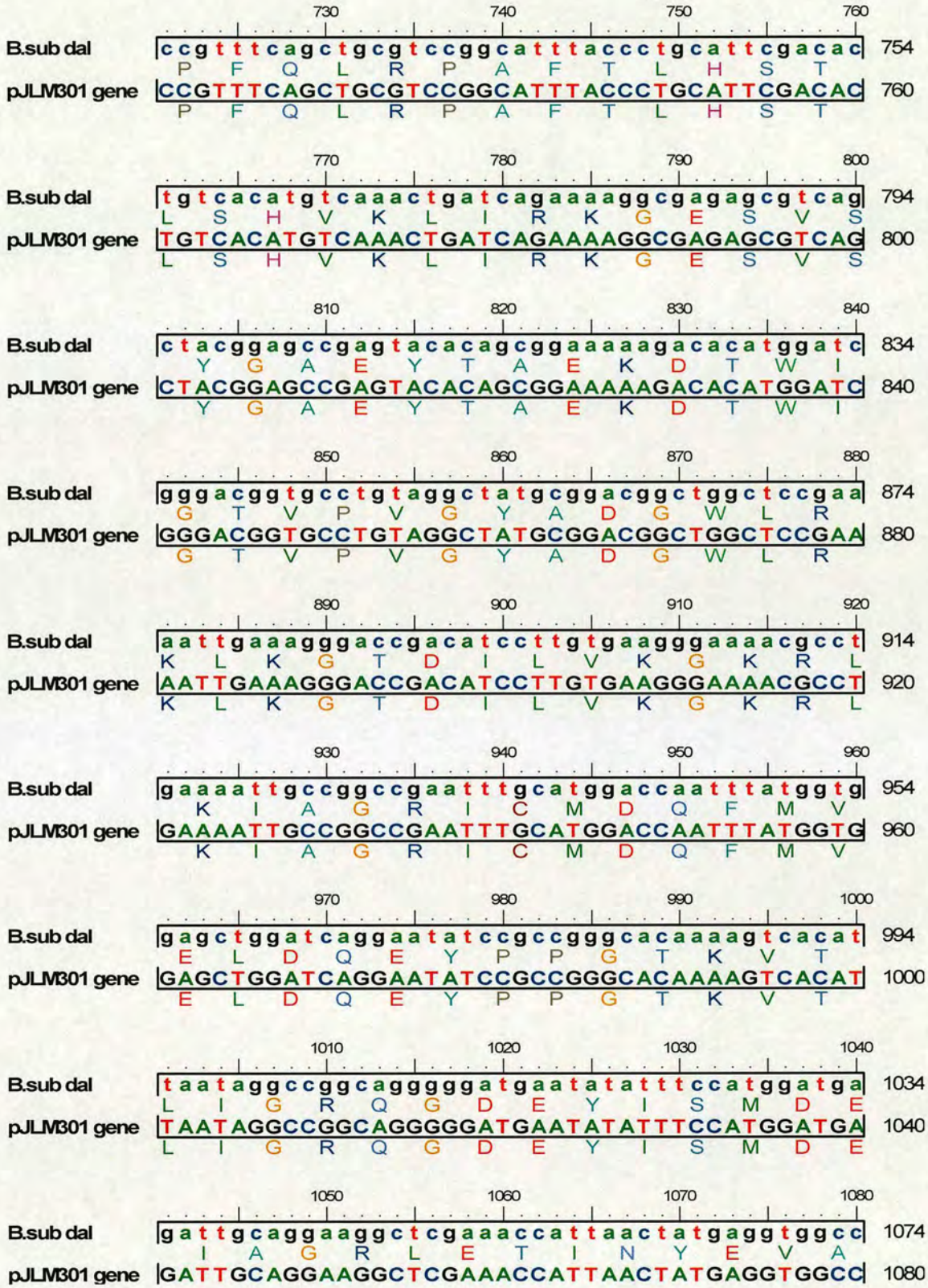

370 380 390 400
M W A M E E V T G A A R A
pJLM201 gene ATGTGGGCCATGGAGGAGGTCACCGGGGCGGCCCGGGCGG 400
M W A M E E V T G A A R A
410 420 430 440
Scoe alr ccgggg tgcccgcgcgcg tgcagctcaaggccgacaccgg 434
A G V P A R V Q L K A D T G
pJLM201 gene CCGGGGTGCCCGCGCGCGTGCAGcTCAAGGCCGACACCGG 440
A G V P A R V Q L K A D T G
450 460 470 480
Scoe alr cctcgggcggggcggctgccagcccggcgcggactgggag 474
L G R G G C Q P G A D W E
pJLM201 gene CCTCGGGCGGGGCGGCTGCCAGCCCGGCGCGGACTGGGAG 480
L G R G G C Q P G A D W E
490 500 510 520
Scoe alr cggctggtcggggccgcgctgcgggccgaggaggagggggc 514
R L V G A A L R A E E E G
pJLM201 gene CGGcTGGTCGGGGCCGCGCTGCGGGCCGAGGAGGAGGGGC 520
R L V G A A L R A E E E G
530 540 550 560
Scoe alr tgctgagagtcaccgggctctggtcccacttcgcctgcgc 554
L L R V T G L W S H F A C A
pJLM201 gene TGCTGAGAGTCACCGGGcTCTGGTCCCACTTCGCCTGCGC 560
L L R V T G L W S H F A C A
570 580 590 600
Scoe alr cgacgagcccggccacccctccatcgccgcccagctcacc 594
D E P G H P S I A A Q L T
pJLM201 gene CGACgAGCCCGGCCACCCCTcCATCGCCGCCcAGcTcACC 600
D E P G H P S I A A Q L T
610 620 630 640
Scoe alr cgtttccgggagatgacggcctacgccgagcagcggggcc 634
R F R E M T A Y A E Q R G
pJLM201 gene CGTTTCCGGGAGATGACGGCCTACGCCGAGCAGCGGGGCC 640
R F R E M T A Y A E Q R G
650 660 670 680
Scoe alr tgcgtcccgaggtgcggcacatcgccaactcgccggccac 674
L R P E V R H I A N S P A T
pJLM201 gene TGCGTCCCGAGGTGCGGCACATCGCCAACTCGCCGGCCAC 680
L R P E V R H I A N S P A T
690 700 710 720
Scoe alr cctcaccctccccgacgcccacttcgacctcgtccggccc 714
L T L P D A H F D L V R P
pJLM201 gene CCTCACCCTCCCCGACGCCCACTTCGACCTCGTCCGGCCC 720
L T L P D A H F D L V R P

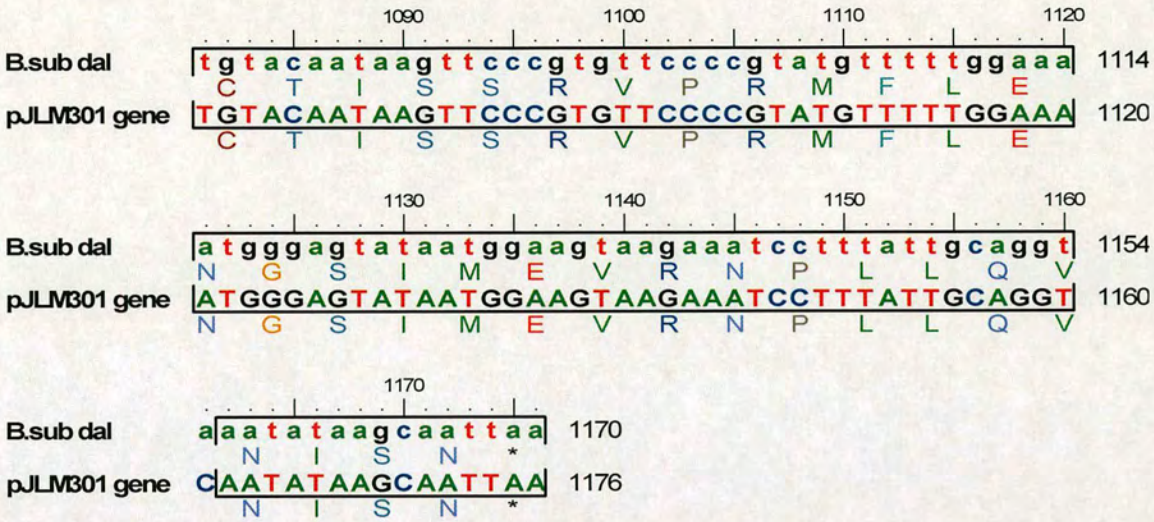

730 740 750 760
Scoe alr ggcatcgcgatgtacggtgtctcaccgagccccgagatcg 754
G I A M Y G V S P S P E I
pJLM201 gene GGCATCGCGATGTACGGTGTCTCACCGAGCCCCGAGATCG 760
G I A M Y G V S P S P E I
770 780 790 800
Scoe alr gcaccccggccgacttcgggctgcgccccgtgatgacgct 794
G T P A D F G L R P V M T L
pJLM201 gene GCACCCCGGCCGACTTCGGGCTGCGCCCCGTGATGACGCT 800
G T P A D F G L R P V M T L
810 820 830 840
Scoe alr ggccgcctcgctggccctggtcaagcaggtccccggcggc 834
A A S L A L V K Q V P G G
pJLM201 gene GGCCGCCTCGCTGGCCCTGGTCAAGCAGGTCCCCGGcGGC 840
A A S L A L V K Q V P G G
850 860 870 880
Scoe alr cacggcgtcagctacgggcaccactacaccacccccggcg 874
H G V S Y G H H Y T T P G
pJLM201 gene CACGGCGTCAGCTACGGGCACCACTACACCACCCCCGGCG 880
H G V S Y G H H Y T T P G
890 900 910 920
Scoe alr agaccaccctcggcctcgtcccgctcggctacgcggacgg 914
E T T L G L V P L G Y A D G
pJLM201 gene AGACCACCCTCGGCCTCGTCCCGCTCGGcTACGCGGACGG 920
E T T L G L V P L G Y A D G
930 940 950 960
Scoe alr catccccccggcacgcctcctccagcggccccgtgctggtc 954
I P R H A S S S G P V L V
pJLM201 gene cATCCCCCCGGCACGCCTCCTCCAGCGGCCCCCGTGCTGGTC 960
I P R H A S S S G P V L V
970 980 990 1000
Scoe alr gacggcaagtggcgcacggtcgcggggcggatcgcgatgg 994
D G K W R T V A G R I A M
pJLM201 gene GACGGCAAGTGGCGCACTGTCGCGGGGCGGATCGCGATGG 1000
D G K W R T V A G R I A M
1010 1020 1030 1040
Scoe alr accagttcgtcgtcgacctgggcggggaccggcccgagcc 1034
D Q F V V D L G G D R P E P
pJLM201 gene ACCAGTTCGTCGTCGACCTGGGCGGGGACCGGCCCGAGCC 1040
D Q F V V D L G G D R P E P
1050 1060 1070 1080
Scoe alr gggcgccgaggcggtgctgttcggaccgggcgaccgcggc 1074
G A E A V L F G P G D R G
pJLM201 gene GGGCGCCGAGGCGGTGCTGTTCGGACCGGGCGACCGCGGC 1080

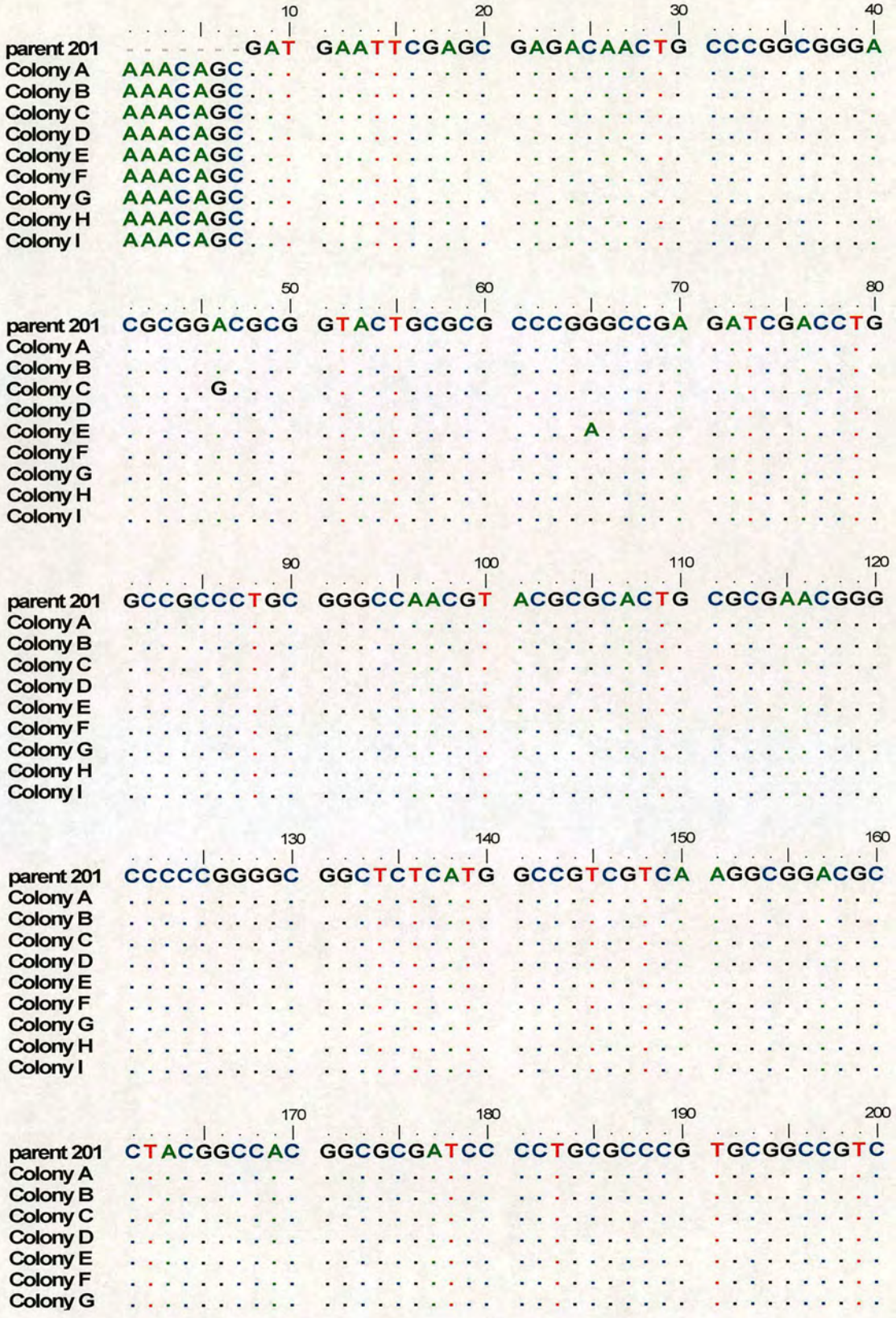

1090
1100
1110
1120
Scoe alr
gagcccaccgccgaggactgggcccaggcggcgggcacca
1114
E P T A E D W A Q A A G T
pJLM201 gene
GAGCCCACCGCCGAGGACTGGGCCCAGGCGGCGGGCACCA
1120
E P T A E D W A Q A A G T
1130
1140
1150
1160
Scoe alr
tcgcgtacgagatcgtcacccggatcggaagccgggtccc
1154
I A Y E I V T R I G S R V P
pJLM201 gene
TCGCGTACGAGATCGTCACCCGGATCGGAAGCCGGGTCCC
1160
I A Y E I V T R I G S R V P
1170
1180
Scoe alr
gcgcgtctacgtcaacgaatga
1176
R V Y V N E *
pJLM201 gene
GCGCGTCTACGTCAACGAATGA
1182
R V Y V N E *

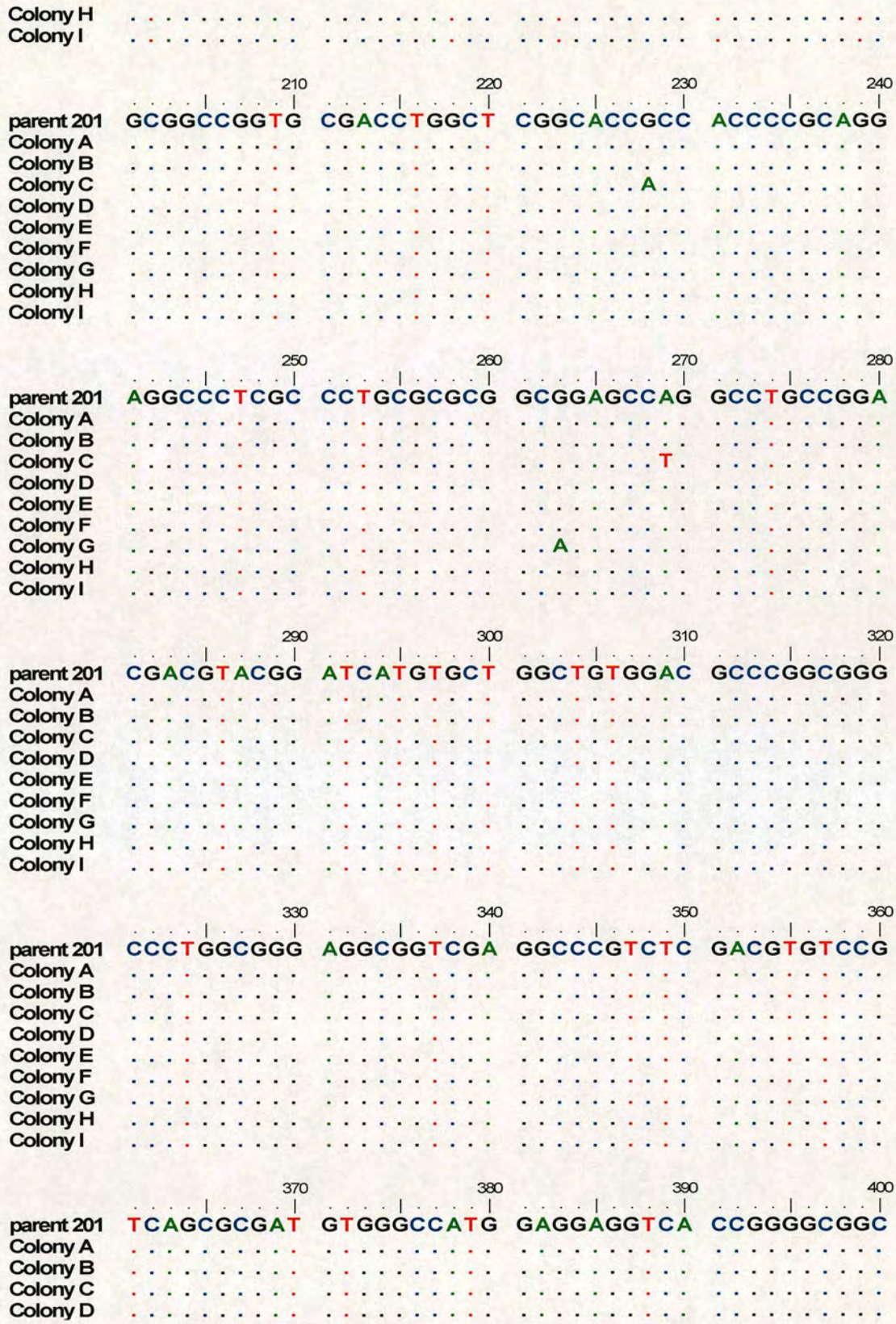

B.sub dal
pJLM301 gene
10 20 30 40
atgagcacaaaacctttttacagagatacgtggg
M S T K P F Y R D T W
ATGAATTCGAGCACAAAACCTTTTTACAGAGATACGTGGG
M N S S T K P F Y R D T W
34
40
50 60 70 80
cggaaattgacttgtccgcgataaaggaaaatgtcagcaa
A E I D L S A I K E N V S N
CGGAAATTGACTTGTCCGCGATAAAGGAAAATGTCAGCAA
A E I D L S A I K E N V S N
74
80
90 100 110 120
tatgaaaaaacatatcggtgaacatgtccacttgatggca
M K K H I G E H V H L M A
TATGAAAAAACATATCGGTGAACATGTCCACTTGATGGCA
M K K H I G E H V H L M A
114
120
130 140 150 160
gttgtgaaagcaaacgcctacgggcatggtgatgcagaaa
V V K A N A Y G H G D A E
GTTGTGAAAGCAAACGCCTACGGGCATGGTGATGCAGAAA
V V K A N A Y G H G D A E
154
160
170 180 190 200
cagcaaaggctgctcttgacgcaggtgcttcatgcttggc
T A K A A L D A G A S C L A
CAGCAAAGGCTGCTCTTGACGCAGGTGCTTCATGCTTGGC
T A K A A L D A G A S C L A
194
200
210 220 230 240
cgtggccatttttggatgaagcgatttcactgcgcaaaaag
V A I L D E A I S L R K K
CGTGGCCATTTTTGGATGAAGCGATTTCACTGCGCAAAAAG
V A I L D E A I S L R K K
234
240
250 260 270 280
ggattgaaggcgcctatattggtgcttggcgcggttcccc
G L K A P I L V L G A V P
GGATTGAAGGCGCCTATATTGGTGCTTGGCGCGGTTCCCC
G L K A P I L V L G A V P
274
280
290 300 310 320
cggagtatgtggcaatcgctgctgagtatgacgtgacctt
P E Y V A I A A E Y D V T L
CGGAGTATGTGGCAATCGCTGCTGAGTATGACGTGACCTT
P E Y V A I A A E Y D V T L
314
320
330 340 350 360
aacaggttattctgttgaatggcttcaggaggcagcccgc
T G Y S V E W L Q E A A R
AACAGGTTATTCTGTTGAATGGCTTCAGGAGGCAGCCCGC
T G Y S V E W L Q E A A R
354
360

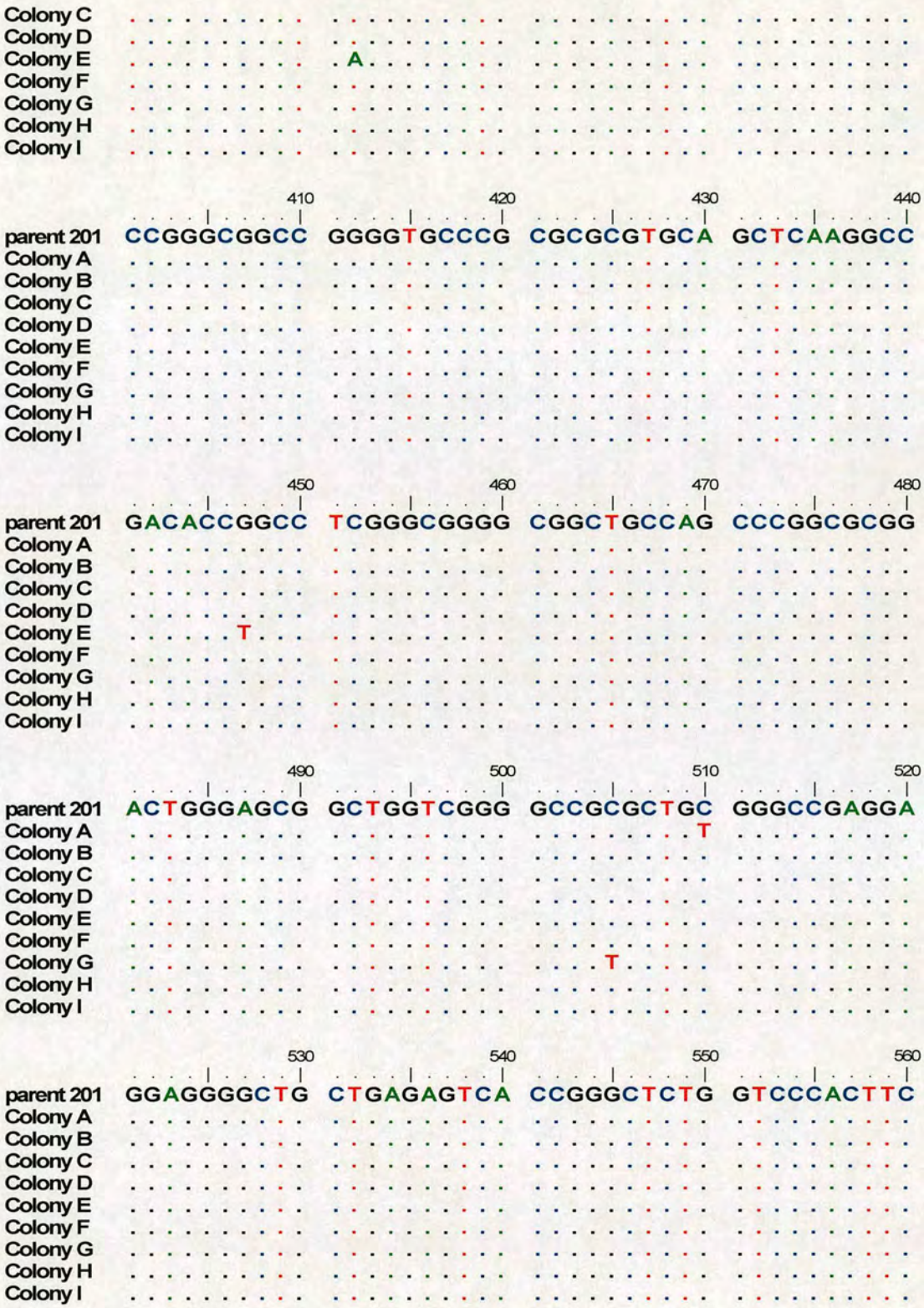

370 380 390 400
H T K K G S L H F H L K V
pJLM301 gene CACACGAAAAAAGGTTCTCTTCATTTTCATCTGAAGGTCG 400
H T K K G S L H F H L K V
410 420 430 440
B.sub dal atacgggggatgaacagacttggtgtaaaaaacagaggaaga 434
D T G M N R L G V K T E E E
pJLM301 gene ATACGGGGATGAACAGACTTGGTGTAAAAACAGAGGAAGA 440
D T G M N R L G V K T E E E
450 460 470 480
B.sub dal agttcagaacgtgatggcaattcttgaccgcaaccctcgt 474
V Q N V M A I L D R N P R
pJLM301 gene AGTTCAGAACGTGATGGCAATTCTTGACCGCAACCCTCGT 480
V Q N V M A I L D R N P R
490 500 510 520
B.sub dal ttaaagtgcaaaggggtatttacccattttgcgacagcgg 514
L K C K G V F T H F A T A
pJLM301 gene TTAAAGTGCAAAGGGGTATTTACCCATTTTGCGACAGCGG 520
L K C K G V F T H F A T A
530 540 550 560
B.sub dal atgaaaaagaaagaggctatttcttaatgcagtttgagcg 554
D E K E R G Y F L M Q F E R
pJLM301 gene ATGAAAAAGAAAGAGGCTATTTCTTAATGCAGTTTGAGCG 560
D E K E R G Y F L M Q F E R
570 580 590 600
B.sub dal ctttaaagagctgattgctccgctgccgttaaagaatcta 594
F K E L I A P L P L K N L
pJLM301 gene CTTTAAAGAGCTGATTGCTCCGCTGCCGTTAAAGAATCTA 600
F K E L I A P L P L K N L
610 620 630 640
B.sub dal atggtccactgcgcgaacagcgccgctggactccggctga 634
M V H C A N S A A G L R L
pJLM301 gene ATGGTCCACTGCGCGAACAGCGCCGCTGGACTCCGGCTGA 640
M V H C A N S A A G L R L
650 660 670 680
B.sub dal aaaaaggcttttttaatgcagtcagattcggcatcggcat 674
K K G F F N A V R F G I G M
pJLM301 gene AAAAAGGCTTTTTTAATGCAGTCAGATTCGGCATCGGCAT 680
K K G F F N A V R F G I G M
690 700 710 720
B.sub dal gtatggccttcgcccgtctgctgacatgtcggacgagata 714
Y G L R P S A D M S D E I
pJLM301 gene GTATGGCCTTCGCCCGTCTGCTGACATGTCGGACGAGATA 720
Y G L R P S A D M S D E I

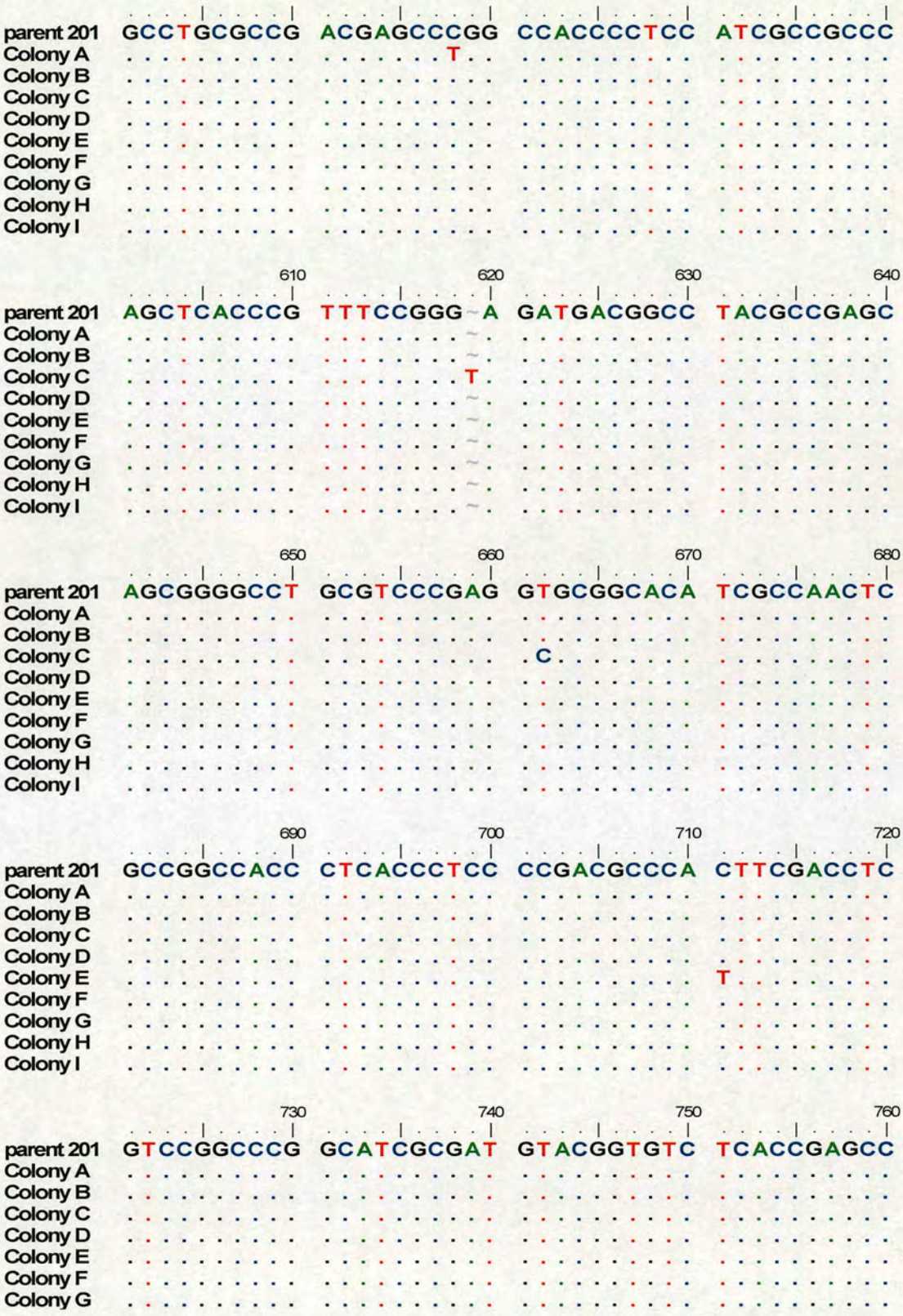

730 740 750 760
B.sub dal ccgtttcagctgcgtccggcatttaccctgcattcgacac 754
P F Q L R P A F T L H S T
pJLM301 gene CCGTTTCAGCTGCGTCCGGCATTTACCCTGCATTCGACAC 760
P F Q L R P A F T L H S T
770 780 790 800
B.sub dal tgtcacatgtcaaactgatcagaaaaggcgagagcgtcag 794
L S H V K L I R K G E S V S
pJLM301 gene TGTCACATGTCAAACTGATCAGAAAAGGCGAGAGCGTCAG 800
L S H V K L I R K G E S V S
810 820 830 840
B.sub dal ctacggagccgagtacacagcggaaaaagacacatggatc 834
Y G A E Y T A E K D T W I
pJLM301 gene CTACGGAGCCGAGTACACAGCGGAAAAAGACACATGGATC 840
Y G A E Y T A E K D T W I
850 860 870 880
B.sub dal gggacggtgcctgtaggctatgcggacggctggctccgaa 874
G T V P V G Y A D G W L R
pJLM301 gene GGGACGGTGCCTGTAGGCTATGCGGACGGCTGGCTCCGAA 880
G T V P V G Y A D G W L R
890 900 910 920
B.sub dal aatttgaaagggaccgacatccttgtgaagggaaaacgcct 914
K L K G T D I L V K G K R L
pJLM301 gene AATTTGAAAGGGACCGACATCCTTGTGAAGGGAAAACGCCT 920
K L K G T D I L V K G K R L
930 940 950 960
B.sub dal gaaaatttgccggccgaatttgcatggaccaatttatggtg 954
K I A G R I C M D Q F M V
pJLM301 gene GAAAATTTGCCGGCCGAATTTGCATGGACCAATTTATGGTG 960
K I A G R I C M D Q F M V
970 980 990 1000
B.sub dal gagctggatcaggaatatccgccgggcacaaaagtcacat 994
E L D Q E Y P P G T K V T
pJLM301 gene GAGCTGGATCAGGAATATCCGCCGGGCACAAAAGTCACAT 1000
E L D Q E Y P P G T K V T
1010 1020 1030 1040
B.sub dal taataggccggcaggggatgaatatatttccatggatga 1034
L I G R Q G D E Y I S M D E
pJLM301 gene TAATAGGCCGGCAGGGGGATGAATATATTTCCATGGATGA 1040
L I G R Q G D E Y I S M D E
1050 1060 1070 1080
B.sub dal gattgcaggaaggctcgaaaccattaactatgaggtggcc 1074
I A G R L E T I N Y E V A
pJLM301 gene GATTGCAGGAAGGCTCGAAACCATTAACTATGAGGTGGCC 1080

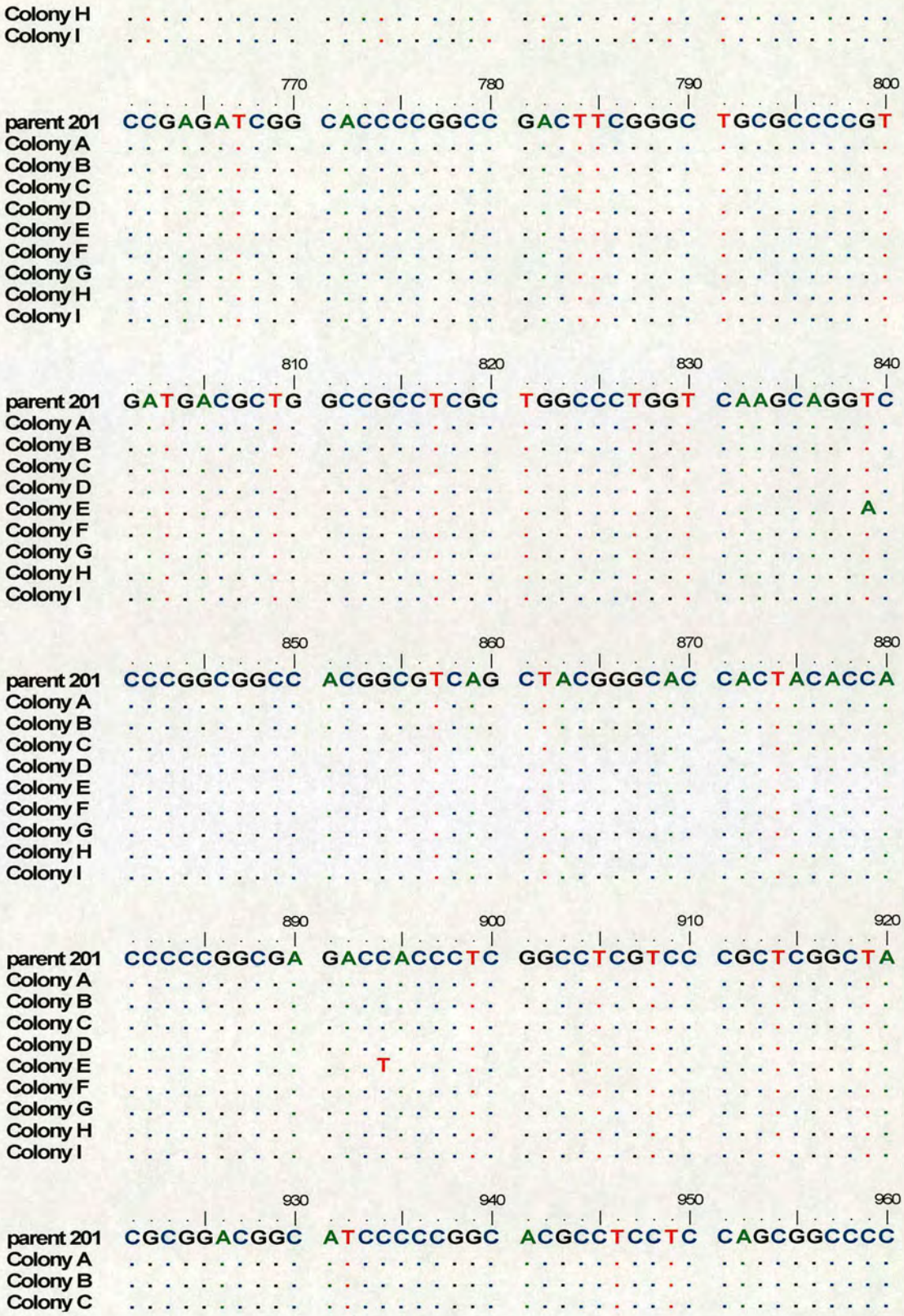

1090
1100
1110
1120
B.sub dal
tgtacaataagttcccgtgttccccgtatgttttttggaaa
1114
C T I S S R V P R M F L E
pJLM301 gene
TGTACAATAAGTTCCCGTGTTCCCCGTATGTTTTTGGAAA
1120
C T I S S R V P R M F L E
1130
1140
1150
1160
B.sub dal
atgggagtataatggaagtaagaaatcctttattgcaggt
1154
N G S I M E V R N P L L Q V
pJLM301 gene
ATGGGAGTATAATGGAAGTAAGAAATCCTTTATTGCAGGT
1160
N G S I M E V R N P L L Q V
1170
B.sub dal
aaatataagcaatttaa
1170
N I S N *
pJLM301 gene
CAATATAAGCAATTAA
1176
N I S N *

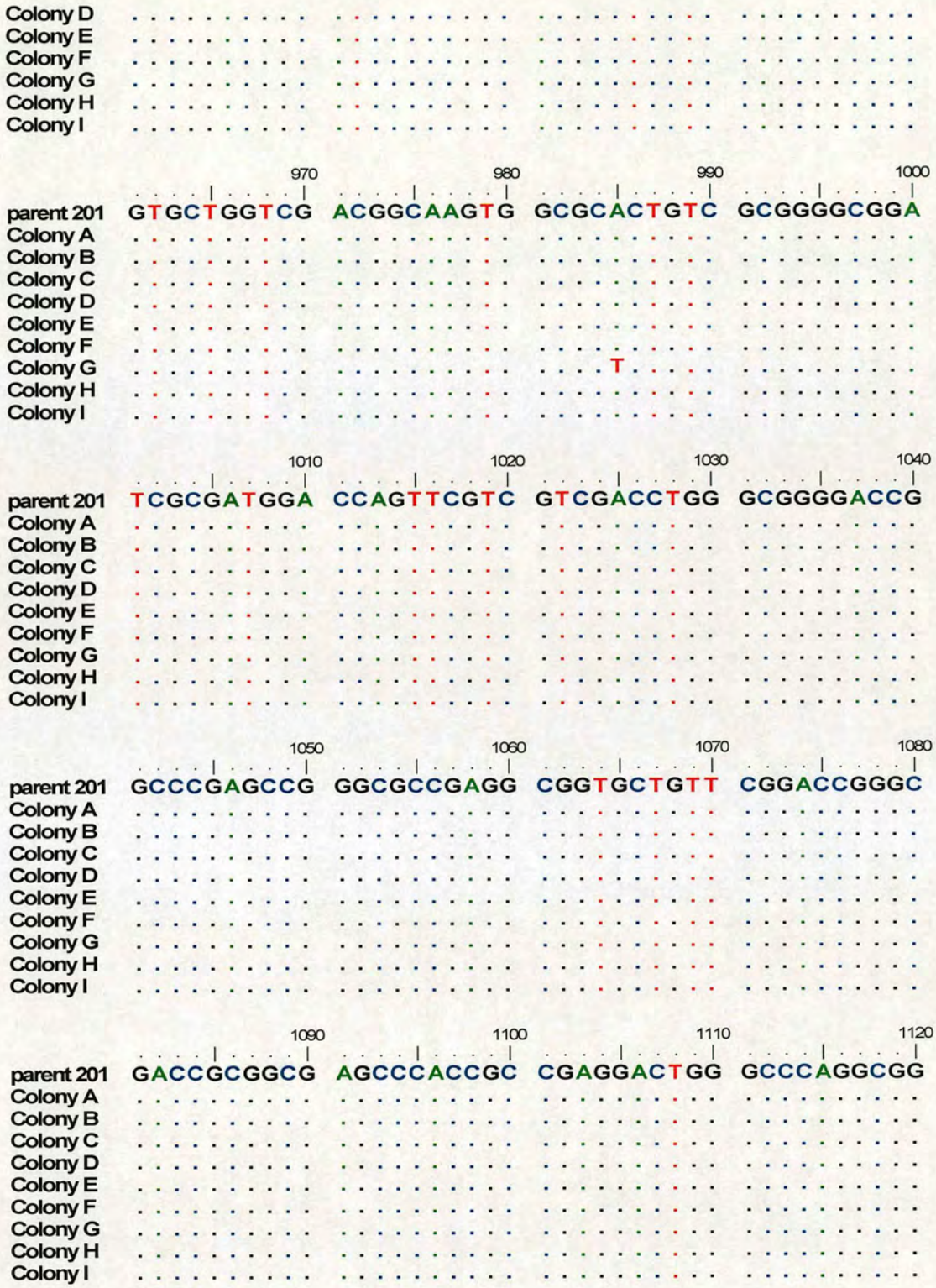

10 20 30 40
parent 201 GAT GAATTCGAGC GAGACAACTG CCCGGCGGGA
Colony A AAACAGC
Colony B AAACAGC
Colony C AAACAGC
Colony D AAACAGC
Colony E AAACAGC
Colony F AAACAGC
Colony G AAACAGC
Colony H AAACAGC
Colony I AAACAGC
50 60 70 80
parent 201 CGCGGACGCG GTACTGCGCG CCCGGGCCGA GATCGACCTG
Colony C G
Colony E A
90 100 110 120
parent 201 GCCGCCCTGC GGGCCAACGT ACGCGCACTG CGCGAACGGG
130 140 150 160
parent 201 CCCCCGGGGC GGCTCTCATG GCCGTCGTCA AGGCGGACGC
170 180 190 200
parent 201 CTACGGCCAC GGCGCGATCC CCTGCGCCCG TGCGGCCGTC

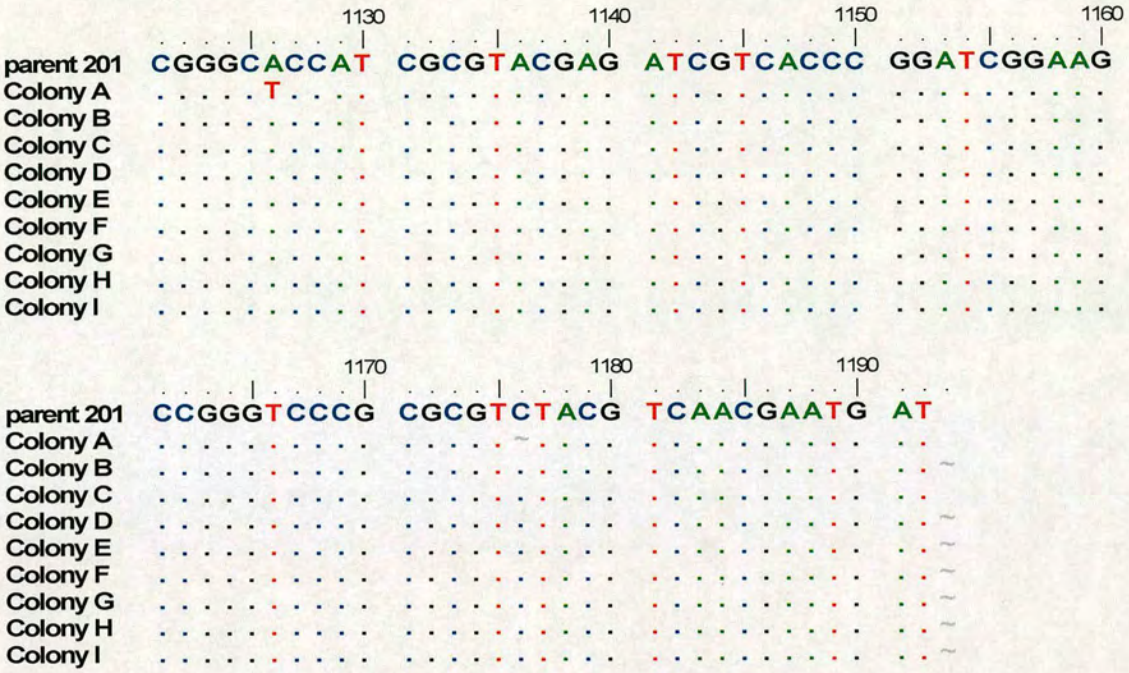

Colony H
Colony I
210
220
230
240
parent 201 GCGGCCGGTG CGACCTGGCT CGGCACCGCC ACCCCGCAGG
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
250
260
270
280
parent 201 AGGCCCTCGC CCTGCGCGCG GCGGAGCCAG GCCTGCCGGA
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
290
300
310
320
parent 201 CGACGTACGG ATCATGTGCT GGCTGTGGAC GCCCGGCGGG
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
330
340
350
360
parent 201 CCCTGGCGGG AGGCGGTCGA GGCCCGTCTC GACGTGTCCG
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
370
380
390
400
parent 201 TCAGCGCGAT GTGGGCCATG GAGGAGGTCA CCGGGGCGGC
Colony A
Colony B
Colony C
Colony D

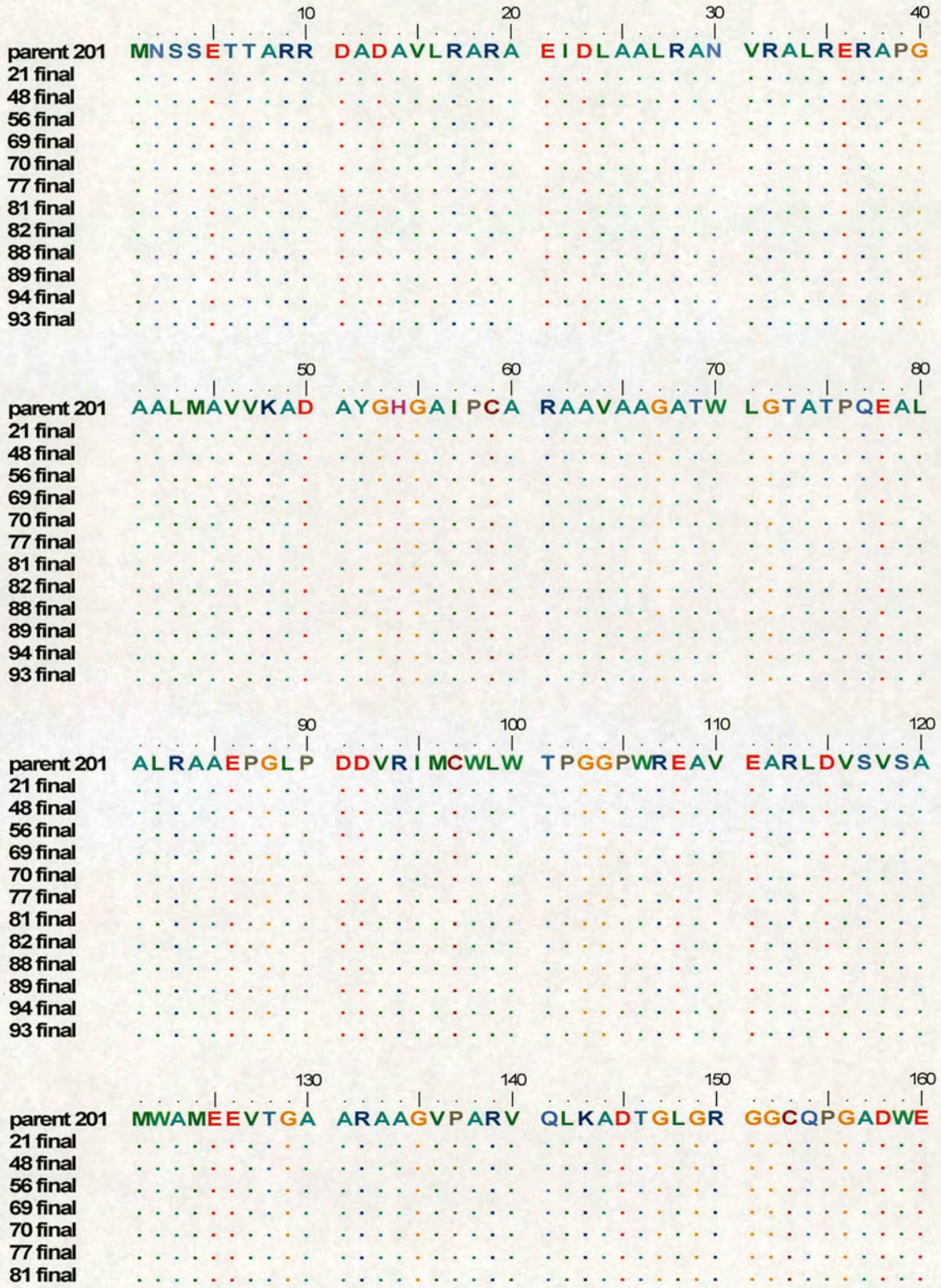

Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
410
420
430
440
parent 201 CCGGGCGGCC GGGGTGCCCG CGCGCGTGCA GCTCAAGGCC
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
450
460
470
480
parent 201 GACACCGGCC TCGGGCGGGG CGGCTGCCAG CCCGGCGCGG
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
490
500
510
520
parent 201 ACTGGGAGCG GCTGGTCGGG GCCGCGCTGC GGGCCGAGGA
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
530
540
550
560
parent 201 GGAGGGGCTG CTGAGAGTCA CCGGGCTCTG GTCCCACTTC
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I

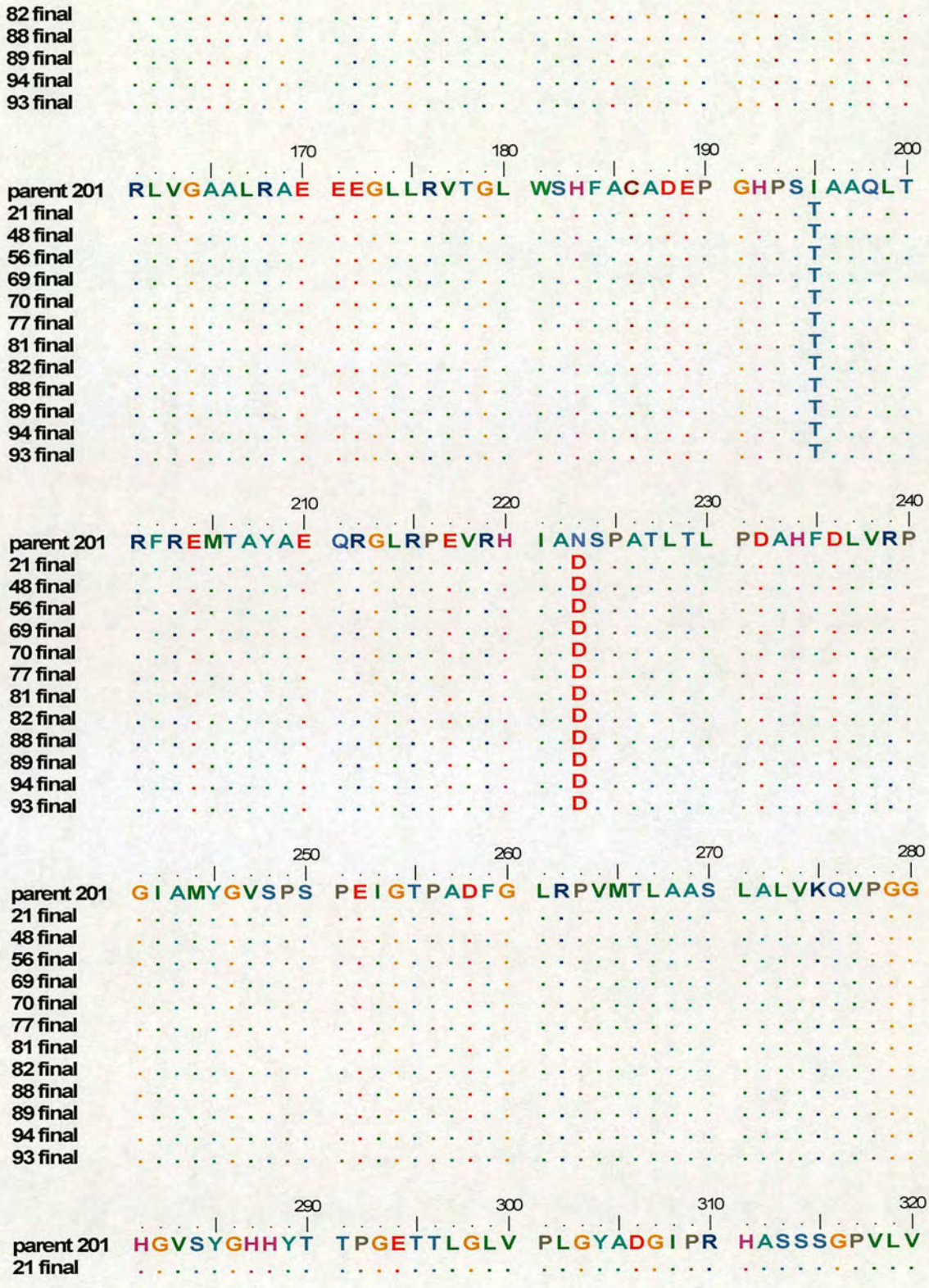

parent 201 GCCTGCGCCG ACGAGCCCGG CCACCCCTCC ATCGCCGCCC
Colony A . . . . . T . . . . . .
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
610 620 630 640
parent 201 AGCTCACCCG TTTCCGGG-A GATGACGGCC TACGCCGAGC
Colony A
Colony B
Colony C . . . . . T . . . . . .
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
650 660 670 680
parent 201 AGCGGGGCCT GCGTCCCGAG GTGCGGCACA TCGCCAACTC
Colony A
Colony B
Colony C . . . . . C . . . . . .
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
690 700 710 720
parent 201 GCCGGCCACC CTCACCCTCC CCGACGCCCA CTTCGACCTC
Colony A
Colony B
Colony C
Colony D
Colony E . . . . . T . . . . . .
Colony F
Colony G
Colony H
Colony I
730 740 750 760
parent 201 GTCCGGCCCG GCATCGCGAT GTACGGTGTC TCACCGAGCC
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G

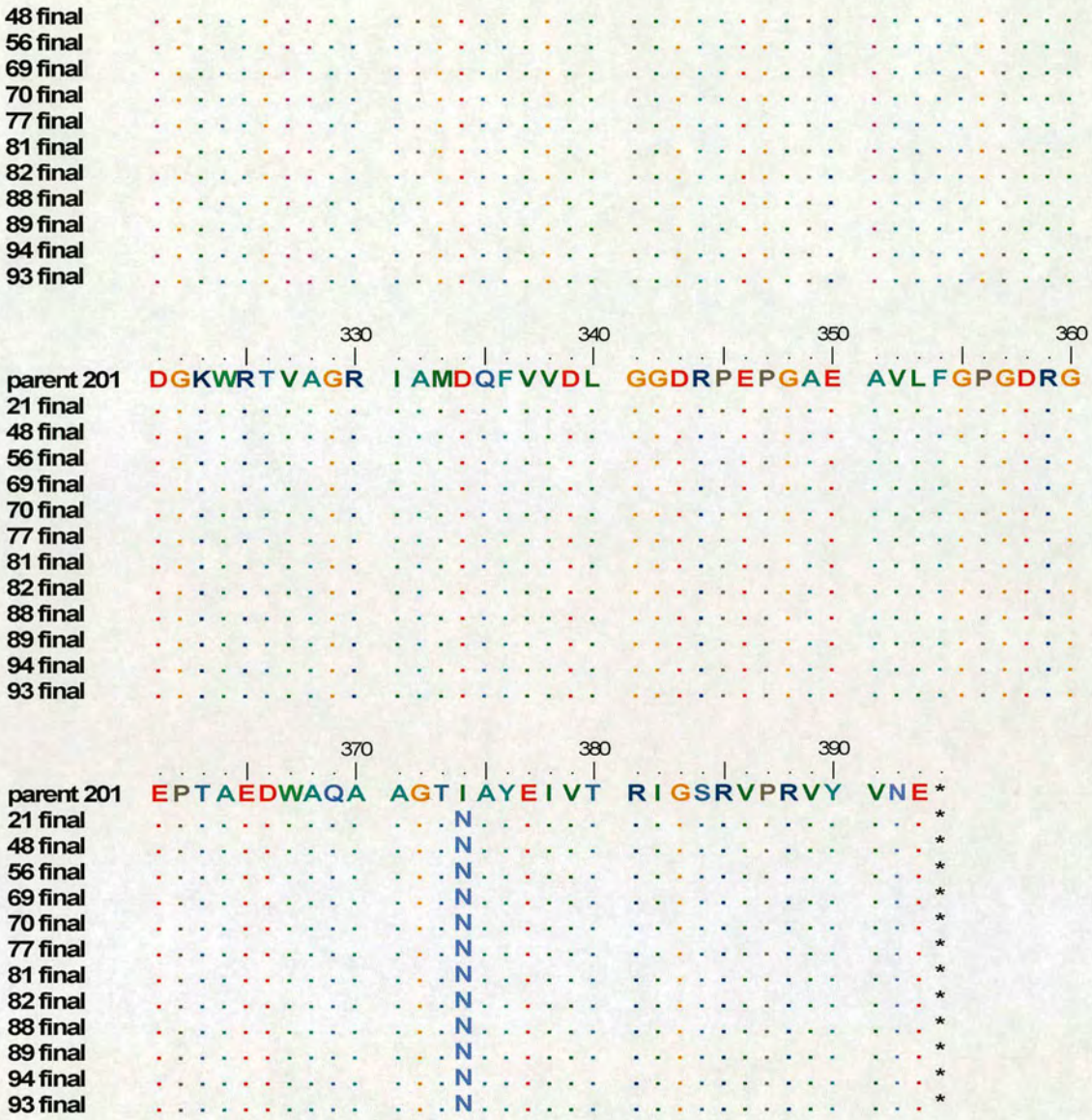

Colony H
Colony I
770
780
790
800
parent 201 CCGAGATCGG CACCCCGGCC GACTTCGGGC TGCGCCCCGT
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
810
820
830
840
parent 201 GATGACGCTG GCCGCCTCGC TGGCCCTGGT CAAGCAGGTC
Colony A
Colony B
Colony C
Colony D
Colony E A
Colony F
Colony G
Colony H
Colony I
850
860
870
880
parent 201 CCCGGCGGCC ACGGCGTCAG CTACGGGCAC CACTACACCA
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
890
900
910
920
parent 201 CCCCCGGCGA GACCACCCTC GGCCTCGTCC CGCTCGGCTA
Colony A
Colony B
Colony C
Colony D
Colony E T
Colony F
Colony G
Colony H
Colony I
930
940
950
960
parent 201 CGCGGACGGC ATCCCCCGGC ACGCCTCCTC CAGCGGCCCC
Colony A
Colony B
Colony C

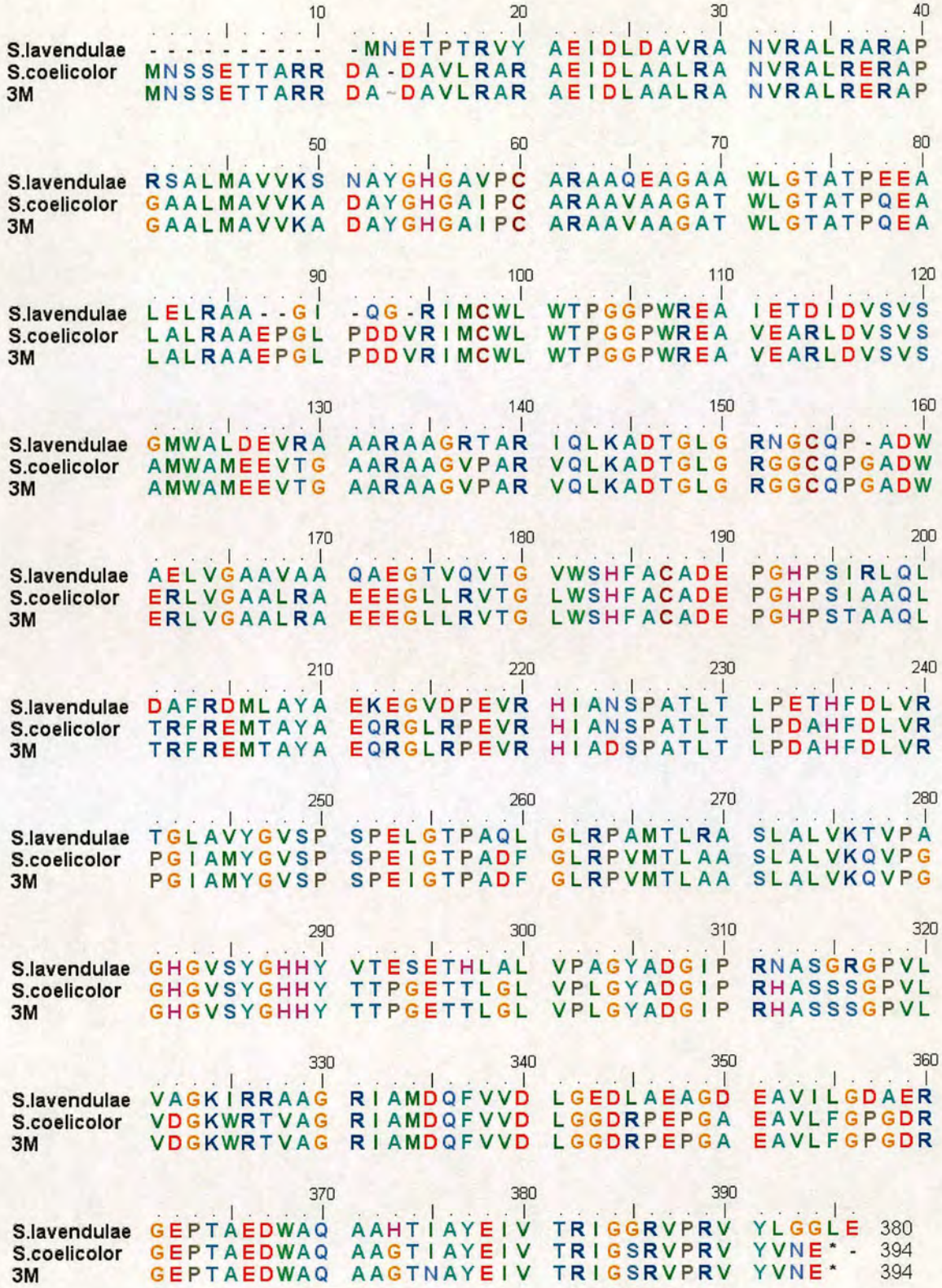

Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
970
980
990
1000
parent 201 GTGCTGGTCG ACGGCAAGTG GCGCACTGTC GCGGGGCGGA
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
1010
1020
1030
1040
parent 201 TCGCGATGGA CCAGTTCGTC GTCGACCTGG GCGGGGACCG
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
1050
1060
1070
1080
parent 201 GCCCGAGCCG GGCGCCGAGG CGGTGCTGTT CGGACCGGGC
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
1090
1100
1110
1120
parent 201 GACCGCGGCG AGCCCACCGC CGAGGACTGG GCCCAGGCGG
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I


1130
1140
1150
1160
parent 201 CGGGCACCAT CGCGTACGAG ATCGTCACCC GGATCGGAAG
Colony A . . . . . T . . . . . .
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I
1170
1180
1190
parent 201 CCGGGTCCCG CGCGTCTACG TCAACGAATG AT
Colony A
Colony B
Colony C
Colony D
Colony E
Colony F
Colony G
Colony H
Colony I


10 20 30 40
parent 201 MNSSETTARR DADAVLRARA EIDLAALRAN VRALRERAPG
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
50 60 70 80
parent 201 AALMAVVKAD AYGHGAIPCA RAAVAAGATW LGTATPQEAL
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
90 100 110 120
parent 201 ALRAAEPGLP DDVRIMCWLW TPGGPWREAV EARLDVSVSA
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
130 140 150 160
parent 201 MWAMEEVTGA ARAAGVPARV QLKADTGLGR GGCQPGADWE
21 final
48 final
56 final
69 final
70 final
77 final
81 final


82 final
88 final
89 final
94 final
93 final
170
180
190
200
parent 201 RLVGAALRAE EEGLLRVTGL WSHFACADEP GHPSIAAQLT
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
210
220
230
240
parent 201 RFREMTAYAE QRGLRPEVRH IANSPATLTL PDAHFDLVRP
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
250
260
270
280
parent 201 GIAMYGVSPS PEIGTPADFG LRPVMTLAAS LALVKQVPGG
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
290
300
310
320
parent 201 HGVSYGHHYT TPGETTLGLV PLGYADGIPR HASSSGPVLV
21 final


48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
330
340
350
360
parent 201 DGKWRTVAGR IAMDQFVVDL GGDRPEPGAE AVLFGPGDRG
21 final
48 final
56 final
69 final
70 final
77 final
81 final
82 final
88 final
89 final
94 final
93 final
370
380
390
parent 201 EPTAEDWAQA AGTIAYEIVT RIGSRVPRVY VNE*
21 final N *
48 final N *
56 final N *
69 final N *
70 final N *
77 final N *
81 final N *
82 final N *
88 final N *
89 final N *
94 final N *
93 final N *


10 20 30 40
S.lavendulae - - - - - - - - - - -MNETPTRVY AEIDLDAVRA NVRALRARAP
S.coelicolor MNSSETTARR DA-DAVLRAR AEIDLAALRA NVRALRERAP
3M MNSSETTARR DA-DAVLRAR AEIDLAALRA NVRALRERAP
50 60 70 80
S.lavendulae RSALMAVVKS NAYGHGAVPC ARAAQEAGAA WLGTATPEEA
S.coelicolor GAALMAVVKA DAYGHGAIPC ARAAVAAGAT WLGTATPQEA
3M GAALMAVVKA DAYGHGAIPC ARAAVAAGAT WLGTATPQEA
90 100 110 120
S.lavendulae LELRAA--GI -QG-RIMCWL WTPGGPWREA IETDIDVSVS
S.coelicolor LALRAAEPGL PDDVRIMCWL WTPGGPWREA VEARLDVSVS
3M LALRAAEPGL PDDVRIMCWL WTPGGPWREA VEARLDVSVS
130 140 150 160
S.lavendulae GMWALDEVRA AARAAGRTAR IQLKADTGLG RNGCQP-ADW
S.coelicolor AMWAMEEVTG AARAAGVPAR VQLKADTGLG RGGCQPGADW
3M AMWAMEEVTG AARAAGVPAR VQLKADTGLG RGGCQPGADW
170 180 190 200
S.lavendulae AELVGAAVAA QAEGTVQVTG VWSHFACADE PGHPSIRLQL
S.coelicolor ERLVGAALRA EEEGLLRVTG LWSHFACADE PGHPSIAAQL
3M ERLVGAALRA EEEGLLRVTG LWSHFACADE PGHPSTAAQL
210 220 230 240
S.lavendulae DAFRDMLAYA EKEGVDPEVR HIANSPATLT LPETHFDLVR
S.coelicolor TRFREMTAYA EQRGLRPEVR HIANSPATLT LPDAHFDLVR
3M TRFREMTAYA EQRGLRPEVR HIADSPATLT LPDAHFDLVR
250 260 270 280
S.lavendulae TGLAVYGVSP SPELGTPAQL GLRPAMTLRA SLALVKTVPA
S.coelicolor PGIAMYGVSP SPEIGTPADF GLRPVMTLAA SLALVKQVPG
3M PGIAMYGVSP SPEIGTPADF GLRPVMTLAA SLALVKQVPG
290 300 310 320
S.lavendulae GHGVSYGHHY VTESETHLAL VPAGYADGIP RNASGRGPVL
S.coelicolor GHGVSYGHHY TTPGETTLGL VPLGYADGIP RHASSSGPVL
3M GHGVSYGHHY TTPGETTLGL VPLGYADGIP RHASSSGPVL
330 340 350 360
S.lavendulae VAGKIRRAAG RIAMDQFVVD LGEDLAEAGD EAVILGDAER
S.coelicolor VDGKWRTVAG RIAMDQFVVD LGGDRPEPGA EAVLFGPGDR
3M VDGKWRTVAG RIAMDQFVVD LGGDRPEPGA EAVLFGPGDR
370 380 390
S.lavendulae GEPTAEDWAQ AAHTIAYEIV TRIGGRVPRV YLGGLE 380
S.coelicolor GEPTAEDWAQ AAGTIAYEIV TRIGSRVPRV YVNE*- 394
3M GEPTAEDWAQ AAGTNAYEIV TRIGSRVPRV YVNE* 394