# Recombinant transketolases with increased synthetic potential

A thesis submitted to the University of London for the degree of Doctor of Philosophy

by

**Sarah L. Bacon**

Department of Biochemistry and Molecular Biology
University College London
Gower Street
London
WC1E 6BT

August 2002

*For A.C.S.*

## Acknowledgements

Firstly I would like to thank my supervisor, Dr John Ward, for continual help and guidance throughout this project, and Dr Helen Hailes of the UCL Chemistry Department for advice on Chemistry matters.

Much appreciation also to the group:- Dunstan, Haitham, Jeannette, Lankani and Lisa for providing an endless source of amp plates, amusement and practical suggestions.

Finally, thank you to the Biotechnology and Biological Sciences Research Council for their support of this work.

# Abstract

This thesis describes the progress made towards the development of recombinant enzymes for use in the catalysis of a carbon-carbon bond-forming reaction. Transketolase (TK) is one such enzyme and has considerable synthetic potential as a stereospecific C-C bond-forming reagent.

Recombinant forms of TK from *Escherichia coli* have been produced, over-expressing to levels of 15-25% of the total cellular protein. The native enzyme is stereospecific for D-enantiomeric, linear sugar aldehyde acceptors. It also dictates the stereochemistry of the two new chiral centres in the reaction product. Furthermore, it has been shown to have activity with a range of simple α-unsubstituted aldehydes, so demonstrating a low specificity for the aldehyde component. Combined with the high stereoselectivity for the C-C bond-forming step, this makes TK an attractive catalyst for asymmetric synthesis.

Using the enzyme crystal structure as a model for their design, a selection of TK mutants were generated with assorted combinations of amino acid changes located within the active site funnel. The synthetic potential of these modified TKs was analysed combinatorially with a range of aromatic aldehydes, using β-hydroxypyruvic acid as the ketol donor. Relaxation or reversal of the TK requirement for a D- α-hydroxyl group would provide even greater scope for its syntheic applications. Variants tailored for these needs were constructed and characterised using D-glyceraldehyde and DL-glyceraldehyde as aldehyde acceptors. In addition, an ion-exchange chromatographic method of purification was developed for a poly-lysine tagged version of TK, affording a 4.5-fold purification of the enzyme. The TK thus produced is available for subsequent investigations into its specific reactivity with a range of substrates.

**Acknowledgements** ..... 3
**Abstract** ..... 4
**Contents** ..... 5
**Abbreviations** ..... 11
**Units** ..... 12
**Symbols for Amino Acids** ..... 12

## Chapter One: Introduction

**1.1 Introduction** ..... 14
**1.2 Enzymes as industrial catalysts** ..... 14
**Figure 1.1** Structures of some non-steroidal anti-inflammatory drugs ..... 18
**Figure 1.2** Aldol Condensation ..... 19
**1.3 Transketolase** ..... 20
1.3.1 Structure ..... 21
**Figure 1.3** *E. coli* transketolase ..... 22
**Figure 1.4** Thiamine pyrophosphate ..... 23
1.3.2 Function ..... 25
**Figure 1.5** The pentose phosphate pathway ..... 26
**Figure 1.6** General transketolase reaction ..... 27
**Figure 1.7** TK reaction with β-HPA as ketol donor ..... 28
**Figure 1.8** Mechanism of action of TPP in TK-catalysed reaction ..... 30
**Figure 1.9** Suggested mechanism for deprotonation of TPP at C(2) ..... 32
**1.4 Use of TK for organic synthesis** ..... 33
**Figure 1.10** Relative rates ($V_{rel}$) for a selection of aldehyde acceptors ..... 36
**1.5 Production of constructs for over-expression of TK** ..... 37
**Table 1.1** TK production in shake flask culture ..... 38
**1.6 Design of TKs with broad industrial potential** ..... 39
**Figure 1.11** Active site of transketolase ..... 39
**1.7 Project aims** ..... 40
1.7.1 Purification of transketolase ..... 40
1.7.2 Expansion of transketolase active site ..... 40
1.7.3 Relaxation of transketolase stereoselectivity ..... 40

## Chapter Two: Equipment and General Techniques

**2.1 Equipment** ..... 43
2.1.1 Chromatographic apparatus ..... 43
2.1.2 Spectroscopic apparatus ..... 43
2.1.3 Additional apparatus ..... 44
**2.2 General Techniques** ..... 45
2.2.1 Linked-enzyme assay ..... 45
**Figure 2.1** Linked-enzyme assay for TK ..... 46
2.2.2 Protein quantitation ..... 47
2.2.3 Alkaline lysis plasmid preparation ..... 48
2.2.4 DNA reaction work up ..... 49
2.2.5 Glycerol freeze stock preparation ..... 49
2.2.6 Competent cell preparation ..... 49
2.2.7 Transformation of DNA ..... 50
2.2.8 Quantification of oligonucleotides and plasmid DNA ..... 50
2.2.9 Restriction endonuclease digests ..... 50

2.2.10 Purification of ssDNA ..... 51
2.2.11 Horizontal agarose gel electrophoresis ..... 52
2.2.12 SDS polyacrylamide gel electrophoresis (SDS-PAGE) ..... 52
2.2.13 DNA ligation reactions ..... 53
2.2.14 Annealing of oligonucleotides to ssDNA ..... 53
2.2.15 Synthesis and recircularisation ..... 53
2.2.16 Polymerase chain reaction (PCR) ..... 54
2.2.17 HPLC analysis ..... 54
2.2.18 Ion exchange chromatography ..... 55
2.2.19 Biotransformations for HPLC analysis ..... 56
2.2.20 Proof of transketolase stereoselectivity ..... 56
2.2.21 Bacterial strain and growth conditions ..... 57
2.2.22 Bacterial extracts preparation ..... 57

## Chapter Three: Expression, purification and characterisation of lysine-tagged *E. coli* TK by ion exchange chromatography

**3.1 Introduction** ..... 59
**3.2 Construction of poly-lysine tagged TK** ..... 59
**Figure 3.1** Scheme for construction of pQR730 ..... 60
**Figure 3.2** Plasmid map of pQR730 ..... 61
**3.3 First stage of purification** ..... 61
**Figure 3.3** Pre-purification step ..... 62
**Figure 3.4** First stage of purification ..... 62
**Table 3.1** Specific activities for first stage of purification ..... 63
**3.4 Second stage of purification** ..... 63
**Figure 3.5** Second stage of purification ..... 64
**Table 3.2** Specific activities for second stage of purification ..... 64
**Figure 3.6** SDS-PAGE showing purified TK ..... 65
**Table 3.3** Purification of TK (from 200ml *E.coli* culture) ..... 65
**Table 3.4** Purification of recombinant TK from *E. coli* JM107 cells ..... 66
**3.5 Structure of purified transketolase** ..... 66
**Table 3.5** Investigation into enzyme structure ..... 67
**Table 3.6** Proportion of apo-:holo-enzyme ..... 67
**3.6 Discussion** ..... 67

## Chapter Four: Applications of modified TK for synthesis using aromatic aldehydes

**4.1 Introduction** ..... 72
**4.2 Initial TK mutants** ..... 72
**Figure 4.1** Plasmid map of pQR706 ..... 73
**Table 4.1** Transketolase Mutants (constructed by Carol French) ..... 74
**Table 4.2** Specific activities for TK mutants ..... 76
**Figure 4.2** General scheme for TK reaction with aromatic aldehydes ..... 77
**4.3 HPLC analysis** ..... 77
**Figure 4.3** Aromatic aldehydes chosen for analysis with modified TK ..... 78
**Figure 4.4** Standard HPLC curve for β-HPA ..... 79
**Figure 4.5** Loss in β-HPA over 18h for each of the 8 constructs ..... 81
**Figure 4.6** Loss in β-HPA over 18h for each of the 6 aromatic aldehydes ..... 81
**4.4 Product extraction and characterisation** ..... 83
4.4.1 NMR analysis ..... 83

**Figure 4.7** $^1$H NMR for pyridoxal ..... 85
**Figure 4.8** $^1$H NMR for pyridoxal reaction ..... 85
**Figure 4.9** $^1$H NMR for 2-pyridinecarboxaldehyde ..... 88
**Figure 4.10** $^1$H NMR for 2-pyridinecarboxaldehyde reaction ..... 88
**Figure 4.11** $^1$H NMR for 3-pyridinecarboxaldehyde ..... 91
**Figure 4.12** $^1$H NMR for 3-pyridinecarboxaldehyde reaction ..... 91
**Figure 4.13** $^1$H NMR for 4-pyridinecarboxaldehyde ..... 93
**Figure 4.14** $^1$H NMR for 4-pyridinecarboxaldehyde reaction ..... 93
**Figure 4.15** $^1$H NMR for 2-furancarboxaldehyde ..... 94
**Figure 4.16** $^1$H NMR for 2-furancarboxaldehyde reaction ..... 94
**Figure 4.17** $^1$H NMR for 2-pyrrolecarboxaldehyde ..... 95
**Figure 4.18** $^1$H NMR for 2-pyrrolecarboxaldehyde reaction ..... 95
**Table 4.3** Conversion efficiencies calculated from NMR integration ..... 96
**Figure 4.19** Electron donation by heteroatom ..... 97
4.4.2 Mass spectrometry analysis ..... 98
**Figure 4.20** Mass spectrum for 2-pyridinecarboxaldehyde ..... 100
**Figure 4.21** Mass spectrum for 2-pyridinecarboxaldehyde reaction ..... 100
**Figure 4.22** Mass spectrum for 3-pyridinecarboxaldehyde ..... 101
**Figure 4.23** Mass spectrum for 3-pyridinecarboxaldehyde reaction ..... 101
**Figure 4.24** Mass spectrum for 4-pyridinecarboxaldehyde ..... 102
**Figure 4.25** Mass spectrum for 4-pyridinecarboxaldehyde reaction ..... 102
**Table 4.4** Mass spectrometry peaks for complexed and uncomplexed reaction products ..... 103
**Table 4.5** Proportion of each TK reaction product complexed to sodium. 103
**Figure 4.26** Complexing of TK reaction products with sodium ions ..... 104
**Table 4.6** Accurate mass data for TK from pQR739/pyridinecarboxaldehyde reaction ..... 106
**4.5 Further mutants** ..... 106
**Table 4.7** Additional TK mutants ..... 106
**Figure 4.27** Scheme for mutant construction ..... 107
**Table 4.8** Specific activities for TK mutants ..... 108
**4.6 Discussion** ..... 109

## Chapter Five: Modification of TK stereoselectivity

**5.1 Introduction** ..... 113
**5.2 Initial TK mutants** ..... 114
**5.3 Reaction with glycolaldehyde** ..... 114
**Figure 5.1** Reaction of β-HPA with glycolaldehyde ..... 115
**Figure 5.2** Standard HPLC curve for β-HPA/GA/L-E ..... 116
**Figure 5.3** Loss in β-HPA/GA and gain in L-E for wtp reaction ..... 118
**Figure 5.4** Loss in β-HPA/GA and gain in L-E for S385*L reaction ..... 118
**Figure 5.5** Loss in β-HPA/GA and gain in L-E for H461*L reaction ..... 119
**Figure 5.6** Loss in β-HPA/GA and gain in L-E for H461*S reaction ..... 119
**Figure 5.7** Loss in β-HPA/GA and gain in L-E for I189A reaction ..... 120
**Figure 5.8** Loss in β-HPA/GA and gain in L-E for I189V reaction ..... 120
**Figure 5.9** Loss in β-HPA/GA and gain in L-E for H258A reaction ..... 121
**Figure 5.10** Loss in β-HPA/GA and gain in L-E for H258D reaction ..... 121
**Table 5.1** TK specific activities for transformation with glycolaldehyde.. 122
**5.4 Reaction with D-glyceraldehyde and DL-glyceraldehyde** ..... 125
**Figure 5.11** Reaction of β-HPA with D-/L-glyceraldehyde ..... 125
**Figure 5.12** Standard HPLC curve for β-HPA/D-gla/DL-gla ..... 127
**Figure 5.13** Loss in β-HPA/D-gla/DL-gla for wtp reaction ..... 128
**Figure 5.14** Loss in β-HPA/D-gla/DL-gla for S385*L reaction ..... 129
**Figure 5.15** Loss in β-HPA/D-gla/DL-gla for H461*L reaction ..... 130
**Figure 5.16** Loss in β-HPA/D-gla/DL-gla for H461*S reaction ..... 131
**Figure 5.17** Loss in β-HPA/D-gla/DL-gla for I189A reaction ..... 132

Figure 5.18 Loss in β-HPA/D-gla/DL-gla for I189V reaction ..... 133
Figure 5.19 Loss in β-HPA/D-gla/DL-gla for H258A reaction ..... 134
Figure 5.20 Loss in β-HPA/D-gla/DL-gla for H258D reaction ..... 135
Table 5.2 TK specific activities for reactions with D-glyceraldehyde and DL-glyceraldehyde ..... 136
Table 5.3 DL-glyceraldehyde specific activities relative to those for D-glyceraldehyde ..... 137
Figure 5.21 Hypothetical picture for relaxation of TK stereoselectivity ..... 138
Figure 5.22 Alternative structures for D-xylulose and L-ribulose ..... 140
5.5 **Further mutants** ..... 141
Table 5.4 TK stereoselectivity mutants ..... 141
Figure 5.23 Hypothetical active site stereoselectivity control showing key enzyme/substrate interactions ..... 142
Table 5.5 Specific activities for TK mutants ..... 143
5.6 **Discussion** ..... 144


## Chapter Six: Experimental


6.1 **Preparation of uracil-containing pQR706/776** ..... 149
Table 6.1 Template ssDNA concentrations ..... 149
6.2 **Attempted construction of pQR780/781/783/784** ..... 149
Table 6.2 Oligonucleotide concentrations and melting temperatures ..... 149
Table 6.3 Mutant construction reaction mixtures ..... 150
6.3 **PCR mutagenesis of TK gene** ..... 150
Table 6.4 Oligonucleotide concentrations and melting temperatures ..... 150
Table 6.5 Template DNA ..... 151
6.4 **Construction of pQR780/781/782/783/784/785/786/787/788/789** ..... 151
Table 6.6 Template/primer combinations for mutant construction ..... 151
Table 6.7 Mutant TK sequencing primers ..... 152
6.5 **Expression of proteins in *E. coli*** ..... 152
6.5.1 Production of soluble protein ..... 152
6.5.2 Protein purification by ion-exchange chromatography ..... 153
6.6 **Biotransformations for HPLC analysis** ..... 153
Table 6.8 Substrates (100mM) for analytical biotransformations ..... 154
6.7 **Extraction of biotransformation products** ..... 154
Table 6.9 $R_f$ values derived from TLC analysis ..... 155


## Chapter Seven: Discussion


7.1 **Introduction** ..... 157
7.2 **Purification of transketolase** ..... 157
7.3 **Expansion of transketolase active site** ..... 159
7.4 **Modification of transketolase stereoselectivity** ..... 169
7.5 **Conclusion** ..... 174

**References** ..... 176

## Appendices


**Appendix One: Media, reagents and buffers** ..... 189

**A1.1 Suppliers of materials** ..... 189

**A1.2 Microbiological techniques** ..... 189

**A1.3 Microbiological strains** ..... 190

**A1.4 Microbiological growth media**

A1.4.1 LB medium ..... 190

A1.4.2 Minimal media for amber strains ..... 191

**A1.5 Solutions for plasmid preparation by alkaline lysis**

A1.5.1 QPP1: Resuspension buffer ..... 191

A1.5.2 QPP2: Alkaline lysis buffer ..... 191

A1.5.3 QPP3: Neutralisation buffer ..... 191

**A1.6 Solutions for preparation of ssDNA**

A1.6.1 Polyethylene glycol (PEG) solution ..... 191

A1.6.2 Sodium acetate solution ..... 192

**A1.7 Enzyme reaction buffers**

A1.7.1 NEB buffer 1 ..... 192

A1.7.2 NEB buffer 2 ..... 192

A1.7.3 NEB buffer 3 ..... 192

A1.7.4 NEB buffer 4 ..... 192

A1.7.5 T4 DNA polymerase buffer ..... 193

A1.7.6 T4 ligase buffer ..... 193

A1.7.7 10x Reaction buffer ..... 193

**A1.8 Nucleic acid buffers and solutions**

A1.8.1 TE (Tris/EDTA) buffer ..... 193

A1.8.2 TBE (Tris/borate/EDTA) buffer ..... 193

A1.8.3 Buffered phenol ..... 194

A1.8.4 Phenol/chloroform ..... 194

A1.8.5 6x agarose gel sample loading buffer ..... 194

A1.8.6 1kb DNA ladder ..... 194

A1.8.7 7.5% acrylamide stock ..... 195

A1.8.8 Polyacrylamide gels ..... 195

**A1.9 SDS-PAGE protein electrophoresis solutions**

A1.9.1 Electrode buffer: stock solution ..... 195

A1.9.2 Electrode buffer: working solution ..... 195

A1.9.3 Stock running gel buffer (0.75M, Tris base pH8.3) ..... 195

A1.9.4 Stock stacking gel buffer (1.25M, Tris base pH6.8) ..... 196

A1.9.5 Ammonium persulphate ..... 196

A1.9.6 Sample loading buffer ..... 196

A1.9.7 Stacking gel (6%) ..... 196

A1.9.8 Separating gel (15%) ..... 196

A1.9.9 Coomassie blue stain solution (0.2% w/v) ..... 197

A1.9.10 Mark12™ wide range protein standard ..... 197

A1.9.11 Fast destain solution ..... 197

A1.9.12 Slow destain solution ..... 197

**A1.10 HPLC buffers**

A1.10.1 $H_2SO_4$ (6mM) ..... 198

**A1.11 FPLC elution buffers**

A1.11.1 Gly-Gly (150mM, pH9.5) ..... 198

A1.11.2 Gly-Gly (150mM, pH7.0) ..... 198

A1.11.3 MES (150mM, pH5.0) ..... 198

A1.11.4 NaCl (1M, pH7.5) ..... 198

**A1.12 FPLC wash buffers**

A1.12.1 NaCl (2M) ..... 199

A1.12.2 NaOH (2M) ..... 199

A1.12.3 75% acetic acid ..... 199

**A1.13 Antibiotics** ..... 199

**A1.14 Miscellaneous solutions**

A1.14.1 10% SDS ..... 199

A1.14.2 1M Tris.HCl ..... 199

A1.14.3 $CaCl_2$ (0.1M)... ....200
A1.14.4 10x annealing buffer....200
A1.14.5 Gly-Gly (100mM, pH7.6)....200
A1.14.6 Gly-Gly (70mM, pH7.6)....200
A1.14.7 Permanganate stain solution ....200
A1.14.8 Anisaldehyde stain solution....200
A1.14.9 IPTG solution (100mM) ....201
A1.14.10 EDTA (0.5M) ....201

**Appendix Two: Sequences of oligonucleotides**....202
**A2.1 Mutagenesis primers**... ....202
**A2.2 Sequencing primers**..... ....203
**A2.3 Primers for construction of additional mutants described in this thesis**....204

**Appendix Three: Nucleotide sequence of the *Escherichia coli* K-12 transketolase (tkt) gene**....205

# Abbreviations

| | |
|---|---|
| **A** | adenine |
| **amp** | ampicillin |
| **β-ME** | β-mercaptoethanol |
| **BSA** | bovine serum albumin |
| **C** | cytosine |
| **cm** | chloramphenicol |
| **D-gla** | D-glyceraldehyde |
| **DL-gla** | DL-glyceraldehyde |
| **DNA** | deoxyribonucleic acid |
| **dsDNA** | double-stranded DNA |
| ***E. coli*** | *Escherichia coli* |
| **EDTA** | ethylenediaminetetraacetic acid |
| **ee** | enantiomeric excess |
| **FPLC** | fast protein liquid chromatography |
| **G** | guanine |
| **GA** | glycolaldehyde |
| **gdDNA** | gapped-duplex DNA |
| **Gly-Gly** | Glycyl-Glycine |
| **HPLC** | high performance liquid chromatography |
| **β-HPA** | β-hydroxypyruvic acid |
| **IPTG** | isopropyl-β-D-thiogalactopyranoside |
| **kan** | kanomycin |
| **kb** | kilo base(s) |
| **L-E** | L-erythrulose |
| **LEA** | linked-enzyme assay |
| **MES** | 4-morpholineethanesulphonic acid |
| **M13 IG** | M13 intergenic region |
| **NMR** | nuclear magnetic resonance |
| **OD** | optical density |
| **O/N** | over night |
| **NADH** | nicotinamide adenine dinucleotide (reduced form) |
| **PCR** | polymerase chain reaction |
| **PEG** | polyethylene glycol |
| **sdm** | site-directed mutagenesis |
| **SDS** | sodium dodecyl sulphate |
| **ssDNA** | single-stranded DNA |
| **T** | thymine |
| **TBE** | Tris, borate, EDTA buffer |
| **TE** | Tris, EDTA buffer |
| **TEMED** | N, N, N', N'-tetramethylethylene diamine |
| **TK** | transketolase |
| **TLC** | thin layer chromatography |
| **TPP** | thiamine pyrophosphate |
| **Tris** | tris(hydroxymethyl)aminoethane |
| **U** | uracil |
| **UV** | ultra violet |
| **Xgal** | 5-Bromo-4-Chloro-3-Indolyl-β-D-Galactopyranosideside |

## Units

| | | | |
|---|---|---|---|
| **°C** | degrees centigrade | **M** | molar |
| **h** | hours | **mol** | moles |
| **kDa** | kilo daltons | **rpm** | revolutions per minute |
| **l** | litre | **sec** | seconds |
| **mg** | milligrams | **%** | percentage |
| **μg** | micrograms | **% v/v** | percentage volume by volume |
| **min** | minutes | **% w/v** | percentage weight by volume |
| **ml** | millilitre | **ppm** | parts per million |
| **mM** | millimolar | **U** | units of enzyme (μmol/min) |

## Symbols for Amino Acids

| | | | | | |
|---|---|---|---|---|---|
| A | Ala | Alanine | M | Met | Methionine |
| C | Cys | Cysteine | N | Asn | Asparagine |
| D | Asp | Aspartic acid | P | Pro | Proline |
| E | Glu | Glutamic acid | Q | Gln | Glutamine |
| F | Phe | Phenylalanine | R | Arg | Arginine |
| G | Gly | Glycine | S | Ser | Serine |
| H | His | Histidine | T | Thr | Threonine |
| I | Ile | Isoleucine | V | Val | Valine |
| K | Lys | Lysine | W | Trp | Tryptophan |
| L | Leu | Leucine | Y | Tyr | Tyrosine |

# Chapter One: Introduction

# CHAPTER ONE: Introduction

## 1.1 Introduction

Production of pharmaceuticals is an extensive industry and the increasing demand for the generation of novel substrates presents major medical and intellectual challenges. Structural requirements for biological activity often dictate the presence of one or more chiral centres in a drug and although many of these compounds are marketed as racemates, the enantiomers must, particularly from the biological point of view, be regarded as individual substances. This is important, since the neglect of stereochemistry in the development and application of drugs and bioactive agents can lead to serious complications and is a source of problems in pharmacokinetics.

As regulating factors concerning the production and marketing of pharmaceuticals become more severe, the importance of the manufacture of homochiral products continues to expand. This has led to an increased move by pharmaceutical companies towards the production of single isomers. Combined with the requirement for shorter, more economic syntheses as a result of enhanced competition within the chemical and pharmaceutical industries, this has extended the significance of enzymes and their industrial potential (Faber & Franssen, 1993). Conventional synthetic organic chemistry has been successful in the production of many pharmaceutical compounds, but has failed to provide a satisfactory solution in numerous cases, particularly concerning such issues as product yield and purity. Hence there is a need for an alternative approach to complement existing traditional synthetic methods. The use of biotransformations to carry out synthetically useful organic reactions is becoming increasingly popular (Davies *et al.*, 1989) and is introduced in the following section.

## 1.2 Enzymes as industrial catalysts

Enzymes are the biological catalysts necessary for the biochemical processes of life. Often an enzyme will not only transform the substrate, or substrates, involved *in vivo* but will (*in vitro*) accept a whole range of substrates and transform these with the same regio- and stereoselectivity, a feature commonly exploited by industry. Biocatalysts frequently carry out to high enantiomeric purity in a single step, a transformation that may require a number of chemical stages. Providing substrates satisfy certain physical and chemical criteria dictated by the conformation of the active site of the enzyme, conversion of a wide range of potentially useful compounds may be accomplished.

These characteristics render certain enzymes of great potential for catalytic roles within chemical syntheses.

Over recent decades industry has increasingly harnessed enzymes for the synthesis of commercially important products. Due to their organic nature, the reaction conditions for functional enzymatic catalysis are constrained within narrow operational limits. However, unlike chemical catalysis, these conditions are frequently relatively mild, being limited to ambient temperatures and pressures and near neutral pH. This is widely exploited for organic synthesis and is of central importance in the modification of chiral drugs, since it allows for selective transformations without jeopardising the existing chiral centres and labile chemical bonds. Additionally, a range of biocatalysts may possibly be used in combination, carrying out processes requiring sequential or parallel enzyme-catalysed conversions.

Enzymatic catalysis is inherently more selective than chemical synthesis, resulting in reduced side reactions, easier separation and potentially higher economic efficiency. In organic synthesis, enantiomers are normally obtained in a ratio of 1:1, whereas the products of biological origin, such as hormones and antibiotics *etc.* are generally obtained in stereospecific form or in a substantial enantiomeric excess of one particular form. Largely for economic reasons, synthetic products such as drugs and insecticides have been widely marketed as racemic mixtures. One of the most valuable synthetic features of enzymes is their ability to discriminate between the individual enantiomers of racemic substrates (Whitesides & Wong, 1985; Jones, 1986), commonly delivering a particular stereoisomer in large enantiomeric excess (% ee). Furthermore, in many cases the stereoselectivity of a given enzyme in organic solvent is identical to that in aqueous solution (Laumen *et al.*, 1988; Alder *et al.*, 1989). Due to an increasing demand for enantiomerically pure compounds, for example those to be utilised as pharmaceutical products, the benefits of enzymes are immense since they can be used to set the stereochemistry of key intermediates. Environmental issues are also in favour of enzymes as they help to reduce the quantity of hazardous compounds, organic solvents, heavy metals and other chemical waste. Especial benefits are found in the field of carbohydrate chemistry, where the inherent multi-functionality of sugars pose an enormous challenge for the organic chemist who has to employ a whole range of protective groups, preventing unwanted reactions involving the hydroxyl, keto, or phosphate groups.

Within an organism, many fundamental processes relevant to drug action, for example inhibition of enzymes, transport across membranes and binding to receptors, depend on the stereochemistry of the drug (Testa & Mayer, 1988; Beckett, 1991). There are numerous examples (Ariens, 1986) in which one enantiomer of a chiral drug has a therapeutically useful action and the other does not, one enantiomer is toxic and the other is safe, or one enantiomer is agonist and the other is antagonist.

Despite these facts, until recently the majority of synthetic chiral drugs were marketed as chiral mixtures (Ariens *et al.*, 1988). However, the proportion of synthetic chiral pharmaceuticals introduced as single enantiomers is expanding rapidly and two major reasons make this growth possible. Firstly the regulatory climate has changed, meaning that chiral drugs will no longer be considered as individual entities but as combinations of drugs (Cayen, 1991; Cartwright, 1991). This will require the implementation of expensive validation studies to acquire pharmacological and toxicological data on each individual enantiomer. Secondly, recent advances in the synthesis of optically pure compounds, for example chemical asymmetric synthesis (Brown & Davis, 1989), catalytic kinetic resolution (Kagan & Fiaud, 1988), resolution of racemates via stereoselective crystallization (Sheldon, 1990) and chiral chromatography (Krstulovic, 1989), have enabled the preparation of larger quantities of single isomers. Among these methods, biocatalytic procedures are playing an increasingly prominent role.

Although chemical and biological syntheses can be viewed as competing technologies, there are many processes, particularly in the food and pharmaceutical industries, where biocatalysis has no chemical competition. It is however, the enantioselective nature of enzyme catalysed reactions which is of greatest interest, particularly for the pharmaceutical industry, not only for the synthesis of new drugs but for improving the formulation of existing drugs, especially drug racemates of which there are over 500 currently marked (Cayen, 1991). With advances in chemoenzymatic technology, asymmetric synthesis (Ward, 1990; Crosby, 1991), stereoselective crystallisation (Gottarelli & Spada, 1991; McBride & Carter, 1991) and chiral analytical techniques (Morrison & Scott, 1985; Ferringa *et al.*, 1986; Cannarsa, 1996), regulatory pressures are driving the pharmaceutical industry towards the manufacture of single isomer drugs. In many cases biotransformation will be the only suitable methodology and it was projected that 80% of all new synthetic drugs would be homochiral by the year 2000 (Margolin, 1993).

Changes in the technology for production of enzymes and for their stabilization and manipulation now make these catalysts attractive for wider use in large-scale organic chemistry. Plant, animal and microbial cells are all sources of industrially useful biocatalysts, with the microbial (bacterial, yeast, and fungal) enzymes typically accounting for 80% of total industrial production (Godfrey & Reichelt, 1983). Microbial cells are relatively simple to grow, exhibit high growth rates and are extremely diverse in species. In addition, this group is receptive to genetic manipulation, which can result in enzyme over-expression (Tien & Tu, 1987) and permits catalyst modification by protein engineering techniques (Wetzel, 1986).

Non-steroidal anti-inflammatory agents (see Figure 1.1) are widely used drugs for treatment of human connective tissue diseases such as arthritis (Shen, 1972), with the activity of agents belonging to the group of 2-arylpropionic acid derivatives residing in the (S)-enantiomers. In the case of naproxen (A), the (S)-enantiomer is 28 times more active *in vivo* than its R-counterpart (Rozowski *et al.*, 1971). Among other nonsteroidal anti-inflammatory drugs are ibuprofen (B) (Ahmar *et al.*, 1989) and suprofen (C) (Sih *et al.*, 1988). Drugs of this group are bulky, sterically hindered acids and intermediates in their production pose considerable steric challenges, a feature typical of many pharmaceutical compounds. This requires consideration in the development of chemoenzymatic syntheses for these compounds and modification of the biocatalyst, to accommodate such substrates, may be necessary.

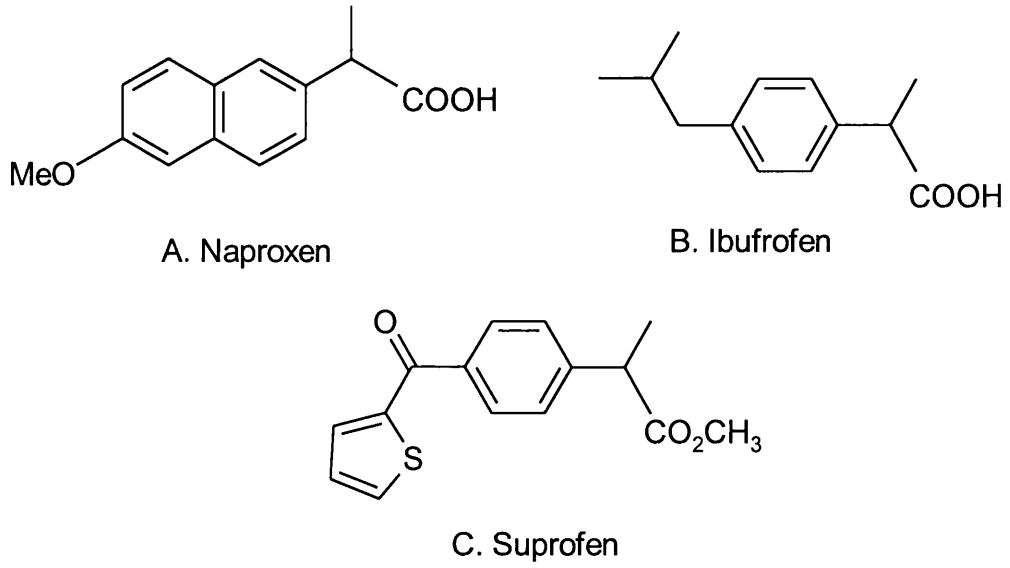



**Figure 1.1** *Structures of some non-steroidal anti-inflammatory drugs*

By varying the stereochemistry of drug molecules it is possible to affect their efficiency, toxicity, and disposition. Advantages of the use of optically pure drugs are well documented and considerable effort is being made to prepare them using all available synthetic techniques. At present, an ideal synthetic technique does not exist, all methods have their advantages and their limitations and the decision regarding the most suitable approach is considered on a case-by-case basis. The area of organic synthesis with enzymes is still in its infancy and among more than 2000 known enzymes, very few (mostly hydrolases) have been used routinely in organic synthesis. Modern methods of protein engineering and industrial microbiology will bring to the market more inexpensive, stable enzymes with broad substrate specificity and high stereoselectivity. In the future, enzymatic and other biocatalytic methods, which now include the promising abzyme technology (Lerner *et al.*, 1991), will successfully compete with chemical techniques, in many cases becoming the method of choice in the preparation of chiral drugs.

A variety of enzymes, mostly lyases and aldolases, have so far been used to synthesise complex sugars, sugar analogues and other biologically important natural compounds (Toone *et al.*, 1989; Wong & Whitesides, 1994). In particular, thiamine-dependent enzymes have been employed for some time as catalysts in chemoenzymatic syntheses (Faber, 1992; Krampitz, 1969). To utilise enzymes for non-physiological reactions and

to yield novel products, a broad substrate spectrum is desirable. Thiamine-dependent enzymes vary in their substrate tolerance from rather strict substrate specificity (phosphoketolases, glyoxylate carboligase) to more permissive enzymes (transketolase [TK], dihydroxyacetone synthase, pyruvate decarboxylase) and therefore differ in their potential for application as biocatalysts.

Formation of carbon-carbon bonds is of central importance to synthetic organic chemistry and may be catalysed by lyases (Van der Werf *et al.*, 1994), including aldolases (Kragl *et al.*, 1992), and transferases such as TK. In classical organic chemistry many types of C-C bond-forming reactions have been developed, for example the Aldol Condensation (Figure 1.2) in which the α carbon atom of one aldehyde or ketone molecule adds to the carbonyl carbon of another (Nielsen & Houlihan, 1968). Also the Diels Alder reaction in which a double bond adds 1,4 to a conjugated diene in a 2+4 cycloaddition (Brieger & Bennett, 1980) so the product is always a six-membered ring, a reaction which is easy and rapid and of very broad scope. However, few of these conversions are selective towards their reagents and have the ability to dictate the stereochemistry of the products. Since enzymes may perform either or both of these tasks they provide an important alternative to such conventional methods.

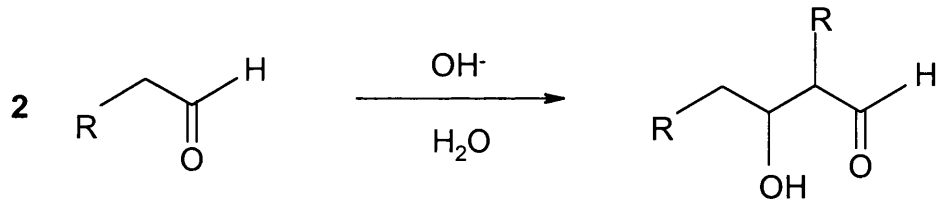



**Figure 1.2** *Aldol Condensation*

One drawback however, has been the availability of commercial enzymes at a good quality, reasonable prices, and a reliable supply from biological sources. With the advent of molecular biology and the use of highly productive recombinant microorganisms, these issues are of lesser significance. Possible improvements to the biotransformations can be achieved by modifying the substrate specificity or other

properties of the enzyme and also by addition of poly-histidine (Schmitt, *et al.*, 1993) or poly-lysine tails (French & Ward, 1994) to improve enzyme isolation and immobilisation. Although for industrial processes it is not always necessary to extract the enzyme in pure form since crude preparations, whole cells, or even cell fragments may be sufficiently catalytically active. Enzymes whose function *in vivo* is the degradation of carbohydrates may prove to be valuable biocatalysts for stereoselective formation of carbon-carbon bonds, TK is one such enzyme and is discussed in greater detail in the following section.

## 1.3 Transketolase

Existence of TK was first published in the early 1950s (de la Haba & Racker, 1952; Horecker & Smrriniotes, 1952) as catalyst for the cleavage of ribulose-5-P, in the presence of ribose-5-P, to glyceraldehyde-3-P (a deviation of its *in vivo* reaction). The enzyme was purified from yeast and named by Racker in 1953 when it was also reported to require thiamine pyrophosphate (TPP) and $Mg^{2+}$ ions as cofactors (Racker *et al.*, 1953; Horecker & Symriniotes, 1953).

Since its discovery much work has been documented on the isolation and purification of TK from a variety of sources, most commonly yeast or spinach. Study of the kinetics of TK reactions and the function of the cofactors and the active site has been essentially confined to the yeast enzyme, for which the pH optimum is widely reported as pH7.6 and Michaelis constants for the native phosphorylated carbohydrate substrates are in the region Km = $10^{-4}$ to $10^{-3}$ M (Datta & Racker, 1961).

The enzyme is composed of two subunits of equal molecular mass (Cavalieri *et al.*, 1975). Individual subunits consist of three domains similar in mass, each corresponding to one structural unit observed by electron microscopy. TK is active as a homodimer, and it is known that the two TPP-binding sites exhibit negative cooperativity (Egan & Sable, 1981). Most of the invariant TK residues are located in the cleft between the two subunits where the cofactors and substrates are bound, suggesting a functional importance.

### 1.3.1 Structure

*E. coli* transketolase (E.C.2.2.1.1) is a homodimer comprising subunits of molecular mass 72kDa, the two subunits combine to form a molecule with a $C_2$ axis of symmetry which can be seen in Figure 1.3 (the individual subunits are coloured red and orange respectively). The resultant enzyme contains two identical active sites symmetrically located at its interface, each subunit contributes residues to both active sites and therefore an individual subunit cannot be catalytically active. The crystal structure of this enzyme has been elucidated and reveals the identity of amino acid residues at the dimer interface and hence those located around the active site (Littlechild *et al.*, 1995). This information enables predictions to be made as to the nature of interactions between TK and its cofactors, TPP and $Mg^{2+}$ ions (Lindqvist *et al.*, 1992; Meshalkina *et al.*, 1997), and of those with its substrates. Cofactors bind in a deep cleft within the active site funnel with residues from both subunits contributing to their binding, the functional unit of TK thus being the assembled dimer. In order to obtain a fully active catalytic site, $Mg^{2+}$ may have to act as a ligand between the enzyme and the TPP molecule. Water molecules also perform a structural role in the binding of TPP (Nikkola *et al.*, 1994), and one of the ligands of the active site magnesium ion is a water molecule. Several inaccessible water molecules are found between the two TPP molecules, where they participate in a hydrogen bond network joining the two cofactors. The two-fold related active sites are located between different domains of the two subunits and TPP prevents the dissociation of apo-transketolase into monomers at low protein concentrations.

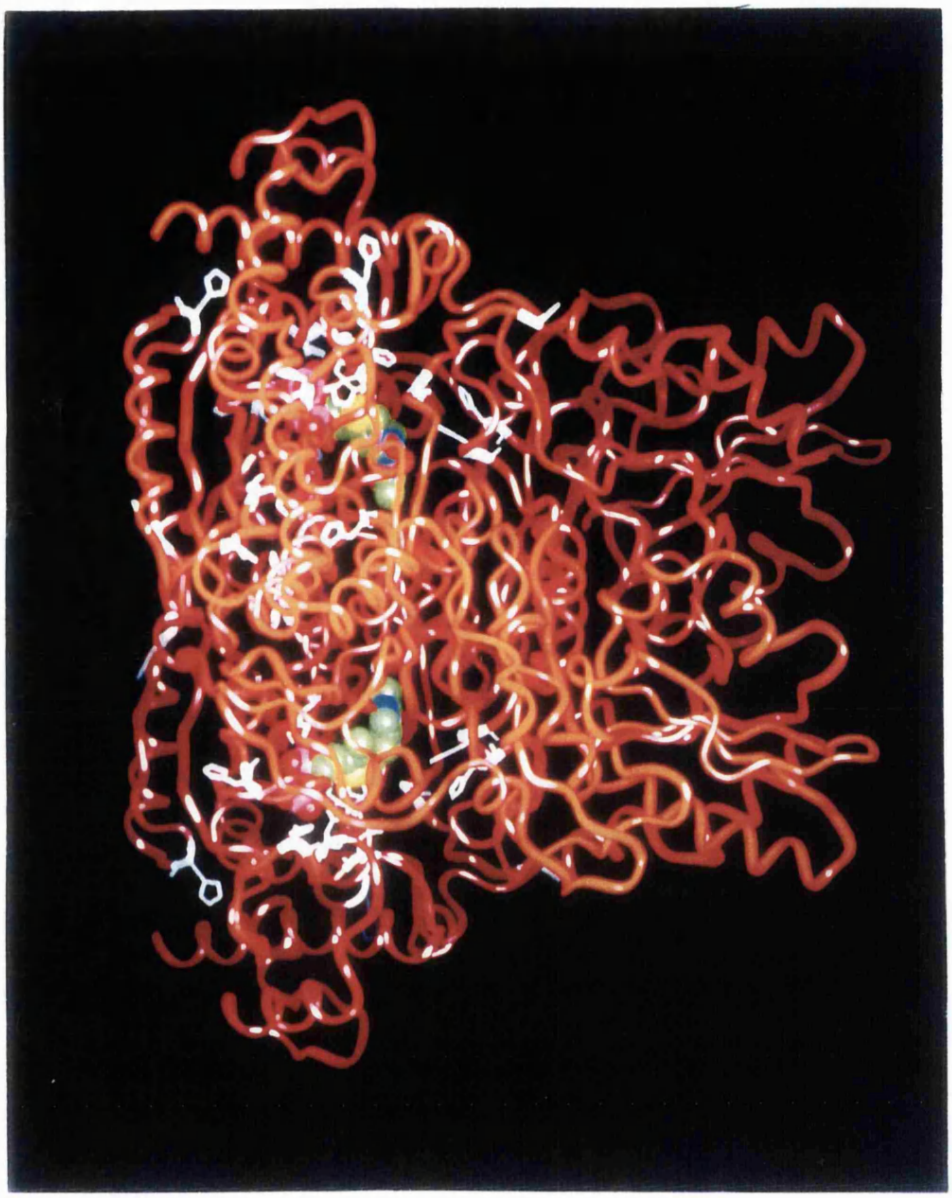


**Figure 1.3** *E. coli transketolase*

The cofactor, vitamin $B_1$ derived thiamine pyrophosphate (see Figure 1.4), is bound at the interface between the two subunits, it serves as coenzyme in many biochemical reactions involving cleavage of the C-C bond adjacent to a keto-group. Whereas non-enzymatic catalysis by thiamine has been rather well studied, enzyme mediated thiamine catalysis is poorly understood (Kluger, 1987), mainly due to the lack of structural information concerning enzyme-cofactor and enzyme-cofactor-substrate interactions.

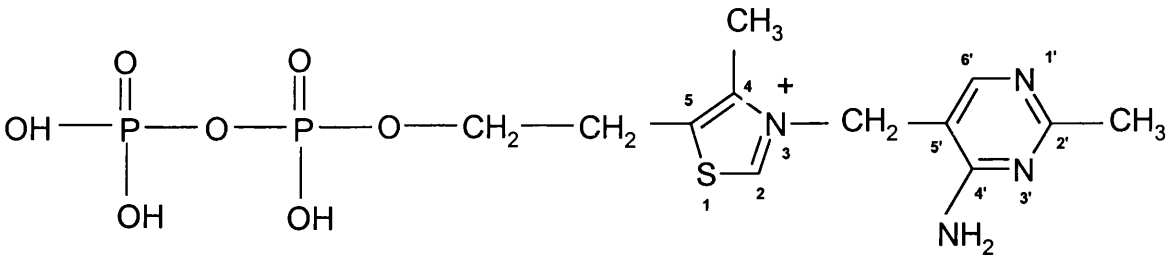



**Figure 1.4** *Thiamine pyrophosphate*

The enzyme subunit is built up of three domains of the α/β type (Lindqvist *et al.*, 1992) with the overall shape of the subunit being a rounded V. The diphosphate moiety of TPP is bound to the enzyme at the carboxyl end of the parallel β-sheet of the N-terminal domain of TK, it interacts with the protein directly through hydrogen bonding and indirectly via the divalent metal ion acting as an additional anchor. The presence of a phosphate binding site for TPP interaction may explain the observation that both sulphate and phosphate inhibit TK reaction at > 5mM concentration (Datta & Racker, 1961; Kotchetov, 1982). The thiazolium ring interacts with residues from both subunits through mainly hydrophobic interactions, whereas the pyrimidine ring is buried in a hydrophobic pocket of the enzyme formed by loops at the carboxyl end of the β-sheet in the middle domain in the second subunit. Structural analysis of TK identifies amino acids critical for cofactor binding and provides mechanistic insights into thiamine catalysis.

As with other TPP-dependent enzymes, TK requires $Mg^{2+}$ for catalytic activity (Kochetov, 1982), however other metal ions such as $Ca^{2+}$, $Mn^{2+}$ and $Co^{2+}$ can replace

the $Mg^{2+}$ ion with varying levels of activity (Heinrich *et al*, 1972). The amino acid sequences of TK from various sources are rather similar to each other, for instance the amino acid sequences of TK from *E. coli* and humans show about 50% sequence identities to the yeast enzyme (Sundstrom *et al.*, 1993; Sprenger *et al.*, 1995). An alignment of 17 available amino acid sequences for TKs from various organisms reveals a total of 54 invariant residues, with many of these amino acids clustered at the active site of the enzyme (Singleton *et al.*, 1996). Crystal structures of the apo- and holo-enzyme (Sundstrom *et al.*, 1992; Lindqvist *et al.*, 1992; Nikkola, 1994), various complexes of TK with TPP analogues (Nilsson *et al.*, 1993; Konig *et al.*, 1994) and a complex of holo-TK with the acceptor substrate erythrose-4-P (Nilsson *et al.*, 1997) are available. In addition, the crystal structures of a number of site-directed mutants have been determined providing further insights into the detail of interactions between TK and its cofactors and substrates (Wikner *et al.*, 1997).

Leading to the bound cofactor is a channel, formed by loop regions of the proteins, which binds and orientates the substrates (Schneider & Lindqvist, 1993). The overall tertiary and quaternary structure of apo-TK is identical to that of the holo-enzyme. No large conformational changes, such as domain-domain rotation or a difference in the packing of the two subunits, are associated with cofactor binding (Sundstrom *et al.*, 1992), these only cause local perturbations with minor rearrangement of some of the side chains. The loops at the active site are flexible in the apo-enzyme thus enabling TPP to reach its binding site. Binding of the cofactor induces defined conformations for these two loops at the active site. One of the loops is directly involved in binding of the cofactors $Mg^{2+}$ and TPP and acts like a flap, closing off the binding site of diphosphate. After binding of the cofactor, residues of this loop form interactions to residues of the loop from the second subunit, these interactions fix the conformation of the two loops from a flexible to a closed conformation. In the apo- structure, these interactions are lost and the loop has a high mobility.

The dimer interface is not particularly smooth and contains a number of inaccessible cavities with one of these hollows being a large, solvent-filled channel spanning the region between the two TPP molecules. This channel contains several conserved glutamic acid residues and in fact, a hydrogen bonding network connects the N1' atom of the pyrimidine ring of one TPP molecule with the corresponding nitrogen atom of the TPP molecule in the second subunit. The purpose of this conserved network could

be to shuffle protons between the two active sites, thus one TPP molecule could sense the state of the other.

TPP binds in a deep cleft lined with conserved residues located on the loop regions which build up the active site and is, except for the C2 atom of the thiazolium ring, totally inaccessible from the outer solution (only the part of the thiazolium ring which is involved in catalysis is exposed to solvent and accessible to substrate molecules). During turnover the C2 atom forms a covalent adduct with the substrate, a mechanism common to all TPP catalysed reactions. The channel is narrow and it is not possible that donor and acceptor substrates can bind simultaneously, consistent with the kinetic Bi Bi Ping Pong mechanism for TK (Kremer *et al.*, 1980).

### 1.3.2 Function

Together with transaldolase, TK creates a reversible link between the two main metabolic pathways, the pentose phosphate pathway (see Figure 1.5) and glycolysis (Boiteux & Hess, 1981). In *E. coli* two branches of the pentose phosphate pathway occur: 1) an oxidative branch in which pentose phosphate is formed from glucose-6-P and 2) a non-oxidative branch in which fructose-6-P and glyceraldehyde-3-P are formed. TK's particular role is involvement in the non-oxidative part of the pentose phosphate pathway, a route representing a method for the production of NADPH (Wood, 1985). Specifically, it initiates the reversible transfer of a two-carbon ketol unit to an aldehyde (Datta & Racker, 1961), forming a new C-C bond and creating an asymmetric centre, the general reaction scheme for which is given in Figure 1.6. *In vivo* TK catalyses the reversible transfer of a two-carbon ketol group from D-ribose-5-P to D-xylulose-5-P thereby generating D-sedoheptulose-7-P (Racker, 1961), since the products contain both a nucleophile and an electrophile the reaction is reversible and equilibrium is established.

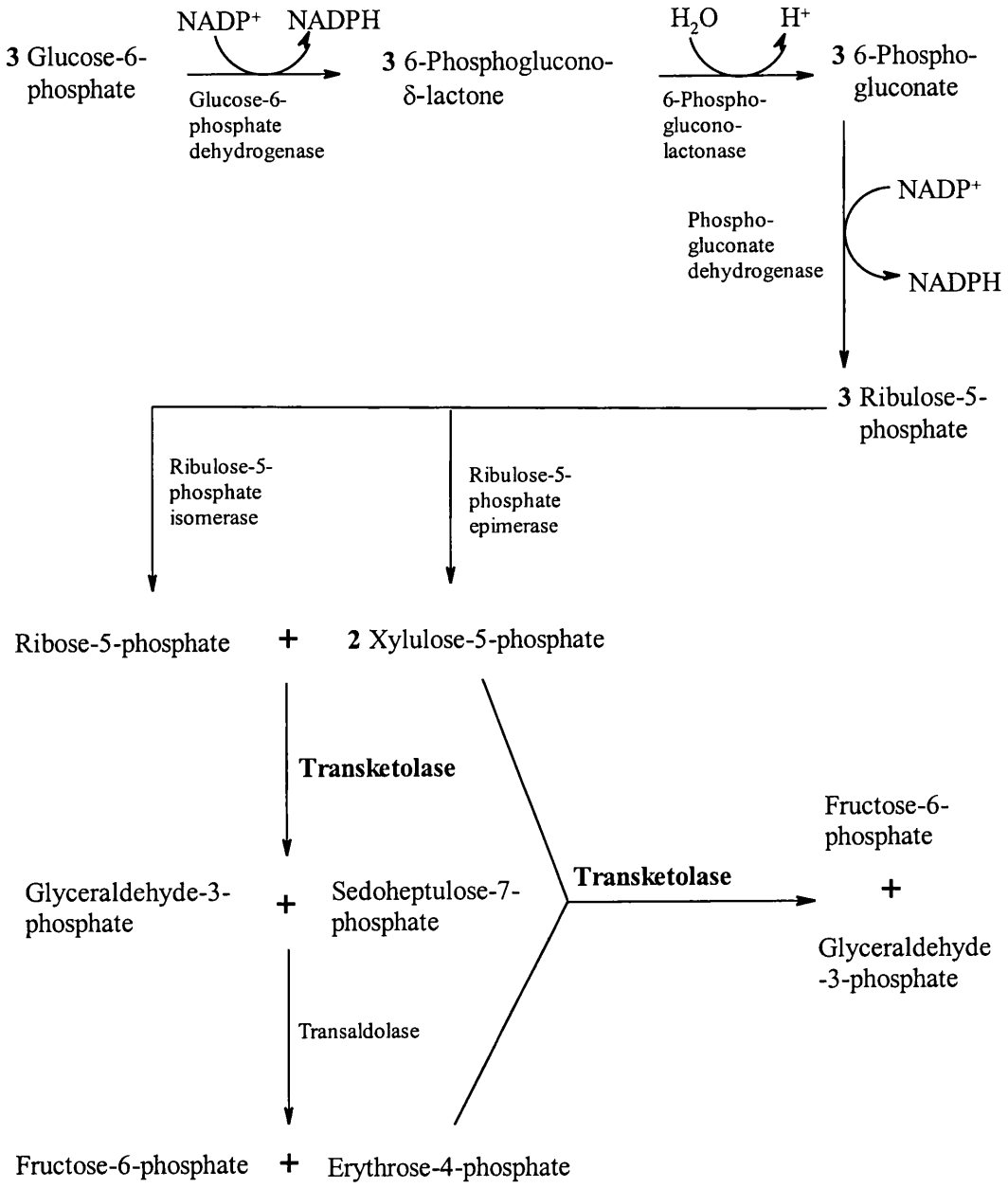



**Figure 1.5** *The Pentose Phosphate Pathway*

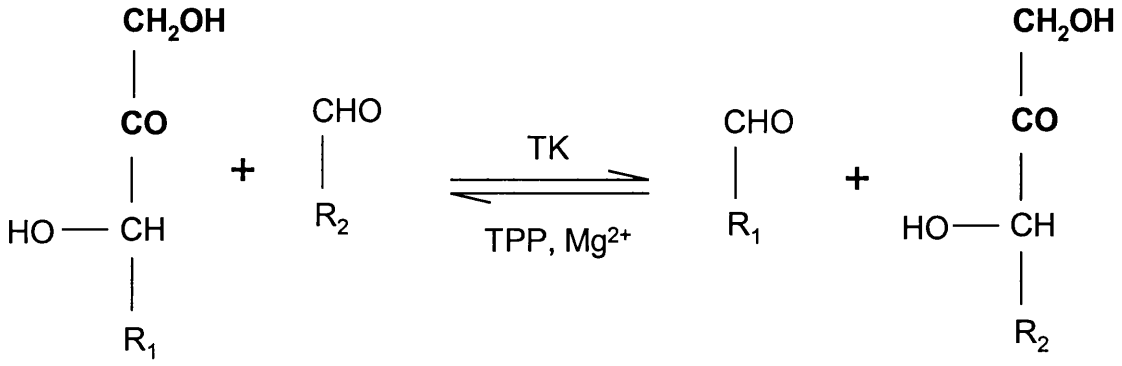



**Figure 1.6** *General transketolase reaction*

The enzyme is involved twice in the pentose phosphate pathway and the essence of these reactions is that excess ribose-5-P, formed by the pentose phosphate pathway, can be completely converted into glycolytic intermediates (*ie.* TK enables the cell to shuttle ribose-5-P and glycolytic intermediates between the two pathways). In bacteria, the enzyme is recruited for the catabolism of pentose sugars (*eg.* D-xylose, D-ribose, L-arabinose), the formation of ribose-5-P (for nucleotides, tryptophan, and histidine biosynthesis), and the provision of erythrose-4-P, a precursor of the shikimic acid and pyridoxine pathways that lead eventually to the formation of the aromatic amino acids, aromatic vitamins, and vitamin $B_6$ (pyridoxine) (Fraenkel, 1987; Zhao & Winkler, 1994). Mutants that lack TK are shikimic-acid auxotrophs or are auxotrophic for a combination of aromatic compounds plus pyridoxine and are unable to grow on pentoses (Iida *et al.*, 1993). Moreover, the pentose phosphate pathway is the only pathway that allows *E. coli* to utilise sugars such as D-xylose, D-ribose, or L-arabinose, which cannot be catabolised by other routes (Fraenkel, 1987; Lin, 1987). It was found that cells disrupted for TKL1 were unable to grow in the absence of aromatic amino acids (Sundstrom *et al.*, 1993) and this is most likely due to the fact that TK is required for the synthesis of erythrose-4-P (Braus, 1991), a precursor of the aromatic amino acids.

In photosynthetic organisms, the enzyme is involved in a second metabolic process. An essential part of the photosynthetic dark reactions is the Calvin cycle, where ribulose-1,5-bisphosphate is regenerated from phosphoglycerate, formed by the action of ribulose bisphosphate carboxylase. One of the enzymes of the Calvin cycle is TK,

catalysing the reaction of fructose-6-P with glyceraldehyde-3-P, yielding xylulose-5-P and erythrose-4-P and the conversion of sedoheptulose-7-P and glyceraldehyde-3-P to ribose-5-P and xylulose-5-P.

By using β-hydroxpyruvic acid (β-HPA) as the ketol donor the consequent loss of $CO_2$ ensures that the TK reaction is irreversible and hence goes to completion, leading to a trihydroxylated compound containing the D-*threo*-configuration (Srere *et al.*, 1958) as can be seen in Figure 1.7. This irreversible reaction feature is useful when conducting biotransformations for detailed analysis since it enables accurate monitoring of reagent concentrations, allowing a direct correlation to be made between the rate of disappearance of β-HPA and the extent of enzymatic conversion.

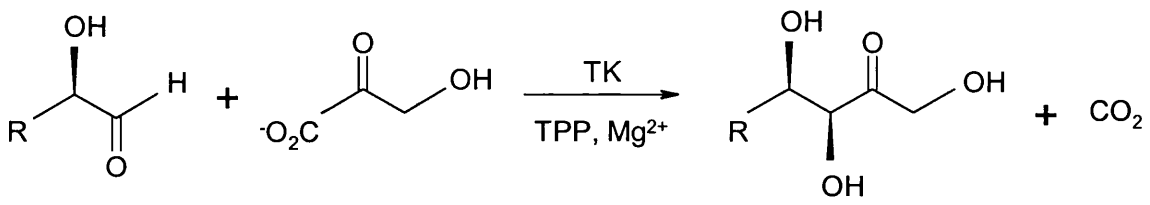



**Figure 1.7** *TK reaction with β-HPA as ketol donor*

TK contains vitamin $B_1$-derived TPP as prosthetic group and requires $Mg^{2+}$ ions for catalytic activity and the structural details of TPP binding and the resulting implications for the enzyme mechanism are a major consideration for work with TK. A second interesting feature is associated with the broad substrate specificity of the enzyme. TK is rather promiscuous with respect to the chain length of its substrates and will transform compounds ranging from a 3-C to a 7-C skeleton. This diverse range of possible substrates is achieved through the architecture of the substrate channel, and in particular the design of the phosphate-binding site. The reacting end of the substrate has to approach the C2 carbon atom of the thiazolium ring, or the α-carbon atom of the dihydroxy-ethyl TPP intermediate. Due to differences in the length of the carbon chain of the linear phospho-sugar substrates, this will result in a variation of the position of their phosphate group in the substrate channel. The location of the phosphate-binding site at the entrance of the channel makes it possible for substrates to reach the active centre irrespective of the chain length, whilst the interactions of the enzyme with the

substrate phosphate group are maintained by the use of two arginine residues, which through their flexible side chains can adjust to these variations in the position of the phosphate group (Lindqvist *et al.*, 1992; Schneider & Lindqvist, 1998).

TPP is the cofactor in a number of enzymatic reactions where a C-C bond adjacent to a carbonyl group is cleaved. Irrespective of the particular reaction catalysed, enzymatic thiamine catalysis proceeds through two common intermediates. One of these intermediates is the ylide form of TPP, resulting from the deprotonation of the C2 atom of the thiazolium ring in the initial step of the catalytic cycle. The second common intermediate is an α-carbanion which is formed after the nucleophilic attack of the ylide on the substrate (Kluger, 1987). The transfer is facilitated by TPP via the mechanism shown in Figure 1.8. In the first step, deprotonated TPP in the enzyme active site couples with the ketol donor giving a Zwitterion covalent intermediate. The Zwitterion then decarboxylates to the enol (second step) which will react with the substrate aldehyde. Once in this form, the bond to the thiazole ring will spontaneously cleave (third step), forming the reaction product and regenerating TPP.

The enol could theoretically react with either the Re-face or the Si-face of the aldehyde, this would result in a mixture of diastereomers at the generated stereo centre (*). However, steric interactions between the enzyme and the reacting substrates ensure that the enol will only react with one of the aldehyde faces (Re-), thus giving the shown stereochemistry at the new stereo centre (*). In this way TK determines product stereochemistry through the nature of its catalytic mechanism.

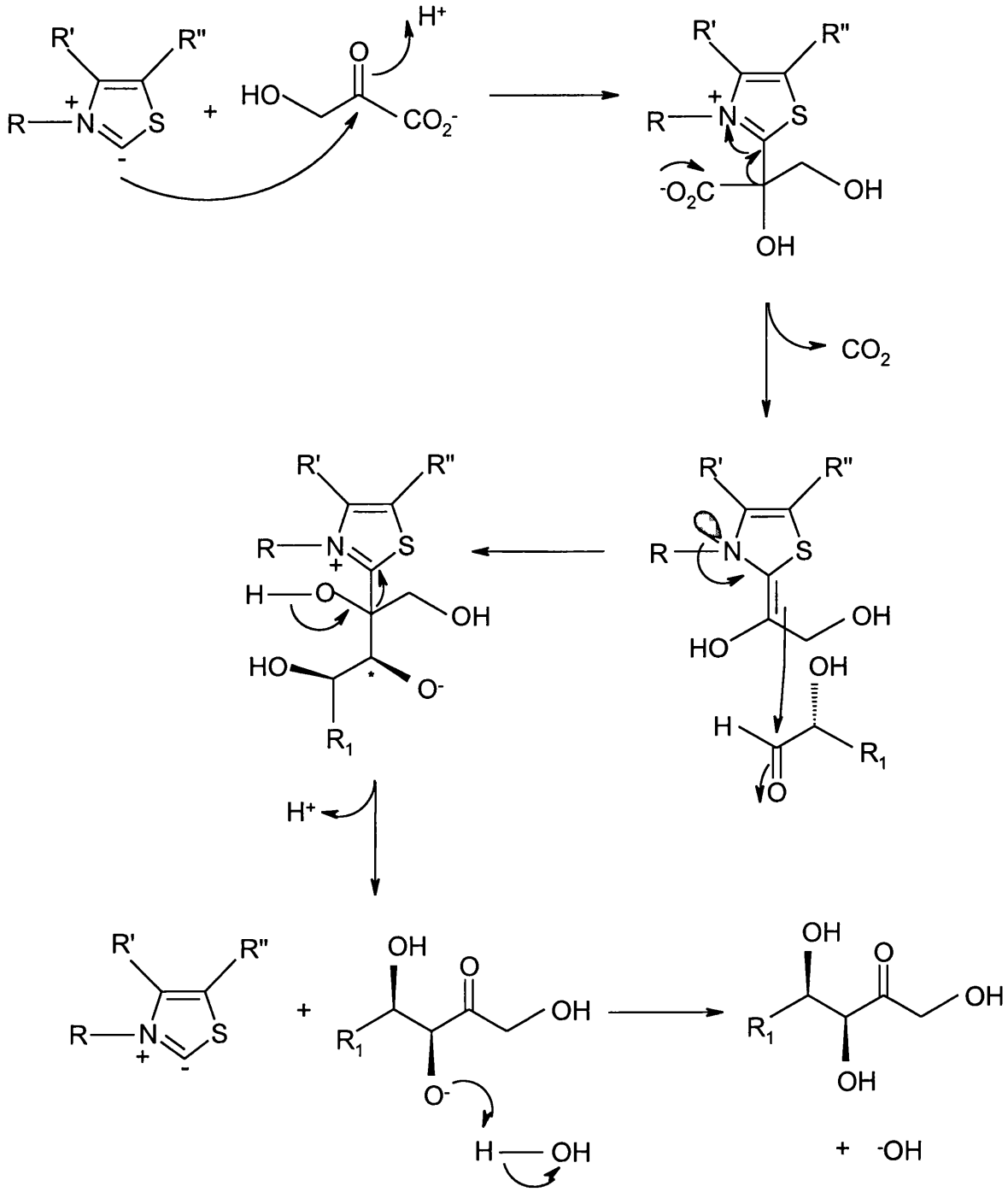



**Figure 1.8** *Mechanism of action of TPP in TK-catalysed reaction* (Kluger, 1990)

The first half of the TK reaction consists of the cleavage of the donor substrate and formation of the first product, an aldose and a covalent intermediate, the α-carbanion of α,β-dihydroxyethyl TPP (Krampitz, 1969). In the second half of the reaction, a nucleophilic attack of the α-carbanion on the acceptor substrate occurs and the second product, a ketose with its carbon skeleton extended by two carbon atoms is formed (in principle, the reverse of the first half of the reaction).

The crystallographic analyses of TK in complex with the coenzyme TPP, or the coenzyme analogue thiamine thiazolon diphosphate, have provided detailed information regarding the mode of coenzyme binding and the interactions of TPP with the enzyme (Lindqvist *et al.*, 1992; Nilsson *et al.*, 1993; Nikkola *et al.*, 1994). In the crystal structure of TK there is no enzymatic base in a position suitable for proton abstraction at the C2 atom of the thiazolium ring of TPP. Instead, a molecular mechanism for proton abstraction has been proposed (see Figure 1.9) that suggests a cofactor-assisted deprotonation (Schneider & Lindqvist, 1993). Essential to this model is the observed hydrogen bond between the side chain of the invariant residue Glu418* and the N1' nitrogen atom of TPP. This interaction has also been found in other TPP-dependent enzymes (Muller & Schulz, 1993; Dyda *et al.*, 1993) and it has been proposed that the mechanism for generating the C2 carbanion, the first step in enzymatic thiamine catalysis, is common to all thiamine-dependent enzymes (Muller *et al.*, 1993; Schneider & Lindqvist, 1993).

During the first half-cycle of catalysis, at least two proton transfer steps have to occur; the C2 carbon atom of the thiazolium ring of TPP has to be deprotonated, and upon cleavage of the donor-TPP product, the C3 hydroxyl group of the substrate will have to be deprotonated to yield the aldose product. In addition, the negative charge developing at the C2 carbonyl oxygen of the donor substrate requires electrostatic stabilisation. The first proton transfer is common to all TPP-dependent enzymes, and compelling evidence has been assembled that this step is catalysed by TPP itself. This step is facilitated by the enzyme in two ways: maintenance of the V-conformation, which positions the 4-$NH_2$ group in the vicinity of the C2 carbon of the thiazolium ring; and protonation of the N1' nitrogen atom of the pyrimidine ring, which generates the 4-imino group. This 4'-imino group can then act as a sufficiently basic group for proton abstraction from the C2 atom of the thiazolium ring. Once formed, the C2 carbanion of the cofactor is able to attack the carbonyl carbon of the donor substrate. For covalent

bond formation between these two groups to occur, stabilisation of the negative charge developing at the carbonyl oxygen is required and it has been speculated that this could be achieved by either the conserved residue His481* or charged 4'-imino group of TPP.

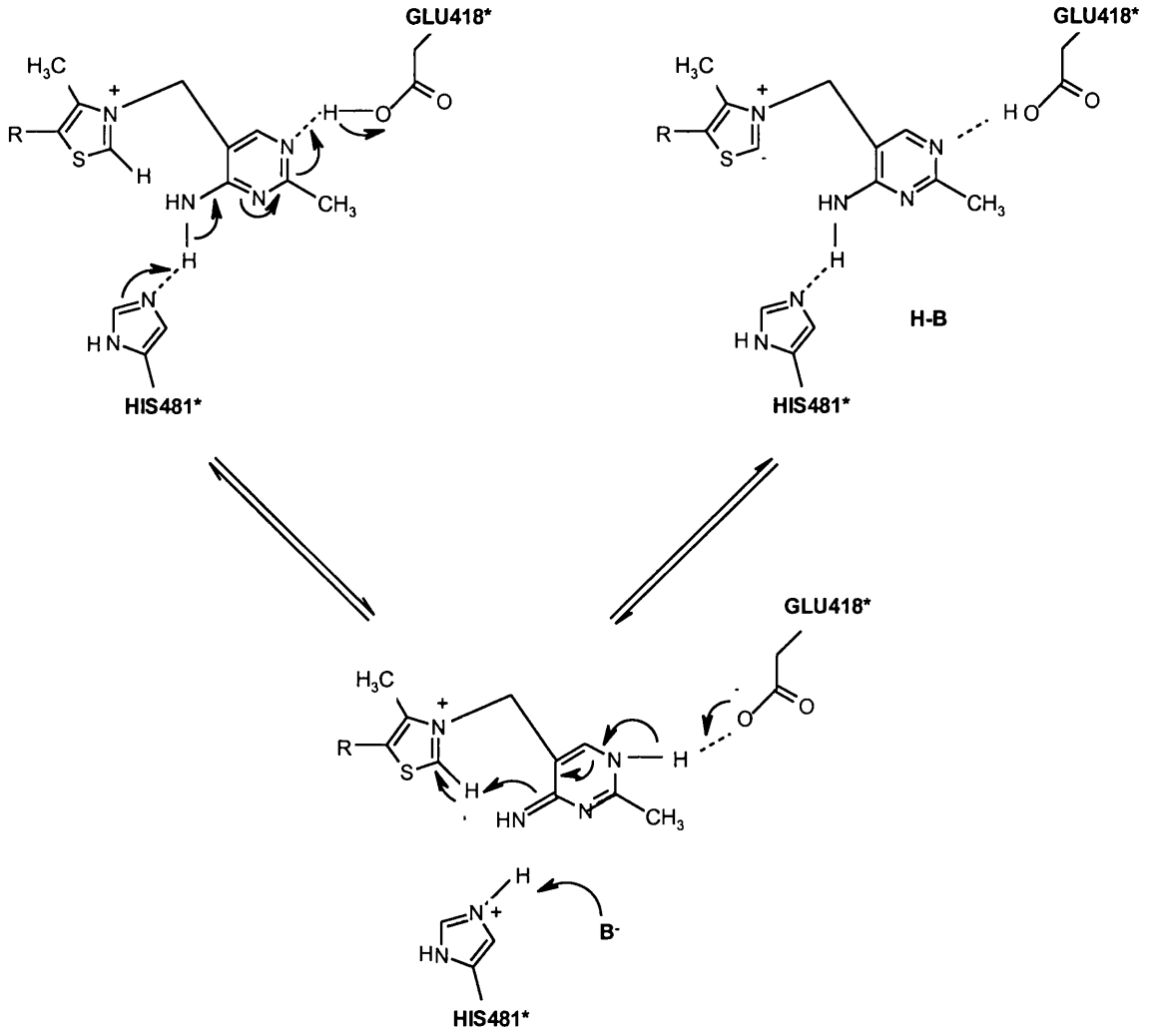



**Figure 1.9** *Suggested mechanism for deprotonation of TPP at C(2)* (Lindqvist *et al.*, 1992)

Replacement of the invariant TK residue Glu418* with glutamine and alanine has been performed (Wikner *et al.*, 1994), this residue forms a hydrogen bond to the N1' nitrogen atom of the pyrimidine ring of the cofactor TPP. Crystallographic analyses of the mutants show that these amino acid substitutions do not induce structural changes beyond the site of mutation and in both cases, the cofactor binds in a manner identical to the wild-type enzyme. However, the E418*Q mutant shows 2% and the E418*A mutant shows about 0.1% of the catalytic activity of wild-type enzyme whereas the affinities of the mutant enzymes for TPP are comparable to that of the native enzyme. The hydrogen bond between the coenzyme and the side chain of Glu418* is thus not required for coenzyme binding but essential for catalytic activity. These findings demonstrate the functional importance of this interaction and support the molecular model for cofactor deprotonation, the first step in enzymatic thiamine catalysis. The base abstracting the proton from the C2 carbon atom is the 4-imino group of the pyrimidine ring. The hydrogen bond between the N1' nitrogen atom and the side chain of Glu418* is essential for conversion of the 4-amino group to the 4-imino group during the catalytic cycle.

Chemical modification of TPP through replacement of the N1' nitrogen by a carbon atom has provided evidence that the N1' nitrogen atom of the pyrimidine ring is indeed required for catalytic activity (Schellenberger, 1967; Golbik *et al.*, 1991). Subsequently it was shown that the binding mode of this cofactor analogue to TK is identical to the binding of TPP (Konig *et al.*, 1994), structural differences in cofactor binding can therefore be ruled out as being responsible for the loss of catalytic function. Replacement of the N3' nitrogen atom of the pyrimidine ring results in a cofactor analogue that is catalytically active. Besides having metabolic importance, TK has important potential applications as a catalyst for industrial organic synthesis (Toone *et al.*, 1989; Kobori *et al.*, 1992) and reasons for this are highlighted below.

## 1.4 Use of transketolase for organic synthesis

Characteristically enzymes exhibit a high degree of regio- and stereoselectivity and specificity (Jones & Beck, 1976; Cannarsa, 1996) and in addition to transforming the native substrates, they often accept a whole range of compounds and convert these with the same precision. In most TPP-dependent reactions the holoenzyme catalyses a carbon-carbon bond cleavage of the substrate, this is followed by condensation of a portion of the cleaved substrate, associated with TPP, with an acceptor to form the

product. The nature of the acceptor determines whether the enzyme is specific in its reaction or more permissive, for synthetic purposes permissive enzymes such as TK are of great value.

In recent years, TK (E.C.2.2.1.1) has attracted much interest as catalyst for chemoenzymatic reactions and consequently has been purified from several sources including spinach, yeast and recombinant *E. coli* strains (Bolte *et al.*, 1987; de la Haba *et al.*, 1955; Draths *et al.*, 1992; Sprenger, 1993; Sprenger *et al.*, 1995). It has been shown that TK from *E. coli* can be easily supplied in amounts surpassing one million units (Lilly *et al.*, 1996; Hobbs *et al.*, 1996).

Initial syntheses involving TK exploited the enzyme in its natural role, in the synthesis of phosphorylated- (Srere *et al.*, 1958) and non-phosphorylated- (Demuynck *et al.*, 1990; Villafranca & Axelrod, 1971) carbohydrates. Lately however, TK has been shown to have a low specificity for the aldehyde component, accepting a wide range of α-hydroxyaldehydes (Effenberger *et al.*, 1992; Kobori *et al.*, 1992). Previously it was believed that only α-hydroxyaldehydes of the D-configuration were possible aldehyde acceptors for TK, but recent reports have depicted the use of simple α-unsubstituted aldehydes (Demuynck *et al.*, 1991; Hobbs *et al.*, 1993; Morris *et al.*, 1996) and nitrosoaromatics (Corbett & Chiplico, 1980; Corbett & Corbett, 1986), albeit with significantly lower rates than for the hydroxylated compounds. Relative rates of reaction for some of the compounds investigated thus far are shown in Figure 1.10 in which rates are given relative to glycolaldehyde, the fastest reacting substrate.

The low reaction rate for DL-glyceraldehyde highlights the enzymes' specificity for α-hydroxyaldehydes of the D-configuration. For longer chain sugars the rate will also be affected by the equilibrium between chain and ring forms, since only the former will have facile entry into the narrow active site funnel. However, relative rate of reaction does not have to be particularly high for useful applications in preparative organic synthesis, another feature contributing to the synthetic potential of TK. Although α-hydroxyaldehydes are the best substrates, simple aldehydes (*eg.* propionaldehyde and pyruvaldehyde) were also found to react, these compounds were converted into products on a preparative scale with subsequent product extraction and structural characterisation using $^{1}H$ NMR. In contrast to the TKs from spinach and yeast (Demuynck *et al.*, 1991) no conversion of aromatic aldehydes such as benzaldehyde or

hydroxybenzaldehyde has been detected with *E. coli* TK (Sprenger *et al.*, 1995; Hobbs *et al.*, 1994). Furthermore, no structures of the reaction products from conversions of aromatic and heterocyclic substrates, with TK from any source, have yet been shown (Schorken & Sprenger, 1998).

Values shown here demonstrate that although wild-type TK is active towards a range of aldehydes, reactivity is still very limited, particularly as the length of the carbon chain increases and as more bulky side chains are introduced. Even though the enzyme has reactivity with α-unsubstituted aldehydes, if an α-hydroxyl group is present then it must be in the D-configuration for effective transformation to occur.

TK can use a large variety of phosphorylated and non-phosphorylated monosaccharides acting as donor and acceptor substrates (Usmanov, 1983; Kobori, 1992), a property that makes the enzyme a useful tool for stereospecific organic synthesis of carbohydrates. Only sugar phosphates with D-*threo*-configuration at the C-3 and C-4 carbon atoms can act as efficient donors in the TK reaction [C-3, C-4 ketone correspond to C-1, C-2 aldehyde]. This high stereoselectivity of the enzyme for the donor substrate can be understood from the observed pattern of hydrogen bonds that the substrate forms at the active site (Nilsson, 1997).

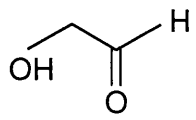
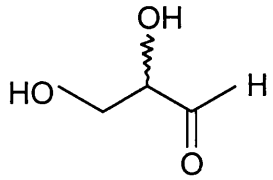
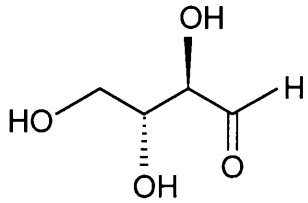
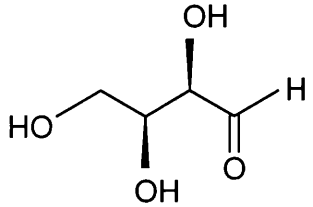
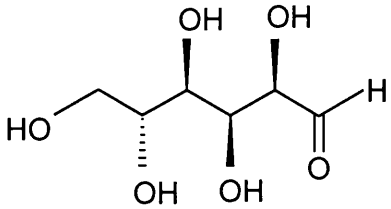
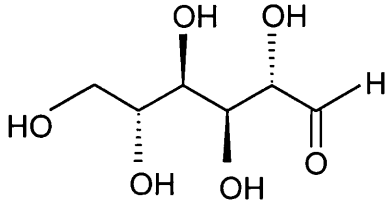
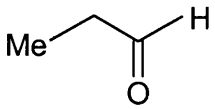
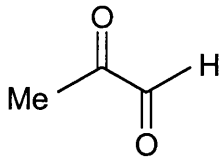
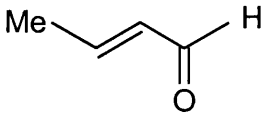
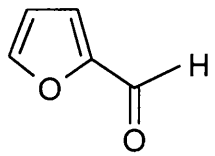
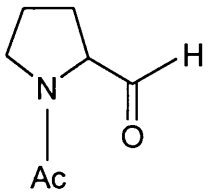
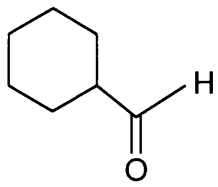
| Aldehyde | $V_{rel}$ | Aldehyde | $V_{rel}$ |
|---|---|---|---|
|  | 100 |  | 41 |
|  | 82 |  | 70 |
|  | 13 |  | 4 |
|  | 24 |  | 21 |
|  | 3 |  | ~0 |
|  | ~0 |  | ~0 |

**Figure 1.10** *Relative rates (Vrel) for TK reaction with a selection of aldehyde acceptors* (Hobbs *et al.*, 1994)

Promising applications of TK for commercial syntheses have been demonstrated with the reported production of the beetle pheromone (+)-*exo*-brevicumin (Myles *et al.*, 1991), the valuable aromatic furaneol (Hequet *et al.*, 1994), and the azo sugar 1,4-imino-D-arabinitol (Ziegler *et al.*, 1988). In contrast, the chemical synthesis of these chiral substances requires multi-step syntheses with complex techniques including the introduction and removal of protective groups. Chiral resolution of racemic aldehyde using TK has also been used to remove the undesired R-enantiomer of several racemic 2-hydroxyaldehydes (Effenberger *et al.*, 1992) as optically active 2-hydroxyaldehydes are reported to be useful for various chemical syntheses (Mulzer, 1991), so less expensive racemic mixtures can be used as starting materials for the synthesis of enantiomerically pure compounds. Furthermore, the unphosphorylated sugar compounds, desired in most cases, are formed directly with TK and the enzyme has also been utilised in the synthesis of certain deoxysugars (Hequet *et al.*, 1994).

That TK does have such relaxed specificity for the aldehyde component, combined with the high stereoselectivity for the C-C bond-forming step, makes this enzyme an extremely useful catalyst for asymmetric synthesis. However, in order for TK to have effective applications in organic synthesis both its substrate range and organic solvent tolerance need to be extended with molecular engineering, the former is the main theme of this project and further details are outlined below.

## 1.5 Production of constructs for overexpression of TK

For many years TK was exclusively obtained from two sources (1) *Saccharomyces cerevisiae* (Datta *et al.*, 1961) and (2) spinach (Villafranca & Axelrod, 1971). Although the synthetic potential of these sources was demonstrated on a small scale (Myles *et al.*, 1991) neither was able to yield large quantities of the enzyme. For this reason work was carried out to develop an organism that could be used to produce greater quantities of TK (Hobbs *et al.*, 1993).

Previously the gene for TK had been cloned into *E. coli* BJ502 (Draths & Frost, 1990) providing a useful starting point for over-expression of the protein. The fragment of DNA encoding the TK gene was excised from the plasmid of *E. coli* (BJ502/pKD112A) and re-introduced into a high copy plasmid vector pUC18 (Vieira & Messing, 1987) in strain JM107, generating two new transformants (JM107/pQR182 and JM107/pQR183), differing in the orientation of the gene with respect to the *lac*

promoter sited on pUC18. These two new transformants showed increased specific activity and both produced TK with approximately four fold higher specific activity than BJ502/pKD112A. Thus a new source of enzyme able to produce TK in large quantities suitable for preparative biotransformations was established.

The TK fragment was subsequently cloned from pQR182 into pBGS18- in both orientations with respect to the *lac* promoter (French & Ward, 1995), generating pQR700 and pQR701, with the TK gene under the control of its own promoter in both constructs. These pBGS based constructs produced similar levels of TK compared to the pUC constructs when grown O/N on complex media. Although the pBGS vector series (Spratt *et al.*, 1986) is similar in copy number to the pUC vector series and contains an identical *lac* promoter region, during fed-batch fermentations the pBGS constructs pQR700 and pQR701 were far more stable in culture than the pUC constructs. This was due to the pBGS vector being selectively maintained by the antibiotic kanamycin, more stable than ampicillin, and only inactivated within the host cell by phosphorylation.

Further to this, the size of the construct was reduced to improve efficiency of TK expression. The TK gene was amplified from pQR182 by PCR, utilising primers designed to remove extraneous fragments of DNA encoding the speB, ORF1 and $cmt^B$ genes and resulting in the generation of pQR706 and pQR711. These genes are still under the control of the TK promoter and use of inducers does not enhance the enzyme level. Both of these constructs produced higher levels of TK than the original fragment (see Table 1.1) and were used as the starting point for construction of all recombinant TKs employed for biotransformations described in this thesis.

| Construct | TK (U/ml) | Specific activity (U/mg) | Yeild (mg/L) |
|---|---|---|---|
| JM107/pQR183 | 1.38 | 3.94 | 53 |
| JM107/pQR182 | 2.07 | 4.22 | 77 |
| JM107/pQR700 | 1.51 | 3.87 | 55 |
| JM107/pQR701 | 2.16 | 4.49 | 82 |
| JM107/pQR706 | 4.50 | 6.43 | 156 |
| JM107/pQR711 | 4.25 | 6.25 | 168 |

**Table 1.1** *TK production in shake flask culture* (French & Ward, 1995)

## 1.6 Design of transketolases with broad industrial potential

The active site of TK appears as a funnel created by residues from both subunits of the enzyme dimer and is illustrated in Figure 1.11, this leads down to the TPP cofactor positioned at the end. Amino acid residues from the two subunits are coloured orange and grey respectively, and residues believed to have particularly important catalytic functions are highlighted in green. The active site channel is able to accept sugars in the open chain linear form and many other small aldehyde acceptors. However, many potentially useful substrates for the pharmaceutical industry contain bulky aromatic or heterocyclic groups. The work undertaken here addressed the question of how TK could be adapted to accommodate such aldehydes, along with other important considerations regarding its potential for organic synthesis such as enzyme purification and modification of its stereoselectivity. These topics constituted the three main areas of this project.

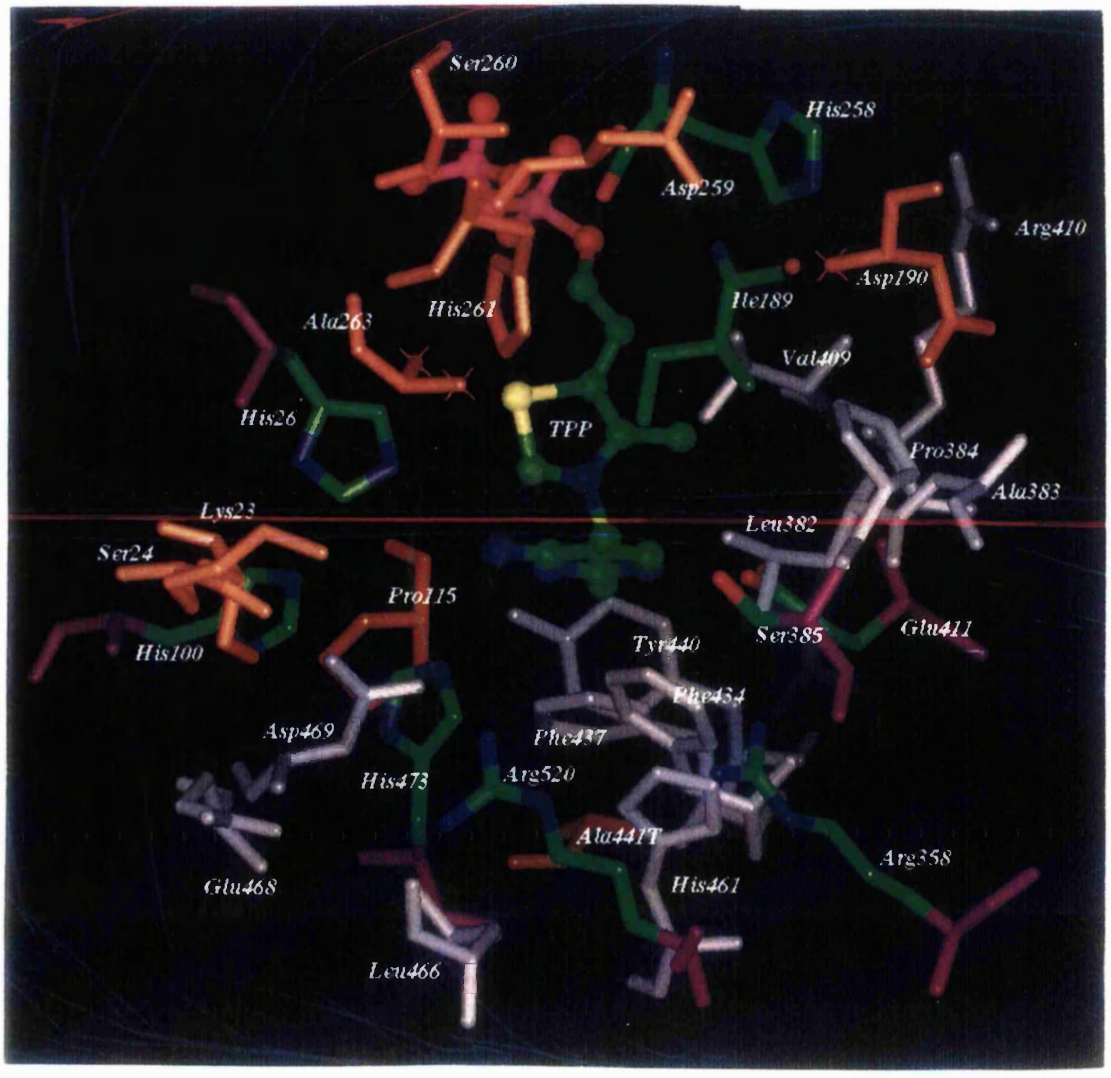



**Figure 1.11** *Active site of transketolase*

## 1.7 Project aims

The research for this project fell into three main areas:-

- Purification of transketolase;
- Expansion of transketolase active site;
- Modification of transketolase stereoselectivity.

Each of these is discussed in more detail below.

### 1.7.1 Purification of transketolase

In order for accurate rate data to be obtained from HPLC analysis, it is necessary for TK to be as pure as possible, thus reducing the possibility of side reactions causing the degradation of reaction substrates. For this reason a version of TK incorporating an N-terminal poly-lysine tail was generated. Work described in this thesis outlines the development of a two-step FPLC method of purification for the poly-lysine tagged enzyme using anion-exchange chromatography.

### 1.7.2 Expansion of transketolase active site

In addressing this question a variety of groups within the active site were altered. From the 3-D structure several amino acid side chains can be identified which could potentially sterically hinder the entry of certain substrates, for example those that were bulky or have substituents falling into this region when the aldehyde is bound. Mutants incorporating replacement of certain larger amino acid residues with alanines, are expected to have the ability to bind substrates containing bulky groups close to the aldehyde moiety. Assorted recombinant TK enzymes were generated and the work described here was that undertaken to characterise these constructs. An additional six mutants, where the active site entrance was opened up, were generated and partially characterised.

### 1.7.3 Modification of transketolase stereoselectivity

TK will add a two-carbon ketol unit from β-HPA to aldehyde acceptors, which unlike the native substrate have no hydroxyl group on the position α- to the aldehyde, a feature potentially extremely useful for biosynthetic applications. However, in the biocatalytic formation of novel sugars and other compounds with an α-hydroxyl group, TK has a strict requirement for the hydroxyl moiety to be in the D-configuration, it would therefore be useful to have a version of the enzyme tailored to add a two-carbon unit to the opposite enantiomer.

Molecular modelling experiments (not conducted as part of this project) highlight residues believed be acting to orient the substrate, allowing only the binding of molecules with an α-hydroxy group of the correct stereochemistry. To investigate these observations, and to provide an enzyme exhibiting the alternative stereospecificity, these key residues were altered using site-directed mutagenesis. It was expected that a mutant with reversal of such residues would exhibit the opposite specificity to that of the native enzyme. For this project, work was undertaken to generate a series of TK mutants designed to have relaxed/reversed stereoselectivity, characterisation using HPLC techniques afforded rate information for these constructs.

# Chapter Two: Equipment and General Techniques

# CHAPTER TWO: Equipment and General Techniques

## 2.1 Equipment

### 2.1.1 Chromatographic apparatus

Thin layer chromatography (TLC) was performed on pre-coated silica gel glass plates (5 x 10cm [Merck Silika gel 60 $F_{254}$]). The plates were visualised using either potassium permanganate solution (A1.14.7) or anisaldehyde solution (A1.14.8), depending on the compounds being analysed. Flash chromatography was performed on Silica gel (Merck 60 $F_{254}$).

Fast protein liquid chromatography (FPLC) was performed with a Pharmacia FPLC system on a MonoQ$^R$ HR5/5 anion-exchange column (1ml).

High performance liquid chromatography (HPLC) was performed on two Aminex$^R$ HPX-87H ion exclusion columns (each 300mm x 7.8mm) in series, maintained at 65°C. A Perkin Elmer ISS-100 HPLC system was used with a Schodex RI-71 refractive index detector. The running buffer was $H_2SO_4$ (6mM) at a flow rate of 0.45ml/min for a run time of 36 minutes.

### 2.1.2 Spectroscopic apparatus

Spectroscopic assays and optical density readings were recorded on a Beckman DU$^R$7500 spectrometer.

Nuclear magnetic resonance (NMR [$^1$H and $^{13}$C]) spectra were recorded on a Varian VXR400 (400MHz) spectrometer. All chemical shifts (δ) are reported in parts per million (ppm) and coupling constants (J) are quoted in Hertz. The results are given as:

$$\delta \text{ (number of atoms, multiplicity, J)}^{\text{assignment}}$$

where the multiplicity is denoted as:

s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet.

Mass analysis was conducted at the London School of Pharmacy on a V G 7070H instrument with a Finnigan Incos II data system. All values are reported as relative abundance (% [vertical]) plotted against m/z (mass of ion [horizontal]) where the most abundant ion is arbitrarily assigned a value of 100 per cent.

### 2.1.3 Additional apparatus

Analytical and preparative electrophoresis of plasmid DNA; DNA restriction fragments and single-stranded bacteriophage DNA was performed on 0.8-2% horizontal agarose gels in GibcoBRL Horizon$^R$58 or Horizon$^R$11.14 electrophoresis tanks. Gels were stained with ethidium bromide (0.5μg/ml). DNA was visualised at 312nm on a UVP, INC transilluminator.

Centrifugation of small samples was performed in a ALC$^R$ microCENTRIFUGE 4204, at a speed of 14,000rpm. Larger samples were centrifuged in either a Sorvall$^R$ RC2-B or a Sorvall$^R$ RC-5B centrifuge equipped with either a SS34 or a GSA rotor according to the volume to be processed, at speeds of up to 10,000rpm.

Polymerase chain reactions (PCR), the annealing of oligonucleotides to template DNA and the construction of gdDNA were conducted in a HYBAID OmniGene thermal cycler.

pH was measured using a Beckman Φ12 pH/ISE meter.

Purified water was obtained from an ELGA OPTION 4 water purifier.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of proteins was carried out in a CAMBRIDGE electrophoresis protein analysis chamber.

Sonication was performed using an MSE Soniprep 150 ultrasonic disintegrator.

## 2.2 General techniques

### 2.2.1 Linked-enzyme assay

The assay for determining TK activities was the linked-enzyme assay method of Heinnich *et al.*, 1972. A suitable dilution of sample was added to 1ml of the enzyme reaction mixture prior to addition of 10μl D-ribose-5-P (30mM). Ribose-5-P was the only substrate added which, in the presence of the linking enzymes and co-factors, cascades through the pathway depicted in Figure 2.1, eventually oxidising NADH in the final step. Reaction was followed by measurement of oxidation of NADH at 340nm. Provided the TK catalysed step was the rate-limiting step in this pathway, the rate of oxidation of NADH was directly proportional to the concentration of TK in the assay. The enzyme reaction mixture contained the following; TPP (0.25mM), $MgCl_2$ (9mM), NADH (disodium salt [0.154mM]), α-glycerophosphate dehydrogenase (0.2U/ml), triosephosphate isomerase (0.2U/ml), phosphoriboisomerase (0.2U/ml), D-ribulose-5-phosphate-3-epimerase (0.2U/ml), Gly-Gly (70mM, pH7.6 [A1.14.6]).

All assays were performed in triplicate and after an initial lag phase a linear decline in rate is observed, this value is used for determination of TK activity. The specific activity values quoted within this thesis are the average values obtained for each TK construct, the error ranges stated give the maximum deviations observed as a percentage of this mean value. As a control, prior to conducting these assays, background oxidation of NADH was measured in the absence of ribose-5-P and also in the absence of recombinant TK. For all TK constructs produced using *E. coli* strain JM107 the background oxidation of NADH under either of these conditions was found to be negligible. All assays were conducted at rt. In each case additional control assays were performed using the TK produced from pQR706/JM107 and that from JM107 cells (with no recombinant protein), these provided standards by which all new TK mutants could be characterised.

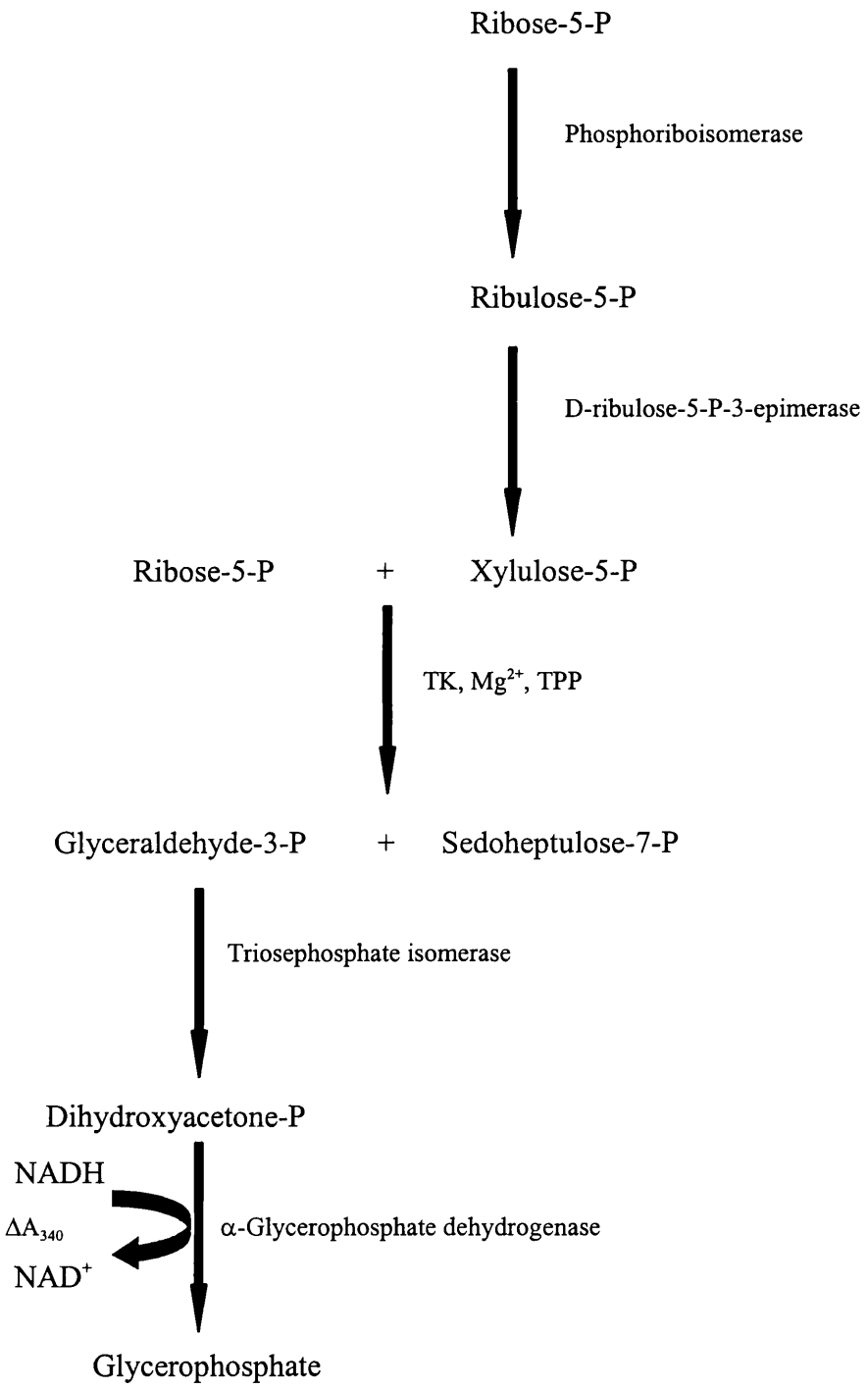



**Figure 2.1** *Linked-enzyme assay for TK*

The relationship between rate of oxidation of NADH ($\Delta A/\Delta t$) and TK concentration are as follows:

Using $c = A/\varepsilon l$ where: c = concentration

A = absorbance

$\varepsilon$ = extinction coefficient

l = light path length

$c/t = (\Delta A/\Delta t)/\varepsilon l$

$c/t = (\mu mol/ml)/min = U/ml = [TK]$ in assay

$(\Delta A/\Delta t) = /min$ = rate of oxidation of NADH

here for NADH: $\varepsilon = 6.3 ml/\mu mol$ and c = 1cm

therefore:

$$[TK] \text{ in assay} = \frac{\text{rate of oxidation of NADH } (/min)}{\varepsilon\ (ml/\mu mol)}$$

$$= (\mu mol/ml)/min$$

$$= U/ml$$

The result from this calculation can be related to the TK concentration in the test sample by taking into account the dilution in the assay:

$$\text{Activity of test sample} = \frac{\text{TK in assay (U)}}{\text{volume of test sample added (ml)}}$$

$$= U/ml$$

### 2.2.2 Protein quantitation

Assay reagent was prepared by adding BIO-RAD protein assay reagent with de-ionised water (1:4, reagent:water). The assay reagent (1ml) was added to the protein solution (5-25μg protein) in a spectrophotometric cuvette and left to stand at rt for at least 5 minutes. After this time the absorbance was measured at 595nm.

The procedure was standardized using BSA as a standard.

### 2.2.3 Alkaline-lysis plasmid preparation

For most purposes, small scale (~1.5ml) preparation of double-stranded DNA from bacterial cultures yielded sufficient material for subsequent manipulations. When more DNA was required, the technique was scaled up appropriately.

1) Bacterial cell culture (5ml) was grown O/N in LB media (A1.4.1). The culture (1.5ml) was harvested by centrifugation at 14,000rpm for 5 minutes, and the supernatant removed by decantation;
2) The cell pellet was resuspended in QPP1 (A1.5.1 [100μl]) by vortexing;
3) Freshly prepared QPP2 (A1.5.2 [200μl]) was added, the suspension mixed by inversion and placed on ice for 5 minutes;
4) Ice cold QPP3 (A1.5.3 [150μl]) was added, the suspension mixed by inversion and placed on ice for 10 minutes;
5) Particulate matter was pelleted by centrifugation at 14,000rpm for 10 minutes and the supernatant transferred to a clean tube;
6) The supernatant was extracted by vortexing twice with an equal volume of buffered phenol (A1.8.3), twice with phenol/chloroform (A1.8.4) and then once with chloroform. After each extraction, the aqueous (top) phase was separated from the organic phase by centrifugation at 14,000rpm for 5 minutes. If the aqueous phase was turbid, the extraction procedure was repeated until a clear solution was obtained;
7) 100% ethanol (3x volume) was added to the aqueous phase along with 3M NaOAc pH5.2 (1/10 volume) in order to precipitate the plasmid DNA. The solution was vortexed, placed on ice for 10 minuted and centrifuged at 14,000rpm for 10 minutes;
8) The supernatant was removed, the nucleic acid pellet washed with 70% (v/v) ethanol (1ml), and dried briefly;
9) The pellet was resuspended in TE (A1.8.1 [50μl]). DNA was stored at -20°C until needed.

### 2.2.4 DNA reaction work up

To remove enzymes and buffer components from DNA reaction mixtures the following procedure was used:

1) The solution was extracted twice with an equal volume of phenol (A1.8.3), twice with phenol/chloroform (A1.8.4) and once with chloroform. Each time vigorous mixing was followed by centrifugation to separate the phases. The aqueous phase was transferred to a fresh tube;
2) 3M NaOAc pH5.2 (1/10 volume) and 100% ethanol (3x volume) were added and the solution mixed;
3) The solution was placed on ice for 10 minutes, and precipitated DNA was pelleted by centrifugation at 14,000rpm for 10 minutes;
4) The pellet was washed with 70% (v/v) ethanol, dried briefly, and resuspended in TE (A1.8.1).

### 2.2.5 Glycerol-freeze stock preparation

A culture of the appropriate strain of *E. coli* was preserved by the addition of 40% glycerol to the culture. 40% glycerol (0.5ml) was added to the culture (0.5ml). Following gentle mixing, the culture was stored at -70°C.

### 2.2.6 Competent cell preparation

1) A single colony from an LB (A1.4.1) plate was used to inoculate sterile LB medium (5ml) and grown O/N at 37°C, 225rpm;
2) 0.2ml from this O/N culture was used to inoculate a flask containing LB medium (20ml). This culture was then grown at 37°C, 225rpm until mid log phase ($OD_{600}$=0.3);
3) Cells were transferred to sterile 1.5ml tubes, incubated on ice for 10 minutes and pelleted by centrifugation (4000rpm, 10 minutes, 4°C);
4) The supernatant was removed by decantation and the bacterial cell pellets resuspended in 1ml 0.1M $CaCl_2$ (A1.14.3);
5) Resuspended cells were incubated on ice for 20 minutes and harvested as before;
6) Cells were resuspended in 200µl 0.1M $CaCl_2$, incubated at 14°C O/N, and used fresh the next day.

### 2.2.7 Transformation of DNA

1) DNA mixture (2μl) was added to the appropriate *E. coli* competent cells (100μl), stirred to mix and incubated on ice for 45 minutes.
2) The mixture was heat shocked at 42°C for 30 seconds followed by incubation on ice for a further 2 minutes.
3) The mixture was then added to LB (5ml) and incubated at 37°C, 225rpm for 1h to recover the cells.
4) Transformation mixture (200μl) was plated onto LB/agar (supplemented with the appropriate antibiotic) and grown at 37°C O/N.

### 2.2.8 Quantification of oligonucleotides and plasmid DNA

Semi-quantitative assessment of the purity and concentration of DNA solutions was made on the basis of relative intensity of ethidium bromide staining on agarose gels (2.2.11). The limit of detection of this method is ~25ng.

Accurate quantification of DNA and oligonucleotides was accomplished by UV spectrophotometry. One OD unit at 260nm corresponds to a concentration of 50μg/ml for dsDNA, 40μg/ml for ssDNA and 20μg/ml for oligonucleotides. Pure DNA was assumed to exhibit an $OD_{260}/OD_{280} > 1.8$.

### 2.2.9 Restriction endonuclease digests

1) Restriction endonuclease digestions were performed using buffers as recommended by the manufacturer (A1.7.1-4) at the prescribed temperature (usually 37°C), in a water bath;
2) Typical digestions contained DNA at concentrations of 10-50μg/ml and 1-10 units of enzyme per μg of DNA;
3) Digestion was monitored by agarose gel electrophoresis (2.2.11) and allowed to proceed to completion. If necessary, incubation was continued and/or additional enzyme was added;
4) The resultant DNA restriction pattern was analysed by gel electrophoresis (2.2.11).

### 2.2.10 Purification of ssDNA

The isolation of single-stranded phagemid DNA was required for mutagenesis. Production of ssDNA by superinfection was accomplished by using M13 helper phage.

1) A single colony of *E. coli* CJ236 carrying the desired phagemid was used to inoculate LB medium (A1.4.1 [5ml]) containing the appropriate antibiotic. The culture was grown at 37°C, 225rpm to an $OD_{250}$=0.7 (~5h), corresponding to ~$5x10^7$-$2x10^8$ cells per ml;
2) An M13 helper phage stock (Norrander *et al.*, 1983) was diluted such that the addition of 5µl to the growing bacterial culture would result in a multiplicity of infection of ~20 (helper phage solution should be ~$2x10^{12}$pfu/ml);
3) Diluted phage stock (5µl) was added to the bacterial culture and incubated for an additional 1h at 37°C, 300rpm;
4) Superinfected culture (2ml) was diluted into LB (50ml) containing the antibiotic corresponding to the phagemid resistance marker plus kanamycin (50µg/ml). The culture was subsequently incubated O/N at 37°C, 300rpm;
5) Cells were removed by centrifugation at 10,000rpm, 4°C for 20 minutes, and the phage-containing supernatant was decanted;
6) 1/20 of the original culture volume of both 3M NaOAc pH7.0 (A1.6.2) and 40% PEG (A1.6.1) were added to the phage supernatant and the solution mixed thoroughly by vortexing;
7) The mixture was incubated on ice for 40 minutes to precipitate phage particles. Phage were harvested by centrifugation at 10,000rpm, 4°C, for 30 minutes;
8) The supernatant was discarded and the phage pellet was resuspended in TE (A1.8.1 [1ml]) and the mixture transferred to a 1.5ml microcentrifuge tube and recentrifuged to remove remaining cell debris;
9) Phage particles were reprecipitated on ice for 45 minutes with 1/20 volumes of 3M NaOAc pH7.0 and 40% PEG. Phage were pelleted as before, and resuspended in TE (0.5ml);
10) The solution was extracted twice with phenol (A1.8.3), once with phenol/chloroform (A1.8.4), and once with chloroform;
11) 3M NaOAc pH5.2 was mixed with the final aqueous preparation, and the ssDNA precipitated by the addition of 100% ethanol (3x volume), and leaving on ice for 30 minutes;

12) DNA was pelleted by centrifugation for 10 minutes and washed with 70% (v/v) ethanol. The pellet was resuspended in TE (50μl) then analysed by agarose gel electrophoresis (2.2.11). Phage DNA was stored at -20°C until required.

### 2.2.11 Horizontal agarose gel electrophoresis

Horizontal agarose gels were cast and run in TBE (A1.8.2) buffer. 0.8-2% agarose (w/v) gels were routinely used. Gels were prepared by dissolving electrophoresis grade agarose in TBE in a microwave oven. Molten agarose was thoroughly mixed, allowed to cool and poured into casting trays. Set gels were loaded into electrophoresis chambers containing sufficient running buffer to cover the gel. DNA samples were mixed with loading buffer (A1.8.5) and loaded into gel wells using a pipette. Electrophoresis was carried out at a constant field strength of 15-20V/cm. DNA was visualised by transillumination with UV light at 312nm.

Commercially supplied weight markers (A1.8.6) were used to size DNA fragments.

### 2.2.12 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to assess the protein composition and weight of samples. The separating gel (A1.9.8) was prepared by mixing all the reagents, with the TEMED and APS (A1.9.5) added last as they begin gel solidification. Once mixed it was poured immediately, a layer of water-saturated butanol was poured on top to give an even layer. This was allowed to set at rt for ~1h. The butanol was discarded and then the stacking gel (A1.9.7) was prepared. Again, the TEMED and APS were added last and this was poured immediately on top of the separating gel. The comb was inserted while the stacking gel was poured on. After setting, the comb was removed and the gel was placed in the tank, filled with electrode buffer (A1.9.2). Prior to loading, the protein samples were denatured by heating to 95°C for 5 minutes with prior addition of sample loading buffer (A1.9.6).

The gels were run at a constant voltage of 35mA until the bromophenol blue dye just ran off the bottom of the gel. The gel was then removed from the tank and stained in coomassie blue staining solution (A1.9.9) for 1h. The gels were destained once in fast

destain (A1.9.11) for 30 minutes, and twice in slow destain (A1.9.12) for 2h (or until the protein bands were sufficiently visible on a white light box).

### 2.2.13 DNA ligation reactions

Typical DNA ligations consisted of DNA restriction fragments (5-50ng/μl), T4 DNA ligase buffer (A1.7.6) and T4 DNA ligase. Ligations were carried out in a volume of 50μl and incubated O/N at 16°C.

Relative amounts of vector and insert varied according to the nature of the ligation being performed. The enzyme was inactivated by addition of 0.5M EDTA (A1.14.10 [1μl]) and heating to 65°C for 10 minutes.

### 2.2.14 Annealing of oligonucleotides to ssDNA and formation of gdDNA

The following was mixed together:

| | |
|---|---|
| T4 DNA polymerase buffer (A1.7.5) | 2.5μl |
| Kinased oligonucleotide | 0.5μl (~150ng) |
| Linear Bluescript | 3μl (~1μg) |
| Template ssDNA | 2μl (~1.5μg) |
| Milli-Q water | 17μl |

The mixture was heated instantly to 95°C, and left there for 5 minutes. It was then cooled to 37°C over 1h to allow the oligonucleotide and the linear Bluescript to find their complementary sequences. The resultant gdDNA was stored briefly on ice before use.

### 2.2.15 Synthesis and recircularisation

The following were mixed together without vortexing so as not to disrupt the oligonucleotide-gdDNA complex.

| | |
|---|---|
| Annealing mixture | 10μl |
| dNTPs (2.5mM) | 2μl |
| T4 DNA ligase buffer (A1.7.6) | 1μl |
| Milli-Q water | 5.5μl |

| | |
|---|---|
| T4 DNA ligase | 0.5μl |
| T7 DNA polymerase | 1μl |

The reaction was incubated at rt for 90 minutes.

### 2.2.16 Polymerase chain reaction (PCR)

The polymerase chain reaction was used for the amplification of DNA fragments, the reaction mixture contained the following:

| | |
|---|---|
| Template DNA | 50ng |
| Forward primer | 125ng |
| Reverse primer | 125ng |
| 10x reaction buffer (A1.7.7) | 5μl |
| dNTP mix | 1μl |
| *PfuTurbo* DNA polymerase* | 1μl |
| Milli-Q water | to 50μl |

* (Lundberg *et al.*, 1991)

Reactions were set up in 0.5ml microcentrifuge tubes. The mixtures were kept on ice while adding reaction components and the enzyme was added last. The PCR was carried out in a thermocycler as follows:

| Segment | Cycles | Temperature (°C) | Time |
|---|---|---|---|
| 1 | 1 | 95 | 30 seconds |
| 2 | 16 | 95<br>55<br>68 | 30 seconds<br>1 minute<br>12 minutes |

The reactions were analysed by gel electrophoresis (2.2.11). DNA was purified from successful reactions using QIAquick columns obtained from QIAGEN.

### 2.2.17 HPLC analysis

The isocratic HPLC system used for this work was that devised by Mitra and Woodley, 1996 and consisted of two Aminex$^{R}$87H columns (7.8 x 300mm from Biorad Laboratories Ltd., Hemel Hempstead, Hertfordshire, UK) in series with 6mM $H_2SO_4$ (A1.10.1) mobile phase running at 0.45ml/min. The columns were maintained at 65°C

and the peaks were detected by a change in refractive index. Samples were diluted in mobile phase to a final concentration of 20mM and filtered through 0.45μm filters prior to injection (10μl). All assays were performed in triplicate and in each case a sample from reaction with recombinant TK produced from pQR706 was run through the HPLC system in order to provide a standard by which the new TK mutants could be assessed. The specific activity values quoted within this thesis are the average values obtained for each TK construct, the error ranges stated give the maximum deviations observed as a percentage of this mean value. Substrate retention times under these conditions were found to be as follows:

| | |
|---|---|
| β-HPA | 21.32 minutes |
| Glycolaldehyde | 32.14 minutes |
| L-erythrulose | 30.85 minutes |
| DL-glyceraldehyde | 29.74 minutes |

The relationship between the rate of degradation of substrate ($\Delta A/\Delta t$) and TK activity were calculated using the following:

$$[TK] \text{ in assay} = (100\mu mol/ml)/min$$
$$= U/ml$$

*ie.* One unit of TK activity during these transformations was defined using the rate of degradation of 100μmol of reaction substrate.

### 2.2.18 Ion exchange chromatography

The MonoQ$^R$ HR5/5 column used for purification was obtained from Pharmacia. The column was equilibrated with 10ml Gly-Gly (150mM, pH7.0 [A1.11.2]). Protein (17.5mg) was applied and the column was washed with 10ml Gly-Gly (150mM, pH7.0). TK was eluted using a linear gradient of NaCl from 0 to 0.5M over 30ml [1st step]. The TK-containing fraction (1ml) was reapplied to the column (equilibrated with MES [100mM, pH5.0 {A1.11.3}]) and TK was eluted with a linear gradient of NaCl from 0 to 1M over 30ml [2nd step]. The chromatography was carried out at 22°C and a

flow rate of 0.5ml/min was used throughout. Column cleaning was performed using 2M NaCl (A1.12.1), 2M NaOH (A1.12.2) and 75% acetic acid (A1.12.3).

### 2.2.19 Biotransformations for HPLC analysis

All reaction substrates were obtained from Sigma Chemical Company, Poole, Dorset, UK. The reaction mixture contained the following: TPP (2.4mM), $MgCl_2$ (0.9mM), β-mercaptoethanol (10mM), aldehyde (100mM), β-HPA (100mM), Gly-Gly buffer (70mM, pH7.6 [A1.14.6]), and TK, in a final volume of 10ml. TK was added last to initiate reaction and the mixture was incubated at 37°C O/N (in the case of sugar reactions, aliquots were taken every minute for the first ten minutes). The reactions were monitored by TLC (ethyl acetate:methanol = 95:5 [0.1μl of reaction mixture was was spotted onto the TLC plate]) and the final reaction mixture analysed using HPLC. The retention factor ($R_f$) for each compound analysed was measured from the developed TLC plates. This figure is the relationship between the distance moved by the compound and the distance moved by the eluting solvent, and is calculated using:

$$R_f = \frac{\text{Distance moved by the product spot}}{\text{Distance moved by the solvent front}}$$

For reactions with aromatic aldehydes, reaction products were extracted using diethyl ether with subsequent structural characterisation conducted using NMR ($^1H$ and $^{13}C$) and mass spectrometry.

### 2.2.20 Proof of TK stereoselectivity

In order to confirm the stereoselectivity of certain TK mutants a reaction was carried out as in (2.2.19) using glycolaldehyde as the aldehyde acceptor. The product (L-erythrulose) was extracted twice using an equal volume of ethyl acetate (10ml). The organic solvent was removed using a rotary evaporator and the resultant mixture was then purified by flash chromatography using a 30cm Silica gel (Merck 60 $F_{254}$) column. The product was eluted using chlorofrom:methanol = 4:1, 0.5ml aliquots were collected and monitored using the TLC techniques described in 2.2.19. Optical rotation data was provided by Helen Hailes of the UCL Chemistry Department.

### 2.2.21 Bacterial strain and growth conditions

TK was prepared from the required *E. coli* strain. Bacteria were grown aerobically at 37°C, 225rpm in LB medium supplemented with 0.1% glucose and the appropriate antibiotic. For pQR730 (poly-lysine tagged TK), lacking the TK promoter, the medium was also supplemented with IPTG (A1.14.9) to a final concentration of 1mM. Cells were harvested by centrifugation for 5 minutes at 14,000rpm.

### 2.2.22 Bacterial extracts preparation

Crude extracts were prepared by disruption of cells by sonication. Cells from 3ml O/N cultures were resuspended in Gly-Gly (100mM, pH7.6 [A1.14.5]) and sonicated (6x 10 second bursts with 10 second intervals). Disrupted cells were centrifuged for 10 minutes at 14,000rpm to remove cell debris. The supernatant thus produced was stored on ice and used for subsequent reactions.

# Chapter Three: Expression, purification and characterisation of lysine-tagged *E. coli* TK by ion exchange chromatography

# CHAPTER THREE: Expression, purification and characterisation of lysine-tagged *E. coli* TK by ion exchange chromatography

## 3.1 Introduction

In order to accurately monitor the reactions of TK with a range of substrates, a relatively pure form of the enzyme is desirable. This reduces the possibility of side-reactions and thus allows more reliable rate data for these transformations to be obtained. This chapter outlines the design of a recombinant TK with an N-terminal affinity tag, and details the development of an ion exchange FPLC method designed for the rapid purification of this protein.

To facilitate the immobilisation of *E. coli* TK a recombinant protein, with an N-terminal affinity tag of ten lysine residues, was engineered by Carol French (French & Ward, 1994). Initial immobilisation experiments had been conducted by Carol French, using Eupergit C and CNBr activated Sepharose 4B, with promising results (French & Ward, 1994). However, prior to this project, no use had been made of the poly-lysine tail for work relating to the purification of TK. Since the poly-lysine tagged TK was already available from this previous work, it was decided to attempt to devise a purification method for this protein using affinity chromatography, exploiting the fact that this version of TK is able to strongly bind a MonoQ$^R$ HR5/5 column.

Bacterial cells (of which over-expressed TK constitutes ca. 10% of total cellular protein) were disrupted by sonication, and the supernatant thus produced was purified in two sequential steps. Firstly a salt gradient at pH7.0 was used, the resultant TK-containing fraction being re-loaded onto the column and subjected to a second salt gradient at pH5.0. Purified enzyme was analysed by SDS-PAGE (2.2.12) and characterised using a linked-enzyme assay (2.2.1). This method affords a 4.5-fold purification of TK for use in subsequent investigations into its specific activity with a range of novel substrates.

## 3.2 Construction of poly-lysine tagged TK

To assist the efficient purification of TK from *E. coli* using ion-exchange chromatography, and to aid enzyme immobilisation, a recombinant protein with an N-terminal affinity tag of ten lysine residues was engineered. Prior attempts by Carol French to generate a poly-histidine tagged version of the enzyme, to be purified using

$Ni^{2+}$ - NTA chromatography (Schmitt, *et al.*, 1993), were unsuccessful. A strategy for the production of this modified TK is given in Figure 3.1. As can be seen in this scheme, the final construct was generated in three sequential steps using a polymerase chain reaction (Saiki, *et al.*, 1988), at each stage the PCR product was cloned into pCRScript SK(+). Sequences of primers used for these reactions are given in Appendix Two (A2.3).

pQR706 (see section 1.5) was used as the starting point for production of the lysine-tagged target vector, pQR730. In the first stage, primers C1TK and LYS2 were used for PCR of the TK gene with addition of a two-lysine tail. The resultant vector, pQR712, was then amplified with C1TK and LYS4 incorporating a further two lysine residues into the tail and producing pQR714. In the final step, C1TK and LYS10 were employed for generation of TK with six additional lysine residues, thus completing construction of the ten-lysine tail. The N-terminal lysine residues inserted at each stage are highlighted in bold-type in the scheme below.

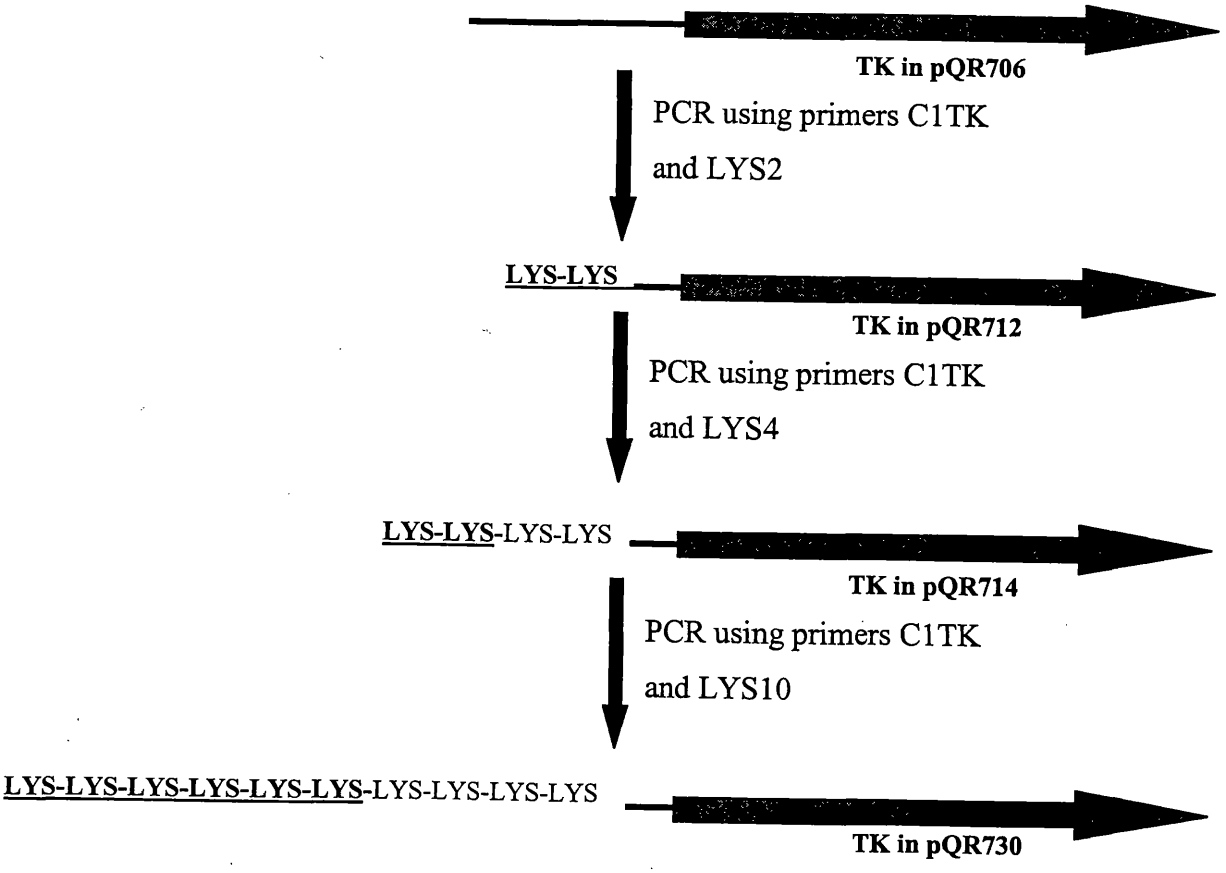



**Figure 3.1** *Scheme for construction of pQR730*

The final expression vector, pQR730, was used to express TK, with an N-terminal affinity tag of ten consecutive lysine residues, in *E. coli* strain JM107. The TK promoter has been removed in this construct and the TK gene is under the control of the *lac* promoter. Cells were grown in LB/amp at 37°C supplemented with IPTG (A1.14.9) to a final concentration of 1mM. Expression level of the enzyme in *E. coli* JM107 cells was 0.137U/ml of crude cell extract supernatant. The plasmid map of pQR730 is given in Figure 3.2.

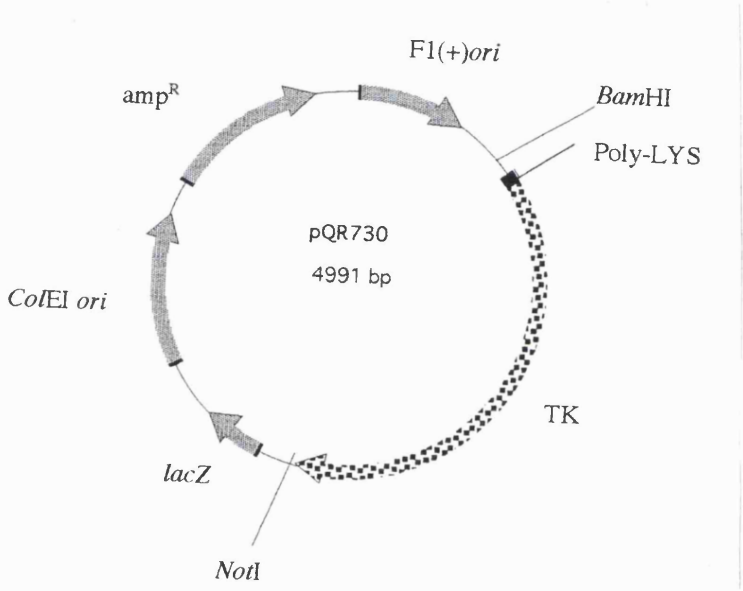



**Figure 3.2** *Plasmid map of pQR730*

### 3.3 First stage of purification

All FPLC took place on a MonoQ$^R$ HR5/5 column at 22°C with a flow rate of 0.5ml/min and was monitored by measuring absorption at 280nm. The first stage of the purification, for which the elution profile is shown in Figure 3.4, involved a NaCl gradient from 0 to 0.5M over 30ml at pH7.0. 1ml of concentrated supernatant (17.5mg protein) was applied to a column pre-equilibrated with Gly-Gly (150mM, pH7.0 [A1.11.2]). The column was washed with 10ml of this buffer, followed by a pre-purification step comprising a pH gradient from 7.0 to 9.5 over 30ml as shown in Figure 3.3, this removes a significant proportion of cellular protein (wild-type TK with

no poly-lysine tail elutes from the column during this stage and is collected in fractions 21 and 22 as represented here). Elution of tagged-TK was afforded by a linear NaCl gradient from 0 to 0.5M over 30ml. All TK (in apo- or holo-form) elutes from the column in a single fraction, number three shown here (1ml, containing 6.5mg total protein).

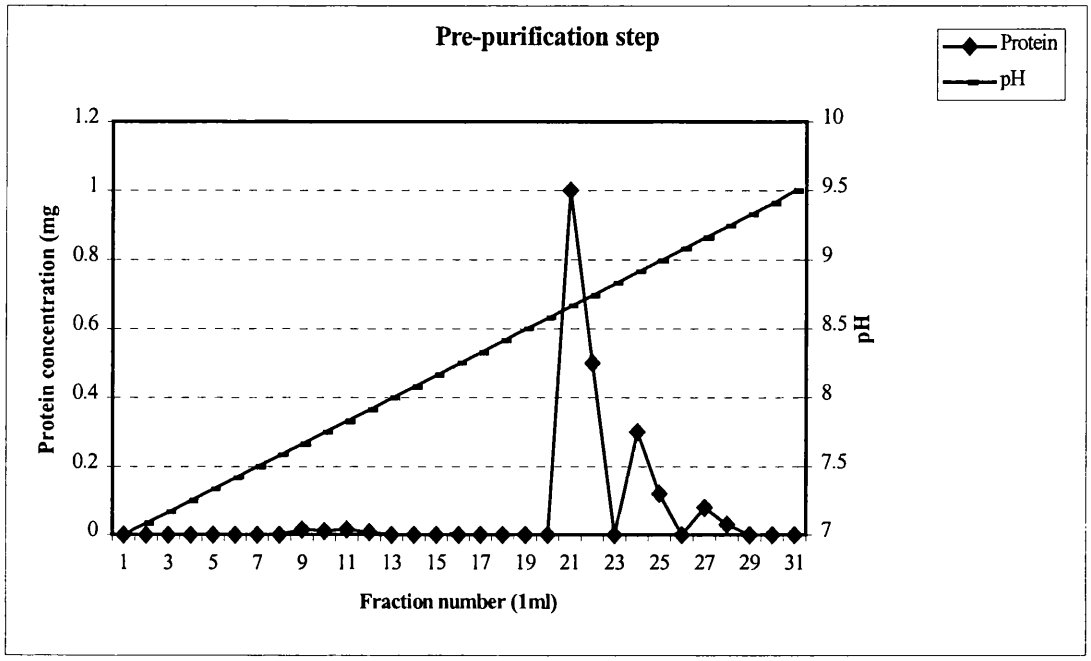



**Figure 3.3** *Pre-purification step (pH gradient 7.0 to 9.5)*

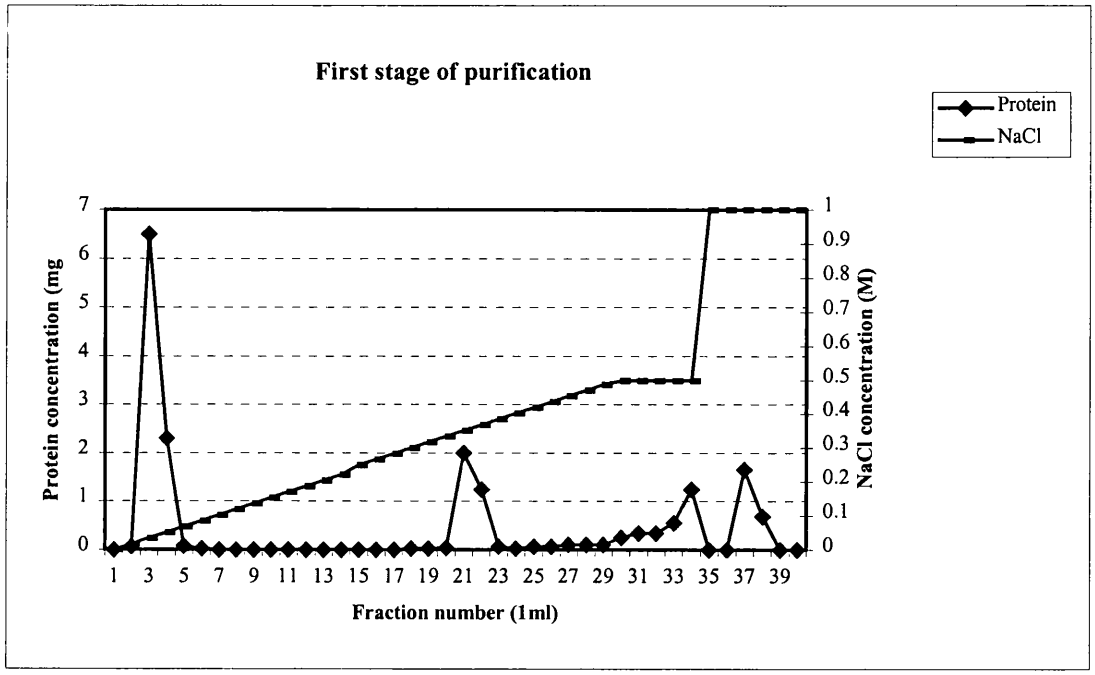



**Figure 3.4** *First stage of purification (NaCl gradient 0 to 0.5M at pH7.0)*

Linked-enzyme assay results showing the activity of fraction 3 and closely surrounding fractions are given in Table 3.1 below. Virtually no activity can be seen for fractions 1, 2, 4 and 5 indicating an absence of TK in these aliquots. Protein was quantified by the Bradford method (Bradford, 1976) and fractions were analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (2.2.12).

| Fraction | Activity (U/ml) | Protein (mg/ml) | Specific Activity (U/mg) |
|---|---|---|---|
| 1 | 0.0000 | 0.00 | 0.000 |
| 2 | 0.0016 | 0.06 | 0.003 |
| 3 | 2.1430 | 6.50 | 0.330 |
| 4 | 0.0048 | 2.30 | 0.002 |
| 5 | 0. 0016 | 0.08 | 0.002 |
| 6 | 0.0000 | 0.02 | 0.000 |
| 7 | 0.0000 | 0.00 | 0.000 |

**Table 3.1** *Specific activities for first stage of purification*

### 3.4 Second stage of purification

To refine the enzyme even further the TK-containing fraction (1ml) from the first purification step was re-applied to the column after equilibration with MES (100mM, pH5.0 [A1.11.3]). The column was washed with 10ml of this buffer, and eluted with a linear gradient of NaCl from 0 to 1M over 30ml. The elution profile for the NaCl gradient at pH5.0 can be seen in Figure 3.5. In this chart, fractions one through five represent the final 5ml of wash buffer prior to the salt gradient. All TK elutes from the column in fraction twelve shown here (1ml, containing 2mg total protein).

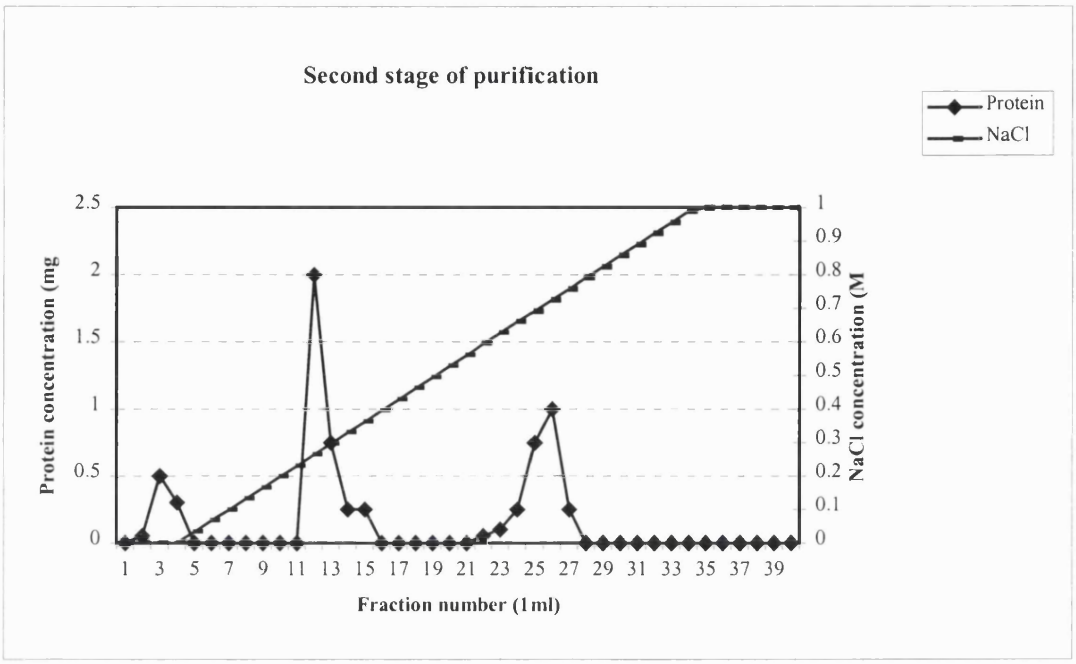



**Figure 3.5** *Second stage of purification (NaCl gradient 0 to 1M at pH5.0)*

Linked-enzyme assay data for fraction 12 along with closely bordering fractions are given in Table 3.2. Very little activity is seen for fractions 10, 11, 13 and 14, suggesting that TK is contained almost exclusively in the single 1ml elution.

| Fraction | Activity (U/ml) | Protein (mg/ml) | Specific Activity (U/mg) |
|---|---|---|---|
| 10 | 0.0000 | 0.00 | 0.000 |
| 11 | 0.0000 | 0.00 | 0.000 |
| 12 | 1.4130 | 2.00 | 0.707 |
| 13 | 0.0063 | 0.75 | 0.008 |
| 14 | 0.0016 | 0.25 | 0.006 |
| 15 | 0.0008 | 0.25 | 0.003 |
| 16 | 0.0000 | 0.00 | 0.000 |

**Table 3.2** *Specific activities for second stage of purification*

Figure 3.6 shows the SDS-polyacrylamide electrophoresis gel containing the purified TK. Electrophoresis was performed on a 15% acrylamide gel and staining was achieved using Coomassie blue (A1.9.9). Details of the gel are as follows: Lane 1, wide range protein standard markers (A1.9.10 [1μg]); Lane 2, final purified protein

(fraction 12 from the second step [15μg]) for which the purity was increased 4.5-fold; Lane 3, native proteins from *E. coli* JM107 cells [25μg]; Lane 4, crude supernatant (JM107/pQR730) with TK at ca. 10% of the total cellular protein [150μg].

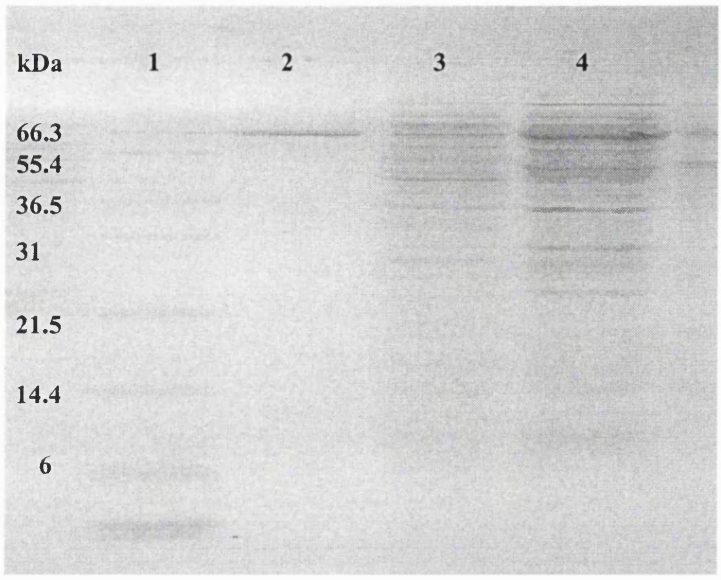



**Figure 3.6** *SDS-PAGE showing purified TK*

Protein was quantified by the Bradford method (2.2.2) and results can be seen in Table 3.3. Values are represented as a percentage of the total cellular protein yield originally applied to the column, and in addition these quantities have been converted into the fold-increase in purity, showing an overall 4.5-fold increase in TK purity.

| Sample | | Protein (mg) | % of total | Purification (fold) |
|---|---|---|---|---|
| Column load | | 17.5 | 100 | 1.000 |
| TK-containing fraction | First step | 6.5 | 37 | 2.113 |
| | Second step | 2.0 | 11 | 4.529 |

**Table 3.3** *Purification of TK (from 200ml E. coli culture)*

Confirmation that this purification procedure preserves the catalytic activity of TK was provided by employment of the linked-enzyme assay throughout these investigations. It appears that overall TK activity decreases at each stage, however it is the specific activities which are of greater importance, these are given as units of enzyme activity

(U) per mg of protein present and are shown in Table 3.4. Given here are the specific activities for TK from the two purification steps along with values for the crude extracts from both pQR706 and pQR730 (the figures given here for TK from pQR730 are those for the crude extract concentrated by 20 times prior to purification). From comparison of these values it can be seen that there is a considerable increase in specific activity as the enzyme is purified (*ie.* as TK constitutes a greater proportion of the protein being analysed).

| Sample | Activity (U/ml) | Protein (mg/ml) | Specific activity (U/mg) |
|---|---|---|---|
| Crude extract (TK from pQR706) | 4.330 | 0.62 | 6.98 |
| Crude extract (TK from pQR730 [conc.]) | 2.730 | 17.50 | 0.156 |
| First step | 2.143 | 6.50 | 0.330 |
| Second step | 1.413 | 2.00 | 0.707 |

**Table 3.4** *Purification of recombinant TK from E. coli JM107 cells*

Recombinant TK from pQR730/JM107 is over-expressed at ca. 10% of the total cellular protein, whereas that from pQR706/JM107 is present at roughly 20% of the total protein. As can be seen from the data given in Table 3.4, the specific activity of recombinant TK produced from pQR730/JM107 is much lower than that of the wild-type protein generated from pQR706 (roughly 48 times). Even in the case of the purified poly-lysine tagged protein, the specific activity is still only a tenth of that for wild-type TK. It may be that the presence of the poly-lysine tail results in the formation of inclusion bodies of over-expressed protein from pQR730. The protein thus produced would have greatly diminished activity since it would be unable to form the quaternary structure necessary for efficient catalytic activity.

### 3.5 Structure of purified TK

LEAs employing a modified reaction mixture (with no TPP, $Mg^{2+}$) show that only a small amount of holo-enzyme elutes from the column, this is highlighted by the data given in Table 3.5. As can be seen here, the specific activity of the purified enzyme from both stages is considerably reduced when the assay is performed in the absence of the cofactors TPP and $Mg^{2+}$ (shaded boxes), suggesting that a considerable proportion of enzyme is present in the apo-form. The majority of the TK, in the inactive apo-form,

is subsequently activated under the conditions of the assay. Analogous tests on crude extract showed that the TK loaded onto the column was almost exclusively in the active holo-form, this was ensured by adding TPP [0.25mM] and $Mg^{2+}$ [9mM] to the buffers used in supernatant preparation (however values given here for the crude extract were also true for TK extracts prepared in the absence of its cofactors).

| Sample | Activity (U/ml) | Protein (mg) | Specific activity (U/mg) |
|---|---|---|---|
| Crude extract -TPP | 2.725 | 17.50 | 0.156 |
| Crude extract +TPP | 2.730 | 17.50 | 0.156 |
| 1st step (fraction 3) -TPP | 0.156 | 6.50 | 0.024 |
| 1st step (fraction 3) +TPP | 2.143 | 6.50 | 0.330 |
| 2nd step (fraction 12) -TPP | 0.025 | 2.00 | 0.013 |
| 2nd step (fraction 12) + TPP | 1.413 | 2.00 | 0.707 |

**Table 3.5** *Investigation into enzyme structure*

Structural information can be extrapolated from this data and is summarised in Table 3.6, this data highlights the growing proportion of apo-enzyme as purification proceeds. The growing presence of this from of the enzyme results from a deficiency of its cofactors under these conditions, intensifying with increasing manipulation of TK.

| Sample | % apo-enzyme | % holo-enzyme | apo- : holo-enzyme |
|---|---|---|---|
| Crude extract | 0.18 | 99.82 | 1:555 |
| 1st step (fraction 3) | 92.72 | 7.28 | 13:1 |
| 2nd step (fraction 12) | 98.23 | 1.77 | 56:1 |

**Table 3.6** *Proportion of apo-: holo-enzyme*

## 3.5 Discussion

The methodology detailed in this chapter provides a technique for a 4.5-fold purification of a poly-lysine tagged version of TK. Purified enzyme is obtained almost exclusively in the inactive apo-form, but is readily converted to the active holo-enzyme

under the enzymes' biotransformation conditions with the addition of its cofactors TPP and $Mg^{2+}$ ions. Purified TK thus produced is then available for further investigations regarding its activity with a range of substrates.

During large-scale biotransformations it is useful if the biocatalyst can be immobilised as this facilitates the recycling of the enzyme, thus increasing its useful life. In an attempt to immobilise TK from a crude extract, a poly-lysine tail was attached to the 5'-end of the TK gene from *E. coli* (French & Ward, 1994). The TK dimer has a number of lysine residues situated on its surface and therefore covalent immobilisation of the wild-type enzyme would be random, accompanied by the risk that it may involve amino acid residues essential for the catalytic activity of the enzyme or those important for substrate binding. The poly-lysine tail was designed to help to orientate the enzyme, by creating a specific region on the protein for immobilisation to take place. The position of the N-terminus with respect to the active site is known from the three dimensional structure of the protein (Littlechild *et al.*, 1995). The tail was added to the 5'-end of the gene as the N-terminus is projected away from the protein, therefore ensuring that the tail is free from the active site region.

Initial immobilisation experiments were conducted by Carol French and yielded promising results (French & Ward, 1994). However, prior to this project, no use had been made of the poly-lysine tail for work relating to the purification of the protein. Since the poly-lysine tagged TK was already available, it was decided to attempt to devise a purification method for the protein using affinity chromatography, exploiting the fact that this version of TK is able to strongly bind a MonoQ$^{R}$ HR5/5 column.

However, a number of limitations were found to be associated with this approach. Firstly, it takes a relatively long time to perform the purification using this method when compared to other types of affinity chromatographic techniques. Secondly, the specific activity of poly-lysine tagged TK was found to be considerably less than that for the non-tagged protein derived from pQR706 (roughly 48 times lower). Even in the case of the final purified protein, the specific activity was still only a tenth of that for the wild-type enzyme. A possible explanation for this is that the presence of the poly-lysine tail results in the formation of inclusion bodies of over-expressed protein from pQR730. The protein thus produced would have much diminished activity, since it would not be in the correct quaternary structure to effectively perform its enzymatic

function. Alternatively the poly-lysine tail may reduce enzyme activity for other reasons, for example altering the local pH or ion concentration around the N-terminus, which could have an effect extending out to the active site. In addition, the recombinant TK produced from pQR730/JM107 was only over-expressed at around 10% of the total cellular protein, whereas that generated from pQR706/JM107 (without the poly-lysine tail) was found to over-express at around 20% of the total cellular protein. Therefore, for a particular volume of crude enzyme extract, roughly twice as much active enzyme was present in the case of the wild-type protein.

Taking all of these factors into consideration, it was concluded that this technique was not as effective for the purification of TK as had been originally planned. Further TK mutants generated during this project (detailed in Chapters Four and Five) were not constructed into this system, and crude enzyme extracts were in fact used for all mutant characterisation work carried out during this project. Numerous attempts by Carol French to produce an N-terminal poly-histidine tagged version of TK, to be purified using $Ni^{2+}$ - NTA chromatography (Schmitt, *et al.*, 1993), were unsuccessful. As a result it was thought that the presence of a poly-histidine tail, in this particular location on the protein, may have been in some way incompatible with TK. It was for this reason that work was undertaken to attempt to purify the poly-lysine tagged protein, which had been previously constructed and was readily available.

In light of the problems encountered using the poly-lysine tagged protein for FPLC purification, it may be worth attempting to construct a C-terminal poly-histidine tagged version of TK for possible future applications in purification work. Although it must be considered that the C-terminus is not projected away from the protein surface as is the case with the N-terminus, and thus a C-terminal tag may in some further way interfere with the enzymatic function. The C-terminus is, however, still some distance away from the active site (Littlechild *et al.*, 1995). Alternatively, advances in PCR techniques since this initial work was conducted in 1994, might now make the construction of the N-terminal histidine-tagged protein possible. Another option to consider would be to attempt to clone the TK gene into a commercial histidine-tag vector.

A further strategy to consider in the case of a situation where a relatively pure form of TK were to be required would be the use of an alternative form of affinity

chromatography, one possibility being employment of cofactor-mimic columns. This type of approach has been successful for enzymes such as dehydrogenases where there is a rapid association/dissociation of the enzyme/cofactor complex, as a large proportion of the enzyme is uncomplexed as it passes through the column and thus is readily bound. In contrast, the TPP cofactor for TK is strongly bound within the active site of the enzyme dimer (Egan & Sable, 1981; Booth & Nixon, 1993), with apparent Km values for the enzyme from *Saccharomyces cerevisiae* of around 1.5μM (Nilsson *et al.*, 1997). Within a crude cell extract the proportion of holo-:apo- enzyme is roughly 40:60 (French & Ward, 1994). Due to these factors, this methodology may not be appropriate for TK since the active holo-enzyme dimers may run straight through the column and thus a large fraction of the TK from the crude extract would be lost. This suggests that a different type of affinity column would be more suitable in this case. One option would be to devise a column incorporating a linear sugar-like end attached to the column matrix. Attachments with carbon chain lengths of about six carbon atoms could be used, ensuring that the substrate-mimic ends are able to reach down into the long active site funnel of TK and thus into the region where catalytic activity occurs. This type of column could also be tailored to function as a type of screening technique, in order to isolate enzymes with specifically altered stereoselectivities such as those described in Chapter Five, this approach will be discussed further in Chapter Seven.

# Chapter Four: Applications of modified TK for synthesis using aromatic aldehydes

# CHAPTER FOUR: Applications of modified TK for synthesis using aromatic aldehydes

## 4.1 Introduction

Conventional synthetic organic chemistry has been successful in the production of numerous pharmaceutical compounds, but fails as a satisfactory method in many cases particularly concerning such issues as product yield and purity. Hence there is a need for an alternative approach to complement existing traditional synthetic methods and for this reason the use of biotransformations, to carry out synthetically useful organic reactions is becoming increasingly popular.

Often an enzyme will not only transform the substrate, or substrates, involved *in vivo* but will accept a whole range of compounds and convert these with the same regio- and stereoselectivity, a feature commonly exploited by industry. However, for effective catalysis the substrate must satisfy certain physical and chemical criteria dictated by the conformation of the active site of the enzyme and since many drugs contain bulky sterically hindered groups, intermediates in their production pose considerable steric challenges. This requires consideration when developing chemoenzymatic syntheses and modification of the biocatalyst, to accommodate such substrates, may be necessary.

Formation of carbon-carbon bonds is of central importance to synthetic organic chemistry and enzymes such as TK, whose function *in vivo* is the degradation of carbohydrates, may prove to be valuable biocatalysts for selective C-C bond formation. Variants of this enzyme with expansion of the active site, designed to accommodate more bulky substrates, would greatly enhance its synthetic applications. A range of TK mutants intended for this purpose were generated, and their transformations with a range of aromatic aldehydes were investigated.

## 4.2 Initial TK mutants

Using the enzyme crystal structure (Littlechild *et al.*, 1995) as a model for their design a selection of TK mutants, with assorted combinations of amino acid changes located within the active site, were generated by Carol French (French & Ward, 1996) using a polymerase chain reaction (all work to characterise these mutants was conducted during this project). It was anticipated that substituting large residues, located at key positions

within the active site tunnel, with alanines would lead to incorporation of cavities at these positions and the widened active site funnels thus produced were expected to permit entry of more bulky aldehyde acceptors.

Mutations were incorporated using the relevant mutagenic primer, as listed in Appendix Two, and modified genes were cloned into pCRScript SK(+). Genes to be over-expressed in *E. coli* Amber suppressor strains were subsequently cloned into pALTER™-Ex2. pQR706 (section 1.5) was used as a starting point for construction of these mutants and a plasmid map for this vector is given in Figure 4.1 below. These constructs were combinatorially tested with a selection of aromatic compounds, and initial screening to identify successful pairings was performed using HPLC techniques.

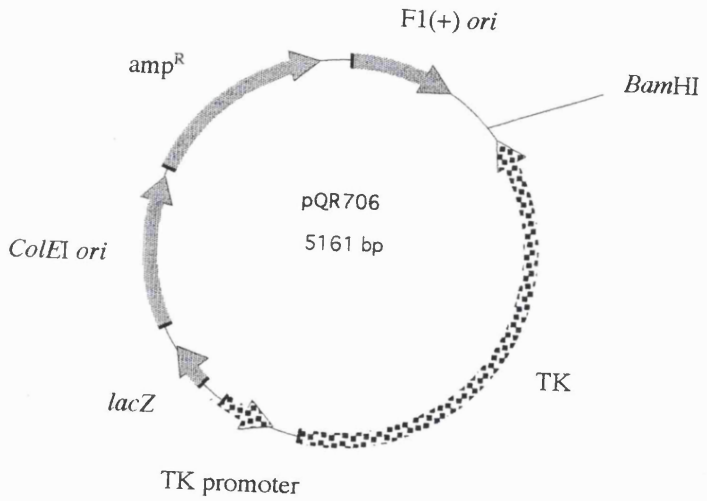



**Figure 4.1** *Plasmid map of pQR706 (French & Ward, 1995)*

Amino acid substitutions were made at locations in which the residues were believed to play a significant role in substrate and/or cofactor binding. An initial set of seven TK mutants were constructed. Three of these alterations were incorporated by replacement of the desired amino acid codon with a stop codon, constructs were subsequently expressed in the relevant Amber-strain (Kleina *et al.*, 1990; Normanly *et al.*, 1990) as part of the INTERCHANGE™ *in vivo* Amber Suppression System. Each strain carries a different amber suppressor tRNA which can be used to insert a specific amino acid at the amber stop position during expression. In this way, the stop codon was translated into the desired amino acid residue within the TK gene. The remaining four mutants were generated by direct amino acid substitutions, base changes for all seven constructs are given in Table 4.1. Although in theory, the amber interchange system enables production of a large number of mutants from a single genetic manipulation, attempts

during this project to generate a wide range of TK mutants using this system were unsuccessful. Efficient transformation of constructs into these strains proved to be extremely difficult and vectors that were effectively transformed commonly produced enzymes exhibiting greatly reduced activity.

Positions of all amino acid residues described here are depicted in Figure 1.11. From this information, along with factors relating to the size, charge, and hydrophilicity of the amino acid side-chains, predictions regarding the roles of these groups in catalysis, substrate binding and orientation of the substrate can be made. I189 is positioned roughly half way down the active site funnel and its large hydrophobic side-chain points towards the TPP cofactor. It has been reported that the main chain oxygen of this residue is involved in the binding of the $Mg^{2+}$ ions (Schneider & Lindqvist, 1998). H461* is located close to the entrance, opposite the two arginine residues R358* and R520* (section 1.3.2), together they create an overall positive charge at the active site opening. It is therefore possible that this residue is also a constituent of the phosphate-binding site. S385* is situated just past the active site entrance and its hydroxyl side chain may be involved in hydrogen bonding interactions with the hydroxyl moiety of the sugar phosphate group. H258 sits midway along the active centre and points towards residue F434* located on the opposite face of the funnel. It is believed that the positively charged side chain of this residue is involved, through hydrogen bonding interactions, in the partial sharing of pairs of electrons with the α-hydroxyl group of the aldehyde acceptor. Together with the bulky hydrophobic side chain of F434*, this may be an important factor in determining the stereoselectivity of TK. Here however attention is focused on TK activity towards aromatic aldehydes, work to investigate factors contributing to TK stereoselectivity is described in Chapter Five.

| *E. coli* strain/construct | Alteration | Base changes |
|---|---|---|
| Amber-LEU/pQR735 | S385*STOP (L) | TCT to TAG |
| Amber-LEU/pQR738 | H461*STOP (L) | CAC to TAG |
| Amber-SER/pQR738 | H461*STOP (S) | CAC to TAG |
| JM107/pQR739 | I189A | ATC to GCC |
| JM107/pQR749 | I189V | ATC to GTC |
| JM107/pQR776 | H258A | CAC to GCC |
| JM107/pQR748 | H258D | CAC to GAC |

**Table 4.1** *Transketolase Mutants (constructed by Carol French)*

Amino acids derived from the second subunit are marked*

These constructs over-expressed TK (using its own promoter), in *E. coli* strain JM107 or the appropriate Amber-strain, with varying levels of enzyme activity. Observed expression levels were adequate for the investigative TK transformations necessary for this research. Initial characterisation was conducted using the linked-enzyme assay method of Heinrich *et al.*, 1972 (2.2.1). Control assays, carried out in the absence of either recombinant TK-containing extract or Ribose-5-P, showed that there was no observable loss of NADH due to side-reactions and so these values were considered an accurate reflection of TK activity. All assays were performed in triplicate, alongside assays for reactions conducted with TK from pQR706/JM107 and that from JM107 cells to be used as standards by which to evaluate the new mutants.

Specific activities of these modified enzymes were calculated and values, along with that for the native protein, are given in Table 4.2. Examination of this data shows that all mutant TKs have reduced activity towards the native sugar substrates relative to the wild-type protein, this would be expected since these amino acid substitutions cause the active site structure to deviate from its optimal conformation for these compounds.

The mutant produced from pQR735(L) [S385*L] had considerably reduced activity towards the native sugars, this decline was possibly due to the larger leucine side chain blocking entry into the active site. Also, there may have been a loss of interactions involving the serine hydroxyl moiety with the phosphate group of these sugars, thus leading to less efficient binding of these substrates. Enzyme produced from pQR738(L) [H461*L] had reduced activity in the LEA and here also, interactions with the phosphate groups of the sugars have been lost. For recombinant TK from pQR738(S) [H461*S] there was diminished activity in the LEA too, however this was not so great as for replacement with leucine. In this case, replacement of the positively charged histidine with the hydroxyl group of serine may mean that hydrogen bonding interactions still occur, but with the phosphate oxygens rather than the hydrogen atoms. In addition, the presence of the hydrophobic side chain of leucine in this usually positively charged region, may result in unfavourable interactions at the active site entrance.

In the case of TK produced from pQR739/749 [I189A/V] there was much lower activity for the open chain sugars and this could be attributed to less efficient binding of the

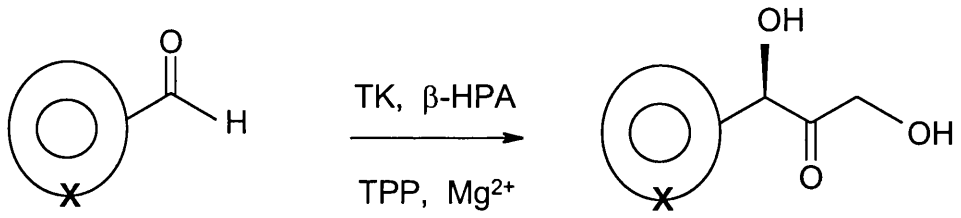
$Mg^{2+}$ ions, thus leading to a more unstable protein dimer. Enzymes produced from pQR776/748 [H258A/D] both had reasonably high activity in the LEA, along with an increased reactivity towards the aromatic aldehydes. It can therefore be concluded that replacement of this residue does not have large deleterious effects on the active centre of TK.

For cases where it was intended to introduce a cavity into the active site funnel the amino acid residue concerned was replaced by an alanine residue rather than a glycine. Since glycine is largely involved in turns within the protein structure, this methodology was chosen so as to reduce the possibility of additional unwanted structural changes occurring within the mutant protein. Furthermore, the introduction of an alanine residue maintains a more hydrophobic environment whilst still generating the desired hole within the active site.

| *E. coli* strain/construct | Activity (U/ml) | Concentration (mg/ml) | Specific Activity (U/mg) | Error (%) |
|---|---|---|---|---|
| JM107/pQR706 | 4.330 | 0.62 | 6.98 | +/- 4.6 |
| Amber-LEU/pQR735 | 0.444 | 0.60 | 0.74 | +/- 1.3 |
| Amber-LEU/pQR738 | 0.508 | 0.60 | 0.85 | +/- 1.4 |
| Amber-SER/pQR738 | 1.317 | 0.71 | 1.85 | +/- 2.1 |
| JM107/pQR739 | 0.635 | 0.52 | 1.22 | +/- 1.5 |
| JM107/pQR749 | 0.286 | 0.52 | 0.55 | +/- 1.1 |
| JM107/pQR776 | 4.030 | 0.67 | 6.01 | +/- 3.6 |
| JM107/pQR748 | 2.620 | 0.71 | 3.69 | +/- 2.8 |

**Table 4.2** *Specific activities for TK mutants*

Initial investigations into the synthetic potential of these modified *E. coli* TKs were conducted using β-hydroxypyruvic acid (β-HPA) as the ketol donor with a range of water-soluble aromatic compounds as aldehyde acceptors, a general reaction scheme for which is given in Figure 4.2 below. Reactions were initially monitored using HPLC, following the loss of β-HPA due to enzymatic conversion. Effective mutant/aldehyde combinations were employed for larger scale biotransformations, with subsequent extraction and characterisation of reaction products.



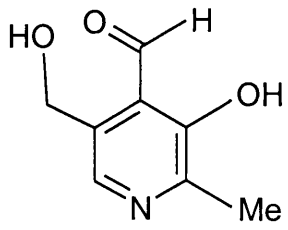

**X = O or N**

**Figure 4.2** *General scheme for TK reaction with aromatic aldehydes*

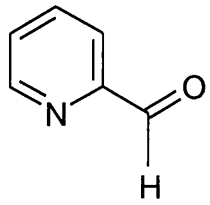
## 4.3 HPLC analysis

In order to screen for enhanced activity towards more bulky substrates, the TK mutants were used to transform the six water-soluble aromatic aldehydes whose structures are shown in Figure 4.3. Selected compounds were commercially available to high purity, and comprised a variety of aromatic/heterocyclic systems. Biotransformations were performed (2.2.19) at substrate concentrations of 100mM (both β-HPA and aldehyde), with TK produced from O/N culture (3ml), in a final reaction volume of 10ml. Reactions were allowed to run for 18h and continual monitoring was conducted using TLC (ethyl acetate:methanol = 95:5).

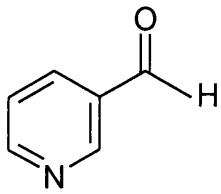
The isocratic HPLC system employed here was that devised by Mitra & Woodley, 1996. It consisted of two Aminex$^{R}$87H columns (7.8 x 300mm) in series, with 6mM $H_2SO_4$ mobile phase running at 0.45ml/min. Columns were maintained at 65°C and peaks were detected by a change in refractive index. Dilution of reaction samples in the acidic mobile phase (pH 2.3) to the concentration required for the assay (ca. 20mM) simultaneously terminated the TK reaction. The substrates were stable under these acidic conditions for the time required for analysis. Samples were filtered through 0.45μm filters prior to injection (10μl).



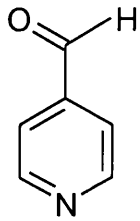

A. Pyridoxal



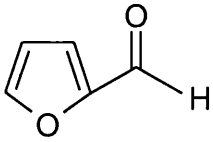

B. 2-Pyridinecarboxaldehyde



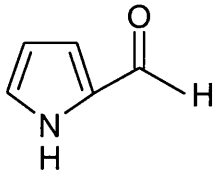

C. 3-Pyridinecarboxaldehyde




D. 4-Pyridinecarboxaldehyde




E. 2-Furancarboxaldehyde




F. 2-Pyrrolecarboxaldehyde

**Figure 4.3** *Aromatic aldehydes chosen for analysis with modified TK*

A standard curve for β-HPA was plotted (see Figure 4.4) to allow HPLC integrated peak areas to be converted into a concentration value (mM), the retention time for β-HPA under these conditions was 21.32 minutes. For each of the concentrations shown (mM) the relevant quantity of β-HPA was run through the HPLC system three times. This chart enabled a direct concentration loss of this substrate, due to enzymatic conversion, to be calculated. Control reactions measuring the loss of β-HPA under these assay conditions, but with reactions run in the absence of crude enzyme extract (with all other reaction components present), were performed and it was found that no appreciable

degradation of this compound occurred under these conditions. Therefore calculated values were considered to be reasonable within the limits given, although controls were not performed using reactions run under these conditions but in the absence of the aromatic substrates A-F (with all other reaction components present). It was found that a number of problems were encountered when using this approach, most noticeably the lack of mass balances found during these assays. These difficulties, along with concerns regarding some of the standard curves obtained for use in concentration calculations, are detailed in the relevant sections of both this chapter and in Chapter Five.

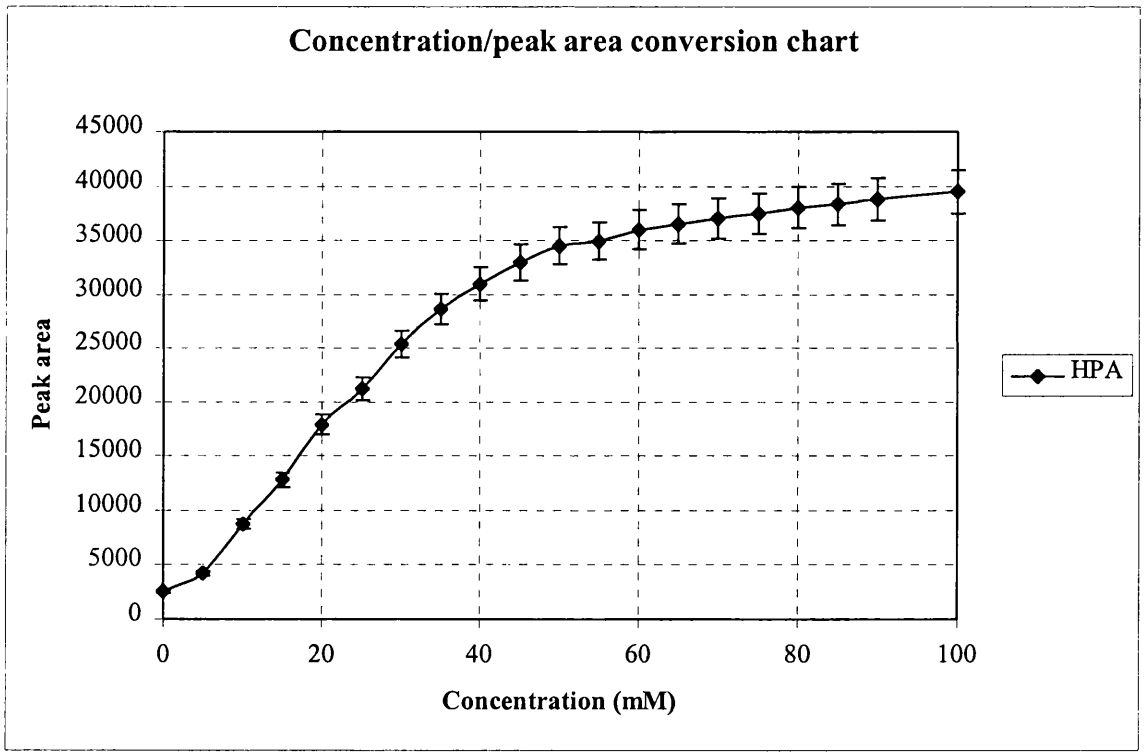



**Figure 4.4** *Standard HPLC curve for β-HPA*

The twin column HPLC system detected only the β-HPA in the reaction mixture and transformations were followed by measurement of the loss of this substrate (see Figures 4.5 and 4.6). As can be seen from these two charts, a wide range of activity levels for the various mutant/aldehyde combinations were observed.

Losses in β-HPA (mM) for each of these reactions are shown in Figures 4.5 and 4.6. Most of the measurements for loss of β-HPA (mM) shown here fall in the region 0-50mM, a region for which the β-HPA standard curve is relatively linear. It was

therefore considered that concentration calculations made using this information were reasonably reliable. In Figure 4.5 data is presented in a form aimed to facilitate comparisons between the various TK mutants. Constructs are shown as eight blocks along the horizontal axis and the six columns within a block each represent one of the aldehydes. From this chart it can be seen that several of the modified TKs, when compared to the native protein, have enhanced activity towards these aromatic compounds. In particular, TK produced from pQR739 [I189A] exhibits much increased transformation efficiency. Introduction of the smaller alanine residue incorporates a hole within the active site funnel, thus allowing the more facile entry of a range of larger substrates.

Figure 4.6 shows identical data, but in a form allowing direct comparisons to be made between the six aldehydes. Here, the six blocks each represent one of the aldehyde substrates, and the eight columns within each of these blocks each represent a particular TK mutant. From this preliminary information it was apparent that aldehydes B through D (*ie.* the three pyridinecarboxaldehyde compounds) displayed the greatest reactivity, a more detailed explanation of these observations is given in section 4.4.

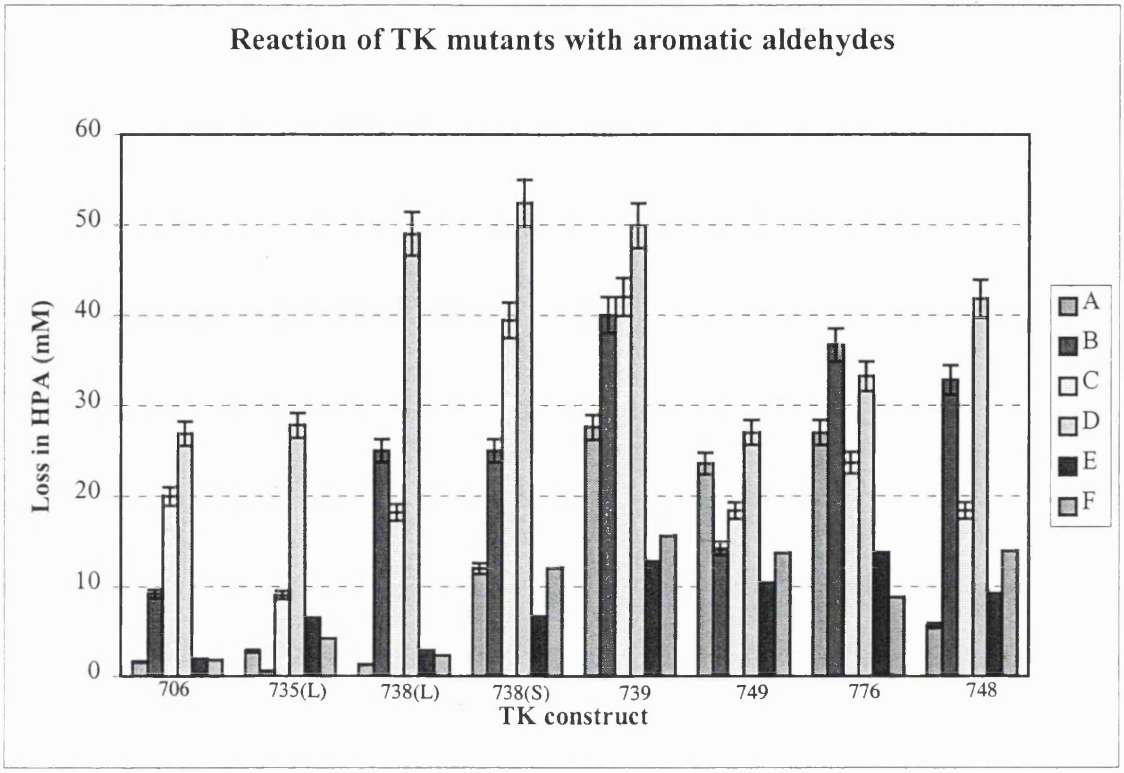



**Figure 4.5** *Loss in β-HPA over 18h for each of the eight TK constructs*

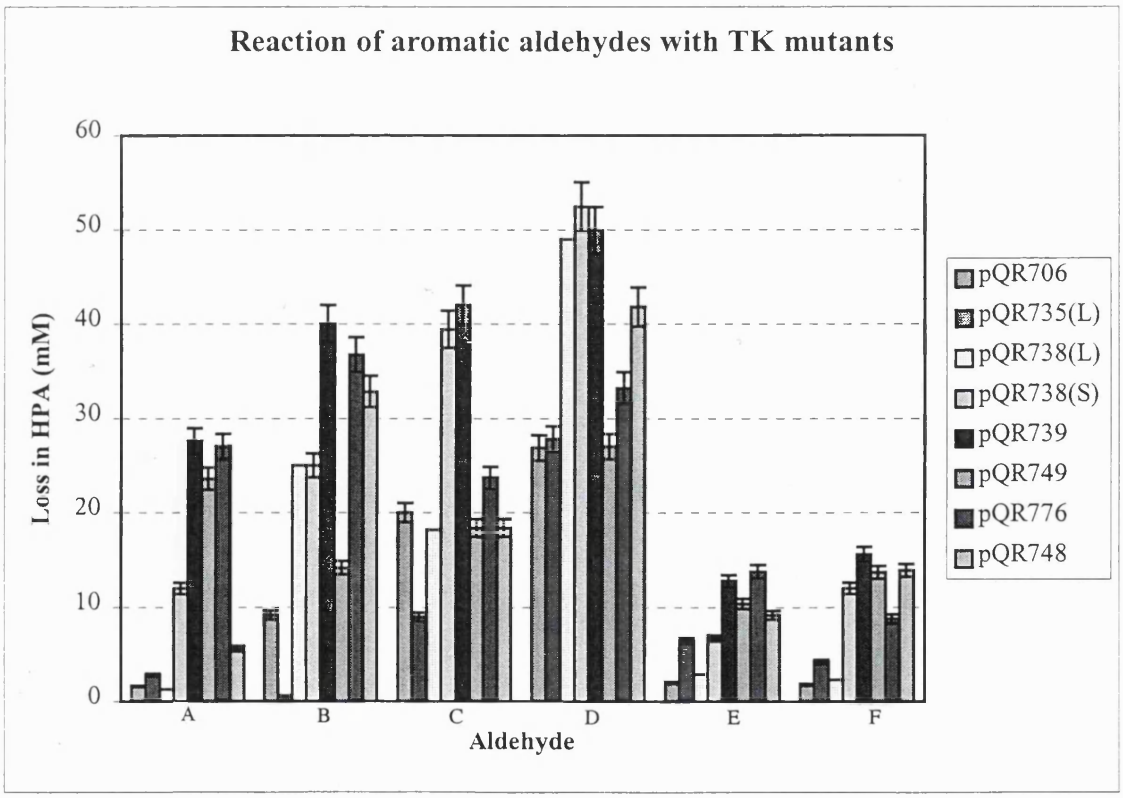



**Figure 4.6** *Loss in β-HPA over 18h for each of the six aromatic aldehydes*

For recombinant TK from pQR739 and pQR749, the observed increase in activity towards a selection of more bulky aldehydes may result from the opening up of the active site funnel. This reduces unfavourable hydrophobic interactions with these compounds and outweighs the decline in activity due to the increased instability of the protein caused by these amino acid substitutions. For mutants derived from pQR776 and pQR748, the substitution of H258 with a residue possessing a smaller side chain also appears to partially open up the active site to more bulky substrates.

Additionally, it was observed that the reactivity of the wild-type TK (from pQR706) towards E and F is very low (2.0 and 1.8mM β-HPA transformed in 18 hours) but several of the mutants have over five times greater reactivity towards these substrates. This is particularly so for enzymes generated from pQR739, pQR749, pQR776 and pQR748. An interesting comparison can be made between TK from pQR738(L) and pQR738(S) where the replacement of a leucine for a serine at the same amino acid position doubles the activity towards C whilst the activity towards B and D remain almost constant. In addition, the activity towards A is increased over 9-fold by changing from leucine to serine at this position whilst the activities towards compounds E and F are increased by 2x and 5x respectively.

Since no reaction of β-HPA was seen in the control, the loss of β-HPA was taken to be equivalent in molar terms to the amount of aldehyde converted, it can be seen that 50% of the D aldehyde is transformed by recombinant TK from pQR738(L), pQR738(S) and pQR739 (at least 25% of this compound was transformed by all of the eight TK constructs and overall it was the most reactive aldehyde). It is also interesting to compare of the activities of TK from pQR739 and pQR749 in which an alanine residue has been replaced by a valine at the same amino acid position. Here the activity of mutant TK from pQR739 towards the three pyridinecarboxaldehydes B, C and D is roughly twice that of the TK from pQR749, whereas their apparent respective activities towards substrates A, E and F were very similar.

Since TK from pQR739 [I189A] (French & Ward, 1996) consistently exhibited enhanced reactivity towards these novel aldehydes relative to the reactions of the native protein, this mutant was selected for more in depth investigations. Biotransformations were conducted on a larger scale (5x) for each of the six aldehydes, using TK from this

particular construct, reaction products were then extracted and structurally characterised using NMR and mass spectrometry.

## 4.4 Product extraction and characterisation

The synthetic potential of the modified enzyme [I189A] was further investigated using β-HPA as the ketol donor and the same six aromatic aldehydes as acceptors. Reactions were monitored using TLC ($R_f$ values were as follows: pyridoxal, 0.44; 2-pyridinecarboxaldehyde, 0.55; 3-pyridinecarboxaldehyde, 0.56; 4-pyridinecarboxaldehyde, 0.51; 2-furancarboxaldehyde, 0.59; 2-pyrrolecarboxaldehyde, 0.60; 2-pyridinecarboxaldehyde reaction product, 0.72; 3-pyridinecarboxaldehyde reaction product, 0.70; 4-pyridinecarboxaldehyde reaction product, 0.67) and products were extracted with diethyl ether. Structural characterisation was provided by NMR and mass spectrometry analysis. 5mg of sample was submitted for $^{1}H$ NMR and mass spectrometry, and 20mg for $^{13}C$ NMR examination. Control spectra were obtained for each of the other reaction mixture components in order to allow these to be deducted from reaction product spectra where necessary.

### 4.4.1 NMR analysis

Although some loss of β-HPA was detected during HPLC analysis for reaction with pyridoxal (Figure 4.5), no observable reaction products were found after extraction and $^{1}H$ NMR characterisation. This is shown in Figures 4.7-8, for which the two NMR spectra are identical (*ie.* no new peaks, representing product material protons, can be seen in the spectrum for the reaction mixture).

HPLC analysis showed small losses in β-HPA for both 2-furancarboxaldehyde and 2-pyrrolecarboxaldehyde. This was originally believed to be due to TK conversion, however these findings were not supported by $^{1}H$ NMR data for these reactions, shown in Figures 4.15-16 and 4.17-18 respectively. No new peaks, due to protons present in reaction products, can be seen in the spectra for TK reactions with these compounds. This suggested that the TK variants investigated thus far were in fact relatively ineffective for transformation of these particular substrates, although TK produced from pQR739, pQR749, pQR776 and pQR748 do show increased transformation of E and F from ca. 2% (wild type TK) to 10-15%. It may have been that this small quantity of product was not extractable in the particular solvent system employed, it would

therefore be worth while in future work to repeat these biotransformations and to investigate the effectiveness of alternative solvent systems for their extraction. Alternatively, it may be that other components from the crude enzyme extract used for these biotransformations are in some way making the extraction process less effective for these particular compounds. A better approach may be to devise a flash chromatographic method to purify the reaction products for subsequent structural characterisation work.

All three pyridinecarboxaldehyde compounds were converted into reaction products which could be extracted and structurally characterised. NMR data for compounds produced from these three biotransformations were as follows:

2-pyridinecarboxaldehyde reaction product

$^{1}$H NMR (400MHz) DMSO-$d_6$ δ: 4.86 (d, 1H, J=-12)$^{F}$, 4.60 (d, 1H, J=-12)$^{G}$, 5.96 (s, 1H)$^{E}$, 7.35 (t, J=6.0, 1H)$^{B}$, 7.55 (d, J=7.2, 1H)$^{C}$, 7.85 (t, J=7.8, 1H)$^{D}$, 8.43 (d, J=6.7, 1H)$^{A}$. $^{13}$C NMR (400MHz) DMSO-$d_6$ δ: 62$^{8}$, 72$^{6}$, 122$^{2}$, 124$^{4}$, 137$^{3}$, 147$^{1}$, 148$^{5}$, 195$^{7}$.

3-pyridinecarboxaldehyde reaction product

$^{1}$H NMR (400MHz) DMSO-$d_6$ δ: 4.82 (d, 1H, J=-12)$^{F}$, 4.62 (d, 1H, J=-12)$^{G}$, 6.05 (s, 1H)$^{E}$, 7.40 (t, J=5.8, 1H)$^{B}$, 7.90 (d, J=7.5, 1H)$^{C}$, 8.45 (d, J=5.9, 1H)$^{A}$, 8.55 (s, 1H)$^{D}$. $^{13}$C NMR (400MHz) DMSO-$d_6$ δ: 62$^{8}$, 77$^{6}$, 124$^{2}$, 135$^{3}$, 142$^{4}$, 148$^{1}$, 150$^{5}$, 194$^{7}$.

4-pyridinecarboxaldehyde reaction product

$^{1}$H NMR (400MHz) DMSO-$d_6$ δ: 4.83 (d, 1H, J=-12)$^{F}$, 4.58 (d, 1H, J=-12)$^{G}$, 5.97 (s, 1H)$^{E}$, 7.45 (d, J=7.5, 2H)$^{B+C}$, 8.45 (d, J=7.5, 2H)$^{A+D}$. $^{13}$C NMR (400MHz) DMSO-$d_6$ δ: 62$^{8}$, 76$^{6}$, 122$^{2+4}$, 144$^{3}$, 150$^{1+5}$, 194$^{7}$.

The $^{1}$H NMR data for reaction of the six aldehydes with TK from pQR739 are shown in Figures 4.7 through 4.18. This information was supported by $^{13}$C NMR analysis of the compounds.

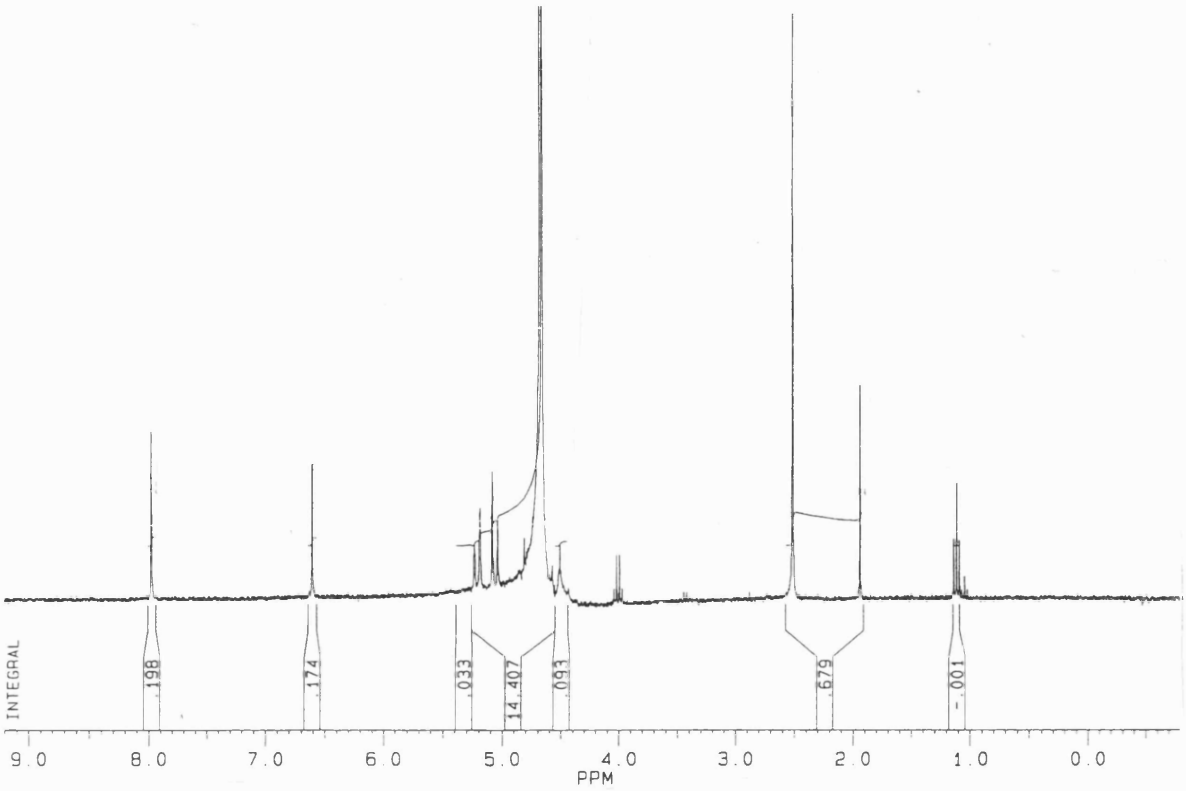



**Figure 4.7** *$^1H$ NMR for pyridoxal*

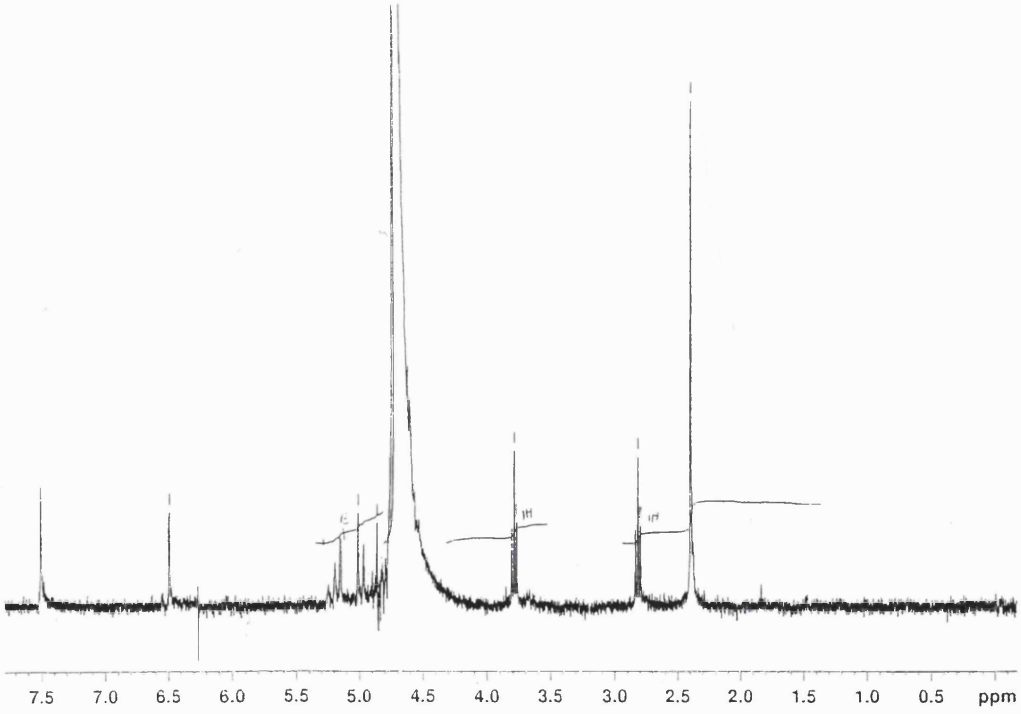
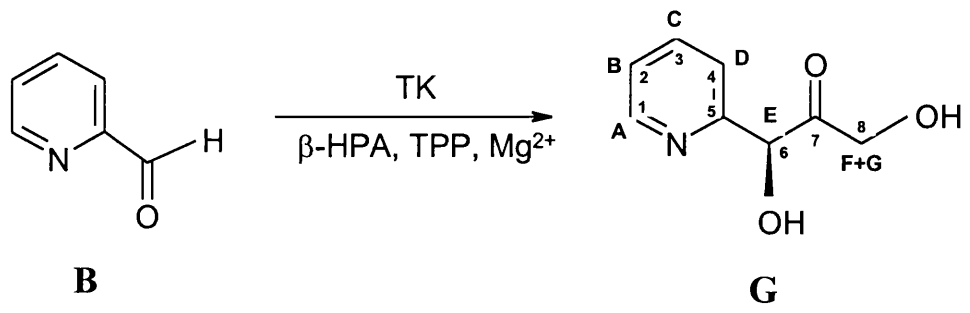


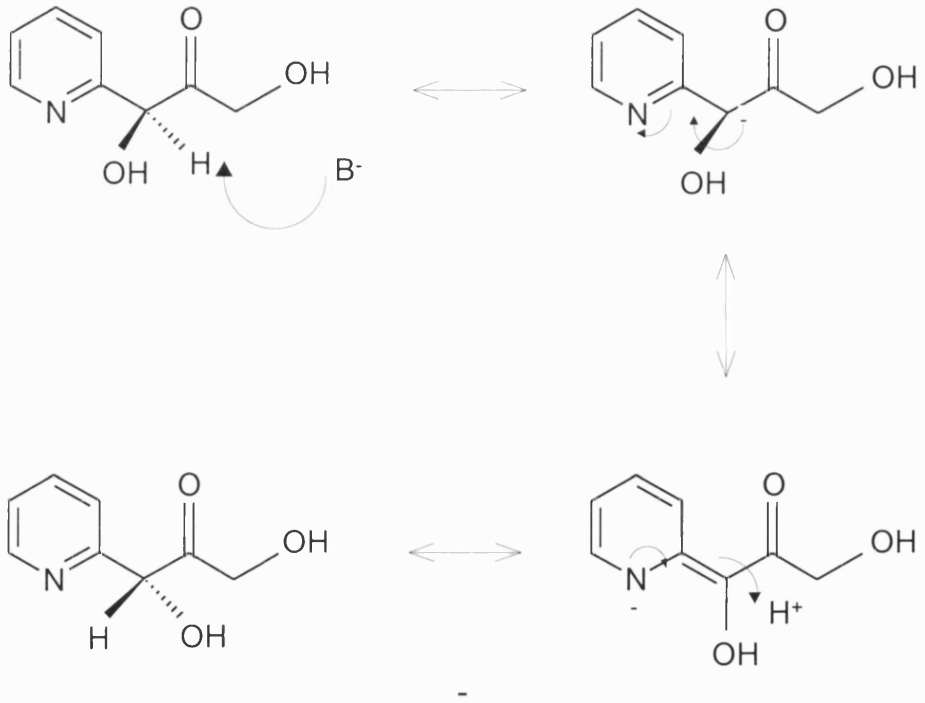

**Figure 4.8** *$^1H$ NMR for pyridoxal reaction*

*TK reaction with 2-pyridinecarboxaldehyde*




An observable product was generated during the TK reaction with this compound, as can be seen in Figure 4.10. Substantial peaks for the starting material (see Figure 4.9) are still present after transformation, but there are an additional seven peaks representing product protons. Three of these are in the region < 6ppm and result from the non-aromatic protons E, F and G. The remaining four, in the region 7-9ppm, represent the new aromatic protons A, B, C and D. These peaks are shifted slightly from those depicting aromatic protons in the starting material since due to a minor alteration in the electronic environment, a small deviation in the nuclear shielding results. This means that the nuclei are affected differently by the magnetic field and so appear to be non-equivalent in the NMR spectrum (this is true also, for reactions with 3-pyridinecarboxaldehyde and 4-pyridinecarboxaldehyde, depicted in Figures 4.12 and 4.14 respectively).

For the biotransformation product, compound **G**, there is the possibility of interconversion between stereoisomers at the newly formed stereo-centre via the mechanism outlined below. The resultant negative charge generated by abstraction of proton E can be readily stabilised by the adjacent aromatic ring, thus in this instance (and similarly for products from TK conversions of aldehydes **C** and **D**) a racemic mixture of product compound may ultimately be obtained.




*Interconversion between stereoisomers for TK/pyridinecarboxaldehyde reaction products*

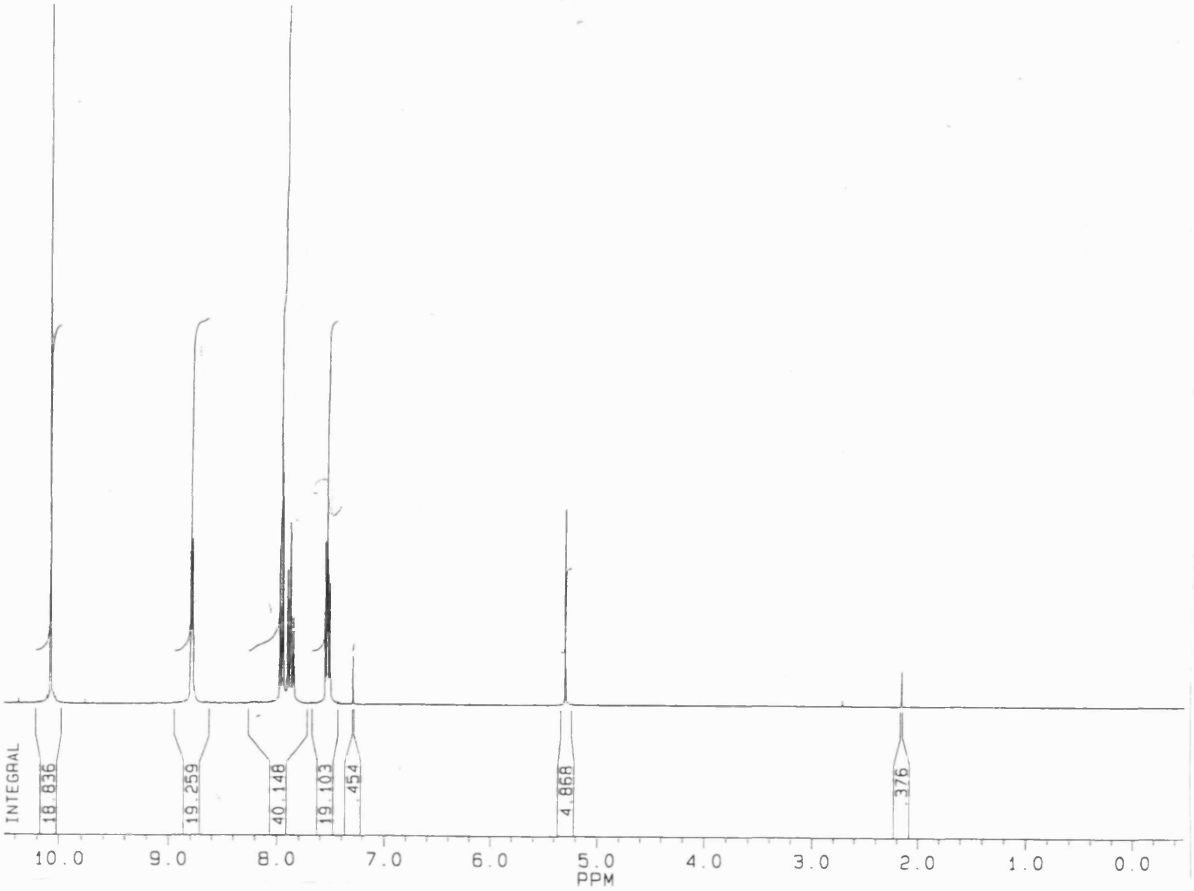



**Figure 4.9** *$^1H$ NMR for 2-pyridinecarboxaldehyde*

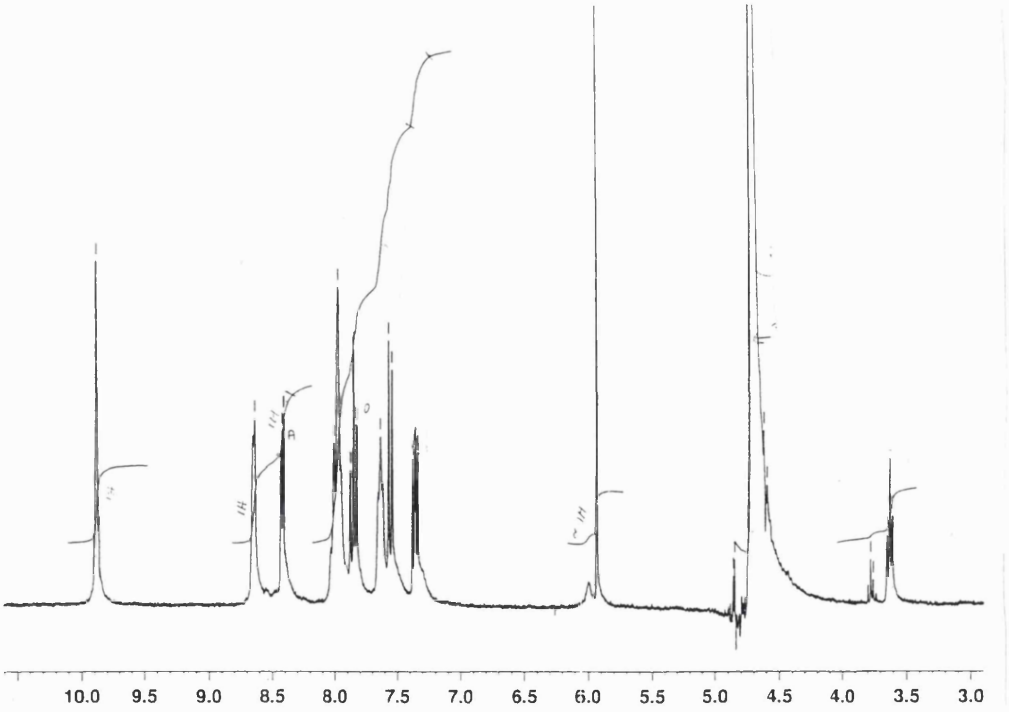
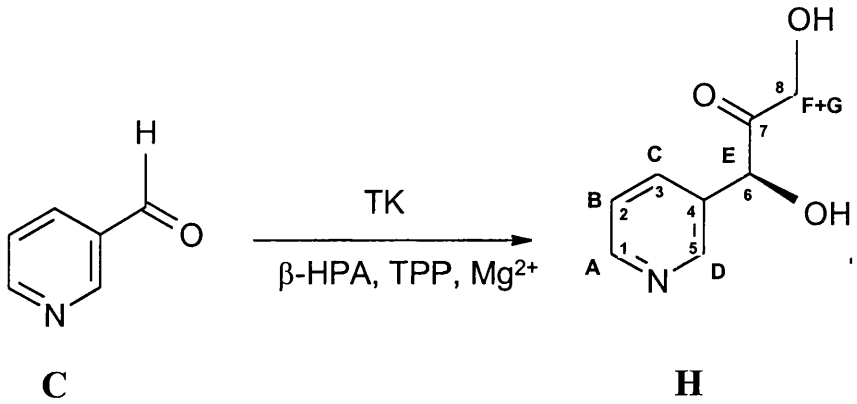


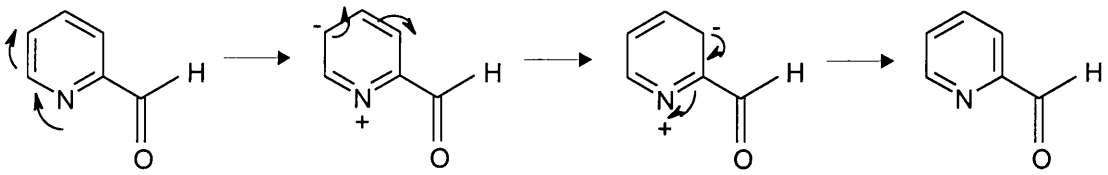

**Figure 4.10** *$^1H$ NMR for 2-pyridinecarboxaldehyde reaction*

*TK reaction with 3-pyridinecarboxaldehyde*

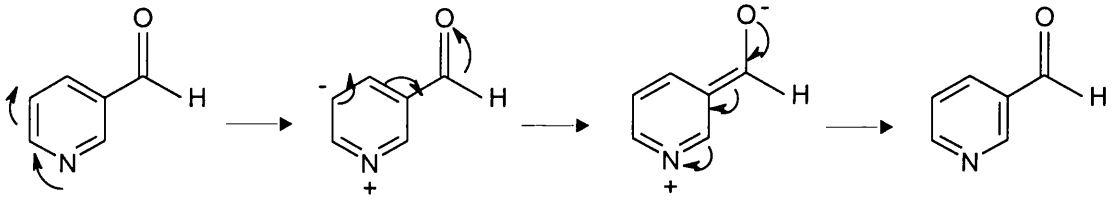



Reaction with this compound also afforded a characterisable product on transformation using TK, although the conversion occurred to a lesser extent than for 2-pyridinecarboxaldehyde. This is highlighted in Figure 4.12 (explanation of new peaks is the same as that for 2-pyridinecarboxaldehyde) where it can be seen that for peaks in the aromatic region (7-9ppm), a greater proportion of starting material remains. This may be attributed to the differing electronic environments within the two compounds as highlighted in the electron delocalisation schemes show overleaf. It can be seen that for the *ortho-* and *para-* compounds the lone pair of electrons from the aromatic nitrogen atom can be delocalised around the aromatic ring, but not onto the aldehyde oxygen atom. For the structure in which the carbonyl group is *meta-* to the aromatic nitrogen atom the carbonyl group may also become involved in the electron delocalisation. When this occurs electron density is pushed onto the carbonyl carbon atom, thus making it less electrophilic and consequently less susceptible to the nucleophilic addition reactions catalysed by TK. This would explain why a lower percentage of *meta-*pyridinecarboxaldehyde was converted into product material than was the case for the corresponding TK-catalysed transformations of *ortho-* and *para-*pyridinecarboxaldehyde.

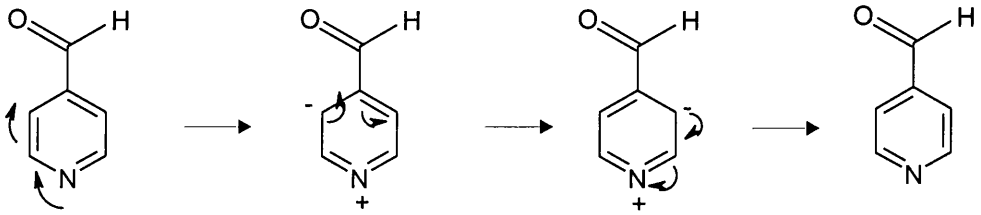
***ortho***




***meta***




***para***




*Electron delocalisation around the aromatic systems of the three pyridinecarboxaldehydes*

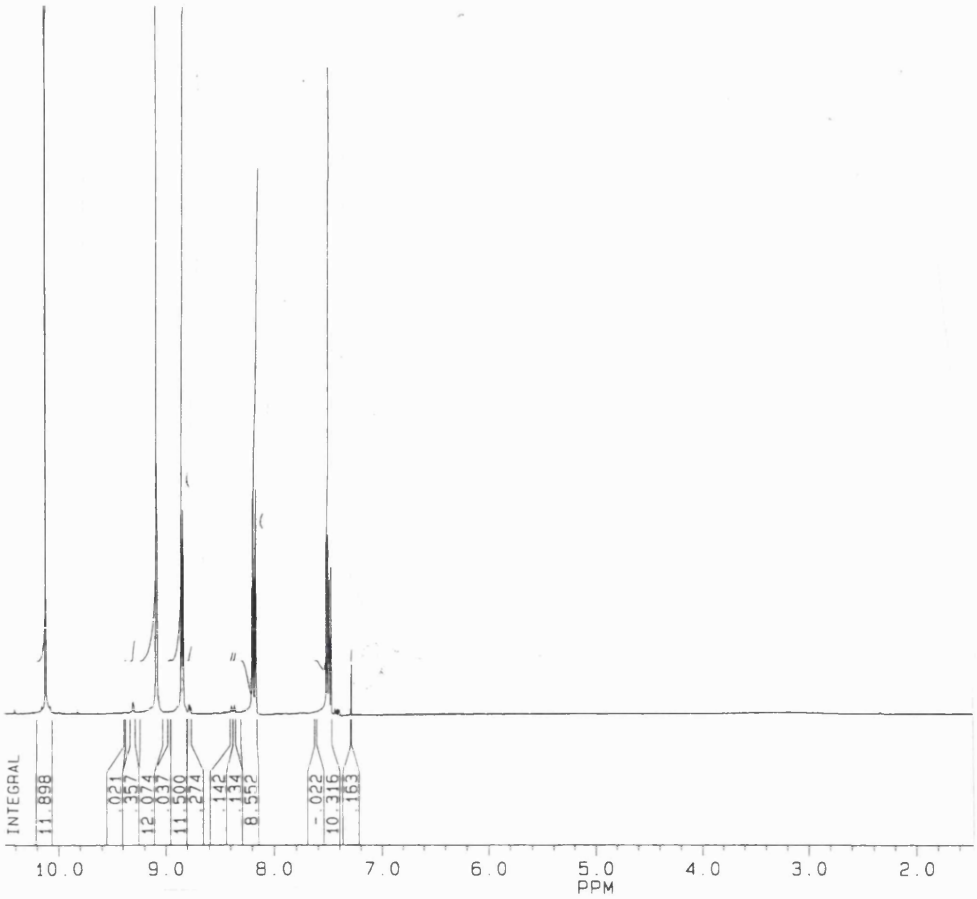



**Figure 4.11** *$^1H$ NMR for 3-pyridinecarboxaldehyde*

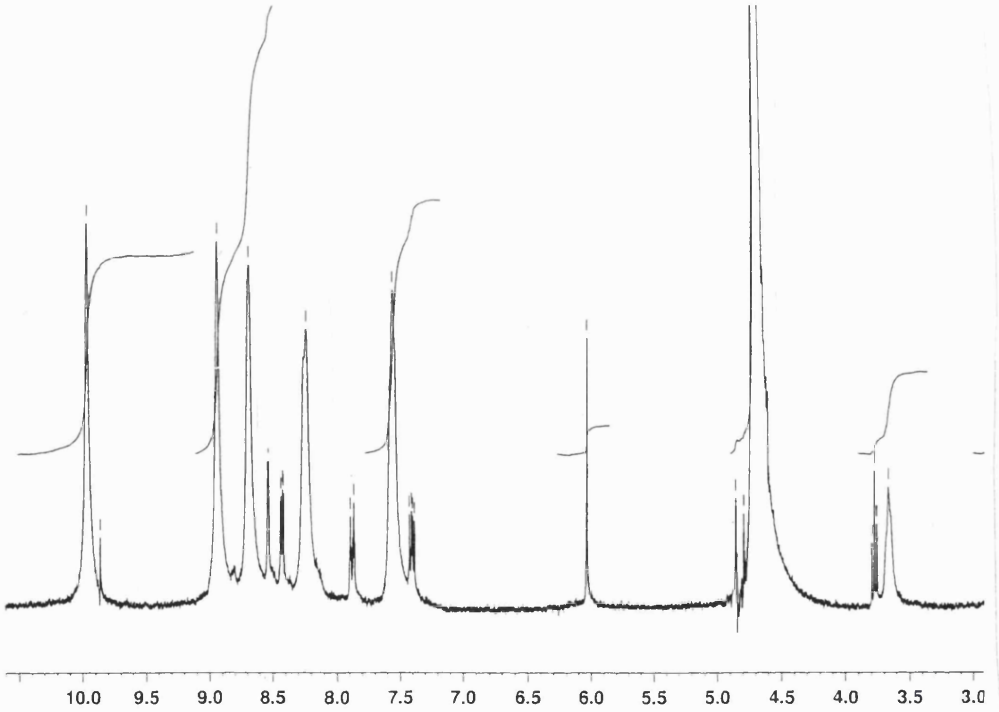
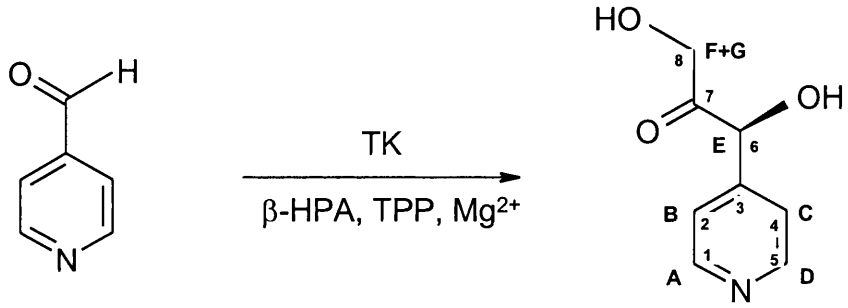



**Figure 4.12** *$^1H$ NMR for 3-pyridinecarboxaldehyde reaction*

*TK reaction with 4-pyridinecarboxaldehyde*




**D** **I**

For this compound a substantial amount of reaction product is observed. As can be seen in Figure 4.14, comparison of the starting-material and product aromatic protons shows that a significantly greater quantity of product material is present. This may be largely explained by the electronic factors previously discussed regarding the differing electron delocalisation patterns for the *ortho*- *meta*- and *para*-compounds. In addition, having the nitrogen lone pair of electrons positioned *para*- to the carbonyl group may be a more favourable arrangement for the particular electronic environment within the TK active site, thus resulting in the apparent greater transformation efficiency of this compound relative to *ortho*-pyridinecarboxaldehyde.

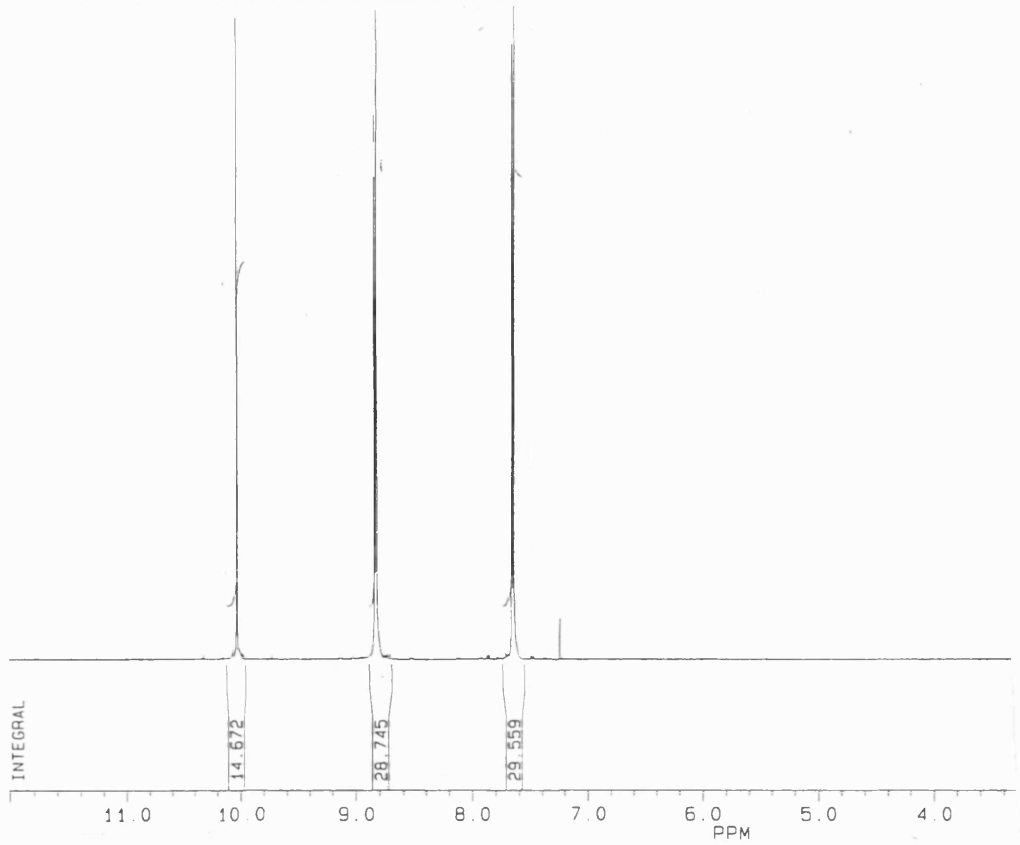



**Figure 4.13** *$^1H$ NMR for 4-pyridinecarboxaldehyde*

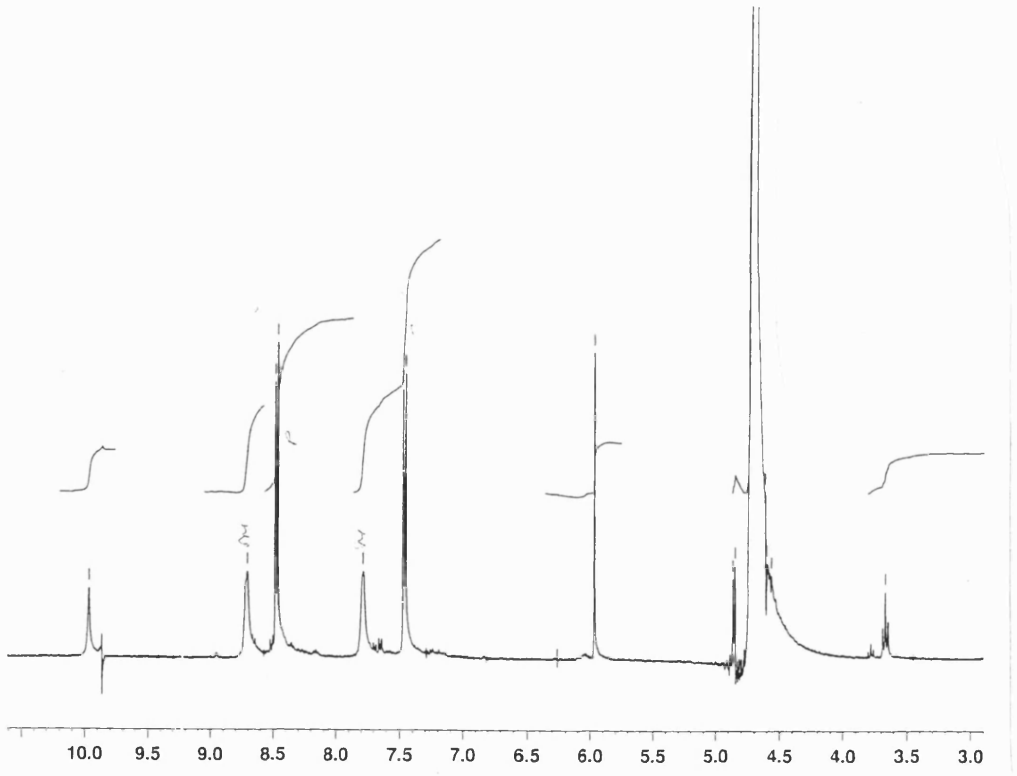



**Figure 4.14** *$^1H$ NMR for 4-pyridinecarboxaldehyde reaction*

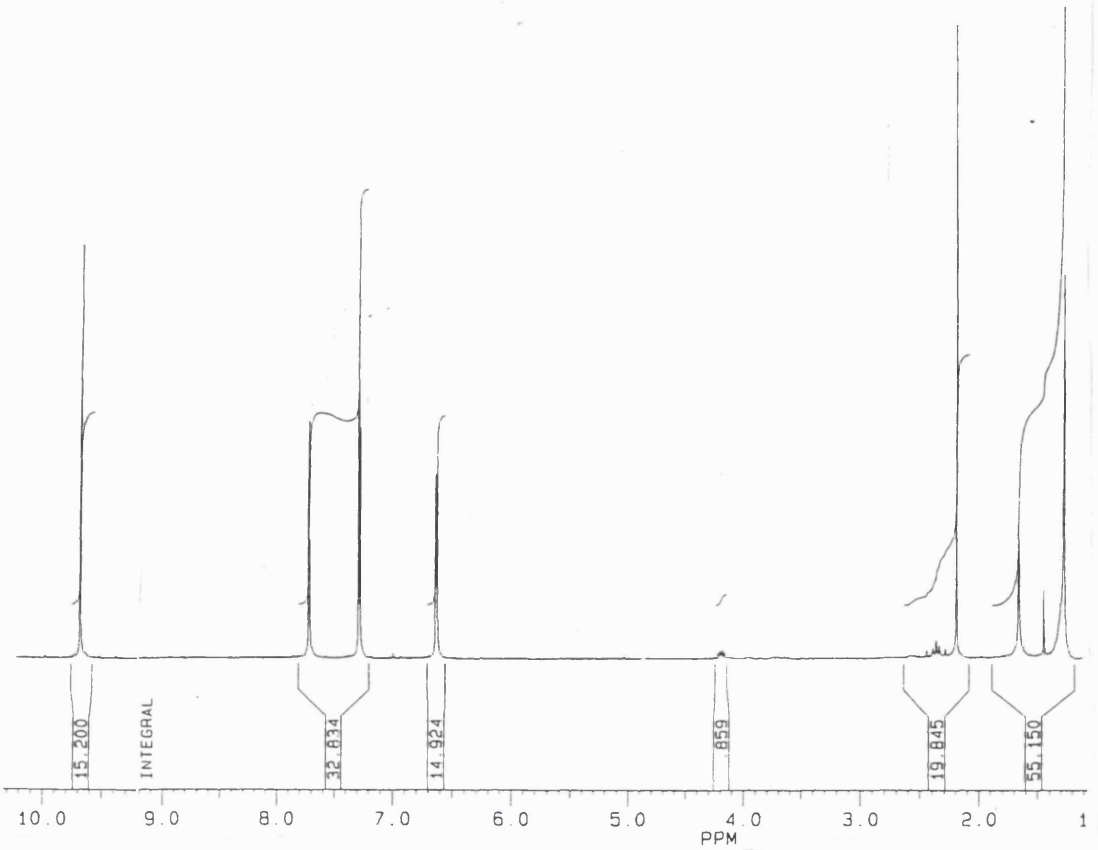



**Figure 4.15** *$^1H$ NMR for 2-furancarboxaldehyde*

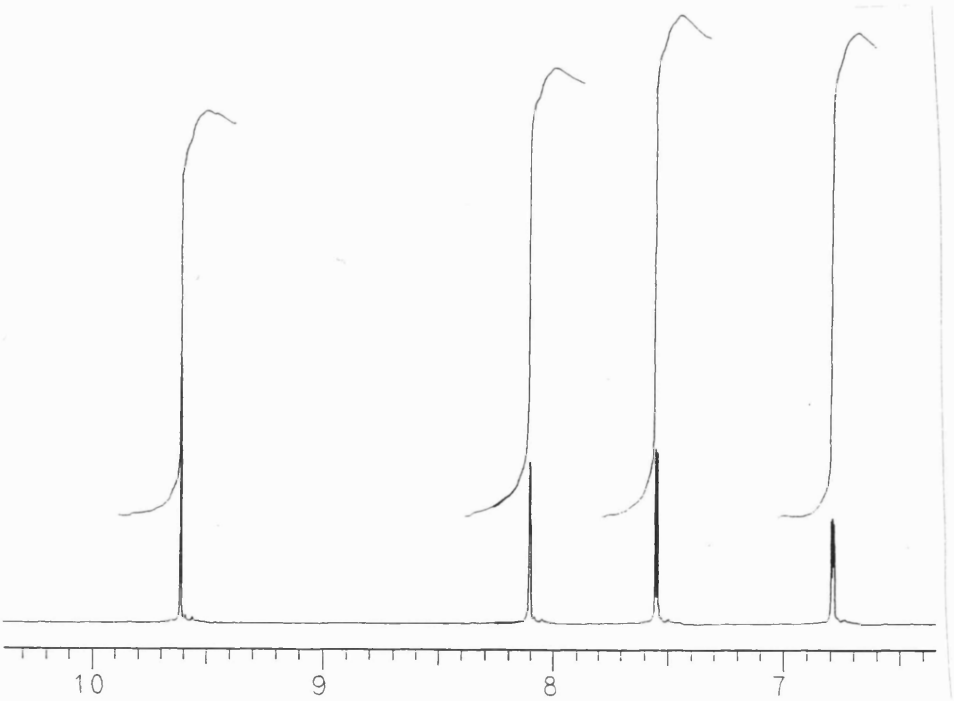



**Figure 4.16** *$^1H$ NMR for 2-furancarboxaldehyde reaction*

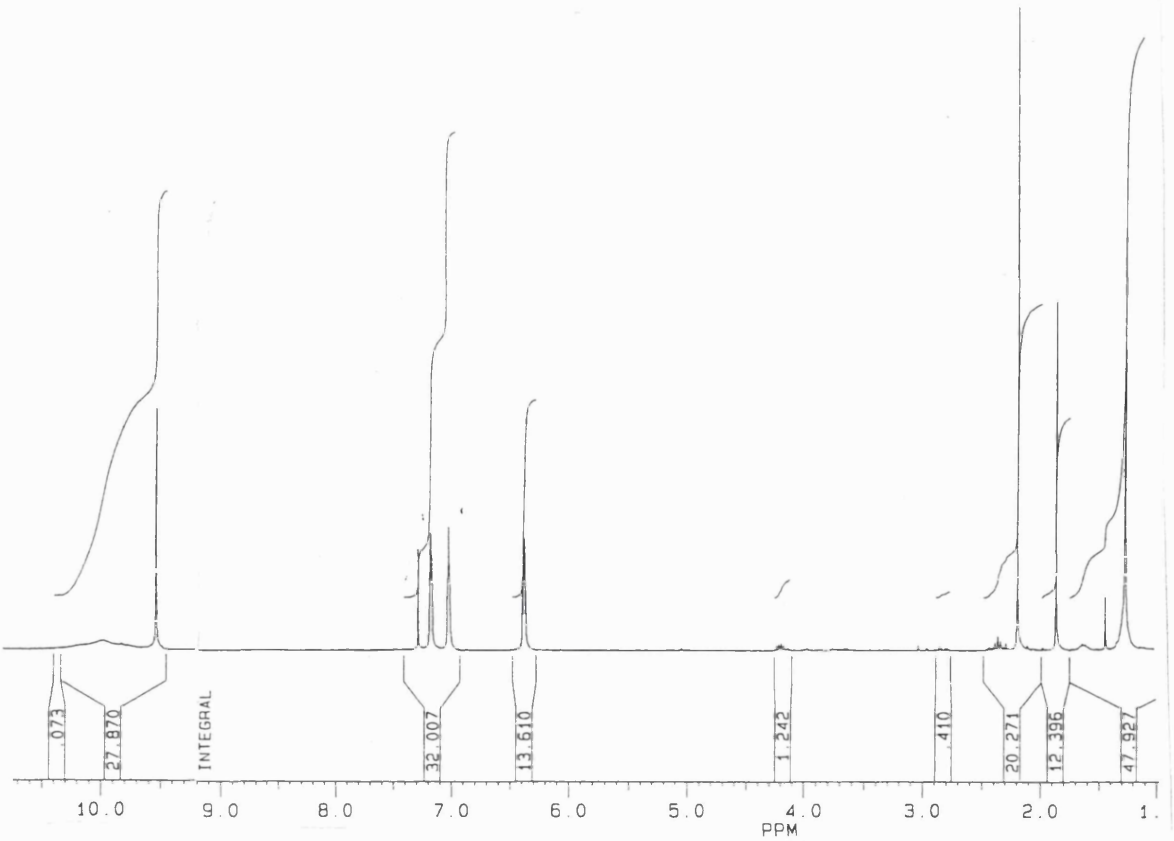



**Figure 4.17** *$^1H$ NMR for 2-pyrrolecarboxaldehyde*

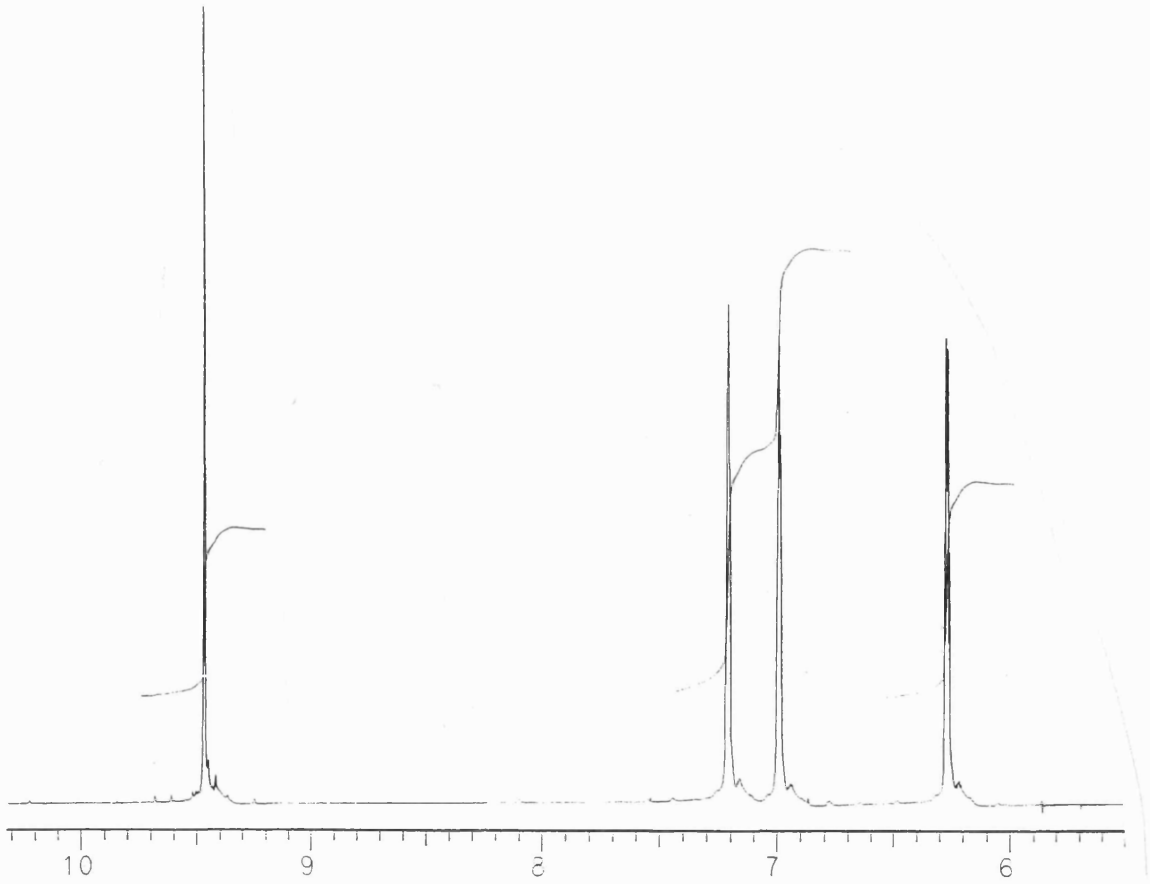



**Figure 4.18** *$^1H$ NMR for 2-pyrrolecarboxaldehyde reaction*

$^1$H NMR data is summarised in Table 4.3. Comparisons have been made between integrated peak areas for starting materials and products and in addition, the information has been presented as a percentage conversion value for each of the three successful transformations. The compound 4-pyridinecarboxaldehyde showed the greatest transformation efficiency with TK produced from pQR739/JM107, this reaction gave an accurate mass product peak (+/- 10ppm) at 167.0587 (calculated value = 167.0580).

| Aldehyde | Aldehyde:Product (NMR integration) | Conversion efficiency (%) |
|---|---|---|
| A. pyridoxal | | |
| B. 2-pyridincarboxaldehyde | 1.00:1.03 | 51 |
| C. 3-pyridinecarboxaldehyde | 3.94:1.00 | 20 |
| D. 4-pyridinecarboxaldehyde | 1.00:3.16 | 76 |
| E. 2-furancarboxaldehyde | | |
| F. 2-pyrrolecarboxaldehyde | | |

**Table 4.3** *Conversion efficiencies calculated from NMR integration*

Reactions of TK from pQR739 with pyridoxal, 2-furancarboxaldehyde, and 2-pyrrolecarboxaldehyde afforded no observable reaction products upon NMR analysis. For pyridoxal this could be due to the large number of side chains attached to the pyridine ring, making this compound still too bulky to be accommodated within the expanded active site funnel. For reactions with 2-furancarboxaldehyde and 2-pyrrolecarboxaldehyde the lack of reactivity may be attributed to the electron-donating ability of the heteroatom in the rings of these compounds (see Figure 4.19). Consequent reduction of the electrophilicity of the aldehyde carbon atom will thus make these two substrates less efficient aldehyde acceptors in the nucleophilic addition reactions catalysed by TK.

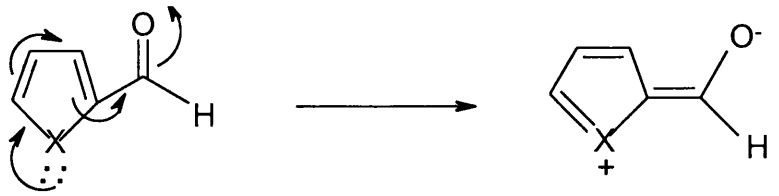



**Figure 4.19** *Electron donation by heteroatom*

However, for all three of these compounds some degree of β-HPA loss was observed during the HPLC analysis of their reactions with TK from pQR739/JM107 (A: 28mM; E:13mM; F:15mM). It may simply be the case that insufficient product material was generated under the reaction conditions employed here to enable detection within the limits of the NMR techniques used. Alternatively it may be that products could not be extracted efficiently under the particular conditions used during this work, in this instance an investigation of alternative extraction procedures would be desirable. Conducting the biotransformations on a larger scale and concentrating the reaction products, or extracting them using a different solvent system, may thus enable their detection within a proton NMR spectrum as was the case for products G, H and I.

Alternatively, it could be that a side reaction is occurring during the reactions used for HPLC analysis and this is what in fact is causing the degradation of β-HPA (rather than the expected TK conversion). Control reactions were conducted in the absence of recombinant TK-containing crude enzyme extract and no appreciable degradation of β-HPA occurred under these conditions. However, controls were not performed in the absence of the aromatic aldehydes and it could be that some component of the crude enzyme extract (other than TK) is causing the β-HPA to break down. Further assays, with implementation of these additional controls, would need to be conducted in order to investigate this possibility. These findings concur with other work performed using this particular assay (Mitra, 1997) in which a loss of β-HPA not relating to TK transformation was observed, resulting in incomplete mass balances. Alternative reaction conditions were subsequently used by Mitra, conducting transformations in the

absence of buffers and utilising acid/alkali titration via a pH-stat. This resulted in a reduction in the undesired degradation of β-HPA and would be another avenue to explore during future work conducted in this area.

To demonstrate that the stereospecificity of the C-C bond forming step was unaffected by the alteration I189A a reaction was conducted using glycolaldehyde and producing L-erythrulose. The L-erythrulose was extracted with ethyl acetate and purified by flash chromatography (chloroform:methanol = 4:1). $[\alpha^{25}]_{578} = +11.4^{o}$ (Lit: $[\alpha^{25}]_{578} = +12^{o}$).

### 4.4.2 Mass spectrometry analysis

Since observable reaction products were found in the $^{1}H$ NMR spectra for TK reaction with 2-pyridinecarboxaldehyde, 3-pyridinecarboxaldehyde and 4-pyridinecarboxaldehyde, these samples were further analysed using mass spectrometry. Results can be seen in Figures 4.20 to 4.25. Peaks of interest are positioned at, or around, the following values:

79 Pyridine ring major ion;
107 Starting material pyridinecarboxaldehyde;
167 TK biotransformation product;
191 TK biotransformation product complexed to sodium.

Peaks closely bordering these (*eg.* those located at 78 and 80) can be assigned to structures in which the appropriate number of protons have been either lost or gained. It should be noted that in many instances, using these particular mass spectrometry techniques, the fragmentation patterns for a particular compound are extremely difficult to predict, consequently two separate spectra obtained from the same reaction sample may differ considerably in appearance. This is because the conditions under which the spectra are produced can never be identical from one run to the next. For this reason, these spectra were used more for qualitative rather than quantitative analysis. Complexation data shown in Table 4.4 was extrapolated from a typical spectrum for each of the compounds, this is given here to demonstrate the likelihood of a complex being formed between each of the product compounds and a sodium ion, and values are not intended to be absolute.

In addition, sometimes a particular compound may be present within the sample but only a very small peak at its mass value, or even no peak at all, can be seen in the mass spectrum. This type of situation could result if the compound has fragmented completely under the conditions used to obtain the spectrum. In these cases, the presence of the expected fragmentation ion peaks can be used to confirm the presence of the particular compound within the sample being analysed. This is highlighted by Figures 4.20 and 4.24, the mass spectra for 2-pyridinecarboxaldehyde and 4-pyridinecarboxaldehyde respectively, in both instances only small peaks at 107 (the mass of the carboxaldehyde compound) are observed, whereas much larger peaks can be seen at 78/79 which is the mass of the pyridine ring (a major ion in the fragmentation pattern for each compound).

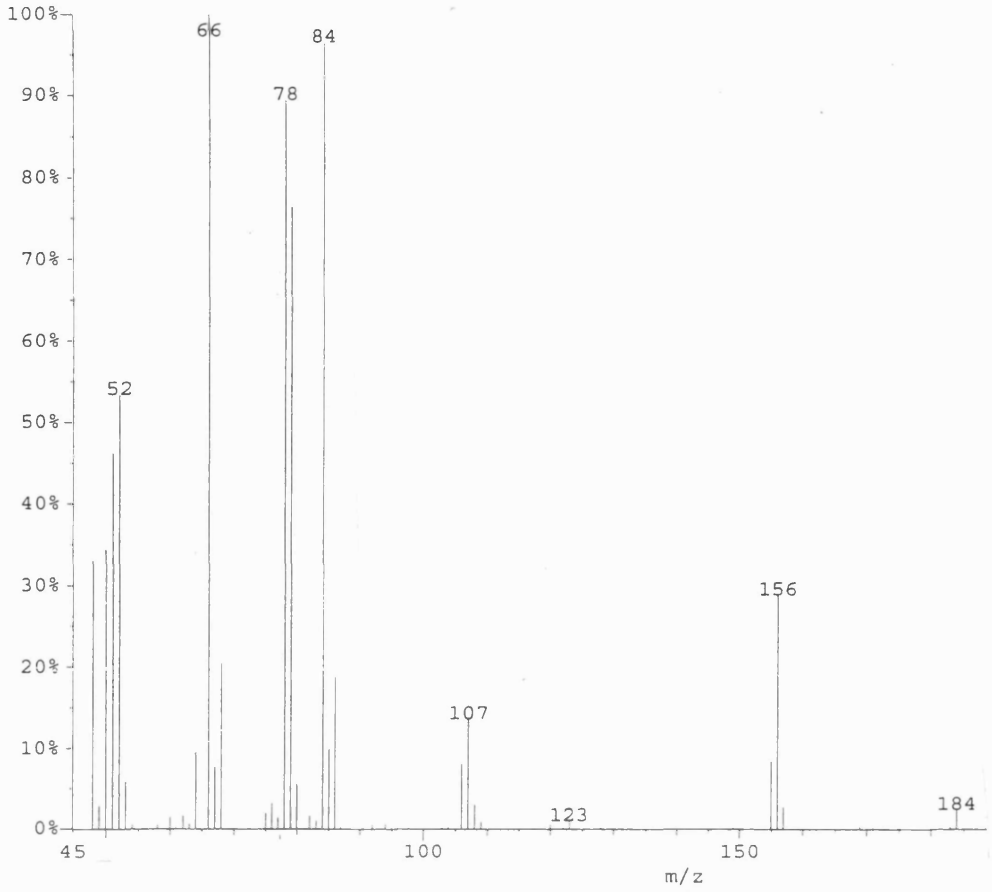



**Figure 4.20** *Mass spectrum for 2-pyridinecarboxaldehyde*

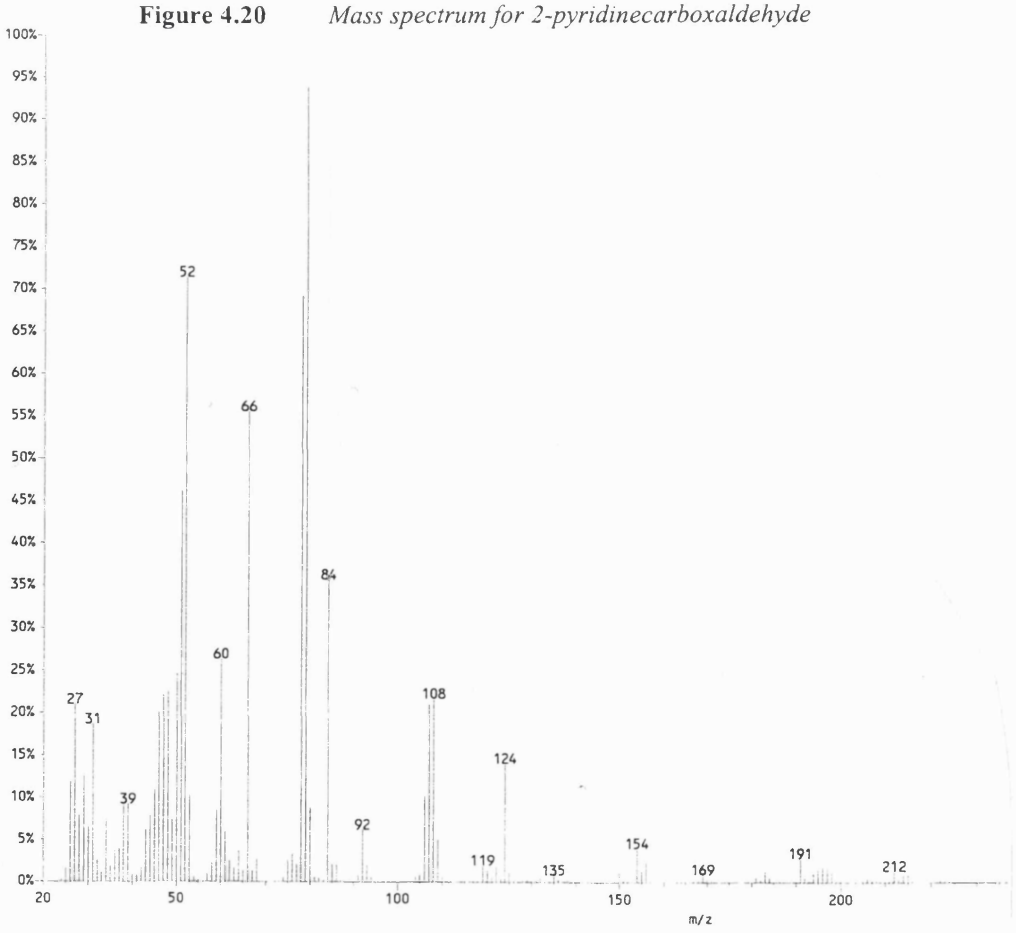



**Figure 4.21** *Mass spectrum for 2-pyridinecarboxaldehyde reaction*

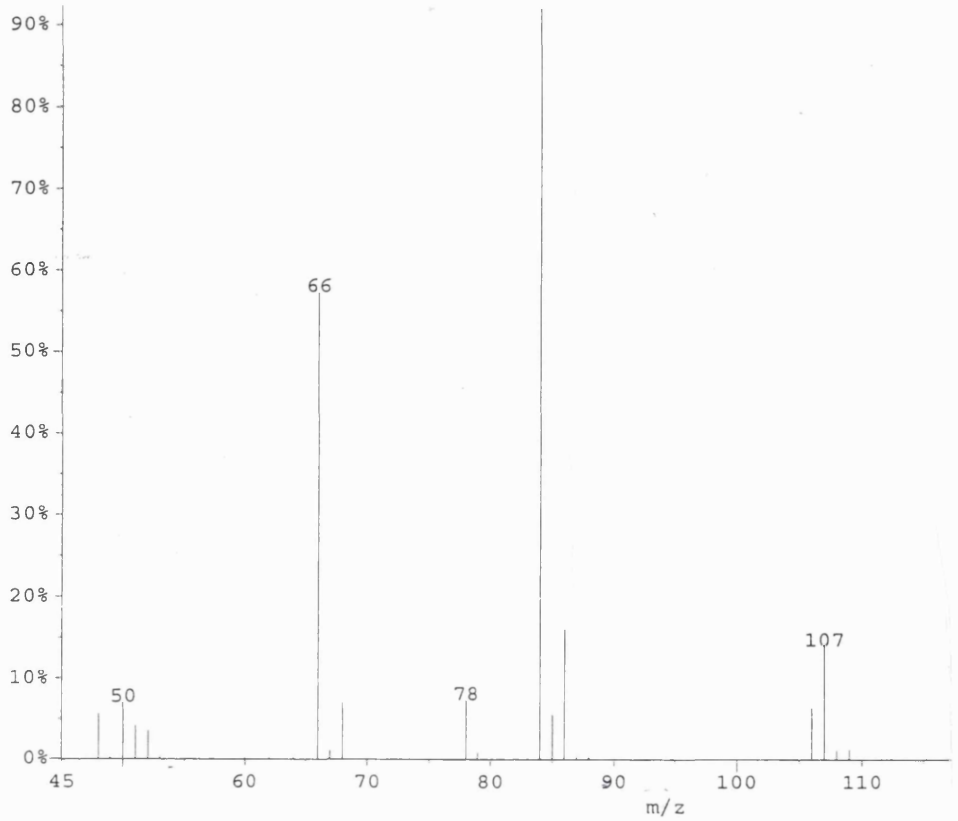



**Figure 4.22** *Mass spectrum for 3-pyridinecarboxaldehyde*

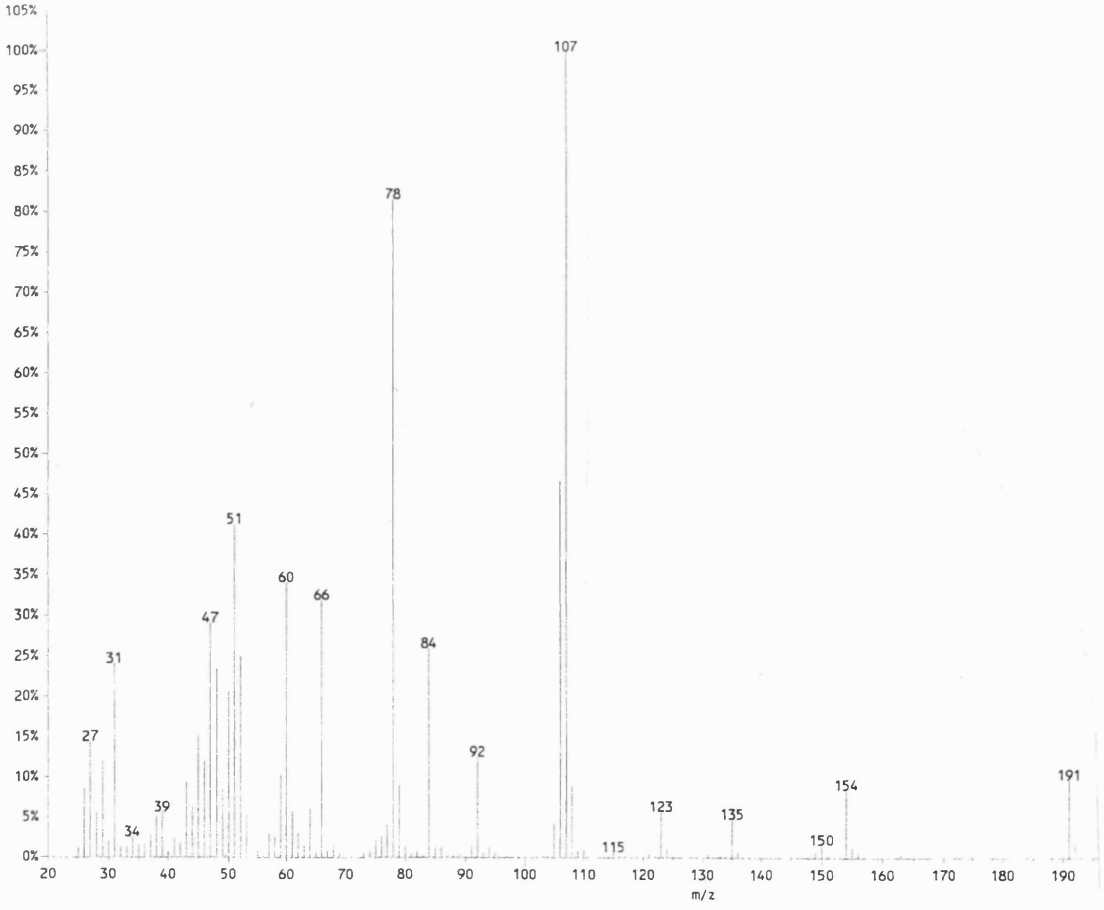



**Figure 4.23** *Mass spectrum for 3-pyridinecarboxaldehyde reaction*

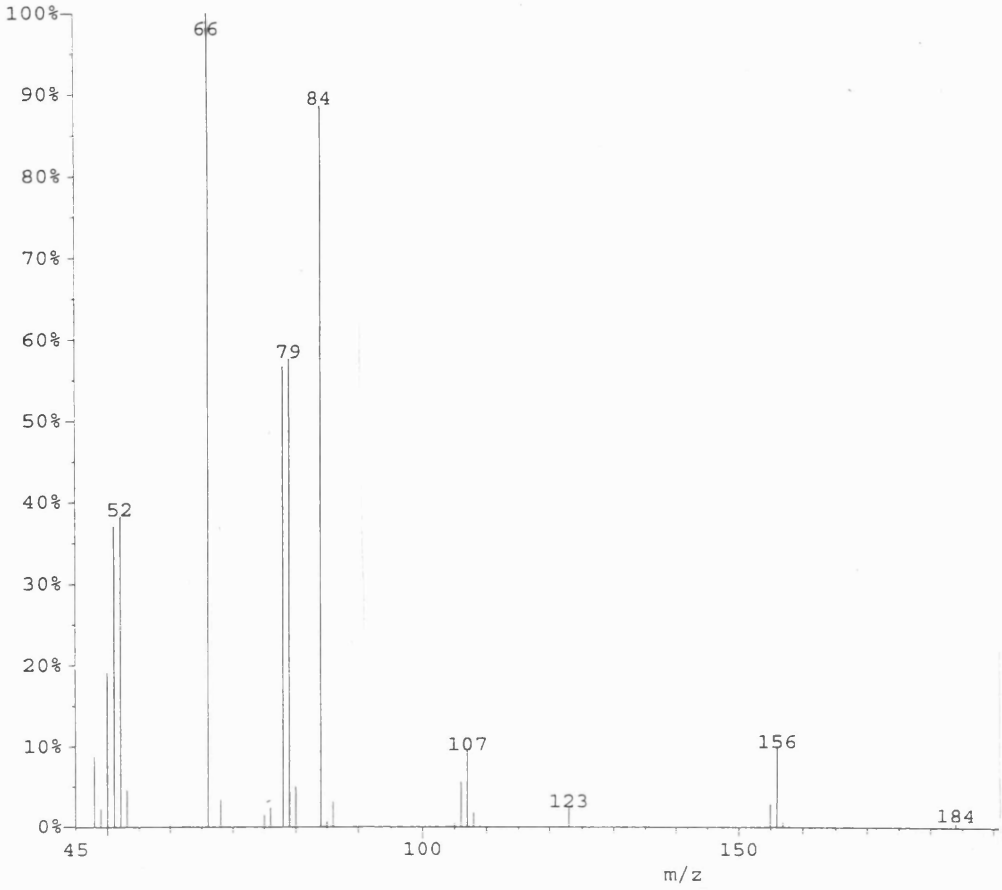



**Figure 4.24** *Mass spectrum for 4-pyridinecarboxaldehyde*

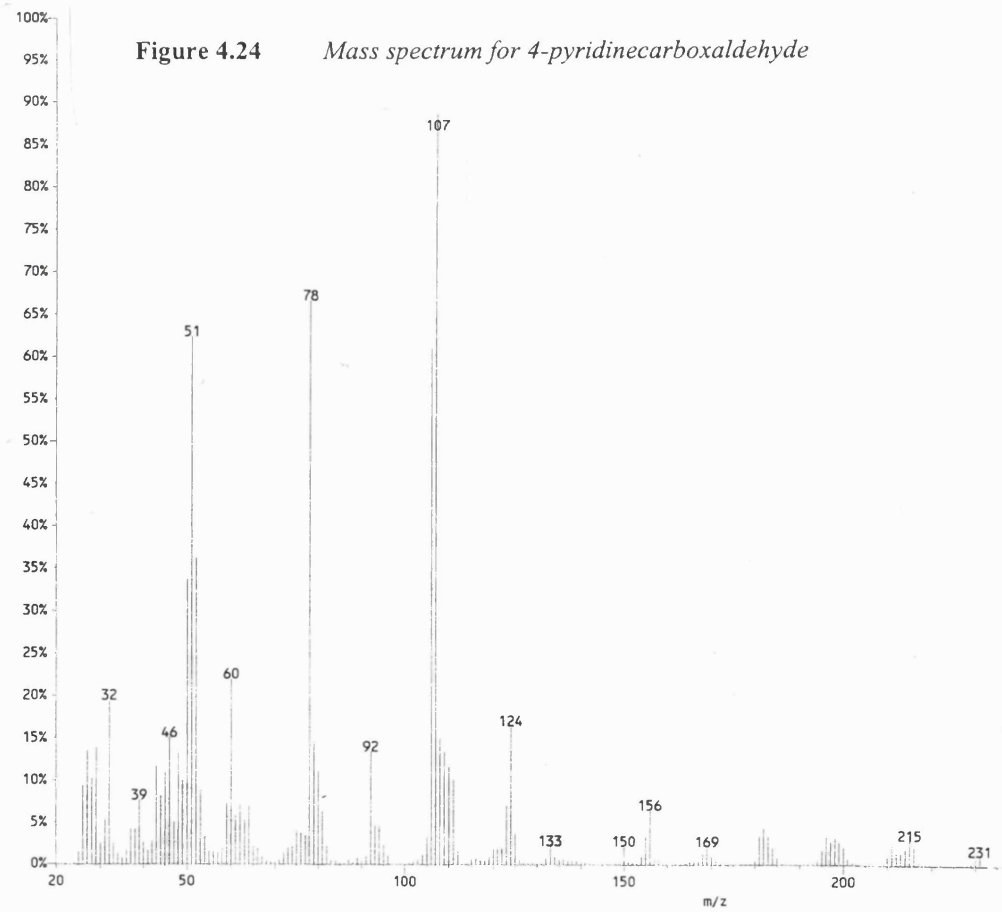



**Figure 4.25** *Mass spectrum for 4-pyridinecarboxaldehyde reaction*

| Aldehyde | Peak (% abundance) | |
|---|---|---|
| | 167 to 170 (uncomplexed product) | 191 to 193 (complexed product) |
| B | 168:- 0.21<br>169:- 0.94<br>170:- 0.26<br>**1.41** | 191:- 2.87<br>192:- 0.50<br>193:- 0.18<br>**3.55** |
| C | | 191:- 9.79<br>192:- 1.69<br>**11.48** |
| D | 167:- 0.43<br>168:-1.25<br>169:- 2.16<br>170:- 0.96<br>**4.80** | |

**Table 4.4** *Mass spectrometry peaks for complexed and uncomplexed TK reaction products*

Peaks located in the region of 191 can be attributed to complexation of the reaction products with sodium ions. Table 4.4 shows the sums of peaks believed to represent uncomplexed product (ca. 167) and complexed product (ca. 191). Comparisons of these two values enables a sample percentage ratio for complexed to uncomplexed material to be calculated and results are given in Table 4.5. Differences in these values can be explained by considering the possible structures for these complexes, as depicted in Figure 4.26.

| Reaction Product | % uncomplexed (X) | % complexed (Y) | X:Y |
|---|---|---|---|
| G | 28.43 | 71.57 | 1.00:2.52 |
| H | 0.00 | 100.00 | 0.00:1.00 |
| I | 100.00 | 0.00 | 1.00:0.00 |

**Table 4.5** *Proportion of each TK reaction product complexed to sodium*

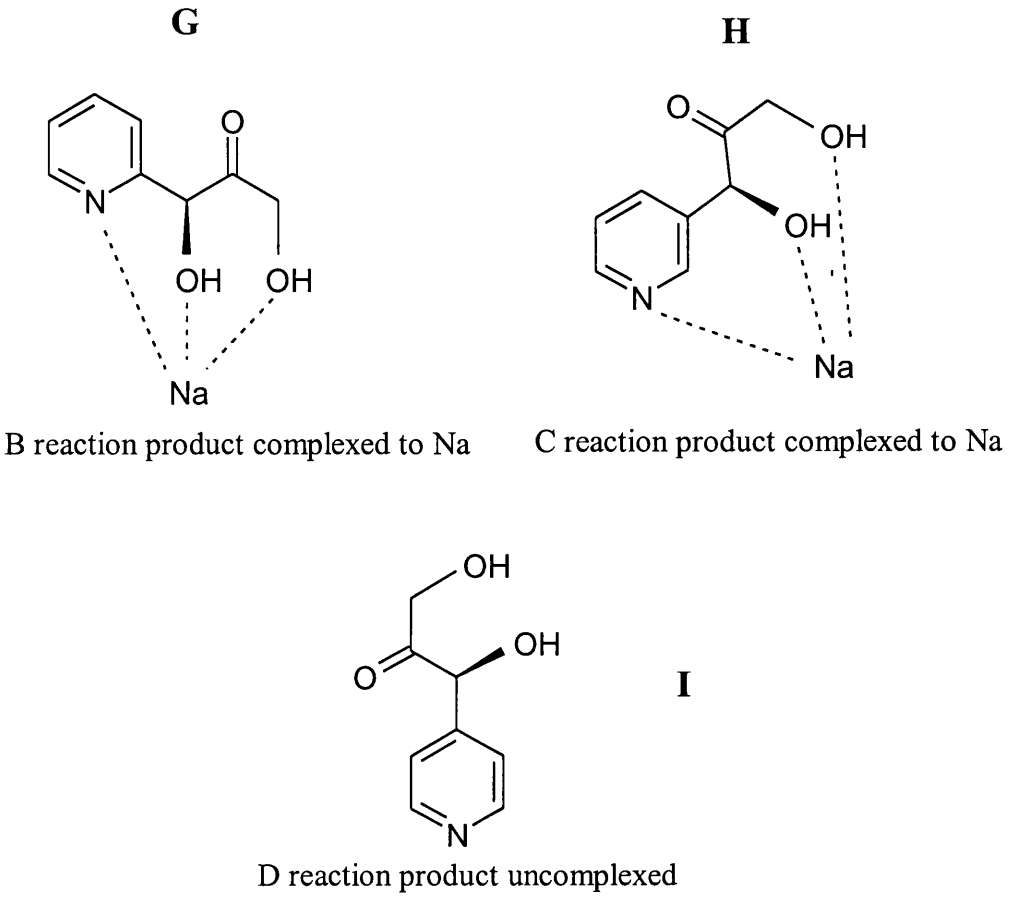



**Figure 4.26** *Complexing of TK reaction products with sodium ions*

Compounds G and H both exhibited complexation with sodium ions, however whereas G is present roughly 70% in the complexed form, H appears to be almost exclusively in its complexed state. These differences may be explained by considering the ring structures produced when the sodium-complex is formed. Complexed G contains one 5-membered ring and one 6-membered ring, whereas the corresponding H analogue possesses two 6-membered rings. The relative stabilities of these two ring forms are dictated by the respective Gibbs Free Energies of their formation.

Change in free energy is given by the equation:

$$\Delta G = \Delta H - T\Delta S \qquad \text{(kJ/mol)}$$

where: $\Delta H$ is the enthalpy change (kJ/mol)
$\Delta S$ is the entropy change (J/Kmol)
T is the temperature (K)

The more negative the value of $\Delta G$ the more stable the resultant structure, therefore a negative value for $\Delta H$ and a positive value for $T\Delta S$ are desired. Upon ring formation, entropy is lost since movement of bonds becomes more restricted (giving a negative value), however, enthalpy is released due to formation of bonding interactions (hence also giving a negative value).

The 5-membered ring will be more strained and so less enthalpy is released as a result of the weaker bonding interactions. In addition this size of ring is less flexible, resulting in a greater loss in entropy. Combined, these factors mean that overall this ring has a more positive $\Delta G$ value and hence is less stable than the 6-membered ring. Therefore, the complexed structure for G will be thermodynamically less favourable then that for H and so will be present in a lower abundance. For substrate I no complexed material was observed. In this case, the position of the side chain on the heterocyclic ring orients the hydroxyl groups away from the heteroatom and thus prevents their complexation with a sodium ion.

Preliminary mass spectrometry data was corroborated by accurate mass analysis and values for reaction of recombinant TK from pQR739 with 2-pyridinecarboxaldehyde are given in Table 4.6. Calculated values were derived using the elemental accurate masses as follows:

| Element | H | C | N | O | Na |
|---|---|---|---|---|---|
| Accurate mass | 1.0078 | 12.0000 | 14.0031 | 15.9949 | 22.9898 |

| Calculated value | Experimental value | Difference (ppm) |
|---|---|---|
| 107.0370 | 107.0367 | -3 |
| 167.0580 | 167.0587 | +7 |
| 191.0556 | 191.0568 | +12 |

**Table 4.6** *Accurate mass data for TK from pQR739/2-pyridinecarboxaldehyde reaction*

[Experimental values are correct to +/- 10ppm]

## 4.5 Further mutants

An additional set of five TK mutants were also engineered (these 5 TK mutants were constructed during this project), these constructs involved replacement of the pair of arginine residues (R358* and R520*) located at the active site entrance. Substitution of these residues with alanines should facilitate further the entry of bulky compounds into the active site funnel, and may also destroy part of the phosphate-binding site believed to be located in this region. Mutants were constructed according to the scheme depicted in Figure 4.27, using a QuikChange™ Site-Directed Mutagenesis Kit. The new TK constructs and their respective amino acid alterations, along with the base substitutions used to implement these changes, are given in Table 4.7. Also shown are the percentage efficiencies for each of the mutagenesis reactions. Numerous attempts to generate these constructs via the method described by Kunkel (Kunkel, 1985; Kunkel *et al.*, 1987) using ssDNA, and also via the intermediate production gdDNA (Kramer & Fritz, 1987), were unsuccessful.

| *E. coli* strain/construct | Alteration (s) | Base Changes | % Mutagenesis efficiency |
|---|---|---|---|
| JM107/pQR739 | I189A | ATC to GCC | |
| JM107/pQR785 | R358*A | CGT to GCT | 100 |
| JM107/pQR786 | R520*A | CGT to GCT | 100 |
| JM107/pQR787 | R358*A<br>R520*A | CGT to GCT<br>CGT to GCT | 75 |
| JM107/pQR788 | I189A<br>R358*A | ATC to GCC<br>CGT to GCT | 100 |
| JM107/pQR789 | I189A<br>R520*A | ATC to GCC<br>CGT to GCT | 75 |

**Table 4.7** *Additional TK mutants*

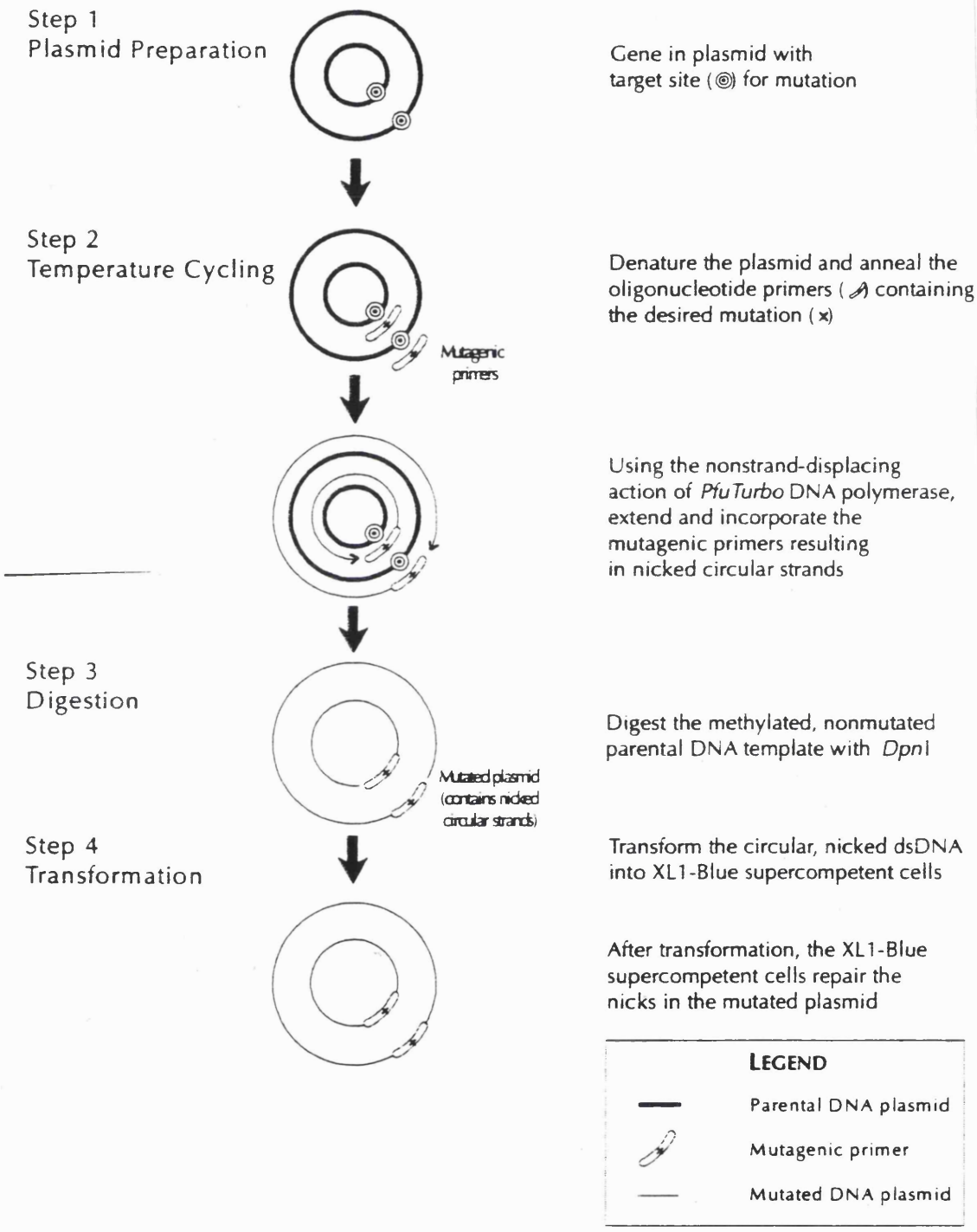



**Figure 4.27** *Scheme for mutant construction (Taken from QuikChange™ Site-Directed Mutagenesis Kit Instruction Manual)*

Over-expression of these vectors was performed in *E. coli* stain JM107 using the TK promoter. Initial characterisation was again conducted using the linked-enzyme assay (2.2.1). All linked enzyme assays were performed in triplicate, alongside assays for reactions conducted with TK from pQR706/JM107 and that from JM107 cells to be used as standards by which to evaluate the new mutants. Specific activities of these modified enzymes were calculated and values, along with those for the native protein (from pQR706) and the previously characterised enzyme from pQR739, are given in Table 4.8. Examination of this data shows that all mutant TKs have reduced activity towards the native substrates relative to the wild-type protein, this would be expected since the active site funnel for each of these constructs deviates from the ideal conformation for transformation of these compounds.

| *E. coli* strain/construct | Activity (U/ml) | Concentration (mg/ml) | Specific activity (U/mg) | Error (%) |
|---|---|---|---|---|
| JM107/pQR706 | 4.330 | 0.62 | 6.98 | +/- 4.6 |
| JM107/pQR739 | 0.635 | 0.52 | 1.22 | +/- 1.5 |
| JM107/pQR785 | 0.507 | 0.57 | 0.89 | +/- 1.2 |
| JM107/pQR786 | 0.444 | 0.58 | 0.77 | +/- 1.3 |
| JM107/pQR787 | 0.264 | 0.60 | 0.44 | +/- 0.9 |
| JM107/pQR788 | 0.381 | 0.54 | 0.71 | +/- 1.3 |
| JM107/pQR789 | 0.317 | 0.53 | 0.60 | +/- 1.1 |

**Table 4.8** *Specific activities for TK mutants*

For all of these new constructs the observed decline in activity may be attributed, in part, to the destruction of the phosphate-binding site located at the active site entrance, this occurs as the two positively charged arginine residues are replaced with alanines. This is true particularly for TK from pQR787 in which the phosphate-binding site has been completely removed. For all five constructs these alterations may result in a loss of bonding interactions with the phosphate groups of the sugars, thus leading to less efficient binding of these substrates. Additionally, for enzymes from pQR788 and pQR789 the replacement of residue I189 with an alanine will contribute to a less efficient binding of the $Mg^{2+}$ ions, leading to a more unstable protein dimer with a consequent reduction in activity.

## 4.6 Discussion

This chapter outlined the work conducted towards the design, generation and characterisation of a range of TK mutants. TK, whose function *in vivo* is the degradation of carbohydrates, has the potential to be a valuable biocatalyst for C-C bond formation. The enzyme has a fairly relaxed specificity for its aldehyde component (Hobbs *et al.*, 1994), and this, combined with the high stereoselectivity for its C-C bond-forming step, enables TK to potentially be an extremely useful catalyst for asymmetric synthesis. The mutants described in this chapter were variants of the enzyme incorporating expansion of the active site funnel, designed in such a way as to accommodate more bulky substrates. Amino acid substitutions were made at locations in which residues were believed to play a key role in substrate and/or cofactor binding.

Preliminary characterisations, conducted using a linked-enzyme assay, showed that all mutant TKs exhibited reduced activity towards the native sugar substrates relative to that of the wild-type protein. This would be expected, since these amino acid substitutions cause the active site structure to deviate from its optimal conformation for these compounds. Initial investigations into the synthetic potential of the modified *E. coli* TKs were conducted using β-HPA as the ketol donor, employing a range of water-soluble aromatic compounds as aldehyde acceptors. In the first instance reactions were monitored using HPLC techniques (Mitra & Woodley, 1996), following the loss of β-HPA due to enzymatic conversion. Effective mutant/aldehyde combinations were then employed for larger scale biotransformations with subsequent extraction and structural characterisation of reaction products.

A number of problems were encountered when using these HPLC assays. For example, reactions of TK produced from pQR739 with pyridoxal, 2-furancarboxaldehyde, and 2-pyrrolecarboxaldehyde afforded no observable reaction products upon NMR analysis. However, for all three of these compounds some degree of β-HPA loss was found during the HPLC assays of their reactions with TK from pQR739. It may simply be the case that insufficient product material was produced, under the particular reaction conditions employed, to enable their detection within the limits of the NMR techniques used. Alternatively it could be that the products generated from these reactions did not extract efficiently into the organic solvent. It would therefore be useful, as part of future experiments in this area, to repeat these biotransformations, and to investigate the

possible effectiveness of alternative solvent systems for the extraction of these particular transformation products. There is also the possibility that other constituents of the *E. coli* crude enzyme extracts employed during these biotransformations were reducing the effectiveness of the extraction procedure. It may be useful, should this be the case, to devise a solid phase chromatographic technique for the purification of these reaction products, thus providing pure material which would be available for subsequent structural characterisation experiments.

Alternatively, it could have been that a side reaction was occurring during the biotransformations used for HPLC analysis, and that this was what in fact was causing the degradation of β-HPA (rather than the expected TK conversion). Control reactions were not performed in the absence of the aromatic aldehydes (but with all other reaction components present) and it may have been that some other compound or enzyme within the crude enzyme extract was causing the β-HPA to break down. Additional reactions, with implementation of these particular controls, would need to be conducted in order to eliminate this possibility. These findings support the idea that it may in fact be worthwhile spending additional time devising an effective purification method for TK, in order to reduce the probability of such unwanted side reactions occurring and so distorting the data obtained from HPLC analysis. Nevertheless, despite the problems encountered, this technique was still a useful screening tool for highlighting interesting mutant/aldehyde combinations to be used for more detailed investigations.

From the results of the initial HPLC screening, the mutant TK produced from pQR739/JM107 appeared to exhibit enhanced reactivity towards a selection of aromatic aldehydes. This mutant has replacement of amino acid residue I189 with an alanine, thus incorporating a hole within the active site funnel. This cavity appears to facilitate the entry of more bulky aldehyde compounds, and especial reactivity was found in the case of the pyridinecarboxaldehyde substrates. Reaction products from these successful biotransformations were extracted and structurally characterised, using NMR and mass spectrometry techniques to confirm that the expected TK conversion had in fact taken place.

The results obtained were very interesting since, in contrast to the TKs from spinach and yeast (Demuynck *et al.*, 1991), no conversion of aromatic aldehydes had previously

been detected with *E. coli* TK (Sprenger *et al.*, 1995; Hobbs *et al.*, 1994). Furthermore, no structures of the reaction products from conversions of aromatic and heterocyclic substrates, with TK from any source, had prior to this project been shown (Schorken & Sprenger, 1998). The next step for work in this area would be to attempt to scale up these reactions in order to achieve conversion of these compounds on a preparative scale.

More detailed analysis, involving the transformation of a selection of bulky aromatic/heterocyclic compounds, is required in order to fully characterise the new TK mutants generated during this project. Although the linked-enzyme assay data provides an insight into the differing behaviour of these constructs towards their native sugar substrates, direct examination using the more bulky aromatic aldehydes is required. As before, initial screening using HPLC assays is needed in order to highlight potentially useful mutant/aldehyde combinations. Products from these transformations can then be subsequently extracted from the reaction mixture and structurally characterised. In addition the creation of an HPLC assay, allowing the separation of the product peaks resulting from reaction of TK on β-HPA and the novel aldehydes would be preferable, as this would show at an early stage that the desired aldehyde transformation is taking place.

Finally, work is required to generate variants of TK having increased tolerance towards organic solvents. This would enable investigations to be made regarding TK activity with a wider range of substrates, for example those containing a thiophene or an indole ring, which are not soluble in aqueous solution. Some work was carried out during this project in an effort to transform these types of compound, using a two-phase biotransformation system, but no conclusive results were obtained.

# Chapter Five: Modification of TK stereoselectivity

# CHAPTER FIVE: Modification of TK stereoselectivity

## 5.1 Introduction

Structural requirements for biological activity often imply the presence of several chiral centres within a drug. Many of these compounds are marketed as racemates, however the enantiomers must, particularly from the biological point of view, be regarded as individual substances. As regulating factors concerning the production and marketing of pharmaceuticals become more severe, the importance of the manufacture of homochiral products continues to grow. As a result this has lead to an increased move by pharmaceutical companies towards the production of single isomers. Biocatalysts frequently carry out to high enantiomeric purity in a single step, a transformation that may require a number of chemical stages and due to the increasing demand for enantiomerically pure pharmaceutical compounds, the benefits of enzymes are enormous since they can be used to set the stereochemistry of key intermediates.

In classical organic chemistry many types of carbon-carbon bond-forming reactions have been developed. However, few of these conversions are selective towards their reagents and have the ability to dictate the stereochemistry of the products. Because enzymes may perform either or both of these tasks, they provide an important alternative to these conventional methods. Since TK has such relaxed specificity for its aldehyde component, combined with a high stereoselectivity for its C-C bond-forming reaction step, it is an extremely useful catalyst for asymmetric synthesis. However, one restriction in its applications is that in the biocatalytic formation of novel substances with an α-hydroxyl group, TK has a strict requirement for the hydroxyl moiety to be in the D-configuration.

It would therefore be extremely useful to have a version of the enzyme tailored for biotransformation of the opposite enantiomer. For this purpose a series of TK variants were constructed in which residues believed to be dictating the stereoselectivity of the enzyme were replaced. These modified enzymes were characterised through transformation of both enantiomers of the sugar glyceraldehyde.

## 5.2 Initial TK mutants

Molecular modelling studies have highlighted residues believed to be acting to orient the substrate, allowing exclusive binding of molecules with an α-hydroxyl group of the correct stereochemistry. H258 sits midway along the active centre pointing towards residue F434* located on the opposite face of the funnel. It is believed that the positively charged side chain of H258 is involved, through hydrogen bonding interactions, in the sharing of pairs of electrons with the α-hydroxyl group of the aldehyde acceptor. Together with the bulky hydrophobic side chain of F434*, this may be an important factor in determining the stereoselectivity of TK. To investigate the precise role of these residues, and to attempt to produce an enzyme exhibiting the alternative specificity, these amino acids were altered using site-directed mutagenesis. It was expected that a mutant with reversal of these residues would exhibit the opposite specificity to that of the native enzyme, whereas a variant in which they have been replaced with alanines should have complete relaxation of the requirement for D-enantiomeric compounds.

Initial TK constructs analysed were those previously described in Chapter Four (section 4.2). Preliminary analytical reactions involving glycolaldehyde (section 5.3) were performed in order to obtain relative rate data for TK conversion of this model substrate. This was followed by transformation of D-glyceraldehyde and DL-glyceraldehyde (section 5.4) in order to indentify enzymes exhibiting altered stereoselectivities. Since the work conducted in Chapter Four indicated that β-HPA may have been degraded by side reactions occurring within the reaction mixture, transformations here were monitored by measuring the loss of glycolaldehyde due to TK conversion.

## 5.3 Reaction with glycolaldehyde

To screen the relative enzymatic activities towards the model substrate glycolaldehyde all eight TK constructs were used to transform this compound, as depicted in the reaction scheme in Figure 5.1. This substrate was nominated as the standard since it exhibited the fastest $V_{rel}$ value, as shown by the rate data given in Figure 1.10. Biotransformations were performed (2.2.19) at substrate concentrations of 100mM (both β-HPA and aldehyde), with TK produced from O/N culture (3ml), in a final reaction volume of 10ml. Reactions were allowed to run for 1h and continual

monitoring was conducted using TLC (ethyl acetate:methanol = 95:5 [$R_f$ values: β-HPA, 0.88; glycolaldehyde, 0.71; L-erythrulose, 0.75]).

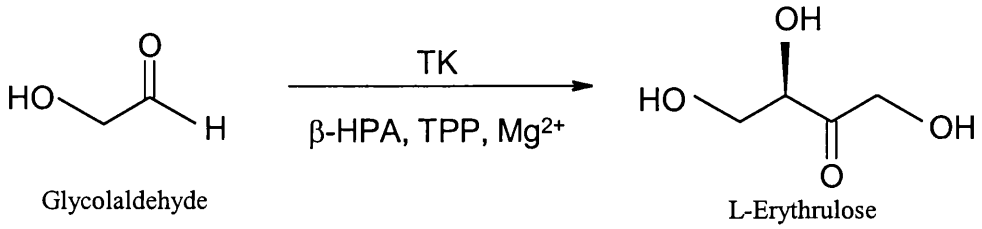



**Figure 5.1** *TK reaction of β-HPA with glycolaldehyde*

The isocratic HPLC system used was that described in section 4.3. Aliquots were taken every minute for the first ten minutes and then after thirty and sixty minutes of TK reaction. Dilution of reaction samples to the required concentration for the assay (ca. 20mM) in the acidic mobile phase (pH 2.3) simultaneously terminated the TK reaction. The substrates were stable under these acidic conditions for the time required for analysis. Samples were filtered through 0.45μm filters prior to injection (10μl). Retention times for each of the compounds under these conditions were: β-HPA, 21.32 minutes; L-erythrulose, 30.85 minutes; glycolaldehyde, 32.14 minutes.

Standard curves for β-HPA, glycolaldehyde (GA) and L-erythrulose (L-E) were plotted (see Figure 5.2) to allow HPLC integrated peak areas to be converted into substrate concentration values (mM). This enabled a direct concentration loss (gain for the reaction product L-erythrulose) of these substrates, due to enzymatic conversion, to be calculated. For each of the concentrations shown (mM) the relevant quantity of β-HPA, glycolaldehyde, or L-erythrulose was run through the HPLC system three times. However, as is evident from Figure 5.2, the relationship between peak area and substrate concentration on these standard curves were not found to be completely linear (this was also true for the standard curves of D-glyceraldehyde and DL-glyceraldehyde [see Figure 5.12]). For this reason there was some doubt regarding the overall accuracy of the substrate concentrations calculated using these standards. It might be the case that using a refractive index measurement is not in fact the best method for detection of

these particular types of compound and it may be beneficial in future work to investigate the possibility of an alternative detection technique such as UV absorption.

Since during the reactions conducted in Chapter Four it was hinted that the β-HPA may have been broken down by non-TK transformations, it was elected to use changes in glycolaldehyde concentration for calculation of rate data here. However as can be seen from the following data charts, the loss of the other reagent, β-HPA, closely mirrored the decline in glycolaldehyde in each case. Control reactions, measuring any loss of glycolaldehyde under the conditions of this assay, but in the absence of crude enzyme extract (with all other reaction components present), were performed. It was found that no appreciable degradation of this compound occurred under these conditions and therefore calculated values were considered to be reasonable within the limits given, although analogous controls were not performed in the absence of β-HPA. It may be that some other constituent of the crude enzyme extract (in addition to the recombinant TK) is causing the glycolaldehyde to break down and is thus distorting the data obtained from these experiments.

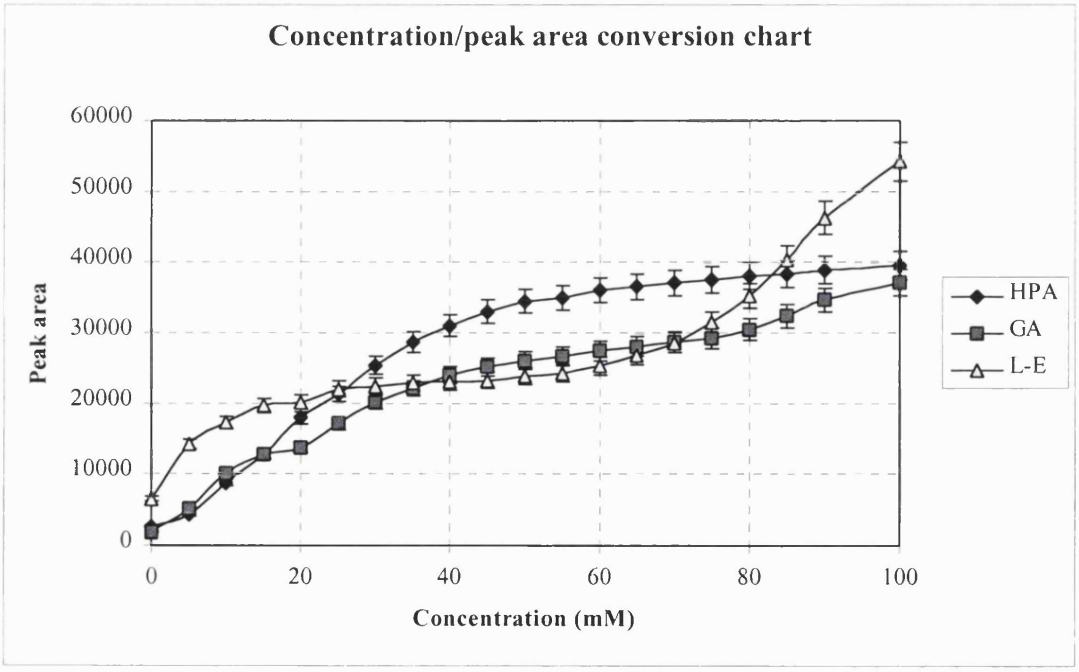



**Figure 5.2** *Standard HPLC curve for β-HPA/glycolaldehyde/L-erythrulose*

The twin column HPLC system detected the β-HPA, glycolaldehyde and L-erythrulose present in the reaction mixture and transformations were followed by measurement of the loss of glycolaldehyde (see Figures 5.3 through 5.10). As can be seen from these eight charts, a wide range of activity levels for these TK mutants with glycolaldehyde were observed.

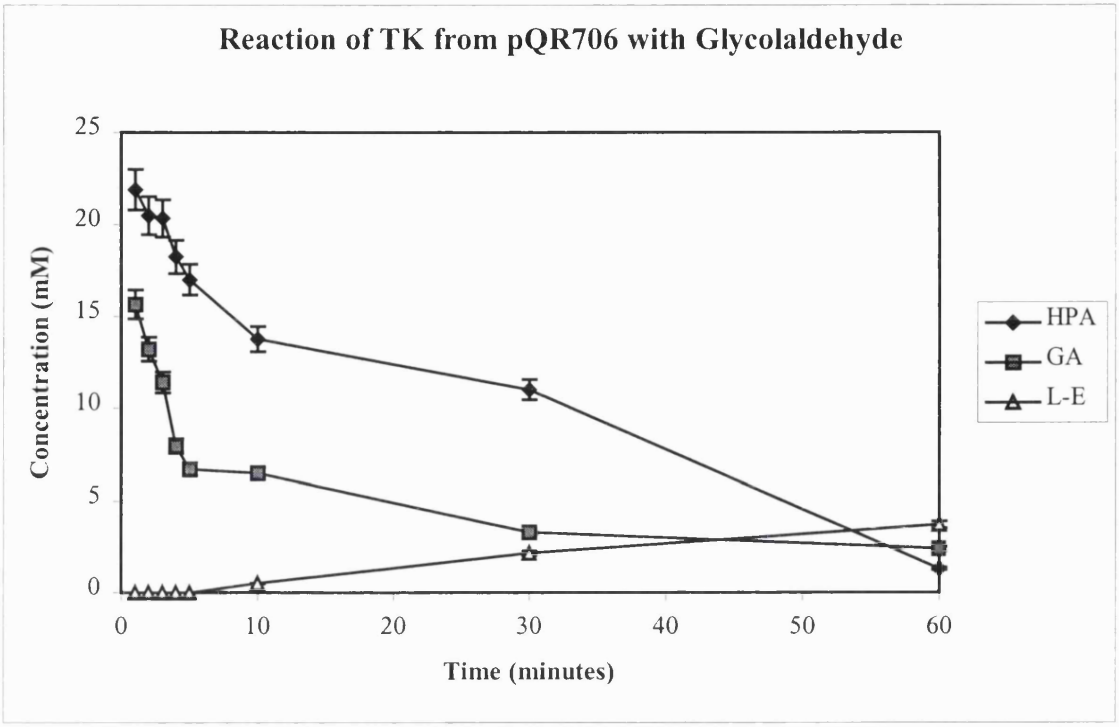



**Figure 5.3** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for wtp reaction*

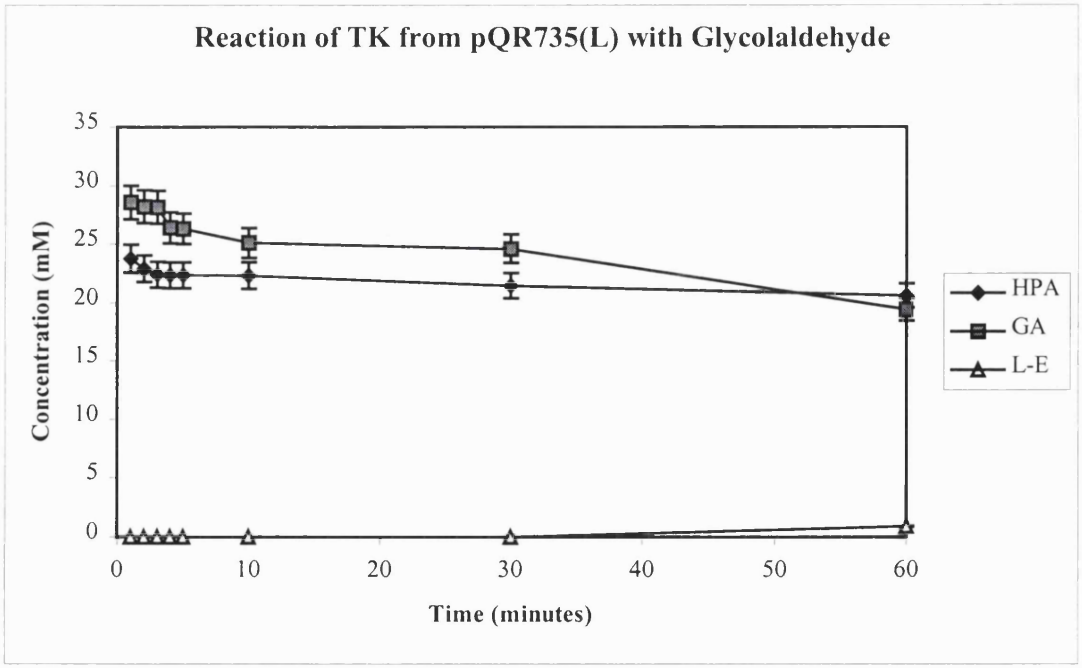



**Figure 5.4** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for S385*L reaction*

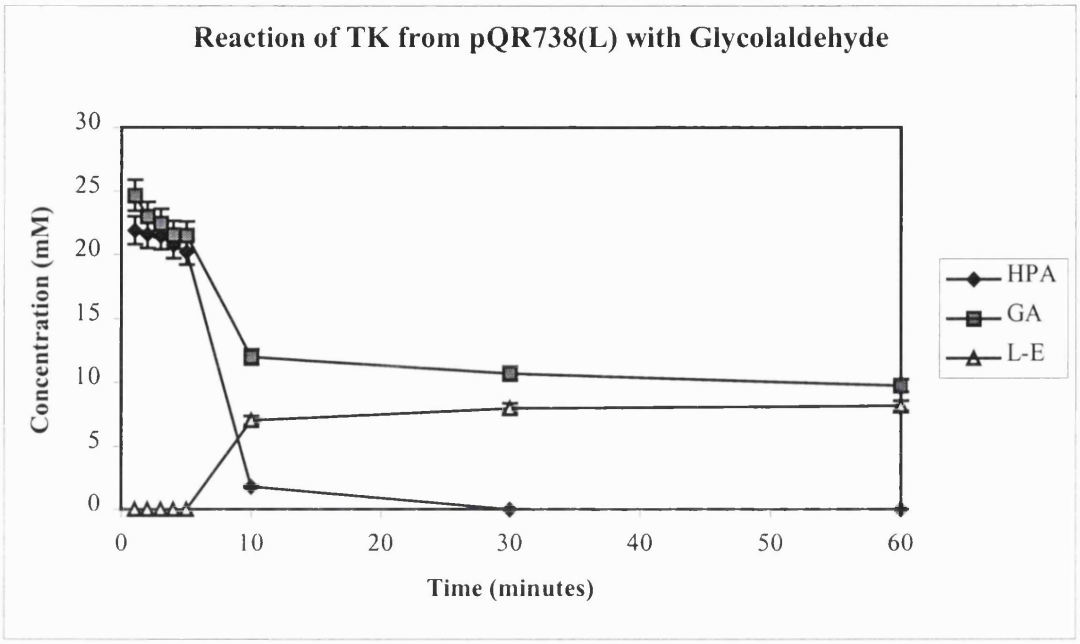



**Figure 5.5** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for H461*L reaction*

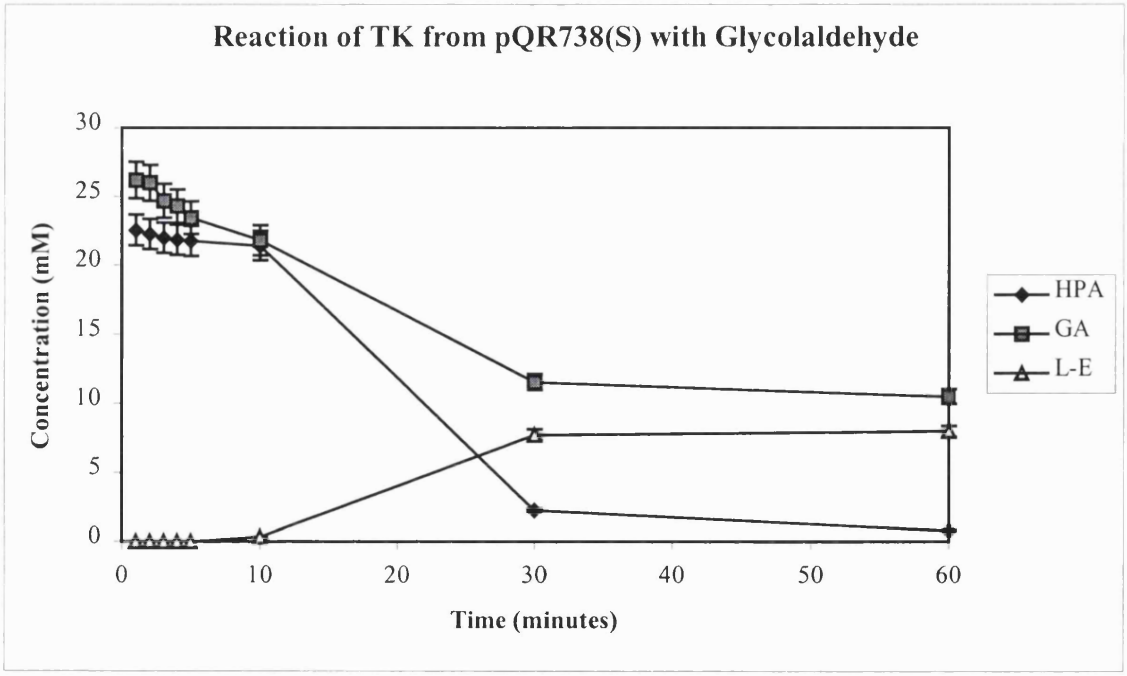



**Figure 5.6** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for H461*S reaction*

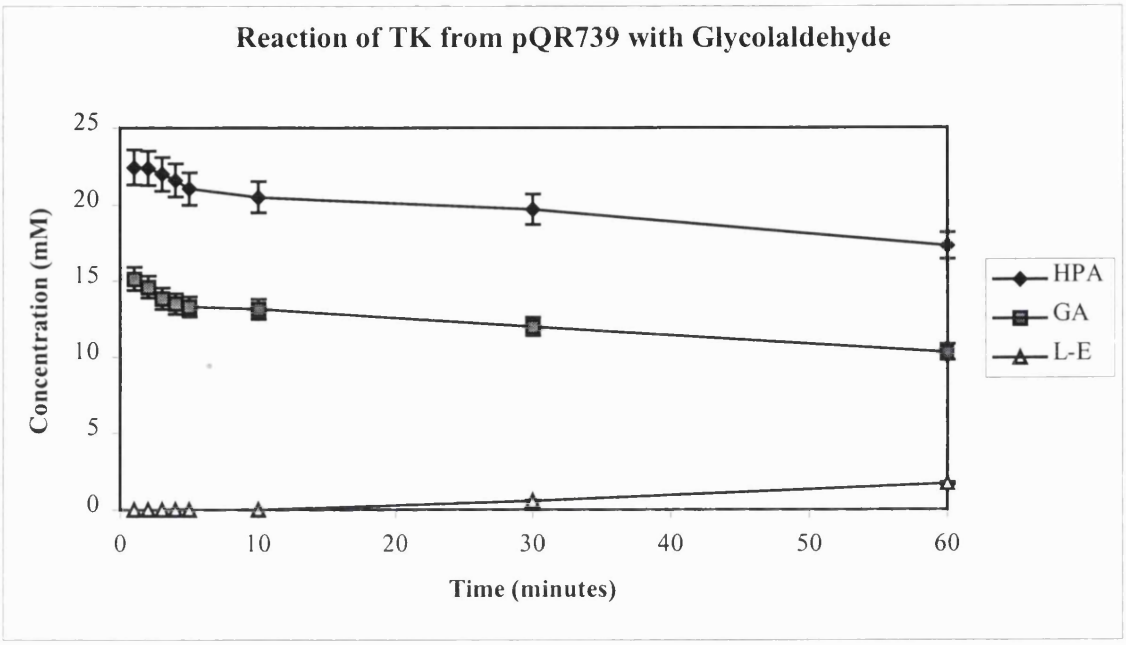



**Figure 5.7** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for I189A reaction*

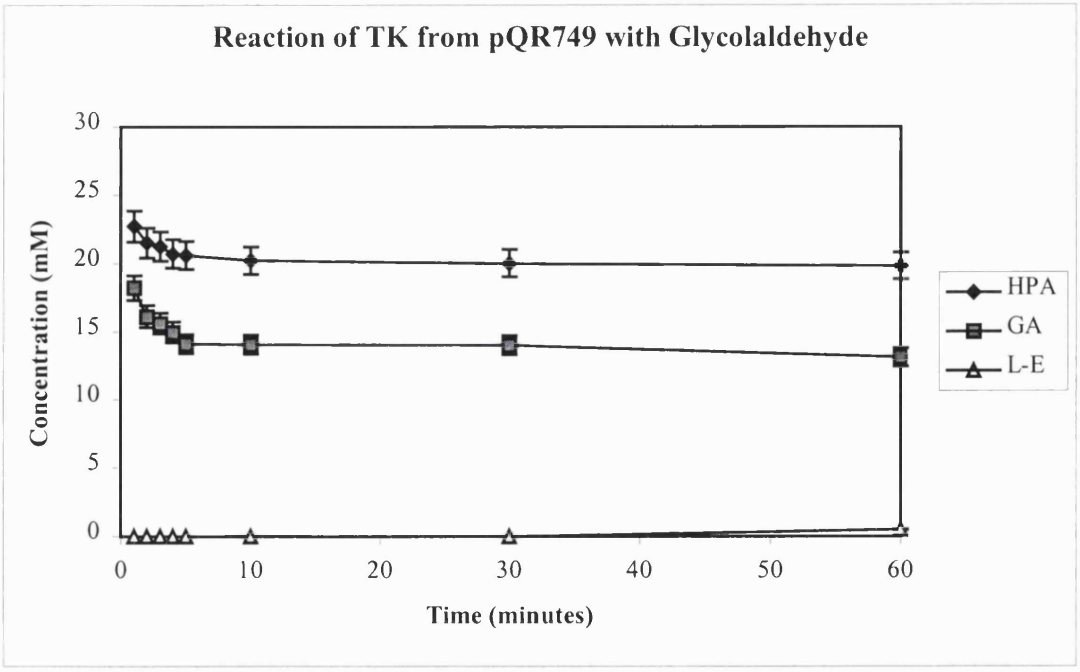



**Figure 5.8** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for I189V reaction*

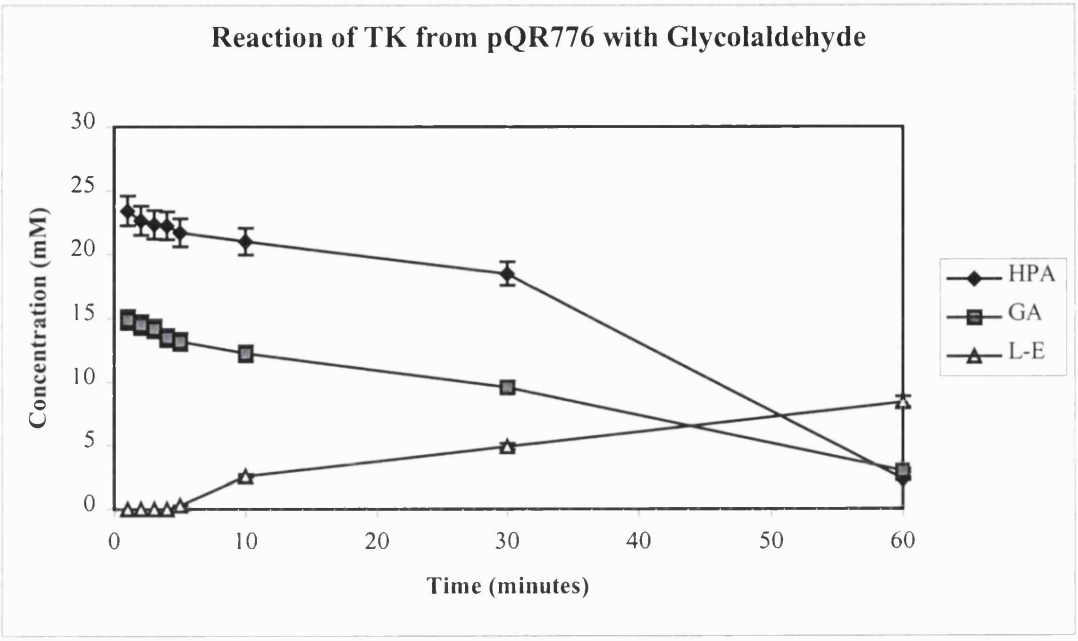



**Figure 5.9** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for H258A reaction*

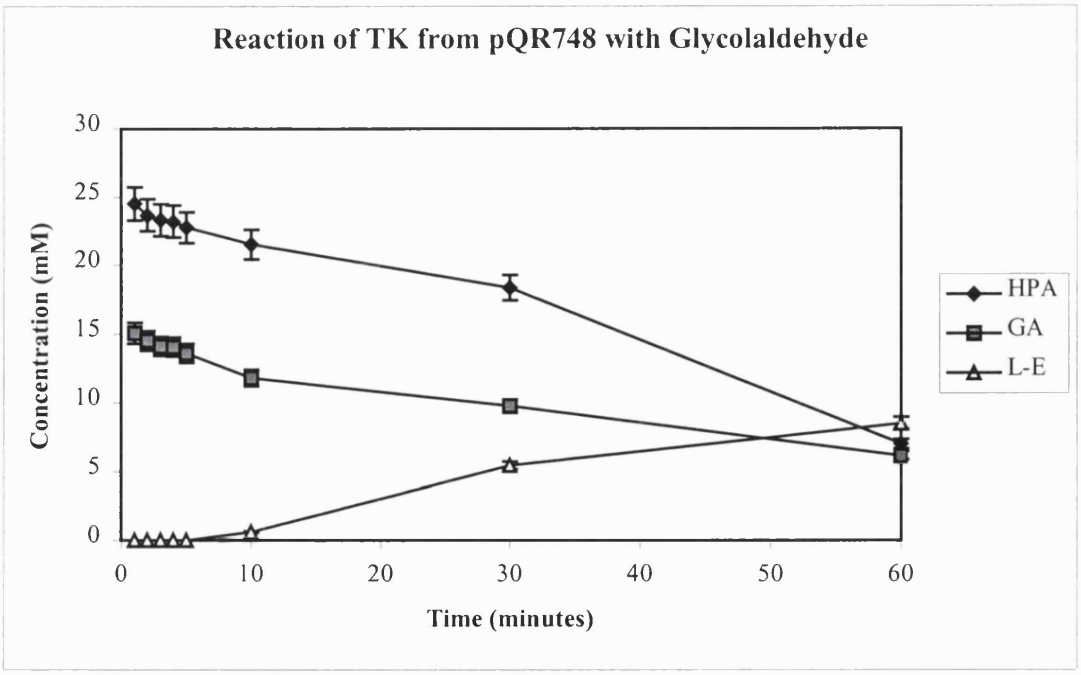



**Figure 5.10** *Loss in β-HPA/glycolaldehyde and gain in L-erythrulose for H258D reaction*

The losses of glycolaldehyde were used to calculate specific activities for the TK conversions and values are given in Table 5.1 below, data has also been converted into specific activities for each of the seven TK mutants compared to that for reaction of the native enzyme produced from pQR706. Examination of these figures shows that all mutant TKs, when compared to the wild-type protein, exhibit reduced activity towards this model sugar substrate. This can be explained since glycolaldehyde resembles the native sugar substrates of TK, the amino acid substitutions present in these constructs cause the active site structure to deviate from its optimal conformation for transformation of these compounds.

| *E. coli* strain/construct | Specific Activity (U/mg) | Error (%) | Specific activity relative to native enzyme |
|---|---|---|---|
| JM107/pQR706 | 6.56 | +/- 5.5 | 1.00 |
| Amber-LEU/pQR735 | 1.22 | +/- 3.4 | 0.19 |
| Amber-LEU/pQR738 | 2.92 | +/- 4.8 | 0.45 |
| Amber-SER/pQR738 | 0.80 | +/- 2.9 | 0.12 |
| JM107/pQR739 | 1.86 | +/- 3.6 | 0.28 |
| JM107/pQR749 | 2.40 | +/- 3.8 | 0.37 |
| JM107/pQR776 | 3.58 | +/- 4.7 | 0.54 |
| JM107/pQR748 | 3.21 | +/- 4.4 | 0.49 |

**Table 5.1** *TK specific activities for transformation with glycolaldehyde**

* Calculated using disappearance of glycolaldehyde data

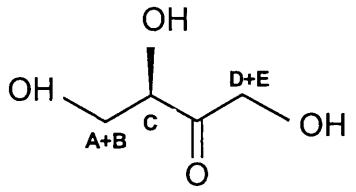
For the alterations S385*L and H461*L, introduction of the hydrophobic side chain of leucine close to the entrance of the active centre may result in unfavourable interactions with this substrate as it enters the active site funnel. In the case of the enzyme containing the alteration H461*S, since there is no phosphate group present in this substrate, the potential gain in hydrogen-bonding interactions which accompany the introduction of the serine residue for the native sugar substrates is not present here, in addition the serine hydroxyl group may act to repel the hydroxyl groups of glycolaldehyde. For the alteration I189A/V, as explained previously this residue is more significantly involved with cofactor binding than with that of the substrate. Its replacement will therefore result in less efficient binding of $Mg^{2+}$ ions and hence less efficient binding of TPP, so giving a less stable enzyme dimer. Substitution of H258 for alanine or aspartic acid has the least effect on TK reactivity with glycolaldehyde. Since this histidine residue is believed to be involved in stereospecific hydrogen-bonding interactions with the α-hydroxyl moiety, which is not attached to a chiral

centre in this case, this alteration does not have such a large deleterious effect on the transformation efficiency for this compound.

To confirm that L-erythrulose was indeed the sole product from these transformations, reaction mixtures were extracted with ethyl acetate and products were structurally characterised using NMR. $^{1}H$ NMR data for the L-erythrulose produced from TK conversion of glycolaldehyde was as follows:

$^{1}H$ NMR (400MHz) $D_2O$ δ: 3.72 (d, 2H, J=6.2)$^{A+B}$, 4.31 (t, 1H, J=6.0)$^{C}$, 4.39 (d, 1H, J=-18,)$^{D}$, 4.49 (d, 1H, J=-18)$^{E}$.

These peaks can be assigned as shown below:




A significant concern regarding the data obtained from these assays related to the lack of mass balances observed for these transformations. Values for the losses of β-HPA and glycolaldehyde, and the corresponding gain in L-erythrulose, should be stoichiometric. As seen in figures 5.3 to 5.10 this was not found strictly to be the case. Also, since the reaction samples were diluted down to 20mM in the acidic mobile phase it would be expected that in the graphs shown in Figures 5.3 to 5.10 the concentrations of the two starting materials (β-HPA and glycolaldehyde) would be at 20mM when the reaction time was equal to zero. The discrepancies observed from these graphs may be attributable to errors resulting when converting peak areas to concentrations using the standard curves (problems regarding these standards were highlighted earlier in this section). In addition pipetting errors, when diluting samples and when loading them onto the HPLC system, may have lead to deviations from the expected values.

Alternatively it could be that some other component of the crude enzyme extract is interfering with the detection of these compounds, this suggests that either purification of enzymes to be used for this type of analysis, or the employment of an alternative detection method, would probably produce more reliable data. Another possible

contributor to the lack of mass balances found during the assays used to follow TK conversion of β-HPA and glycolaldehyde relates to the fact that commercial glycolaldehyde is purchased as a crystalline dimer, with the dimer subsequently dissociating into the reactive monomers within the biotransformation mixture. It may be that the dimer is not being fully dissociated under the conditions employed here, and that any residual dimer present in the solution is interfering with, or distorting, the detection of the glycolaldehyde monomer.

Although in most cases the loss of β-HPA was closely mirrored by the corresponding loss of glycolaldehyde, and periods of greater β-HPA/glycolaldehyde losses were generally accompanied by some L-erythrulose increases, these changes were not observed to be rigidly stoichiometric as would be expected. This cast doubt over whether accurate rate information could in fact be obtained using these assays. Nevertheless they do still provide useful information regarding the relative reactivities of each of the TK mutants towards glycolaldehyde, and therefore can be used as a tool to aid in the selection of mutants for more detailed investigations. These assays can be employed as a type of screening process, for analysis of large numbers of new TK mutants in order to identify recombinant enzymes possessing novel reactivities.

Having determined the presence of a suitable level of enzymatic activity towards this type of substrate under the biotransformation conditions required for this HPLC assay, work was undertaken to investigate the stereoselectivities of these constructs by using them to transform D-glyceraldehyde and DL-glyceraldehyde. This work is detailed in the following section.

## 5.4 Reaction with D-glyceraldehyde and DL-glyceraldehyde

The stereoselectivity of the seven modified TKs was investigated, again using β-HPA as the ketol donor, with D-glyceraldehyde and DL-glyceraldehyde as aldehyde acceptors. TK conversions for D-glyceraldehyde and L-glyceraldehyde, showing their respective biotransformation products, are depicted in Figure 5.11. DL-glyceraldehyde was chosen as a substrate for these investigations due to the non-availability of a commercial source of L-glyceraldehyde. Levels of activity for L-glyceraldehyde were then deduced from graphs showing 100% drop for D-glyceraldehyde and 50% drop for DL-glyceraldehyde with wild-type TK. Biotransformations were performed (2.2.19) at substrate concentrations of 100mM (β-HPA and aldehyde), with TK produced from O/N culture (3ml), in a final reaction volume of 10ml. As before, reactions were allowed to run for 1h and were monitored using TLC (ethyl acetate:methanol = 95:5 [$R_f$ values: β-HPA, 0.88; D-glyceraldehyde, 0.77; DL-glyceraldehyde, 0.77]).

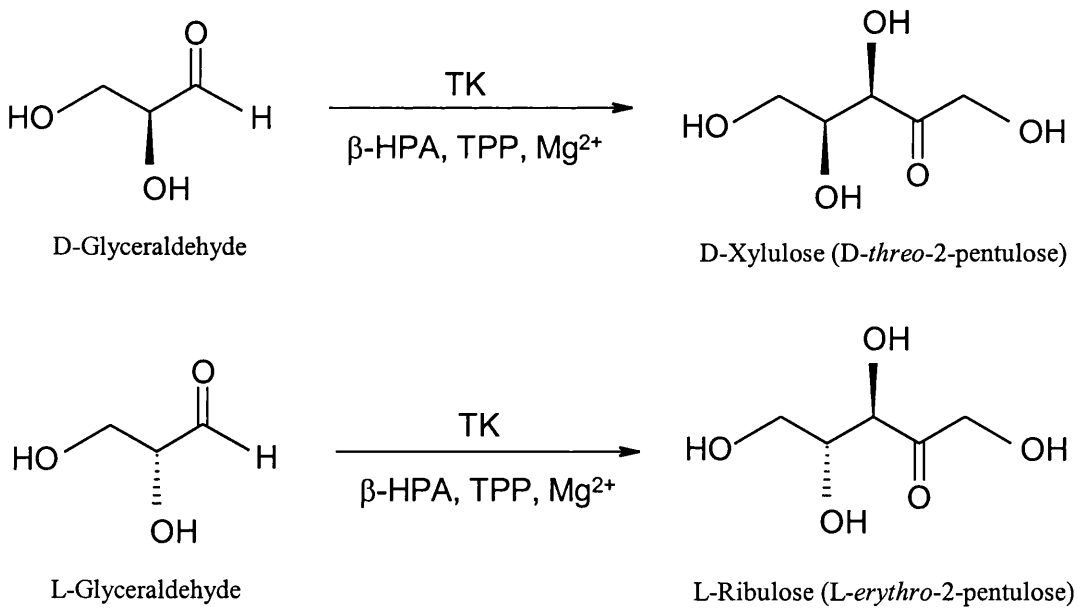



**Figure 5.11** *TK reaction of β-HPA with D-/L-glyceraldehyde*

The isocratic HPLC system used for these experiments was identical to that employed in section 5.3. In this case aliquots were taken every minute for the first ten minutes and then after fifteen, thirty, forty-five and sixty minutes of TK transformation. Dilution of reaction samples to the required concentration for the assay (ca. 20mM) in

the acidic mobile phase (pH 2.3) simultaneously terminated the TK reaction. The substrates were stable under these acidic conditions for the time required for analysis. Samples were filtered through 0.45μm filters prior to injection (10μl).

Standard curves for β-HPA, D-glyceraldehyde (D-gla) and DL-glyceraldehyde (DL-gla) were plotted (see Figure 5.12) to allow HPLC integrated peak areas to be converted into concentration values (mM). Retention times for each of the compounds under these conditions were: β-HPA, 21.32 minutes; D-glyceraldehyde, 29.74 minutes; DL-glyceraldehyde, 29.74 minutes. This enabled a direct concentration loss of these substrates, due to enzymatic conversion, to be calculated. As explained in section 5.3, there was some concern regarding the accuracy of peak area to concentration conversions made using these standard curves, and the same limitations are applicable here. For these reactions changes in glyceraldehyde concentration were chosen for calculation of rate data, again because of the problems encountered with β-HPA in Chapter Four. Control reactions, measuring the loss of glyceraldehyde under these assay conditions, but in the absence of crude extract, were performed and it was found that no appreciable degradation of this compound occurred under these conditions. Therefore, calculated values were considered to be reasonable within the limits given although the concerns relating to mass balances discussed in section 5.3 apply here also.

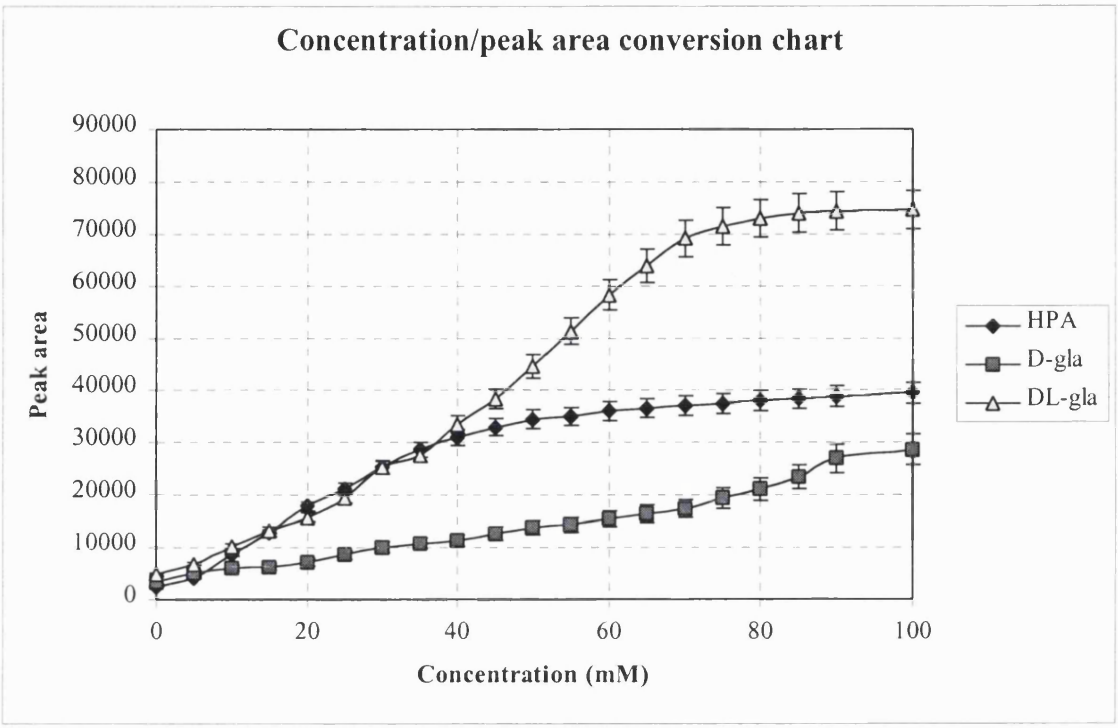
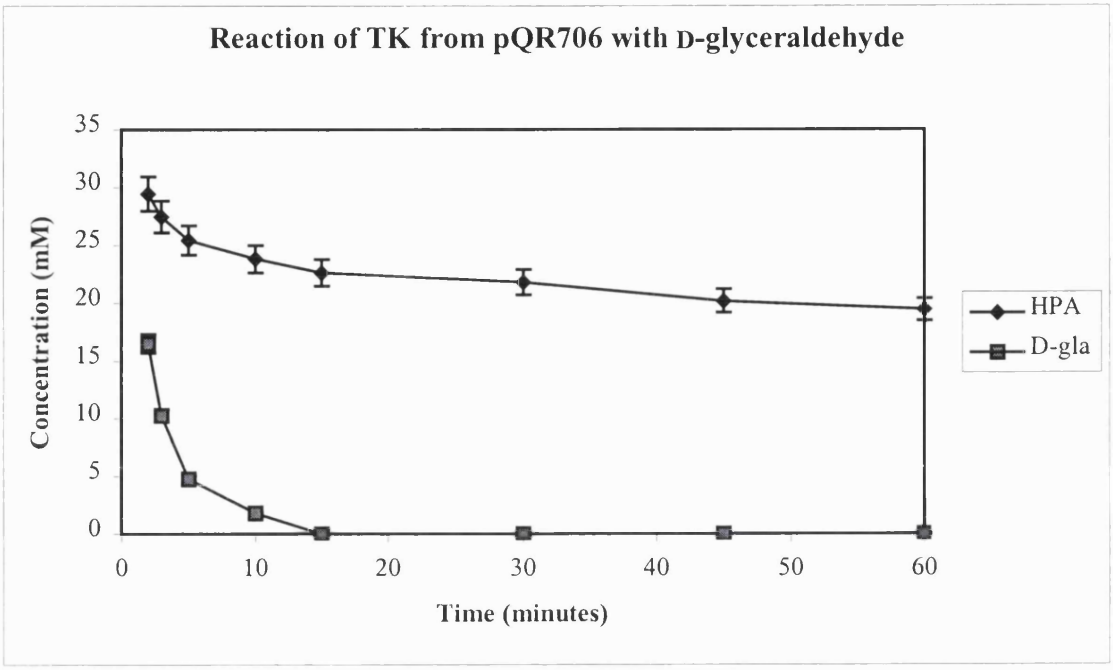



**Figure 5.12** *Standard HPLC curve for β-HPA/D-glyceraldehyde/DL-glyceraldehyde*

Separate biotransformations were performed for each of the TK constructs with both D-glyceraldehyde and DL-glyceraldehyde. The twin column HPLC system detected all substrates present in the reaction mixtures and transformations were followed by measurement of the loss of glyceraldehyde (see Figures 5.13 through 5.20). As can be seen from these sixteen charts, the various TK mutants exhibited a wide range of activity levels with these two substrates, thus indicating possible variations in the stereoselectivities of these enzymes. Reaction products from these biotransformations were not effectively detected during these assays and it could be that some other component of the crude enzyme extract employed for these reactions interfered with their detection. This further supports the argument that it might be preferable to use a pure form of TK when conducting these transformations, and also suggests that investigations into the possibility of an alternative detection method may be beneficial.



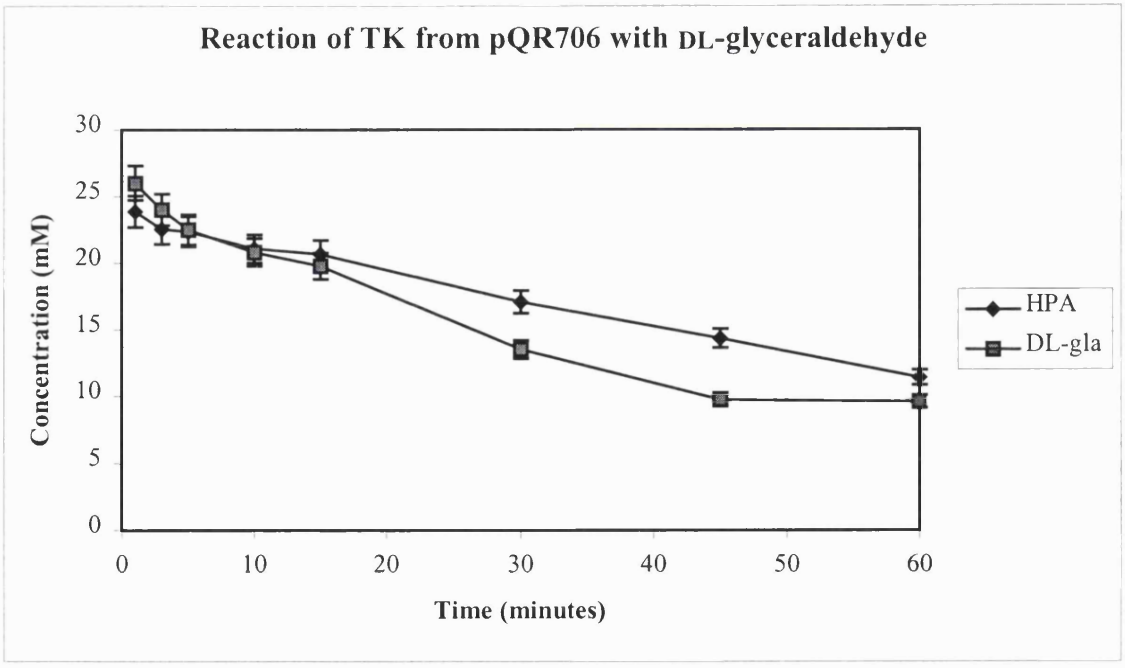
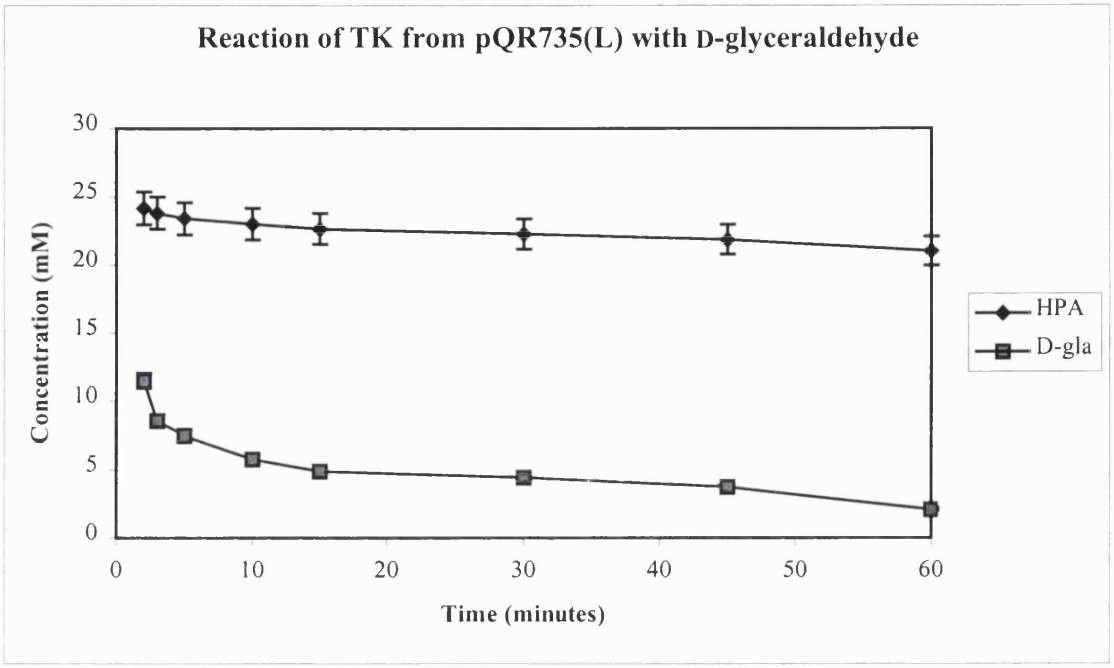




**Figure 5.13** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for wtp reaction*



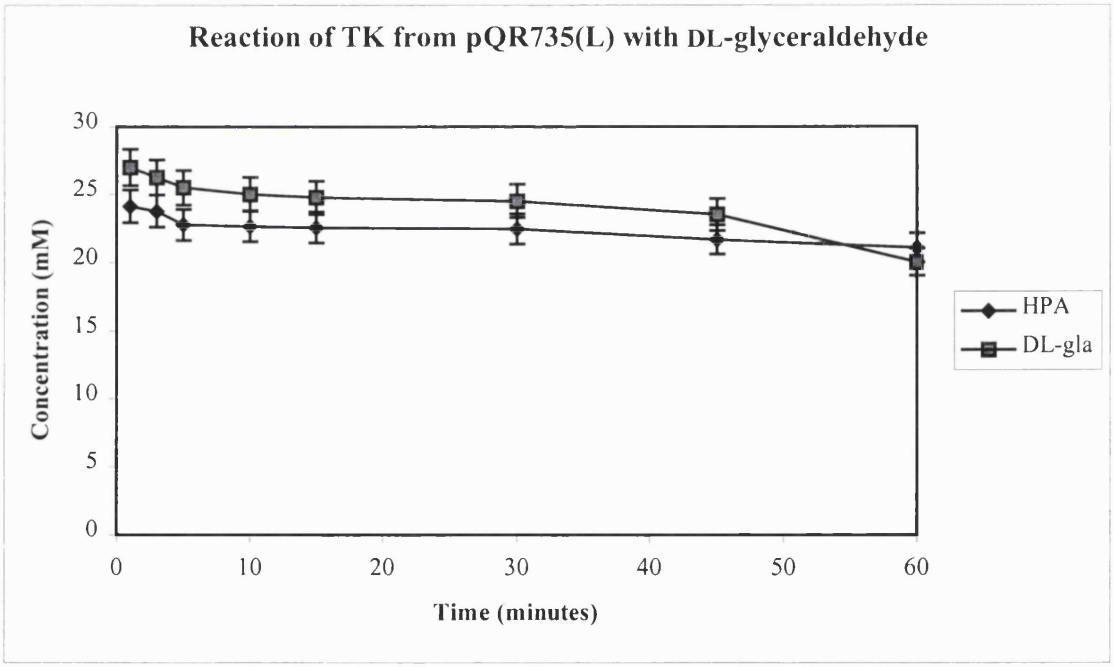
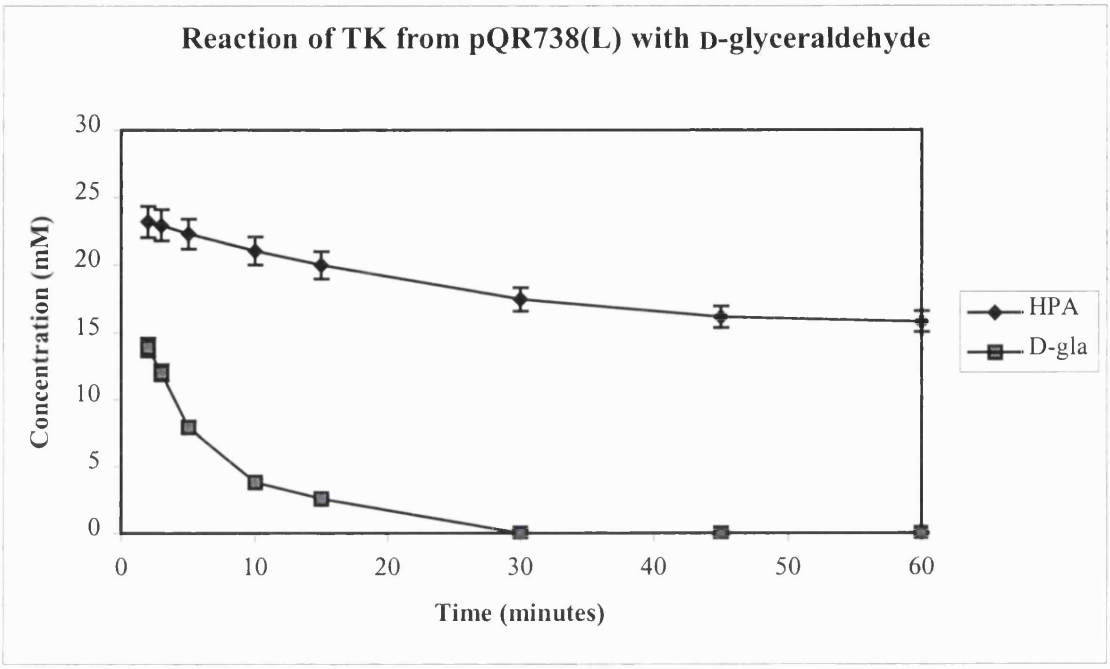




**Figure 5.14** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for S385*L reaction*



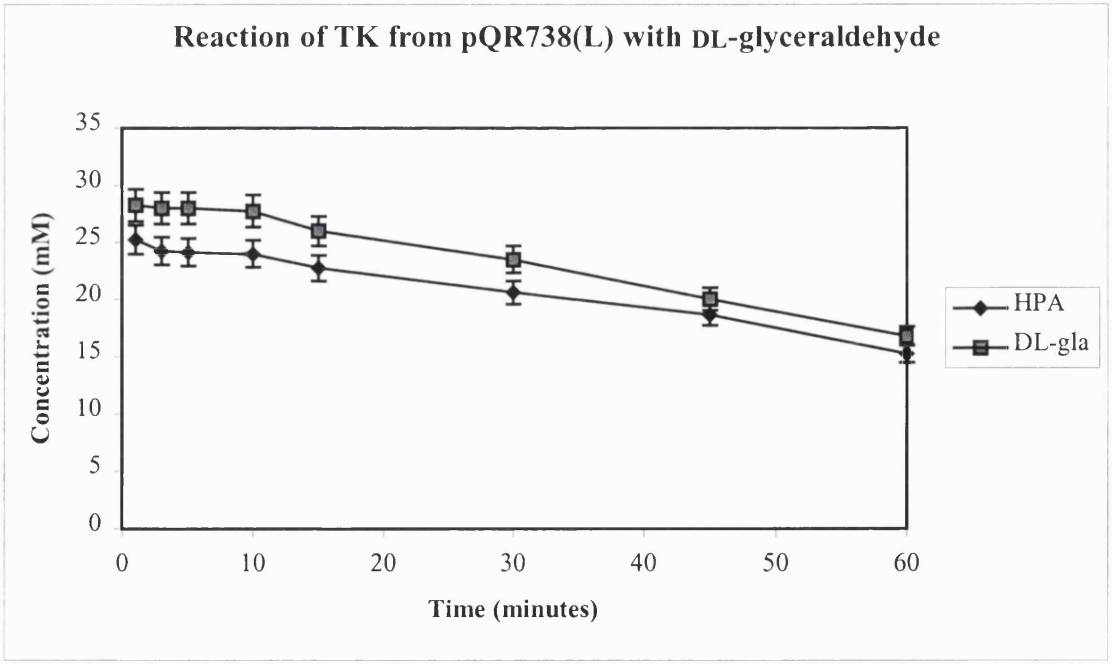
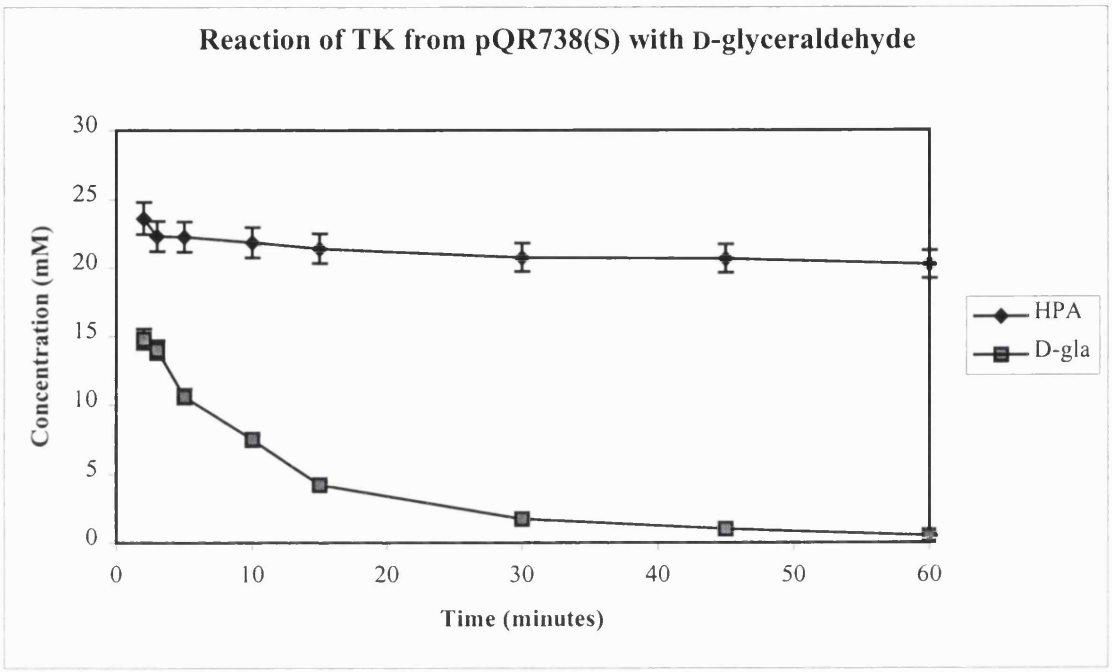




**Figure 5.15** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for H461*L reaction*



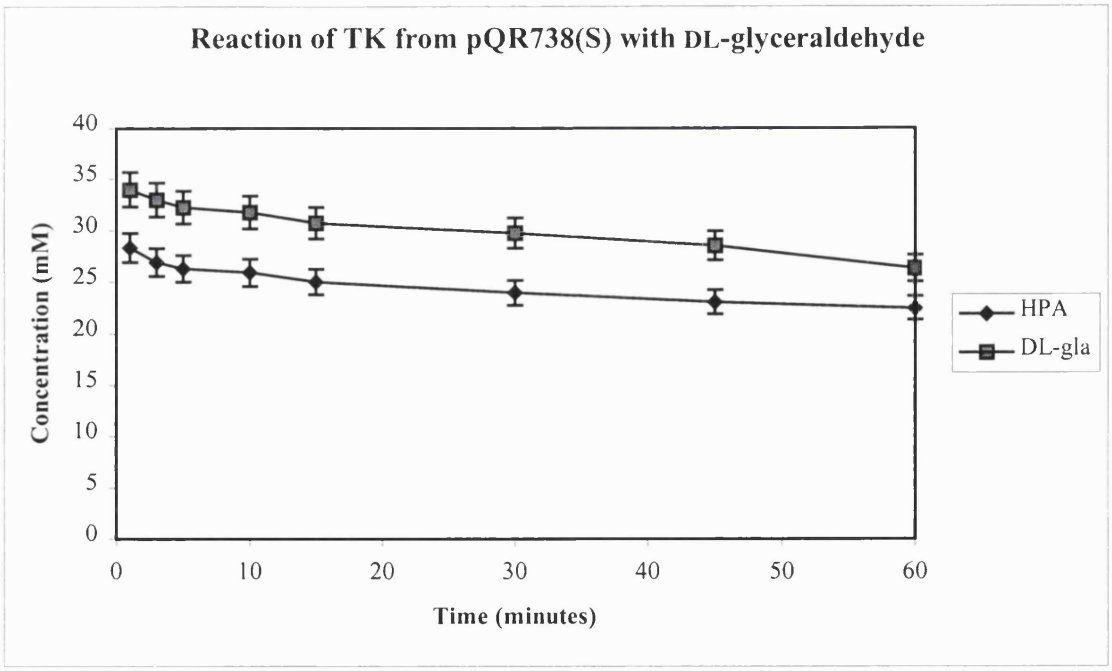
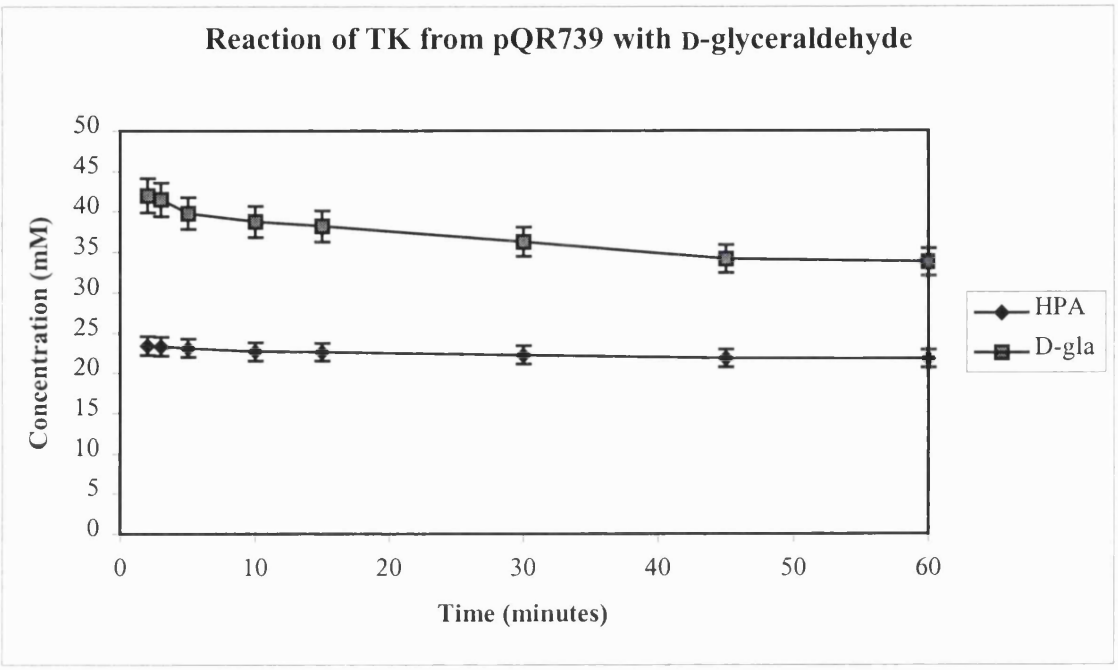




**Figure 5.16** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for H461*S reaction*



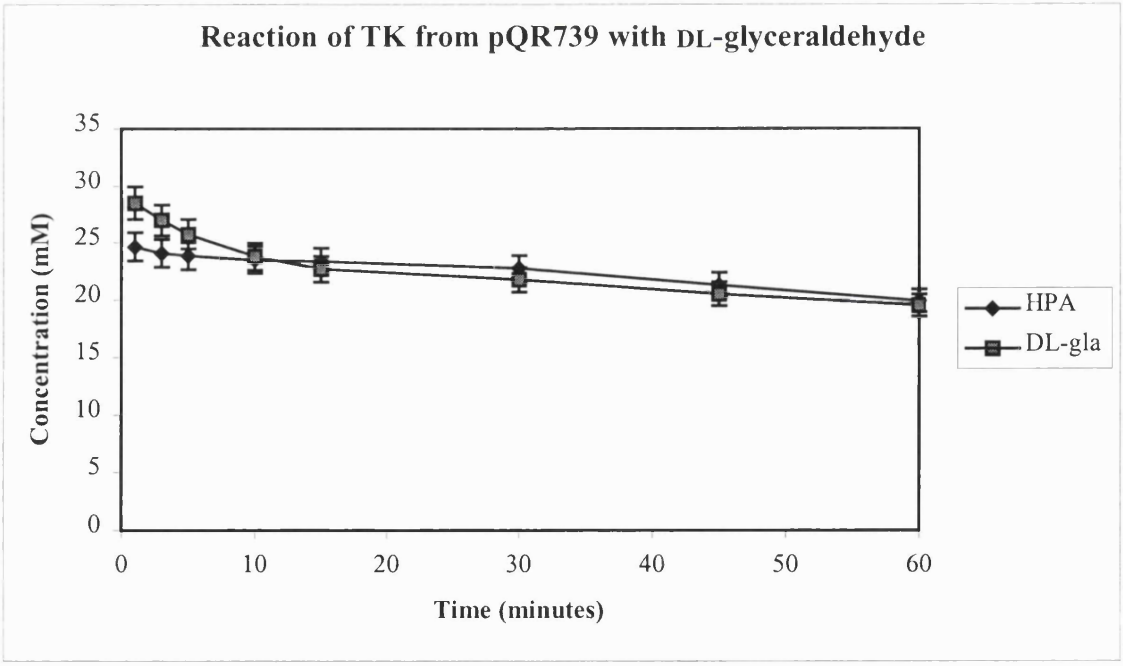
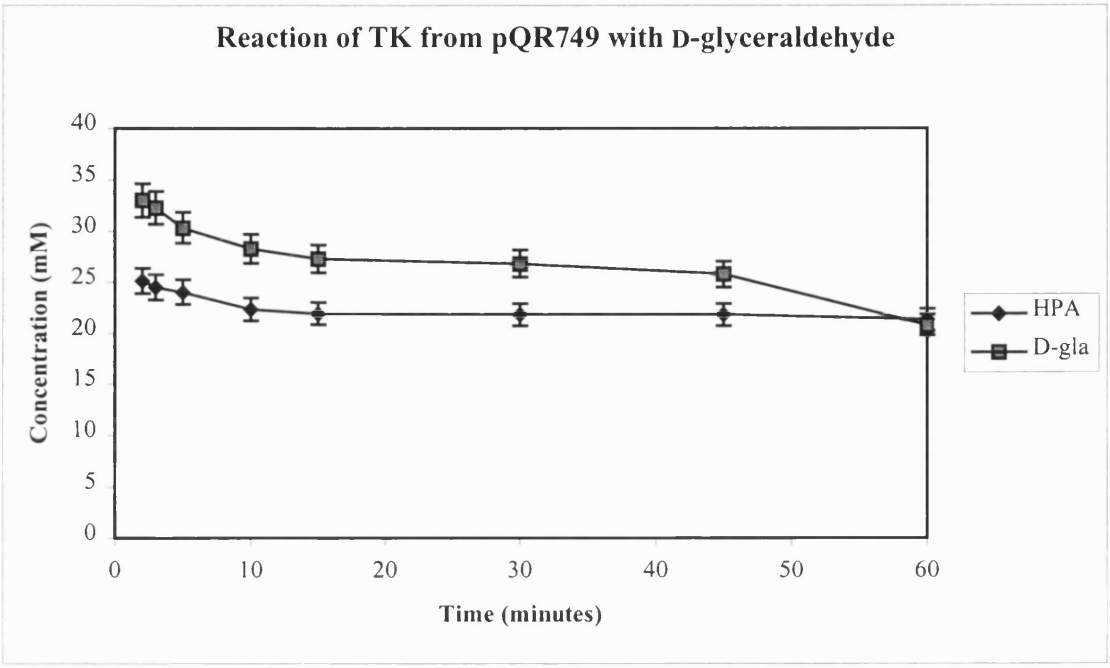




**Figure 5.17** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for I189A reaction*



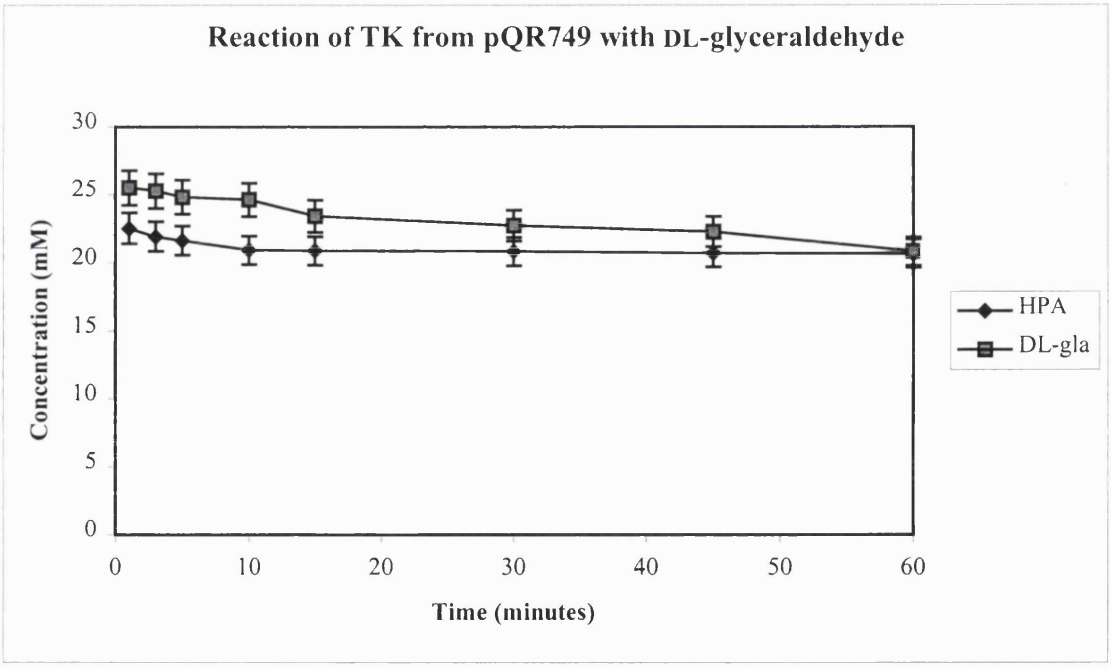
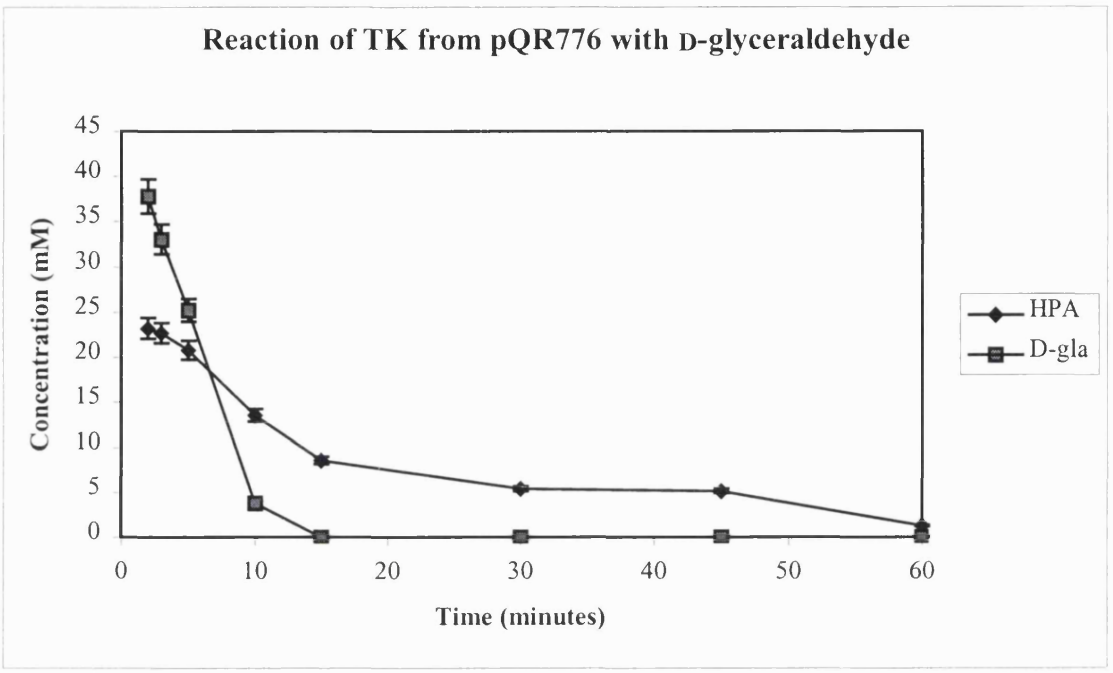




**Figure 5.18** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for I189V reaction*



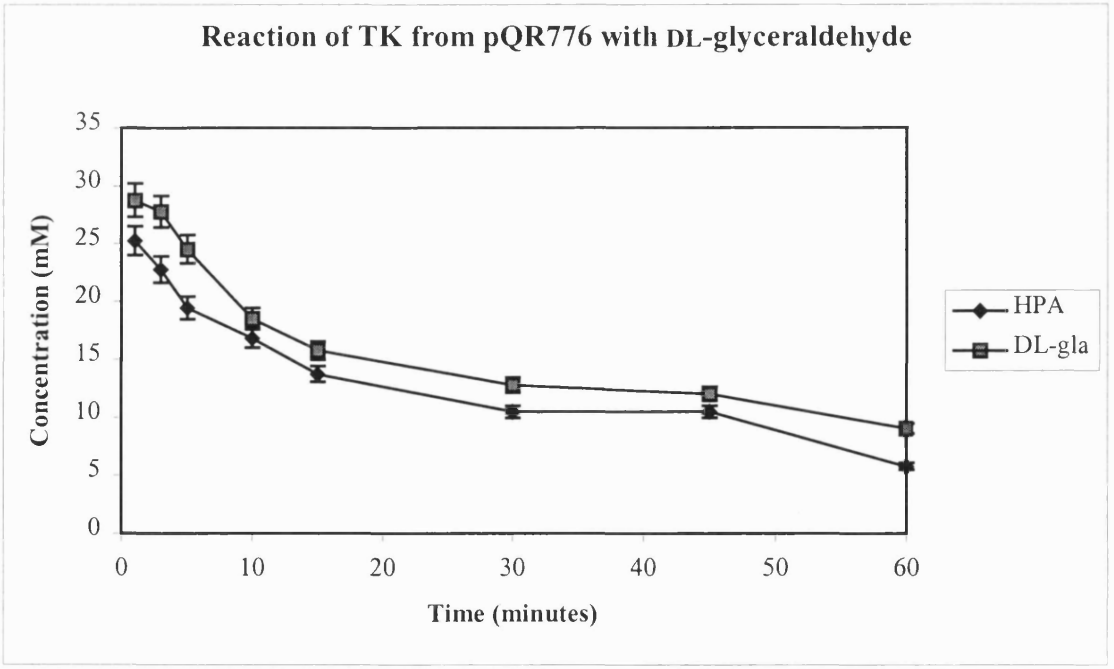
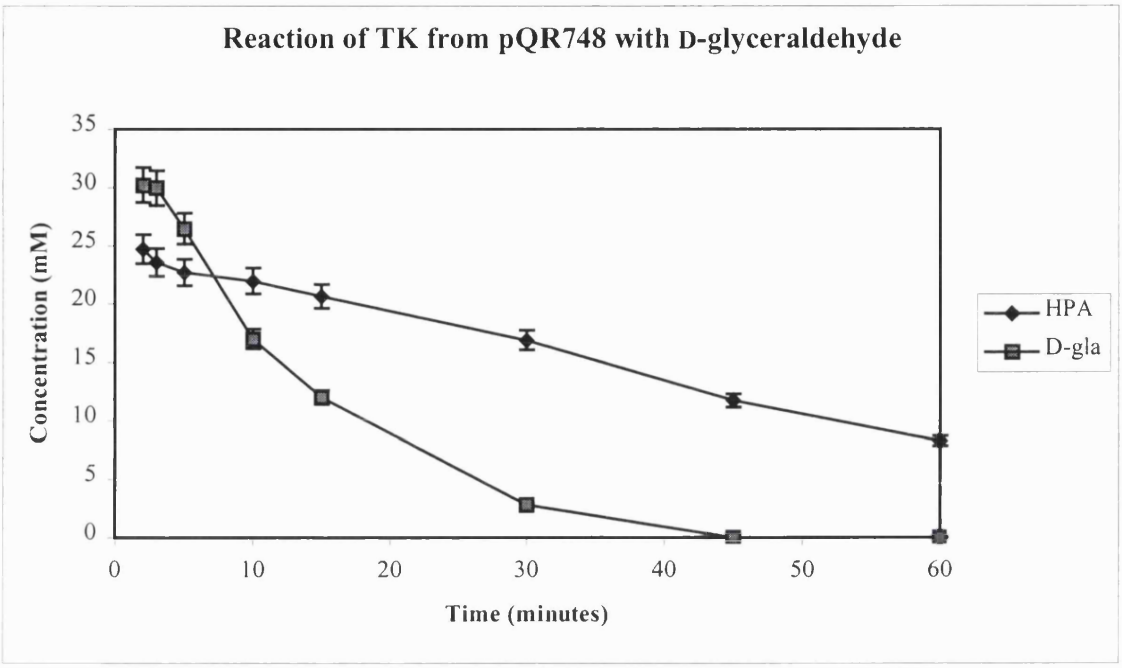




**Figure 5.19** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for H258A reaction*



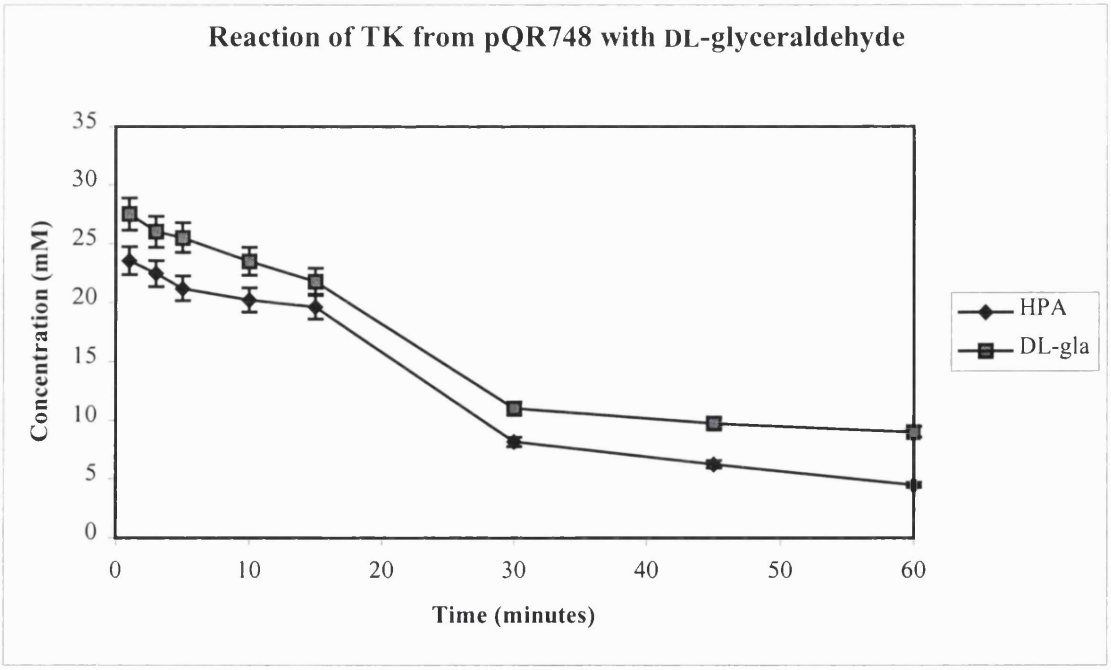




**Figure 5.20** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for H258D reaction*

| *E. coli* strain/construct | Specific Activity (U/mg) | | | |
|---|---|---|---|---|
| | D-glyceraldehyde | Error (%) | DL-glyceraldehyde | Error (%) |
| JM107/pQR706 | 4.69 | +/- 5.6 | 1.98 | +/- 3.1 |
| Amber-LEU/pQR735 | 0.95 | +/- 2.8 | 0.87 | +/- 2.2 |
| Amber-LEU/pQR738 | 2.05 | +/- 3.6 | 1.38 | +/- 3.3 |
| Amber-SER/pQR738 | 1.20 | +/- 3.4 | 1.25 | +/- 2.9 |
| JM107/pQR739 | 1.56 | +/- 3.7 | 0.83 | +/- 1.9 |
| JM107/pQR749 | 2.37 | +/- 4.4 | 1.06 | +/- 2.6 |
| JM107/pQR776 | 3.33 | +/- 4.7 | 6.16 | +/- 5.2 |
| JM107/pQR748 | 2.21 | +/- 4.1 | 3.70 | +/- 4.8 |

**Table 5.2** *TK specific activities for reactions with D-glyceraldehyde and DL-glyceraldehyde**

* Calculated using disappearance of glyceraldehydes data

Loss of glyceraldehyde data was converted into specific activities, and analysis of these values gave an indication as to possible relaxation/reversal of the stereoselectivity of each of the enzymes tested. Specific activities for each mutant are given in Table 5.2 and relative specific activities for reaction of each of the constructs with D-glyceraldehyde and DL-glyceraldehyde are given in Table 5.3. It is interesting to note from Table 5.3 that all of these constructs, to varying degrees, have an increased specific activity with DL-glyceraldehyde when compared to the wild-type enzyme produced from pQR706. However TK from pQR776 [H258A] (and to a lesser extent that from pQR748 [H258D]) appeared to have the greatest potential activity towards the L-enantiomer. Hence this particular construct was selected for further biotransformations, with more detailed product analysis using NMR.

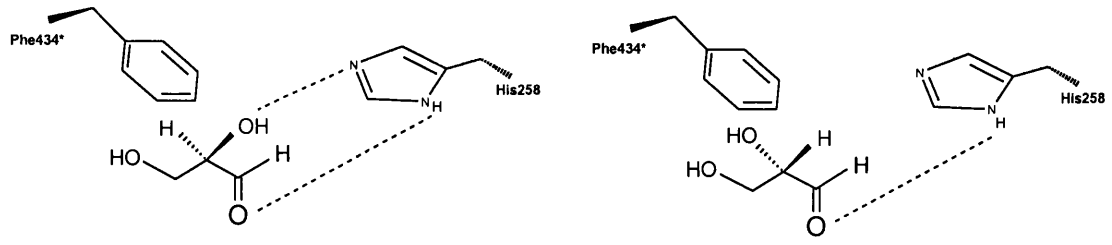
As was explained in section 5.3 there was some concern here regarding the lack of mass balances found during these assays, and again there is doubt as to whether accurate rate data can in fact be deduced from these assays. Nevertheless, relative activity information for each of the TK mutants compared to the wild-type enzymatic conversions of these substrates may be deduced, in addition to relative activity information comparing transformations of D-glyceraldehyde and DL-glyceraldehyde for each construct. As with the information obtained during experiments described in section 5.3 for reactions with glycolaldehyde, the provisional activity data obtained here may be used as a basis for the selection of potentially useful TK mutants to be

employed for more detailed investigations regarding their potential to transform these compounds.

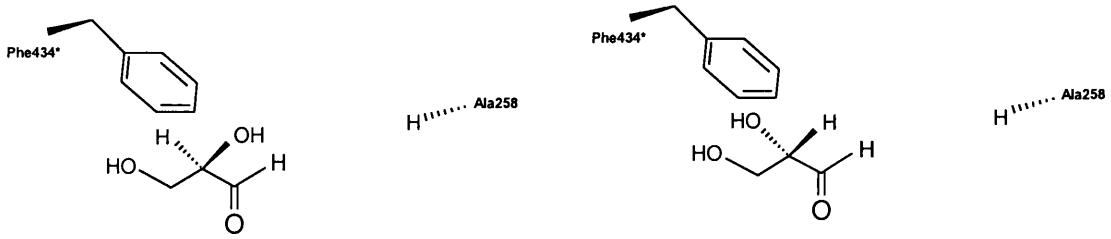
| *E. coli* strain/construct | Relative Specific Activities | |
|---|---|---|
| | D-glyceraldehyde | DL-glyceraldehyde |
| JM107/pQR706 | 1.00 | 0.42 |
| Amber-LEU/pQR735 | 1.00 | 0.92 |
| Amber-LEU/pQR738 | 1.00 | 0.67 |
| Amber-SER/pQR738 | 1.00 | 1.04 |
| JM107/pQR739 | 1.00 | 0.53 |
| JM107/pQR749 | 1.00 | 0.45 |
| JM107/pQR776 | 1.00 | 1.85 |
| JM107/pQR748 | 1.00 | 1.67 |

**Table 5.3** *DL-glyceraldehyde specific activities relative to those for D-glyceraldehyde*

Introduction of the much smaller alanine residue (which also lacks hydrogen bonding ability) into this region was anticipated to have these sorts of effects. This is highlighted by the diagrams shown in Figure 5.21. This is a hypothetical picture derived from viewing 3-D graphics of the TK active site and is not based on computerised model building.




**TK from pQR706: Wild-type transketolase**




**TK from pQR776: Transketolase with partial relaxation of stereoselectivity**

**Figure 5.21** *Hypothetical picture for partial relaxation of TK stereoselectivity*

$^{1}H$ and $^{13}C$ NMR analysis for reaction of recombinant TK from pQR776 with DL-glyceraldehyde showed the presence of both D-xylulose and L-ribulose, the transformation products for D-glyceraldehyde and L-glyceraldehyde respectively. $^{13}C$ NMR data for these two compounds was as follows:

D-xylulose

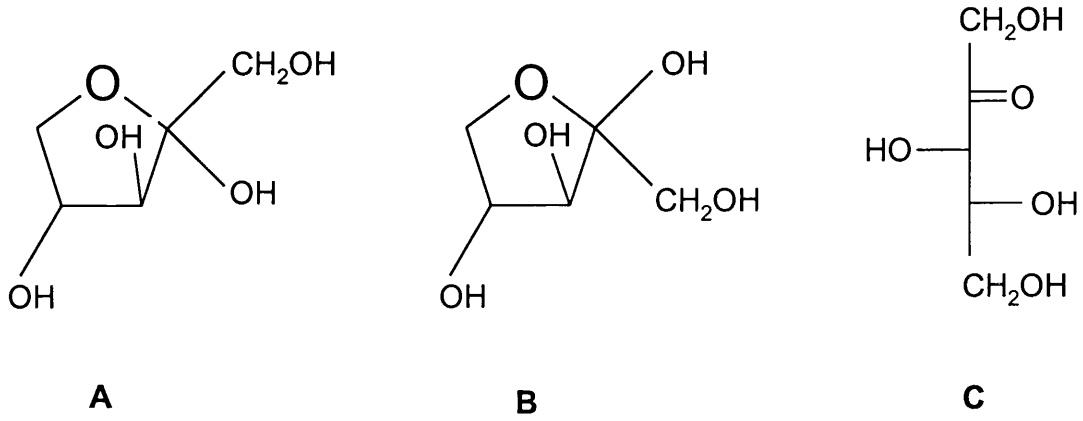
$^{13}C$ NMR (400MHz) DMSO-$d_6$ δ: $64^{5C}$, $65^{1A}$, $66^{1B}$, $69^{1C}$, $72.5^{5B}$, $74.75^{5A}$, $75^{4C}$, $78^{4B}$, $78.5^{4A}$, $79^{3C}$, $79.5^{3B}$, $84^{3A}$, $105.5^{2B}$, $107^{2A}$, $210^{2C}$.

L-ribulose

$^{13}C$ NMR (400MHz) DMSO-$d_6$ δ: $61^{5C}$, $63.5^{1B}$, $64^{1A}$, $67^{1C}$, $71.5^{5B}$, $72^{5A}$, $72.5^{4A}$, $73^{3A}$, $74^{4B}$, $75^{4C}$, $75.5^{3C}$, $76^{3B}$, $103^{2B}$, $106^{2A}$, $213^{2C}$.

The large number of carbon peaks for each compound results from the fact that each of these compounds can exist in three separate tautomeric forms as depicted in Figure 5.22, a solution will contain the acyclic keto form and two furanose forms. The corresponding carbon atoms for each tautomer will be present in slightly different electronic environments and so their nuclei will be differently affected by a magnetic field, thus resulting in a total of fifteen carbon signals for each compound. This data is consistent with previously assigned NMR spectra for these structures (Demuynck *et al.*, 1990; Vuorinen & Serianni, 1990).

Tautomers of D-xylulose (D-*threo*-2-pentulose)




Tautomers of L-ribulose (L-*erythro*-2-pentulose)

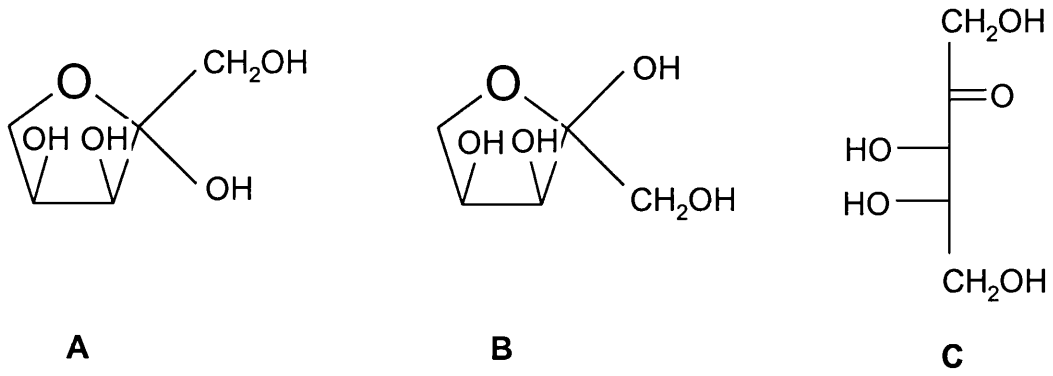



**Figure 5.22** *Alternative structures for D-xylulose and L-ribulose*

## 5.5 Further mutants

Since the work conducted thus far appeared to support the theory that residues H258 and F434* play an important role in determining the stereoselectivity of TK, a further set of five TK mutants were engineered (these 5 TK mutants were constructed during this project). New constructs were generated using a QuikChange™ Site-Directed Mutagenesis Kit according to the scheme depicted in Figure 4.27. These constructs comprised a series of enzyme variants, as shown in Table 5.4, designed to exhibit a gradual relaxation of the requirement for a D-enantiomeric α-hydroxyl group. Also given are the percentage efficiencies for each of the mutagenesis reactions. The range thus produced culminates in the construction of pQR782, which produces a TK variant with reversal of these two key residues.

| *E. coli* strain/construct | Alteration (s) | Base changes | % Mutagenesis efficiency |
|---|---|---|---|
| JM107/pQR776 | H258A | CAC to GCC | |
| JM107/pQR780 | F434*A | TTC toGCC | 100 |
| JM107/pQR781 | H258A<br>F434*A | CAC to GCC<br>TTC to GCC | 25 |
| JM107/pQR782 | H258F<br>F434*H | CAC to TTC<br>TTC to CAC | 25 |
| JM107/pQR783 | H258F | CAC to TTC | 50 |
| JM107/pQR784 | F434*H | TTC to CAC | 50 |

**Table 5.4** *TK stereoselectivity mutants*

Two important mutants in this series are shown in the hypothetical diagrams depicted in Figure 5.23, these pictures highlight the predicted interactions between the side chain of the residue in position 434* and the groups located on the α-carbon atom for each of the enantiomers of glyceraldehyde. The enzyme from pQR781 has substitution of both H258 and F434* with alanine residues, this was expected to cause complete relaxation of TK stereoselectivity since the majority of bonding and hydrophobic interactions in the region of the α-hydroxyl group will have been destroyed. These interactions include hydrophobic repulsive interactions involving F434* and the other substituent of the α-carbon atom, tending to destabilize the reaction intermediates. TK from pQR782 incorporates the exchange of residues H258 and F434*, with an intended reversal of the stereoselectivity of the enzyme.


**TK from pQR706: Wild-type transketolase**


**TK from pQR781: Transketolase with complete relaxation of stereoselectivity**


**TK from pQR782: Transketolase with reversed stereoselectivity**

**Figure 5.23** *Hypothetical active site stereoselectivity control showing key enzyme/substrate interactions*

Since it was previously understood that the H258 and F434* residues faced each other on opposite sides of the active site funnel, this guided the choice of these residues for mutagenesis. However, a recent, much higher resolution structure shows that H258 is further away from its believed position in the active site funnel than previously thought. Nevertheless, this does not alter the logic behind the choice of F434* which is in the same place as in the original 3-D structure. The F434* side chain protrudes into the active site funnel, substituting this with alanine and with histidine does alter activity and the previous explanation outlined for this change of activity does still apply. Mutants having alteration of H258, to alanine and to phenylalanine, also have an altered phenotype although the reason for the change in activity may not be as simple as previously thought. Since this residue can now be placed further away from the α-hydroxyl group of the aldehyde acceptor, at present a conclusive explanation for the observed reactivity difference is not available. In addition, the latest refined TK active site structure highlights another histidine residue, H261, which is located within the active site funnel in the proximity of the α-hydroxyl group and would be an interesting amino acid to alter in subsequent work.

Over-expression of these mutant transketolases was performed in *E. coli* strain JM107 using the TK promoter. Initial characterisation was conducted using the linked-enzyme assay (2.2.1). Specific activities of these modified enzymes were calculated and values, along with those for the native protein (from pQR706) and the previous stereoselectivity mutant (from pQR776), are given in Table 5.5. Examination of this data shows that, as expected, all mutant TKs have reduced activity towards the native sugar substrates relative to the wild-type protein.

| *E. coli* strain/construct | Activity (U/ml) | Concentration (mg/ml) | Specific activity (U/mg) | Error (%) |
|---|---|---|---|---|
| JM107/pQR706 | 4.330 | 0.62 | 6.98 | +/- 4.6 |
| JM107/pQR776 | 4.030 | 0.67 | 6.01 | +/- 3.6 |
| JM107/pQR780 | 3.982 | 0.65 | 6.13 | +/- 3.4 |
| JM107/pQR781 | 4.181 | 0.61 | 6.85 | +/- 4.3 |
| JM107/pQR782 | 0.526 | 0.63 | 0.83 | +/- 2.4 |
| JM107/pQR783 | 1.432 | 0.62 | 2.31 | +/- 3.5 |
| JM107/pQR784 | 1.305 | 0.63 | 2.07 | +/- 3.2 |

**Table 5.5** *Specific activities for TK mutants*

The very small reductions in activity when either, or both, of H258 and F434* are replaced with alanine can be explained by the fact that these changes cause an opening up of the active site in, or close to, the region of substrate binding. Although certain bonding interactions may be lost due to these alterations, no significant unfavourable hydrophobic interactions are introduced, and so these TK variants still have a high activity towards the native sugars. Mutants generated from pQR782, pQR783 and pQR784 demonstrated considerably reduced activity in the linked-enzyme assay, for TK from pQR783 and pQR784 activity levels are reduced to a third of that for the wild-type protein. For the enzymes from pQR782 and pQR784, the replacement of F434* with a hisitdine residue will introduce a charged side chain into this region, this may result in unfavourable interactions with the D-enantiomeric sugars used in this assay. In the case of TK from pQR783 the observed changes in specific activity can, at present, be less readily explained due to the uncertainty regarding the precise location of H258.

## 5.6 Discussion

In classical organic chemistry many types of C-C bond-forming reaction have been developed, although few of these conversions are selective towards their reagents and have the ability to dictate the stereochemistry of the products. Since TK has reasonably relaxed specificity for its aldehyde component, combined with the high stereoselectivity for its C-C bond-forming step, it is potentially an extremely useful catalyst for asymmetric synthesis. However, one current restriction in its applications is that in the biocatalytic transformation of novel substrates with an α-hydroxyl group, TK has a strict requirement for the hydroxyl moiety to be in the D-configuration (Demuynck *et al.*, 1991). Some reports have depicted the use of simple α-unsubstituted aldehydes as substrates for TK (Hobbs *et al.*, 1993; Morris *et al.*, 1996), however these are accompanied by significantly lower conversion rates than for the hydroxylated compounds. Contained within this chapter was the work carried out in an attempt to generate a series of TK variants, each designed to have a slightly different modification of the wild-type stereospecificity. These modified enzymes were characterised through transformation of both enantiomers of the sugar glyceraldehyde.

Molecular modelling studies have highlighted residues believed to be acting to orient the substrate, allowing exclusive binding of molecules with an α-hydroxyl group of the correct stereochemistry. The two residues, H258 and F434*, initially believed to be involved in dictating the stereoselectivity of the wild-type enzyme were substituted in

various combinations, and initial investigations using the D-enantiomeric native sugar substrates as part of the linked-enzyme assay were conducted. It was expected that a mutant with reversal of these two residues would exhibit the opposite stereoselectivity to that of the native enzyme, whereas a variant in which they have both been replaced with alanines would, in theory, display a complete relaxation of the requirement for D-enantiomeric substrates.

As a continuation of this work, the mutants were screened to determine their relative activities towards the model substrate glycolaldehyde, employing β-HPA as the ketol donor. Initially reactions were monitored using a HPLC assay, following the loss of glycolaldehyde due to enzymatic conversion. However, a number of problems were encountered when using these HPLC analyses, most noticeably the lack of mass balances found during the assays (this was also true of assays for later transformations involving the two enantiomers of glyceraldehyde). Although in most cases the loss of β-HPA was closely mirrored by the corresponding loss of glycolaldehyde, and periods of greater β-HPA/glycolaldehyde loss were generally accompanied by some increase in L-erythrulose, these changes were not found to be rigidly stoichiometric as would be expected. This cast doubt over whether accurate rate data could, in fact, be obtained from these assays.

However, this type of analysis did still provide useful information regarding the relative reactivities of each of the TK mutants towards these sugar compounds, and was an important screening tool for the selection of mutants to be used for more detailed analysis. Control reactions were not performed in the absence of the aromatic aldehydes (but with all other reaction components present) and it may have been that some other compound or enzyme within the crude enzyme extract was causing the reaction substrates to break down. Additional reactions, with implementation of these particular controls, would need to be conducted in order to eliminate or to confirm this possibility. Further to this, there may have been sufficient quantities of alternative sugar ketol donors, or aldehyde acceptors, in the added TK crude enzyme extract to have lead to competing reactions, thus resulting in some of the non-stoichiometric observations found during HPLC analysis. If, after further investigations, this were found to be the case, then purification of the TK enzyme before use in biotransformations would be the best way to eliminate these problems.

Having determined the presence of a suitable level of enzymatic activity towards glycolaldehyde under the biotransformation conditions required by the HPLC assay, work was then undertaken to investigate the stereoselectivity of these constructs by using them to transform D-glyceraldehyde and DL-glyceraldehyde. Again reactions were initially monitored using HPLC techniques, by measuring the loss of the aldehyde component. Upon analysis of these results, the enzyme produced from pQR776 [H258A] appeared to have the greatest potential activity towards the L-enantiomer. Hence this particular construct was selected for further biotransformations, with more detailed reaction product analysis being conducted using NMR techniques. Introduction of the much smaller alanine residue (which also lacks hydrogen bonding ability) at this location was anticipated to have these sorts of effects, since it results in the introduction of a cavity in the region of the active site believed to be involved in dictating the stereoselectivity.

As the work on the initial TK mutant produced from pQR776 appeared to support the theory that residues H258 and F434* played an important role in determining the stereoselectivity of TK, a further set of five TK mutants were engineered. These new constructs comprised a series of enzyme variants, designed to exhibit a gradual relaxation of the requirement for a D-enantiomeric α-hydroxyl group. The range culminated in the construction of pQR782, which produces a TK variant with reversal of these two key residues.

In order to determine the precise stereoselectivity of the new TK mutants a more detailed analysis, using the HPLC screening approach, is required. Although the linked-enzyme assay data thus far obtained provides an insight into the differing behaviour of these constructs towards their native D-enantiomeric sugar substrates, direct examination using the individual enantiomers of a particular compound is needed. As before, reactions could be screened using HPLC techniques, and from this data promising TK constructs could be selected for further investigation including the structural characterisation of their reaction products.

Since it was previously understood that residues H258 and F434* faced each other on opposite sides of the active site funnel, this guided the choice of these residues for mutagenesis. However, a much higher resolution structure recently produced shows that H258 is further away from its originally believed position in the active site funnel.

However, this does not affect the reasoning behind the choice of F434*, which remains in the same place as in the original 3-D structure. Mutants having an alteration of H258, to alanine and to phenylalanine, do have an altered phenotype although the reason for the change in activity may be more complicated previously thought. Since this residue can now be placed further away from the region believed to be occupied by the α-hydroxyl group, at present a conclusive explanation for the observed reactivity difference is not available. Further to this, the latest TK active site structure highlights another histidine residue, H262, which is located within the active site funnel in the proximity of the α-hydroxyl group. The precise role of this residue would be a key area to investigate as part of future TK stereoselectivity experiments.

# Chapter Six: Experimental

# CHAPTER SIX: Experimental

All media, reagents, and buffers are contained in Appendix One. General techniques are described in Chapter Two. The sequences of all oligonucleotides are contained within Appendix Two.

## 6.1 Preparation of uracil-containing pQR706/776

1) pQR706/776 (2μl at 50ng/μl) was used to transform (2.2.7) *E. coli* strain CJ236 (A1.3 [100μl]), placed on LB/amp agar and grown O/N at 37°C;
2) A single colony was grown in LB/amp (5ml) O/N at 37°C, 225rpm;
3) This was used to make glycerol freeze stocks (2.2.5);
4) U-containing single-stranded pQR706/776 was prepared (2.2.10), the DNA resuspended in TE (50μl) and quantified (2.2.8). Concentrations for template ssDNA are given in Table 6.1.

| Template ssDNA | Concentration (μg/ml) |
|---|---|
| pQR706 | 0.75 |
| pQR776 | 1.25 |

**Table 6.1** *Template ssDNA concentrations*

## 6.2 Attempted construction of pQR780/781/783/784

1) The oligonucleotides H258A, H258F, F434A and F434H (A2.1) were purchased from Pharmacia at known concentrations (see Table 6.2);
2) The 3kb linear pBluescript fragment (*Eco*RI/*Nco*I) and the relevant oligonucleotide were annealed to the template ssDNA (2.2.14);
3) The second strand was synthesized (2.2.15);
4) The resultant dsDNA was worked up (2.2.4) and resuspended in TE (10μl) of which 5μl was used to transform JM107 cells (A1.3 [100μl]). Transformed cells were plated onto LB/amp and grown O/N at 37°C.

| Oligonucleotide | Concentration (ng/μl) | $T_m$ (°C) |
|---|---|---|
| H258A | 110.2 | 86.6 |
| H258F | 112.3 | 82.4 |
| F434A | 84.1 | 82.9 |
| F434H | 144.1 | 80.4 |

**Table 6.2** *Oligonucleotide concentrations and melting temperatures*

The combinations (and quantities) of template ssDNA, linear pBluescript and mutagenic oligonucleotide used for the production of each construct are given in Table 6.3.

| Construct | Template ssDNA (μl) | Linear pBluescript (μl) | Oligonucleotide (μl) |
|---|---|---|---|
| pQR780 | pQR706 (2) | 3 | F434A (1.73) |
| pQR781 | pQR776 (1.2) | 3 | F434A (1.82) |
| pQR783 | pQR706 (2) | 3 | H258F (1.81) |
| pQR784 | pQR706 (2) | 3 | F434H (1.87) |

**Table 6.3** *Mutant construction reaction mixtures*

Numerous attempts to generate TK mutants using this method were unsuccessful and plasmid DNA from colonies analyzed by sequencing failed to show incorporation of the desired mutagenic oligonucleotide. It was therefore decided to construct mutants using a polymerase chain reaction.

## 6.3 PCR mutagenesis of TK gene

Forward and reverse oligonucleotides H258A, H258F, F434A, F434H, R358A and R520A (A2.1) were purchased from Pharmacia at known concentrations (see Table 6.4).

| Oligonucleotide | Concentration (ng/μl) | $T_m$ (°C) |
|---|---|---|
| H258A (FOR) | 110.2 | 86.6 |
| H258A (REV) | 110.2 | 86.6 |
| H258F (FOR) | 112.3 | 82.4 |
| H258F (REV) | 225.6 | 82.4 |
| F434A (FOR) | 84.1 | 82.9 |
| F434A (REV) | 95.9 | 82.9 |
| F434H (FOR) | 144.1 | 80.4 |
| F434H (REV) | 42.1 | 80.4 |
| R358A (FOR) | 83.1 | 79.1 |
| R358A (REV) | 93.9 | 79.1 |
| R520A (FOR) | 76.9 | 78.8 |
| R520A (REV) | 67.9 | 78.8 |

**Table 6.4** *Oligonucleotide concentrations and melting temperatures*

Template DNA was purified from 1.5ml O/N cultures using QIAquick columns obtained from QIAGEN and diluted to a final concentration of 50 ng/μl. Any variations from the wild-type sequence are shown in Table 6.5.

| Template DNA | Amino acid alteration |
|---|---|
| pQR706 | wtp |
| pQR776 | H258A |
| pQR784 | F434*H |
| pQR785 | R358*A |
| pQR739 | I189A |

**Table 6.5** *Template DNA*

## 6.4 Construction of pQR780/781/782/783/784/785/786/787/788/789

The standard reaction mixture below (total 50μl) was used in each case, as part of the QuikChange Site-Directed Mutagenesis Kit purchased from STRATAGENE:

| | |
|---|---|
| Template DNA | 50ng |
| Forward primer | 125ng |
| Reverse primer | 125ng |
| 10x reaction buffer (A1.7.7) | 5μl |
| dNTP mix | 1μl |
| *PfuTurbo* DNA polymerase (2.5U/μl) | 1μl |
| Milli-Q water | to 50μl |

PCR was performed (2.2.16), this was followed by addition of the restriction endonuclease *Dpn*I (10U) to each reaction mixture. The restriction digests were incubated at 37°C for 1h and each was analyzed (10μl) by horizontal gel electrophoresis (2.2.11). The combinations (and quantities) of Template DNA, Forward Primer and Reverse Primer used for the production of each construct are given in Table 6.6.

| Construct | Template DNA (μl) | Olignucleotide (μl) | |
|---|---|---|---|
| | | Forward | Reverse |
| pQR780 | pQR706 (1) | F434A (1.5) | F434A (1.25) |
| pQR781 | pQR776 (1) | F434A (1.5) | F434A (1.25) |
| pQR782 | pQR784 (1) | H258F (1.1) | H258F (0.55) |
| pQR783 | pQR706 (1) | H258F (1.1) | H258F (0.55) |
| pQR784 | pQR706 (1) | F434H (0.9) | F434H (2.5) |
| pQR785 | pQR706 (1) | R358A (1.5) | R358A (1.25) |
| pQR786 | pQR706 (1) | R520A (1.6) | R520A (1.8) |
| pQR787 | pQR785 (1) | R520A (1.6) | R520A (1.8) |
| pQR788 | pQR739 (1) | R358A (1.5) | R353A (1.25) |
| pQR789 | pQR739 (1) | R520A (1.6) | R520A (1.8) |

**Table 6.6** *Template/primer combinations for mutant construction*

Each reaction mixture (1μl) was initially transformed into Epicurian Coli$^R$ XL1-Blue supercompetent cells (A1.3 [50μl]). Subsequently, all constructs were transformed into *E. coli* JM107 cells to enable direct activity comparisons to be made with existing systems. All amino acid alterations were confirmed by sequencing (performed by Oswel) in both directions, sequencing primers are detailed in Appendix Two (A2.2) and the specific primers used for analysis of each construct are given in Table 6.7.

| Construct | Forward primer | Reverse primer |
|---|---|---|
| pQR780 | F434FOR | F434REV |
| pQR781 | F434FOR | F434REV |
| pQR782 | F434FOR | F434REV |
| pQR783 | H258FOR | H258REV |
| pQR784 | F434FOR | F434REV |
| pQR785 | R358FOR | R358REV |
| pQR786 | R520FOR | R520REV |
| pQR787 | R358FOR | R358REV |
| pQR788 | R358FOR | R358REV |
| pQR789 | R520FOR | R520REV |

**Table 6.7** *Mutant TK sequencing primers*

## 6.5 Expression of proteins in *E. Coli*

### 6.5.1 Production of soluble protein

1) *E. coli* strain JM107 (A1.3) was transformed with the relevant plasmid and plated on LB/amp agar. A single colony was used to inoculate LB/amp (5ml) and the cultures grown O/N at 37°C, 225rpm;
2) The cells were harvested by centrifugation at 14,000rpm for 5 minutes;
3) The cell pellet was resuspended in 1ml of Gly-Gly buffer (100mM, pH7.6 [A1.14.5]);
4) Sonication was performed using 6 x 10 second bursts at a power setting of 9, with 10 second cooling periods between the bursts;
5) The cell debris was removed by centrifugation at 14,000rpm for 10 minutes.

### 6.5.2 Protein purification by ion-exchange chromatography

1) O/N culture (200ml) was grown (IPTG added to a final concentration of 1mM). Cells were harvested by centrifugation at 6000rpm for 15 minutes;
2) Cells were washed by resuspending in 10ml Gly-Gly buffer (100mM, pH7.6) and harvested by centrifugation at 10,000rpm for 10 minutes;
3) Cells were again resuspended in 10ml Gly-Gly buffer (100mM, pH7.6);
4) Sonication was performed using 6 x 10 second bursts at a power setting of 9, with 10 second cooling periods between the bursts;
5) The cell debris was removed by centrifugation at 10,000rpm for 30 minutes and the supernatant was filtered through 0.45μm filters. TK samples were stored at 4°C (any unused samples were discarded after 2 weeks);
6) Supernatant (1ml [ca. 20mg protein]) was loaded onto a MonoQ$^R$ HR5/5 anion-exchange column. A pH gradient from 7.0 to 9.5 was run in order to remove a proportion of the cellular protein;
7) This was followed by a NaCl gradient from 0 to 0.5M at pH7.0. Fractions were analyzed using a linked-enzyme assay (2.2.1) and the BIO-RAD protein assay (2.2.2). The TK containing fraction was re-applied to the MonoQ$^R$ HR5/5 column;
8) The second stage of purification involved a NaCl gradient from 0 to 1M at pH5.0. Fractions were analyzed using the linked-enzyme assay and the BIO-RAD protein assay.

All FPLC chromatography was conducted at 22°C, with a flow rate of 0.5ml/min.

## 6.6 Biotransformations for HPLC analysis

The reaction mixture (10ml) contained the following:

| | | |
|---|---|---|
| 1. | TK (3ml O/N culture in 1ml Gly-Gly [100mM] ca.4U) | 1ml |
| 2. | TPP (final concentration 2.4mM) | 0.011g |
| 3. | $MgCl_2.6H_2O$ (final concentration 0.9mM) | 0.0018g |
| 4. | β-mercaptoethanol (10mM) | 7.8μl |
| 5. | Ketone donor (final concentration 100mM) | 0.11g |
| 6. | Aldehyde acceptor (final concentration 100mM) | Table 6.8 |
| 7. | Gly-Gly buffer (final concentration 100mM) | 8.6ml |

The quantities of substrates necessary to obtain the desired final concentration of 100mM required in the reaction mixture are given in Table 6.8.

| Compound | Mass added to reaction mixture (10ml) |
|---|---|
| β-hydroxpyruvic acid | 0.11g |
| D-glyceraldehyde | 0.09g |
| DL-glyceraldehyde | 0.09g |
| Glycolaldehyde | 0.06g |
| L-erythrulose | 120μl (0.12g) |
| 2-furancarboxaldehyde | 96μl (0.096g) |
| 2-pyridinecarboxaldehyde | 107μl |
| 3-pyridinecarboxaldehyde | 107μl |
| 4-pyridinecarboxaldehyde | 107μl |
| 2-pyrrolecarboxaldehyde | 0.095g |
| Pyridoxal | 0.203g |

**Table 6.8** *Substrates (100mM) for analytical biotransformations*

1) Prior to reaction components 1, 2, 3, 4, and 7 were mixed and incubated at 25°C for 20 minutes to ensure that a high proportion of holo-enzyme was present (4 is a reducing agent and prevents oxidation of cysteine residues involved in interactions with TPP);
2) Reactions were conducted at 37°C and aliquots (200μl) were taken at suitable time intervals;
3) Mobile phase (6mM $H_2SO_4$ [800μl]) was added to the aliquots immediately. The acid present was sufficient to quench the reactions.

## 6.7 Extraction of biotransformation products

1) Biotransformations were carried out as in section 6.6, with continuous monitoring using TLC (ethyl acetate:methanol = 95:5), $R_f$ values for each of the aromatic compounds studied are given in Table 6.9;
2) After 18h, reaction mixtures (10ml) were extracted three times with an equal volume of diethyl ether;
3) Organic solvent was removed using a rotary evaporator, the resulting mixture was then freeze dried to remove any remaining traces of water;
4) Samples were suspended in an appropriate deuterated solvent and analysed using NMR ($^1H$ and $^{13}C$) and mass spectrometry.

| Compound | Retention Factor |
|---|---|
| A. Pyridoxal | 0.44 |
| B. 2-pyridinecarboxaldehyde | 0.55 |
| C. 3-pyridinecarboxaldehyde | 0.56 |
| D. 4-pyridinecarboxaldehyde | 0.51 |
| E. 2-furancarboxaldehyde | 0.59 |
| F. 2-pyrrolecarboxaldehyde | 0.60 |
| G. TK reaction product from B | 0.72 |
| H. TK reaction product from C | 0.70 |
| I. TK reaction product from D | 0.67 |

**Table 6.9** *$R_f$ values derived from TLC analysis*

# Chapter Seven: Discussion

# CHAPTER SEVEN: Discussion

## 7.1 Introduction

During early studies involving TK the enzyme was obtained exclusively from two sources (1) *Saccharomyces cerevisiae* (Datta *et al.*, 1961) and (2) spinach (Villafranca & Axelrod, 1971). Although the synthetic potential of these sources was demonstrated on a small scale (Myles *et al.*, 1991), neither was able to yield large quantities of the enzyme. In response to this, work was carried out to develop an organism that could be used to produce greater quantities of TK for employment in larger scale biotransformations (Hobbs *et al.*, 1993).

The TK gene was cloned into pBGS18- in both orientations with respect to the *lac* promoter (French & Ward, 1995), generating pQR706 and pQR711 with the TK gene under the control of its own promoter in both constructs. Both of these expression systems yielded high levels of TK and thus a new source of the enzyme, able to produce TK in quantities sufficient to perform preparative biotransformations, had been established. These constructs were used as the starting material for the production of all recombinant TKs employed for the biotransformations described in this thesis. The research conducted during this project fell into three main categories as follows:-

- Purification of transketolase;
- Expansion of transketolase active site;
- Modification of transketolase stereoselectivity.

The findings resulting from experiments undertaken in each of these areas are discussed separately in greater detail below.

## 7.2 Purification of transketolase

The first area of work performed as part of this project, concerned experiments directed towards the development of a purification method for a modified version of TK. When conducting large-scale biotransformations it is useful if the biocatalyst can be immobilised, since this facilitates the recycling of the enzyme and so reduces wastage and increases its useful life. In an attempt to immobilise TK from a crude extract, a poly-lysine tail was attached to the 5'- end of the gene from *E. coli* (French & Ward, 1994).

Immobilisation experiments yielding promising results (French & Ward, 1994), had previously been conducted by Carol French. Although prior to this project no work relating to the purification of the protein, exploiting the characteristics associated with the poly-lysine tail, had been performed. It was decided that, since the poly-lysine tagged version of TK had already been engineered and was readily available, this construct would be used as a starting point in the formulation of a purification procedure using affinity chromatography.

The purification methodology ultimately devised (as detailed in Chapter Three) involved ion exchange chromatography, and made use of the fact that the presence of an N-terminal poly-lysine tail enables the protein to strongly bind a MonoQ$^{R}$HR5/5 column. Although a certain amount of purified TK was obtained using this method, a number of limitations were found to be associated with this approach. The first of these shortcomings concerns the time required to perform the purification, as this method takes a relatively long time to carry out when compared to other affinity chromatographic purification techniques, for example $Ni^{2+}$ - NTA chromatography Schmitt *et al.*, 1993). A further drawback was that the specific activity of the poly-lysine tagged TK was found to be significantly lower than that of the non-tagged protein derived from pQR706. One possible reason for the reduction in activity is the formation of inclusion bodies of the over-expressed tagged protein. The protein thus produced exhibits significantly inferior activity since it is no longer in the optimal quaternary structure to effectively perform its catalytic function. Alternatively, enzymatic activity may be reduced in this construct for other reasons. The tail may, for example, alter the local pH or ion concentration around the N-terminus and this could have effects extending out to the active site ultimately resulting in a reduction in activity.

Considering all of these factors, it was concluded that the effectiveness of this particular technique for the purification of TK was not as great as had been originally planned. The additional TK mutants generated during this project (as detailed in Chapters Four and Five) were not constructed into the poly-lysine system, and crude extracts were in fact used for all subsequent characterisation work. Although, as was discussed during the relevant chapters, the use of crude enzyme extracts was found to have its own associated problems, particularly for those reactions where monitoring using HPLC techniques was required. These observations, which leant weight to the

argument that a pure form of the enzyme would be preferable, are discussed later in this chapter.

In light of the difficulties encountered using the poly-lysine tagged protein for FPLC purification, and of those associated with the use of crude enzyme extracts for TK biotransformations, it may be beneficial during future research to look into other potential purification methods for the enzyme. One such possibility would be to construct a C-terminal poly-histidine tagged version of TK, however it should be noted that the C-terminus is not projected away from the protein surface as is the case with the N-terminus. Thus, although it is still some distance from the active site (Littlechild *et al.*, 1995), a C-terminal tag may in some additional way interfere with the enzymatic function. Another option would be to attempt to clone the TK gene into a commercial histidine-tag vector.

A different type of approach would be to devise a procedure to purify the untagged protein. One strategy to consider would be the development of a sugar-mimic type column, which could be tailored to suit the binding environment of the active site of TK. This type of column would have an additional advantage in that it could be adapted in order to enable the selective binding of enzymes with particular stereoselectivities. For example during the type of work described in Chapter Five, in which mutant enzymes were constructed designed to have modified stereoselectivities, it would be extremely useful to have a technique whereby an enzyme specific for a particular enantiomer could be isolated at the purification stage. This could be achieved by incorporating sugar-like ends onto the column matrix with either a D- or an L- hydroxyl group at the key stereocentre, so enabling only enzymes specific for that particular enantiomer to bind the column and thus be purified.

### 7.3 Expansion of transketolase active site

Commonly enzymes can achieve transformation not only of the substrate or substrates employed *in vivo*, but will (*in vitro*) accept a wide selection of substrates and transform these with the same regio- and stereoselectivity (Jones & Beck, 1976; Cannarsa, 1996). Providing substrates comply with certain physical and chemical criteria dictated by the conformation of the active site of the enzyme, conversion of a wide range of potentially useful compounds may be accomplished. A variety of enzymes, mostly lyases and aldolases, have so far been used to synthesise complex sugars, sugar analogues, and

other biologically important natural compounds (Toone *et al.*, 1989; Wong & Whitesides, 1994). In particular, thiamine-dependent enzymes have been used for some time as catalysts in chemoenzymatic syntheses (Faber, 1992; Krampitz, 1969).

*E. coli* transketolase (E.C.2.2.1.1) is a homodimer comprising subunits of molecular mass 72kDa with the two subunits combining to form a molecule with a $C_2$ axis of symmetry. In terms of previous work conducted specifically relating to the generation of mutant transketolases with extended substrate ranges such as those described in this thesis, not a great deal of research has been undertaken in this field. However, considerable work has been performed towards the elucidation of the crystal structures of TK from various sources including those from *S. cerevisiae* (Schneider *et al.*, 1989; Nikkola *et al.*, 1994) and *E. coli* (Littlechild *et al.*, 1995). These crystal structures are extremely useful as they reveal the identity of amino acid residues at the dimer interface, and hence those located around the active site. Information derived from these structures also enables predictions to be made as to the nature of interactions between TK and its cofactors, TPP and $Mg^{2+}$ ions (Lindqvist *et al.*, 1992; Meshalkina *et al.*, 1997), and of those with its substrates.

In addition, a substantial amount of research has been performed involving mutagenesis within the active site, incorporating amino acid changes designed to investigate the interactions of TK with TPP and $Mg^{2+}$ ions and to obtain mechanistic information regarding the TK catalysed transformations. Crystal structures of the apo- and holo-enzyme (Sundstrom *et al.*, 1992; Lindqvist *et al.*, 1992; Nikkola, 1994), and those of various complexes involving TK with TPP analogues (Nilsson *et al.*, 1993; Konig *et al.*, 1997) are also available. Furthermore, the crystal structures of a number of site-directed mutants have been determined, providing additional insights into the details of interactions between TK and its cofactors (Wikner *et al.*, 1997).

Other studies have focused on the transformation potential of wild-type TK. It has been found that the native enzyme is active towards a range of aldehydes (Effenberger *et al.*, 1992; Kobori *et al.*, 1992; Hobbs *et al.*, 1994), although activity is very limited in many cases, particularly as the length of the carbon chain increases and as more bulky side chains are incorporated into the structures. However, this would probably be expected since all of this work involved unmodified enzymes, and not ones specifically engineered to be reactive towards these novel substrates. Chapter Four of this thesis

outlines work conducted towards the design, generation and characterisation of a range of TK mutants. Incorporated into these mutants are amino acid changes intended to cause expansion of the active site funnel and thus permit entry and accommodation of more bulky substrates. Amino acid substitutions were made at locations in which residues were believed to play a key role in substrate and/or cofactor binding.

Molecular modelling information, combined with a knowledge of factors relating to the size, charge and hydrophilic nature of the amino acid side chains, enables predictions to be made regarding the roles of particular amino acid residues in catalysis and substrate binding. The first residue chosen for investigation during this project was I189 positioned roughly midway along the active site funnel, its large hydrophobic side chain can be observed to point towards the TPP cofactor. Previous work has reported that the main chain oxygen of this residue is involved in the binding of the $Mg^{2+}$ ions (Schneider & Lindqvist, 1998). During the linked-enzyme assays performed on TK produced from pQR739 [I189A] and pQR749 [I189V] significantly reduced activity towards the open chain sugars was found. This may be attributable to the less efficient binding of the $Mg^{2+}$ ions in these constructs, leading to a more unstable and hence less catalytically effective protein dimer.

Another amino acid residue selected for alteration was S385* which is situated close to the active site entrance. It has been reported that the hydroxyl side chain of this residue may be involved in hydrogen bonding interactions with the hydroxyl moiety of the sugar phosphate group (Schneider & Lindqvist, 1998). During these studies it was found that the mutant produced from pQR735(L) [S385*L] exhibited significantly diminished activity towards the native sugar substrates. This decline may have been due to a loss of interactions, existing for the wild-type protein, involving the serine hydroxyl moiety and the phosphate group of these sugars. This could result in less efficient binding of these particular substrates and hence a deterioration of the transformation efficiency.

Also replaced in the original set of TK mutants, was residue H461*. This amino acid is located close to the entrance of the active site opposite two arginine residues R358* and R520*, together they create an overall positive charge at the active site opening. It may be the case that this residue is also a constituent of the phosphate-binding site described by Lindqvist *et al.*, 1992. The enzyme variant incorporating the amino acid alteration

H461*L was found to have reduced activity with the native sugar substrates and this may be due, in part, to a loss of stabilising interactions at the active site entrance involving the phosphate groups of these sugars.

In the case of recombinant TK generated from pQR738(S), in which residue H461* has been replaced by serine, diminished activity in the linked-enzyme assay was also observed. However, the decrease in reactivity was not found to be so great as when this residue is replaced by leucine. Here replacement of the positively charged histidine with the hydroxyl group of serine may not completely destroy the hydrogen bonding interactions, in fact they may now occur with the phosphate oxygens rather than the hydrogen atoms. An additional factor contributing to the reduction in activity for the H461*L construct may be that the presence of the hydrophobic side chain of leucine, in this usually positively charged region, could result in the introduction of unfavourable interactions at the active site entrance.

The final amino acid to be substituted as part of the initial mutagenesis work was H258. This residue was selected for alteration since it was initially believed to be located in the region of the active site thought to be involved in determining the stereoselectivity of the enzyme. It was understood that the positively charged side chain of this residue is involved, through hydrogen bonding interactions, in the partial sharing of pairs of electrons with the α-hydroxyl group of the aldehyde acceptor. Coupled with the effects of the bulky hydrophobic side chain of F434*, this was considered to potentially play an important role in determining the stereoselectivity of TK. Enzymes produced from pQR776 [H258A] and pQR748 [H258D] both exhibited reasonably high activity during initial characterisation studies using the linked-enzyme assays. It was therefore concluded that replacement of this residue, in the manner described, did not appear to have significant deleterious effects on the active centre of TK regarding its ability to transform the native sugar substrates.

In summary, the preliminary characterisations of these initial TK mutants, conducted using a linked-enzyme assay, showed that all variant enzymes displayed a reduced activity towards the native sugar substrates when compared to that of the wild-type protein. These findings would be expected, since the amino acid substitutions incorporated into these proteins would cause the active site to deviate from its optimal structure for accommodation and transformation of these compounds. The

conformational changes may cause a reduction in the binding efficiency of cofactors TPP and $Mg^{2+}$ ions, thus leading to less stable, and hence less reactive, protein dimers. Alternatively they may result in a loss of important bonding interactions with the sugar substrate thus reducing transformation efficiency, or conversely they may introduce certain unfavourable interactions with the cofactors or substrates with a consequent decline in reactivity.

If β-hydroxypyruvic acid (β-HPA) is employed as the ketol donor during TK transformations, $CO_2$ is lost during the conversions ensuring that reactions are irreversible and hence go to completion (Srere *et al.*, 1958). This is an important reaction characteristic when conducting biotransformations for detailed analysis, since it facilitates the accurate monitoring of reagent concentrations. This allows, in theory, a direct correlation to be made between the rate of disappearance of β-HPA and the extent of enzymatic conversion. For this reason, β-HPA was employed as the ketol donor for all biotransformations involving novel aldehyde substrates performed during this project.

Initial investigations into the synthetic potential of the modified *E. coli* transketolases, with a selection of water-soluble aromatic compounds employed as aldehyde acceptors, were conducted. Reactions were monitored using the HPLC techniques described by Mitra & Woodley, 1996, by following the loss of β-HPA due to enzymatic conversion. For recombinant TK from pQR739 and pQR749, the observed increase in activity towards a number of these more bulky aldehydes may result from the opening up of their active site funnels. This could possibly cause a reduction in the extent of unfavourable hydrophobic interactions with these bulky compounds, outweighing any decline in activity resulting from an increased instability of these proteins also caused by the amino acid substitutions. Since the TK produced from pQR739 [I189A] (French & Ward, 1996) consistently appeared to exhibit enhanced reactivity towards the novel aromatic aldehydes (relative to the corresponding reactions of the wild-type protein), this mutant was selected for more in depth investigations involving structural characterisation of product molecules using NMR and mass spectrometry techniques.

Reactions of TK from pQR739 with pyridoxal, 2-furancarboxaldehyde, and 2-pyrrolecarboxaldehyde afforded no observable reaction products upon NMR analysis. For pyridoxal this could be due to the large number of functional groups attached to the

pyridine ring with the result that this compound is, still, too sterically hindered to be accommodated within the expanded active site funnel. For reactions with 2-furancarboxaldehyde and 2-pyrrolecarboxaldehyde, the lack of reactivity may be attributable to the electron-donating ability of the heteroatom in the rings of these compounds (see Figure 4.19). As electron density is pushed onto the carbonyl carbon atom, this leads to a reduction in the electrophilicity at this centre, thus making these two substrates less efficient acceptors in the nucleophilic addition reactions catalysed by TK.

It is important to note that a number of problems were encountered when obtaining data from the preliminary HPLC assays used to monitor these biotransformations. For example, reactions of TK produced from pQR739 with pyridoxal, 2-furancarboxaldehyde, and 2-pyrrolecarboxaldehyde yielded no observable reaction products upon NMR analysis. However, for all three of these compounds some degree of β-HPA loss was found to occur during HPLC analysis (see Figures 4.5 and 4.6). This may have been attributable to the fact that insufficient product material was produced, when performing reactions on the scale employed during these studies, to enable their detection within the limits of the NMR techniques available. Another possible explanation may have involved a side reaction (catalysed by one of the other enzymes present within the crude enzyme extract) occurring within the biotransformation mixtures used for HPLC analysis, it could have been this transformation that was causing the degradation of β-HPA rather than the expected TK conversion. Alternatively, some other constituent of the crude enzyme extract may have in some way interfered with the extraction process, thus making it a less effective procedure for these particular compounds. As discussed at the end of Chapter Four, development of a method for purification of the enzyme would be one possible way to circumvent these latter two problems. A number of alternative approaches towards the purification of TK were considered in the discussion at the end of Chapter Three. As a further option it would also be useful to look into the possible use of flash chromatography, in order to purify biotransformation products before they are submitted for analysis using NMR and mass spectrometry.

Control reactions were not performed in the absence of the aromatic aldehydes (but in the presence of all other reaction components) and it may have been that some other compound or enzyme within the crude enzyme extract was causing the β-HPA to break

down. Further transformations, with implementation of these additional controls, would need to be conducted during future work in order to either eliminate this as a possibility or to quantify its effects. This would involve performing reactions in which all reaction components (except for the aldehyde acceptor) were present, and monitoring the mixtures using HPLC to look for any background degradation of β-HPA. Despite its current limitations as a quantitative analytical tool, this technique did still prove to be a useful screening technique for highlighting interesting mutant/aldehyde combinations to be used for subsequent more detailed structural investigations.

From the initial HPLC screening experiments, the mutant TK produced from pQR739/JM107 appeared to exhibit enhanced reactivity towards a number of the aromatic aldehydes. This mutant has replacement of amino acid residue I189 with an alanine, with a resultant expansion of the active site funnel. The cavity thus produced appears to facilitate the entry of more bulky aldehyde compounds, which are ineffective substrates with the wild-type protein. Especial reactivity for this new mutant was found in the case of the pyridinecarboxaldehydes. Reaction products from these successful biotransformations were extracted, and structurally characterised using NMR and mass spectrometry techniques to confirm that the expected TK conversion had in fact taken place.

An important point to be considered, when performing these particular biotransformations, is that it may not be possible to control the stereochemistry at the new chiral centre generated. This is a result of the potential for inter-conversion between the two possible isomers of the product, occurring via an anionic intermediate which can be readily stabilised by the adjacent aromatic ring (see Chapter Four, page 87). Nevertheless, the results obtained from structural characterisation experiments on these reaction products were very promising. *E. coli* TK had previously been found to exhibit activity with a range of different aldehyde acceptors (Hobbs *et al.*, 1994), although no reactivity had been reported during transformations with cyclic or aromatic compounds (see Figure 1.10). In addition, structures of reaction products from conversions of aromatic and heterocyclic substrates with TK, from any source, had not previously been confirmed (Schorken & Sprenger, 1998). As a further step it would be interesting in future work to attempt to scale up these successful transformations in an attempt to convert these compounds on a preparative scale.

An additional set of five TK mutants were also engineered during this project, these constructs incorporated replacement of the pair of arginine residues (R358* and R520*) located at the active site entrance. The anticipated effects of substituting these residues with alanines was the further facilitation of the entry of bulky compounds into the active site funnel, it was also considered that these changes may result in destruction of part of the phosphate-binding site thought to be located in this area. The location of the phosphate-binding site at the entrance of the channel is important since it makes it possible for substrates to reach the active site irrespective of their chain length. The interactions of the enzyme with the substrate phosphate group are maintained by use of the two arginine residues which, through their flexible side chains, can adjust to these variations in the position of the phosphate group (Lindqvist *et al.*, 1992). Previous work performed on the yeast enzyme by Schneider & Lindqvist, 1998 has shown that replacement of these residues by alanine yields mutant enzymes with considerable residual activities, but with large increases in the Km values for phosphorylated substrates. This is consistent with the described role of these residues in binding the phosphate group of substrates, and also provides a molecular explanation as to why non-phosphorylated sugars are generally poor substrates for wild-type TK.

For all of these new TK constructs some decline in activity towards the native sugar substrates was observed, this may be attributed, in part, to the destruction of the phosphate-binding site located at the active site entrance. This would be particularly true for the enzyme produced from pQR787 in which both of the arginine residues have been replaced by alanine. For all five constructs, these alterations may be resulting in a loss of bonding interactions with the phosphate groups of the sugars and so leading to less efficient binding of these substrates. Additionally, enzymes produced from pQR788 and pQR789, having replacement of residue I189 with an alanine, may be experiencing less efficient binding of the $Mg^{2+}$ ions leading to a more unstable protein dimer and consequently even further diminished activity.

An attractive area to explore as part of future work would be to compare the reactivities of these mutant transketolases towards a series of phosphorylated sugars and the corresponding non-phosphorylated analogues. This would provide an insight into the importance of the phosphate-binding site in determining the activity levels of TK, and also the relative importance of each of the arginine residues as contributors to this site. In addition, it would be useful to investigate the reactivities of these new constructs

with a wider range of aldehydes, involving for example direct examination of reactions with the aromatic aldehydes shown to be active with the initial set of TK mutants. As was done previously, preliminary screening using HPLC would be required in order to highlight potentially useful mutant/aldehyde combinations. Products from these transformations would then require extraction and structural characterisation as before.

As an extension to this work it would be useful to develop variants of the enzyme, with substitution of amino acid residues on the protein surface, which are capable of exhibiting increased tolerance towards organic solvents. Enzymes of this type would be extremely useful as they would open up the possibility for transformation of a much wider range of aldehyde acceptors, including for example substrates such as thiophenes or indoles which are not soluble in aqueous solution. A number of techniques may be employed in order to generate such proteins, these would include random mutagenesis and DNA shuffling experiments. However, a method of screening new mutants, with the ability to highlight promising TK constructs generated by these means, would need to be developed alongside this approach.

In addition, it would be interesting as part of future work to make further amino acid alterations using site-directed mutagenesis. This could include, for example, the generation of mutants in which the active site funnel is expanded even more, thus allowing entry of still larger substrate molecules. Mutants in which residues such as F434*, F437* and L382*, which all have fairly large side chains pointing into the active site channel, have been replaced with alanines may exhibit these sorts of characteristics. Enzyme variants incorporating each of these substitutions along with the alteration I189A, may have a sufficiently widened active site to enable the entry of molecules such as indoles or pyridoxal which were previously found not to be reactive.

Further work relating to the stereoselectivity of the enzyme would also be a useful area to pursue. Mutants engineered with the changes F434*A and H261A would be expected to have considerable relaxation of the requirement for D-enantiomeric substrates, whereas a variant incorporating the substitutions F434*H and H261F ought to show a preference for substrates containing an α-hydroxyl group of the L-configuration. These anticipated alterations in the phenotype are derived from information provided by the latest refined 3-D structures of the active site and are discussed further in Section 7.4. Additionally, it can be seen from the enzyme crystal

structure that the side chain of residue D469* also points into the active site funnel in the region believed to be occupied by the α-hydroxyl group. The negatively charged side chain of D469* may, therefore, also be involved in hydrogen bonding interactions with the α-hydroxyl group of the substrate and thus play a role in determining the stereoselectivity of the enzyme. This residue would consequently be another important one to investigate during future mutagenesis work.

As mentioned earlier, it would be of interest to conduct further research into the precise nature of the phosphate-binding site of TK, and thus the amino acid residues contributing to the binding of the substrate phosphate group. From the crystal structure it can be seen that four positively charged residues H461*, R358*, R520* and H473*, along with the negatively charged S385*, are all situated in the region of the phosphate-binding site at the entrance to the active site funnel. These residues may all in some way, interact with the phosphate moiety through hydrogen bonding interactions. One possibility during future experiments would be to replace each of these residues with alanines, with the aim of a complete destruction of this site, and to monitor comparative reactivities of a selection of phosphorylated and non-phosphorylated compounds. Also of interest would be an investigation into the implications of these changes on the Km values for a range of phosphorylated substrates, when compared to their Km values with the native protein.

Another area to investigate further would involve the replacement of a number of charged residues within the active site with hydrophobic residues such as valine or leucine. Although hydrophobic, the side chains of these residues would not be too bulky so as to severely block the entry of substrates into the channel, as would be a potential problem with substitution using an amino acid such as phenylalanine. Replacement of residues such as H258, H261, S385* and D469*, all containing charged side chains which point into the active site funnel, would be interesting choices to alter in this way. These types of modifications to the active site funnel may permit the entry, and thus the transformation, of relatively hydrophobic compounds such as the thiophenes which have previously exhibited no reactivity with TK. Although additional alterations to the protein surface, increasing the solvent tolerance of the enzyme, may also be required in order to achieve high levels of reactivity towards these types of non-water soluble compound.

Finally, it may be worthwhile looking into the consequences of introducing a greater number of charged moieties into the active site. For example replacement of residues, containing side chains pointing into the active site, with amino acids such as serine, aspartic acid, or glutamic acid would result in an overall negatively charged funnel region and may enhance reactivity towards substrates containing amine groups. Reactions involving these types of compounds are important for the production of many pharmaceutical intermediates, as these structures are often amine containing. Conversely, substituting such residues with histidine, lysine, or arginine would increase the overall positive charge within the active site funnel, and thus may increase the transformation efficiency of compounds containing, for example, acid groups.

If a specific substrate were to be identified as especially important in the production of a particular pharmaceutical intermediate then it may be possible, by considering all of the above factors, to engineer a version of the enzyme possessing an active site specifically tailored to provide optimal binding and hence maximum transformation efficiency for that particular compound. Features such as the size, charge, and solvent preference of the substrate would all need to be taken into account when designing such an enzyme.

Another extension of the research already carried out on TK would be to look into the possibility of two-step biotransformations. For example, the TK transformation product could, itself, be the substrate for a second enzyme. One potential candidate for such reactions would be transaminase, the second step in this particular case would be the addition of a chiral amine group onto the TK product. This approach would be useful for the generation of molecules, which could be employed as intermediates in the production of compounds, with structures similar to those found in a number of pharmaceutical compounds such as serotonin and chloramphenicol.

## 7.4 Modification of transketolase stereoselectivity

TK has previously been shown to be effective for the transformation of a wide range of phosphorylated and non-phosphorylated monosaccharides acting as donor and acceptor substrates (Usmanov, 1983; Kobori, 1992), a characteristic making the enzyme a useful tool for stereospecific synthesis of carbohydrates. However, one of the limitations associated with use of the wild-type enzyme as shown by Demuynck *et al.*, 1991, was that only sugar phosphates with a D-*threo*-configuration at the C-3 and C-4 carbon

atoms can act as efficient donors. This high stereoselectivity of the enzyme for the donor substrate was explained by Nilsson, 1997, and can be understood from the observed pattern of hydrogen bonds formed by the substrate at the active site. The low reaction rate found for DL-glyceraldehyde by Hobbs *et al.*, 1994 (see Figure 1.10) further highlights the enzymes' preference for α-hydroxyaldehydes of the D-configuration. Some reports have, however, depicted the transformation of simple α-unsubstituted aldehydes (Hobbs *et al.*, 1993; Morris *et al.*, 1996), although in each of these cases significantly lower conversion rates were observed than for the hydroxylated compounds. Chapter Five details the work conducted during this project towards attempts to generate a series of TK variants, each designed to have a slightly different modification of the wild-type stereospecificity. The stereoselectivities of these modified enzymes were investigated via transformations of both enantiomers of the sugar glyceraldehyde.

Residues believed to be acting to orient the substrate, allowing exclusive binding of molecules with an α-hydroxyl group of the correct stereochemistry, have been highlighted by molecular modelling experiments. The residues H258 and F434*, located midway along the active site funnel on opposite faces and pointing towards each other, were initially believed to play an important role in determining the stereoselectivity of the enzyme. This was thought to occur via a combination of hydrogen bonding interactions between the α-hydroxyl group and the side chain of H258, in addition to the steric effects resulting from having the bulky hydrophobic phenyl group of F434* situated in this particular location. To further investigate the precise roles of these residues, and to attempt to generate a version of TK exhibiting reversed specificity, these amino acids were altered using site-directed mutagenesis. The expected outcome of these changes was that a mutant having reversal of these two key residues would exhibit the opposite specificity to that of the native enzyme. Whereas a variant in which H258 and F434* have been replaced with alanines should show considerable relaxation of the requirement for D-enantiomeric compounds, and would thus be anticipated to exhibit an increased reactivity towards a racemic mixture of a particular substrate.

As a first step, the TK mutants were screened to determine their relative activities towards the model substrate glycolaldehyde. As was the case for work involving transformation of the aromatic aldehydes described in Chapter Four, β-HPA was again

used as the ketol donor. Initially reactions were monitored using the HPLC assays, by following the loss of glycolaldehyde. Examination of results from these assays showed that all mutant transketolases exhibited reduced activity towards this model substrate, when compared to the reactions of the wild-type protein. These observations were not unexpected since glycolaldehyde resembles the native sugar substrates of TK. The amino acid substitutions present in these new TK constructs cause the active site structure to deviate from its optimal conformation for transformation of these types of compound, thus reducing the conversion efficiency.

As a consequence of the alterations S385*L and H461*L, with the accompanying introduction of the hydrophobic side chain of leucine close to the entrance of the active centre, this may result in unfavourable interactions with the substrate as it enters the active site funnel. For the mutant with alteration H461*S, as there is no phosphate group present in glycolaldehyde, the potential gain in hydrogen bonding interactions which accompany introduction of the serine residue for the native sugar substrates is not present here. In addition the serine hydroxyl group may act to repel the hydroxyl groups of this substrate. In the case of I189, its side chain is more significantly involved in interactions with the cofactor rather than with substrate binding. Its replacement with alanine or valine results in less efficient binding of $Mg^{2+}$ ions and hence less effective binding of TPP, so giving a less stable and thus potentially less reactive enzyme dimer. Substitution of H258 with alanine or aspartic acid appears to have the least detrimental effect on TK reactivity with glycolaldehyde. Since this histidine residue was believed to be involved in stereospecific hydrogen bonding interactions with the α-hydroxyl moiety of substrates, and this particular group is not attached to a chiral centre in glycolaldehyde, this alteration was not expected to have an especially large deleterious effect on the transformation efficiency for this compound.

As was the case for work relating to expansion of the active site described in Chapter Four, a number of problems were encountered when performing the HPLC analysis of these reactions, most noticeable was the lack of mass balances found during these assays (this was also true of assays for transformations discussed later in this chapter involving the two enantiomers of glyceraldehyde). Although on the whole, the loss of β-HPA was closely mirrored by the corresponding loss of glycolaldehyde, and periods of β-HPA/glycolaldehyde loss were generally accompanied by some degree of L-erythrulose increase, these changes were not found to be completely stoichiometric as

would be expected. This raised some concern as to whether accurate rate data could in fact be obtained using these assays. These findings concur with other work performed using this particular assay (Mitra, 1997) in which the loss of reaction substrates, not resulting from TK conversions, was observed to result in incomplete mass balances. Alternative reaction conditions were subsequently employed by Mitra, with biotransformations conducted in the absence of buffers utilising acid/alkali titration via a pH-stat. This would perhaps be a useful methodology to employ during any further work involving these types of assays, since the modified transformation conditions were found to result in a reduction in the undesired degradation of reaction substrates.

Another alternative would be to use the D- and L-enantiomers of a compound other than glyceraldehyde. For example using substrates with a longer R group, which should still be detectable using the HPLC assay previously employed, but which should not be as susceptible to unwanted side reactions catalysed by other enzymes. This is the case since the further the structures deviate from that of the native substrate, the less likely it is that they will be transformed by competing reactions. A decline in the occurrence of these competing reactions, along with a reduction in the amount of substrate degradation resulting from the buffer conditions employed for biotransformations, would ensure that any substrate losses observed during the HPLC assays more accurately reflected the extent of their conversion by TK.

In order to eliminate completely the possibility of competing side reactions by other enzymes, and also the potential reduction in efficiency for extractions of reaction products by other components of the crude enzyme extracts, a pure from of TK would be desirable. This again suggests that it would be beneficial in future work to devise an effective purification technique for TK. The sugar-mimic columns described earlier in this chapter would be extremely useful in the case of these reactions, since columns tailored to purify enzymes with specific stereoselectivities could potentially be used to yield pure TK which is reactive exclusively towards either the D- or the L-enantiomer of a particular compound.

Once a suitable level of enzymatic activity towards glycolaldehyde had been established, work was then undertaken to investigate the stereoselectivities of the new TK constructs by using them to transform D-glyceraldehyde and DL-glyceraldehyde. As before reactions were initially monitored using HPLC techniques, by measuring the

loss of the aldehyde component. Analysis of the data obtained during these assays highlighted the fact that the enzyme produced from pQR776 [H258A] appeared to have the greatest potential activity towards the L-enantiomer. Hence this particular construct was selected for further biotransformations, accompanied by more detailed analysis of product structures using NMR techniques. Introduction of the much smaller alanine residue, also lacking hydrogen bonding ability, into this region was anticipated to have these sorts of effects, since it results in the introduction of a cavity into the region of the active site believed to be involved in binding the α-hydroxyl group of the substrate.

The theory that residues H258 and F434* played an important role in determining the stereoselectivity of TK appeared to be supported by the experiments conducted on the TK mutant from pQR776, as a result a further set of five TK mutants were engineered. These new constructs comprised a series of enzyme variants designed to exhibit a gradual relaxation of the requirement for a D-enantiomeric α-hydroxyl group. This new enzyme range culminated in the production of pQR782, which yields a TK variant in which these two key residues have been reversed. Although the linked-enzyme assay data obtained for these constructs provides an insight into the differing behaviour of these enzymes towards the native D-enantiomeric sugar substrates, direct examination using the individual enantiomers of a specific compound is needed. As discussed earlier it would perhaps be advisable to try an alternative substrate, possessing a longer carbon chain, for these investigations. If non-purified enzymes are used these alternative substrates may be less susceptible to competing reactions which could distort any subsequent HPLC data obtained.

As the initial crystallographic data available during this work indicated that residues H258 and F434* played important roles in determining the stereoselectivity of TK, this strongly influenced their selection for mutagenesis. However since the mutant construction experiments were undertaken, a much higher resolution structure has become available. This new structure suggests that the residue H258 is in fact further away from its initially believed position within the active site funnel, although the residue F434* remains in the same place as in the originally available 3-D structure. As a result of this development, the reasons for the altered phenotype of mutants having replacement of H258 with alanine and with phenylalanine may not be as simple to explain as was previously thought. As this residue has now been placed further away from the region occupied by the α-hydroxyl group, a conclusive explanation for the

observed reactivity differences is not currently available. A further observation from the latest refind TK active site structure is that another histidine residue, H262, can now be seen to be located within the active site funnel in the proximity of the α-hydroxyl group. This residue would be an interesting one to alter in subsequent work, initially replacing it with an alanine or a phenylalanine and investigating any resultant modifications to the stereoselectivity of the enzyme.

## 7.5 Conclusion

In conclusion, although considerable time was expended during this project towards the development of a purification methodology for TK, an effective purification technique was not ultimately devised. During subsequent experiments it was discovered that the use of crude enzyme extracts for performing TK biotransformations was not in fact desirable. The presence of other extract components resulted in potential problems concerning the detection of substrates using HPLC assays, and also with the extraction of reaction products into organic solvents. It also leads to the possibility of competing reactions, which can distort the HPLC data obtained. It would, therefore, be advantageous as part of future TK work in this area to investigate the effectiveness of alternative purification techniques for the enzyme.

Promising results were obtained upon investigation of the reactions of a selection of TK active site mutants with a range of aromatic aldehyde acceptors. One mutant in particular was especially efficient in the transformation of these compounds, and NMR and mass spectrometry techniques were used to confirm the structures of reaction products from these conversions. Information derived from characterisation experiments involving these mutant enzymes, along with that provided by 3-D crystal structures of the active site, can be used as a starting point for the design of additional mutants, modelled in such a way as to enhance their synthetic potential even further.

Research conducted into the modification of the stereoselectivity of TK also showed some promising preliminary results, including the potential relaxation of the requirement for a D-enantiomeric substrate by one of the mutants generated. Further to this a whole series of additional mutants were engineered, each designed to exhibit a slightly different modification of the wild-type stereoselectivity. More detailed characterisation of these new constructs will provide valuable information on which to base the design of future TK variants in this series.

# References

Ahmar, M., Girard, C., Bloch, R., 1989, "Enzymatic resolution of methyl 2-alkyl-2-arylacetates", *Tetrahedron Lett.*, **30**, 7053-7056.

Alder, U., Breitgoff, D., Klein, P., Laumen, K., Schneider, M., 1989, "Enzymatic ester hydrolysis and synthesis – Two approaches to cycloalkane derivatives of high enantiomeric purity", *Tetrahadron Lett.*, **30**, 1793-1796.

Ariens, E. J., 1986, "Stereochemistry: A source of problems in medicinal chemistry", *Med. Res. Rev.*, **6**, 451-466.

Ariens, E. J., Wuis, E. W., Veringa, E. J., 1988, "Stereoselectivity of bioactive xenobiotics", *J. Biochem. Pharmacol.*, **37**, 9-18.

Beckett, A., 1991, "Chirality and its importance in Drug Development", *Biochem. Soc. Trans.*, **19**, 443-446.

Boiteux, A., Hess, B., 1981, "Design of glycolysis", *Phil. Trans. R. Soc. Lond. B.*, **293**, 5-22.

Bolte, J., Demuynck, C., Samaki, H., 1987, "Utilisation of enzymes in organic chemistry: transketolase catalysed synthesis of ketoses", *Tetarahedron Lett.*, **28** (45), 5525-5528.

Booth, C. K., Nixon, P. F., 1993, "Reconstitution of holotransketolase is by a thiamin-diphosphate-magnesium complex", *Eur. J. Biochem.*, **218**, 261-265.

Bradford, M. M., 1976, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding", *Anal. Biochem.*, **72**, 248-254.

Braus, G. H., 1991, "Aromatic Amino Acid Biosynthesis in the Yeast *Saccharomyces cerevisae*: a model system for the regulation of a Eukaryotic Biosynthetic Pathway", *Microbiol. Rev.*, **55**, 349-370.

Brieger, G., Bennett, J. N., 1980, "The Intramolecular Diels-Alder Reaction", *Chem. Rev.*, **80**, 63-97.

Brown, J. M., Davies, S. G., 1989, "Chemical asymmetric synthesis", *Nature*, **342**, 631-636.

Cannarsa, M. J., 1996, "Single enantiomer drugs: new strategies and directions", *Chemistry & Industry*, 20th May 1996, 374-378.

Cartwright, A. C., 1991, "Regulatory implications: an authority's perspective", *Biochem. Soc. Trans.*, **19**, 465-467.

Cavalieri, S. W., Neet, K. E., Sable, H. Z., 1975, "Enzymes of pentose biosynthesis. The quaternary structure and reacting form of transketolase from bakers' yeast", *Arch. Biochem. Biophys.*, **171**, 527-532.

Cayen, M. N., 1991, "Racemic mixtures and single isomers", *Chirality*, **3**, 94-98.

Corbett, M. D., Chipko, B. R., 1980, "Comparative aspects of hydroxamic acid production by thiamine-dependent enzymes", *Bioorg. Chem.*, **9**, 273-287.

Corbett, M. D., Corbett, B. R., 1986, "Effects of ring substituents on the transketolase-catalysed conversion of nitroso aromatics to hydroxamic acids", *Biochem. Pharm.*, **35**, 3613-3621.

Crosby, J., 1991, "Synthesis of optically active compounds – a large scale perspective", *Tetrahedron*, **47**, 4789-4846.

Datta, A. G., Racker, G. A., 1961, "Mechanism of action of transketolase", *J. Biol. Chem.*, **236** (3), 617-623.

Davies, H. G., Green, R. H., Kelly, D. R., Roberts, S. M., 1989, In: *Biotransformations in Preparative Organic Chemistry*, Academic Press, London.

Demuynck, C., Bolte, J., Hecquet, L., Samaki, H., 1990, "Enzymes as reagents in organic chemistry: transketolase-catalysed synthesis of D-[1,2-$^{13}C_2$]xylulose", *Carbohydr. Res.*, **206**, 79-85.

Demuynck, C., Bolte, J., Hecquet, L., Dalmas, V., 1991, "Enzyme-catalysed synthesis of carbohydrates: synthetic potential of transketolase", *Tetrahedron Lett.*, **32** (36), 5085-5088.

Draths, K. M., Frost, J. W., 1990, "Synthesis Using Plasmid-Based Biocatalysis: Plasmid Assembly and 3-Deoxy-D-*arabino*-heptulosonate Production", *J. Am. Chem. Soc.*, **112**, 1657-1659.

Draths, K. M., Pompliano, D. L., Conley, D. L., Frost, J. W., Berry, A., Disbrow, G. L., Staversky, R. J., Lievense, J. C., 1992, "Biocatalytic synthesis of aromatics from D-glucose: the role of transketolase", *J. Am. Chem. Soc.*, **114**, 3956-3962.

Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., Jordan, F., 1993, "Catalytic centers in the thiamine diphosphate dependent enzyme pyruvate decarboxylase at 2.4A resolution", *Biochemistry*, **32**, 6165-6170.

Effenberger, F., Null, V., Ziegler, T., 1992, "Preparation of optically pure L-2-hydroxyaldehydes with yeast transketolase", *Tetrahedron Lett.*, **33** (36), 5157-5160.

Egan, R. M., Sable, H. Z., 1981, "Transketolase kinetics", *J. Biol. Chem.*, **256** (10), 4877-4883.

Faber, K., 1992, In: *Biotransformations in Organic Chemistry*, Springrt Verlag, Berlin, 204-220.

Faber, K., Franssen, M. C. R., 1993, "Prospects for the increased application of biocatalysts in organic synthesis", *TiBtech.*, **11** (11), 461-470.

Ferringa, J., Strijveen, P, Kellogg, C., 1986, "E. E. determination without chiral auxillary compounds", *J. Org. Chem.*, **51**, 5484-5486.

Fraenkel, D. G., 1987, "Glycolysis, pentose phosphate pathway, and Entner-Doudoroff pathway". In: Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., Umbarger, H. E., (eds) *Escherichia coli* and *Salmonella typhimuriumi* cellular and molecular biology, vol 1. American Society for Microbiology, Washington DC, 142-150.

French, C., Ward, J. M., 1994, Enzyme immobilization by genetically attached poly-lysine affinity tails. In: *Proceedings of 3rd International Symposium*, "Perspectives on protein engineering and complementaty technologies". Oxford University.

French, C., Ward, J. M., 1995, "Improved production and stability of *E. Coli* recombinants expressing transketolase for large scale biotransformation", *Biotechnology Letters*, **17** (3), 247-252.

French, C., Ward, J. M., 1996, "Production and modification of *E. coli* transketolase for large-scale biocatalysis", *Annal. N. Y. Acad. Sci.*, **799**, 11-18.

Godfrey, T., Reichelt, J., 1983, In: *Industrial Enzymology.* The Nature Press, New York, USA.

Golbik, R., Neef, H., Hubner, G., Kotchetov, G. A., Schellenberger, A., 1991, "Function of the aminopyrimidine part in thiamine pyrophosphate enzymes", *Bioorg. Chem.*, **19**, 10-17.

Gottarelli, C., Spada, G. P., 1991, "Spontaneous resolution of 2,2'-dimethoxy-1,1'-binaphthalene", *J. Org. Chem.*, **56**, 2096-2098.

de la Haba, G., Racker, E., 1952, "Metabolism of ribose-5-phosphate and ribulose-5-phosphate", *Fed. Proc.*, **11**, 201.

de la Haba, G., Leder, I. G., Racker, E., 1955, "Crystalline transketolase from bakers' yeast: isolation and properties", *J. Biol. Chem.*, **214**, 409-426.

Hecquet, L., Bolte, J., Demuynck, C., 1994, "Chemoenzymatic synthesis of 6-deoxy-D-fructose and 6-deoxy-L-sorbose using transketolase", *Tetrahedron*, **50** (29), 8677-8684.

Heinrich, P. C., Stefan, H., Janser, P., Wiss, O., 1972, "Studies on the reconstitution of apo-transketolase with thiamine pyrophosphate and analogues of the coenzyme", *Eur. J. Biochem.*, **30**, 533-541.

Hobbs, G. R., Lilly, M. D., Turner, N. J., Ward, J. M., Willetts, A. J., Woodley, J. M., 1993, "Enzyme catalysed carbon-carbon bond formation: Use of transketolase from *Escherichia coli*", *J. Chem. Soc. Perkin. Transactions*, **1**, 165-166.

Hobbs, G. R., Lilly, M. D., Turner, N. J., Ward, J. M., Woodley, J. M, 1994, "Development of transketolase as a catalyst for carbon-carbon bond formation", *ECB6: Proceedings of the 6th European Congress in Biotechnology*, 463-466.

Hobbs, G. R., Mitra, R. K., Chauhan, R. P., Woodley, J. M., Lilly, M. D., 1996, "Enzyme-catalysed carbon-carbon bond formation: large-scale production of *Escherichia coli* transketolase", *J. Biotechnol.*, **45** (2), 173-179.

Horecker, B. L., Smyrniotes, P. Z., 1952, "The enzymatic formation of sedoheptulose phosphate from pentose phosphate", *J. Am. Chem. Soc.*, **74**, 2123.

Horecker, B. L., Smyrniotes, P. Z., Klenow, H., 1953, "The formation of sedoheptulose-phosphate from pentose-phosphate", *J. Biol. Chem.*, **205**, 661-682.

Iida, A., Teshiba, S., Mizobuchi, K., 1993, "Identification and characterisation of the *tktB* gene encoding a second transketolase in *Escherichia coli* K-12", *J. Bacteriology*, **175** (17), 5375-5383.

Jones, J. B., 1986, "Enzymes in Organic Synthesis", *Tetrahedron*, **42**, 3351-3403.

Jones, J. B., Beck, J. F., 1976, "Asymmetric syntheses and resolutions using enzymes", *Tech. Chem.*, **10**, 107-141.

Kagan, H. B., Fiaud, J. C., 1988, "Kinetic Resolution", *Top. Stereochem.*, **18**, 249-330.

Kleina, L. G., Marson, J-M., Normanly, J., Abelson, J., Miller, J. H., 1990, "Construction of *Escherichia coli* Amber Suppressor tRNA Genes. II Synthesis of Additional tRNA Genes and Improvement of Suppressor Efficiency", *J. Mol. Biol.*, **213**, 705-717.

Kluger, R., 1987, "Thiamin diphosphate: A mechanistic update on enzymatic and non-enzymatic catalysis of decarboxylation", *Chem. Rev.*, **87**, 863-876.

Kluger, R., 1990, "TPP derivatives in C-C bond formation", *Chem. Rev.*, **90**, 1151-1155.

Kobori, Y., Myles, D. C., Whitesides, G. M., 1992, "Substrate specificity and carbohydrate synthesis using transketolase", *J. Org. Chem.*, **57**, 5899-5907.

Kochetov, G. A., 1982, "Transketolase from yeast, rat liver, and pig liver", *Methods Enzymol.*, **90**, 209-223.

Konig, S., Schellenberger, A., Neef, H., Schneider, G., 1994, "Specificity of coenzyme binding in thiamin diphosphate-dependent enzymes. Crystal structure of yeast transketolase in complex with analogs of thiamin diphosphate", *J. Biol. Chem.*, **269** (14), 10879-10882.

Kragl, U. P., Appel, D., Gygax, O., Ghisalba, C., Wandrey, 1992. "Aldolases for use in carbohydrate synthesis: enzymatic reaction engineering as a tool for process optimization". In: *Biochemical Engineering for 2001*. Proc. Of Asia-Pacific

Biochemical Engineering Conference. Furusaki, S., Endo, I., Matsuno, R., (eds). Springer-Verlag, Tokyo, 84-87.

Kramer, W., Fritz, H. J., 1987, "Oligonucleotide-directed construction of mutations via gapped duplex DNA", *Methods in Enzymology*, **154**, 350-367.

Krampitz, L. O., 1969, "Catalytic functions of thiamine pyrophosphate", *Annu. Rev. Biochem.*, **38**, 213-241.

Kremer, A. B., Egan, R. M., Sable, H. Z., 1980, "The active site of transketolase", *J. Biol. Chem.*, **255** (6), 2405-2410.

Krstulovic, A. M., (ed.), 1989, In: *Chiral Separations by HPLC.* Ellis Norwood Ltd., Chichester.

Kunkel, T. A., 1985, "Rapid and efficient site-specific mutagenesis without phenotypic selection", *Proc. Natl. Acid. Sci. USA*, **82**, 488-492.

Kunkel, T. A., Roberts, J. D., Zakour, R. A., 1987, "Rapid and efficient site-specific mutagenesis without phenotypic selection", *Meth. Enzymol.*, **154**, 367-382.

Laumen, K., Breitgoff, D., Schneider, M. P., 1988, "Enzymatic preparation of enantiomerically pure secondary alcohols. Ester synthesis by irreversible acyl transfer using a highly selective ester hydrolase from *Pseudomonas sp.*; an attractive alternative to ester hydrolysis", *J. Chem. Soc., Chem. Commun.*, **10** (22), 1459-1461.

Lerner, R. A., Benkovic, S. J., Schultz, P. G., 1991, "At the Crossroads of Chemistry and Immunology: Catalytic Antibodies", *Science*, **252**, 659-667.

Lilly, M. D., Chauhan, R., French, C., Gyamerah, M., Hobbs, G. R., Humphrey, A., Isupov, M., Littlechild, J. A., Mitra, R. K., Morris, K. G., Rupprecht, M., Turner, N. J., Ward, J. M., Willets, A. J., Woodley, J. M., 1996, "Carbon-carbon bond synthesis: The impact of rDNA technology on the production and use of *E. coli* transketolase", *Ann. N. Y. Acad. Sci.*, **782**, 513-525.

Lin, E. C. C., 1987. "Dissimilatory pathways for sugars, polyols, and carboxylates". In: Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., Umbarger, H. E., (eds) *Escherichia coli* and *Salmonella typhimuriumi* cellular and molecular biology, vol 1. American Society for Microbiology, Washington DC, 244-284.

Lindqvist, Y., Schneider, G., Ermler, U., Sundstrom, M., 1992, "Three-dimensional structure of transketolase, a thiamine diphosphate dependant enzyme, at 2.5A resolution", *EMBO J.*, **11** (7), 2373-2379.

Littlechild, J. A., Turner, N. J., Hobbs, G. R., Lilly, M. D., Rawas, A., Watson, H. C., 1995, "Crystallisation and preliminary X-ray crystallographic data with *E. Coli* transketolase", *Acta. Crystal.*, **51**, 1074-1076.

Lundberg, K. S., Shoemaker, D. D., Adams, M. W. W., Short, J. M., Sorge, J. A., Mathur, E. J., 1991, "High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiousus*", *Gene*, **108**, 1-6.

Margolin, A. L., 1993, "Enzymes in the synthesis of chiral drugs", *Enzyme Microb. Technol.*, **15**, 266-280.

McBride, J. M., Carter, R. L., 1991, "Spontaneous resolution by stirred crystallisation", *Angew. Chem. Ed. Engl.*, **30**, 293-295.

Meshalkina, L., Nilsson, U., Wikner, C., Kostikowa, T., Schneider, G., 1997, "Examination of the thiamin diphosphate binding site in yeast transketolase by site-directed mutagenesis", *Eur. J. Biochem.*, **244** (2), 646-652.

Mitra, R. K., 1997, "Carbon-Carbon Bond Formation: Reactor selection for TK-catalysed biotransformations", *PhD Thesis*, University College London, U.K.

Mitra, R. K., Woodley, J. M., 1996, "A useful assay for transketolase in asymmetric synthesis", *Biotechnology Techniques*, **10** (3), 167-172.

Morris, K. G., Smith, M. E. B., Turner, N. J., Lilly, M. D., Mitra, R. K., Woodley, J. M., 1996, "Transketolase from *Escherichia coli*: A practical procedure for using the biocatalyst for asymmetric carbon-carbon bond synthesis", *Tet. Asymm.*, **7**, 2185-2188.

Morrison, R., Scott, T., 1985, In: *Asymmetric Synthesis*, vol 4. Academic Press, New York, USA.

Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F., Schneider, G., 1993, "A thiamine diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase", *Structure*, **1** (2), 95-103.

Muller, Y. A., Schulz, G. E., 1993, "Structure of the thiamine- and flavin-dependent enzyme pyruvate oxidase", *Science*, **259** (5097), 965-967.

Mulzer, J., 1991, In: *Organic synthesis highlights*. Altenbach and Braun. Eds.: 243-250. VCH V erlags Gesellschaft, Weinheim, Germany.

Myles, D. C., Andrulis, P. J., Whitesides, G. M., 1991, "A transketolase-based synthesis of (+)-*exo*-brevicomin", *Tetrahedron Lett.*, **32** (37), 4835-4838.

Nielsen, Houlihan, 1968, "The Aldol Condensation", *Org. React.*, **16**, 1-438.

Nikkola, M., Lindqvist, Y., Schneider, G., 1994, "Refined structure of transketolase from *Saccharomyces cerevisae* at 2.0A resolution", *J. Mol. Biol.*, **238**, 387-404.

Nilsson, U., Lindqvist, Y., Kluger, R., Schneider, G., 1993, "Crystal structure of transketolase in complex with thiamine thiazone diphosphate, an analogue of the reaction intermediate at 2.3A resolution", *FEBS Lett.*, **326** (1-3), 145-148.

Nilsson, U., Meshalkina, L., Lindqvist, Y., Schneider, G., 1997, "Examination of substrate binding in TPP-dependent transketolase by crystallography and site-directed mutagenesis", *The Journal of Biological Chemistry*, **272** (3), 1864-1869.

Normanly, J., Kleina, L. G., Masson, J-M., Abelson, J., Miller, J. H., 1990, "Construction of *Escherichia coli* Amber Suppressor tRNA Genes. III Determination of tRNA Specificity", *J. Mol. Biol.*, **213**, 719-726.

Norrander, J., Kempe, T., Messing, J., 1983, "Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis", *Gene*, **26**, 101-106.

Racker, E, 1961, In: *The enzymes*, Boyer, P. D., Lardy, H., Myrzback, K. (Eds), vol **5**, pg 397, Academic Press, New York.

Racker, E., de la Haba, G., Leder, I. G., 1953, "TPP, a coenzyme of transketolase", *J. Am. Chem. Soc.*, **75**, 1010-1011.

Roszowski, A. P., Rooks, W. H., Tomolinis, A. J., Miller, L. M., 1971, "Anti-Inflammatory and analgetic properties of d-2-(6'-methoxy-2'-naphthyl)-propionic acid (naproxen)", *J. Pharmacol. Exp. Ther.*, **179**, 114-123.

Saiki, R. K., Gelfand, |D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., Erlich, H. A., 1988, "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase", *Science*, **239**, 487-491.

Schellenberger, A., 1967, "Structure and Mechanism of Action of the Active Centre of Yeast Pyruvate Decarboxylase", *Angew. Chem. Int. Ed. Engl.*, **6**, 1024-1035.

Schmitt, J., Hess, H., Stunnenberg, H. G., 1993, "Affinity purification of histidine-tagged proteins", *Molecular Biol. Reps.*, **18** (3), 223-230.

Schneider, G., Lindqvist, Y., 1993, "Enzymatic thiamine catalysis: Mechanistic implications from three dimensional structure of transketolase", *Bio-organic Chemistry*, **21** (1), 109-117.

Schneider, G., Lindqvist, Y., 1998, "Crystallography and mutagenesis of transketolase: mechanistic implications for thiamine catalysis", *Biochim. Biophys. Acta*, **1385** (2), 387-398.

Schneider, G., Sundstrom, M., Lindqvist, Y., 1989, "Preliminary Crystallographic Data for Transkelolase from Yeast", *J. Biol. Chem*, **264** (36), 21619-21620.

Schorken, U., Sprenger, G. A., 1998, "Thiamin-dependent enzymes as catalysts in chemoenzymatic syntheses", *Biochim. Biphys. Acta.*, **1385** (2), 229-243.

Sheldon, R., 1990, "Industrial synthesis of optically active compounds", *J. Chem. Ind.*, 212-219.

Shen, T. Y., 1972, "Perspectives in Nonsteroidal Anti-inflammatory Agents", *Angew. Chem. Int. Ed. Engl.*, **11**, 460-472.

Sih, C. J., Gu, Qu-M., Fulling, G., Wu, S.-H., Reddy, D. R., 1988, "Resolution and properties of suprofen", *Dev. Ind. Microb.*, **29**, 221-229.

Singleton, C. K., Wang, J. J., Shan, L., Martin, P. R., 1996, "Conserved residues are functionally distinct within transketolasees of different species", *Biochemistry*, **35** (49), 15865-15869.

Spratt, B. G., Hedge, P. J., teHeesen, S., Edelman, A., Broome-Smith, J. K., 1986, "Kanamycin-resistant vectors that are analogues of plasmids pUC8, pUC9, pEMBL8, pEMBL9", *Gene*, **41**, 337-342.

Sprenger, G. A., 1993, "Nucleotide Sequence of the *E. Coli* K-12 transketolase (tkt) gene", *Biochim. Biophys. Acta*, **1216**, 307-310.

Sprenger, G. A., Schorken, U., Sprenger, G., Hermann, S., 1995, "Transketolase A from *Escherichia coli* K-12: Purification and properties of the enzyme from recombinant strains", *Eur. J. Biochem.*, **230**, 525-532.

Srere, P. A., Cooper, J. R., Tabachnick, M., Racker, E., 1958, "The oxidative pentose phosphate cycle. I. Preparation of substrate and enzymes", *Arch. Biochem, and Biophys.*, **74**, 295-305.

Sundstrom, M., Lindqvist, Y., Schneider, G., Hellman, U., Ronne, H., 1993, "Yeast TKL1 Gene encodes a transketolase that is required for efficient Glycolysis and Biosynthesis of Aromatic Amino Acids", *J. Biol. Chem.*, **268** (32), 24346-24352.

Sundstrom, M., Lindqvist, Y., Schneider, G., 1992, "3-D structure of apo-transketolase: flexible loops at the active site enable cofactor binding", *FEBS lett.*, **313** (3), 229-231.

Testa, B., Mayer, M., 1988, "Stereoselective drug metabolism and its significance in drug research", *Prog. Drug Res.*, **32**, 249-303.

Tien, M., Tu, C. P. D., 1987, "Cloning and sequencing of cDNA for a ligase from *Phanaerochaete chrysogenum*", *Nature*, **326**, 520-523.

Toone, E. J., Simon, E. S., Bednarski, M. D., Whitesides, G. M., 1989, "Enzyme-catalysed synthesis of carbohydrates", *Tetrahedron*, **45** (17), 5365-5422.

Usmanov, R. A., Kochetov, G., 1983, "Function of the arginine residue in the active centre of baker's yeast transketolase", *Biochemistry (Engl. Transl. Biokhimiia),* **48** (5), 772-781.

Van der Werf, M. J., Van der Tweel, W. J. J., Kamphuis, J., Hartmans, S., De Bont, J. A. M., 1994, "The potential of lyases for the industrial production of optically active compounds", *TiBtech*, **12**, 95-103.

Vieire, J., Messing, J., 1987, "Producing of single-stranded plasmid DNA", *Methods in enzymology,* **123**, 3-11.

Villafranca, J. J., Axelrod, B., 1971, "Heptulose synthesis of non-phosphorylated aldoses and ketoses by spinach transketolase", *J. Biol. Chem.*, **246** (10), 3126-3131.

Vuorinen, T., Serianni, A. S., 1990, "Synthesis of D-*erythro*-2-pentulose and D-*threo*-2-pentulose and analysis of the $^{13}C$- and $^{1}H$-nmr spectra of the 1-$^{13}C$- and 2-$^{13}C$ substituted sugars", *Carbohydrate Research*, **209**, 13-31.

Ward, R. S., 1990, "Non-enzymatic asymmetric transformations involving symmetrical bifunctional compounds", *Chem. Soc. Rev.*, **19**, 1-19.

Wetzel, R., 1986, "Engineered proteins with improved characteristics", *Protein Engineering*, **1** (2), 3-8.

Whitesides, G. M., Wong, C.-H., 1985, "Enzymes as Catalysts in Synthetic Organic Chemistry", *Angew. Chem. Int. Ed. Engl.*, **24**, 617-638.

Wikner, C., Nilsson, U., Meshalkina, L., Udekwu, C., Lindqvist, Y, Schneider, G., 1997, "Identification of catalytically important residues in yeast transketolase", *Biochemistry*, **36** (50), 15643-15649.

Wikner, C., Meshalkina, L., Nilsson, U., Nikkola, M., Lindqvist, Y., Sundstrom, M., Schneider, G., 1994, " Analysis of an invariant cofactor-protein interaction in thiamin diphosphate-dependent enzymes by site-directed mutagenesis. Glutamic acid 418 In transketolase is essential for catalysis", *J. Biol. Chem.*, **269** (51), 32144-32150.

Wong, C., Whitesides, G. M., 1994, In: *Enzymes In Synthetic Organic Chemistry.* Elsevier Science, Amsterdam, Holland.

Wood, 1985, In: *The pentose phosphate pathway.* Academic Press, Orlando.

Zhao, G., Winkler, M. E., 1994, "An *Escherichia coli* K-12 tktA tktB mutant deficient in transketolase activity requires pyridoxine (vitamin B6) as well as the aromatic amino acids and vitamins vital for growth", *J. Bacteriol.*, **176** (19), 6134-6138.

Ziegler, T., Straub, A., Effenberger, F., 1988, "Enzyme-catalysed synthesis of 1-deoxymannojirimycin, 1-deoxynojirimycin and 1,4-dideoxy-1,4-imino-D-arabinitol", *Angew. Chem. Int. Ed. Engl.*, **27**, 716-717.

# Appendices

# APPENDIX ONE: Media, Reagents and Buffers

## A1.1 Suppliers of materials

General chemicals and biochemicals used to make buffers and reagents were of the highest purity available and were obtained from Aldrich, Gibco-BRL, or Sigma.

Oligonucleotides were purchased from Amersham Pharmacia Biotech.

Sequencing was performed by Oswel.

QuikChange™ Site-Directed Mutagenesis kit was purchased from STRATAGENE.

INTERCHANGE™ Amber Suppressor *in vivo* Mutagenesis System was purchased from Promega.

Dehydrated Bacto™ Tryptone, Bacto™ Yeast, Bacto™ Agar and other ingredients for cell culture media were purchased from Difco. Water for media preparation and all other uses was purified using a Millipore Milli-Q Plus system with a 0.4μm filter attached to the outlet.

Agarose was purchased from Gibco-BRL.

Molecular weight markers for DNA gels were purchased from Gibco-BRL. Those for protein gels were purchased from NOVEX™.

M13 helper phage, restriction endonucleases, DNA modifying enzymes and 10x buffer concentrates were purchased from Gibco-BRL, Pharmacia, New England Biolabs and Boehringer-Mannheim.

Easigel™ for SDS-PAGE protein gels was obtained from Flowgen. Acrylamide was purchased from Severn Biotech Ltd.

N,N,N',N'-tetramethylethylenediamine (TEMED) and ammonium persulphate were obtained from Bio-Rad.

QIAquick gel extraction kits and QIAprep spin miniprep kits were purchased from QIAGEN.

## A1.2 Microbiological Techniques

Standard sterile technique was maintained throughout. Media and equipment were sterilized by autoclaving at 121°C for 20 minutes. Thermolabile solutions were sterilized by filtration through 0.22μm filters. Plates were poured and streaked, and media inoculated, with attendant flame to create an updraught and to sterilize inoculation loops.

All aqueous solutions were made up to the specified volume by the addition of de-ionized water.

## A1.3 Microbiological strains

*E. coli* strains used in this study

| *E. coli* strain | Genotype |
|---|---|
| XL1 | F' Tn *10proA*$^+$*B*$^+$ *lacI*$^q$ *Δ(lacZ)M15/recA1 endA1 gyrA96* (Nal$^r$) *thi hsdR17* ($r_K^-$ $m_K^+$) *supE44 relA1 lac* |
| JM107 | F' *traD36 lacI*$^q$*Δ(lacZ)M15 proA*$^+$*B*$^+$*1* e14$^-$(McrA$^-$) *Δ(lac-proAB) thi gyrA96* (Nal$^r$) *endA1 hsdR17*($r_K^-$ $m_K^+$) *relA1 supE44* |
| CJ236 | F' *cat* (=pCJ105; M13$^s$Cm$^r$)/*dut ung1 thi-1 relA1 spoT1 mcrA* |
| Amber-SER, -LEU, -LYS, -TYR | *ara Δ(lac-pro) gyrA metB argE*$^{am}$ *rpoB, thi* |

Upon receipt, all bacterial strains were streaked out on the appropriate medium and a single colony isolated and maintained for further experimentation frozen at -70°C in 20% glycerol.

## A1.4 Microbiological growth media

Solid media were prepared by supplementing liquid media with 2% agar (w/v) prior to autoclaving. All media were allowed to cool to below 50°C before the addition of thermolabile substances such as antibiotics.

**A1.4.1 LB (Luria Bertani) Type I medium:** General-purpose bacterial cell culture medium;

Per litre:

| | |
|---|---|
| Tryptone | 10.0g |
| Yeast Extract | 5.0g |
| NaCl | 10.0g |
| Glucose | 1.0g |

For LB Type II medium no glucose was added. Autoclave at 121°C for 20 minutes.

### A1.4.2 Minimal media for amber strains

| Strain | Minimal media components (per litre) | |
|---|---|---|
| SER, TYR, LYS, LEU | 0.1M Proline (10ml) | M9 salts |
| | 0.1M Methionine (10ml) | 20% glucose (10ml) |
| | 0.1M Thiamine-HCl (10ml) | 1M $CaCl_2$ (0.1ml) |
| | 1M $MgSO_4$ (2ml) | |

## A1.5 Solutions for plasmid preparation by alkaline lysis

### A1.5.1 **QPP 1:** Resuspension buffer;

Per litre:

| | |
|---|---|
| 1.0M glucose | 50ml |
| 0.5M EDTA pH8.0 | 20ml |
| 1.0M Tris.HCl (A1.14.2) | 25ml |

Adjust to pH8.0 with HCl.

### A1.5.2 **QPP 2:** Alkaline lysis solution;

Per 10ml:

| | |
|---|---|
| 5.0M NaOH | 0.4ml |
| 10% (w/v) SDS (A1.14.1) | 1.0ml |

Add the NaOH to the water prior to addition of the SDS to prevent solution precipitation. Prepare fresh before use.

### A1.5.3 **QPP 3:** Neutralisation solution;

Per litre:

| | |
|---|---|
| 5.0M Potassium acetate | 600ml |
| Glacial acetic acid | 115ml |

## A1.6 Solutions for preparation of single-stranded DNA

### A1.6.1 Polyethylene glycol (PEG) solution:

| | |
|---|---|
| PEG 8000 | 40% (w/v) |

Autoclave at 121°C for 20 minutes.

### A1.6.2 Sodium acetate solution:

Per 100ml:

| | |
|---|---|
| NaOAc | 8.2g |

Adjust to pH7.0 with NaOH. Autoclave at 121°C for 20 minutes.

## A1.7 Enzyme reaction buffers (supplied by the manufacturer)

### A1.7.1 NEB buffer 1:

| | |
|---|---|
| Bis.Tris.Propane.HCl | 10mM |
| $MgCl_2$ | 10mM |
| DTT (pH7.0) | 1mM |

### A1.7.2 NEB buffer 2:

| | |
|---|---|
| Tris.HCl | 10mM |
| $MgCl_2$ | 10mM |
| NaCl | 50mM |
| DTT (pH7.9) | 1mM |

### A1.7.3 NEB buffer 3:

| | |
|---|---|
| Tris.HCl | 50mM |
| $MgCl_2$ | 10mM |
| NaCl | 100mM |
| DTT (pH7.9) | 1mM |

### A1.7.4 NEB buffer 4:

| | |
|---|---|
| Tris acetate | 20mM |
| Mg acetate | 10mM |
| K acetate | 50mM |
| DTT (pH7.9) | 1mM |

**A1.7.5 10x T4 DNA polymerase buffer:**

| | |
|---|---|
| NaCl | 500mM |
| Tris.HCl pH7.9 at 25°C | 100mM |
| $MgCl_2$ | 100mM |
| DTT | 10mM |

**A1.7.6 10x T4 DNA ligase buffer:**

| | |
|---|---|
| Tris.HCl pH7.5 | 500mM |
| MgCl2 | 100mM |
| DTT | 100mM |
| ATP | 10mM |

**A1.7.7 10x Reaction Buffer:**

| | |
|---|---|
| KCl | 100mM |
| $(NH_4)_2SO_4$ | 100mM |
| Tris-HCl pH8.8 | 200mM |
| $MgSO_4$ | 20mM |
| 1% Triton$^R$ X-100 | |
| 1mg/ml nuclease-free bovine serum albumin (BSA) | |

## A1.8 Nucleic acid buffers and solutions

**A1.8.1 TE (Tris/EDTA) buffer:** For DNA suspension;

| | |
|---|---|
| Tris.HCl (A12.2) | 10mM |
| EDTA | 1mM |

Adjust to pH8.0 with HCl.

**A1.8.2 TBE (Tris/borate/EDTA) buffer:** For agarose gel electrophoresis;

| | |
|---|---|
| Triszma base | 90mM |
| Boric acid | 90mM |
| EDTA | 2mM |

Adjust to pH8.3.

Gels were stained after running, using ethidium bromide solution at 0.5μg/ml.

**A1.8.3 Buffered Phenol:**

| | |
|---|---|
| Molten phenol | 500ml |
| 8-Hydroxyquinoline | 0.5g |

**A1.8.4 Phenol/chloroform:**

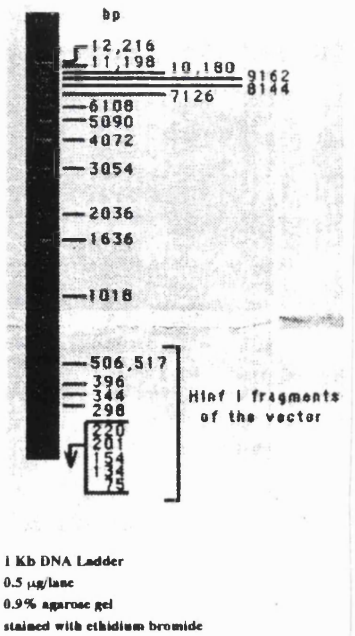
Per 100ml:

| | |
|---|---|
| Buffered phenol (A1.8.3) | 50ml |
| Chloroform | 50ml |

**A1.8.5 6x Agarose gel sample loading buffer:**

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

**A1.8.6 1kb DNA ladder** (Gibco-BRL):




Picture of 1kb ladder, taken from Gibco-BRL product catalogue. 2μl was loaded per lane.

**A1.8.7 7.5% acrylamide stock:** (7.5% of 30% acrylamide and 7.5% of glycerol);

Per 100ml:

| | |
|---|---|
| 5x TBE | 20ml |
| 30% acrylamide | 25ml |
| 50% glycerol | 15ml |

**A1.8.8 Polyacrylamide gels:**

| Gel size | 7.5% acrylamide (ml) | 10% APS (µl) | TEMED |
|---|---|---|---|
| small | 10 | 100 | 10 |
| medium | 40 | 400 | 40 |
| large | 50 | 500 | 50 |

10% APS was made fresh before use.

## A1.9 SDS-PAGE protein electrophoresis solutions

**A1.9.1 Electrode buffer: stock solution**

Per litre:

| | |
|---|---|
| Glycine | 150g |
| Trizma base | 30g |

Store at rt.

**A1.9.2 Electrode buffer: working solution**

Per litre:

| | |
|---|---|
| 10% SDS (A1.14.1) | 10ml |
| Stock electrode buffer | 100ml |

Store at rt.

**A1.9.3 Stock running gel buffer (0.75M Tris base pH8.3)**

Per 500ml:

| | |
|---|---|
| Trizma base | 45.43g |

Adjust to pH8.3 using HCl. Store at 4°C.

### A1.9.4 Stock stacking gel buffer (1.25M Tris base pH6.8)

Per 250ml:

| | |
|---|---|
| Trizma base | 37.86g |

Adjust to pH6.8 using HCl. Store at 4°C.

### A1.9.5 Ammonium persulphate

Per ml:

| | |
|---|---|
| Ammonium persulphate | 80mg |

Prepare fresh before use.

### A1.9.6 Sample loading buffer

Per 46ml:

| | |
|---|---|
| SDS | 1g |
| Glycerol | 10ml |
| Bromophemol blue | 2.5mg |
| 1/10x stacking gel buffer | 33.5ml |

Store in 1ml aliquots at rt. Add 87µl β-mercaptoethanol immediately before use.

### A1.9.7 Stacking gel (6%)

| | |
|---|---|
| Acrylamide | 6ml |
| Stacking buffer | 3ml |
| $H_2O$ | 21ml |
| 10% SDS (A1.14.1) | 300µl |
| APS (A1.9.5) | 300µl |
| TEMED | 15µl |

### A1.9.8 Separating gel (15%)

| | |
|---|---|
| Acrylamide | 20ml |
| Running buffer | 20ml |
| 10% SDS (A1.14.1) | 200µl |
| APS (A1.9.5) | 400µl |
| TEMED | 20µl |

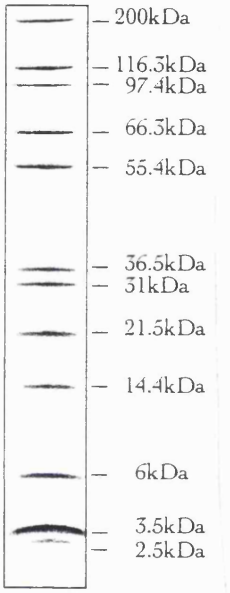
**A1.9.9 Coomassie blue stain solution (0.2% w/v)**

Per litre:

| | |
|---|---|
| Stain | 2g |
| Methanol | 500ml |
| Glacial acetic acid | 100ml |

Mix thoroughly and filter before storing at rt.

**A1.9.10 Mark12™ Wide Range Protein Standard** (NOVEX™):




Picture of Mark12™ Wide Range Protein Standard, taken from NOVEX™ product catalogue. 15µl was loaded per lane.

**A1.9.11 Fast destain solution**

| | |
|---|---|
| Methanol | 40% (v/v) |
| Glacial acetic acid | 10% (v/v) |

**A1.9.12 Slow destain solution**

| | |
|---|---|
| Methanol | 10% (v/v) |
| Glacial acetic acid | 10% (v/v) |

## A1.10 HPLC buffers

### A1.10.1 6mM $H_2SO_4$

Per litre:

| | |
|---|---|
| $H_2SO_4$ | 320µl |

Degas and filter through 0.45µm filters.

## A1.11 FPLC elution buffers (degass and filter to 0.22µm)

### A1.11.1 Gly-Gly (pH9.5, 150mM)

Per litre:

| | |
|---|---|
| Gly- Gly | 19.8g |

Adjust to pH9.5 and autoclave at 121°C for 20 minutes.

### A1.11.2 Gly-Gly (pH7.0, 150mM)

Per litre:

| | |
|---|---|
| Gly-Gly | 19.8g |

Adjust to pH7.5 and autoclave at 121°C for 20 minutes.

### A1.11.3 MES (pH5.0, 150mM)

Per litre:

| | |
|---|---|
| MES | 19.52g |

Adjust to pH5.0 and autoclave at 121°C for 20 minutes.

### A1.11.4 NaCl (pH7.5, 1M)

Per litre:

| | |
|---|---|
| NaCl | 58.44g |

Adjust to pH7.5 and autoclave at 121°C for 20 minutes.

## A1.12 FPLC wash buffers

### A1.12.1 2M NaCl

Per 10ml:

NaCl 1.17g

### A1.12.2 2M NaOH

Per 10ml:

NaOH 0.8g

### A1.12.3 75% Acetic acid

Per 10ml:

Glacial acetic acid 7.5ml

## A1.13 Antibiotics

| Antibiotic | Sock solution (mg/ml) | Working solution (μg/ml) | Solvent |
|---|---|---|---|
| ampicillin | 100 | 100 | $H_2O$ |
| kanamycin | 10 | 10 | $H_2O$ |
| chloramphenicol | 20 | 20 | 80% EtOH |
| tetracycline | 12.5 | 12.5 | 80% EtOH |

Filter-sterilize then store in 1ml aliquots at -20°C.

## A1.14 Miscellaneous solutions

### A1.14.1 10% SDS

Per 100ml:

SDS 10g

Adjust to pH8.0 and autoclave at 121°C for 20 minutes.

### A1.14.2 1M Tris.HCl

Per 100ml:

Tris.HCl 15.76g

Adjust to pH7.4 and autoclave at 121°C for 20 minutes.

**A1.14.3 0.1M $CaCl_2$**

Per 100ml:

| | |
|---|---|
| $CaCl_2$ | 1.47g |

Autoclave at 121°C for 20 minutes.

**A1.14.4 10x Annealing buffer**

| | |
|---|---|
| Tris.HCl (pH7.5) | 200mM |
| $MgCl_2$ | 20mM |
| NaCl | 500mM |

**A1.14.5 Gly-Gly (100mM, pH7.6)**

Per 100ml:

| | |
|---|---|
| Gly-Gly | 1.32g |

Adjust to pH7.6 and autoclave at 121°C for 20 minutes.

**A1.14.6 Gly-Gly (70mM, pH7.6)**

Per 100ml:

| | |
|---|---|
| Gly-Gly | 0.924g |

Adjust to pH7.6 and autoclave at 121°C for 20 minutes.

**A1.14.7 Permanganate TLC stain solution**

| | |
|---|---|
| $KMnO_4$ | 3.0g |
| $K_2CO_3$ | 20.0g |
| 5% aq. NaOH | 5ml |

**A1.14.8 Anisaldehyde TLC stain solution**

| | |
|---|---|
| Anisaldehyde | 15.0g |
| EtOH | 250ml |
| conc. $H_2SO_4$ | 2.5ml |

### A1.14.9 IPTG solution (100mM)

Per 10ml:

| | |
|---|---|
| IPTG | 0.238g |

Store in 1ml aliquots at -20°C.

### A1.14.10 0. 5M EDTA

Per 100ml:

| | |
|---|---|
| EDTA | 18.612g |

Adjust to pH8.0 with NaOH pellets. Autoclave at 121°C for 20 minutes.

# APPENDIX TWO: Oligonucleotide Sequences

## A2.1 Mutagenesis primers (5'-phosphorylated and FPLC purified)

**A2.1.1** H258A (FOR) CAC to GCC

5'- CAAAGCCGGTACC**GCC**GACTCCCACGGTGC-3' (915 to 944bp)

**A2.1.2** H258A (REV) GTG to GGC

5'-GCACCGTGGGAGTC**GGC**GGTACCGGCTTTG-3' (944 to 915bp)

**A2.1.3** F434A (FOR) TTC to GCC

5'-CGTACACCTCCACC**GCC**CTGATGTTCGTGG-3' (1442 to 1471bp)

**A2.1.4** F434A (REV) GAA to GGC

5'-CCACGAACATCAG**GGC**GGTGGAGGTGTACG-3' (1471 to 1442bp)

**A2.1.5** H258F (FOR) CAC to TTC

5'-CAAAGCCGGTACC**TTC**GACTCCCACGGTGC-3' (915 to 944bp)

**A2.1.6** H258F (REV) GTG to GAA

5'-GCACCGTGGGAGTC**GAA**GGTACCGGCTTTG-3' (944 to 915bp)

**A2.1.7** F434H (FOR) TTC to CAC

5'-CGTACACCTCCACC**CAC**CTGATGTTCGTGG-3' (1442 to 1471bp)

**A2.1.8** F434H (REV) GAA to GTG

5'-CCACGAACATCAG**GTG**GGTGGAGGTGTACG-3' (1471 to 1442bp)

**A2.1.9** R358A (FOR) CGT to GCT

5'-CGAAAATCGCCAGC**GCT**AAAGCGTCTCAG-3' (1214 to 1242bp)

**A2.1.10** R358A (REV) ACG to AGC

5'-CTGAGACGCTTT**AGC**GCTGGCGATTTTCG-3' (1242 to 1214bp)

**A2.1.11** R520A(FOR) CGT to GCT

5'-CTGATCCTCTCC**GCT**CAGAACCTGGCG-3' (1702 to 1728bp)

**A2.1.12** R520A (REV) ACG to AGC

5'-CGCCAGGTTCTG**AGC**GGAGAGGATCAG-3' (1728 to 1702bp)

## A2.2 Sequencing primers

**A2.2.1** H258FOR

5'-GGTCATGACGCGGCATC-3' (808 to 824bp)

**A2.2.2** H258REV

5'-CTGGCCTGCTTCTTTCG-3' (1059 to 1043bp)

**A2.2.3** F434FOR

5'-CAACGAAGATGCTGCGG-3' (1344 to 1360bp)

**A2.2.4** F434REV

5'-CTCAACCGGCTGGTGAG-3' (1587 to 1571bp)

**A2.2.5** R358FOR

5'-CGAAAGAAGCAGGCCAG-3' (1043 to 1059bp)

**A2.2.6** R358REV

5'-CCGCAGCATCTTCGTTG-3' (1360 to 1344bp)

**A2.2.7** R520FOR

5'-CTCACCAGCCGGTTGAG-3' (1571 to 1587bp)

**A2.2.8** R520REV

5'-CAGTTCAACTTCTGAAC-3' (1844 to 1829bp)

## A2.3 Primers for construction of additional mutants characterised in this thesis (constructed by Carol French)

**A2.3.1** C1TK (5'-phosphorylated)

5'-CGCGCTCTAGAGGTACCGTCCGGCGGTA-3'

**A2.3.2** LYS2 (5'-phosphorylated)

5'-TGGAGTCAAAATG**AAAAAA**GGTGGTTCCTCACGTAAA-3'

**A2.3.3** LYS4 (5'-phosphorylated)

5'-TGGAGTCAAAATG**AAAAAAAAAAAA**GGTGGTTCCTCACGTAAA-3'

**A2.3.4** LYS10 (5'-phosphorylated)

5'-GTCAAAATG**AAGAAGAAGAAGAAGAAGAAAAAAAAAAAA**GGTGGT-3'

**A2.3.5** H258D CAC to GAC

5'-AAAGCCGGTACC**GAC**GACTCCCACGGT-3' (916 to 942bp)

**A2.3.6** I189A ATC to GCC

5'-AACGGTATTTCT**GCC**GATGGTCACGTT-3' (709 to 735bp)

**A2.3.7** I189V ATC to GTC

5'-AACGGTATTTCT**GTC**GATGGTCACGTT-3' (709 to 735bp)

**A2.3.8** H461STOP CAC to TAG

5'-ATGGTTTACACC**TAG**GACTCCATCGGT-3' (1525 to 1551bp)

**A2.3.9** S385STOP TCT to TAG

5'-GACCTGGCGCCG**TAG**AACCTGACCCTG-3' (1297 to 1322bp)

# APPENDIX THREE: Nucleotide sequence of the *Escherichia coli* K-12 transketolase (tkt) gene

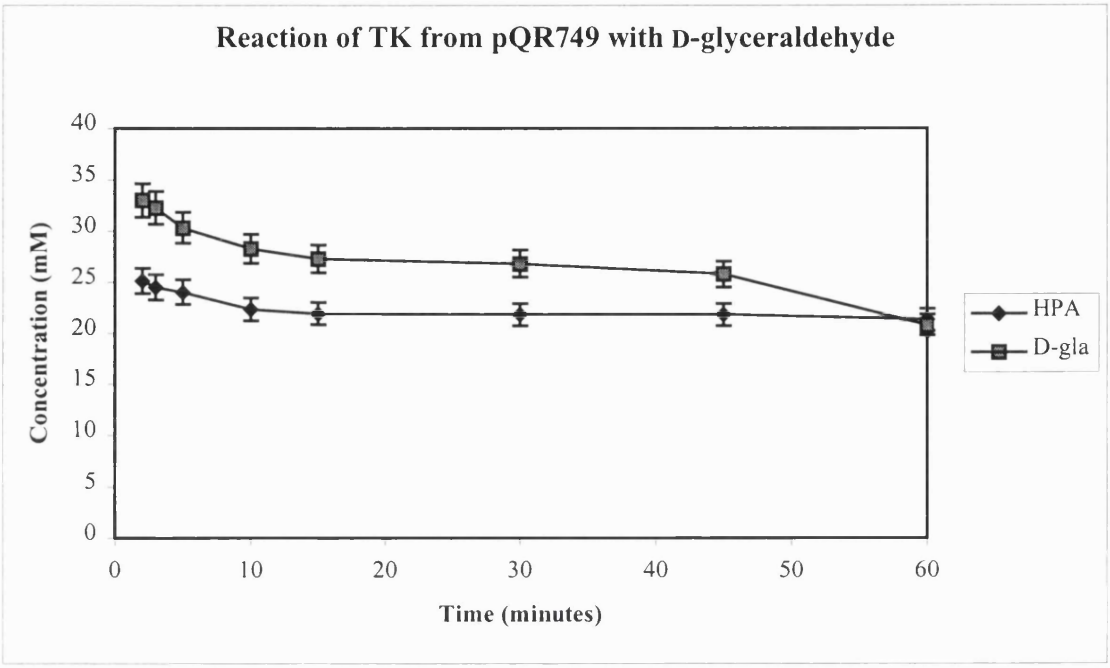
```
   1 gatccagaga tttctgaagc ggcaaaagga tgttccatgt acatgacgcg cggcttgcgg
  61 taaattgttg gcaaattttt cccggcgtag cccaaaacgc gctgtcgtca agtcgttaag
 121 ggcgtgccct tcatcatccg atctggagtc aaaatgtcct cacgtaaaga gcttgccaat
 181 gctattcgtg cgctgagcat ggacgcagta cagaaagcca aatccggtca cccgggggcc
 241 cctatgggta tggctgacat tgccgaagtc ctgtggcgtg atttcctgaa acacaacccg
 301 cagaatccgt cctgggctga ccgtgaccgc ttcgtgctgt ccaacggcca cggctccatg
 361 ctgatctaca gcctgctgca cctcaccggt tacgatctgc cgatggaaga actgaaaaac
 421 ttccgtcagc tgcactctaa aactccgggt cacccggaaa gtggggttac accgctgggt
 481 gtggaaacca ccaccggtcc gctgggtcag ggtattgcca acgcagtcgg tatggcgatt
 541 gcagaaaaaa cgctggcggc gcagtttaac cgtccgggcc acgacattgt cgaccactac
 601 acctacgcct tcatgggcga cggctgcatg atggaaggca tctcccacga agtttgctct
 661 ctggcgggta cgctgaagct gggtaaactg attgcattct acgatgacaa cggtatttct
 721 atcgatggtc acgttgaagg ctggttcacc gacgacaccg caatgcgttt cgaagcttac
 781 ggctggcacg ttattcgcga catcgacggt catgacgcgg catctatcaa acgcgcagta
 841 gaagaagcgc gcgcagtgac tgacaaacct tccctgctga tgtgcaaaac catcatcggt
 901 ttcggttccc cgaacaaagc cggtaccCAC gactcccacg gtgcgccgct gggcgacgct
 961 gaaattgccc tgacccgcga acaactgggc tggaaatatg cgccgttcga aatcccgtct
1021 gaaatctatg ctcagtggga tgcgaaagaa gcaggccagg cgaaagaatc cgcatggaac
1081 gagaaattcg ctgcttacgc gaaagcttat ccgcaggaag ccgctgaatt tacccgccgt
1141 atgaaaggcg aaatgccgtc tgacttcgac gctaaagcga aagagttcat cgctaaactg
1201 caggctaatc cggcgaaaat cgccagcCGT aaagcgtctc agaatgctat cgaagcgttc
1261 ggtccgctgt tgccggaatt cctcggcggt tctgctgacc tggcgccgtc taacctgacc
1321 ctgtggtctg gttctaaagc aatcaacgaa gatgctgcgg gtaactacat ccactacggt
1381 gttcgcgagt tcggtatgac cgcgattgct aacggtatct ccctgcacgg tggcttcctg
1441 ccgtacacct ccaccttcct gatgttcgtg gaatacgcac gtaacgccgt acgtatggct
1501 gcgctgatga aacagcgtca ggtgatggtt tacacccacg actccatcgg tctgggcgaa
1561 gacgggccga ctcaccagcc ggttgagcag gtcgcttctc tgcgcgtaac cccgaacatg
1621 tctacatggc gtccgtgtga ccaggttgaa tccgcggtcg cgtggaaata cggtgttgag
1681 cgtcaggacg gcccgaccgc actgatcctc tcccgtcaga acctggcgca gcaggaacga
1741 actgaagagc aactggcaaa catcgcgcgc ggtggttatg tgctgaaaga ctgcgccggt
1801 cagccggaac tgattttcat cgctaccggt tcagaagttg aactggctgt tgctgcctac
1861 gaaaaactga ctgccgaagg cgtgaaagcg cgcgtggtgt ccatgtcgtc taccgacgca
1921 tttgacaagc aggatgctgc ttaccgtgaa tccgtactgc cgaaagcggt tactgcacgc
1981 gttgctgtag aagcgggtat tgctgactac tggtacaagt atgttggcct gaacggtgct
2041 atcgtcggta tgaccacctt cggtgaatct gctccggcag agctgctgtt tgaagagttc
2101 ggcttcactg ttgataacgt tgttgcgaaa gcaaaagaac tgctgtaatt agcatttcgg
2161 gtaaaaaagg tcgcttcggc gacctttttt attaccttga taatgtccgt ttgcggacca
2221 agcaatagat aaggcgtgtt gtagatcaca aatatttata tgcaataaat atcaattatg
2281 taatatgcat cacgatatgc gtattgacat ttgttgttat atctataact caatgttata
2341 taagaaatta ataattcact gttttcaaaa caaccggttt ccctgctcaa ttgctttcat
2401 taaaccgcac gagtatttta tttttaccgc cggacggtac c
```

**TK gene: 154-2148**

Underlined/bold triplets are the start and stop codons.

Sequences altered in constructs engineered and characterised for work described in this thesis are highlighted in bold-type.

Sequences altered in additional constructs, generated by Carol French and characterised in work described in this thesis, are highlighted in underlined italics.



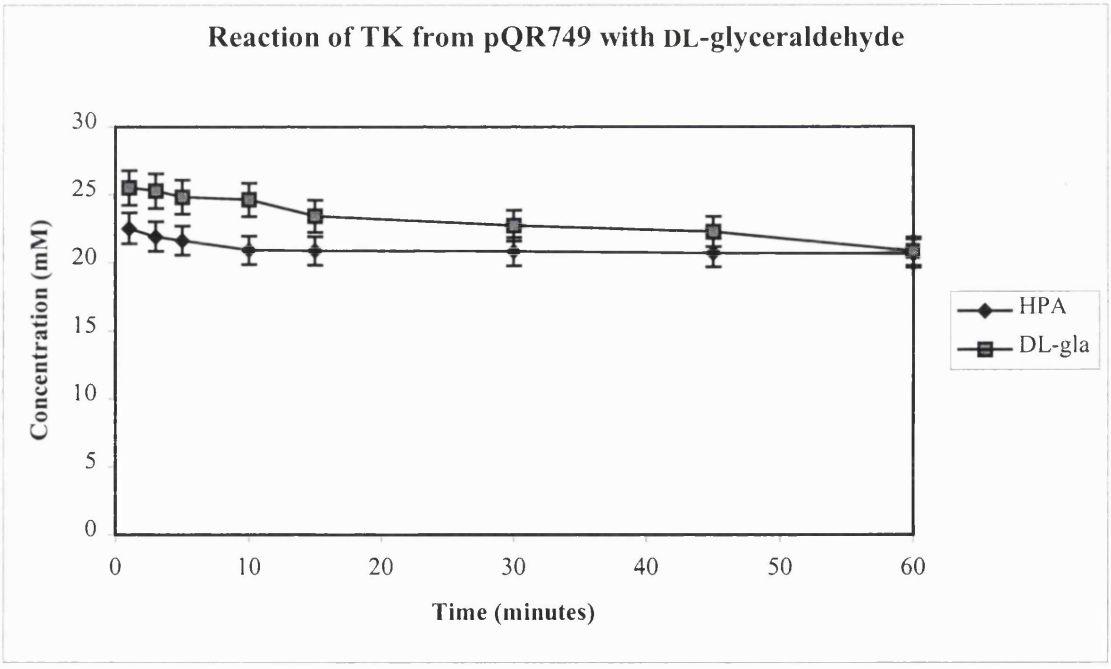




**Figure 5.18** *Loss in β-HPA/D-glyceraldehyde/DL-glyceraldehyde for I189V reaction*